



**HAL**  
open science

# Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés

Guillaume Artigue

► **To cite this version:**

Guillaume Artigue. Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés. Hydrologie. Université Montpellier 2 Sciences et Techniques du Languedoc, 2012. Français. NNT: . tel-01883033

**HAL Id: tel-01883033**

**<https://hal.science/tel-01883033>**

Submitted on 27 Sep 2018

**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.



## **THESE**

Présentée à l'Université Montpellier II  
Sciences et Techniques du Languedoc

Pour obtenir le grade de docteur

**Spécialité : *Eaux Continentales et Sociétés***

**École Doctorale : *Systèmes Intégrés en Biologie, Agronomie, Géosciences,  
Hydrosciences et Environnement (SIBAGHE)***

# **Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés.**

Présentée et soutenue publiquement par :

**Guillaume ARTIGUE**

**Le 3 décembre 2012**

**Directeurs de thèse : Monsieur Séverin Pistre et Madame Anne Johannet.  
Encadrement : Madame Valérie Borrell.**

## **Jury**

**Monsieur Vazken Andréassian : rapporteur.**

**Madame Françoise Prêteux : rapporteur.**

**Madame Marie-George Tournoud : examinatrice, présidente du jury.**

**Madame Caroline Wittwer : examinatrice.**

**Madame Valérie Borrell : encadrante, invitée.**

**Monsieur Roger Moussa : invité.**

**Madame Anne Johannet : co-Directrice de thèse.**

**Monsieur Séverin Pistre : co-Directeur de thèse.**



*« [...] La pluie tombe dans le fleuve ;  
le fleuve se jette dans la mer  
et pourtant la mer n'est jamais rassasiée. [...] »  
« L'homme à la tête ronde » ; Spartacus – Arthur Koestler, 1939.*



## **Remerciements**

J'ai un grand plaisir à remercier Monsieur Séverin Pistre, mon Directeur de Thèse pour son implication constructive dans les relations entre Hydrosociences Montpellier et notre équipe de l'Ecole des Mines d'Alès, ainsi que pour ses relectures et ses conseils clairs et avisés.

Egalement, je remercie Madame Anne Johannet, ma co-Directrice de Thèse pour ses nombreux conseils et ses relectures attentives, en particulier concernant les réseaux de neurones.

L'encadrement de Madame Valérie Borrell d'Hydrosociences Montpellier m'a beaucoup aidé en contribuant aux réflexions menées à propos des problématiques hydrologiques et de modélisation ; qu'elle en soit ici chaleureusement remerciée.

Je suis également très honoré que Madame Françoise Prêteux et Monsieur Vazken Andréassian aient accepté d'être les rapporteurs de ce travail ; qu'ils soient ici remerciés de leur contribution décisive.

Pour l'honneur qu'elle me fait d'examiner ce travail, je remercie Madame Marie-George Tournoud ainsi que Monsieur Roger Moussa, qui m'a de plus fait le plaisir de participer à mes Comités de Thèse.

Pour son accueil dans le laboratoire je tiens à remercier Monsieur José-Marie Lopez-Cuesta, Directeur du Centre des Matériaux de Grande Diffusion, ainsi que Monsieur Eric Garcia-Diaz, Directeur du Pôle Matériaux et Structures du Génie Civil. Je remercie aussi Monsieur Marc Vinches, animateur de l'équipe Modélisation des Matériaux, des Structures et des Hydrosystèmes pour ses conseils et ses relectures.

De même, je remercie pour son accueil Monsieur Eric Servat, Directeur du Laboratoire Hydrosociences Montpellier.

Je remercie vivement le SCHAPI qui est à l'origine de ces travaux et du cadre opérationnel dans lequel ils ont pu se dérouler, plus particulièrement Madame Caroline Wittwer, qui de plus examinera mon travail, et Monsieur Arthur Marchandise, qui ont été de bon conseil lors de plusieurs étapes clé des travaux. Je remercie aussi l'ensemble des partenaires du projet BVNE et du projet ANR FLASH pour les discussions constructives que nous avons pu avoir, en particulier avec Monsieur Pierre Roussel, qui a d'ailleurs participé à mes Comités de Thèse. Je remercie également Météo-France pour la fourniture de données météorologiques et encore une fois le SCHAPI, pour la fourniture de données hydrométéorologiques.

Je n'oublie pas le Syndicat Mixte du Galeizon et la Mairie de Cendras (30) qui ont fourni les données concernant le retour d'expérience de septembre 2002 sur ce bassin ; qu'ils en soient remerciés. De même, merci à l'Observatoire Hydrométéorologique Cévennes Vivarais (OHMCV) pour les données rendues disponibles pour les bassins non jaugés.

Merci à Pierre-Alain Ayrat pour son aide dans la recherche de données et ses conseils avisés. Merci aussi à Jean-François Didon-Lescot de l'UMR Espace pour ses conseils et orientations lors du début des travaux.

Je remercie l'ensemble de l'équipe Modélisation des Matériaux, des Structures et des Hydrosystèmes pour son accueil et sa bonne humeur. Merci à Monsieur Bernard Vayssade pour nos échanges sur l'hydrologie et les épisodes anciens, Line Kong-A-Siou, Audrey Bornancin-Plantier et Virgile Taver, doctorants ou post-doctorants de l'équipe.

Merci également aux stagiaires qui sont passés par le laboratoire et ont beaucoup apporté par leurs travaux et leur sympathie : Coralie Guillot, Kévin Cros et Yann Visserot. Merci aux élèves de l'Ecole des Mines d'Alès, qui au cours de leurs différentes missions en laboratoire ont apporté leur contribution à ces travaux.

Je remercie tous les membres du Laboratoire Hydrosociences Montpellier pour leur accueil chaque fois que cela a été nécessaire ainsi que pour leur sympathie, en particulier Mathieu Coustau qui a terminé son doctorat récemment.

Merci à Monsieur Dominique Bertin de la société Géonosis pour la création et le perfectionnement constant du logiciel RNF pro.

Plus personnellement, je remercie Veronica pour sa patience et son affection, chacune sans limite. Je remercie également mes parents qui m'ont toujours soutenu et judicieusement conseillé sur tous les plans. Je remercie aussi ma sœur Sophie ainsi que Tony Lo-Pinto, sans oublier Nathan et ses frères. Je remercie également ma grand-mère ainsi que mes oncles et tantes.

Je remercie mes collègues de Keraunos SAS, Messieurs David Dumas, Pierre Mahieu et Emmanuel Wesolek, pour leur bonne humeur mais aussi pour leur compréhension et leur patience face à une charge de travail parfois encombrante dans notre projet commun.

Je remercie tous mes amis pour leur bonne humeur et leur soutien et notamment : Alain et Virginie à Domancy, Cyrille à Montpellier, Emmanuel et Marion à Paris, Grégory à Chambéry, Mathieu et Lindsey en Australie, Nicolas et Cyrielle à Belfort, Thibaut à Paris...

## Table des matières

Introduction .....	13
1. Eléments de contexte et problématique .....	15
1.1 Caractérisation du territoire.....	16
1.1.1 Topographie et hydrographie.....	16
1.1.2 Géologie.....	20
1.1.3 Occupation des sols.....	22
1.1.4 Aspects démographiques et sociétaux .....	23
1.2 Aspects climatiques et hydrométéorologiques.....	25
1.2.1 Climatologie générale.....	25
1.2.2 Episodes pluvieux intenses.....	28
1.2.2.1 Situations à fortes précipitations sur la zone .....	28
1.2.2.2 Climatologie des épisodes pluvieux intenses .....	30
1.2.2.3 De la complexité de la relation climat - hydrologie.....	32
1.3 La Gestion du risque d'inondations .....	33
1.4 Le projet BVNE Gardons .....	36
1.5 Données hydrométéorologiques.....	39
1.5.1 Disponibilité et qualité des données de hauteur d'eau et de débit .....	39
1.5.1.1 Disponibilité.....	39
1.5.1.2 Qualité.....	41
1.5.2 Disponibilité et fiabilité des données de pluie.....	41
1.6 Position de la problématique.....	42
1.7 Conclusions .....	43
2. Dynamique hydrologique et modélisation des crues rapides.....	45
2.1 Mécanismes de la réponse hydrologique .....	46
2.1.1 Interface atmosphère-bassin : .....	46
2.1.1.1 Influence de la végétation.....	46
2.1.1.2 Influence des cultures .....	47
2.1.1.3 Influence des zones artificielles .....	48
2.1.2 Incidence directe de la pluie sur le bassin.....	49
2.1.1.1 Ruissellements de surface .....	49
2.1.1.2 Processus de sub-surface.....	51
2.1.1.3 Stockage et pertes .....	52

2.1.2	Écoulements.....	52
2.1.3	Conclusions.....	54
2.2	Vers la prévision des crues éclair .....	54
2.2.1	Modélisation hydrologique.....	54
2.2.1.1	Quelques définitions .....	54
2.2.1.2	Taxonomie des modèles hydrologiques .....	56
2.2.1.3	Qualités attendues d'un modèle hydrologique .....	57
2.2.1.4	Évaluer la performance d'un modèle.....	58
2.2.1.5	Conclusions .....	60
2.2.2	Mise en œuvre des modèles hydrologiques pour la simulation et la prévision.....	60
2.2.3	Mise en œuvre des modèles hydrologiques sur des bassins non jaugés .....	62
2.2.3.1	Procédé de « généralisation » .....	62
2.2.3.2	Limites de la démarche.....	64
2.3	Conclusions .....	64
3.	Les réseaux de neurones formels appliqués à la prévision .....	67
3.1	Généralités sur les réseaux de neurones formels .....	67
3.1.1	Le neurone .....	67
3.1.2	Réseaux de neurones.....	68
3.2	Propriétés fondamentales des réseaux de neurones.....	70
3.2.1	L'approximation universelle .....	70
3.2.2	La parcimonie.....	70
3.3	Apprentissage des réseaux de neurones .....	71
3.3.1	Minimisation de la fonction de coût.....	71
3.3.2	Problème du surajustement .....	71
3.3.3	Estimer l'erreur de généralisation : la validation croisée .....	73
3.3.4	Sélection de la complexité.....	74
3.3.5	La sélection de variables.....	75
3.4	Prévision par réseaux de neurones.....	76
3.4.1	Généralités.....	76
3.4.2	Application à l'hydrologie.....	77
3.4.2.1	Adaptation des modèles aux conditions locales.....	77
3.4.2.2	Le traitement de la question de la généralisation .....	78
3.5	Conclusions .....	78

4.	Prévision des crues du Gardon à Mialet.....	81
4.1	Présentation du bassin versant du Gardon à Mialet .....	82
4.2	Données disponibles .....	85
4.2.1	Extraction et traitement.....	85
4.2.1.1	Principes de base et définition d'un seuil de sélection des évènements.....	85
4.2.1.2	Base de données fournie dans le cadre de BVNE Gardons.....	86
4.2.1.3	Sélection des évènements.....	88
4.2.2	Description des évènements .....	89
4.2.2.1	Description générale.....	89
4.2.2.2	Précipitations : cumuls et intensités .....	90
4.2.2.3	Débits .....	95
4.2.2.4	Relation entre pluie et débit.....	97
4.3	Synthèse de prédicteurs neuronaux.....	99
4.3.1	Sélection de variables ou dimensionnement des fenêtres temporelles d'entrée.....	104
4.3.1.1	Méthode experte de sélection de variables.....	105
4.3.1.2	Méthode empirique de sélection de variables .....	107
4.3.2	Architecture optimale .....	111
4.3.3	Régularisation par l'arrêt précoce .....	112
4.4	Résultats en prévision en l'absence de prévisions de pluie.....	113
4.4.1	Modèle statique dirigé : cas du bassin bien jaugé .....	114
4.4.2	Modèle récurrent non dirigé : cas du bassin mal jaugé.....	117
4.4.2.1	Prévisions déterministes .....	118
4.4.2.2	Prévisions ensemblistes.....	119
4.4.3	Utilisation de l'humidité issue de la couche racinaire du modèle SIM.....	121
4.5	Conclusions .....	122
5.	Généralisation aux bassins non jaugés.....	125
5.1	Démarche proposée.....	126
5.2	Méthodologie développée.....	127
5.2.1	Rappel des principales contraintes liées à la problématique .....	127
5.2.2	Analyse des données et des prévisions brutes : vers une méthode de correction des débits prévus.....	128
5.2.2.1	Sélection et aperçu des bassins cibles .....	128
5.2.2.2	Sélection et description des évènements retenus .....	130
5.2.2.3	De l'importance de la dimension temporelle et des amplitudes .....	133

5.2.2.4	Synthèse.....	138
5.2.3	Méthode de correction des débits prévus .....	138
5.2.3.1	Formules empiriques de prédétermination des débits spécifiques de pointe .....	138
5.2.3.2	Etablissement d'une règle de prédétermination des débits spécifiques importants pour les Cévennes .....	140
5.2.3.3	Etablissement d'une règle de détermination de coefficient correctif des débits prévus	141
5.3	Application aux bassins cibles.....	142
5.3.1	Le ruisseau du Lyonnais.....	143
5.3.1.1	Présentation rapide du bassin.....	143
5.3.1.2	Evènements et résultats .....	144
5.3.1.3	Analyse des résultats obtenus .....	145
5.3.2	L'Alzon.....	145
5.3.2.1	Présentation rapide du bassin.....	145
5.3.2.2	Evènements et résultats.....	146
5.3.2.3	Analyse des résultats obtenus .....	147
5.3.3	La Salandre .....	148
5.3.3.1	Présentation rapide du bassin.....	148
5.3.3.2	Evènements et résultats.....	148
5.3.3.3	Analyse des résultats obtenus .....	150
5.3.4	L'Amous .....	150
5.3.4.1	Présentation rapide du bassin.....	150
5.3.4.2	Evènements et résultats.....	151
5.3.4.3	Analyse des résultats obtenus .....	152
5.3.5	L'Auzonnet.....	152
5.3.5.1	Présentation rapide du bassin.....	152
5.3.5.2	Evènements et résultats.....	153
5.3.5.3	Analyse des résultats obtenus .....	155
5.3.6	Le Galeizon amont .....	155
5.3.6.1	Présentation rapide du bassin.....	155
5.3.6.2	Evènements et résultats.....	156
5.3.6.3	Analyse des résultats obtenus .....	157
5.3.7	La Ganière .....	157
5.3.7.1	Présentation rapide du bassin.....	157

5.3.7.2	Evènements et résultats .....	158
5.3.7.3	Analyse des résultats .....	160
5.3.8	L'Avène.....	161
5.3.8.1	Description rapide du bassin .....	161
5.3.8.2	Evènements et résultats .....	162
5.3.8.3	Analyse des résultats .....	163
5.3.9	Le Galeizon à Cendras .....	163
5.3.9.1	Description rapide du bassin .....	163
5.3.9.2	Evènements et résultats .....	164
5.3.9.3	Analyse des résultats .....	164
5.3.10	Le Gardon à Saumane.....	165
5.3.10.1	Description rapide du bassin.....	165
5.3.10.2	Evènements et résultats .....	165
5.3.10.3	Analyse des résultats .....	167
5.3.11	La Baume.....	167
5.3.11.1	Description rapide du bassin.....	167
5.3.11.2	Evènements et résultats .....	168
5.3.11.3	Analyse des résultats .....	169
5.3.12	L'Ardèche.....	170
5.3.12.1	Description rapide du bassin.....	170
5.3.12.2	Evènements et résultats .....	171
5.3.12.3	Analyse des résultats .....	172
5.3.13	Le Chassezac .....	172
5.3.13.1	Présentation rapide du bassin .....	172
5.3.13.2	Evènements et résultats .....	173
5.3.13.3	Analyse des résultats .....	175
5.3.14	Le Gardon à Anduze .....	175
5.3.14.1	Description rapide du bassin.....	175
5.3.14.2	Evènements et résultats .....	176
5.3.14.3	Analyse des résultats .....	178
5.4	Conclusions .....	178
5.4.2	Analyse générale .....	179
5.4.3	Apports et limites.....	181

Conclusion générale.....	183
Références bibliographiques.....	187
Annexes.....	195
Liste des figures.....	307
Liste des tableaux.....	313
Liste des photos.....	315
Résumé.....	318
Summary.....	318

## Introduction

Le risque hydrométéorologique est l'une des principales menaces naturelles en France et plus particulièrement dans les régions méditerranéennes. Ces régions ont été durement éprouvées ces dernières décennies tant sur le plan humain que matériel : catastrophe de Nîmes en 1988, de Vaison-la-Romaine en 1992, de l'Aude en 1999, du Gard en 2002 ou encore du Var en 2010... Ces deux dernières catastrophes ont chacune causé plus de 20 victimes et un milliard d'euros de dégâts. A chaque fois, les zones impactées sont soumises à des crues rapides (moins de 10 heures de temps de réponse) et d'ampleur importante que l'on appelle *crues éclair*. Ces réactions hydrologiques extraordinaires sont liées à des épisodes pluvieux intenses et souvent très localisés, par nature difficiles à prévoir.

Face à ces enjeux, l'hydrologie et la météorologie se doivent de développer des solutions pour permettre la prévention de ces risques et la protection des populations. Cette protection passe notamment par l'alerte précoce, pour laquelle les prévisions hydrologiques sont indispensables. Il s'agit là d'un vrai enjeu scientifique étant donné que les pluies qui sont à l'origine de ces crues sont très difficilement prévisibles et que les temps de réponse des systèmes hydrologiques concernés sont très brefs. De plus, les processus physiques qui relient pluie et débit sont difficiles à quantifier et lorsqu'ils le sont, difficiles à généraliser aux échelles de temps et d'espace nécessaires, ce qui ne facilite pas leur modélisation.

C'est au regard de ces éléments que le projet BVNE (Bassin Versant Numérique Expérimental) Gardons a été initié par le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) en 2006, pour contribuer à l'amélioration des prévisions sur des bassins capables de ce type de réactions. L'objectif de la démarche est d'alimenter la carte de vigilance hydrologique *vigicrues* disponible sur internet ([www.vigicrues.gouv.fr](http://www.vigicrues.gouv.fr)) avec des prévisions à la fiabilité optimisée.

C'est dans ce contexte que s'inscrivent les travaux de ce mémoire : leur finalité est de développer des modèles de prévision hydrologique pour les crues éclair permettant d'améliorer l'anticipation des inondations à des fins de gestion des situations critiques et d'information préventive des populations. Cette anticipation doit fonctionner dans différents environnements de mesures hydrométéorologiques : sur les bassins où l'on dispose de chroniques de débit et d'observations en temps réels (bassins jaugés), sur les bassins où l'on dispose de chroniques de débit mais pas d'observations fiables en temps réel (bassins mal jaugés) et enfin et surtout, sur les bassins où l'on ne dispose ni de chroniques de débit, ni d'observations en temps réel (bassins non jaugés).

Le défi relevé dans ce travail de thèse a été de réaliser des prédicteurs statistiques opérant sur tous ces types de bassins, qu'ils disposent de mesures de débits ou en soient complètement dépourvus. Pour atteindre cet objectif, *a priori* paradoxal, nous commencerons par étudier le contexte géographique en caractérisant aussi finement que possible la zone d'étude, située dans le massif des Cévennes, réputé pour ses crues dévastatrices générées par des épisodes pluvieux extrêmes. Les aspects climatiques et hydrologiques seront abordés et nous pourrions, au regard des mesures prises pour mitiger les risques d'inondations, comprendre à quel point la prévision des crues occupe une position centrale dans leur gestion. Le projet BVNE et ses motivations seront présentés en détail et des considérations sur la qualité et la disponibilité des données hydrométéorologiques liées à ces questions seront développées. Cette approche nous permettra de formuler clairement la problématique en vue des chapitres suivants.

Par la suite, nous nous attacherons à décrire les processus à l'origine de ces réactions hydrologiques, depuis l'interface entre atmosphère et bassin jusqu'à l'exutoire. Nous présenterons alors quelles solutions de modélisation des crues rapides pour la prévision ont été mises en place, que ce soit sur des bassins jaugés ou non jaugés. L'état de la connaissance des processus hydrologiques nous conduira à justifier l'utilisation de l'apprentissage statistique par réseaux de neurones pour la résolution de notre problématique.

Ce sont justement les réseaux de neurones qui seront présentés par la suite. Leurs propriétés fondamentales et leur capacité d'apprentissage seront mises en perspective avec les difficultés liées au surajustement et aux méthodes existantes pour l'éviter. Les applications de cette famille de modèles à l'hydrologie seront enfin recensées.

Nous appliquerons alors ces modèles à un bassin situé au cœur de la zone d'étude : le bassin du Gardon à Mialet. Après une sélection rigoureuse de variables, nous mettrons en œuvre les méthodes de limitation de la complexité décrites précédemment pour élaborer plusieurs types de modèles : statique dirigé par les débits observés (cas du bassin bien jaugé), récurrent non dirigé déterministe (cas du bassin mal jaugé) ou récurrent non dirigé ensembliste. Nous pourrons ainsi présenter des prévisions, jusqu'à l'horizon du temps de réponse du bassin, sur différents événements, depuis le plus intense de la base de données jusqu'à des événements complexes, en l'absence totale d'hypothèse sur les pluies futures. La robustesse de ces prévisions sera évaluée en vue de la généralisation aux bassins non jaugés.

C'est bien cette dernière étape qui fera l'objet du chapitre 5 dans lequel la méthode de généralisation sera élaborée et décrite, les bassins visés par la généralisation présentés et les résultats sur 14 bassins et 35 événements détaillés. Nous pourrons alors tirer des conclusions sur le fonctionnement de cette méthode de généralisation tant par rapport à la superficie des bassins de test (de 4km<sup>2</sup> à 545km<sup>2</sup>), que par la provenance de données pluviométriques (radars ou pluviomètres), ou par le type de fonctionnement hydrologique du bassin.

Enfin nous verrons quelles sont les perspectives de recherche révélées par les résultats de ces travaux, des perspectives nombreuses, passionnantes et qui ouvrent la voie de la prévision opérationnelle statistique pour les bassins versants non jaugés.

## 1. Eléments de contexte et problématique

Ce chapitre va s'attacher à définir le contexte et les motivations des présents travaux de recherche. En premier lieu nous présenterons l'environnement naturel de la zone d'étude, caractérisée par une hétérogénéité de relief assez marquée entre les plaines et les garrigues, les vallées cévenoles et les hauts plateaux lozériens ou ardéchois. La démographie présente également une grande variabilité puisque quelques villes d'importance (comme Alès ou Aubenas) voisinent avec des territoires quasi inhabités. L'occupation du territoire sera ensuite mise en parallèle avec les aléas hydrométéorologiques qui le menacent. En effet le Gard, la Lozère et l'Ardèche sont soumis à des épisodes de pluies diluviennes se déclenchant souvent sur les contreforts sud du relief. Ces épisodes de pluies intenses donnent lieu à des crues aussi spectaculaires que dévastatrices, conduisant régulièrement à des décès et à des dégâts économiques parfois supérieurs au milliard d'euros pour un seul évènement. L'importance des enjeux économiques et humains ainsi révélée nous conduira à évoquer les mesures de gestion du risque d'inondation et nous permettra de mettre en évidence la place centrale qu'occupe la prévision hydrologique dans leur amélioration.

Face à ce risque inondation, les pouvoirs publics se sont organisés pour répondre, tant au niveau de l'alerte, disponible sur internet par l'intermédiaire du site *vigicrues*, qu'au niveau scientifique. Ainsi, le projet BVNE (Bassin Versant Numérique Expérimental) du SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations), qui vise à améliorer les prévisions hydrologiques à des fins d'information préventive et d'aide à la décision pour les autorités, sera présenté.

Même si l'effort des pouvoirs publics est tout à fait remarquable depuis de nombreuses années, proposant ainsi une veille sur le risque crue de plus de 20 000 km de cours d'eau en France, nous verrons en dernière partie que tous les cours d'eau potentiellement à risque de notre territoire d'étude ne sont pas pourvus d'une station de mesure de débit. Se pose alors le problème de la réalisation de systèmes d'alerte pour ces nombreux ruisseaux et rivières susceptibles de générer des crues dévastatrices. Un élément déterminant dans le contexte de ces travaux est donc la disponibilité et la qualité des données de pluie, de hauteur d'eau et de débit.

L'ensemble de ces questions est à l'origine de la problématique qui sera finalement développée et qui porte sur la prévision des crues éclair dans la zone d'étude, quelle que soit la qualité et la quantité des données disponibles.

## 1.1 Caractérisation du territoire

La zone d'étude se situe dans le sud de la France et s'étend sur deux régions (Languedoc-Roussillon et Rhône-Alpes) et trois départements (Gard, Lozère, Ardèche). Elle constitue une grande partie du massif des Cévennes et du Vivarais, ainsi que la partie septentrionale de la plaine du Languedoc (figure 1). Il s'agit d'un territoire à dominante montagneuse et à fort tropisme méditerranéen.

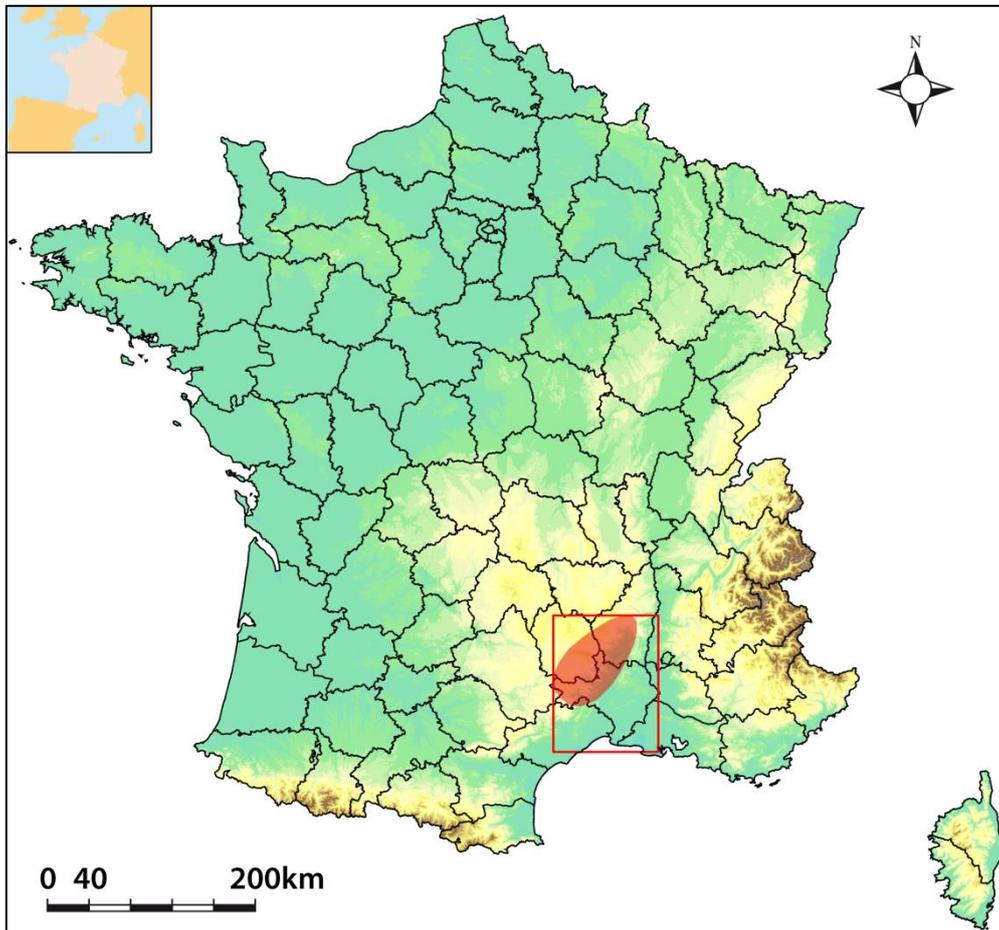


Figure 1: localisation de la zone d'étude en France (en rouge). Le cadre rouge représente l'emprise approximative des cartes suivantes. Altitude d'après IGN BD ALTI 250, départements d'après IGN BD Geofla.

### 1.1.1 Topographie et hydrographie

La zone d'étude est centrée sur l'escarpement cévenol qui assure la transition entre les plateaux du Massif-Central et les plaines du Languedoc.

Les sommets cévenols constituent la bordure méridionale du Massif-Central. La plupart d'entre eux dépassent les 1000m, certains dépassent 1500m (Mont-Lozère 1699m, Mont-Aigoual 1567m, Serre-de-la-Croix-de-Bauzon 1538m...). Si les versants nord sont marqués par des pentes relativement douces, avec les plateaux ardéchois, la Margeride ou les Causses, les versants sud sont nettement plus abrupts et la transition entre les crêtes et le nord de la plaine du Languedoc se fait en seulement une cinquantaine de kilomètres. C'est donc au déboucher des vallées cévenoles que le relief s'adoucit, entre plateaux et collines des garrigues (figure 2).

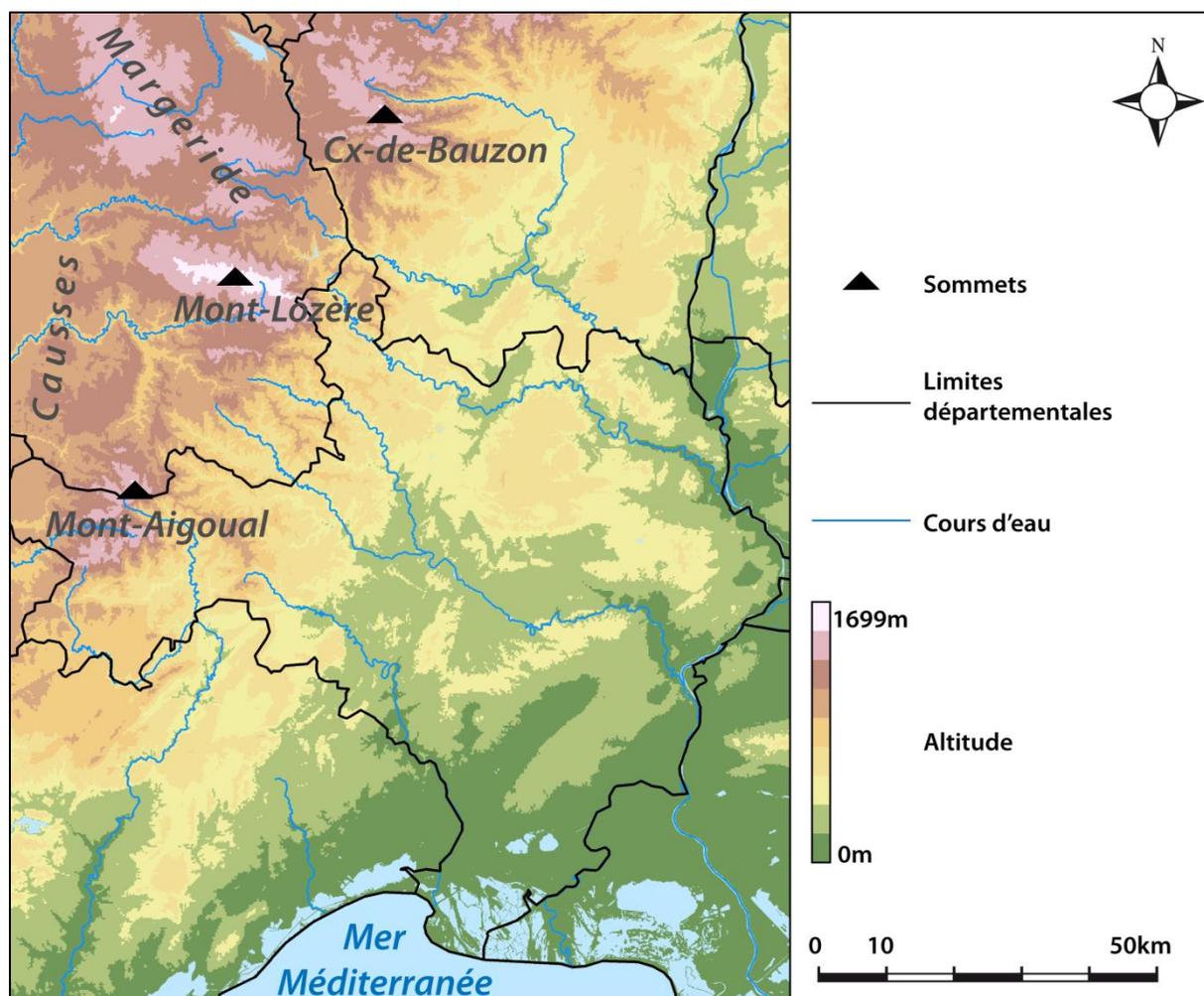


Figure 2: topographie de la zone d'étude. Altitudes d'après ASTER GDEM, un produit of METII and NASA, départements d'après IGN BD Geofla, cours d'eau d'après IGN BD Carthage.

Cette rapide transition entre montagne et plaine se ressent à l'observation des pentes de versant sur la zone visée (figure 3). On constate en effet que sur les versants nord, les pentes sont relativement douces, à l'exception de quelques profondes vallées (Gorges du Tarn, de la Jonte, Vallée du Lot...) et bordures de Causse. En revanche, en versants sud, les pentes sont très fortes jusqu'aux plaines, où les gorges du Gardon, de la Cèze ou de l'Ardèche concentrent la grande majorité des pentes fortes.

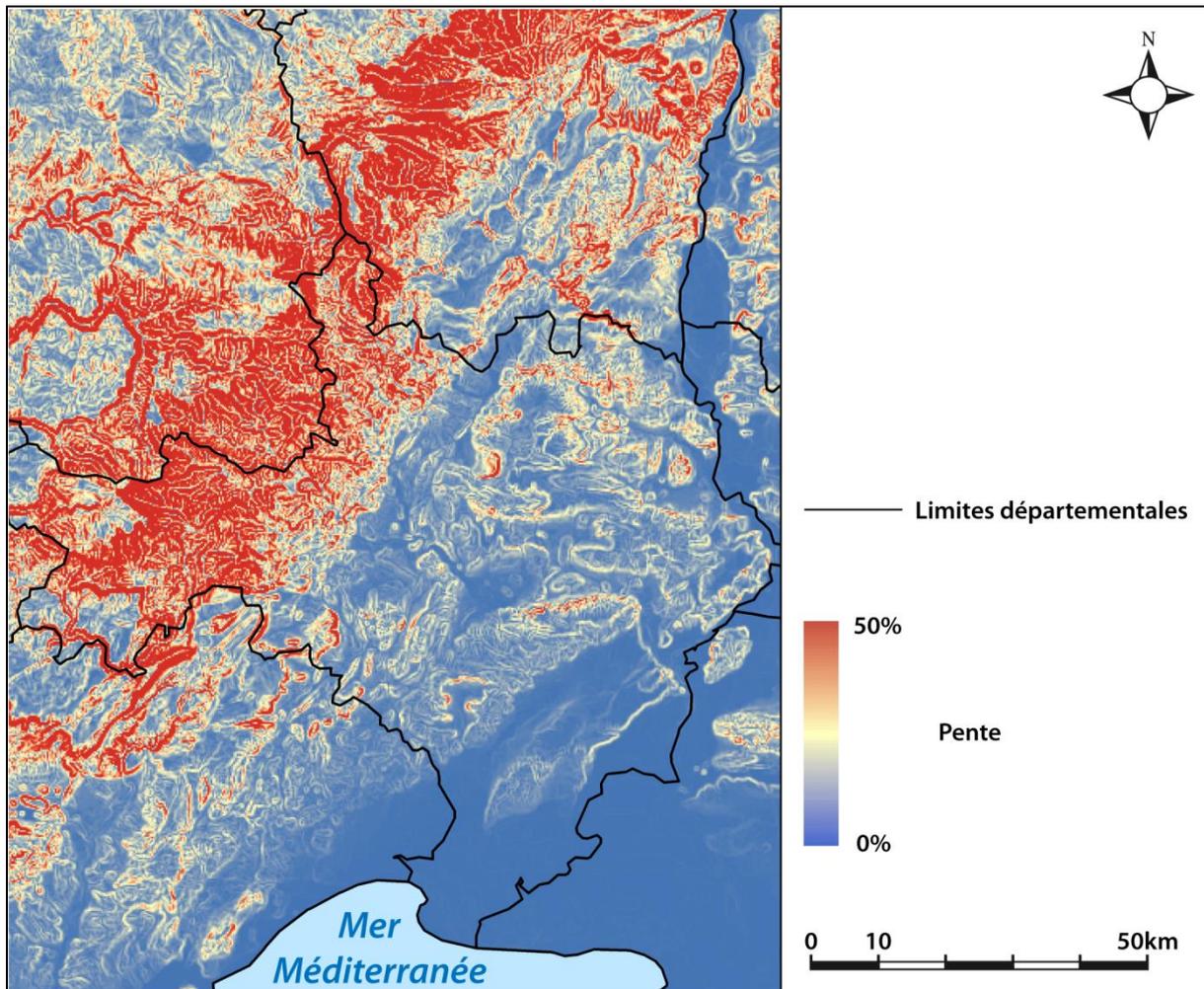


Figure 3: pentes de la zone d'étude. Altitudes utilisées pour le calcul des pentes : IGN BD ALTI 250, limites départementales d'après IGN BD Geofla.

Concernant l'hydrographie, on distingue essentiellement deux zones marquées par des tracés différents : les Cévennes et la plaine.

Les versants méditerranéens du massif cévenol sont marqués par une forte densité de drainage qui témoigne du faible volume des circulations souterraines. Les pentes fortes induisent par ailleurs des tracés de cours d'eau relativement linéaires (figure 4).

Des barrages ont été implantés sur le Gardon et la Cèze (photos 1 et 2) à des fins d'écrêtement des crues (sur le Gardon) éventuellement combinées à des fonctions de soutien d'étiage (sur la Cèze). Sur le Chassezac et l'Ardèche, il s'agit pour l'essentiel de barrages hydroélectriques et de soutien d'étiage.



Photo 1: à gauche : barrage écrêteur de crues et de soutien d'étiage de Sénéchas, sur la Cèze, à sa confluence avec l'Homol ;  
Photo 2: à droite : barrage écrêteur de crues de Sainte-Cécile-d'Andorge, sur le Gardon d'Alès. Photos : Artigue, 2007.

A l'inverse, en plaine, les cours d'eau évoluent dans un relief moins tranché, plus complexe. Ils donc sont plus sinueux, notamment lorsqu'ils sont dirigés par la tectonique. C'est le cas pour la Cèze au niveau de l'escarpement de la faille de Barjac, pour l'Amous avec la faille des Cévennes ou encore pour l'Alauzène avec les plissements pyrénéens des garrigues du Mont Bouquet.

La densité de drainage, nettement moins importante, peut devenir particulièrement faible à l'aplomb des plateaux karstiques, alors que les cours principaux serpentent au fond des gorges incisées dans ces plateaux (Gorges du Gardon, de la Cèze et de l'Ardèche) (figure 4).

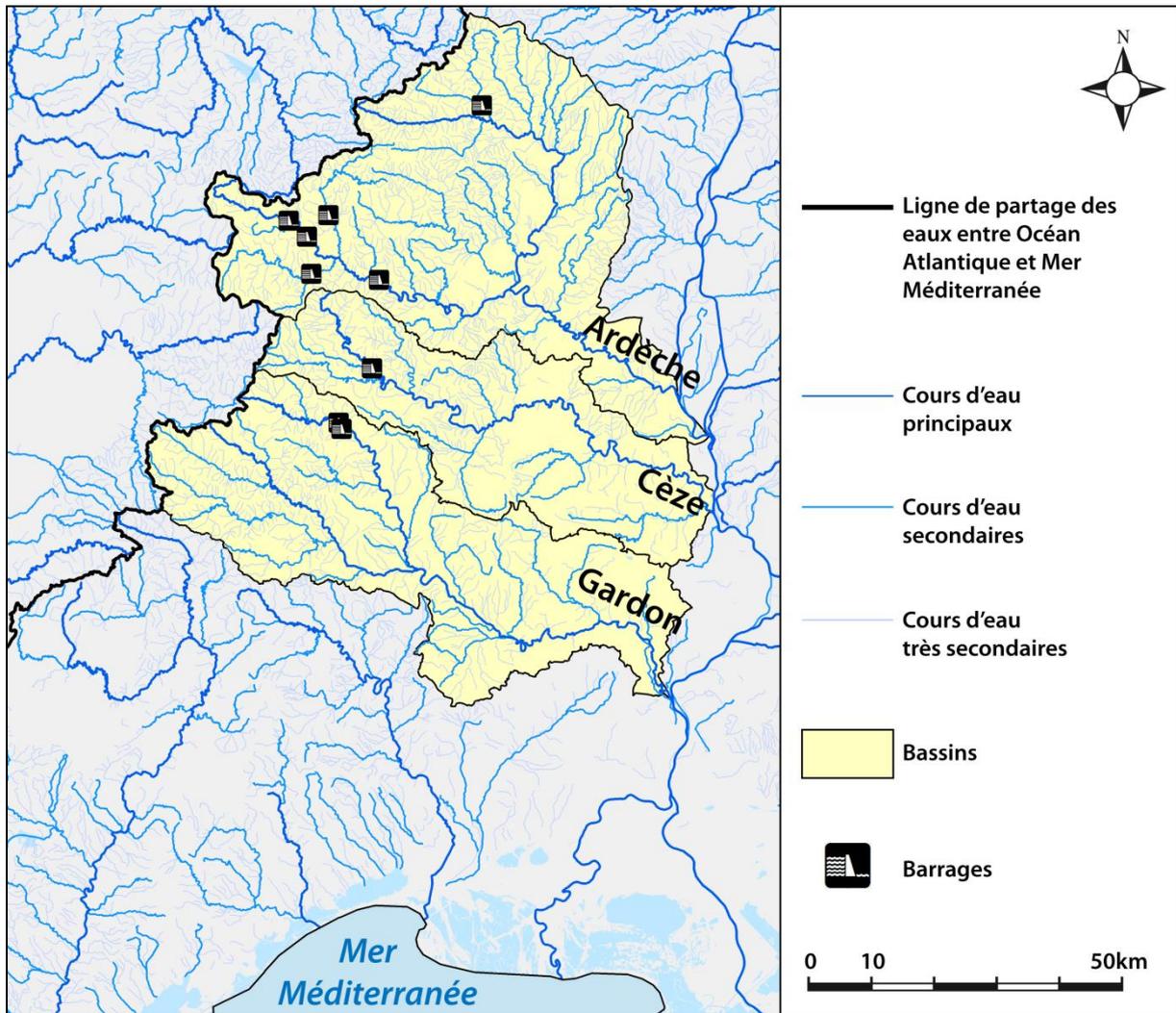


Figure 4: hydrographie de la zone d'étude. Cours d'eau, limites de bassin et ligne de partage des eaux d'après IGN BD Carthage. Les cours d'eau principaux sont de classification 1 et 2, les cours d'eau secondaires 3, 4 et 5 et les cours d'eau très secondaires de classification supérieure (cf. Tableau 1).

### 1.1.2 Géologie

L'histoire géologique de la zone est structurée selon quelques phases aisément identifiables. Dans les Cévennes, la présence des terrains métamorphiques (micaschistes et paragneiss essentiellement) datés de 340 Millions d'années (Ma) (Faure, 1995) témoigne de l'orogénèse hercynienne dans la région. Il en est de même pour la mise en place tardi-orogénique de roches magmatiques, de type granitique, au niveau des Massifs du Mont-Aigoual – Saint-Guiral – Liron (au sud-ouest), du Mont-Lozère (au nord-ouest de la zone) ou du Tanargue dans la partie ardéchoise. La fin de cette orogénèse est marquée par l'activation d'un grand décrochement tardi-hercynien entre Alès et l'est du Massif du Mont-Lozère (faille de Villefort-Soustelle) puis par les dépôts continentaux du Permien (notamment à l'est de la Montagne-Noire) dont la rareté témoigne d'une phase essentiellement érosive.

Cette pénélplanation aboutira à l'immersion de la zone, d'abord partielle au Trias, puis complète dès le Lias terminal. Les dépôts marins se succèdent alors jusqu'au Crétacé, à la fin duquel débute l'orogénèse pyrénéenne.

Cette dernière ne concernera que très marginalement les Cévennes, la faille des Cévennes accommodant la déformation d'axe sud-ouest - nord-est assez efficacement en jouant en décrochement sénestre normal.

C'est à l'Eocène que cette phase compressive connaîtra son paroxysme. On y observe de nombreuses configurations de dépôts continentaux syntectoniques et l'activation de décrochements sénestres de direction sub-cévenole. Les vastes plis (Mont-Bouquet par exemple) dans les séries du Crétacé témoignent eux aussi de cette importante compression.

L'ouverture du bassin du sud-est en liaison avec le rifting généralisé en Europe de l'ouest, à l'Oligocène, fera rejouer nombre de ces décrochements en faille normale et ouvrira des bassins remplis de dépôts oligocènes, avec parfois des accommodations importantes, comme dans le graben d'Alès.

La transgression miocène semble ne pas avoir pénétré trop loin dans les terres. Une phase de volcanisme de type strombolien ou phréato-magmatique concernera l'Ardèche à cette période (entre -5 et -10Ma). L'orogénèse alpine a, quant à elle, eu des effets surtout indirects, en provoquant des mouvements verticaux assez importants sur la bordure cévenole.

Enfin, les importantes variations climatiques du Quaternaire se ressentent en de nombreux points : modelé périglaciaire et reprise des creusements karstiques, entamés pour l'essentiel à la fin du Crétacé, en sont les principaux marqueurs (figure 5).

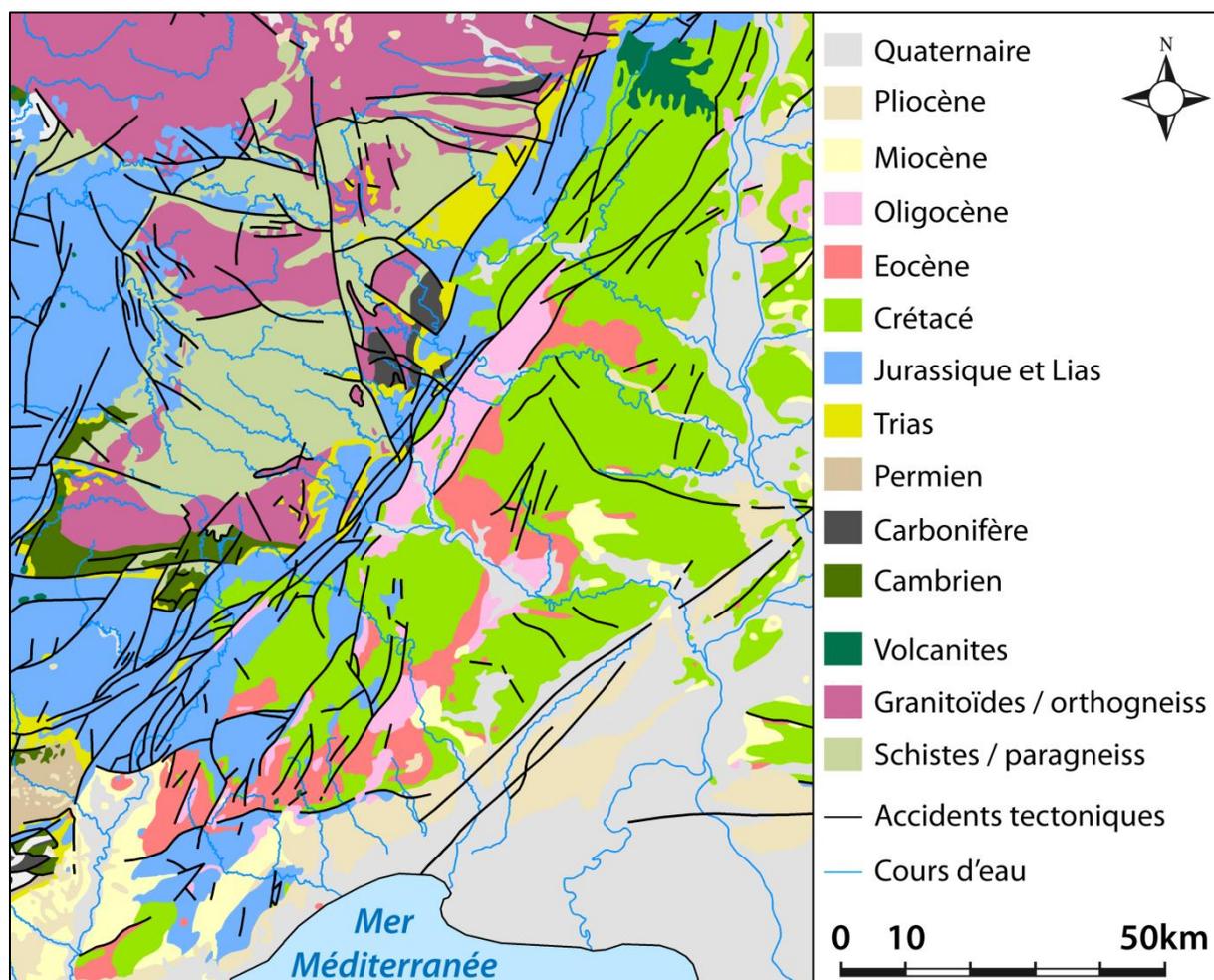


Figure 5: carte géologique simplifiée de la zone d'étude. Géologie d'après cartes géologiques de la France 1/50 000, 1/250 000 et 1/1 000 000, BRGM, feuilles diverses ; cours d'eau d'après IGN BD Carthage.

D'un point de vue lithologique, on distingue deux grands ensembles :

- La plaine languedocienne et la bordure cévenole : les terrains y sont essentiellement carbonatés et argileux (calcaires, marnes et argiles). Les séries carbonatées marines du Crétacé et du Jurassique sont fréquemment karstiques, en particulier dans le Crétacé moyen (Barrémien à faciès Urgonien) et dans le jurassique supérieur (Kimméridgien et Thitonien). Elles présentent donc une porosité globale très importante et à l'affleurement, une perméabilité souvent élevée. Ces séries compétentes sont entrecoupées de bassins où les dépôts tendres sont nettement plus imperméables (photo 3).
- Les Cévennes : les terrains sont majoritairement cristallins, réputés globalement imperméables et peu poreux. Dans le domaine métamorphique, les schistes et les paragneiss présentent un horizon d'altération parfois important alors que les granites et gneiss sont, dans les rares zones presque planes, profondément arénisés. D'une manière générale, une fois passées les altérations de surface, le terrain est effectivement peu poreux et bien souvent imperméable (photo 4).



Photo 3: à gauche : terrains karstiques dans les gorges de la Cèze, sur la commune de Méjannes-le-Clap (30), photo Artigue, 2007

Photo 4: à droite : affleurement de micaschistes au cœur des Cévennes, sur la commune de Saint-André-de-Valborgne (48). Photos : Artigue, 2007.

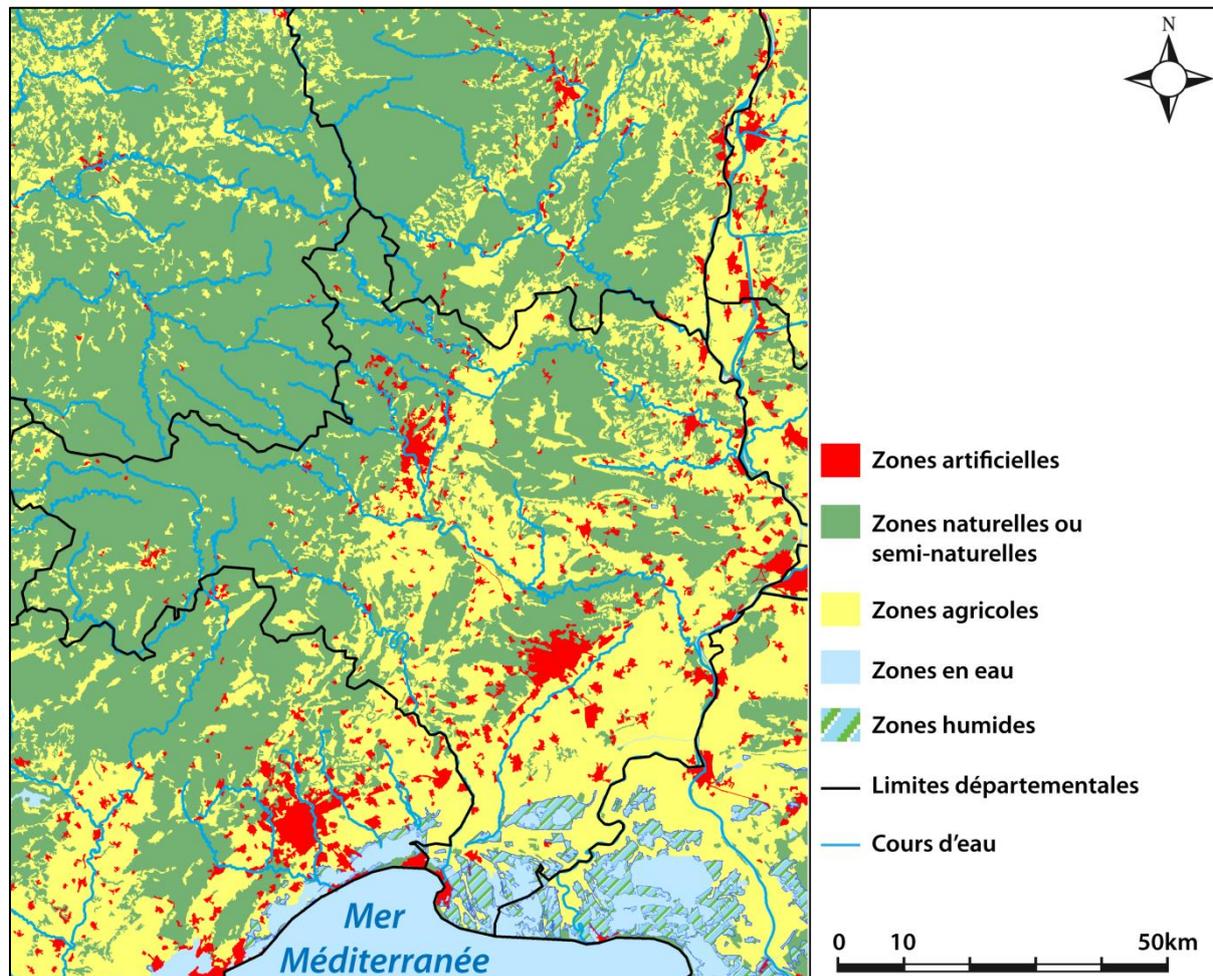
### 1.1.3 Occupation des sols

L'environnement de la zone d'étude est essentiellement rural. La majorité des occupations de sols se divise entre milieux naturels et agricoles.

La plaine languedocienne septentrionale ainsi que les fonds de vallée du piémont sont des territoires agricoles alors que les versants cévenols sont essentiellement forestiers, dédiés à un agropastoralisme peu dense ou semi-naturel.

Les espaces artificialisés sont pour la plupart concentrés autour des agglomérations d'Alès et d'Aubenas. On les retrouve plus ponctuellement en fonds de vallées, là où sont implantés quelques grands villages.

On notera que les milieux humides (zones en eau ou zones humides) sont presque entièrement absents de la zone d'étude. Les retenues des barrages sur le Gardon d'Alès, le Chassezac et la Cèze constituent en effet les seules étendues d'eau, d'une surface réduite (figure 6).



#### 1.1.4 Aspects démographiques et sociétaux

La population se répartit inégalement sur ce territoire accidenté (figure 7). La proximité de l'eau semble avoir été, comme c'est souvent le cas, un facteur déterminant dans l'implantation des principaux centres de densité.

Logiquement, comme pour l'occupation des sols, ce sont les agglomérations d'Aubenas et d'Alès qui concentrent la majorité de la population. Sortie de ces deux bassins de vie en piémont, la densité est moindre dans la plaine. Elle devient très faible en Cévennes où seules quelques bourgades rassemblent plus d'un millier d'habitants.

C'est d'ailleurs en grande partie en piémont que l'on rencontre des densités supérieures à la moyenne nationale de 97 habitants par kilomètre carré.

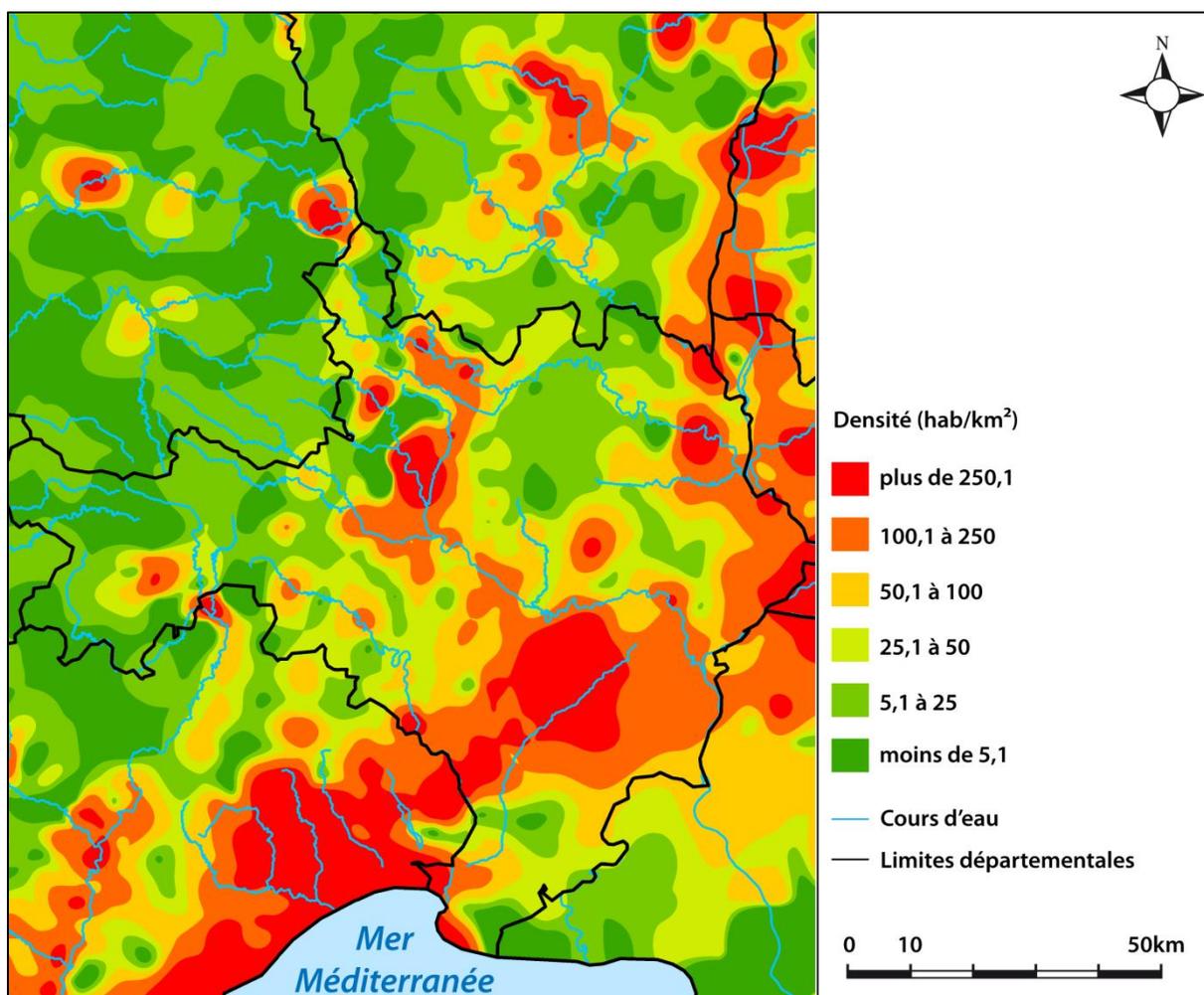
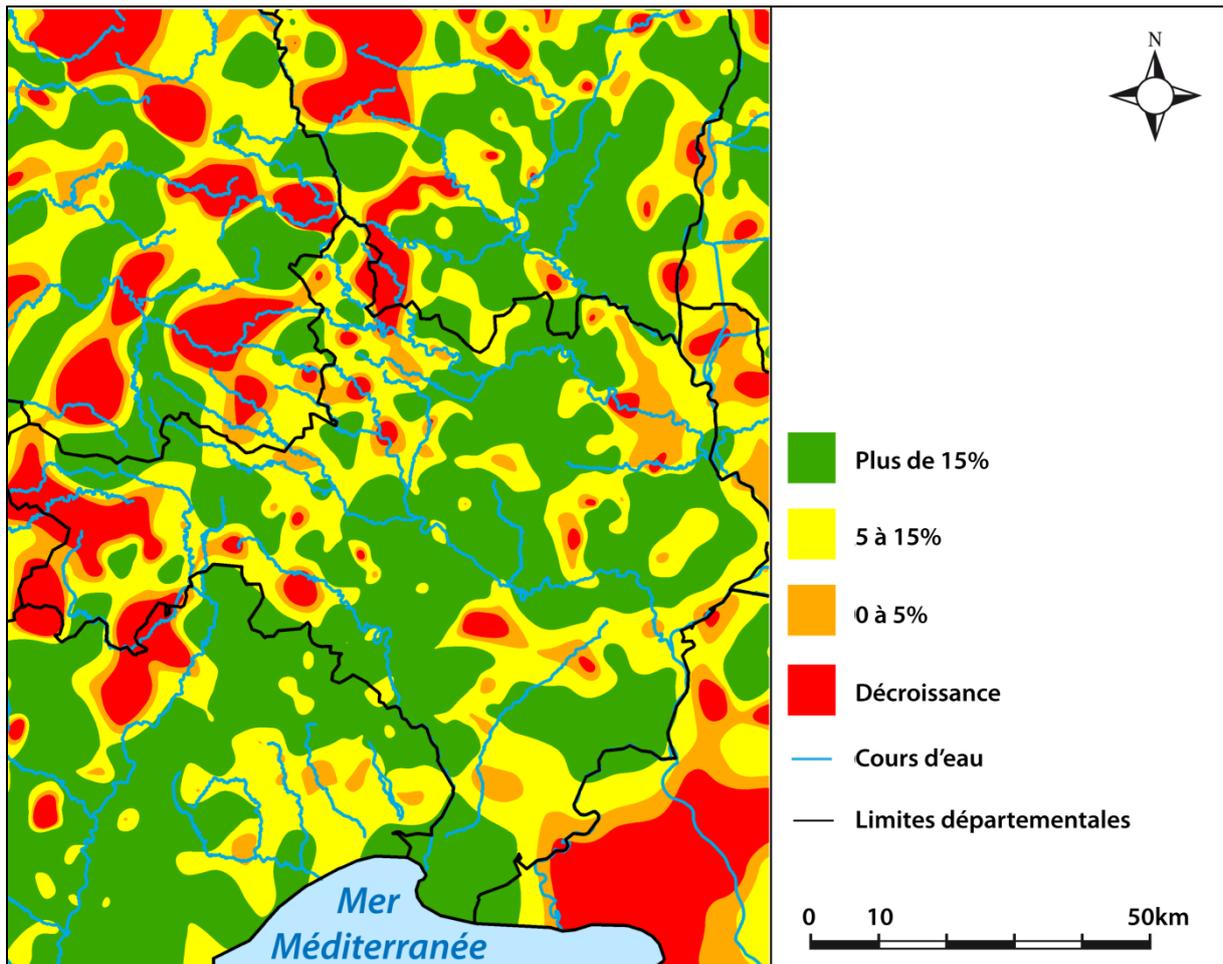


Figure 7: carte de densité de la population sur la zone d'étude. Lissage de type « spline » à partir des données communales. Densité d'après IGN RGC 2010 ; limites départementales d'après IGN BD Geofla ; cours d'eau d'après IGN BD Carthage.

Globalement, la zone est attractive puisque les départements de l'Ardèche, du Gard et de la Lozère ont respectivement vu leur population croître de 0,95%, 1,21% et 0,51% en moyenne et par an entre 1999 et 2008 selon les chiffres de l'INSEE. La croissance démographique est donc, au pire, quasi-égale à la moyenne française (Lozère), plus souvent supérieure.

Cependant, on remarque aussi que cette croissance est inégalement répartie sur le territoire concerné (figure 8). Les régions de relief semblent nettement moins attractives que la plaine où, même si certains centres urbains se dépeuplent (certainement par développement de l'habitat péri-urbain), la croissance démographique est généralement forte.



## 1.2 Aspects climatiques et hydrométéorologiques

### 1.2.1 Climatologie générale

Une rapide analyse des normales de quelques stations de la zone fait clairement apparaître trois ensembles climatiques relativement bien définis :

- la plaine et les versants sud des Cévennes,
- les crêtes cévenoles,
- les plateaux, en retrait des crêtes.

Les deux premiers ensembles constituent le principal théâtre de fréquents épisodes pluvieux intenses qui seront décrits dans la section 1.2.2. Le dernier ensemble marque la fin de l'influence méditerranéenne, tant sur les extrêmes que sur les moyennes, thermiques ou pluviométriques.

La transition entre les zones d'influence méditerranéenne dominante, et celles presque dépourvues de cette influence, est floue et dépend des situations météorologiques (Miquel, 2001). Le découpage schématique choisi ici privilégie la variabilité inter-mensuelle de précipitations, une variabilité forte étant assez caractéristique d'un climat méditerranéen. Par ailleurs, cette variabilité témoigne souvent de la survenue fréquente d'épisodes pluvieux

intenses pendant les mois les plus humides et de longues sécheresses pendant les mois les plus secs. Ces deux éléments sont des marqueurs reconnus du climat méditerranéen.

Les trois ensembles climatiques sont décrits ici :

- Le premier ensemble concerne la partie septentrionale de la plaine du Languedoc et les contreforts cévenols. A l'instar de la station de Saint-Christol-lès-Alès (30) (125m) dont les températures et les précipitations moyennes sont présentées dans la figure 9, on y observe des températures chaudes en été et plutôt douces en hiver, surtout pour les maximales, n'excluant cependant pas des gelées fréquentes et parfois fortes. Le régime pluviométrique est particulièrement irrégulier avec un rapport voisin de quatre entre le mois le plus arrosé (octobre avec 150mm à Saint-Christol) et le moins arrosé (juillet avec 40mm à Saint-Christol). Les pluies sont abondantes (environ 1000mm par an) et l'ensoleillement parmi les plus élevés de France (environ 2300h par an). Ce climat est donc caractéristique d'un arrière-pays méditerranéen.

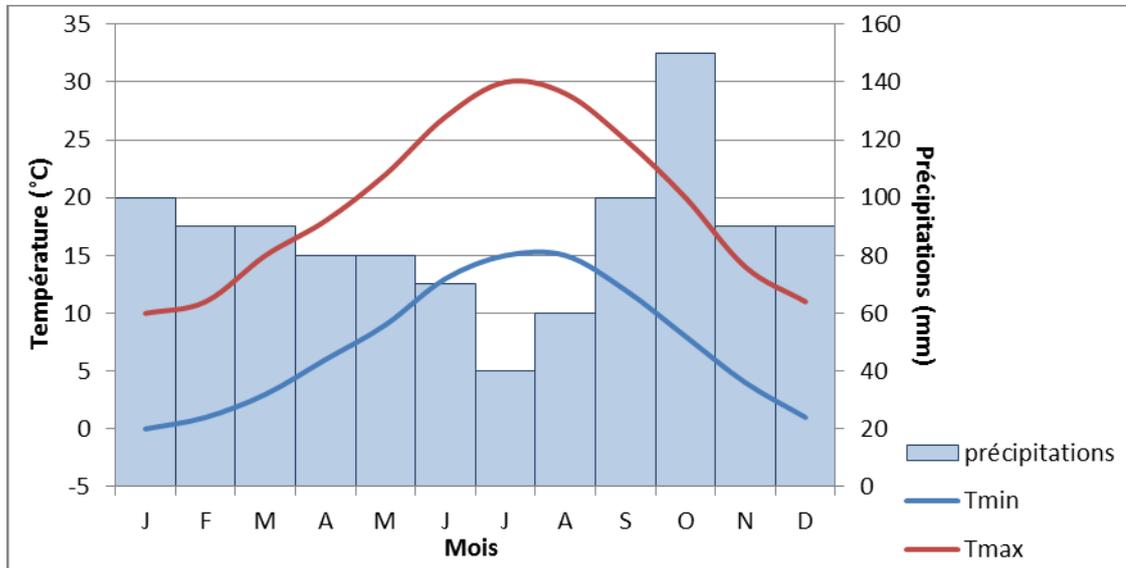


Figure 9: climatologie simplifiée des températures et des précipitations à la station de Saint-Christol-lez-Alès (30) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990).

En remontant les pentes des Cévennes ou en s'enfonçant dans les vallées, les températures fléchissent de quelques degrés et les précipitations deviennent progressivement plus abondantes. Par exemple, il tombe annuellement presque 500mm de plus à Saint-Jean-du-Gard (30) qu'à Saint-Christol-lez-Alès, avec une moyenne de températures déjà légèrement inférieure (un demi degré de moins).

- Le second ensemble est représenté par les sommets cévenols. Les conditions y sont la plupart du temps très rudes avec, par exemple, pour la station du Mont-Aigoual (Valleraugue, 30), en moyenne et par an, 2280mm de pluie, 276 jours de vent violent, 170 jours de pluie, 68 jours de neige et 149 jours de gelée sur la période 1951-1980 (Kessler et Chambraud, 1990). Les températures moyennes y sont froides et le record absolu de température maximale n'atteint même pas 29°C au Mont-Aigoual. On conserve toutefois des caractéristiques méditerranéennes dans l'ensoleillement plutôt important pour une station de montagne (en moyenne 5h40 par jour) ainsi que dans la répartition toujours irrégulière des précipitations (figure 10).

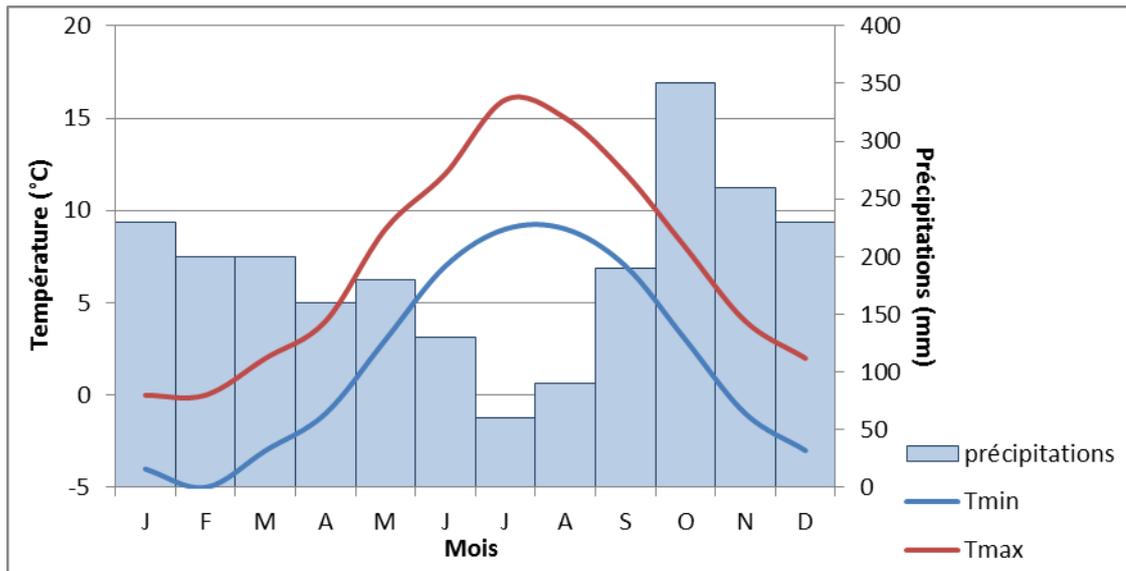


Figure 10: climatologie simplifiée des températures et des précipitations à la station du Mont-Aigoual (Valleraugue, 30) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990).

Les trois grands massifs de la zone, à savoir le Mont-Aigoual, le Mont-Lozère et le Mont-Tanargue, et les crêtes qui les relient, connaissent tous trois des conditions similaires à celles décrites dans le paragraphe précédent. Il s'agit là de l'une des zones les plus arrosées de France. Les cumuls annuels atteignent par exemple 2000mm à Saint-Maurice-de-Ventalon (48) à proximité du Mont-Lozère (contreforts de la Montagne du Bougès et du Mont-Lozère) ou encore 2200mm à Loubaresse (07) sur les pentes du Tanargue.

- Le troisième ensemble marque un franc déclin de l'influence méditerranéenne et se situe sur les versants nord du massif. Si les premiers kilomètres depuis les crêtes présentent des caractéristiques assez proches de celles décrites dans les deux premiers paragraphes, le changement est plus important lorsque l'on arrive sur les plateaux lozériens et ardéchois. A l'image de la station d'Issanlas (07) (1220m), les pluies deviennent bien plus régulières, les épisodes pluvieux intenses nettement plus rares. Les températures sont froides : les maximales moyennes peinent à dépasser les 20°C et les minimales moyennes sont négatives environ 5 mois par an. L'ensoleillement est par ailleurs nettement moindre. La conjonction des influences montagnardes et océaniques rend ce climat majoritairement fort différent de celui des versants sud (figure 11).

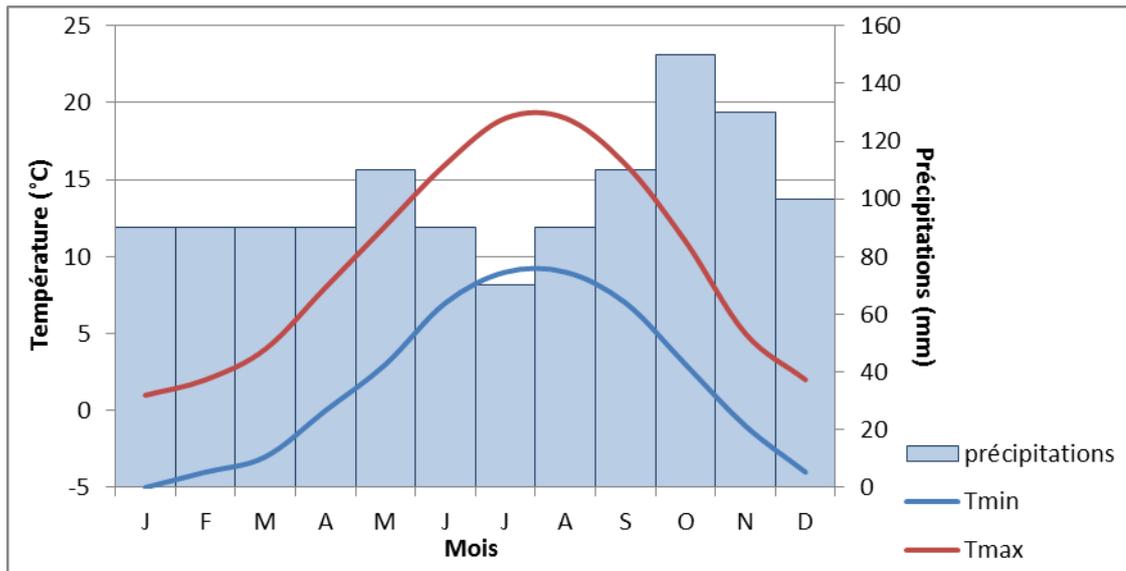


Figure 11: climatologie simplifiée des températures et des précipitations à la station d'Issanlas (07) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990).

En somme, la région d'étude représente une zone de transition entre le climat méditerranéen pur des plaines languedociennes et sa version presque entièrement altérée par l'altitude et les influences océaniques, en retrait des crêtes.

### 1.2.2 Episodes pluvieux intenses

Les franches tendances méditerranéennes de la zone d'étude se manifestent tout particulièrement lors des épisodes pluvieux intenses qui frappent régulièrement la zone. Comme nous le verrons dans cette section, la position géographique favorable à ce type d'évènements climatiques se combine avec des caractéristiques topographiques bien souvent aggravantes.

#### 1.2.2.1 *Situations à fortes précipitations sur la zone*

Un épisode pluvieux diluvien se définit comme un épisode durant lequel la hauteur de précipitations dépasse 200mm en 24h ou moins (Jacq, 1996). Dans la plupart des situations, ces phénomènes sont de nature convective.

L'observation des épisodes pluvieux que subit la zone permet de dégager trois grandes situations météorologiques qui y conduisent :

- Les épisodes diluviens provoqués par un forçage orographique : la convection à l'origine de ce type d'épisodes est initiée par le soulèvement des masses d'air douces et humides venant généralement du quadrant sud-est (figure 12). La mise en place de ce type de flux est la plupart du temps à mettre en relation avec l'approche par l'ouest d'un profond thalweg associé à un creusement dépressionnaire sur les Iles-Britanniques. La convection est souvent peu profonde, dans des conditions d'instabilité parfois très faibles, d'où la faible activité électrique observée la plupart du temps. On observe assez fréquemment ce type d'épisodes sur la zone d'étude, ce qui lui vaut le nom d' « épisode cévenol ». Cependant, l'arc cévenol n'est pas l'unique théâtre de ces épisodes. En revanche, la fréquence à laquelle ils y sont observés est très nettement supérieure à toutes les autres zones de relief du pourtour méditerranéen.

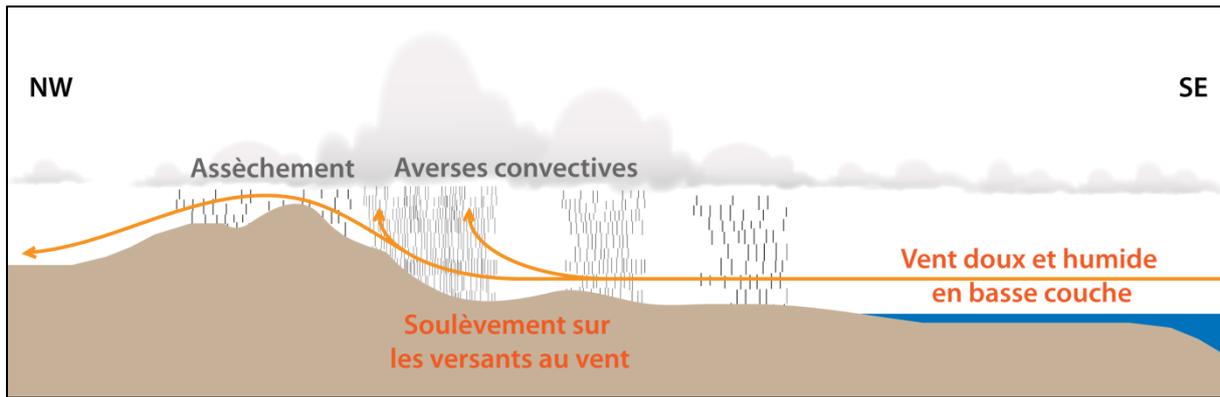


Figure 12: représentation schématique d'un épisode cévenol en coupe entre Golfe du Lion et Cévennes.

On y observe généralement des précipitations modérées à fortes, sur un à plusieurs jours, l'intensité allant souvent crescendo au fil du temps. Les cumuls peuvent dépasser 400mm en 24h mais les intensités horaires dépassent plus rarement les 50mm.

Ce fut notamment le cas lors de l'épisode d'octobre 2008 sur les Cévennes avec, entre autres, 485mm en 24h à Mialet (30) et des lames horaires proches de 50mm sur l'ensemble du massif.

- Les épisodes diluviens provoqués par des structures convectives à régénération arrière : dans ces cas-là, les forçages orographiques peuvent jouer un rôle mais ne sont pas les seuls éléments déclencheurs du phénomène. Les contextes météorologiques sont généralement assez typiques : un profond thalweg d'altitude approche de la zone, s'isole en cut-off associé à une goutte froide d'altitude ; les flux de secteur sud rapides et à courbure cyclonique s'enclenchent ; le phasage d'un contexte convergent au sol avec un contexte divergent en altitude est généralement l'élément déclencheur.

On observera des systèmes convectifs stationnaires ou à déplacement lent qui pourront être multicellulaires à régénération arrière, quasi linéaires (QLCS, *Quasi Linear Convective Systems*), ou encore des orages dits en « V », la pointe du V concentrant les plus fortes précipitations.

Les cumuls observés sont souvent de plusieurs centaines de millimètres et les intensités horaires peuvent largement dépasser les 100mm. Ces phénomènes sont généralement courts (quelques heures) mais peuvent parfois se prolonger plus longtemps, ce qui rend la situation hydrologique catastrophique.

Les cas de l'Aude en novembre 1999 (620mm en 36h dont 112mm en 1h à Lézignan-Corbières) ou encore du Gard en septembre 2002 (687mm en 24h. dont 137mm en 1h à Anduze) en sont des exemples typiques.

- Les épisodes diluviens composites : les deux types d'épisodes précédemment évoqués sont issus de situations météorologiques voisines à l'échelle synoptique (grande échelle). Au sein d'un épisode, c'est-à-dire sur un temps assez bref (quelques jours au maximum), et sur une étendue géographique limitée, ces deux dynamiques ne sont pas incompatibles et parfois même simultanées. C'est ce qui a pu être observé lors de l'épisode de début novembre 2011 qui a concerné une bonne partie de l'arc méditerranéen.

Les configurations de grande échelle qui amènent ces différentes situations à fort potentiel précipitant sont assez bien connues. En revanche, la grande variabilité spatiale et temporelle des pluies n'autorise qu'une prévisibilité médiocre à fine échelle (par exemple, le bassin versant). En effet, les cumuls observés sur 24h peuvent différer de plusieurs centaines de millimètres entre deux points séparés de quelques kilomètres.

### 1.2.2.2 Climatologie des épisodes pluvieux intenses

Un rapide regard sur la figure 13 fait clairement apparaître que la quasi-totalité des épisodes diluviens observés en France l'est à proximité de la Méditerranée.

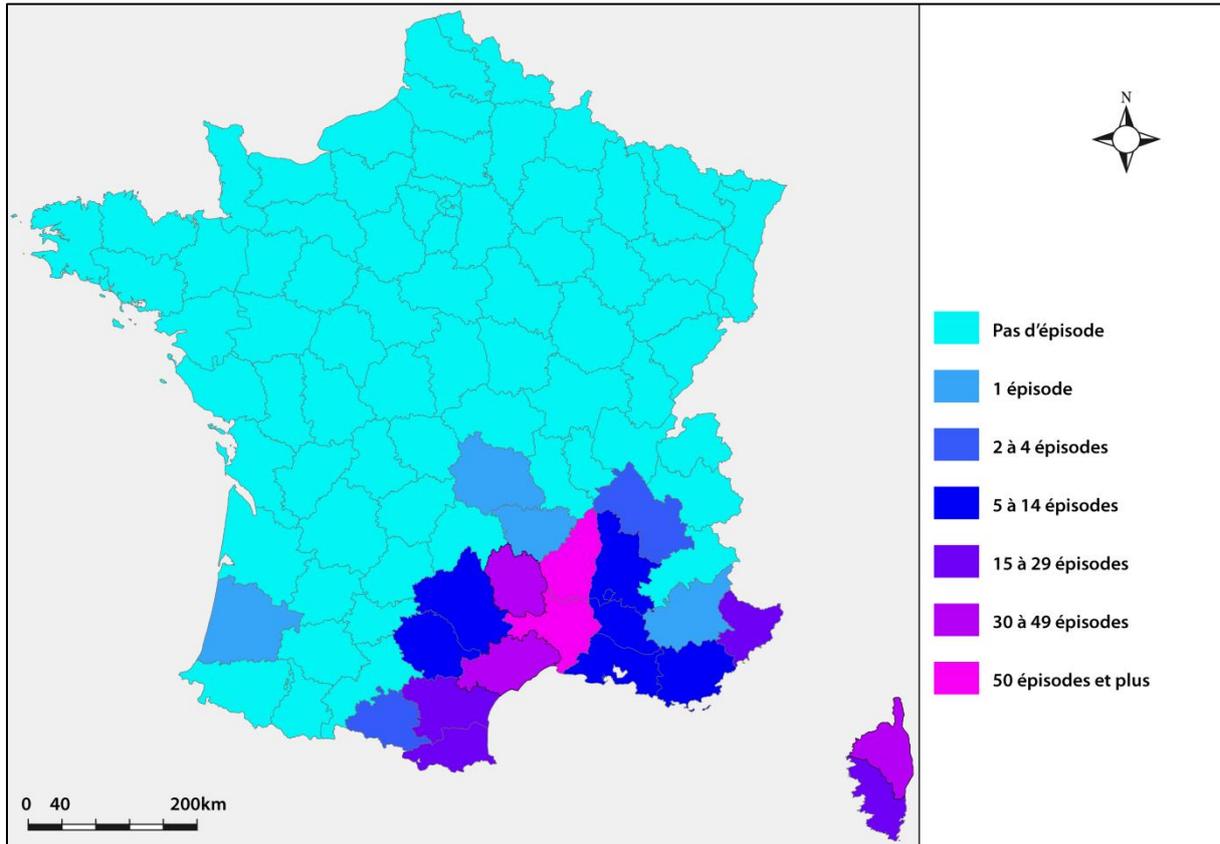


Figure 13: nombre d'épisodes (par département) de plus de 200mm en une journée climatologique de 1958 à 2010 inclus (Météo France, 2012).

On constate aussi que l'est du Rhône est nettement moins concerné par ces épisodes. C'est en effet en basse vallée du Rhône occidentale, dans les Cévennes et dans la plaine languedocienne que les épisodes sont les plus nombreux. En dehors des départements du Gard, de l'Ardèche, de la Lozère et de l'Hérault, les épisodes sont assez fréquents en Corse (surtout en Haute-Corse), dans les Alpes-Maritimes et entre Aude et Pyrénées-Orientales.

D'un point de vue temporel, la figure 14 montre une claire prédominance des épisodes automnaux. Toutefois, on remarque que si certaines périodes de l'année sont particulièrement calmes, il n'en existe aucune qui soit exempte de risque. C'est ce qui a été observé en juin 2010 dans le Var, en juin 1930 dans les Cévennes ou encore sur le Vaucluse en juillet 1991 (Météo France, 2012).

Une étude des 23 jours durant lesquels le cumul journalier a dépassé 200mm sur la période 2002-2006 a par ailleurs montré que les épisodes pluvio-orageux duraient en moyenne un peu

moins de 30 heures, se déclenchaient plutôt l'après-midi et se dissipaient plutôt en deuxième partie de nuit (Boudevillain et al., 2009).

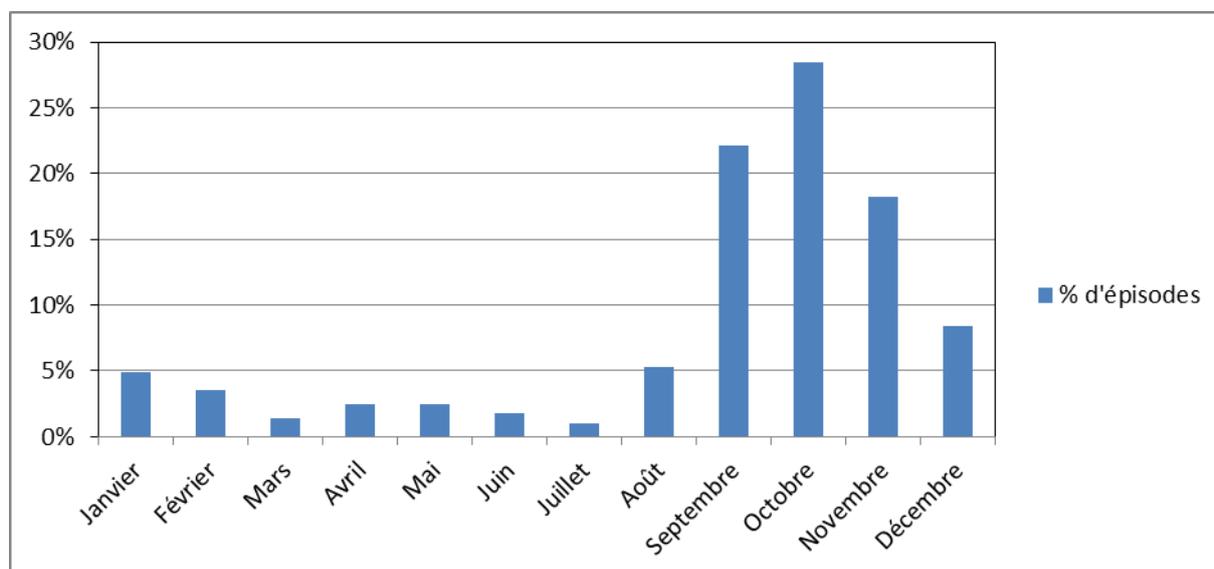


Figure 14: répartition mensuelle des épisodes de plus de 200mm en une journée climatologique sur la France et durant la période 1958-2010 (Météo France, 2012).

Enfin, la figure 15 montre qu'on ne retrouve pas de schéma pluriannuel clair aux échelles spatiales et temporelles considérées. En revanche, il se dégage une tendance à l'augmentation du nombre d'épisodes annuels puisque la pente de la droite de régression linéaire est de 0,1. Cependant, même sur une période de plus de 50 ans, la population est trop limitée pour en tirer des conclusions. Il s'agit là d'évènements intenses et par définition peu fréquents. Il conviendrait d'avoir un recul nettement plus important pour conclure au sujet de l'évolution du nombre d'épisodes annuel.

A ce sujet, il faut noter une activité nettement plus importante que la normale entre 1992 et 1997 qui contribue fortement à faire évoluer la moyenne récente vers des valeurs plus fortes.

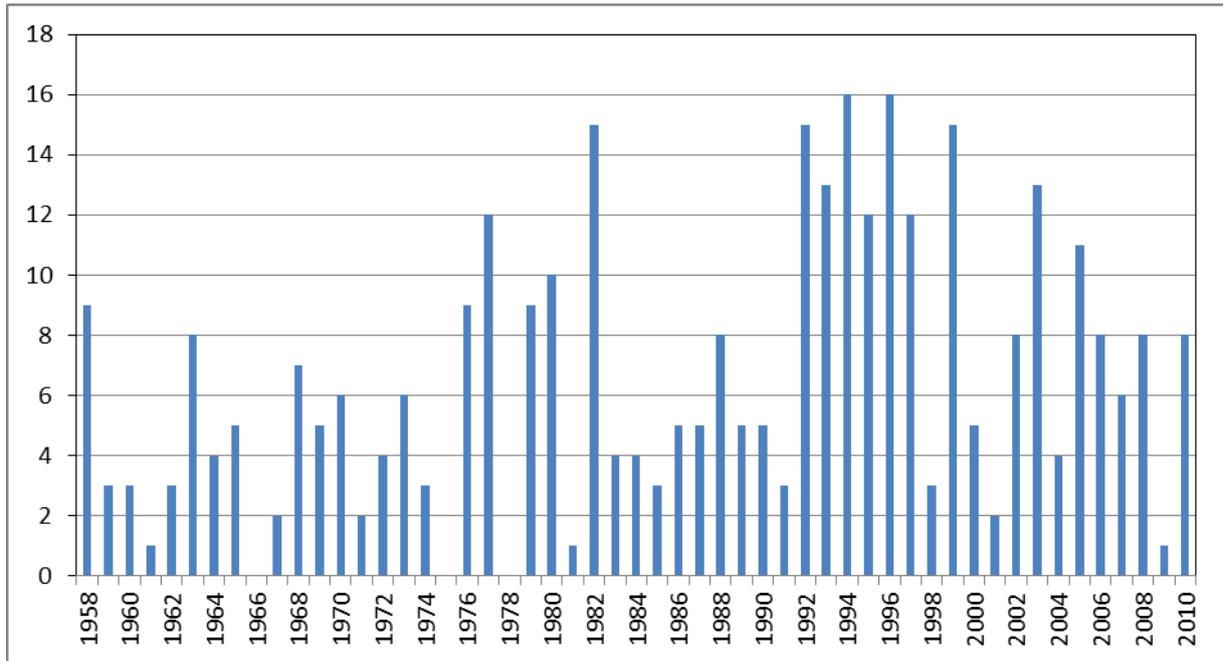


Figure 15: nombre d'épisodes de plus de 200mm en une journée climatologique par an sur la France et par département, durant la période 1958-2010 (Météo France, 2012).

### 1.2.2.3 De la complexité de la relation climat - hydrologie

D'après les données de débit moyen disponibles sur les cours d'eau de la zone d'étude, le régime hydrologique est typiquement pluvial. Le climat se répercute fortement sur l'hydrologie. Les longues sécheresses estivales, marquées par de fortes évapotranspirations, donnent lieu à des diminutions marquées du débit alors que les pluies automnales et hivernales contribuent à regonfler les cours d'eau (figure 16).

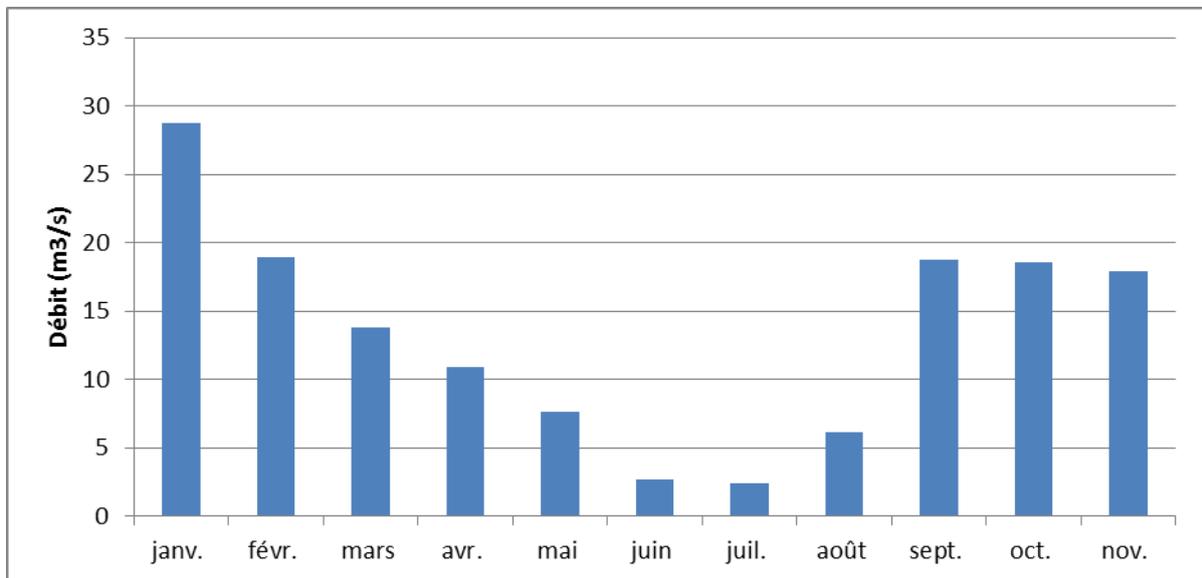


Figure 16: écoulements mensuels (naturels) calculés sur 43 ans pour le Gardon d'Anduze à Anduze (surface du bassin : 545km<sup>2</sup>). Données de débit d'après Banque Hydro.

Les épisodes diluviens qui frappent la zone provoquent alors d'importantes crues, difficilement prévisibles, de temps de montée courts, de débits spécifiques importants, et frappant des bassins de taille petite à modérée. Il s'agit donc « selon le terme proposé par l'Association Internationale des Sciences Hydrologiques de crues « éclair » » (Gaume, 2002).

À Ners (30), sur le Gardon (bassin versant de 1100km<sup>2</sup>), les débits d'étiage peuvent descendre en dessous d'un mètre cube par seconde (0,25 à Ners en 1836 selon la notice de la carte géologique du BRGM, Anduze au 1/50 000) et atteindre rapidement, lors de la survenue d'intenses épisodes pluvieux comme en septembre 2002, des valeurs impressionnantes de 7000 mètres cube par seconde. Ce débit correspond peu ou prou aux premiers débits de vigilance orange crue sur l'aval du Rhône (Tarascon, Avignon)...

Les cours d'eau de taille plus modeste sont, quant à eux, taris lors de périodes de sécheresse (Galeizon, Auzonnet, Avène...) selon la notice de la carte géologique du BRGM, Alès au 1/50 000, mais peuvent dépasser en crue des débits spécifiques de 10 mètres cube par seconde et par kilomètre carré (Gaume, 2002).

La relation entre climat et hydrologie peut ici paraître simple : débits très bas lors de longues sécheresses avec fortes évapotranspirations et crues éclair lors d'épisodes de pluie brefs et intenses.

Pourtant, que ce soit en étiage ou en crue, il est aisé de constater au regard des chroniques d'observations que la mesure des précipitations et de l'évapotranspiration ne permet pas de présumer du débit qui sera observé. En réalité, la relation entre climat et hydrologie dépend aussi beaucoup des caractéristiques du bassin versant, la rendant éminemment complexe. Nous aurons l'occasion d'approfondir cet aspect dans le deuxième chapitre.

### **1.3 La Gestion du risque d'inondations**

Les réactions hydrologiques évoquées dans la section précédente représentent un risque fort pour les populations et les infrastructures. De façon générale, il est couramment admis que le risque résulte d'un croisement géographique entre un aléa et des enjeux plus ou moins vulnérables.

Comme le fait remarquer Leone (2007), la notion de catastrophe est généralement définie comme le franchissement d'un seuil d'endommagement, conditionné par le dépassement des capacités de réponse et de relèvement d'une société.

Plus spécifiquement, cette définition, assez universelle, se traduit en France par la survenue de décès liés à l'évènement déclencheur de la situation critique et par le coût engendré par l'évènement (indemnisations de sinistres, remplacement ou réparation des biens non assurés, préjudice économique, coût des mesures d'urgence...). En ce sens, le territoire français a connu de nombreux cas d'épisodes catastrophiques liés aux crues éclair.

L'épisode des 8 et 9 septembre 2002 a frappé la zone d'étude, plus précisément le piémont cévenol. A la fois exceptionnellement vaste (5000km<sup>2</sup> touchés par plus de 200mm de précipitations, trois cours d'eau principaux en crue majeure : le Vidourle, la Cèze, le Gardon) et exceptionnellement intense (près de 700mm en 24h, dont souvent, plus de 100mm en 1h) (Météo France, 2012), cet évènement a profondément marqué la zone impactée : 24 victimes seront dénombrées, 419 communes sinistrées (Météo France, 2012) et le montant des dégâts a été estimé à 1,2 milliard d'euros (Huet et al., 2003).

On peut aussi citer, hors de la zone d'étude, l'épisode audois de novembre 1999 qui a fait 35 victimes et pour lequel le montant des dommages a pu dépasser, dans certaines zones rurales, 15 000€ par habitant (Vinet, 2008). Les catastrophes d'octobre 1940 dans les Pyrénées-Orientales, septembre et octobre 1958 dans la zone d'étude, octobre 1988 à Nîmes ont aussi fait

payer un lourd tribut aux territoires impactés. Plus récemment, l'épisode varois de juin 2010, a entraîné la mort de 23 personnes et fait 1 milliard d'euros de dégâts (FFSA, 2011).

Les coûts engendrés par les catastrophes naturelles sont de plus en plus importants (Chneiweis, 2008) diminuant mécaniquement l'acceptabilité des événements catastrophiques.

Si, dans certaines zones, la succession d'évènements intenses a tenu les populations et les infrastructures à distance respectable des cours d'eau, d'une manière générale, la pression démographique qui s'exerce sur les régions méditerranéennes a conduit à des implantations d'enjeux en zone inondable. Par exemple, l'Observatoire du Risque Inondation dans le Gard recense près de 39% de la population départementale vivant en zone inondable. Ce chiffre atteint 45% sur le bassin du Vidourle et 58% sur les bassins du Vistre et du Rhony réunis.

Face à cette problématique, il existe différents types de parades qui peuvent se montrer plus ou moins efficaces suivant la façon dont elles sont mises en œuvre :

- Action sur l'aléa : protection des enjeux :

L'action sur l'aléa est essentiellement représentée dans le risque inondations par les barrages, les digues et le ralentissement dynamique.

Ces aménagements structurels peuvent être extrêmement efficaces s'ils sont bien dimensionnés et bien situés mais, à l'inverse, ils peuvent générer un danger très important s'ils sont mal conçus, mal entretenus ou sous-dimensionnés.

L'accumulation d'eau qui se fait derrière l'infrastructure, si elle est libérée brusquement par une rupture ou une brèche, génère une sur-inondation forcément encore plus préjudiciable. S'il est difficile de relever les cas où ces ouvrages ont fourni une bonne protection, les cas où il n'en a rien été sont toujours marquants : en septembre 2002, la rupture d'une digue à Aramon (30) a permis à une vague de déferler sur une ville théoriquement protégée, donc peu organisée pour répondre à une telle crise. Aramon, bien qu'éloignée de l'épicentre des précipitations, fut la commune la plus marquée avec 1654 logements sinistrés, soit plus de 40% du total, selon Ledoux Consultants (2003).

On imagine par ailleurs quelles pourraient être les conséquences de la rupture d'un barrage écrêteur de crues pendant une forte sollicitation (plusieurs centaines de milliers de mètres cube stockés).

Ces ouvrages sont très coûteux, tant à la construction qu'à l'entretien, et ne produisent un bénéfice que très localisé pour les digues, un peu plus étendu pour les barrages, sans pour autant solutionner la totalité du problème.

Les digues tendent d'ailleurs à accélérer le courant et à supprimer l'action d'étalement et d'écrêtement (Salomon, 1997), mais aussi, en supprimant les crues faibles et moyennes, elles font sortir de la mémoire collective la réalité sur les crues du cours d'eau concerné (Vinet, 2010). Enfin, en hautes eaux, les accumulations d'eau à l'arrière des digues peuvent se révéler aussi préjudiciables qu'une crue pour les enjeux concernés. Cela n'empêche pas le territoire français de compter près de 8200km de digues fluviales (Rode, 2012).

L'essentiel des précipitations de l'épisode de septembre 2002 a frappé en aval du barrage de Sainte-Cécile-d'Andorge (photo 2) ce qui n'a pas permis à ce dernier de protéger la ville d'Alès des inondations qu'elle a subies. Il en est de même pour le barrage de Caramany (66) qui, selon Vinet (2010), n'a pu empêcher la crue de l'Agly puisque les pluies sont essentiellement tombées sur un affluent en aval, le Verdoube. Contrairement au cas des digues, les barrages purement écrêteurs de crues sont rares en France.

- Action sur l'implantation des enjeux : la restriction urbanistique :

Dès 1935, les Plans de Surface Submersible ont été prévus par la loi (article 2 du décret-loi du 30 octobre 1935). Ils avaient pour but de permettre à l'administration de s'opposer à toute action ou ouvrage pouvant gêner le libre écoulement des eaux ou la conservation des champs d'inondation. La loi était alors axée sur la préservation du milieu naturel lié à la rivière.

D'évolutions en intégrations dans les Plans d'Occupation des Sols, le législateur en est arrivé aux Plans de Prévention des Risques (PPR) en 1995, donnant bien souvent par là même aux anciens documents une valeur de PPR.

Les PPR délimitent les zones exposées aux risques et prévoient des mesures de protection et de sauvegarde en fonction de cette exposition. Dans les faits, l'application de ces plans sur le terrain se traduit par une réglementation stricte sur la constructibilité : zones constructibles sans risque prévisible, zones constructibles peu risquées avec aménagements spécifiques et zones inconstructibles aux bâtis existants uniquement modifiables dans le sens d'une réduction de la vulnérabilité.

Cette mise en pratique est jugée par beaucoup éloignée de l'esprit initial de la loi. En particulier, sont souvent pointés la logique définitive d'instruction des PPR, les manques des politiques de perspective d'amélioration des conditions de sécurité par la sauvegarde ou la protection ou encore le manque de concertation (Roustan, 2007).

Quoi qu'il en soit, là où l'existant n'est pas concerné, ces mesures permettent de stopper la multiplication des enjeux en zone inondable ce qui constitue déjà un avantage considérable.

- Action sur la vulnérabilité des enjeux :

La notion de vulnérabilité n'est clairement définie qu'en fonction de l'approche et de l'échelle choisies (Ledoux, 2006). Il existe plusieurs types de vulnérabilités qui peuvent être réduites de façon ciblée.

Dans la logique des prescriptions des PPR, la réduction de la vulnérabilité concerne presque exclusivement le bâti (étage aux habitations, vides sanitaires, rehaussement des installations électriques...).

La réduction de la vulnérabilité économique est quant à elle essentiellement assurée par les entreprises qui souhaitent mettre en place des procédures limitant l'impact des épisodes les plus sévères sur leurs activités.

Enfin, la vulnérabilité humaine est tout particulièrement ciblée depuis quelques années par deux types de mesures : la planification de la gestion de crise et l'information préventive. Au travers des lois Risques de 2003 et de Modernisation de la Sécurité Civile de 2004, ces deux thématiques ont été mises au cœur des politiques de gestion des risques.

L'information préventive des populations passe par une communication optimisée sur la prévision des situations à risque : c'est, depuis 2001, le rôle des vigilances météorologiques et depuis 2006, nous le verrons dans la section 1.4, le rôle des vigilances hydrologiques. L'information du grand public est assurée par le biais des médias et par le biais de sites internet.

Par ailleurs, au-delà du fait que les PPR sont consultables en mairie, les communes qui y sont soumises ont l'obligation d'organiser, tous les deux ans, une réunion publique exposant les risques sur le territoire communal et les catastrophes qu'a connues la commune. Enfin, le Document d'Information Communal sur les Risques Majeurs doit être disponible en mairie et sa distribution aux administrés est recommandée. Il détaille les comportements individuels de sécurité pour faire face aux situations de crise.

Les mesures permettant d'aider la gestion de crise sont quant à elles représentées à l'échelle communale par les Plans Communaux de Sauvegarde qui permettent d'organiser l'alerte des populations et de fournir aux administrés une première assistance, voire un secours en cas de danger imminent. On comprendra qu'une prévision hydrologique efficace est l'une des conditions du succès de ces procédures.

Dans sa vocation de veiller à la sécurité des citoyens, l'Etat a mis en place de nombreuses mesures. L'ordre dans lequel elles ont été présentées ici correspond à peu près à l'ordre dans lequel elles ont été privilégiées au fil du temps. En réalité, il apparaît que les avantages de chacune de ces méthodes se complètent et qu'aucune d'entre elles ne devrait être privilégiée au détriment des autres.

Cependant, il est important de remarquer qu'un territoire soumis à un aléa (atténué par un ouvrage de protection) qui menace des enjeux (limités en nombre et en vulnérabilité par la restriction urbanistique) et dont les populations sont bien informées ne peut faire l'objet d'une protection optimale qu'à l'aide de prévisions de l'aléa qui permettent, le cas échéant, de gérer au mieux la crise.

#### **1.4 Le projet BVNE Gardons**

Après la création des vigilances météorologiques (en 2001) et la formalisation des mesures de sauvegarde communales (en 2004), les prévisions hydrologiques se sont naturellement imposées comme une nécessité pour les zones soumises aux aléas d'inondation.

Initié en 2006 par le SCHAPI, le projet BVNE Gardon a pour objet de réunir plusieurs équipes de recherche travaillant sur la prévision des crues éclair. Le but de la démarche est de comparer les prévisions issues des différents modèles sur des bassins et des données communes. Au moment de la rédaction de ce mémoire, outre notre contribution au titre de l'Ecole des Mines d'Alès, les équipes engagées dans ce projet sont les suivantes :

- Bureau de Recherches Géologiques et Minières (BRGM, équipes de Montpellier et Orléans),
- Institut national de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture, ex-CEMAGREF Centre national du Machinisme Agricole, du Génie Rural, des Eaux et des Forêts (IRSTEA, centre d'Antony, équipe Hydrosystèmes et Bioprocédés),
- Environnements DYnamiques et TERRitoires de Montagne (EDYTEM, équipe Dynamiques Actuelles des Milieux de Montagne),
- Hydrosiences Montpellier (HSM, thème Risques Hydrologiques liés aux Aléas Extrêmes),
- Institut de Mécanique des Fluides de Toulouse (IMFT, équipe Modélisation Hydrologique Distribuée),
- Laboratoire d'étude des Transferts en Hydrologie et Environnement (LTHE),

- Météo France.

Les travaux visent à alimenter la carte de vigilance crues dispensée par le SCHAPI sous la responsabilité du Ministère en charge du développement durable et disponible sur [www.vigicrues.gouv.fr](http://www.vigicrues.gouv.fr) (Figure 17).

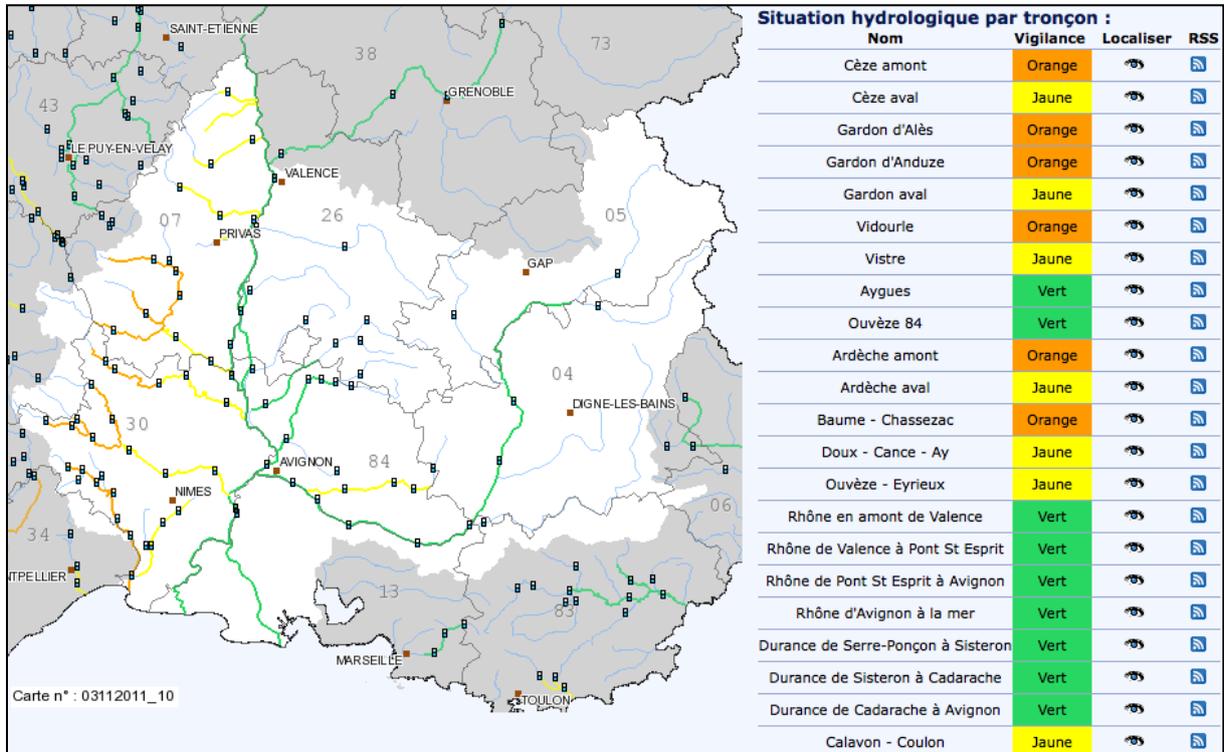


Figure 17: exemple de carte de vigilance crues le 3 novembre 2011 sur le territoire du SPC Grand Delta (en blanc). Les tronçons réglementaires sont colorés en fonction du niveau de risque qui pourrait être atteint dans les prochaines 24 heures.

Une présentation de la vigilance « crues » est disponible sur le site internet dédié (MEDDTL, 2012) :

« La vigilance "crues", mise en place en juillet 2006, est fondée sur les mêmes principes que la vigilance météorologique produite par Météo-France. Son objectif est d'informer le public et les acteurs de la gestion de crise en cas de risque de crues sur les cours d'eau surveillés par l'Etat, dans le cadre de sa mission réglementaire de surveillance, de prévision et de transmission de l'information sur les crues.

La vigilance "crues" est destinée à informer tous les publics intéressés, particuliers ou professionnels, sous une forme simple et claire. Elle est notamment destinée aux pouvoirs publics en charge de la sécurité civile (préfets, maires, etc.), qui déclenchent les alertes lorsque cela est nécessaire et mobilisent les moyens de secours.

Chaque cours d'eau inclus dans le dispositif de la vigilance "crues" apparaît sur la carte de vigilance. Ces cours d'eau sont le plus souvent découpés en tronçons. A chaque tronçon est affectée une couleur : vert, jaune, orange ou rouge selon le niveau de vigilance adapté pour faire face au danger susceptible de se produire dans les 24 heures à venir.

La carte se présente sous une vue nationale ou sous des vues locales accessibles par clic sur la zone du Service de Prévision des Crues concerné (SPC). La carte nationale est accompagnée d'un bulletin d'information à l'échelle nationale. Les cartes locales sont accompagnées de bulletins d'information sur les domaines des Services de Prévision des Crues correspondants. Ces bulletins précisent la chronologie et l'évolution des crues, en qualifiant l'intensité et fournissant (si possible) des prévisions chiffrées pour les stations de référence. Ils contiennent également une indication des conséquences possibles, ainsi que

*des conseils de comportement définis par les pouvoirs publics en fonction du danger susceptible de se produire.*

*Depuis les cartes locales, il est possible en cliquant sur une station hydrologique de visualiser le graphique ou le tableau contenant les dernières hauteurs d'eau ou les derniers débits mesurés à cette station. Des niveaux de crues de référence à la station peuvent être également affichés.*

*La carte de vigilance "crues", les bulletins et les données mesurées sont disponibles en permanence. La carte est actualisée deux fois par jour à 10h et à 16h. En période de crues, les bulletins peuvent être réactualisés plus fréquemment. Par ailleurs, si un changement significatif intervient, cartes et bulletins peuvent être réactualisés à tout moment.*

*La durée de validité de la couleur d'un tronçon est de 24h à compter de la date d'actualisation de la carte. Lorsque les conditions hydrologiques le permettent, les bulletins peuvent également contenir des informations sur l'évolution prévue au-delà de 24 heures. »*

Selon ce même site, les quatre niveaux de vigilance sont définis par les conséquences possibles et conseils de comportements associés à l'importance de la crue prévue :

*« **Rouge** : Risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.*

*Conséquences possibles : Des inondations très importantes sont possibles y compris dans les zones rarement inondées. Les conditions de circulation peuvent être rendues extrêmement difficiles sur l'ensemble du réseau routier ou ferroviaire. Des coupures d'électricité plus ou moins longues peuvent se produire. Des phénomènes de rupture ou de débordement de digues peuvent se produire.*

*Conseils de comportement : Mettez-vous à l'abri et suivez strictement les consignes de sécurité des pouvoirs publics. Evitez tout déplacement. Tenez-vous informé de l'évolution de la situation (radio, etc...). Veillez à la protection des biens susceptibles d'être inondés ou emportés (mobilier, produits toxiques, appareil électriques, etc...). Coupez les réseaux si nécessaire (électricité, gaz, eau).*

***Orange** : Risque de crue génératrice de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes.*

*Conséquences possibles : Des inondations importantes sont possibles. Les conditions de circulation peuvent être rendues difficiles sur l'ensemble du réseau et des perturbations peuvent affecter les transports ferroviaires. Des coupures d'électricité peuvent se produire. Les digues peuvent être fragilisées ou submergées.*

*Conseils de comportement : Mettez-vous à l'abri. Limitez tout déplacement sauf si absolument nécessaire et conformez-vous à la signalisation routière. Tenez-vous informé de l'évolution de la situation (radio, etc...). Veillez à la protection des biens susceptibles d'être inondés ou emportés (mobilier, produits toxiques, appareil électriques, etc...).*

***Jaune** : Risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées.*

*Conséquences possibles : Tenez-vous informé de la situation. Les premiers débordements peuvent être constatés. Certains cours d'eau peuvent connaître une montée rapide des eaux.*

*Conseils de comportement : Soyez vigilant si vous vous situez à proximité d'un cours d'eau ou d'une zone habituellement inondable. Conformez-vous à la signalisation routière.*

***Vert** : Pas de vigilance particulière requise. »*

La nature même des crues méditerranéennes implique des contraintes opérationnelles fortes. L'indispensable anticipation des mesures de sauvegarde prises en temps de crise sur le terrain donne tout son sens à l'amélioration de l'efficacité des modèles de prévision. Une fois l'épisode

pluvieux intense entamé, les réseaux sont mis à rude épreuve : les déplacements sont difficiles et parfois impossibles, les télécommunications sont coupées tout comme l'électricité, l'eau devient impropre à la consommation... Il convient donc d'avoir la capacité d'anticiper la situation critique autant par l'élaboration préalable de procédures de sauvegarde que par l'amélioration des capacités de prévision de l'aléa.

En somme, le projet BVNE participe à l'amélioration des outils d'information préventive à destination des populations, mais aussi des outils d'aide à la décision pour une conduite optimale des procédures de gestion de crise par les autorités.

## **1.5 Données hydrométéorologiques**

### **1.5.1 Disponibilité et qualité des données de hauteur d'eau et de débit**

#### *1.5.1.1 Disponibilité*

De façon récurrente, l'hydrologie se heurte au problème de la disponibilité des données de hauteur d'eau et de débit et ce, par trois aspects différents :

- Les stations limnimétriques peuvent être emportées ou endommagées par les crues extrêmes, comme ce fut le cas en septembre 2002 à Anduze. Parfois, c'est le pont qui sert de support à la station de mesure qui cède, comme en juin 2010 dans le Var. Dans d'autres cas, les transmissions de données sont rendues difficiles par les conditions climatiques, en particulier durant les épisodes pluvio-orageux les plus sévères, c'est à dire au moment où l'information est plus que jamais nécessaire à la bonne marche des services de prévision et d'alerte.
- La longueur nécessaire de l'historique des données impose d'attendre plusieurs années pour obtenir une chronique exploitable. Par ailleurs, l'historique nécessaire est d'autant plus important que la fréquence des évènements recherchés est faible. Ceci est donc particulièrement vrai pour les évènements extrêmes. Enfin, le positionnement des stations doit rester stable dans le temps, sans modification majeure du lit de la rivière. Par exemple, la réfection d'un pont ou le déplacement d'une station (même de quelques dizaines de mètres) obligent souvent à reprendre la chronique à zéro, ce qui dégrade la continuité des données.
- La question de l'existence même des stations de mesure se pose également. Comme le montrent les deux cartes de la figure 18, on peut noter que certains cours d'eau relativement importants ne sont pas pris en compte dans les tronçons réglementaires de vigicrues qui affichent la majorité des stations de mesure limnimétrique existantes.

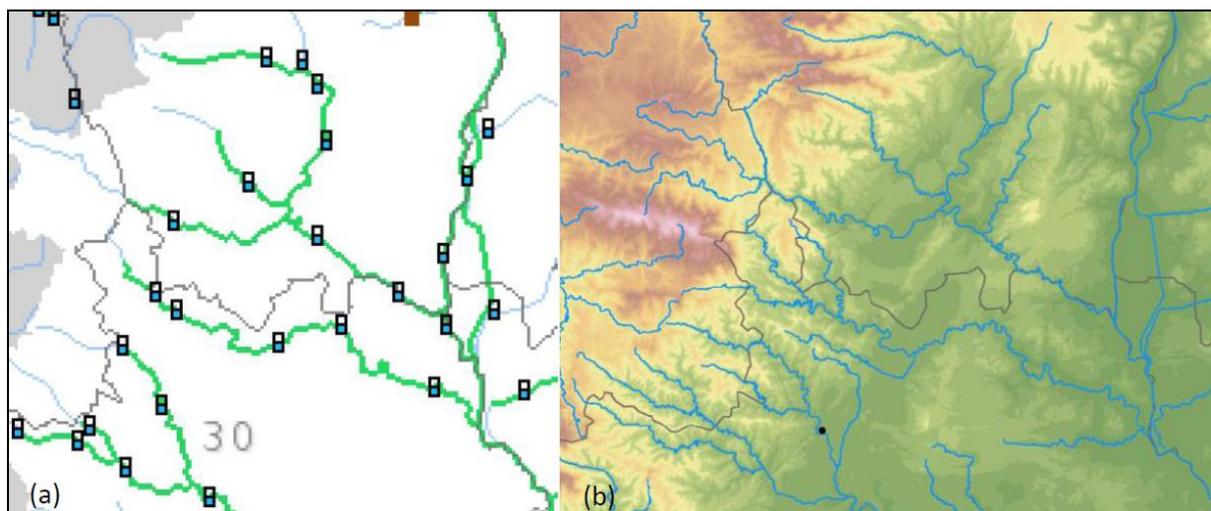


Figure 18: (a) : exemple de carte de vigilance crue sur le nord du Gard et le sud de l'Ardèche, les tronçons en vert sont dits tronçons réglementaires et sont dotés d'un niveau de vigilance, ici minimal. (b) : ensemble des cours d'eau de catégorie trois et moins du nord du Gard et du sud de l'Ardèche (conception Artigue 2010, sources : (a) : vigicrues.gouv.fr ; (b) : altitude d'après ASTER GDEM, un produit de METII and NASA, cours d'eau d'après IGN BD CARTHAGE, limites départementales d'après IGN BD GEOFLA).

Le réseau national des tronçons réglementaires a été défini sur des critères d'enjeux et de faisabilité des prévisions à 24 heures. Il en ressort que seuls les cours d'eau principaux sont représentés. Ont été représentés les cours d'eau classifiés trois et moins sur la figure 18(b). Le tableau 1 donne cette classification (IGN, 2006).

Tableau 1: classification des cours d'eau de la BD Carthage (IGN, 2006).

Classification	Description
1	Tout cours d'eau d'une longueur supérieure à 100km ou tout cours d'eau se jetant dans une "embouchure logique" et d'une longueur supérieure à 25km.
2	Tout cours d'eau d'une longueur comprise entre 50 et 100km ou tout cours d'eau se jetant dans une "embouchure logique" et d'une longueur supérieure à 10km.
3	Tout cours d'eau d'une longueur comprise entre 25 et 50km
4	Tout cours d'eau d'une longueur comprise entre 10 et 25km
5	Tout cours d'eau d'une longueur comprise entre 5 et 10km
6	Tous les autres cours d'eau hormis ceux issus de la densification du réseau
7	Cours d'eau issus de la densification du réseau

Ce niveau de couverture ne doit pas masquer la complexité du problème ici-soulevé et l'on peut trouver au moins deux raisons à cet état de fait :

- Les cours d'eau de taille modeste, même jaugés, présentent des temps de réponse très brefs ne permettant que peu de recul pour que l'alerte atteigne une efficacité suffisante. La prévision de la crue nécessite alors une prévision fine et précise de la pluie - dont l'hydrométéorologie ne dispose pas à ce jour - pour passer outre le faible

temps de réponse et autoriser suffisamment d'anticipation pour proposer à temps le niveau de vigilance adapté.

- Certains cours d'eau ne sont pas jaugés, ne permettant pas, quelle que soit leur taille, de disposer de chroniques de données suffisantes pour faire fonctionner de façon directe des modèles de prévision des crues. La prévisibilité y est donc médiocre : on ne peut utiliser que les pluies observées et prévues pour évaluer une éventuelle réaction hydrologique, sans pour autant connaître la teneur exacte de la relation pluie-débit.

#### *1.5.1.2 Qualité*

Les systèmes de mesure de hauteur d'eau sont relativement fiables compte tenu des conditions auxquelles ils sont soumis. En effet, dans les bases de données disponibles dans le cadre du projet BVNE Gardons, le taux de données manquantes ou erronées sur 17 ans ne dépasse pas 10%. Cependant, plus de la moitié de ces données manquantes ou erronées sont enregistrées pendant les crues. Les questions de qualité des données sont en réalité essentiellement dues à la conversion des hauteurs d'eau en débit par le biais des courbes de tarage.

Le problème concerne en premier lieu les débits élevés pour lesquels il est très difficile de procéder à un jaugeage : les conditions d'accès au site peuvent être très difficiles durant un épisode intense, le risque peut s'avérer important pour les personnels en charge des manipulations pendant le déroulement même de l'opération. De plus, les crues sont généralement brèves et parfois nocturnes ou dominicales.

La modification de la topographie du lit, si elle est majeure, peut entraîner la reprise à l'initial de la chronique de données (cf. 1.5.1.1). Si elle est mineure, suite à une crue qui a fortement érodé ou déposé sur le site ou suite à des aménagements dans le lit, il est nécessaire de réactualiser la courbe de tarage pour obtenir à nouveau une évaluation correcte des débits.

Ainsi, si l'utilisation des débits est souvent préférée étant donné qu'elle apporte, contrairement à la hauteur d'eau, une information volumique, elle entraîne fatalement un certain nombre de difficultés supplémentaires pour l'obtention de données fiables.

#### 1.5.2 Disponibilité et fiabilité des données de pluie

Les données de pluie sont plus aisées à obtenir. L'installation de pluviomètres accompagne bien souvent l'implantation d'une station hydrométrique.

Au-delà de la fiabilité intrinsèque de l'appareil, le positionnement du pluviomètre en fonction de son environnement est un élément capital (à distance suffisante des obstacles, suffisamment haut par rapport à la surface qui le supporte...). Pour les pluviomètres isolés sur le bassin-versant, en particulier en milieu rural, il est assez aisé de trouver des secteurs suffisamment dégagés et adaptés à ces exigences. En revanche, les pluviomètres couplés aux stations hydrométriques, qui sont souvent implantées en ville au plus près d'un pont, sont forcément plus délicats à installer correctement : multiplicité des bâtiments aux alentours, végétation dense dans les zones humides de fond de vallée, distance contrainte entre la station hydrométrique et le pluviomètre...

Un autre facteur d'incertitude tient aux conditions climatiques elles-mêmes. D'abord, le vent perturbe la mesure pluviométrique par modification de la surface d'interception. L'angle d'attaque des gouttes sur l'intérieur du cône, s'il est trop plat, génère des rebonds dont il est difficile d'évaluer l'impact sur la mesure. Dans d'autres cas, c'est la neige qui vient perturber la

mesure. De fortes quantités d'eau solide peuvent alors précipiter sur les sommets, sans que le pluviomètre ne relève quoi que ce soit. En cas de brusque redoux et de hausse de la limite pluie-neige, l'apport d'eau au système hydrologique peut être beaucoup plus important que ce que la mesure laisse penser.

Une fois considéré le problème du positionnement de l'appareil par rapport à son environnement immédiat et celui du procédé de mesure, reste à résoudre le problème de la localisation du pluviomètre dans le bassin versant. Le nombre de pluviomètres nécessaires à une bonne mesure peut être très important dans le contexte climatique défini dans la partie 1.2. En effet, avec une surface d'interception de  $400\text{cm}^2$ , le pluviomètre ne peut être réellement représentatif que du proche environnement de son point de mesure.

Pour éviter les problèmes ici évoqués, l'utilisation de données de réflectivité fournies par les radars météorologiques peut être une bonne solution. Bachoc et al. (2011) remarquent les progrès effectués par la combinaison des données issues des radars et des pluviomètres mais notent aussi que de nombreux progrès supplémentaires restent à réaliser pour obtenir des lames d'eau réalistes, tant sur les pluviomètres que sur les radars ou les méthodes permettant de les calibrer. Ces méthodes peuvent parfois s'avérer encore lourdes pour les utilisations opérationnelles. De plus, les radars sont pour le moment inopérants dans les zones à fort relief. Dans le domaine de l'hydrologie urbaine, Emmanuel et al. (2012) remarquent d'ailleurs que l'évaluation du champ de pluie par le radar est largement meilleure dans des conditions de pluies stratiformes que dans des conditions de pluies convectives. Les auteurs concluent que les données issues du radar ne sont pas aujourd'hui suffisamment fiables pour une mesure du champ de pluie dans le contexte de l'hydrologie urbaine.

Enfin, l'utilisation de données radar pose le problème de la continuité des séries de données. L'amélioration qu'ils peuvent effectivement apporter une fois bien calibrés doit, c'est l'objectif, décrire nettement mieux que les pluviomètres le champ de pluie sur le bassin versant. Ces techniques sont encore récentes ainsi, les méthodes qui nécessitent le plus de volume de données doivent souvent se contenter des informations fournies par les pluviomètres.

Avec les nouvelles générations de radars et l'amélioration du traitement en temps réel des données, de substantielles avancées sont toutefois à attendre dans les années qui viennent. Actuellement, sont testés en Provence-Alpes-Côte-d'Azur des radars à bande X qui ouvrent de nouvelles perspectives en zone de relief, de même que les disdromètres utilisés pour déterminer la phase des précipitations (Bachoc et al., 2011).

## **1.6 Position de la problématique**

Indépendamment des verrous techniques liés à la disponibilité et la qualité des données hydrométéorologiques, toutes les zones potentiellement exposées devraient pouvoir bénéficier de prévisions de débit efficaces. Même si l'idéal serait de disposer de données de qualité pour tous les bassins, pour atteindre cet objectif, il s'agit d'être en mesure de fournir des prévisions sur les cours d'eau à réaction rapide :

- **jaugés**, lorsque la disponibilité des données est adaptée à une utilisation opérationnelle,
- **mal jaugés**, lorsque la transmission ou la capture des données pose passagèrement problème pour différentes raisons, à l'échelle historique ou à l'échelle de l'évènement,

- **non jaugés**, lorsque qu'aucun appareil de mesure du débit n'est installé.

C'est dans ce cadre que s'inscrivent les travaux de recherche dont ce mémoire fait l'objet. Il s'agit de mettre au point un modèle de prévision des crues en absence de prévision de pluie, adaptable à chaque contexte de mesure ici décrit, et permettant aux autorités (gestion de crise) et aux populations (information préventive) d'anticiper les inondations rapides générées par ces crues.

## **1.7 Conclusions**

Le contexte des travaux menés dans le cadre de cette thèse a été défini dans ce chapitre.

Nous avons pu comprendre toutes les spécificités de ce territoire méditerranéen escarpé, au carrefour d'influences climatiques diverses et en croissance démographique. Plus particulièrement, nous avons pu explorer les extrêmes d'un climat qui, historiquement, a tendu alternativement à attirer ou à chasser les populations des environs des cours d'eau au gré des périodes de sécheresse, des épisodes pluvieux, des besoins commerciaux ou du tourisme.

A l'échelle nationale, nous avons vu que la gestion du risque d'inondations utilise un arsenal de mesures pour la prévention, la protection et la gestion de crise. Ces mesures s'appliquent tout particulièrement à notre zone d'étude. Au travers de la description de ces mesures et de la mise en place de la carte Vigicrues et du projet BVNE Gardons, nous avons pu mesurer à quel point la prévision hydrologique était la pierre angulaire d'une gestion intégrée de ce risque.

Face aux problèmes de disponibilité et de qualité des données hydrométéorologiques et à l'importance croissante des enjeux tant humains que financiers, nous nous proposons de mener à bien l'élaboration d'un modèle de prévision des crues qui fonctionne sans prévision de pluie et qui peut s'adapter à la majorité des contextes de mesure.



## **2. Dynamique hydrologique et modélisation des crues rapides**

Ce chapitre présente des éléments de compréhension des processus hydrologiques et de la modélisation des crues pour la prévision ou pour la simulation.

D'abord, seront présentés à travers un schéma descendant, de la pluie au drain, les processus mis-en-jeu dans la relation entre pluie et débit. Ces processus sont multiples et difficilement quantifiables, même relativement les uns aux autres, rendant la dynamique hydrologique complexe et non-linéaire. De l'influence de la végétation à la description des processus de transfert sur les versants, les éléments susceptibles d'influencer la dynamique hydrologique seront explicités au fil d'une première partie.

Face aux problématiques mises en avant dans le chapitre 1 et à cette complexité, après avoir fourni des définitions sur les modèles hydrologiques, résumé les qualités qui en sont attendues et montré comment évaluer leur performance, nous effectuerons dans une deuxième partie un tour d'horizon des diverses mises en œuvre des modèles hydrologiques pour la prévision des crues rapides, notant ainsi les difficultés rencontrées par ces travaux, de la sensibilité au changement d'échelle, à l'indisponibilité de variables critiques comme l'humidité des sols ou les pluies futures. De ce fait il apparaît que même les modèles à base physique doivent être soumis à un calage préalable de leurs paramètres, limitant ainsi leur mise en œuvre sur des bassins non jaugés.

Afin de pallier cette difficulté, nous présenterons plusieurs procédés mis en œuvre par la communauté des hydrologues, parmi lesquels le procédé de « régionalisation » des paramètres d'un modèle.

En synthèse, et face aux nombreuses inconnues que pose encore la génération des crues rapides, nous présenterons comment les modèles statistiques ont tout leur intérêt. Parmi ceux-ci nous proposerons l'utilisation des réseaux de neurones pour la prévision des crues éclair en s'approchant autant que possible des conditions opérationnelles.

## 2.1 Mécanismes de la réponse hydrologique

De façon synthétique, on peut résumer le processus de transformation de la pluie en débit dans le schéma présenté en annexe 1. Ce processus peut se décomposer en plusieurs sous-processus qui peuvent intervenir de manière complémentaire, simultanée, ou successive selon la configuration du bassin versant. Ces processus élémentaires sont les suivants :

- l'interception,
- l'égouttement,
- la rétention superficielle,
- l'évapotranspiration,
- le ruissellement de surface,
- l'infiltration et la percolation,
- l'exfiltration,
- les écoulements hypodermiques ou souterrains par transfert de pression,
- les écoulements hypodermiques par les macropores,
- et enfin, le stockage.

Nous allons décrire dans ce chapitre les mécanismes intervenant dans la réponse hydrologique à la pluie, en nous focalisant sur chaque grande étape classiquement identifiée et, afin de limiter notre propos, en nous plaçant dans le cadre des crues rapides qui nous intéressent dans ce mémoire, et donc des pluies fortes.

### 2.1.1 Interface atmosphère-bassin :

Les premiers des phénomènes, par ordre chronologique, concernent l'interface entre l'atmosphère et le bassin. Cette interface n'est pas nettement délimitée et ses caractéristiques varient en fonction de l'occupation des sols faisant intervenir différents types de processus.

#### 2.1.1.1 *Influence de la végétation*

A l'interface entre l'atmosphère et les terrains, la végétation conditionne beaucoup les tout premiers processus observés :

- l'interception, c'est-à-dire la captation et la retenue d'eau par la végétation,
- l'évapotranspiration, c'est-à-dire le transfert d'eau depuis le sol vers l'atmosphère par évaporation et transpiration des plantes,
- l'égouttement et l'écoulement le long des tiges et troncs, c'est-à-dire le transfert de l'eau de la canopée au sol.

La quantité d'eau directement interceptée par le couvert végétal est très variable. Elle dépend de la structure de l'épisode pluvieux (intensité, intermittence entre périodes d'averses et d'accalmies), de la quantité de pluie, de la période considérée, du type de végétation et de son organisation. Meunier (1996) conclut des différentes études consultées que « *la partie aérienne de la forêt ne peut en aucun cas expliquer une réduction importante des valeurs de pluie au sol pour des épisodes de pluies abondantes (quelques dizaines de mm), et encore moins pour des épisodes d'orages très violents (quelques centaines de mm)* ». Par rapport à notre propos qui concerne les crues rapides provoquées par des épisodes convectifs diluviens, l'interception joue

donc un rôle négligeable (événements relativement brefs à fortes intensités, souvent ventés et se produisant hors des périodes de températures élevées).

Si l'on considère l'influence de l'évapotranspiration, celle-ci peut être très importante sur une saison, mais à l'échelle d'un épisode qui dure tout au plus quelques jours et comparée aux cumuls de précipitations observés dans la zone d'étude, elle n'est pas significative. Elle a toutefois un impact indirect en contribuant à l'assèchement des terrains et donc en modifiant la dynamique des premiers événements de la saison humide : c'est là toute la question de la variabilité de l'état hydrique initial du bassin versant sur laquelle nous reviendrons à plusieurs reprises dans ce mémoire.

Si l'évapotranspiration et l'interception ont un rôle mineur dans la genèse des crues éclair, on ne peut cependant conclure que la végétation est sans impact. Ainsi, dans une revue des recherches françaises sur l'impact de la forêt méditerranéenne sur l'hydrologie, Cosandey et al. (2005) montrent que considérer que l'impact de la forêt sur les crues est nul serait aussi erroné que de considérer qu'il est très important. Les auteurs considèrent que l'influence de la végétation est probablement aussi importante que celle des incertitudes inhérentes à la mesure hydrologique (dont nous avons vu qu'elles pouvaient être très importantes au chapitre 1). Les principales conclusions de Cosandey et al. (2005) sont que : (i) la présence de végétation, quelle qu'elle soit, modifie la dynamique des crues par rapport à un sol nu, jamais boisé, (ii) le sol de la forêt joue un rôle très important, si bien qu'une déforestation a un petit impact sur l'hydrologie, du fait de la persistance du sol forestier, (iii) l'impact du reboisement, s'il existe, est si faible qu'il est masqué par la seule variabilité climatique.

C'est après avoir potentiellement subi l'évapotranspiration et l'interception que l'eau va s'égoutter et/ou s'écouler le long des tiges et des troncs pour rentrer en contact avec le bassin à proprement parler.

#### *2.1.1.2 Influence des cultures*

Selon Cosandey et al. (2003), les cultures favorisent le ruissellement et l'imperméabilisation par la formation de semelles de labour dues au passage fréquent d'engins agricoles. Gaume (2002) cible dans les régions méditerranéennes les vignes, qui, en plus d'être assez imperméabilisées par les semelles de labour et les terrains argileux, ont souvent leurs sillons organisés dans le sens de la pente (photo 5).

Les plantations en tant que telles n'ont pas plus d'influence que la végétation naturelle : on y observera l'interception, l'évapotranspiration, l'égouttement et l'écoulement le long des tiges et des troncs. En revanche, l'organisation des alternances sol nu /cultures et les tassements réalisés par les engins facilitent le ruissellement de surface.



Photo 5 : vignoble en terrain argileux avec sillons dans le sens de la pente, Saint-Hilaire-de-Brethmas (30) ; importants ruissellements suite à un orage qui a laissé 62mm dont 42mm en 25 minutes le 24/09/2012. Photo : Artigue 2012.

### 2.1.1.3 Influence des zones artificielles

Les zones urbanisées constituent la majorité des zones artificielles. Les terrains artificialisés sont très imperméables et donc très favorables aux ruissellements. En témoignent les importantes inondations par ruissellement urbain qui surviennent dès que des pluies intenses et prolongées frappent une agglomération.

En revanche, à l'échelle du bassin versant, l'impact est assez faible. Selon Cosandey et al. (2003), l'urbanisation doit couvrir une proportion du bassin très importante pour avoir un impact sensible sur la réponse hydrologique, ce qui, nous l'avons vu dans le chapitre 1, est rarement le cas dans les zones qui nous intéressent. En définitive, une réelle aggravation des conditions hydrologiques ne peut avoir lieu que sur des bassins presque exclusivement urbanisés ou plus localement, en ville, par accélération de la saturation des réseaux d'évacuation d'eaux pluviales.

En Cévennes, on trouve par ailleurs de nombreux aménagements anciens : les terrasses agricoles, dont les divers types portent les noms patois de *faïsses* ou de *bancels* (photo 6), ainsi que les *tancats*, modestes retenues bâties sur les petits cours d'eau, témoignent du travail considérable d'aménagement effectué par les habitants pendant des siècles. Martin et al. (2008) ont étudié le rôle hydrologique de ces aménagements et en déduisent qu'aucun d'entre eux ne joue un rôle de mitigation des crues.

Quoi qu'il en soit, le Code de l'Environnement, au travers de la loi 2006-1772 du 30/12/2006 sur l'eau et les milieux aquatiques prévoit que soient compensées les imperméabilisations résultant des aménagements et infrastructures nouveaux. La mission inter-services de l'eau de l'Hérault (département voisin et représentatif de la zone d'étude) préconise au choix :

- une compensation à hauteur de 100 litres par mètre carré imperméabilisé,
- la vérification d'une protection centennale et un débit de fuite calé sur l'occurrence biennale existante.
- une non-aggravation des débits par rapport à la situation existante.

Le volume retenu est celui de la contrainte la plus forte.



Photo 6: *faïsses* près de Gravières (07) en aval du Chassezac ; Photo : Artigue 2012.

### 2.1.2 Incidence directe de la pluie sur le bassin

Une fois passée l'interface, là où elle existe, la pluie entre en contact avec le bassin. Cette partie décrit les processus qui se produisent au niveau de ce contact.

#### 2.1.1.1 *Ruissellements de surface*

Le processus de ruissellement est lié, de façon plus ou moins directe, à la capacité d'infiltration du sol. On en distingue plusieurs types qui vont être énoncés ici.

Le premier, et probablement le plus connu, est le ruissellement dit « hortonien » en référence à Robert E. Horton qui a développé cette théorie en 1933.

Le ruissellement de type hortonien implique une saturation du sol par le haut. Il exige que la capacité d'infiltration du sol soit dépassée par l'intensité pluvieuse. L'étude initiale menée par Horton portait sur des bassins versants semi-arides avec peu de sol en milieu naturel (Ayrat, 2005). Il faut aussi noter qu'un certain nombre de zones d'ombre persistent : Beven (2004) s'interroge, au regard des archives d'Horton, sur l'existence d'une contribution souterraine plus importante qu'envisagée sur les zones étudiées en détail par l'auteur dans les années 1930.

Quoi qu'il en soit, la méthode est difficilement transposable à d'autres bassins, en particulier les bassins cévenols qui possèdent une végétation assez dense et parfois même quasi impénétrable.

On peut tout de même observer ce ruissellement dans certaines zones bien particulières. Par exemple, selon Soltner (1990), les sols limoneux manquant d'argile et de calcium favorisent le ruissellement hortonien grâce à la formation d'une croûte de battance issue du réarrangement des éléments de la surface du sol.

Ambroise (1998), Gaume (2002), Cosandey et al. (2000), Cosandey (1994), Obled et al. (1999), entre autres, se montrent très critiques envers cette hypothèse de génération du ruissellement. Ils notent que l'on n'observe que peu de ruissellements sous les épisodes pluvieux intenses, ce que nous pouvons confirmer, en particulier en milieu cévenol. Ils opposent aussi des études géochimiques montrant que les écoulements en crue contiennent souvent une majorité d'eau ayant transité par le sol. Elles ne peuvent donc pas provenir d'un ruissellement de surface (Gaume, 2002).

Cappus (1960) puis Hewlett et al. (1967) décrivent un autre type de ruissellement, avec une saturation qui se produit cette fois par le bas. Dans ce cas, le bas de l'horizon de sol se sature d'abord, la saturation gagnant ensuite, comme pour un remplissage, la surface. Là, elle empêche les nouveaux apports pluviométriques de s'infiltrer et provoque du ruissellement. On parlera d'écoulement sur surfaces saturées. Selon Musy et al. (2009), les pentes faibles et les sols peu épais sont favorables à ce type de saturation « par le dessous ». Dunne et al. (1970) concluent d'expériences dans de petits bassins du Vermont (Etats-Unis-d'Amérique) que les zones favorables au ruissellement de surface se situent près des drains et sont alimentées par les écoulements hypodermiques. Le phénomène décrit s'apparente à une saturation par le dessous, l'intensité pluvieuse n'étant pas suffisante pour provoquer du ruissellement sans la saturation du sol par les apports hypodermiques du versant.

Le ruissellement par exfiltration est la restitution d'eaux souterraines lorsqu'une nappe affleure, chargée par les abondantes précipitations. Il ne doit pas être confondu avec les exurgences karstiques, ou les sources temporaires à régime pluvial, qui donnent généralement lieu à des écoulements concentrés. Dans le contexte cévenol qui nous intéresse ici, cette exfiltration est assez rare du fait de la rareté des nappes en tant que telles (pentes fortes, perméabilité très faible passés les premiers horizons altérés).

Enfin, peu prégnante dans la dynamique hydrologique des crues rapides, la rétention superficielle représente la fraction de pluie qui ne sortira du système hydrologique (par l'un des processus ici décrits) que plus ou moins longtemps après l'épisode pluvieux.

On peut remarquer que dans la catégorie des ruissellements de surface, seul le ruissellement par exfiltration serait à même d'expliquer la forte teneur en eau ayant transité par le sol. Pourtant, la morphologie des terrains cévenols n'est pas adaptée à une telle hypothèse. Par ailleurs, le fait que le ruissellement soit très rarement observé sous les épisodes cévenols intenses (hors zones agricoles, sols nus et zones artificielles) met à mal l'idée d'une contribution importante du ruissellement de surface.

Dans tous les cas et pour diverses raisons, aucun des types de ruissellement de surface ici évoqués ne semble en mesure d'expliquer les crues observées en Cévennes. Il faut donc poursuivre notre investigation en direction des écoulements se produisant sous la surface.

### *2.1.1.2 Processus de sub-surface*

Nous admettrons dans cette partie que le terme « sub-surface » concerne les écoulements ayant lieu dans le sol mais ne concerne pas les transferts profonds. Musy et al. (2009) remarquent que les notions liées à l'écoulement de sub-surface sont encore loin de faire l'unanimité parmi la communauté des hydrologues.

On peut penser que si le ruissellement de surface ne suffit pas à expliquer les crues observées, c'est donc que l'eau s'infiltre. Dans ce cas, les transferts de sub-surface doivent entrer en compte pour conduire l'eau vers les drains et provoquer les écoulements concentrés.

Si cette hypothèse est vérifiée, cela signifie que la conductivité latérale est supérieure à la conductivité verticale, ou en tout cas suffisamment importante pour éviter un ruissellement avec saturation par le dessous. De même, la conductivité verticale doit être suffisamment forte pour éviter le ruissellement par dépassement de la capacité d'infiltration maximale.

Logiquement, on trouve ce type de processus dans des terrains à pente importante avec un substratum peu perméable et des sols suffisamment épais pour éviter une saturation par le dessous. Ces conditions correspondent à peu près à celles rencontrées dans le massif cévenol.

Les auteurs recensent plusieurs moyens de mener à bien ce transfert latéral.

En premier lieu, l'effet piston (écoulement de subsurface par transfert d'énergie) est avancé par certains auteurs comme Musy et al. (2009), Hewlett et al. (1963) ou encore Rasmussen et al. (2000). Selon eux, ce mécanisme suppose qu'une impulsion d'eau reçue par le versant est propagée vers l'aval par le biais d'une onde de pression. Une exfiltration immédiate se produit alors à l'aval. C'est là l'origine de l'image du piston ou de la seringue.

Dans ce cas, il faut supposer que la vitesse de transfert de l'eau dans le versant est faible, alors que l'onde de pression est rapide et provoque un afflux d'eau immédiat dans le drain, en tout cas au pied du versant. La vitesse apparente de l'eau est donc rapide.

Cette hypothèse a l'avantage d'être en mesure d'expliquer la rapidité de la réaction de certains bassins de notre zone d'étude ainsi que d'être cohérente avec les différents constats sur la teneur en eau « ancienne » des eaux de crue.

En revanche, il faut que la capacité de stockage du sol soit faible pour que l'impulsion soit d'un volume équivalent à l'exfiltration qui aura lieu en bas de versant (Musy et al., 2009).

Les auteurs évoquent aussi les écoulements par macropores. Musy et al. (2009) notent qu'ils se voient attribuer le rôle d'accélération de la recharge de nappe en favorisant le déclenchement de l'effet piston. La présence de macropores dans les sols est liée à la végétation, à la faune (fouisseurs), à la structure du sol ou à la fracturation. Marchandise (2007) note que les dynamiques d'écoulement dans les macropores s'apparentent à celles des conduits et qu'ils sont utilisés préférentiellement par l'eau par minimisation de la déperdition d'énergie.

A l'échelle du versant, la bonne conductivité induite par les macropores nécessite un degré de connectivité important pour accéder à la rivière rapidement, du moins à une zone saturée qui y contribue. Il apparaît très complexe de quantifier l'impact réel de ces macropores puisqu'ils sont souterrains et de natures variées.

Enfin, les contributions de sub-surface par intumescence de nappe ou par écoulement de retour ne seront pas détaillées ici en raison de leur incompatibilité avec les régimes hydrologiques et les morphologies de notre zone d'étude.

### *2.1.1.3 Stockage et pertes*

Au-delà des pertes par interception et évapotranspiration qui sont, on l'a vu, négligeables dans les cas d'épisodes brefs et diluviens, d'autres types de perte peuvent être observées.

Pour éliminer tout de suite les moins importantes, on peut citer l'évaporation directe à la surface de l'eau des drains. Pendant un évènement hydrométéorologique intense, la surface en eau augmente mais c'est aussi durant ces épisodes que l'ensoleillement, l'humidité de l'air et le point de rosée sont très défavorables à l'évaporation. D'ailleurs, à ces échelles de temps (l'heure), le phénomène est forcément marginal au regard des volumes précipités.

De même, les transferts à un autre bassin versant sont très marginaux dans la zone qui nous intéresse, du fait de la faible perméabilité du substratum. Les quelques terrains très perméables ne sont pas suffisamment vastes et continus pour permettre des transferts notables dans la zone d'étude.

Les pertes des drains vers les aquifères souterrains non karstiques se produisent surtout en plaine où la nappe est suffisamment vaste. Sur les crues rapides des reliefs cévenols, ce phénomène est négligeable.

Les pertes d'origine karstique qui affectent certains drains sont, quant à elles, limitées par le diamètre des conduits et fissures qui les autorisent. Lorsqu'une rivière roule plusieurs centaines, voire plusieurs milliers de mètres cube par seconde, il est difficile d'imaginer que des pertes de quelques mètres cube par seconde affectent significativement le débit de pointe.

Même si ce n'est pas le cas en Cévennes, les aquifères karstiques peuvent, lorsqu'ils occupent une part importante du bassin, modifier sa dynamique hydrologique. Les terrains très perméables affleurants alimentent alors le réservoir souterrain avant de contribuer aux écoulements de surface, s'il pleut suffisamment, ou si le réservoir est suffisamment rempli. C'est par exemple le cas du karst du Lez, dans l'Hérault, dont l'aquifère source soumis à des pluies intenses de type méditerranéen peut montrer des comportements très variés : stockages momentanés, restitutions différées, fonctionnements de trop-pleins... (Kong A Siou, 2011) (Coustau et al. 2011).

### 2.1.2 Ecoulements

Les drains, en tant que chenaux découverts, concentrent les écoulements. Une fois les premiers d'entre eux atteints, l'eau commence à participer aux écoulements de crue.

Nous avons présenté que, bien que difficilement quantifiable, la contribution des versants dans notre zone d'étude était essentiellement représentée par des apports de sub-surface. Hewlett et al. (1963) montrent qu'un écoulement de sub-surface en régime non saturé peut devenir l'écoulement de base si les terrains présentent de fortes pentes. Dans ce cas, une zone saturée de taille réduite est située en pied de versant et est alimentée en continu par les écoulements non saturés du versant.

Ainsi les écoulements seraient générés dans les premiers drains puis, rapidement, dans les drains principaux où la crue se propage de surcroît. Plusieurs facteurs vont conditionner la réponse à partir de ce stade.

En tout premier lieu, la géométrie du bassin conditionne fortement le type de réponse, plus ou moins brève et élevée, que le bassin va fournir à son exutoire (figure 19). On comprend ici que

les affluents d'un bassin de forme ronde apporteront de façon plus rapprochée dans l'espace leur contribution au drain principal, alors que les affluents d'un bassin de forme allongée apporteront leur contribution de façon plus dispersée, induisant un « aplatissement » de l'hydrogramme de crue. Ainsi, à volume écoulé égal, le bassin ne répond pas de la même manière suivant sa forme.

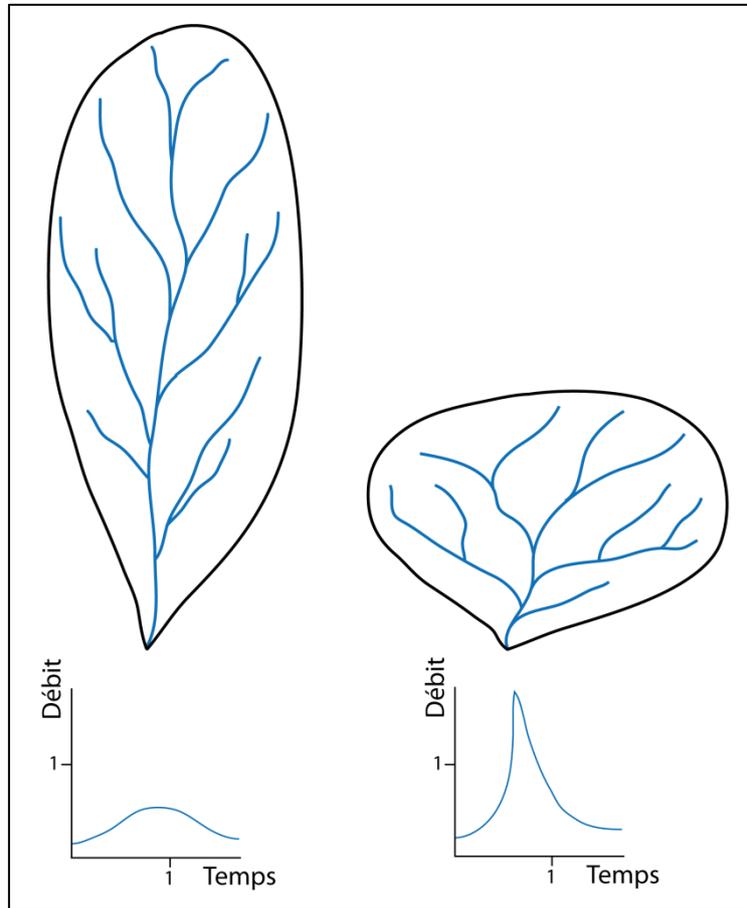


Figure 19: aspect de l'hydrogramme observé en sortie de deux bassins de formes différentes soumis à des précipitations semblables (d'après Musy et al., 2009).

C'est donc logiquement que la densité de drainage va, elle aussi, jouer un rôle important. Elle représente la longueur des drains du réseau hydrographique par unité de surface. Une densité de drainage élevée sera un facteur d'accélération de la réponse puisque le temps de transfert dans les drains est bien souvent inférieur à celui observé dans les versants. Si les drains sont plus nombreux, le cheminement de l'eau jusqu'à ces derniers est plus court donc la réponse accélérée. Par ailleurs, une forte densité de drainage témoigne d'une certaine imperméabilité des terrains et se rencontre souvent dans les zones de forte pente.

En sus de la forme du bassin et de l'organisation de son réseau hydrographique, sa surface joue un rôle important, en particulier dans les débits de pointe qu'il peut fournir. De nombreuses lois existent pour relier surface et débits ou débits spécifiques de pointe (Réménieras, 1986). Nous aurons l'occasion de les détailler plus bas mais on notera simplement que ces lois sont généralement spécifiques à une région avec sa morphologie et sa pluviosité annuelle.

Notons ici que les apports de la pluie directement en surface du cours d'eau ne sont que très faiblement contributifs, le rapport entre la surface en eau et la surface du bassin étant la plupart

du temps très faible pour les petits bassins versants qui nous intéressent ici. Dans le cas de grands bassins avec de vastes plaines d'inondations, l'apport est nettement plus conséquent mais ce n'est pas le cas pour les bassins de la zone d'étude.

La propagation qui s'effectue dans le drain en lui-même verra se produire un amortissement dont l'ampleur peut être très variable. Elle dépend en effet de plusieurs facteurs comme la morphologie et de la topographie du chenal d'écoulement, ou encore la rugosité des matériaux qui composent le chenal. Dans ce cadre, les aménagements anthropiques tels que les digues ou les barrages modifient la dynamique de cette propagation. Les digues offrent des surfaces régulières au courant, tendant à l'accélérer, et resserrent le lit augmentant ainsi la hauteur d'eau. Les barrages tendent, quant à eux, à casser la dynamique de pic observée dans les crues rapides, ce qui est, dans certains cas, plus qu'un simple comportement, mais leur fonction principale.

Les lois de l'hydraulique permettent d'évaluer correctement ces processus de propagation en tenant compte des différents paramètres évoqués ici, mais comme dans le cas de l'hydrologie, elles sont plus fiables dans un contexte d'abondance des mesures.

### 2.1.3 Conclusions

Nous avons pu voir qu'il est possible d'identifier presque un à un les processus et les caractéristiques qui influent sur la relation entre pluie et débit à l'exutoire. De la plus haute feuille de la canopée au chenal d'écoulement le plus aval, une explication existe pour chaque information apportée par l'observation.

Le point commun à tous les phénomènes ici évoqués est la difficulté éprouvée par la communauté scientifique à quantifier leur importance, ne serait-ce que relative. Il faut dire que la tâche est compliquée, on l'a vu dans le chapitre 1, par la difficulté à effectuer des mesures *in situ* fiables mais aussi par la multiplicité des phénomènes à prendre en compte. Rien n'indique en effet qu'un phénomène prédominant dans un bassin le sera dans l'autre. En fait, les questions soulevées ici mettent surtout le doigt sur le problème de la grande variabilité des comportements hydrologiques en fonction des caractéristiques du bassin versant mais aussi en fonction de l'échelle que l'on considère. La combinaison de tous ces mécanismes rend par ailleurs la relation entre pluie et débit typiquement non linéaire.

## 2.2 Vers la prévision des crues éclair

Les enjeux humains et financiers mis en jeu dans la problématique de la prévision des crues éclair sont, on l'a vu, très importants. L'idée que les processus hydrologiques doivent être compris et modélisés pour parvenir à fournir une estimation fiable de la relation entre pluie et débit s'est rapidement imposée comme une nécessité.

### 2.2.1 Modélisation hydrologique

#### 2.2.1.1 *Quelques définitions*

Un modèle se définit, selon l'UNESCO et l'OMM (1992) comme la « *représentation, sous une forme quelconque, d'un objet, d'un processus ou d'un système. En hydrologie, un modèle est la plupart du temps une représentation mathématique d'un bassin, d'un système d'eau, d'une série de données, etc.* »

Dans ce mémoire nous nous limiterons aux modèles mathématiques qui sont les plus nombreux et les plus commodes à concevoir grâce aux capacités de calculs des ordinateurs. Dans ce contexte, on peut schématiser un modèle et ses éléments comme présenté dans la figure 20.

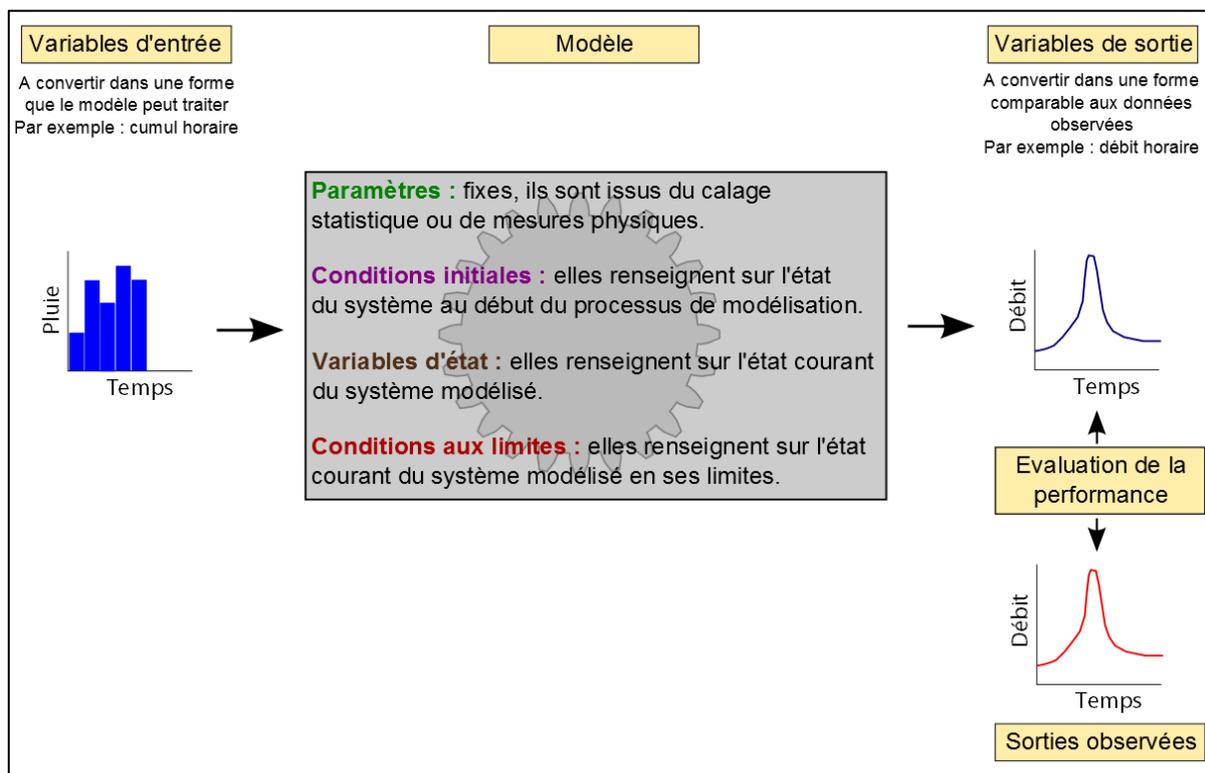


Figure 20: représentation schématique d'un modèle hydrologique mathématique pluie-débit.

De nombreux auteurs fournissent des définitions sur les différents éléments utilisés en modélisation hydrologique (Gaume, 2003 ; Borrell, 2004 ; Mathevet, 2005 ; Marchandise, 2007 ; etc.). A partir de ces définitions et en ajoutant quelques apports, nous avons pu dresser la liste suivante :

- Variables d'entrée : ce sont les données d'entrée du modèle. Dans le cas qui nous intéresse, il s'agit essentiellement de mesures de pluie mais il peut y avoir des données d'évapotranspiration, de températures,... Le pas de temps de la mesure est pertinemment choisi en fonction des contraintes opérationnelles et de la variabilité temporelle du phénomène observé. A cette fin, une analyse de l'évolution de la variance du signal permet de définir le pas de temps maximal pour lequel aucune perte d'information n'est subie. Ce travail permet de limiter la quantité de données au strict nécessaire.
- Variables de sortie : ce sont les données de sortie du modèle. Dans le cas qui nous intéresse, il s'agit de débits, de débits spécifiques ou de hauteurs d'eau. Les données de débit sont généralement préférées à celles de hauteur d'eau étant donné qu'elles sont moins locales (leur relation au lieu de mesure est moins forte) et qu'elles apportent une information volumique liée à l'état du bassin versant.
- Variables d'état : ce sont des variables qui représentent l'état du système hydrologique. Les variables de sortie peuvent être des variables d'état : il est par exemple tout à fait possible de considérer le débit à l'exutoire comme une information d'état du bassin versant. Elles peuvent aussi être des variables d'entrée, si la variable est connue tout au

long du processus, par mesure ou par résultat d'une autre modélisation (humidité des sols, piézométrie, résistivité...).

- Paramètres : ils sont le liant des différentes variables en tant qu'éléments des équations constitutives du modèle. Ils sont fixés dans le cas de modèles non adaptatifs, soit par mesure *in situ* ou par prédétermination arbitraire s'ils ont une signification physique. S'ils n'ont pas de signification physique, ils sont fixés par calage.
- Conditions aux limites : elles représentent les processus d'échanges dits latéraux aux frontières du système hydrologique (flux évaporé, évapotranspiré, intercepté, pertes souterraines ...), à son entrée (par exemple, le débit entrant si seule une portion de bassin est modélisée) ou encore à sa sortie (par exemple, niveau aval imposé par les marées pour un fleuve ou niveau du confluent pour un affluent).
- Conditions initiales : elles représentent l'état dans lequel se trouve le système au début de la période modélisée (teneur hydrique des sols, niveau de remplissage d'une retenue...).

### 2.2.1.2 Taxonomie des modèles hydrologiques

Il est particulièrement délicat de proposer une classification des modèles hydrologiques, cette dernière étant très dépendante de l'angle d'approche. De nombreux auteurs se sont essayés à cet exercice. Nous choisissons ici de présenter la vision de Borrell (2004), confirmée par des références complémentaires :

- Stochastique / déterministe : dans un modèle stochastique, l'une au moins des variables de forçage, des variables d'état ou des paramètres est une variable aléatoire. Par voie de conséquence, la (les) variable(s) de sortie(s) est ou sont une (des) variable(s) aléatoire(s) (Gaume, 2003).  
Un modèle déterministe associe systématiquement à un jeu de variables d'entrée et de paramètres une sortie unique.  
Les modèles déterministes sont les plus utilisés en hydrologie, en particulier pour la prévision des crues rapides. En effet, selon Borrell (2004), les modèles stochastiques sont le plus souvent établis à partir de longues chroniques temporelles pluie-débit, qui font généralement défaut dans le cas des crues éclair. Les modèles déterministes semblent donc être les plus adaptés à notre problématique. Ceci dit, le déterminisme a aussi ses limites : selon Montz et al. (2002), la prévision des crues éclair ne peut être décrite par une approche déterministe à cause de la complexité des processus mis en jeu et par voie de conséquence du manque de connaissance sur ces processus.
- Global / Distribué : un modèle global intègre toutes les propriétés du bassin versant et des forçages, et les résume de telle sorte que la relation soit directe entre pluie et débit. Ces modèles simples trouvent leurs limites dans les bassins qui possèdent des spécificités locales très influentes sur le comportement hydrologique général.  
Un modèle distribué s'attache à tenir compte des hétérogénéités présentes dans le bassin en distribuant les fonctionnements hydrologiques de différentes zones pour obtenir un débit estimé à l'exutoire. Le fonctionnement de ces modèles est plus lourd étant donné qu'ils nécessitent soit des données géographiques soit des mesures *in situ*.
- A base physique / empirique : la transition entre ces deux types de modèles est très progressive puisqu'elle dépend de leur niveau de conceptualisation. Un modèle à base physique pur est fondé uniquement sur les équations de base de la physique (Gaume, 2003).

Un modèle purement empirique n'utilise pas les lois de la physique et est calé à l'aide de chroniques d'observations. C'est le cas des modèles par apprentissage statistique comme les réseaux de neurones (cf. chapitre 3). Cependant, l'absence d'équations physiques dans le modèle ne signifie pas qu'aucun comportement physique du système modélisé n'est pris en compte. Il est possible, comme nous le verrons dans le cas de la sélection de variables, de diriger un modèle empirique vers un type de fonctionnement en soignant la façon dont on lui présente les variables de forçage.

- Évènementiel / continu : un modèle évènementiel utilise une fenêtre temporelle de fonctionnement spécifique à un évènement hydrologique : crue lente ou rapide, étiage... Il est spécialisé dans le type d'évènement pour lequel il a été conçu. Un modèle continu fonctionne en permanence, ce qui implique qu'il doit être performant aussi bien sur les étiages prononcés que sur les crues éclair, et ce en l'espace de quelques jours, voire de quelques heures. Ceci est en particulier vrai dans les régions méditerranéennes. A ce sujet, Beven (1989) rappelle qu'il est utopique de vouloir modéliser le fonctionnement de l'ensemble des relations pluie-débit tant elles sont complexes. Les crues éclair représentent une part temporelle infime du fonctionnement d'un modèle continu et il paraît effectivement bien délicat d'arriver à modéliser simultanément les processus de long terme et les processus très rapides mis en jeu dans les crues éclair.
- Modèle conceptuel : le terme est souvent employé pour définir des modèles qui ne tiennent pas strictement compte de l'état et des processus physiques du bassin. En fait, comme Borrell (2004), nous considérons que tout modèle est conceptuel par essence puisqu'il est une représentation abstraite d'un système. En définitive, c'est plutôt le niveau de conceptualisation, entre modèle à base physique et empirique, qu'il peut être intéressant de maîtriser.

A travers cette revue synthétique des qualificatifs donnés aux modèles hydrologiques, on a pu mettre en évidence qu'il était difficile, et pas toujours pertinent, de les catégoriser. L'exercice a pour principal intérêt de présenter une revue des définitions des termes employés en hydrologie. On notera cependant que le passage d'un type de modèle à l'autre est souvent progressif, des versions hybrides étant fréquemment employées.

### 2.2.1.3 Qualités attendues d'un modèle hydrologique

Comme Mathevet (2005) le fait, on peut relever dans la littérature de référence les qualités attendues pour un modèle hydrologique (Nash et Sutcliffe, 1970, Bergström et al., 1973 ; Linsley, 1982 ; Klemeš, 1986 ; Bergström, 1991 ; Lindström et al., 1997 ; etc.) :

- Le modèle se base sur des fondements scientifiques sains,
- Les données nécessaires sont en accord avec leur disponibilité,
- La complexité du modèle doit être limitée autant que faire se peut,
- Le modèle doit être général, correctement validé, transposable et robuste,
- Le modèle doit être compréhensible de ses utilisateurs.

Cet état des qualités que l'on peut attendre d'un modèle hydrologique est cohérent avec la nécessité, pour un modèle de prévision, qu'il soit couplé ou non à un autre modèle, d'être parfaitement opérationnel en vue d'une utilisation en conditions réelles.

#### 2.2.1.4 Evaluer la performance d'un modèle

Evaluer la performance d'un modèle peut se faire en mesurant l'écart entre estimation et observation. Etant donné que la modélisation hydrologique peut avoir plusieurs objectifs, comme prévoir un pic de crue, une durée de décroissance, ou encore un débit de base, l'utilisateur a la possibilité d'utiliser plusieurs critères de performance. Certains d'entre eux sont construits pour se spécialiser sur certaines parties de l'hydrogramme, d'autres sont conçus pour évaluer l'ensemble de la période modélisée avec le même poids.

Le critère le plus utilisé est le critère de Nash (Nash et Sutcliffe, 1970), ou coefficient de détermination  $R^2$ . Il peut s'écrire de la manière suivante :

$$R^2 = 1 - \frac{\sum_{k=1}^n (Q_p^k - Q_s^k)^2}{\sum_{k=1}^n (Q_p^k - \overline{Q_p})^2}$$

Où :

$Q_p^k$  est le débit mesuré sur le processus à l'instant  $k$ ,

$Q_s^k$  est le débit estimé (simulé ou prévu) à l'instant  $k$ ,

$\overline{Q_p}$  est le débit mesuré moyen sur l'ensemble des  $n$  couples sur lesquels porte le critère.

Ce critère doit son utilisation fréquente en hydrologie au fait qu'il peut être utilisé sur des événements d'intensités très différentes ou sur des bassins de débit moyen également très différents. On remarque que maximiser le critère de Nash (il ne peut être supérieur à 1) revient à minimiser l'écart quadratique moyen entre les sorties estimées et les sorties observées. Lorsque le critère de Nash vaut 1, l'erreur quadratique moyenne est nulle, lorsqu'il vaut 0 la sortie du modèle a un critère de Nash équivalent à ce qu'aurait donné la moyenne du débit sur l'intervalle de temps considéré. Un critère de Nash qui prend des valeurs négatives témoigne de l'incapacité du modèle à fournir une meilleure estimation que la simple moyenne des débits observés. Pour les crues rapides, on considère souvent qu'à partir de 0,8, le critère de Nash est satisfaisant.

Cependant comme ce mémoire porte sur la modélisation et la prévision des crues, nous devons soulever une limite du critère de Nash bien connue : l'erreur calculée est relative à tout l'événement pris en compte ; ainsi un débit simulé comportant une forte erreur sur le pic de crue mais une très bonne prise en compte de la montée de crue et de la récession, donnerait un bon score de Nash alors qu'à l'évidence ce modèle n'est pas satisfaisant.

Pour une meilleure prise en compte du pic dans l'évaluation de la performance, on peut utiliser un autre critère plus adapté aux crues rapides : le critère de hauteur  $C_h$  (Toukourou, 2009), dont la formule est présentée ici :

$$C_h = \frac{1}{n_{pics}} \sum_{j=1}^{n_{pics}} \left( \frac{1}{5} \sum_{k_j-3}^{k_j+1} \frac{\min Q_j^p, Q_j^s}{\max Q_j^p, Q_j^s} \right)$$

Où :

$Q_j^p$  est le débit observé pour le pic  $j$ ,

$Q_j^s$  est le débit simulé ou prévu pour le pic  $j$ ,

$n_{pics}$  est le nombre de pics de l'épisode et  $k_j$  est l'instant du pic de crue  $j$ .

Plus ce critère est proche de 1, meilleure est l'estimation des pics. Le critère de hauteur ne donne pas d'information sur les pics estimés par le modèle mais non observés. Il convient donc de l'utiliser avec précaution, notamment en contrôlant visuellement la simulation ou la prévision effectuée.

Un décalage temporel de quelques pas de temps entre la sortie estimée et la sortie observée peut conduire le critère de Nash à des valeurs proches de 1, comme le fait par exemple remarquer Moussa (2010). Or, si l'amplitude est un facteur important, la synchronisation temporelle l'est tout autant. Il faut éviter de produire une prévision dite naïve, qui consisterait simplement en une reproduction à l'instant  $k+1$  de l'observation à l'instant  $k$ . Ceci est particulièrement important pour les modèles qui utilisent des données de débit observées durant la phase de calage. Pour prendre en compte cette difficulté, on peut utiliser le critère de persistance (Kitadinis et al., 1980) :

$$C_p = 1 - \frac{\sum_{k=1}^n (Q_p^{k+l} - Q_s^{k+l})^2}{\sum_{k=1}^n (Q_p^{k+l} - Q_p^k)^2}$$

Où :

$Q_p^{k+l}$  est le débit observé à l'instant de l'horizon de prévision  $l$ ,

$Q_s^{k+l}$  est le débit prévu à l'instant de l'horizon de prévision  $l$ ,

$Q_p^k$  est le débit observé à l'instant de la prévision  $k$ .

Si la prévision est parfaite, le critère vaut 1. Si l'on n'obtient que la prévision naïve, le critère est nul. Si le critère est négatif, cela signifie que la prévision est particulièrement mauvaise, n'apportant aucune nouvelle information (Toukourou, 2009). Cette affirmation est quelque peu sévère étant donné que l'on doit pouvoir retrouver pour ce critère des observations proches de celles de Moussa (2010) concernant le critère de Nash.

L'ampleur du pic simulé peut aussi être évaluée en calculant quel pourcentage du pic observé il représente. Simple rapport entre débit estimé et observé, on peut l'appeler *PPD* (*Percentage of the Peak Discharge*) et l'exprimer en pourcents. Il faut aussi évaluer quel est le décalage temporel du pic simulé avec le pic observé.

On peut imaginer un critère plus directement représentatif de la bonne prévision du niveau maximal dans les bons délais. Par exemple, on peut introduire le *SPPD* (*Synchronous Percentage*

of the Peak Discharge) qui, exprimé en pourcentage, représente le rapport entre le débit simulé et le débit observé à l'instant du pic observé  $k_{max}$  :

$$SPPD = 100 \frac{Q_s^{k_{max}}}{Q_p^{k_{max}}}$$

Où :

$Q_s^{k_{max}}$  est le débit simulé ou prévu à l'instant  $k_{max}$  du pic observé,  
 $Q_p^{k_{max}}$  est de débit observé à l'instant  $k_{max}$  du pic.

A l'opérationnel, d'autres critères plus qualitatifs peuvent intervenir comme le franchissement de seuils de débit ou de hauteur d'eau correspondant à l'inondation par le cours d'eau d'enjeux particuliers. Plus locaux et subjectifs, ils doivent être associés à une bonne connaissance du terrain.

#### 2.2.1.5 Conclusions

Pour simuler correctement les processus que nous avons présentés dans la section 2.1, l'hydrologue a imaginé différents types de modèles. Chacun peut avoir un objet différent mais les qualités attendues de ces modèles sont communes. Pour évaluer la performance de ces modèles face à l'objet qui est le leur, des critères de performance sont introduits. Ils permettent de juger de façon objective de la qualité des estimations fournies par le modèle.

Munis de ces outils méthodologiques ou mathématiques, il nous reste à présenter la mise en œuvre de ces modèles sur les bassins rapides du type de ceux de la zone d'étude.

#### 2.2.2 Mise en œuvre des modèles hydrologiques pour la simulation et la prévision

La mise en œuvre des modèles hydrologiques sur des bassins à réponse rapide se heurte à la question de l'indisponibilité des données de pluie prévue à des échelles de temps très courtes (l'heure) et pour des dimensions d'espace très faibles (quelques dizaines de kilomètres). Ce problème n'est pas encore résolu et il faut continuer à considérer la prochaine valeur de pluie selon des hypothèses variées parmi lesquelles : (i) la pluie future nulle, ce qui est évidemment l'hypothèse la plus optimiste, (ii) la pluie future constante (égale à la valeur actuelle), ou (iii) la pluie future égale à la « pluie parfaite » ; ce dernier concept ne peut s'appliquer que lorsque l'on rejoue des événements passés et que l'on se place en mode de prévision ou de simulation avec la disposition des meilleurs prévisions possibles : celles qui ont effectivement été observées. Ce dernier mode n'est pas utilisable en conditions opérationnelles.

La grande majorité des modèles utilisés se fonde sur des hypothèses physiques spécifiques et n'ont pas été conçus, à l'origine, pour répondre à la problématique très particulière des crues éclair. En conséquence, autant pour caractériser cette relation pluie-débit particulière que pour mieux la comprendre, les auteurs ont commencé par appliquer les modèles existants à la simulation des crues éclair, c'est à dire avec un horizon de prévision nul.

Sans être exhaustifs, on peut citer quelques-uns des auteurs qui ont cherché à adapter ces modèles à cette problématique. Volontairement, la mise en œuvre des réseaux de neurones dans la littérature n'apparaîtra pas ici mais dans une section du chapitre 3 qui leur est dédié.

Par exemple, Trambly et al. (2010) ont fait des simulations avec le modèle *Soil Conservation Service Curve Number* (SCS-CN) sur un petit bassin situé sur la zone d'étude et ont pu montrer l'importance de la prise en compte de l'état d'humidité des sols à l'initial dans cette problématique. Dans le but d'améliorer la compréhension du phénomène à partir de simulations, Gaume (2002) a utilisé ce même modèle pour reconstruire des crues passées.

TOPMODEL (Beven et al., 1979 ; Beven et al., 1995) est basé sur la distribution de paramètres topographiques qui influencent la réponse hydrologique. En utilisant ce modèle, Lelay et Saulnier (2007) ont pu montrer que la variable la plus importante à prendre en compte était la distribution spécifique des pluies, ce qui a le mérite de souligner l'importance de la quantité et de la qualité des données de précipitations.

Moussa et al. (2007) utilisent MODSPA sur le bassin versant du Gardon d'Anduze avec un pas de temps variable qui permet de modéliser à la fois les processus rapides et ceux de plus long terme. Avec une calibration multi-sites, ce modèle spatialisé sur la base d'informations géographiques conduit à des simulations intéressantes (critères de Nash proches de 0,9).

En prévision, c'est-à-dire avec un horizon de prévision positif et sans prévision de pluie en entrée, peu d'auteurs proposent des solutions viables à l'opérationnel. Borrell et al. (2005) utilisent MARINE, un modèle physique distribué basé sur les caractéristiques des terrains (topographie, occupation des sols...). Ce modèle donne des simulations de grande qualité mais ne fonctionne pas de façon aussi satisfaisante en prévision, essentiellement du fait d'une grande sensibilité aux étapes d'initialisation des variables naturelles. Ce problème est fréquemment rencontré. Les valeurs d'humidité estimées à partir de relevés de piézomètres ou d'un indice d'humidité peuvent alors être utilisées, apportant une amélioration des résultats pour Coustau (2011) qui a mis en œuvre un modèle de type SCS sur un bassin karstique.

Des travaux plus synthétiques ont été réalisés sur le bassin versant du Gardon d'Anduze. Marchandise (2007) y compare plusieurs types de modèles pluie-débit. L'auteur conclut que, d'une manière générale, les paramètres des modèles ne représentent pas bien la grandeur qu'ils sont supposés représenter. Il remarque aussi des difficultés pour la prise en compte de l'humidité et de l'épaisseur des sols, notamment au niveau de leur variabilité dans le temps et dans l'espace. En d'autres termes, une mesure ponctuelle (par exemple de perméabilité ou de niveau piézométrique) n'est pas forcément représentative des conditions sur l'ensemble du bassin et/ou durant l'ensemble de l'épisode.

Le couplage des modèles météorologiques avec les modèles hydrologiques fournit des résultats intéressants sur de grands bassins produisant des crues assez lentes (Bartholmes et Todini, 2005). Ces auteurs obtiennent plus de 80% du pic à neuf jours d'échéance sur le Pô en Italie septentrionale. On comprend en effet que la variabilité spatiale et temporelle des précipitations est moins primordiale dans le fonctionnement d'un grand bassin que dans le fonctionnement d'un petit. Justement, dans le cas des crues éclair, la faible taille des bassins affectés et les intensités précipitantes très élevées de systèmes convectifs de petite taille ne permettent pas de prévisions météorologiques suffisamment précises pour autoriser un couplage intéressant avec l'hydrologie. A n'en pas douter et comme le font remarquer Younis et al. (2008), des bénéfices certains seraient obtenus d'un tel couplage si la possibilité en était donnée. Déjà, Dolciné et al. (2001) avaient pu montrer l'apport des estimations de pluie futures (à partir de la télédétection) à la prévision des crues sur le Gardon d'Anduze. Malheureusement, les épisodes pluvieux étaient faibles (voire stratiforme pour l'un d'entre eux) et les crues tout autant, puisqu'elles ne

dépassaient pas le niveau de vigilance verte... Les auteurs faisaient toutefois remarquer que les prévisions de pluie n'étaient que d'une utilité limitée pour les très courts horizons de prévision. Alfieri et al. (2011) ont aussi testé une méthode impliquant la prévision de pluie mais par la modélisation météorologique. Les résultats montrent des probabilités de franchissement de seuil cohérentes mais sur un seul évènement, sur le Gardon d'Anduze, en novembre 2008. De l'aveu même des auteurs, la méthode aurait besoin de gagner en robustesse en étant testée sur d'autres épisodes.

En l'absence de prévisions de pluie, une démarche intéressante sur le plan scientifique est souvent adoptée, il s'agit des prévisions présentées comme « pluie parfaites ». Dans ce cas on utilise un modèle de prévision en l'alimentant avec les pluies effectivement mesurées sur le bassin (qui comme nous l'avons vu dans le chapitre 1, sont néanmoins potentiellement assez éloignées des pluies réellement tombées, rendant le qualificatif « parfaites » assez imagé). Ce mode de prévision permet de mieux appréhender quelles sont les erreurs issues du modèle hydrologique lui-même et quelles sont celles qui sont issues du manque d'information sur la pluie future. Il va de soi que le mode de prévision en pluie parfaite n'a aucune forme d'utilité à l'opérationnel puisqu'il met en jeu des variables futures dont le prévisionniste ne bénéficie pas en conditions réelles.

### 2.2.3 Mise en œuvre des modèles hydrologiques sur des bassins non jaugés

La prévision des crues consiste à trouver une relation entre la pluie tombée jusqu'à un temps  $k$  et le débit observé à un temps  $k+l$ , (rappelons que  $l$  est l'horizon de prévision ou *lead time*). Si l'absence de prévision de pluie, ou une prévision de qualité insuffisante, est une difficulté dont on imagine pouvoir s'accommoder, par exemple en limitant l'horizon de prévision au temps de réponse du bassin étudié, l'absence de mesure de débit est plus difficile à compenser pour l'ensemble des modèles nécessitant un calage préalable.

Pourtant, comme cela a été présenté dans le premier chapitre et même si « *une mauvaise mesure vaut toujours mieux qu'un magnifique calcul* » (Rodier et al., 1984), il est nécessaire de répondre au besoin de prévision de l'aléa pour les cours d'eau non jaugés.

L'importance de la question est reconnue par la communauté des hydrologues, par exemple au travers du projet PUB (*Prediction on Ungauged Basins*) initié en 2002 par l'Association Internationale des Sciences Hydrologiques. Elle est également de plus en plus demandée par les acteurs territoriaux ou la protection civile au fur et à mesure que la vigilance crue sur les bassins jaugés se fait plus efficace

#### 2.2.3.1 *Procédé de « généralisation »*

La grande majorité des auteurs s'appuie sur le procédé de généralisation pour produire des simulations et des prévisions sur des bassins non jaugés.

Le procédé de généralisation à des bassins non jaugés consiste à caler et valider un modèle sur un bassin jaugé, puis à généraliser ce fonctionnement à d'autres bassins, non jaugés ou considérés comme tel pour l'expérience, pour une nouvelle validation. On peut soit considérer que les paramètres évoluent en fonction du bassin visé, soit qu'ils sont régionalement homogènes.

Merz et al. (2006) distinguent ces deux types de méthodes de généralisation utilisées par les auteurs :

- L'une basée sur la proximité des bassins l'un avec l'autre : on considère que deux bassins voisins auront un comportement proche. Les paramètres des modèles sont régionalisés de telle sorte qu'ils soient valables sur l'ensemble des bassins d'une zone. Les limites spatiales de la régionalisation restent floues (Merz et al., 2006). De nombreuses informations spatialisées sont nécessaires si l'on veut définir un domaine géographique de validité. Les auteurs qui utilisent cette méthode peuvent obtenir des résultats d'assez bonne qualité si la zone est suffisamment homogène. C'est le cas pour Hundecha et al. (2002) qui après avoir généralisé les paramètres du modèle HBV-IWS à 101 sous-bassins rhénans obtiennent des critères de Nash de l'ordre de 0,9. De même, Goswami et al. (2006) utilisent des données régionales moyennées pour simuler les écoulements avec une efficacité pour le moins inégale (de 0,1 à 0,9 en  $R^2$ ). En considérant qu'un bassin de comportement voisin ne nécessite pas de nouveau calage et donc en jouant sur l'adaptabilité du modèle, notamment aux conditions d'infiltration, Borrell et al. (2002) fournissent des simulations de bonne qualité sur un évènement moyen dans le voisinage de notre zone d'étude (bassin du Vidourle), même si un seul évènement est testé, sans évaluation autre que visuelle de la performance. Yadav et al. (2007) recherchent plus à relier les paramètres des modèles à une dynamique hydrologique qu'aux caractéristiques du bassin (en intégrant ainsi également la dynamique des précipitations). De cette façon, il régionalise des contraintes sur le comportement hydrologique. Cette liaison semble fonctionner sur les 30 bassins anglais testés puisque les fourchettes de prévision englobent la plupart du temps le débit observé. Les auteurs recommandent, comme c'est souvent le cas dans la littérature, d'utiliser le moins possible de paramètres pour préserver une efficacité optimale. Notons que cet auteur définit un domaine de validité pour sa méthode. Gibbs et al. (2011) proposent une méthode de régression permettant de régionaliser les paramètres dans les zones où peu de bassins jaugés sont disponibles. En mode non spatialisé, malgré, des résultats parfois corrects en termes de critère de Nash (jusqu'à 0,75), des décalages temporels et d'amplitude importants sont visuellement notables.
- L'autre est basée plus directement sur les caractéristiques des bassins : on adapte le modèle hydrologique calé sur un bassin jaugé à un autre bassin en faisant évoluer des éléments du modèle en fonction des principales caractéristiques du bassin. Les difficultés de cette approche sont évoquées par Fiorentino et Iacobellis (2006) qui mettent en avant la connaissance insuffisante des processus qui grève le reste de la démarche. Les auteurs reconnaissent, comme beaucoup, la grande utilité des données issues de la télédétection mais aussi la difficulté à les intégrer dans les modèles. Campbell (2006) propose d'ailleurs devant cette difficulté de conceptualiser davantage une partie du modèle (modèle physico-statistique). Il s'agit là d'introduire des statistiques là où les paramètres physiques sont difficiles à relier à des caractéristiques du bassin.

Dans ce cadre, l'utilisation de données géographiques pour l'étude des caractéristiques des bassins s'est généralisée avec l'avènement des techniques relatives aux Systèmes d'Information Géographiques. Oudin et al. (2010) notent, quoi qu'il en soit, que les variations des descripteurs traditionnels des bassins, qui sont généralement des descripteurs physiques, ne correspondent pas bien aux variations des comportements hydrologiques et préconisent de s'intéresser de façon plus approfondie aux propriétés des sous-sols des bassins.

Les réseaux de neurones ont eux aussi été utilisés dans le cadre des travaux sur les bassins non jaugés. Ces travaux seront présentés dans le chapitre suivant qui porte spécifiquement sur les réseaux de neurones.

### 2.2.3.2 Limites de la démarche

Les limites des expériences menées sur les bassins non jaugés sont les mêmes que celles menées sur les bassins jaugés : en l'absence de prévisions de pluie, la prévision est limitée dans le temps et en qualité; la transposition des paramètres par généralisation simple ou régionalisation se heurte à leur représentativité réelle vis-à-vis des caractéristiques du bassin. Ces paramètres ne sont en effet pas réellement physiques et sont plus souvent obtenus par calage statistique que par mesure *in situ*, dans des limites définies par la plausibilité des mesures. Il devient donc d'autant plus complexe d'imaginer ce que pourrait donner une adaptation de ces paramètres (calés sur un bassin jaugé) à un bassin non jaugé en fonction des différences de caractéristiques entre les deux bassins...

Cependant, on trouve des simulations, rarement des prévisions, de bonne qualité mais, trop souvent, le domaine de validité n'est pas clairement défini et interdit une évaluation sérieuse de la robustesse de la démarche.

Face à ces limites, le CEMAGREF (maintenant IRSTEA) en partenariat avec Météo-France ont développé la méthode AIGA (Adaptation d'Information Géographique pour l'Alerte en crue) qui utilise les lames d'eau radar et pluviomètres en temps réel, une expertise humaine, des estimations de déficit hydrique, une évaluation de l'aléa dû à l'écoulement hydrologique induit par la lame d'eau et des informations statistiques hydrométéorologiques spatialisées. Le croisement de ces informations permet d'afficher une échelle de gravité de l'évènement en cours en anticipation de quelques heures (Lavabre et Grégoris, 2005). L'expertise humaine et l'utilisation d'une échelle de gravité en lieu et place de débits prévus sont des démarches originales résolument tournées vers une utilisation opérationnelle.

Naturellement, se pose aussi le problème de la vérification des prévisions fournies. Les travaux sur des bassins jaugés utilisés comme des bassins non jaugés (en ignorant les données de débit) ne sont pas une fin puisqu'il faut bien, à un moment donné, procéder à des vraies prévisions sur des bassins non jaugés, mais il s'agit d'un bon procédé pour valider le modèle. Ainsi, une validation sur un grand nombre d'évènements et un grand nombre de bassins peut délivrer un bon critère de confiance.

## 2.3 Conclusions

En présentant les principaux processus dans l'ordre du trajet de l'eau dans le bassin versant, nous avons pu toucher du doigt la complexité de la genèse des crues. Une grande variété de modèles hydrologiques, chacun basé sur une ou plusieurs hypothèses physiques spécifiques, témoigne de l'intérêt porté par la communauté hydrologique à la fois à la problématique des cyndiniques et à celle de l'hydrologie, toutes deux intimement liées aux crues éclair.

Si les modèles hydrologiques peuvent produire des simulations réalistes, il n'en reste pas moins que les hypothèses qu'ils impliquent ne sont pas toutes vérifiées tant la difficulté à quantifier l'impact respectif des processus mis en jeu est grande. La difficulté à obtenir les mesures physiques en sus de la pluie et du débit (déjà difficiles à obtenir précisément) s'ajoute au

manque de fiabilité du changement d'échelle que l'on peut pratiquer à partir de ces mesures. Dans ce contexte, la représentativité des paramètres et des variables d'état n'est pas garantie.

Synthétiquement, plusieurs éléments sont cependant acquis : (i) la pluie fait augmenter le débit de façon non linéaire, (ii) l'état hydrique des terrains avant et durant l'évènement joue un rôle majeur dans la relation pluie-débit mais n'est pas connue de manière fiable sur l'ensemble du bassin, (iii) la mesure de la pluie est plus pertinente quand elle est spatialisée.

Parmi toutes les approches de modélisation, il en est une qui ne fait pas d'hypothèse sur le fonctionnement du bassin, il s'agit de l'approche statistique qui permet d'identifier une fonction mathématique à partir des signaux de pluie et de débit. En particulier les modèles neuronaux de type *boîte noire* qui procèdent par apprentissage statistique semblent bien adaptés à ce contexte : ils pourraient tirer toutes les informations qui leur sont nécessaires des chroniques de pluie et de débit présentées et pourraient, par sélection de variables, être dirigés dans le sens des quelques processus physiques les mieux appréhendés.

Rappelons que le projet BVNE Gardons a pour vocation d'aboutir à l'élaboration de modèles de prévision opérationnels. Pour cette raison, et en l'état des connaissances actuelles, nous nous proposons d'étudier l'élaboration de modèles hydrologiques statistiques, fondés sur les réseaux de neurones, afin de réaliser la prévision des crues en vue d'une anticipation des inondations sur les bassins mal jaugés dans un premier temps, et non jaugés dans un second temps. Nul doute que l'apprentissage d'un modèle statistique pour un bassin versant dépourvu de mesures de débit est un défi scientifique ambitieux. Cependant, avant de présenter les méthodes utilisées à cette fin, le chapitre suivant proposera une description des modèles neuronaux et de leurs utilisations en prévision des crues.

C'est ainsi que nous choisissons une solution qui nous paraît pouvoir remplir notre objectif rapidement et efficacement, sans pour autant perdre de vue que les travaux d'identification et de quantification des processus restent essentiels à l'hydrologie et permettront dans l'avenir une meilleure compréhension des phénomènes impliqués.



### 3. Les réseaux de neurones formels appliqués à la prévision

Les réseaux de neurones formels forment une classe de modèles non-linéaires paramétrés. Ici, ils sont mis à profit pour des tâches de régression.

L'analogie biologique dans la terminologie employée dans le domaine (neurone, coefficient synaptique, apprentissage) est liée à un souci de compréhension des mécanismes généraux des réseaux de neurones et ne doit pas masquer le fait qu'il est ici question d'apprentissage statistique, purement artificiel. Ainsi, les termes « *neurone* » et « *réseaux de neurones* » désigneront de façon plus simple et respectivement « *neurone formel* » et « *réseau de neurones formels* ».

Les tâches de prévision de la relation pluie-débit par réseaux de neurones s'inscrivent dans le cadre d'une modélisation de type « boîte noire », c'est-à-dire dans laquelle on ne connaît du processus que les entrées et les sorties, sans certitude sur le fonctionnement du système.

Dans ce chapitre, après avoir fourni un certain nombre de définitions relatives aux neurones et aux réseaux de neurones, nous aborderons leurs propriétés fondamentales, mises en perspective avec les besoins de la problématique. Dans un second temps, nous nous intéresserons aux capacités d'apprentissage des réseaux de neurones et verrons quelles réponses peuvent être apportées aux problématiques liées au surajustement.

Enfin, nous effectuerons un tour d'horizon des méthodes employées dans la littérature pour fournir des prévisions hydrologiques par réseaux de neurones.

#### 3.1 Généralités sur les réseaux de neurones formels

##### 3.1.1 Le neurone

Un *neurone* est un opérateur mathématique qui applique une transformation non-linéaire de ses entrées vers ses sorties. En pratique, un neurone effectue une somme pondérée par des coefficients synaptiques de ses entrées. On appelle cette somme le *potentiel*  $v$  du neurone, et les coefficients sont les *paramètres* du modèle. Ce potentiel est transformé par la fonction non linéaire d'activation pour délivrer une sortie (figure 21).

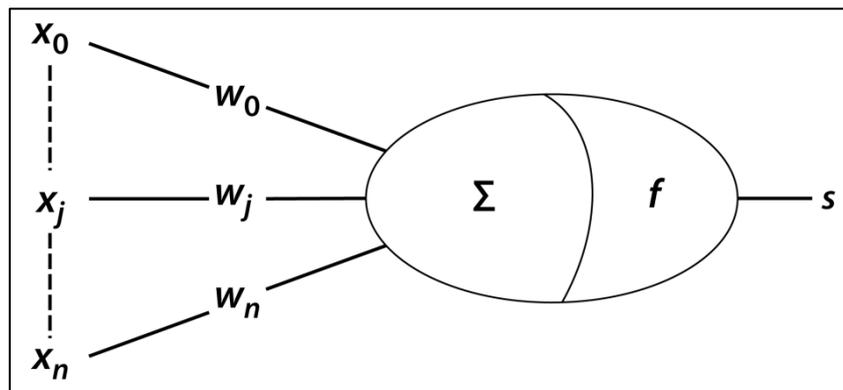


Figure 21: représentation schématique d'un neurone formel.  $x$  représente le vecteur des variables d'entrée,  $w$  la matrice des paramètres, ou coefficients synaptiques,  $f(.)$  est la fonction d'activation et  $s(v)$  la sortie du neurone.

Nous pouvons donc écrire :

$$s = f(x_0, \dots, x_n; w_0, \dots, w_n)$$

Ou plus précisément :

$$s = f(v) = f\left(\sum_{j=0}^n w_j \cdot x_j\right)$$

Où  $v$  est le potentiel du neurone.

Dans les présents travaux, le neurone sera défini par cette relation.

On distingue plusieurs types de fonctions d'activation : fonction échelon (essentiellement utilisée pour la classification), fonction d'activation linéaire (en général l'identité) et la fonction sigmoïde, souvent implémentée par Actg, dont l'intérêt est qu'elle est continue, dérivable, à valeurs bornées et non locale puisqu'elle ne tend pas vers 0 quand la valeur absolue du potentiel tend vers l'infini.

### 3.1.2 Réseaux de neurones

Les neurones peuvent être assemblés pour former un *réseau*. Le réseau est donc une composition de fonctions neurones dotée d'une architecture définie et composée selon la fonction souhaitée. On distingue en général deux types d'architectures : les réseaux à une seule couche de neurones et les réseaux multicouches ; dans ce dernier type de réseau, on distingue deux catégories de neurones :

- Les neurones cachés : leurs sorties ne sont pas des sorties du réseau.
- Les neurones de sortie : leurs sorties sont les sorties du réseau, c'est à dire qu'elles correspondent aux valeurs des sorties mesurées sur le processus.

Les réseaux à une couche sont limités à des fonctions linéaires ou de type seuil, tandis que les réseaux multicouches peuvent effectuer quasiment toute fonction non-linéaire comme nous le présenterons dans la section 3.2.

Selon leur manière de prendre le temps en considération, on distingue par ailleurs deux catégories de réseaux :

- Un réseau de neurones est dit *statique* si le temps n'a pas de rôle fonctionnel. Ainsi il réalise une ou plusieurs fonctions de l'ensemble de ses variables exogènes (variables issues de mesures extérieures). Sans variation de ces variables dans le temps, les sorties ne peuvent varier. Dans ces réseaux la propagation de l'information n'effectue pas de cycle, c'est pourquoi on les appelle aussi « non bouclés ». Par exemple, la figure 22 représente le réseau statique appelé *perceptron multicouche*.

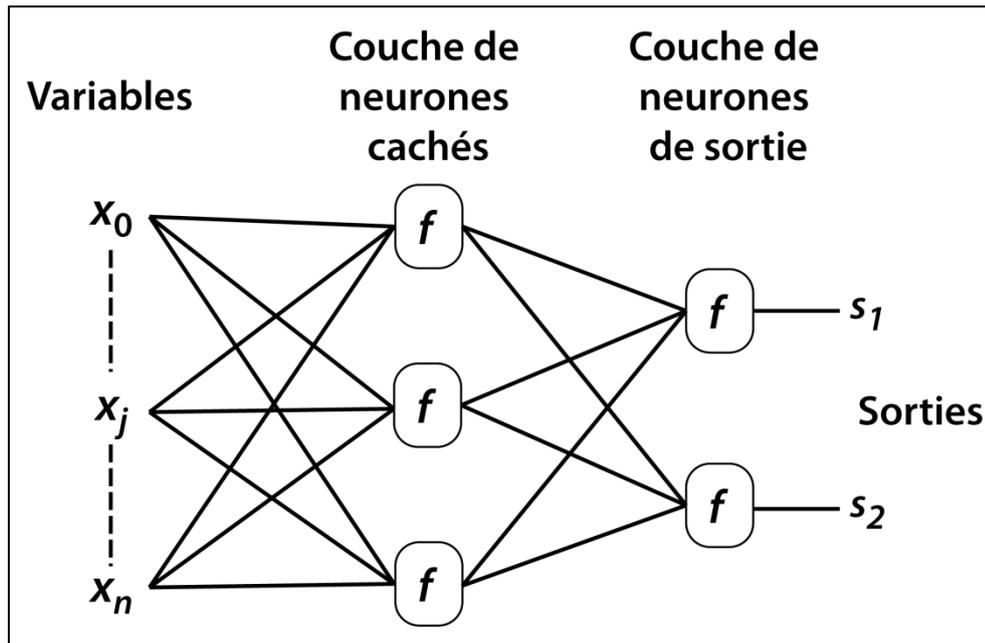


Figure 22: représentation schématique du perceptron multicouche : ici, réseau non bouclé à une couche de neurones cachés.

- Un réseau de neurones est dit *dynamique* s'il réalise une ou plusieurs fonctions de ses variables mais aussi d'une ou plusieurs valeurs passées de ses sorties (sorties des neurones cachés ou de sortie) du réseau. On différenciera alors les variables d'entrée, exogènes, des variables issues du réseau lui-même, dites récurrentes. Ainsi, même si les variables d'entrée sont toutes constantes dans le temps, les sorties du réseau sont susceptibles de varier car les variables récurrentes peuvent alors être non constantes. La représentation graphique de ce type de réseau est cyclique ; on parlera de réseau « bouclé » (figure 23). L'*ordre* du modèle récurrent est égal au nombre de retards pris en compte dans les variables récurrentes.

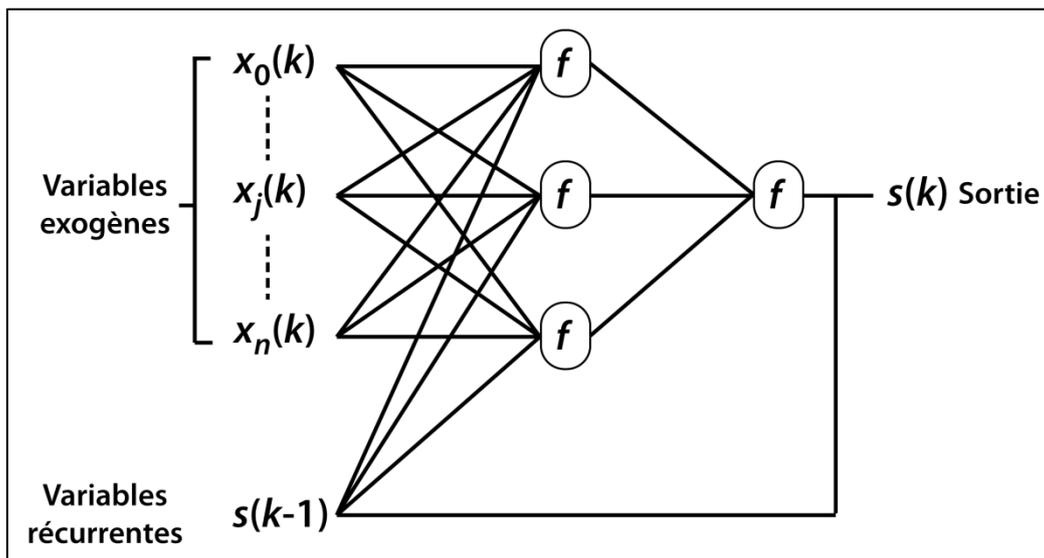


Figure 23: représentation schématique d'un réseau de neurones bouclé d'ordre 1.

## 3.2 Propriétés fondamentales des réseaux de neurones

Les réseaux de neurones, en tant que famille de fonctions non linéaires paramétrées, sont mis en œuvre pour des tâches de modélisation, de prédiction et de classification. Nous présentons dans cette section quelles sont les propriétés fondamentales qui leur permettent d'accomplir ces tâches.

### 3.2.1 L'approximation universelle

La propriété d'approximation universelle des réseaux de neurones a été démontrée par Cybenko (1989) et Funahashi (1989). Elle peut être énoncée comme suit : «*Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation bornée, et un neurone de sortie linéaire.*» (Hornik et al, 1989).

Cette propriété a été démontrée pour le perceptron multicouche possédant des neurones à fonction d'activation sigmoïde sur la couche cachée et, pour l'identification de fonction, un neurone linéaire en sortie. La précision de l'identification est ajustée par le nombre de neurones cachés. Cette propriété fondamentale est une preuve d'existence de solution, mais ne fournit pas de méthode pour calculer les paramètres du modèle. Ce type de réseau est appelé « *approximateur universel* » (figure 24). Il s'agit d'une fonction non linéaire de ses variables et de ses paramètres.

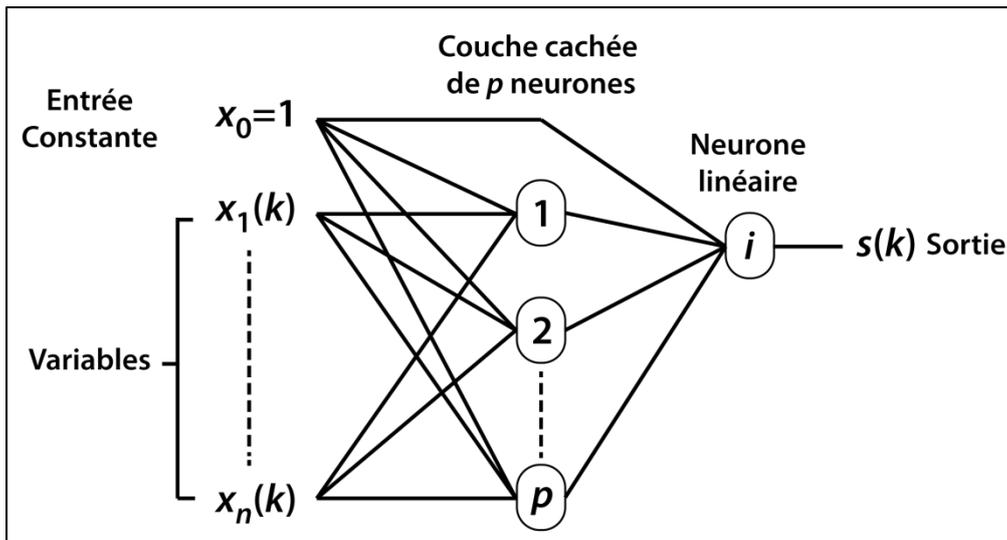


Figure 24: représentation schématique de l'approximateur universel : un réseau de neurones à  $n+1$  entrées,  $p$  neurones et un neurone de sortie linéaire. La variable appelée « entrée constante » est nécessaire pour assurer que la sortie puisse être non nulle même dans le cas où toutes les variables sont nulles.

### 3.2.2 La parcimonie

Barron (1993), a pu démontrer une seconde propriété fondamentale pour l'utilisation des réseaux de neurones, ainsi, si la fonction réalisée par le réseau dépend des paramètres ajustables de façon non linéaire, elle est plus parcimonieuse que si elle dépend linéairement des paramètres. Il a précisé cette propriété en ajoutant qu'un approximateur non linéaire est d'autant plus parcimonieux, par rapport à un approximateur linéaire, que le nombre de variables augmente.

Si l'on ne s'intéresse qu'au nombre de paramètres, les réseaux de neurones sont nettement moins parcimonieux que la plupart des modèles hydrologiques. L'avantage de la parcimonie se ressent principalement en comparaison avec d'autres approximateurs statistiques non-linéaires.

### **3.3 Apprentissage des réseaux de neurones**

L'apprentissage des réseaux de neurones consiste à calculer les paramètres  $w$  du réseau de neurones de sorte que l'on minimise une fonction de coût (c'est-à-dire que l'on pratique une optimisation). Ces données constituent un ensemble : *l'ensemble d'apprentissage*.

#### **3.3.1 Minimisation de la fonction de coût**

Il existe de nombreux types de fonctions de coût. Dans le cas qui nous intéresse, on appelle *fonction de coût* la fonction représentant l'écart quadratique entre les valeurs observées au niveau du processus  $s^p$  et celles estimées par le modèle  $s$ . Il existe plusieurs méthodes permettant de minimiser cette fonction de coût.

Comme la sortie du réseau de neurones dépend non linéairement des entrées et des paramètres, il n'est pas possible de la minimiser de manière algébrique ; on utilise donc pour cela des méthodes itératives. Parmi ces dernières une grande famille de méthodes est celle utilisant le gradient de la fonction de coût.

Pour pratiquer l'apprentissage, la première étape consiste à calculer le gradient de la fonction de coût à l'aide, pour les réseaux multicouches, de l'algorithme de la rétropropagation (Rumelhart et al., 1986.).

Lorsque le gradient est calculé, il reste à mettre en œuvre un algorithme itératif de modification des paramètres. Il existe des méthodes du premier ordre, qui ne prennent en compte que les dérivées premières de la fonction de coût (le gradient), et des méthodes du second ordre qui prennent en compte les dérivées secondes par le biais de la matrice hessienne :

- Les méthodes du premier ordre modifient de façon itérative les paramètres proportionnellement au gradient de la fonction de coût. Le coefficient de proportionnalité utilisé peut-être fixe ou variable.
- Les méthodes du second ordre sont connues pour être beaucoup plus efficaces. De nombreuses méthodes existent ; parmi elles, le choix se fait essentiellement en considérant le nombre de paramètres qui peut rendre le calcul prohibitif pour certaines d'entre elles si ce dernier nombre est trop élevé. La méthode de Levenberg-Marquardt sera utilisée dans les présents travaux. Elle utilise une approximation de la matrice hessienne calculée à partir des produits des dérivées premières (Bender, 1996 et Press et al., 1992). L'utilisation de cette méthode est avantageuse en termes de temps de calcul car elle ne rend pas nécessaire le calcul des dérivées secondes de la fonction de coût. En revanche elle nécessite d'inverser la matrice hessienne qui a comme dimension le nombre de paramètres et peut donc être sujette à des limitations si ce nombre devient trop élevé.

#### **3.3.2 Problème du surajustement**

L'objectif de la mise en œuvre des réseaux de neurones est de rechercher la meilleure approximation de la fonction de régression, compte-tenu des données disponibles. On obtiendra cette approximation en postulant un réseau de neurones et en ajustant ses paramètres pour que

la fonction mise en œuvre par le réseau de neurones s'ajuste aux données présentées dans l'ensemble d'apprentissage.

En conséquence, un modèle trop complexe risque de ne pas s'ajuster à la fonction de régression mais à la somme de cette fonction et du bruit aléatoire présent dans les données. Ce dernier élément peut, on l'a vu dans le chapitre 1, être particulièrement important dans les données hydrométéorologiques. Si le modèle s'ajuste au bruit, il perdra en capacité de généralisation puisqu'il dépendra de la réalisation particulière du bruit dans l'ensemble d'apprentissage. Le modèle apprend alors « par cœur » l'ensemble d'apprentissage, il est incapable de restituer une réponse différente de ce qu'il a appris. On parlera de modèle « *surajusté* ».

En revanche, un modèle trop peu complexe ne sera pas suffisamment flexible pour s'ajuster correctement à la fonction de régression.

Il est donc nécessaire de trouver un juste milieu entre une flexibilité suffisante et un surajustement minimal. On appelle ce problème le « dilemme biais-variance » (Geman et al., 1992). Il est le principal écueil de l'apprentissage statistique et s'exprime mathématiquement de la manière suivante :

Supposant que les observations  $Y(\mathbf{x})$  sur le système à modéliser soient la somme d'une fonction de régression  $q(\mathbf{x})$  et d'une variable aléatoire  $B$  d'espérance mathématique nulle et de variance  $\sigma^2$  qui modélise toutes les perturbations qui affectent la mesure, nous pouvons écrire :

$$Y(\mathbf{x}) = q(\mathbf{x}) + B$$

Dans ce cas, l'erreur théorique de prévision, pour un vecteur  $\mathbf{x}$  donné, c'est-à-dire l'espérance mathématique du carré de la différence entre observation et prévision du modèle, calculée sur tous les ensembles d'apprentissage possibles, est la somme de trois termes (Geman et al., 1992) :

- un terme qui dépend seulement des données : la variance  $\sigma^2$  de la variable  $B$ . Il s'agit de la variance du bruit présent dans les données,
- la variance du modèle qui est calculée sur tous les ensembles d'apprentissage possibles,
- le biais du modèle, qui est le carré de l'espérance mathématique de la différence entre la fonction de régression et la prévision du modèle, calculé, là encore, sur tous les ensembles d'apprentissage possibles.

Comme ces trois termes sont positifs, l'erreur du modèle obtenue par apprentissage ne peut pas être inférieure à l'incertitude sur les données, c'est à dire à la variance  $\sigma^2$  de la variable  $B$ . De plus,  $\sigma^2$  étant fixé pour un bruit  $B$ , si la variance augmente le biais diminue, et inversement.

Ainsi, le biais et la variance évoluent inversement selon la complexité du modèle :

Le biais diminue avec l'augmentation de la complexité : le modèle s'adapte de plus en plus à l'ensemble d'apprentissage. Dans le même temps, la variance augmente rendant le modèle de plus en plus sensible aux détails de l'ensemble d'apprentissage.

La somme de ces deux termes présente un minimum qu'il convient de trouver pour que le modèle offre à la fois une bonne capacité de généralisation et une bonne souplesse. Au-delà de ce minimum, le modèle est surajusté. Ainsi l'erreur sur l'ensemble d'apprentissage ne permet pas d'estimer l'erreur de généralisation.

Ces résultats ne sont exacts que si l'on prend en compte une grande population. De ce fait, ils ont surtout une valeur théorique mais ne sont pas applicables en tant que tel à l'opérationnel.

En pratique, pour concevoir un modèle par apprentissage statistique, il est recommandé de concevoir des modèles de complexité croissante et estimer l'erreur de généralisation sur chacun d'eux. On sélectionnera alors le modèle qui possède les meilleures propriétés de généralisation. Il est celui qui réalise le meilleur compromis entre biais et variance. Le problème est d'autant plus difficile à traiter que l'ensemble d'apprentissage est restreint (Gallinari et Cibas, 1999).

On comprend donc que l'évaluation de l'erreur de généralisation est un des éléments essentiels dans la conception du modèle. Il existe des méthodes qui permettent de l'estimer de façon fiable.

### 3.3.3 Estimer l'erreur de généralisation : la validation croisée

Si l'on évalue l'erreur du modèle avec l'erreur d'apprentissage, on s'expose à ne pas tenir compte du problème du surajustement puisque la différence entre erreur d'apprentissage et erreur de généralisation croît avec la complexité.

La validation croisée (Stone, 1974) permet d'estimer l'erreur de généralisation en soumettant le modèle à des validations multiples avec différentes portions de la base de données. Elle consiste à diviser l'ensemble d'apprentissage en  $D$  sous-ensembles, puis à effectuer l'apprentissage sur  $D-1$  ensembles de la base un nombre  $D$  de fois, en n'excluant jamais deux fois le même ensemble. La moyenne de la valeur de la fonction de coût obtenue pour chacun de ces apprentissages est la valeur qui doit être minimisée (figure 25).

*« La procédure de validation croisée est appliquée à des modèles de complexité croissante, jusqu'à la détection du surajustement, qui se traduit par le fait que le score de validation croisée augmente, alors que l'erreur sur l'ensemble d'apprentissage diminue. On sélectionne alors le réseau pour lequel la plus petite valeur du score de validation croisée a été obtenue »* (Dreyfus et al., 2008).

On calcule donc le score de validation croisée  $V$  pour chaque combinaison de  $D-1$  sous-ensembles de la façon suivante :

$$V = \sqrt{\frac{1}{D} \sum_{E_p=1}^D J(E_p)^2}$$

Où :

$E_p$  est le sous ensemble de données de numéro  $p$  avec  $p = 1$  à  $D$  et comportant  $N_p$  exemples,

$J(E_p)$  est la fonction de coût calculée sur le sous-ensemble  $E_p$ ,



statistique (Mckay, 1992) mais Dreyfus et al. 2008 montrent que l'on peut très bien y parvenir de façon heuristique.

Concernant l'efficacité de ces méthodes, Toukourou et al. (2009) montrent que l'arrêt précoce est plus efficace que la modération des poids pour la problématique des crues éclair sur le bassin-versant du Gardon d'Anduze.

Kong-A-Siou et al. (2012) ont décrit l'évolution du score de validation croisée comparativement sans aucune méthode de régularisation, avec modération des poids ou avec arrêt précoce (figure 26).

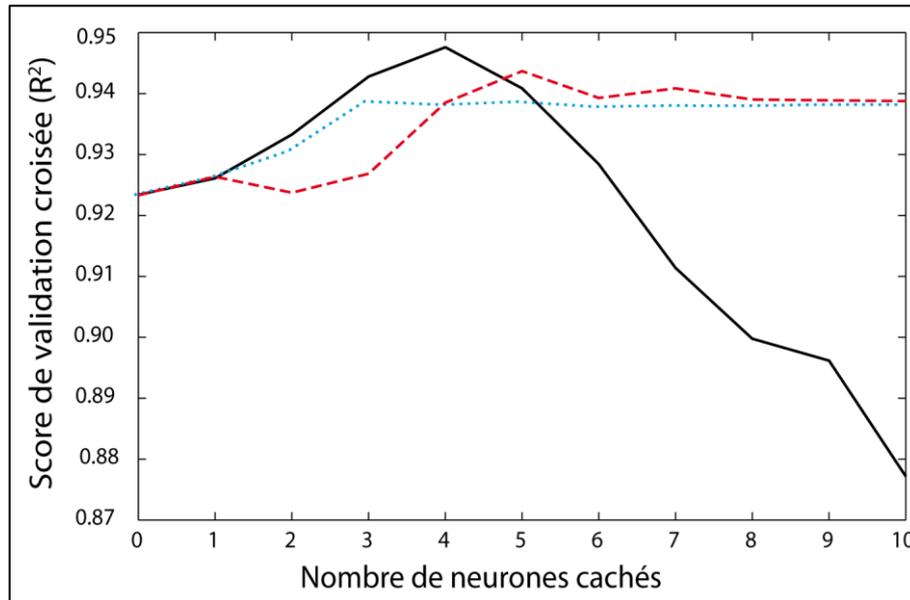


Figure 26: score de validation croisée en fonction du nombre de neurones cachés en trait plein : sans régularisation ; en tirets rouges : avec la modération des poids ; en pointillés bleus : avec l'arrêt précoce. D'après Kong-A-Siou et al., 2012.

Avec la modération des poids comme avec la validation croisée simple les auteurs notent qu'un optimum de complexité se dégage pour un certain nombre de neurones cachés. Ils remarquent aussi que la modération des poids et l'arrêt précoce conduisent à une réduction de la sensibilité du modèle à la complexité induite par le nombre de neurones cachés.

Les auteurs concluent que l'optimum proposé par l'arrêt précoce est plus avantageux. Même si en valeur de score de validation croisée, il est moins élevé, il intervient avec moins de neurones sur la couche cachée ce qui rend le modèle plus parcimonieux.

L'étude suggère qu'il en soit que ces deux méthodes de régularisation sont équivalentes pour des modèles très complexes.

Une fois la complexité rigoureusement sélectionnée, on effectue l'apprentissage sur toutes les données ayant servi à la validation croisée, puis on estime les performances du modèle sur un ensemble de test indépendant, dont les éléments n'ont pas été utilisés auparavant.

### 3.3.5 La sélection de variables

La sélection de variables consiste à choisir un sous-ensemble de variables parmi les variables candidates choisies par les experts comme pouvant être pertinentes pour la modélisation du processus. On notera que la sélection de variables a également pour objet

de diminuer la complexité du modèle puisqu'elle permet de diminuer le nombre de paramètres de ce modèle. Dans le cas où les variables qui influencent le processus sont bien connues, on peut pratiquer cette sélection uniquement de manière experte, à partir des connaissances. En revanche, si le processus est moins bien connu ou si plusieurs variables comportent la même information, mais au travers de méthodes d'observation différentes, on doit mettre en œuvre un processus de sélection de variables qui a pour objet de diminuer la complexité du modèle.

Pour mener à bien un processus de sélection de variables, il faut définir un critère de pertinence pour examiner l'importance des variables. D'après Yacoub et al. (2001), ceci peut être fait par une orthogonalisation de Gram-Schmidt ou par des méthodes heuristiques. Toukourou et al. (2011) ont montré que ce pouvait également être réalisé par le score de validation croisée. Il faut aussi définir un seuil qui permet d'accepter ou de rejeter la variable considérée. Beaudeau et al. (1994) remarquent à l'utilisation de réseaux de neurones pour la prévision de crues turbides qu'une analyse poussée des données hydrométriques disponibles est nécessaire à la sélection d'une architecture aussi simple que possible.

Il peut exister plusieurs jeux de variables de performances équivalentes. En raison du temps de calcul que cela impliquerait, il est souvent impossible de tous les tester. Cependant, le choix d'un jeu performant peut aussi se faire par la connaissance, même relativement limitée, des propriétés du système considéré.

Nous avons vu dans le chapitre 2 que le détail de certains des processus de la genèse des crues rapides restait inconnu. Toutefois, les connaissances actuelles permettent de se faire une idée des grandes lignes de la dynamique hydrologique d'un bassin en se référant aux données hydrométéorologiques observées (relation fortes intensités pluviométriques avec hausse du débit, prise en compte de cumuls antérieurs de pluie pour évaluer l'humidité initiale...). Ainsi la validation croisée pourra être utilisée autant pour effectuer la sélection du modèle le moins complexe, que pour effectuer la sélection de variables.

## **3.4 Prévision par réseaux de neurones**

### **3.4.1 Généralités**

Grâce aux propriétés énoncées dans ce chapitre, les réseaux de neurones ont été très utilisés en hydrologie. La relation pluie-débit, et généralement les phénomènes hydrologiques sont considérés comme dynamiques, non linéaires et souvent difficiles à modéliser ; l'application de l'apprentissage statistique à ce champ scientifique est donc logique (Toukourou, 2009).

Plus particulièrement en ce qui concerne la prévision, celle-ci n'étant pas la vocation première des modèles hydrologiques, il apparaît cohérent d'utiliser des méthodes de type boîtes noires qui peuvent intégrer d'autres phénomènes que les phénomènes purement hydrologiques, (comme par exemple l'absence de données météorologiques futures ou des relations statistiques) en intégrant même une certaine anticipation sur ces données.

De plus, Abrahart et See (2007) montrent que les réseaux de neurones sont adaptés à la modélisation des processus non linéaires mis en jeu dans les relations hydrologiques, à condition qu'ils soient utilisés dans des conditions appropriées (avec une sélection de variables et une régularisation rigoureuses).

Enfin, du fait même du manque de connaissance sur les processus engagés dans le fonctionnement de certains systèmes hydrologiques, l'utilisation de modèles de type boîte noire peut être envisagée très logiquement.

La prévision peut être définie comme l'estimation de la sortie d'un processus physique à un instant postérieur à l'instant présent discret  $k$ . L'horizon de prévision est le laps de temps entre l'instant de prévision et l'instant de cette estimation. Dans le cas d'un modèle discret, comme ce sera le cas ici, l'horizon de prévision est un multiple de la période d'échantillonnage.

Dreyfus et al. (2008) proposent une synthèse de la méthode de conception d'un prédicteur neuronal :

- Le préalable en est l'établissement d'un modèle postulé du comportement du processus non-linéaire pour lequel on souhaite obtenir des prévisions en fonction de la connaissance du mode d'action des perturbations non mesurables.
- Vient ensuite la sélection de variables appropriées pour les entrées et les sorties.
- Puis, vient le choix de la représentation du modèle : soit état (modélisation des conditions dans le système) soit entrée-sortie (modélisation de la sortie à partir des entrées).
- Enfin, il faut choisir l'horizon de prévision souhaité en fonction des variables d'entrée futures disponibles, puis l'ordre du modèle s'il est récurrent et, enfin sa complexité.

### 3.4.2 Application à l'hydrologie

Compte tenu de la présence de bruits et d'incertitudes très importants dans les données hydrologiques, les auteurs se sont heurtés à l'écueil du surajustement et ont cherché à appliquer des méthodes permettant de le limiter. Par exemple, Coulibaly et al. (2000) utilisent l'arrêt précoce alors que Sudheer et al. (2002) s'appuient sur la validation croisée. Malheureusement ces deux études ne font pas l'objet d'une validation sur un ensemble indépendant de l'ensemble d'apprentissage. Les résultats sont donc très probablement surévalués puisque le modèle n'est pas poussé à généraliser à des événements inconnus. Rien ne prouve en effet qu'il se comporterait aussi bien sur des événements inconnus : sa capacité de généralisation est mise en doute.

#### 3.4.2.1 *Adaptation des modèles aux conditions locales*

Les auteurs ont cherché à adapter les boîtes noires neuronales aux conditions locales. Par différentes méthodes (sélection de variables pertinentes, sous-modèles captant un comportement spécifique...) ils arrivent à optimiser les performances de leurs modèles.

Corzo et Solomatine (2007) travaillent à l'optimisation du fonctionnement des réseaux de neurones pour la prévision des crues. Les auteurs utilisent une combinaison de modèles locaux spécialisés pour permettre d'appréhender simultanément les différents processus hydrologiques à l'œuvre dans une crue (débit de base, croissance, décroissance). L'approche donne des résultats satisfaisants sur un bassin assez rapide italien, un autre anglais et un bassin assez lent népalais aux horizons respectifs de deux heures, une heure et un jour.

En créant une typologie des différentes situations hydrométéorologiques, Toth (2009) montre qu'il est possible d'améliorer les performances d'un système de prévision du débit basé sur plusieurs réseaux de neurones. Cette expérience de sélection de variables permet d'utiliser de façon dynamique un « sous modèle » approprié à la situation.

Pereira Filho et Dos Santos (2006) modélisent un bassin versant très urbanisé avec des réseaux de neurones, aidés par des données issues de la télédétection. Ils constatent une grande sensibilité du modèle à la qualité des données ainsi qu'à la quantité d'évènements en apprentissage. Il apparaît nécessaire de disposer d'une base d'apprentissage importante pour tirer le meilleur de l'apprentissage statistique. Ce constat n'est pas surprenant mais il a le mérite de poser des limites à l'utilisation de réseaux de neurones sur des bassins non-jaugés.

Araghinejad et al. (2011) utilise les réseaux de neurones en ensembles pour produire des prévisions hydrologiques probabilistes. Là aussi, un important travail sur les variables est effectué et les résultats qui en découlent sont intéressants.

#### *3.4.2.2 Le traitement de la question de la généralisation*

Le traitement de la capacité de généralisation des prédicteurs neuronaux fait l'objet d'attentions inégales de la part des auteurs.

Schmitz et Cullman (2008) résolvent artificiellement le problème de l'extrapolation (généralisation à des évènements plus intenses) en le supprimant : ils placent l'évènement le plus intense de la base dans l'ensemble d'apprentissage. Cette méthode peut avoir le mérite de fonctionner lorsque la chronique des débits observés est si importante qu'il y a de fortes chances que l'on ait déjà connu et enregistré l'une des crues les plus fortes qui soit.

Concernant les crues rapides, Sahoo et al. (2006) pratiquent la prévision des crues éclair et de la qualité de l'eau avec des réseaux de neurones sur de petits bassins hawaïens. Sans utiliser de méthode de régularisation, les résultats obtenus sont de bonne qualité, certainement grâce à l'absence d'erreur significative dans la mesure pluviométrique sur des bassins de petite taille. Il serait probablement plus difficile d'appliquer cette méthode à de plus grands bassins. Piotrowski et al. (2006) n'utilisent, eux non plus, pas de méthode de régularisation et obtiennent une généralisation de mauvaise qualité à l'ensemble de test.

Enfin, dans notre zone d'étude, Toukourou (2009) utilise les réseaux de neurones formels pour la prévision de crues éclair sur le bassin versant du Gardon d'Anduze. Les données de débit observées jusqu'à l'instant de prévision sont introduites en complément des précipitations, en tant que variable d'état, dans un modèle statique alors appelé « dirigé ». Les résultats obtenus ont nécessité l'emploi de méthodes de régularisation (arrêt précoce, validation croisée, sélection de variables) et se sont avérés tout à fait intéressants, même sur les évènements les plus intenses. Il s'agit là de la seule méthodologie à avoir produit, à ce jour, de bons résultats sur des crues éclair.

D'une manière générale, Abrahart et al. (2012) soulignent que la recherche a tendance à stagner dans ce domaine, se contentant d'applications de méthodes connues à d'autres jeux de données ou d'autres zones d'étude. Ils incitent les chercheurs à se concentrer sur de nouvelles méthodes et sur une plus grande intégration de ces modèles dans les méthodes existantes de l'hydrologie.

## **3.5 Conclusions**

Nous avons pu détailler dans ce chapitre les principales propriétés des modèles à réseaux de neurones. Les méthodes de régularisation indispensables à un apprentissage statistique de qualité ont été détaillées et comparées. Nous avons mesuré toute leur importance dans les applications menées par les auteurs.

Ces derniers abordent la question de la prévision des crues par réseaux de neurones de deux manières : soit en adaptant leur modèle aux conditions locales, soit en essayant d'améliorer sa capacité de généralisation.

Dans la zone d'étude, il a été montré que l'amélioration de la capacité de généralisation conduit à de meilleurs résultats, plus robustes. Renforcés dans notre démarche par ces résultats, plus qu'une adaptation rigide à des conditions spécifiques, nous chercherons à améliorer encore cette capacité de généralisation pour diriger, par la suite, plus aisément nos travaux vers les bassins non jaugés. C'est tout l'objet du chapitre suivant qui va s'attacher à mettre en œuvre les prédicteurs neuronaux sur un bassin cévenol rapide.



## 4. Prévision des crues du Gardon à Mialet

Ce chapitre propose de mettre en application la prévision des crues par réseaux de neurones sur le bassin versant du Gardon de Mialet à Mialet. La prévision sera mise en œuvre pour deux contextes de mesure hydrométéorologique : celui d'un bassin bien jaugé et celui d'un bassin mal jaugé.

Dans un premier temps, le bassin sera rapidement présenté. Ses principales caractéristiques seront énoncées, notamment sa topographie accidentée et sa géologie typiquement cévenole.

Nous explorerons ensuite la base de données fournie dans le cadre de l'étude. Ce sera l'occasion de définir les conditions d'extraction et de traitement des informations hydrométéorologiques, puis de décrire les événements, de façon générale, par le biais des précipitations, du débit et de leur relation.

Comme introduit dans les précédents chapitres, nous allons nous intéresser ici à réaliser la prévision des crues avec un modèle neuronal et sans prévision des pluies futures afin de correspondre au mieux à la demande du projet BVNE de se situer dans un contexte opérationnel. Le mode de sélection de l'architecture du modèle neuronal, tel que décrit au chapitre 3, sera décliné sur ce bassin versant particulier et les processus de son optimisation seront décrits. Les architectures optimales pour chaque contexte de mesure seront alors présentées.

La mise en œuvre du modèle sur plusieurs événements et à plusieurs horizons de prévision sera effectuée, en séparant notamment le cas d'un bassin bien jaugé, où nous verrons que les modèles statiques alimentés par les débits observés passés sont les plus efficaces, et le cas d'un bassin où les mesures de débit peuvent être manquantes en temps réel. Dans ce deuxième cas, un modèle récurrent prenant en compte les débits estimés par le modèle sera mis en œuvre.

Compte tenu de l'importance des données météorologiques sur l'estimation du débit et de la grande incertitude associée à ces données, ce chapitre présente également une analyse de la sensibilité de l'estimation du débit à des variations aléatoires d'un bruit ajouté aux valeurs de précipitation. Un « ensemble » d'événements météorologiques est donc constitué, pour lequel un ensemble de débits peut être calculés, et leurs performances sont ici analysées. Une évaluation de l'apport des entrées de l'humidité issue du modèle SIM (Habets et al., 2008) sera aussi effectuée.

Enfin, l'évaluation des capacités de prévision des modèles récurrents et non récurrents sera effectuée d'une part en regard de critères opérant globalement sur l'hydrogramme et sur le pic et d'autre part en considérant le niveau de vigilance que les hydrogrammes prévu permettent de définir.

La plupart des résultats de ce chapitre ont fait l'objet d'une publication dans la revue *Natural Hazards and Earth System Sciences* (Artigue et al., 2012) (annexe 2) ainsi que d'une présentation et d'une publication dans les actes du *Third World Congress On Nature And Biologically Inspired Computing* (à Salamanque, Espagne, les 19-21 octobre 2011) (Artigue et al., 2011).

#### 4.1 Présentation du bassin versant du Gardon à Mialet

Le Gardon de Mialet est un cours d'eau cévenol, tributaire du Gardon d'Anduze, traversant successivement les départements de la Lozère et du Gard (figure 27)

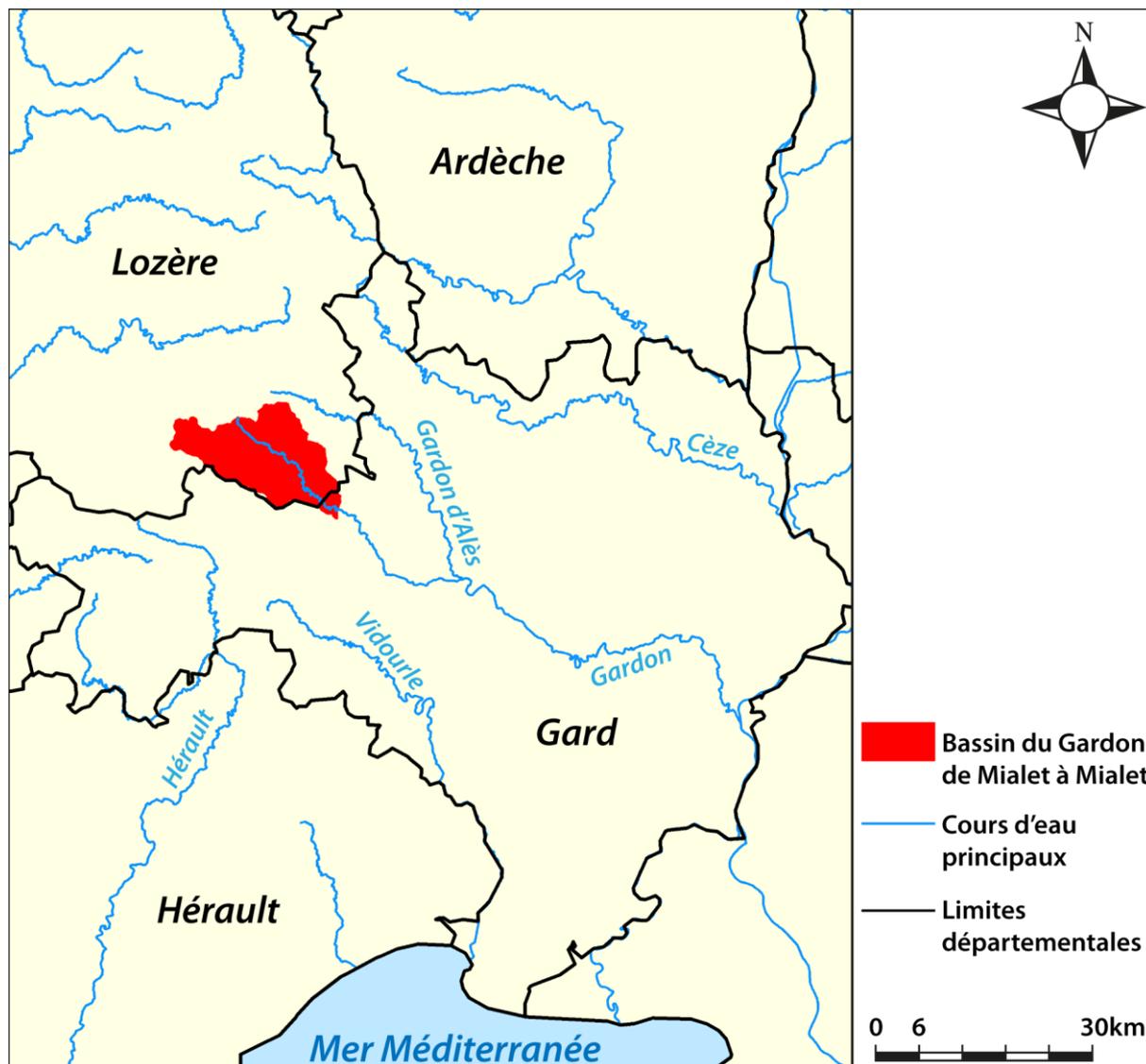


Figure 27: situation du bassin versant du Gardon de Mialet dans la zone d'étude. Cours d'eau d'après IGN BD Carthage ; départements d'après IGN BD Geofla.

Si l'on considère la station hydrométrique de Mialet comme exutoire, le Gardon de Mialet est court (11km) et son bassin est de taille moyenne (220km<sup>2</sup>). Il résulte en fait de la confluence entre le Gardon de Sainte-Croix en rive droite et le Gardon de Saint-Martin-de-Lansuscle en rive gauche (lui-même notamment grossi du Gardon de Saint-Germain-de-Calberte, en rive gauche). Ses altitudes s'échelonnent de 169m à l'exutoire à 1162m au Mont Mars, point culminant du bassin.

L'écoulement général se fait du nord-ouest vers le sud-est dans des vallées d'abord rectilignes en amont, où les pentes sont très fortes, plus sinueuses en aval, où les pentes diminuent. Ces pentes sont, pour les versants, très fortes avec 33% en moyenne. Le drain le plus long (Gardon de Sainte-Croix puis Gardon de Mialet) mesure 29km de longueur et a une pente moyenne de 3%.

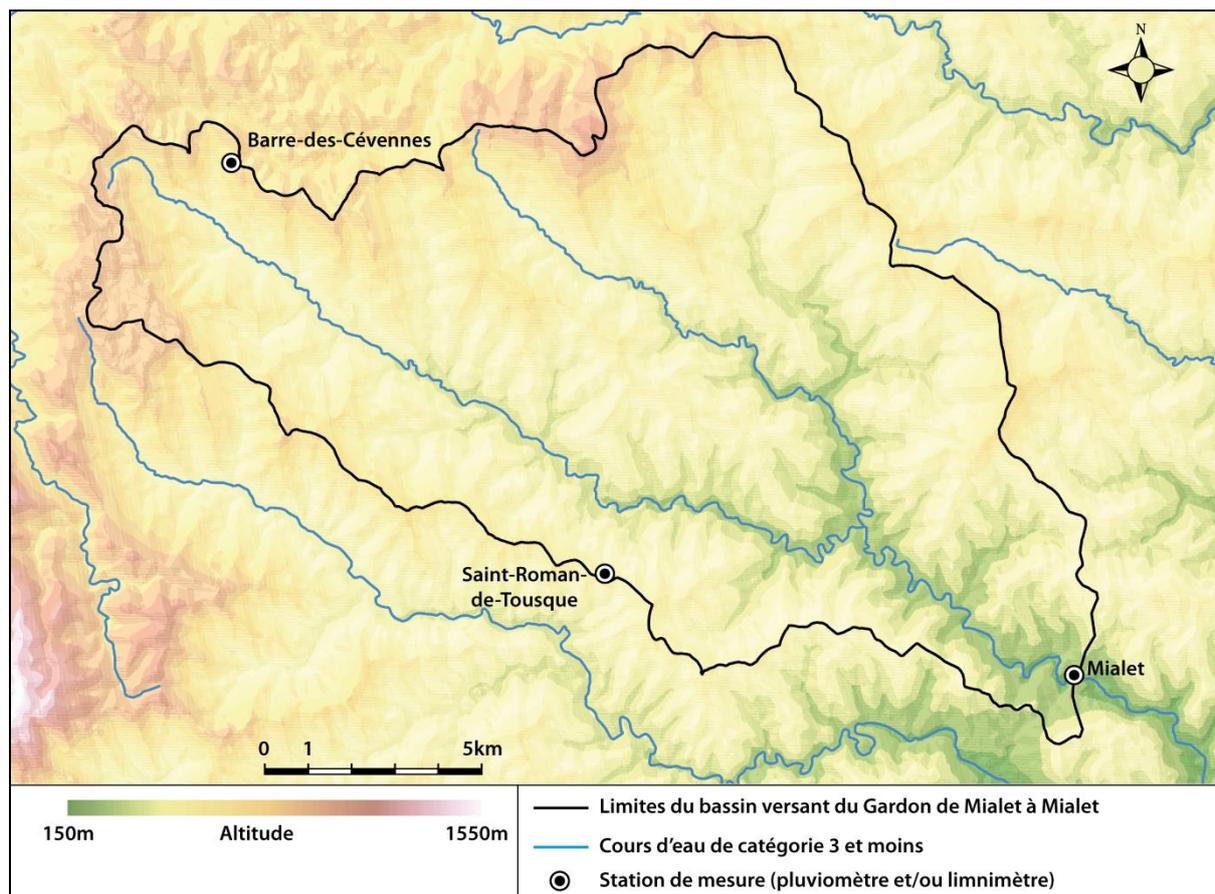


Figure 28: carte topographique et hydrographique simplifiée du bassin versant du Gardon de Mialet à Mialet. Altitudes d'après ASTER GDEM, un produit de METII et NASA ; Hydrographie d'après IGN BD CARTHAGE.

Trois sites remarquables au sens des travaux qui nous intéressent sont localisés dans la figure 28 :

- Mialet (30), 169m d'altitude, à l'exutoire du bassin, est équipé d'une station limnimétrique et d'un poste pluviométrique.
- Saint-Roman-de-Tousque (commune de Moissac-Vallée-Française, 48), 630m d'altitude, dans le centre-ouest du bassin, est équipé d'un poste pluviométrique (photo 7).
- Barre-des-Cévennes (48), 930m d'altitude, au nord-ouest du bassin, est équipé d'un poste pluviométrique.



Photo 7: pluviomètre de Saint-Roman-de-Tousque ; photo : Artigue, 2011.

D'un point de vue géologique, le bassin est assez homogène. Il comporte en effet plus de 95% de séries métamorphiques appelées « *schistes des Cévennes* », composées de séricitoschistes, micaschistes et, parfois, de gneiss. Seul l'extrême aval et l'extrême nord-ouest sont respectivement occupés par des granites et par les séries carbonatées karstiques secondaires perchées de la *can* (cause de petite taille) de l'Hospitalet.

Le bassin est faillé de failles subverticales décrochantes dextres pour la plupart de direction N180 et N090E. Seule une faille presque horizontale, soulignée par une mince couche de gneiss, marque un ancien chevauchement datant de l'orogénèse hercynienne dans le nord-est du bassin ainsi que la base de la série schisteuse du mont Mars, non distinguée des autres séries schisteuses ici.

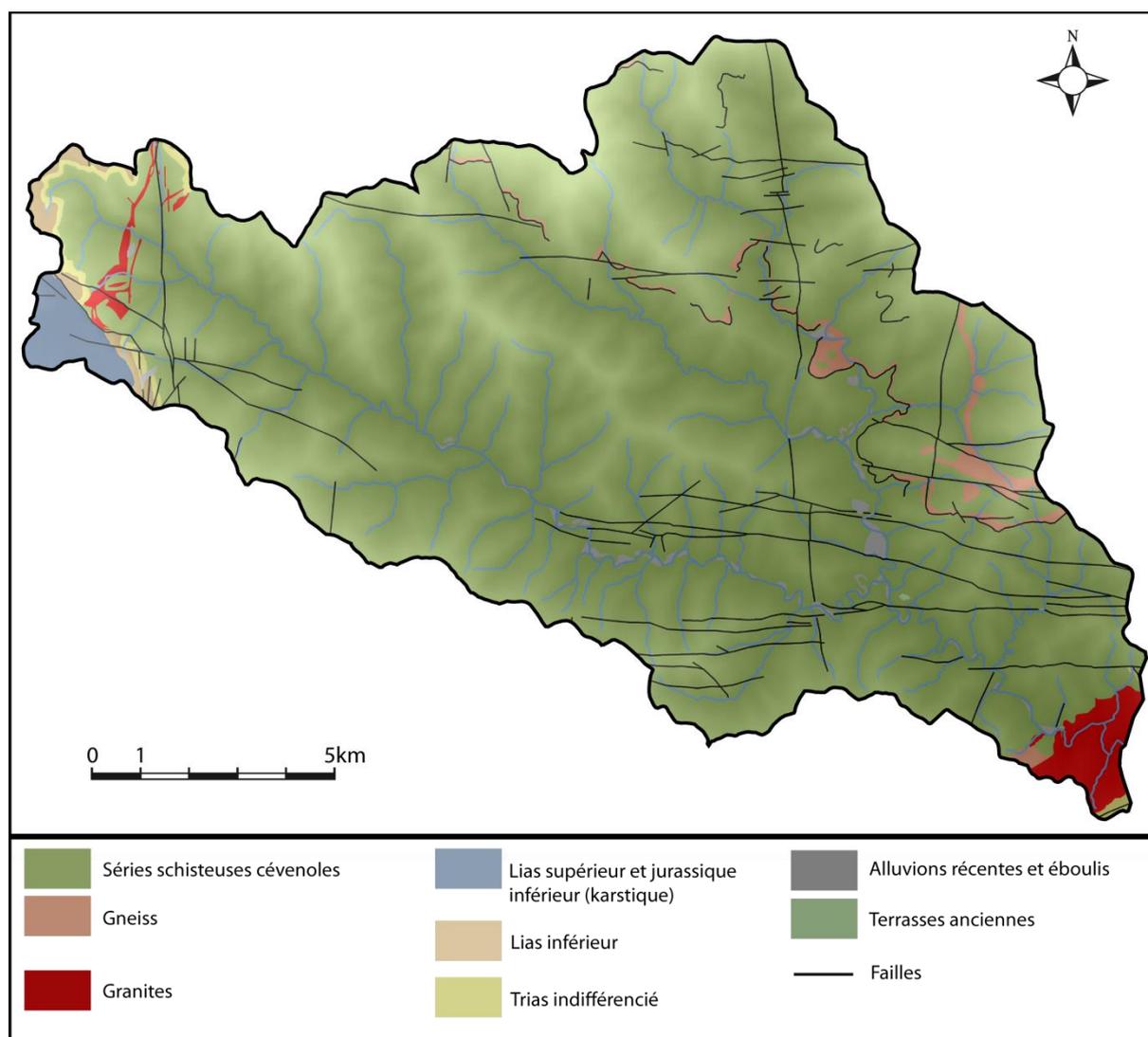


Figure 29: carte géologique simplifiée du bassin versant du Gardon de Mialet à Mialet. Géologie d'après carte géologique de la France au 1/50 000, BRGM ; Cours d'eau d'après IGN BD Carthage ; effet de relief d'après ASTER GDEM, un produit de METII et NASA.

D'un point de vue hydrogéologique, aucun aquifère important n'est présent sur le bassin. Selon la notice de la carte géologique au 1/50 000 du BRGM (feuille 911), les parties schisteuse et granitique se distinguent par des ressources souterraines médiocres du fait de leur imperméabilité. Les sources sont nombreuses mais ne fournissent que des débits très modestes. Seule une pluviométrie importante en toute saison (mais irrégulière) sur l'amont assure un

soutien au débit des cours d'eau. Les rares plaines alluviales offrent les réserves les plus importantes mais sont assujetties au débit, pour le moins variable, du cours d'eau auquel elles sont associées. Concernant la *can* de l'Hospitalet, la présence supposée d'une petite nappe au cœur du léger synforme du plateau permet sans doute de réguler le débit des sources par rapport aux volumes précipités. Ce karst n'étant que partiellement situé sur le bassin, on pourra d'autant plus considérer son impact sur l'hydrologie de surface comme très faible. Enfin, le karst de Mialet se situe en aval de l'exutoire et est donc sans impact.

D'une manière générale, le bassin se caractérise au final par des capacités de stockages faibles sur le long terme. Cette remarque ne veut pas dire que le bassin n'est pas capable de retarder les écoulements, en stockant temporairement de l'eau sur le bassin, mais que les réserves souterraines en eau, permanentes ou presque, sont faibles.

Les sols sont rocailleux, généralement fins, bien que l'épaisseur des horizons soit difficile à déterminer en raison de l'altération avancée des schistes sous-jacents.

Enfin, l'occupation des sols est assez homogène. Selon Ayrat (2005), 92% du bassin est recouvert d'espaces naturels (essentiellement des châtaigniers, des conifères, des forêts mixtes et des garrigues, boisées ou non) alors que les 8% restants se partagent surtout entre les sols nus et les zones artificielles.

Ce bassin est donc typiquement cévenol : il présente de fortes pentes, est peu urbanisé et son sous-sol est essentiellement cristallin. Sa situation au cœur de la zone cévenole, à mi-chemin entre les bassins partagés avec la plaine et les bassins de haute altitude des massifs granitiques, lui confère un climat très représentatif du secteur. Par ailleurs, comme nous allons le voir dans la section suivante, la disponibilité d'une importante chronique de données hydrométéorologiques entre 1992 et 2008 va nous permettre d'extraire une grande quantité d'évènements météorologiques. Le choix du bassin versant du Gardon de Mialet à Mialet pour les premiers essais de prévision des crues se justifie donc aisément par l'ensemble des points ici résumés.

## **4.2 Données disponibles**

Cette partie présente les données extraites de celles fournies dans le cadre du projet BVNE. Nous décrirons les méthodes d'extraction (sélection, seuils, interpolation...) puis détaillerons les caractéristiques de la base de données en trois parties : pluie, débit et relation pluie-débit.

Concernant l'ensemble des évènements présentés dans les chapitres 4 et 5, le mode de numérotation, les dates et les provenances sont indiquées en annexe 3.

### **4.2.1 Extraction et traitement**

#### *4.2.1.1 Principes de base et définition d'un seuil de sélection des évènements*

L'extraction et le traitement des données ont été pensés pour répondre à la problématique de prévision de crues rapides par apprentissage statistique. Cette problématique impose deux principes :

- La modélisation doit être événementielle, comme nous l'avons vu dans la section 2.2.1.2.
- Le nombre d'évènements doit être suffisamment important pour alimenter une base d'apprentissage représentative des différents comportements envisageables sur le bassin.

Afin de disposer d'une base de données comportant une large variété d'épisodes (particulièrement les épisodes peu réactifs à cause des sols secs), nous avons sélectionné les

événements sur la pluie et non sur le débit ou la hauteur d'eau, afin notamment d'éviter le déclenchement d'alertes intempestives. Cela permet en effet de traiter les événements pour lesquels les précipitations, bien que relativement importantes, ne font pas réagir significativement le cours d'eau (sols secs).

Le seuil de sélection choisi est de 100mm en 48 heures glissantes sur l'un au moins des trois pluviomètres. Ainsi, l'un au moins des trois pluviomètres présentés dans la section 4.1 doit être touché par un tel cumul.

Ce seuil peut paraître assez bas pour une zone qui subit de très importants cumuls avec une fréquence exceptionnelle en France métropolitaine. Il appelle deux remarques :

- Ce seuil est généralement celui à partir duquel l'on voit se produire une réaction, même faible, du système hydrologique, au moins lorsque le bassin est très humide.
- Pour diminuer le taux de fausses alertes, un modèle basé sur l'apprentissage statistique doit aussi disposer d'événements assez faibles pour « apprendre » les cas où la réaction hydrologique est atténuée par les conditions sur le bassin.

#### 4.2.1.2 Base de données fournie dans le cadre de BVNE Gardons

Pour le projet BVNE, le SCHAPI a fourni 17 années de données presque continues sur le bassin versant du Gardon de Mialet à Mialet, ses trois pluviomètres et son limnimètre. Les données sont disponibles de 1992 à 2008 mais durant cette période, la période d'échantillonnage évolue : de une heure avant 2002 à cinq minutes après. Les présents travaux s'effectuent à un pas de temps de 30 minutes d'une part par contrainte du projet BVNE et d'autre part, parce que cette période est estimée comme correcte pour représenter les signaux impliqués (Toukourou, 2009).

Il a donc fallu ré-échantillonner l'ensemble des données de la manière suivante :

- Pour la pluie avant 2002, il a fallu ajouter un pas de temps aux données brutes, tout en conservant, bien sûr, le cumul total. La méthode suivante a été imaginée :  
On dispose de deux valeurs de cumul horaire  $P_k$  et  $P_{k+2}$ . Nous cherchons à obtenir  $P_{k+1}$  et  $P'_{k+2}$  le nouveau cumul correspondant en  $k+2$ . Avec des données non cumulatives, nous aurions simplement réalisé une interpolation linéaire et fait la moyenne entre  $P_k$  et  $P_{k+2}$  pour obtenir  $P_{k+1}$ . Cette méthode aurait évidemment faussé le cumul dans le sens d'une augmentation notable.

$$P_{k+1} \neq P_i = (P_k + P_{k+2}) / 2$$

Où  $P_i$  est la pluie interpolée avec la méthode où les cumuls totaux ne sont pas conservés. Afin d'éviter ce problème, nous introduisons un facteur  $C$  qui tient compte du cumul observé initialement en  $k+2$  sur une heure. La forme du signal est conservée mais le signal est multiplié par le coefficient  $C$ .

$$P_{k+2} = (P_{k+2} + P_i) C$$

On a donc :

$$C = P_{k+2} / (P_{k+2} + P_i)$$

D'où l'on peut déduire :

$$P_{k+1}=P_iC$$

Et :

$$P'_{k+2}=P_{k+2}C$$

Où  $P'_{k+2}$  est le nouveau cumul en  $k+2$ .

- Pour la pluie à partir de 2002 jusqu'en 2008, il a suffi de faire une somme des cumuls en cinq minutes toutes les demi-heures pour obtenir un cumul semi-horaire.
- Pour les hauteurs d'eau avant 2002, une moyenne entre les deux hauteurs horaires successives a donné la hauteur intermédiaire qui manquait. Nous sommes bien conscients que l'échantillonnage initial à une heure peut être insuffisamment fréquent selon les événements considérés mais la méthode adoptée ne dégrade pas l'information par rapport à ce qui existait en pas de temps horaire.
- Pour les hauteurs d'eau à partir de 2002 et jusqu'en 2008, nous avons choisi de ne pas effectuer la moyenne des hauteurs de la demi-heure précédente mais de conserver la hauteur maximale de la précédente demi-heure, ceci afin de s'assurer de conserver le maximum observé. Cette méthode peut induire un décalage temporel maximal de 25 minutes mais par opposition, l'utilisation de la moyenne de la dernière demi-heure n'aurait pas apporté une information plus précise temporellement et aurait souvent, même légèrement, sous-estimé le pic.

Les travaux de Toukourou (2009) montrent qu'il est possible de travailler avec les hauteurs d'eau. Cependant, comme la finalité de ces travaux est la prévision sur des bassins peu ou pas jaugés, l'information volumique apportée par le débit pourra s'avérer nécessaire. De plus, les lois d'estimation des débits de pointe en fonction de différentes caractéristiques du bassin ne tiennent pas compte des hauteurs. Ainsi, même si elles apportent nécessairement un biais, nous avons utilisé les courbes de tarage fournies par le SCHAPI dans le cadre du projet BVNE pour obtenir les débits à la station limnimétrique de Mialet.

Les équations de conversion utilisées sont les suivantes (où  $H$  est la hauteur d'eau et  $Q$  le débit) :

- Au-dessous de 3,83m d'eau, on utilisera :

$$Q = 10,024H^2 + 3,9707H + 2,17348$$

- Au-dessus de 3,83m d'eau, on utilisera :

$$Q = 12,154H^2 + 5,470H - 34,59$$

Ces équations correspondent aux courbes de tarage présentées dans la figure 30.

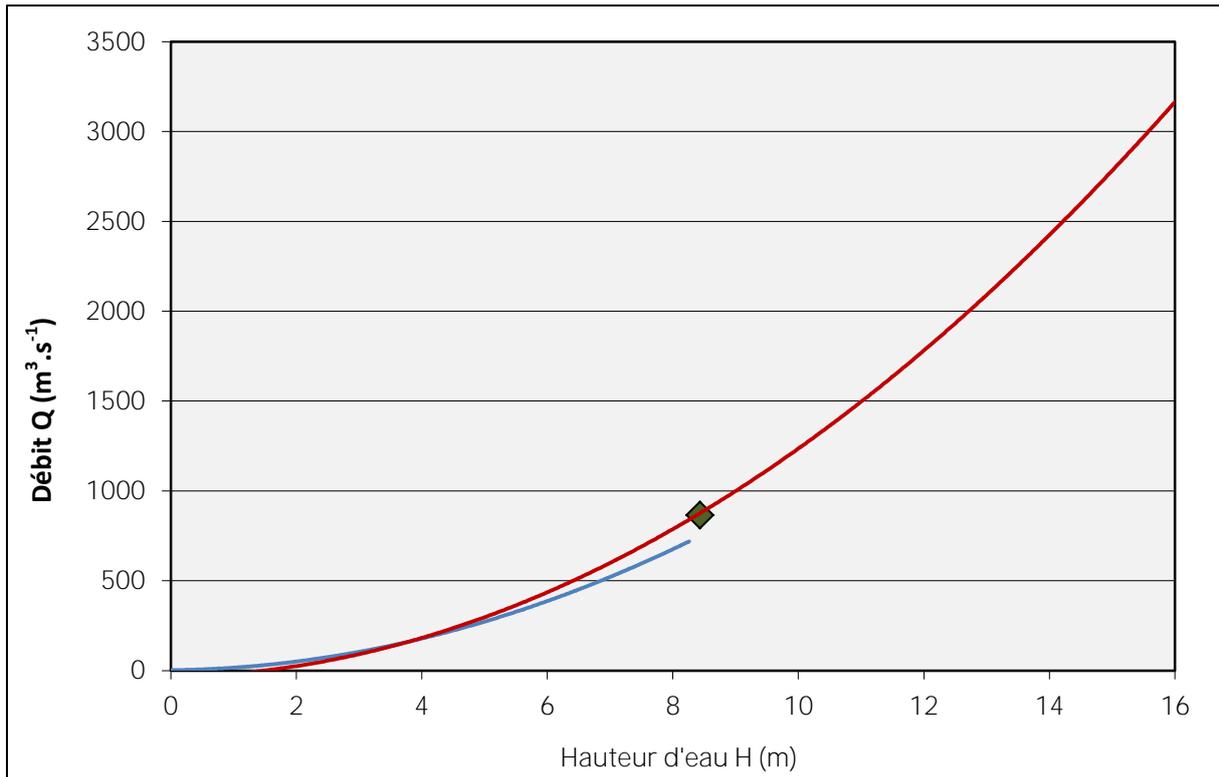


Figure 30: courbes de tarage de la station limnimétrique de Mialet pour les faibles hauteurs (bleu) et pour les hauteurs élevées (rouge). Le marqueur vert représente l'estimation du débit de la crue de septembre 2002 effectuée par SIEE en 2003.

Enfin, il convient de noter que le découpage des évènements se fait de façon très rigoureuse pour le début de l'évènement : la prise en compte se fait dès le premier basculement d'auget d'un pluviomètre (0,2mm). Pour la fin de l'évènement, la prise en compte se fait à la fin de la réaction hydrologique significative, c'est-à-dire lorsque l'on observe un franc aplatissement de la courbe de récession. Il n'est pas nécessaire de définir une règle précise puisqu'il s'agit simplement de la récession à l'arrêt des pluies : ce n'est ni la période la plus difficile à modéliser, ni la plus importante pour la prévision des crues.

#### 4.2.1.3 Sélection des évènements

En utilisant ces méthodes, ce sont environ 80 évènements qui ont pu être extraits de la base de données brute. Après critique, ont été éliminés ceux dont les données présentaient des lacunes ou des incohérences manifestes.

Il en est resté 58 évènements que nous allons présenter dans la partie suivante et qui sont répartis entre 1992 et 2008 tel qu'indiqué dans la figure 31.

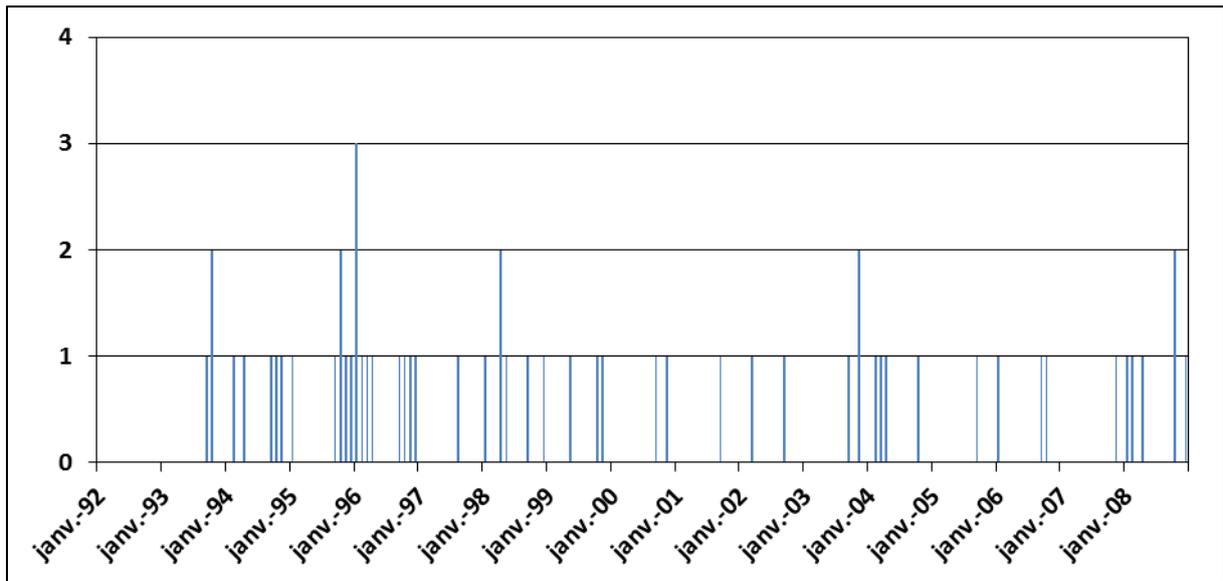


Figure 31 : répartition des événements conservés entre 1992 et 2008.

#### 4.2.2 Description des événements

##### 4.2.2.1 Description générale

Tout d'abord, nous nous sommes intéressés à la répartition des événements au cours de l'année. La durée de référence climatologique pour des données continues est de 30 ans. Pour des événements intenses, elle doit forcément être plus longue. En conséquence, une période de 17 ans n'est évidemment pas suffisante. Ainsi, même si les conclusions que nous en tirons ne peuvent être que partielles, elles sont néanmoins intéressantes à considérer.

La répartition mensuelle des événements sélectionnés dans l'année a été étudiée ; elle se caractérise par un pic automnal principal et deux pics secondaires au printemps et en hiver (figure 32).

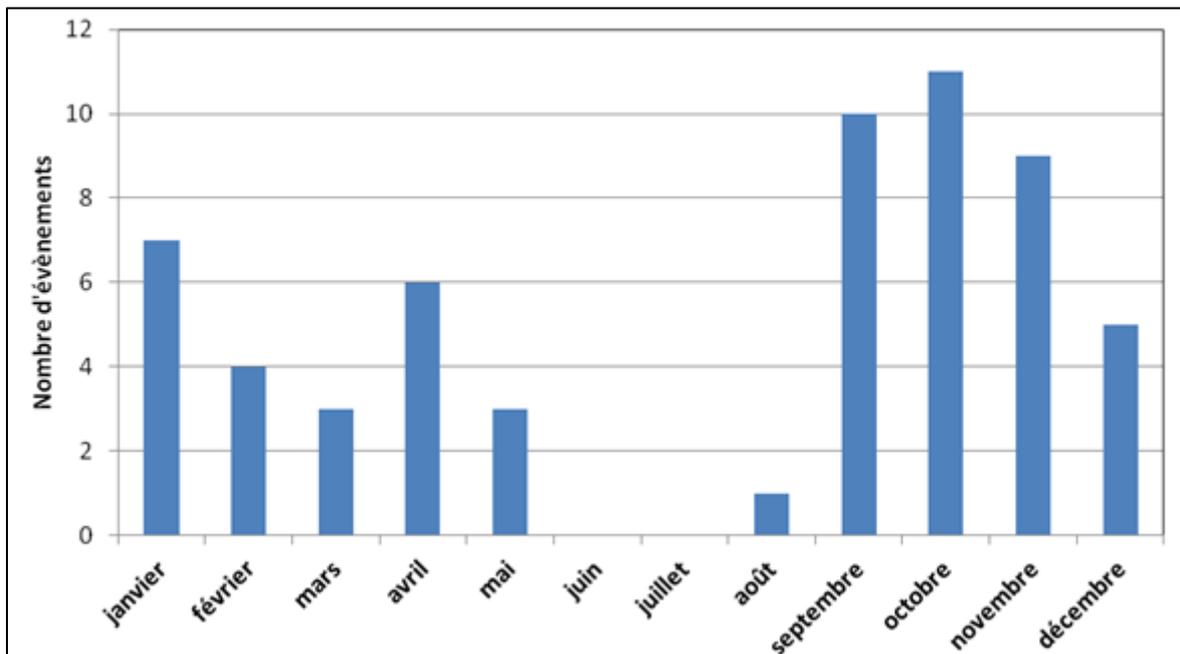


Figure 32: répartition mensuelle des événements de la base de données.

Les mois de septembre, octobre et novembre représentent à eux seuls 52% des évènements. Quant aux deux pics secondaires, étant donné qu'ils sont moins marqués, ils sont peu robustes d'un point de vue statistique.

On constate aussi que les trois mois d'été sont dénués d'épisodes sur cette période, à l'exception d'un épisode qui est survenu à la toute fin du mois d'août. Là encore, la courte période ne doit pas masquer le fait que des évènements intenses peuvent survenir en été comme cela a été le cas dans le Var en juin 2010, à Nîmes le 29 août 1399 ou le 9 août 1873 (metamiga.free.fr) ou encore le 7 août 1964 où l'on relèvera près de 200mm à Remoulins (Jacq, 1994). Cet état de fait est particulièrement préjudiciable en termes de gestion du risque inondations : en été, la population augmente fortement auprès des cours d'eau (fréquentation maximale des terrains de camping notamment).

La figure 33 présente la répartition des durées des épisodes.

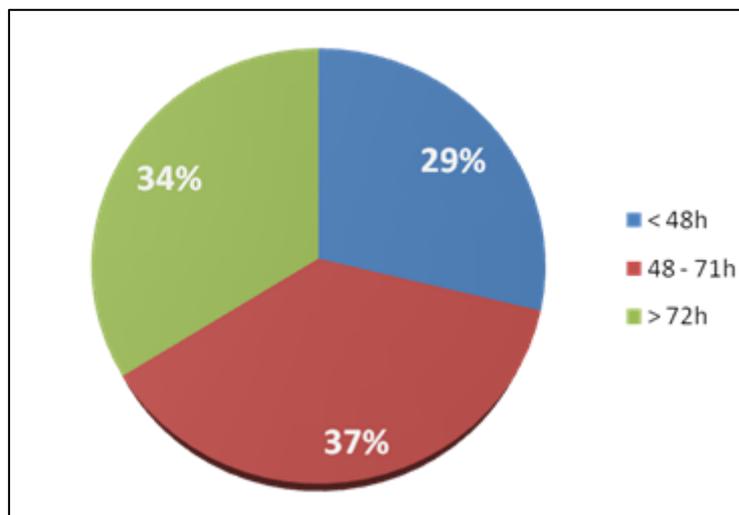


Figure 33: durées des épisodes par classes.

La répartition est à peu près égale entre les trois classes choisies. On observe toutefois une légère prédominance des évènements d'une durée de deux à trois jours et une légère lacune d'évènements de moins de 48h. La durée moyenne des évènements de la base est de 66h, la durée médiane de 60h et l'écart type, important, est de 29h (44% de la moyenne). L'évènement le plus long dure 143h, le plus court 26h.

Les évènements sont donc généralement courts, ce qui est conforme à un caractère méditerranéen avec des épisodes pluvieux brefs et intenses et des réactions hydrologiques rapides.

#### 4.2.2.2 Précipitations : cumuls et intensités

Le test des doubles cumuls permet de mettre en valeur une dérive de fonctionnement des pluviomètres en traçant les cumuls de précipitations cumulatifs d'un pluviomètre par rapport à un autre. Généralement appliqué sur des précipitations annuelles ou mensuelles, on peut le pratiquer sur des cumuls d'épisodes. C'est ce qui a été fait ici (figure 34) et le test ne montre aucune rupture notable de linéarité sur l'ensemble de la période de 17 ans au cours de laquelle sont répartis les épisodes étudiés. On peut donc penser qu'il n'y a pas de dérive majeure de la mesure pluviométrique sur les stations concernées.

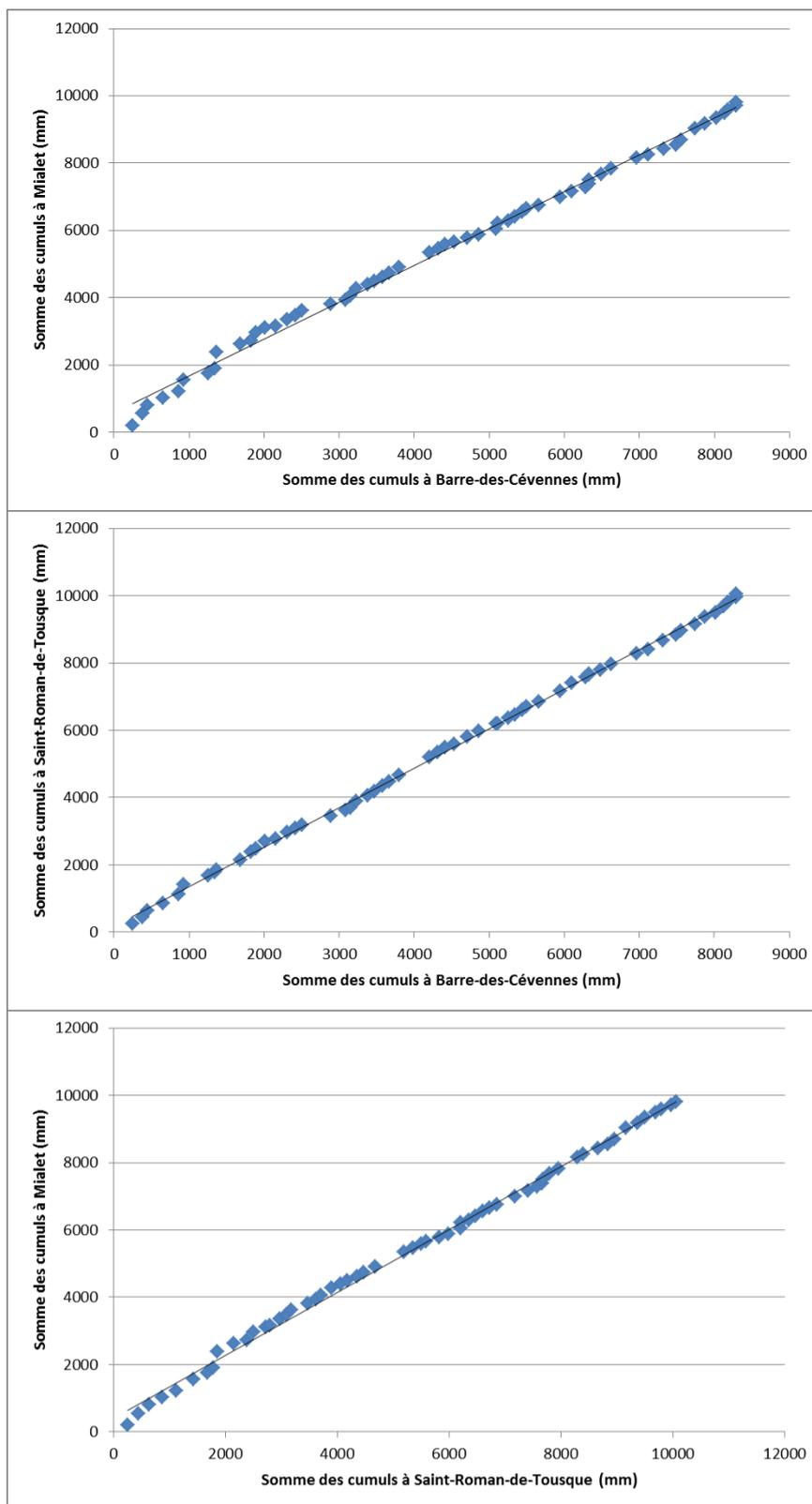


Figure 34: doubles cumuls appliqués à chaque couple de pluviomètres du bassin. La corrélation linéaire entre la droite de régression et les points est à chaque fois supérieure à 0.99.

La figure 35 présente trois histogrammes des cumuls observés sur les trois pluviomètres du bassin par classes de 100mm.

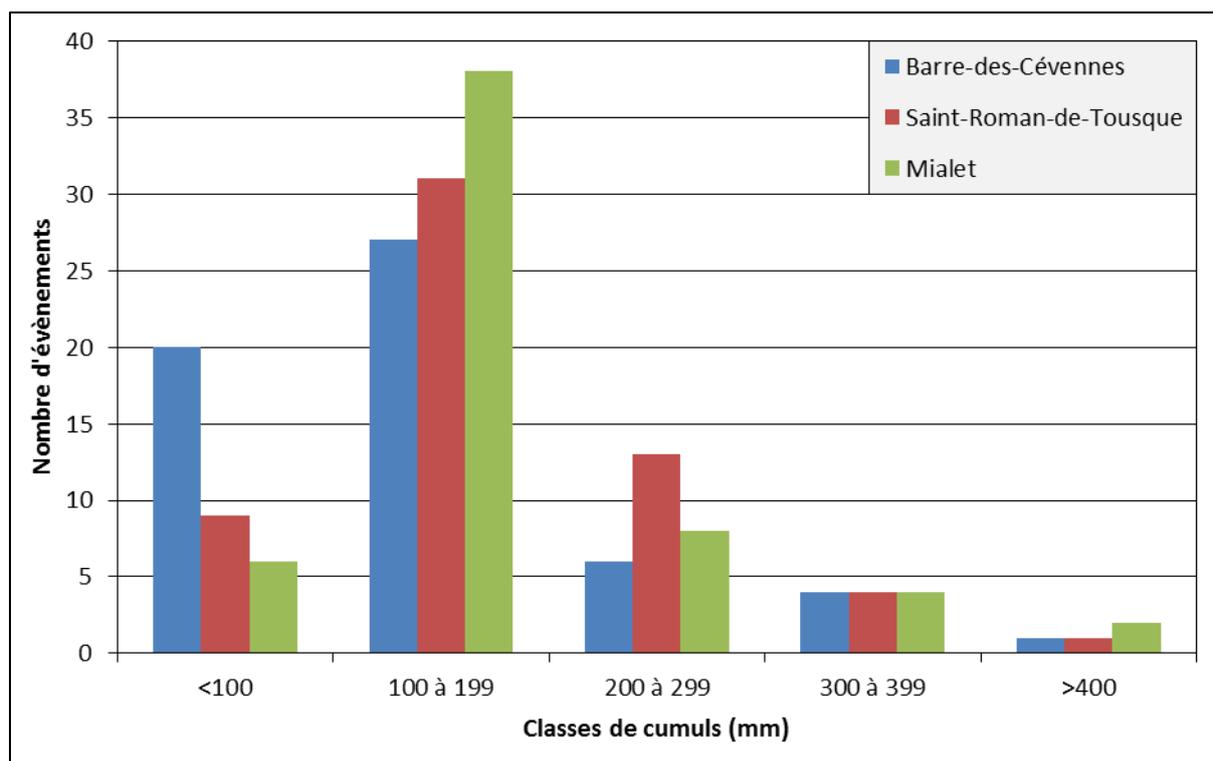


Figure 35: histogramme de la répartition des cumuls par classes de 100mm pour les trois pluviomètres du bassin.

Il ressort clairement que la majorité des cumuls se concentre entre 100 et 200mm. Cependant, quelques différences entre les pluviomètres sont à noter et sont soulignées ici par le tableau 2 :

Tableau 2: quelques statistiques sur les cumuls enregistrés sur les différents pluviomètres, pour les évènements sélectionnés

Cumul	Barre des Cévennes	Saint Roman de Tousque	Mialet
<b>Maximum</b>	410mm	523mm	485mm
<b>Minimum</b>	0mm	1mm	40mm
<b>Moyenne</b>	143mm	173mm	169mm
<b>Médiane</b>	134mm	165mm	141mm
<b>Ecart Type</b>	89mm	86mm	88mm

D'abord, Barre-des-Cévennes enregistre généralement des cumuls plus faibles que ceux des deux autres pluviomètres. Cette observation est cohérente avec l'analyse climatique du chapitre 1, qui place cette station plus près du climat des crêtes, avec une tendance méditerranéenne moins affirmée.

Saint-Roman-de-Tousque enregistre, de peu, les plus forts cumuls avec 4mm en moyenne et 24mm en médiane de plus que Mialet. Dans tous les cas, l'écart-type est fort, plus de 50% de la moyenne et de la médiane, ce qui ne confère pas à ces dernières une robustesse très élevée. Dans ce cas, sur seulement 17 ans et 58 épisodes, on conclura que Barre-des-Cévennes est généralement moins arrosé au cours des épisodes pluvieux intenses.

La corrélation linéaire des mesures relevées aux différents pluviomètres a été étudiée. Elle confirme le caractère localisé des précipitations convectives qui frappent les Cévennes. Les corrélations des pluies semi-horaires observées sur tous les épisodes entre les trois couples de pluviomètres sont toutes très mauvaises (figure 36).

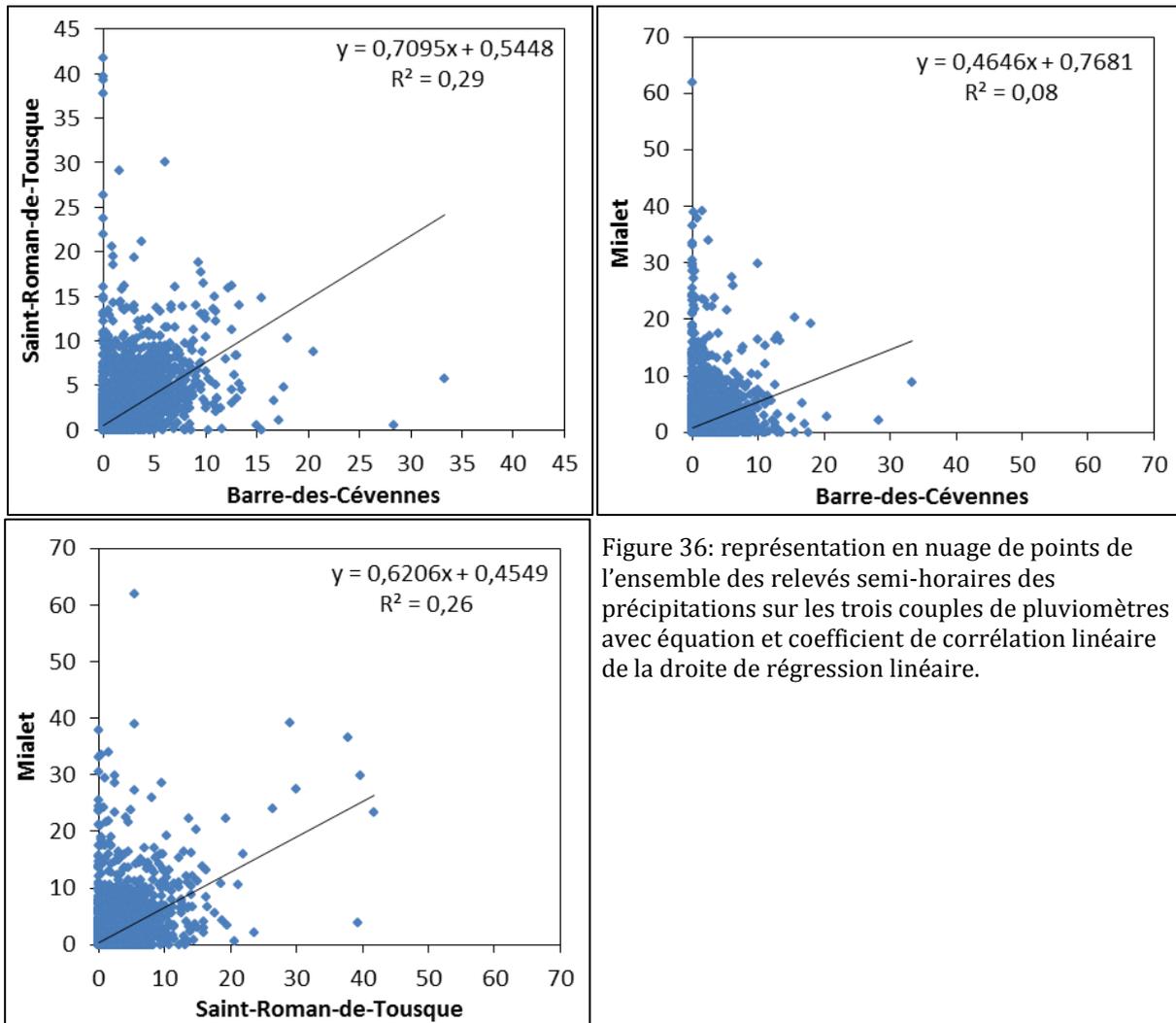


Figure 36: représentation en nuage de points de l'ensemble des relevés semi-horaires des précipitations sur les trois couples de pluviomètres avec équation et coefficient de corrélation linéaire de la droite de régression linéaire.

En revanche, la même étude a été effectuée sur les cumuls par épisode et bien que la population soit nettement plus faible, il a pu être montré que les cumuls totaux observés à Saint-Roman-De-Tousque n'étaient pas totalement indépendants de ceux de Barre-des-Cévennes (figure 37).

Saint-Roman-de-Tousque est plus proche de Mialet que de Barre-des-Cévennes mais Saint-Roman-de-Tousque est à une altitude (630m en crête) plus proche de celle de Barre-des-Cévennes (930m en crête) que de celle de Mialet (169m en fond de vallée). De là à conclure que l'altitude a plus d'influence que la localisation sur les cumuls observés lors des épisodes sur ce bassin, il y a un pas que (i) la faible profondeur statistique disponible, (ii) la faiblesse du coefficient de corrélation linéaire (0.64) et (iii) la faible représentativité des mesures ponctuelles de pluie ne nous permettent pas de franchir.

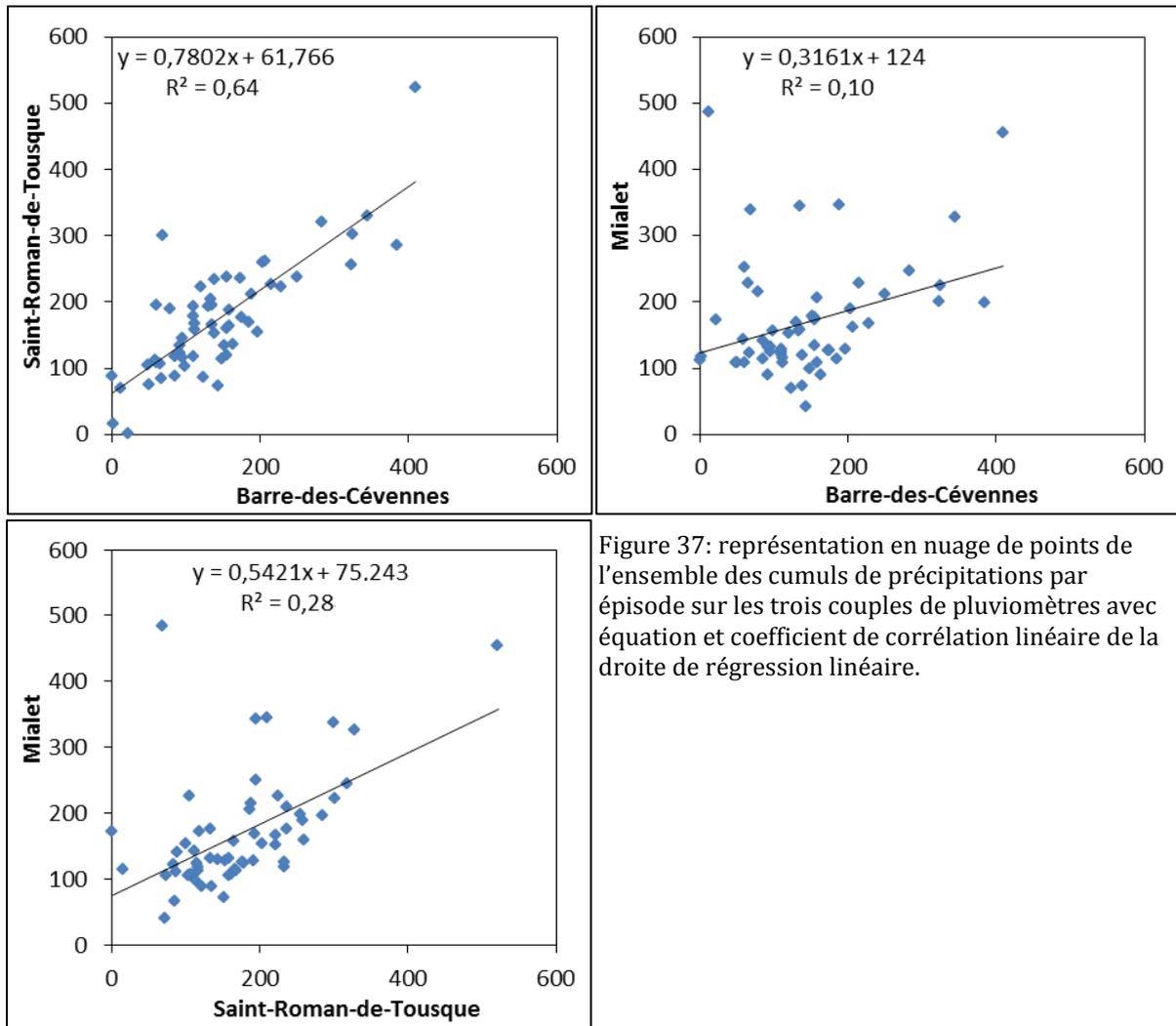


Figure 37: représentation en nuage de points de l'ensemble des cumuls de précipitations par épisode sur les trois couples de pluviomètres avec équation et coefficient de corrélation linéaire de la droite de régression linéaire.

Enfin, la figure 38 présente les intensités moyennes sur une demi-heure et une heure pour chaque station. Les valeurs maximum, moyenne et médiane atteintes par ces variables sur l'ensemble de la base sont détaillées.

Comme pour les cumuls, on constate que Barre-des-Cévennes connaît généralement des intensités plus faibles que les autres stations. C'est, du reste, Mialet qui connaît généralement les intensités les plus élevées, plaçant Saint-Roman-de-Tousque en valeur intermédiaire.

Ces observations sont cohérentes avec les calculs de Ceresetti (2011) qui montrent que la probabilité de fortes intensités pluvieuses augmente lorsque l'on s'éloigne du relief dans la zone d'étude. Pour les maximas, on obtient d'ailleurs des valeurs encore nettement plus élevées que celles de Mialet un peu plus en aval, à Anduze (jusqu'à 137mm/h en moyenne sur une heure et 153mm/h en moyenne sur une demi-heure) ou encore à Cardet, encore plus aval, qui enregistra 146mm en une heure glissante en septembre 2010.

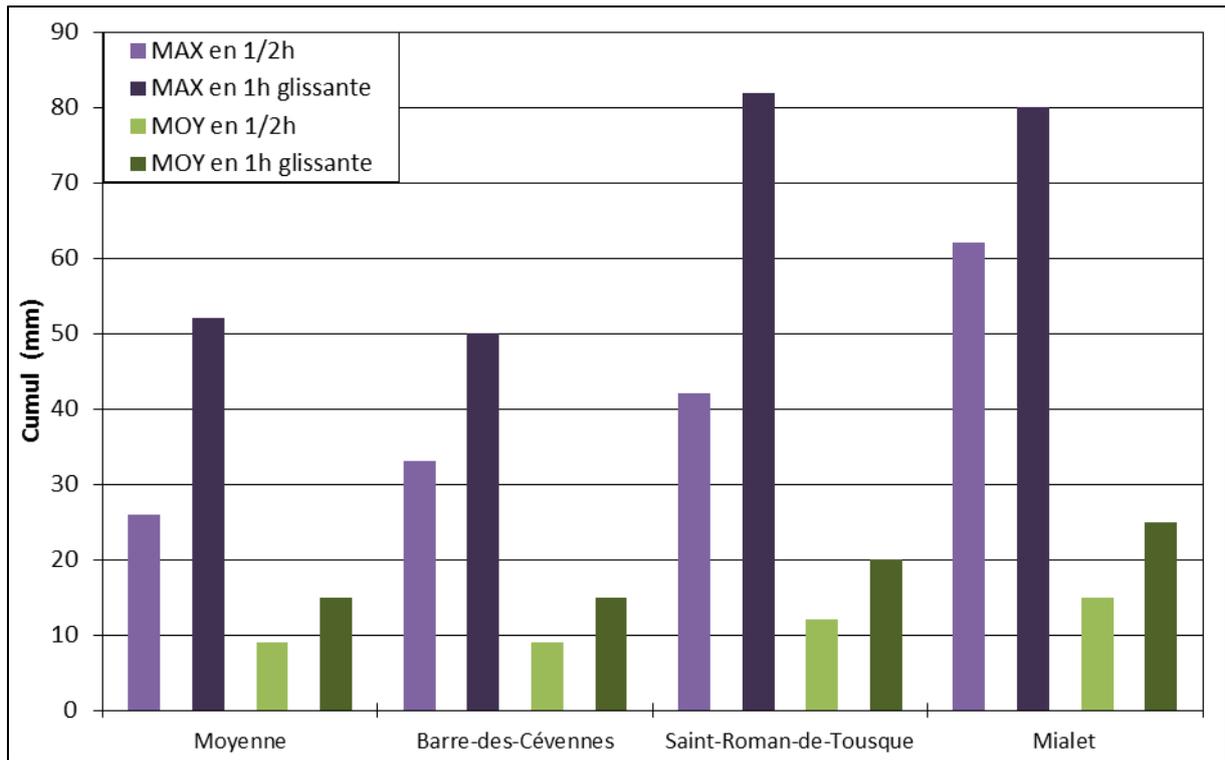


Figure 38: représentation des intensités moyennes sur 1h et 1/2h maximales, moyennes et médianes sur l'ensemble de la base pour chaque poste pluviométrique.

En conclusion de cet aperçu des caractéristiques des précipitations sur la base de données sélectionnée, nous pouvons retenir que :

- Les cumuls peuvent être très importants ponctuellement sur le bassin considéré. Ils peuvent aussi, mais plus rarement, l'être simultanément sur les trois pluviomètres. D'une manière générale, ils sont moins importants sur l'amont du bassin.
- Les pluies sont particulièrement variables dans le temps et dans l'espace. Il est difficile de trouver une corrélation entre les pluviomètres que ce soit au pas de temps élémentaire ou sur un épisode.
- Les intensités peuvent être très fortes sur les deux pluviomètres aval mais dépassent rarement la moitié du maximum des deux autres pluviomètres sur le pluviomètre amont de Barre-des-Cévennes.

#### 4.2.2.3 Débits

Dans ces travaux, nous considérerons comme évènement intense tout évènement présentant un débit spécifique de pointe de plus de  $1 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ . Il s'agit, peu ou prou, à la station de Mialet, d'une crue d'occurrence deux ans, selon le règlement d'information sur les crues.

La figure 39 présente un histogramme des débits spécifiques observés à la station de Mialet ( $220 \text{ km}^2$ ).

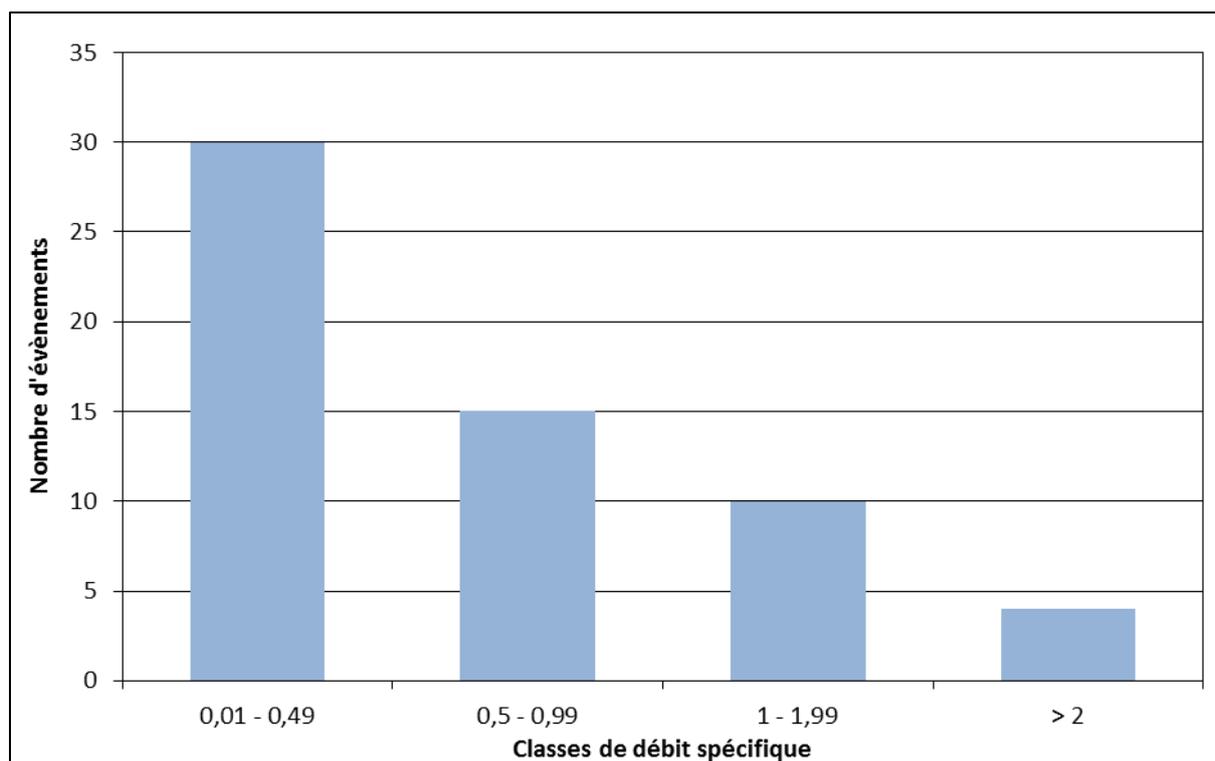


Figure 39: histogramme de la répartition des débits spécifiques de pointe à la station de Mialet.

La base comporte une grande majorité d'évènements de débit de pointe inférieur à  $1 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ . De fait, les épisodes faibles ( $< 0,5 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ) représentent 51% du total des évènements sélectionnés, les épisodes modérés ( $0,5$  à  $0,99 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ) 25%, les épisodes forts ( $1$  à  $1,99 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ) 17% et les épisodes très forts ( $> 2 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ) 7%.

Plus généralement, les débits de pointe s'échelonnent de  $6$  à  $819 \text{ m}^3 \cdot \text{s}^{-1}$ , pour une moyenne de  $161 \text{ m}^3 \cdot \text{s}^{-1}$ , une médiane bien inférieure, à  $109 \text{ m}^3 \cdot \text{s}^{-1}$  et un écart-type presque égal à la moyenne, de  $153 \text{ m}^3 \cdot \text{s}^{-1}$ .

La forme de l'hydrogramme peut, pour les évènements à pic unique, être déterminée par le rapport du débit maximal au débit moyen sur l'évènement. Les évènements présentant une forme de pic typique des crues éclair verront ce rapport s'élever fortement alors que ceux, plus plats, qui ne sont ni très rapides, ni très amples, verront leur rapport rester faible.

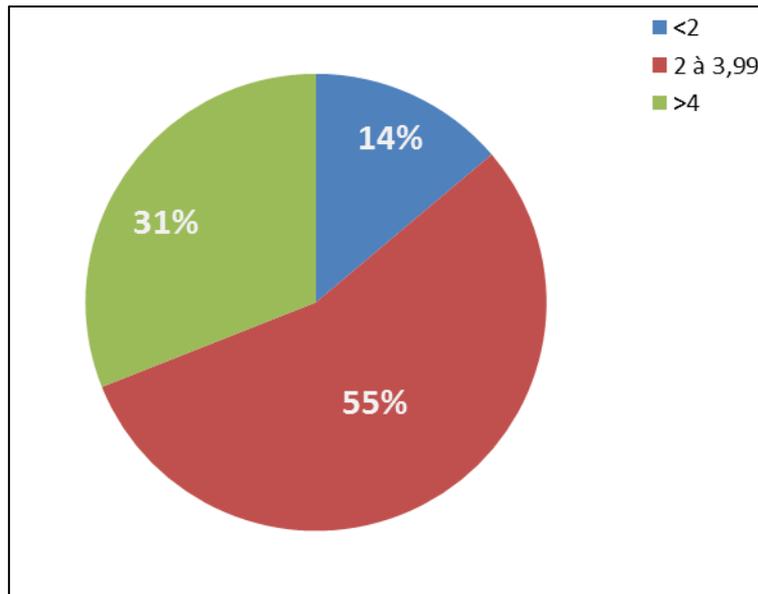


Figure 40: répartition en trois classes des rapports de forme des événements de la base.

Un petit tiers des événements a un rapport de forme élevé, témoignant de pics de crue marqués. Moins de 15% sont relativement plats alors que plus de la moitié se situe dans des proportions moyennes (figure 40). La figure 41 présente les hydrogrammes des deux rapports de forme extrêmes : l'évènement n°230 avec un rapport de 1,2 et l'évènement n°3 avec un rapport de 7,81.

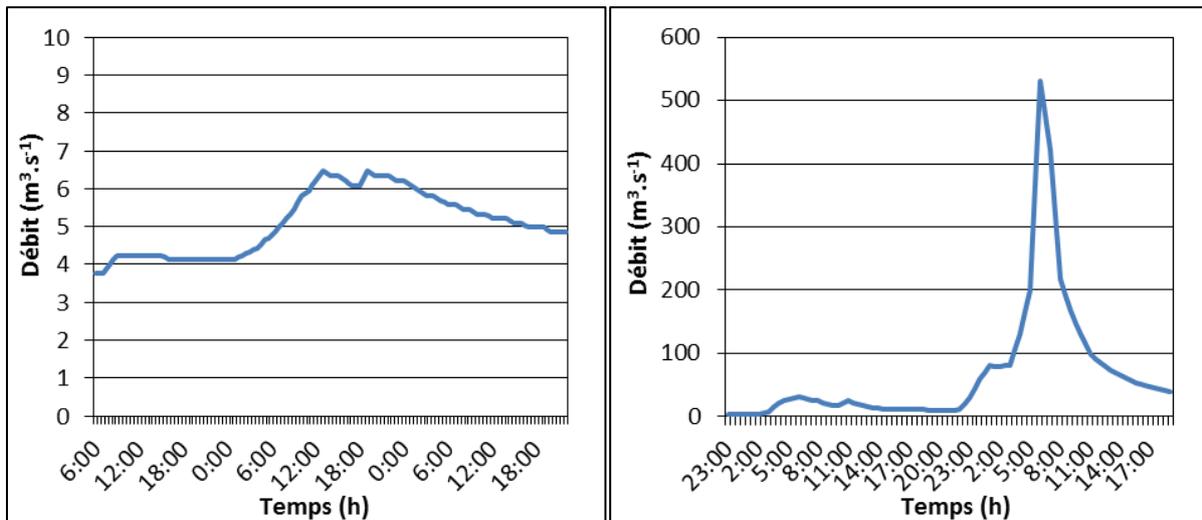


Figure 41: hydrogrammes des deux événements présentant les rapports de forme extrêmes : à gauche, l'évènement n°230 et à droite l'évènement n°3.

Ainsi, une grande variété d'hydrogrammes est disponible. Les formes particulières sont autant de marqueurs de réactions hydrologiques différentes provoquées par des pluies tout aussi variables.

#### 4.2.2.4 Relation entre pluie et débit

Les coefficients de corrélation linéaire qui ont été calculés entre cumuls totaux moyennés sur les trois pluviomètres, ou pour chaque pluviomètre un à un, et le débit de pointe n'ont pas atteint de valeurs significatives (tableau 3).

Tableau 3: coefficients de corrélation linéaire entre les débits de pointe et les pluies en cumuls et en intensités horaires et semi-horaires.

Cumul	Coefficient de corrélation	Intensité maximale 0,5h	Coefficient de corrélation	Intensité maximale 1h	Coefficient de corrélation
Barre-des-Cévennes	0,39	Barre-des-Cévennes	0,34	Barre-des-Cévennes	0,39
Saint-Roman-de-Tousque	0,59	Saint-Roman-de-Tousque	0,54	Saint-Roman-de-Tousque	0,63
Mialet	0,49	Mialet	0,35	Mialet	0,41
Moyenne	0,58	Moyenne	0,50	Moyenne	0,59

Aucun des résultats ici présentés ne peut prétendre donner une indication sérieuse sur une quelconque relation linéaire entre les différents éléments considérés.

De fait, avec cette observation, nous mettons le doigt sur la non-linéarité globale du système hydrologique. Ce dernier répond aux stimuli avec une grande complexité, adaptant plus sa réaction au schéma de succession des pluies durant l'évènement et à l'état initial du bassin qu'à des valeurs d'entrées brutes comme l'intensité maximale ou le cumul. Il peut donc être vu comme un système non linéaire dont la réponse dépend du point de fonctionnement (ou état). Pour illustrer ceci, nous pouvons observer la figure 42 qui présente le coefficient d'écoulement des évènements moyenné par mois de l'année ainsi que la figure 43 qui représente l'évènement n°212.

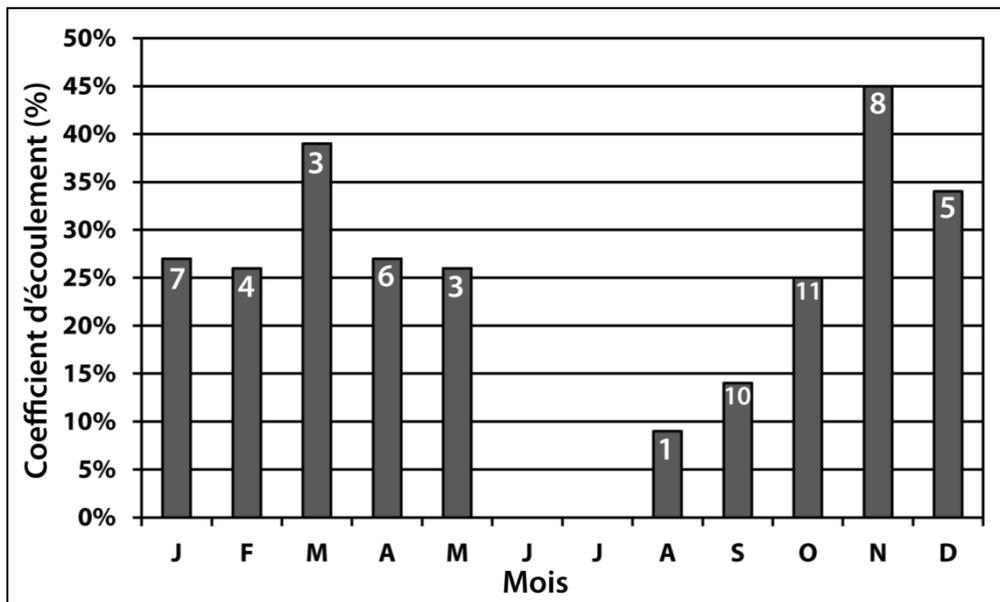


Figure 42: coefficient d'écoulement moyen par mois (la population prise en compte pour la moyenne est indiquée dans les barres).

Comme on peut le voir, le coefficient d'écoulement augmente au fil des mois en fin d'été. Cette évolution est représentative des conditions initiales sèches des bassins après la sécheresse estivale et les températures élevées.

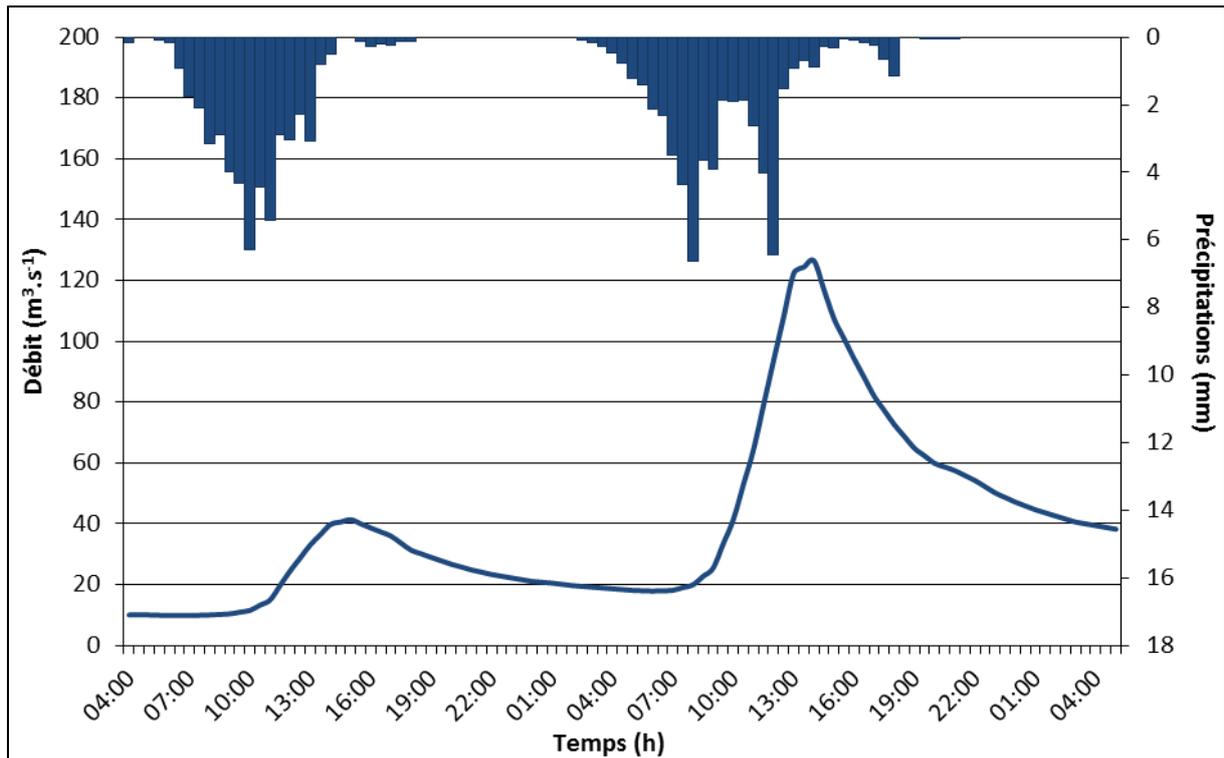


Figure 43: hydrogramme et hyétogramme de l'évènement n°212.

Ici, l'on constate que pour deux averses presque identiques en volume moyen sur le bassin, la réaction est plus de trois fois plus importante pour la deuxième, témoignant de l'importance de l'évolution de l'état d'humidité du bassin dans l'appréhension de la relation pluie débit, en plus de la connaissance des conditions d'humidité initiales. Par ailleurs, la deuxième réaction est simple alors que le signal de pluie est double.

Ces constats sont à mettre en parallèle avec les éléments sur la dynamique de genèse des crues vus dans le chapitre 2. Ils sont l'une des preuves de la complexité des processus mis en jeu, des processus difficiles à quantifier relativement les uns aux autres, et en absolu ; des processus difficiles à recenser, difficiles à mesurer en temps réel et enfin, dont les mesures sont difficilement généralisables. C'est dans ce cadre que l'utilisation de prédicteurs neuronaux basés sur l'apprentissage statistique de la relation causale entre pluie et débit paraît être l'une des solutions les plus appropriées à la problématique.

### 4.3 Synthèse de prédicteurs neuronaux

Suite à la présentation des propriétés et de la synthèse des modèles à apprentissage statistique pour la prévision dans le chapitre 3, il est apparu logique d'utiliser ce type d'approche pour modéliser une relation non linéaire mal connue telle que la prévision du débit en l'absence de prévision de pluie. Rappelons les principales étapes de la démarche de synthèse d'un modèle neuronal décrite en section 3.4 :

- L'établissement d'un modèle postulé du comportement du processus non-linéaire pour lequel on souhaite obtenir des prévisions. Compte tenu des éléments connus de la bibliographie, il apparaît clairement que le modèle doit recevoir en entrées les précipitations ainsi que les débits précédents comme information d'état (Toukourou, 2009).

- Pour éviter le surajustement, la sélection de variables doit être réalisée afin de limiter la complexité du modèle et ainsi préserver ses propriétés de généralisation. Cette étape va consister en la définition précise des fenêtres temporelles transcrivant l'historique de précipitations qui alimentent les modèles ainsi que l'historique des débits précédents (observés ou estimés) décrivant l'état du bassin. Schématiquement, on peut comprendre que le modèle recevant des entrées de précipitations élevées, mais un débit faible correspond à un bassin non encore saturé en eau. A l'inverse, si les précipitations sont élevées et que les débits précédents sont importants et en augmentation, le bassin est en phase de saturation et l'eau est rapidement transférée au cours d'eau. Ainsi le nombre d'entrées de débit peut être interprété comme une information riche qui indique au modèle l'état d'humidité du bassin et la phase de l'hydrogramme (montée de crue si la différence entre deux débits successifs est positive, phase de récession dans le cas contraire).
- Une fois les variables d'entrées déterminées il convient de fixer la complexité du modèle, c'est-à-dire le nombre de neurones cachés.

Ces ajustements doivent être réalisés pour chacun des modèles conçus. Deux types de modèles sont proposés pour cette application :

- Un modèle statique alimenté par les pluies et les débits spécifiques mesurés aux pas de temps précédents tel que présenté en section 3.1 figure 22 : c'est le cas du bassin correctement jaugé où les mesures permettent d'obtenir une bonne base d'apprentissage et une information fiable en temps réel. Ces modèles sont réputés être les plus efficaces pour la simulation ou la prévision des débits comme des hauteurs d'eau (Kong A Siou, 2011).
- Un modèle récurrent alimenté par les pluies et les débits spécifiques estimés aux pas de temps précédents, tel que présenté en section 3.1 figure 23. Dans le cas où l'information de débit n'est pas disponible, il est alors naturel d'utiliser le débit spécifique estimé par le réseau lui-même. Ces modèles sont alors récurrents et plus difficiles à utiliser car ils peuvent être instables. Ce type de modèle est le seul utilisable pour les bassins mal jaugés où les mesures permettent d'obtenir une bonne base d'apprentissage mais ne sont pas assez fiables en temps réel.

Ces deux types de modèles vont donc être mis en œuvre et leurs résultats respectifs pourront être comparés.

Les expériences ont été menées pour des horizons de prévision allant de  $k+1$  à  $k+4$  (rappelons que  $k$  est le temps discrétisé), soit une demi-heure à deux heures et sur quatre événements différents qui seront décrits dans cette partie. La modélisation est une modélisation pluie-débit spécifique dont les sorties sont systématiquement corrigées en débits à la sortie des modèles. Cette manipulation facilitera, nous le verrons, le travail sur les bassins non jaugés.

Dans un premier temps, les expériences conduites sur de nombreuses architectures différentes ont conduit à la réalisation d'un modèle non-linéaire récurrent standard, notamment pour ses propriétés d'approximation universelle et de parcimonie (cf section 3.2), dont nous rappelons l'architecture (figure 44). Les autres architectures ont toutes montré des résultats inférieurs à ceux de ce modèle.

Les variables d'entrées sont les historiques de précipitations relevés pour les stations de Mialet, Saint-Roman-de-Tousque, Barre-des-Cevennes, ainsi que le cumul moyen de pluie tombé sur le

bassin depuis le début de l'évènement. Le cumul moyen a été ajouté en entrée afin d'apporter une information sur le niveau d'humidité du bassin. En effet, afin d'optimiser la complexité du modèle, il n'est pas possible d'appliquer en entrée l'historique total des précipitations ; ces dernières sont donc appliquées au travers d'une fenêtre temporelle glissante de durée relativement courte par rapport à la durée de l'évènement. En revanche la mémoire de la quantité de pluie tombée depuis le début de l'évènement peut être appliquée de manière parcimonieuse sous la forme du cumul moyen de précipitation. De la même manière que les fenêtres temporelles de précipitation l'entrée cumul doit être ajustée précisément. Ici, la dimension des fenêtres temporelles a été sélectionnée sur le score de validation croisée au terme de plans d'expériences voués à trouver la meilleure combinaison.

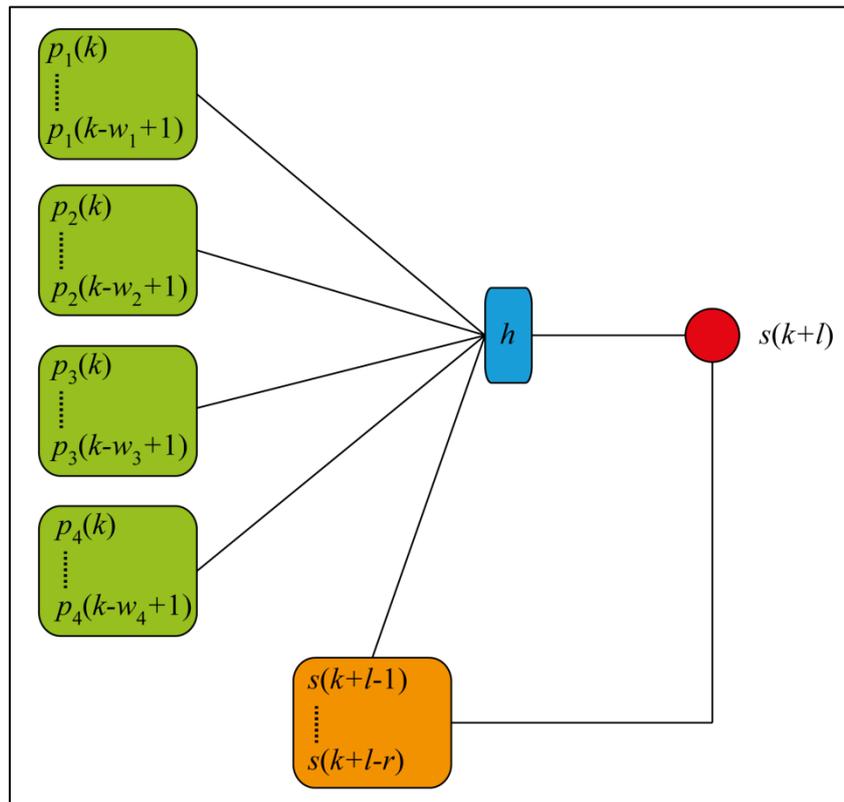


Figure 44: modèle utilisé dans les premières simulations. En vert : les entrées ; en bleu : la couche cachée ; en orange : le débit estimé (entrée récurrente) ; en rouge la sortie.

Les notations y 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 (c'est-à-dire, en quelque sorte, la dimension de la fenêtre temporelle d'entrée récurrente des débits estimés),
- $l$  correspond à l'horizon de prévision visé (*lead-time*),
- $s(k+l)$  est la valeur de la fonction de sortie estimée par le modèle (débit spécifique).

Sont représentées en vert les fenêtres temporelles d'entrée, en bleu la couche cachée du réseau, en rouge la sortie du réseau et en orange, l'éventuelle entrée récurrente.

On peut ainsi écrire l'équation du modèle :

$$s(k+l) = g_{NN}(s(k+l-1), \dots, s(k+l-r), u(k), u(k-1), \dots, u(k-w+1))$$

où  $s$  est le débit estimé,  $k$  le temps discrétisé,  $\mathbf{u}$  le vecteur des variables de forçage  $r$  l'ordre du modèle,  $w$  la longueur de la fenêtre temporelle d'entrée,  $l$  l'horizon de prévision,  $g_{NN}$  la fonction non linéaire implémentée par le modèle.

Les meilleurs résultats obtenus par cette architecture, qui était la meilleure parmi de nombreuses autres, n'ont pas été à la hauteur des attentes dictées par la problématique. Par exemple, pour le modèle dont l'optimisation de la complexité a mené à l'architecture présentée en tableau 4, le tableau 5 présente les meilleurs résultats obtenus par ce modèle sur l'évènement qui a donné les meilleures performances (l'évènement n°13).

Tableau 4 : indication de l'architecture du premier type de modèle utilisé.

Horizon ( $l$ )	Cumul moyen ( $w_4$ )	Barre-des- Cevennes ( $w_3$ )	Saint-Roman- de-Tousque ( $w_2$ )	Mialet ( $w_1$ )	Couche cachée ( $h$ )	Ordre ( $r$ )
$k+1$	1	10	7	6	2	1
$k+2$	2	10	7	4	3	2
$k+3$	2	10	6	5	3	2
$k+4$	2	9	7	4	4	2

Rappelons que les critères de Nash et le *SPPD* ont été définis au paragraphe 2.2.1. Le critère de Nash n'est autre que le coefficient de détermination, tandis que le *SPPD*, exprimé en pourcentage, représente le rapport entre le débit simulé et le débit observé à l'instant du pic observé.

Tableau 5: meilleurs résultats obtenus avec un modèle non linéaire simple sur l'évènement n°13.

Horizon	Nash	<i>SPPD</i>
$k+1$	0,6	50%
$k+2$	0,71	44%
$k+3$	0,73	42%
$k+4$	0,56	43%

Les critères de performance sont médiocres. En réalité, si la non-linéarité est importante dans la relation entre pluie et débit, une part plus ou moins importante de linéarité existe. Comme ce premier modèle comportait un faible nombre de neurones sur la couche cachée après optimisation (souvent deux neurones cachés), nous en avons déduit que la partie linéaire du processus était importante à prendre en compte et que le modèle n'y arrivait pas.

Comme cela a pu être réalisé dans beaucoup d'approches de modélisation, une partie de relation linéaire et une partie de relation non linéaire peuvent être distinguées. Nous avons pratiqué des simulations pour un modèle totalement linéaire et, au regard des résultats en mode non-linéaire et en mode linéaire, avons choisi de combiner un modèle linéaire avec un autre, non linéaire, pour permettre à la couche de neurones cachés de se concentrer de façon exclusive sur la partie non-linéaire de la relation. Les résultats ont alors été meilleurs et présentent suffisamment d'intérêt pour être décrits dans ce qui suit.

Notons qu'avec le formalisme des réseaux de neurones, il est très aisé de réaliser un modèle linéaire puisqu'il suffit de construire une architecture à un seul neurone linéaire. Cette facilité est un avantage non négligeable pour de telles pistes de recherche.

Le modèle proposé explicitant les contributions linéaires et non linéaires, à comparer au perceptron multicouche classique (figure 44) est présenté dans la figure 45 tandis que celui correspondant au modèle récurrent est présenté figure 46.

Les notations  $y$  sont identiques à celles de la figure 44, à ceci près que  $q$  correspond au débit observé.

Sont toujours représentées en vert les fenêtres temporelles d'entrée, en bleu la couche cachée du réseau, en rouge la sortie du réseau et en orange, l'éventuelle entrée récurrente.

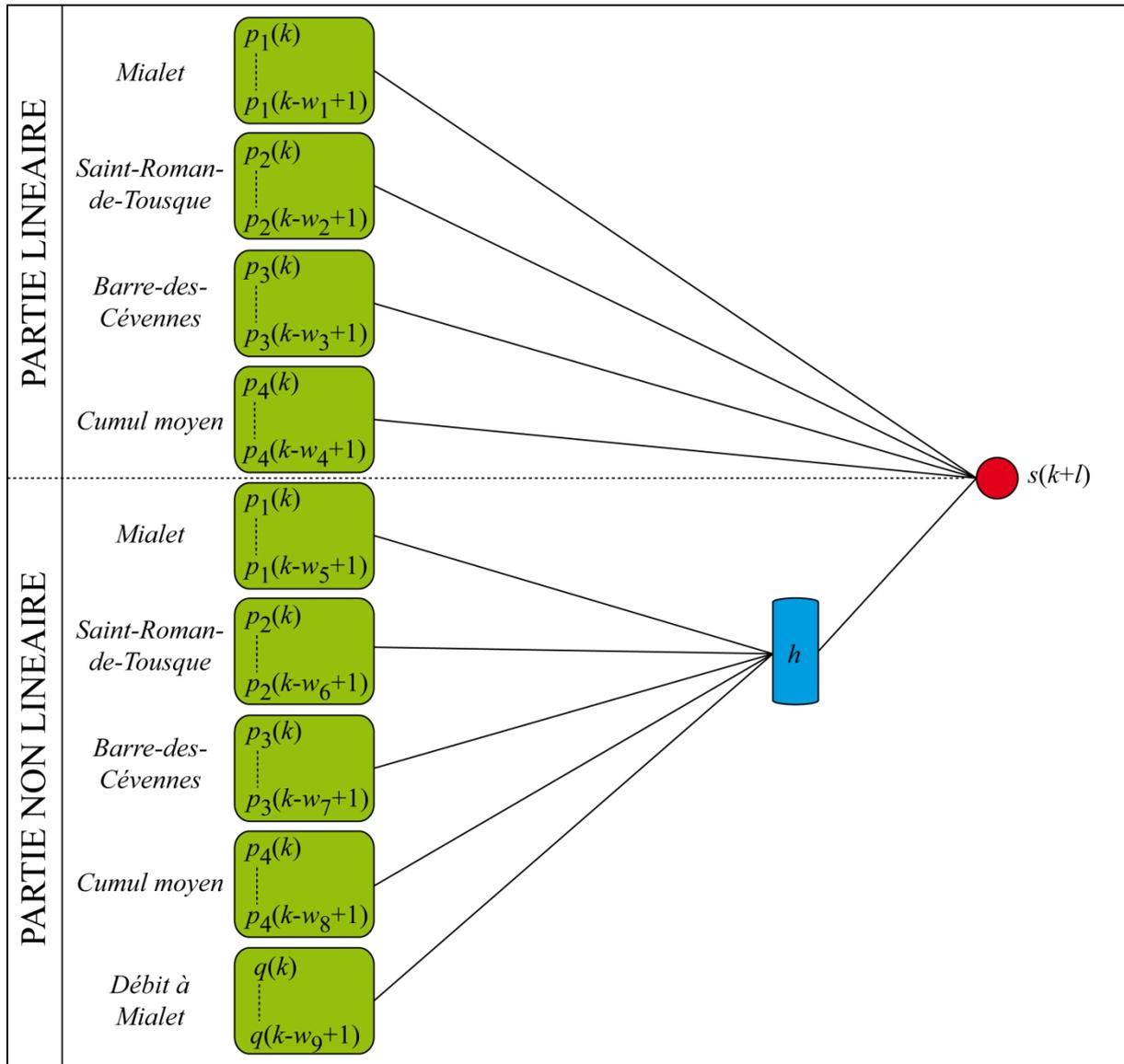


Figure 45: modèle statique utilisé dans la prévision des crues du Gardon de Mialet.

Ce modèle peut s'écrire avec des notations identiques aux précédentes :

$$s(k+l) = g_{LIN}(u(k), u(k-1), \dots, u(k-w_{LIN}+1)) + g_{NN}(u(k), u(k-1), \dots, u(k-w_{NN}+1))$$

où les indices  $lin$  correspondent aux vecteurs et variables affectés à la partie linéaire et les indices  $nn$  correspondent aux vecteurs et variables affectés de la partie non linéaire.  $g_{LIN}$  est la fonction implémentée par la partie linéaire du modèle. Dans ce cas, les vecteurs  $\mathbf{u}$  incluent aussi les débits observés aux  $w_9$  pas de temps précédents. Ces derniers pourraient être considérés

comme l'ordre du modèle si l'équation représentait le système hydrologique mais dans le cas d'un modèle statique, ils demeurent de simples variables.

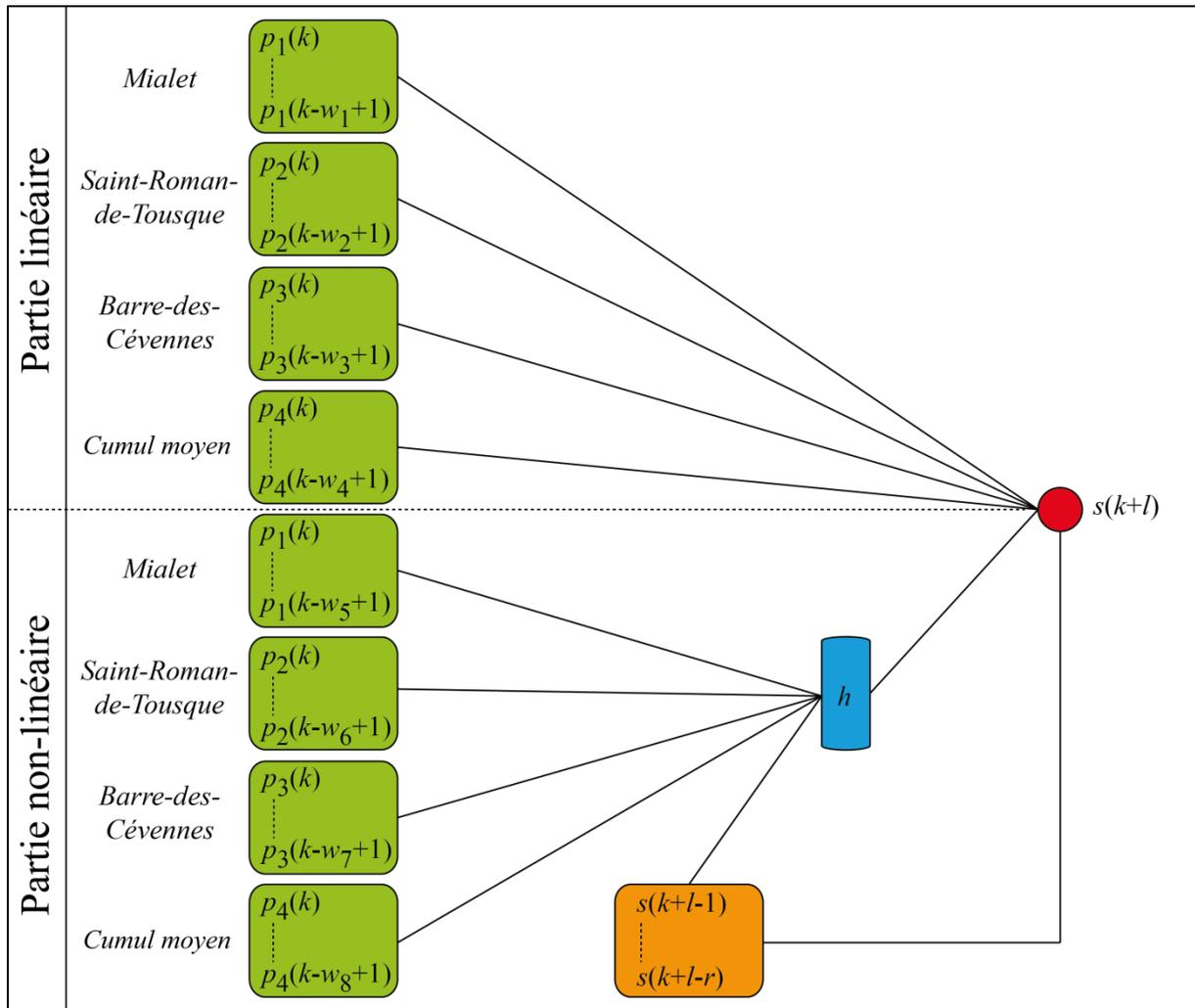


Figure 46: modèle récurrent utilisé dans la prévision des crues du Gardon de Mialet.

Ce modèle peut s'écrire, avec des notations identiques aux précédentes :

$$s(k+l) = g_{LIN}(u(k), u(k-1), \dots, u(k-w_{LIN}+1)) + g_{NN}(s(k+l-1), \dots, s(k+l-r), u(k), u(k-1), \dots, u(k-w_{NN}+1))$$

Dans les deux cas, les entrées sont dédoublées pour alimenter séparément une partie linéaire et une partie non-linéaire avec des fenêtres temporelles différentes. La seule différence notable entre les deux modèles tient à l'utilisation des débits précédemment observés (jusqu'à  $k$ ) dans le modèle statique (figure 45) ou précédemment estimés (jusqu'à  $k-1$ ) dans le modèle récurrent (figure 46). Pour ce dernier modèle, le nombre de sorties estimées utilisées correspond à l'ordre du modèle.

#### 4.3.1 Sélection de variables ou dimensionnement des fenêtres temporelles d'entrée

Nous présentons dans ce paragraphe la méthode choisie pour sélectionner les fenêtres temporelles de précipitations (pour chaque pluviomètre) et de débit (observé ou estimé selon le modèle non récurrent ou récurrent choisi). Comme présenté au chapitre 2, la sélection de ces fenêtres va permettre de minimiser la dimension du vecteur des entrées, et par là, de minimiser

le nombre de paramètres du modèle, diminuant ainsi sa complexité et permettant donc d'augmenter ses capacités de généralisation. Comme pour toute démarche de sélection de variables, principalement deux approches peuvent être adoptées : (i) la méthode experte qui utilise les connaissances que l'on a sur le processus physique et (ii) la méthode empirique qui ajuste itérativement le meilleur modèle par essais successifs. Ces deux types de méthodes, complémentaires, ont été mises en œuvre dans ce travail. Elles sont présentées dans la partie suivante.

#### *4.3.1.1 Méthode experte de sélection de variables*

Dans le système hydrologique, le temps qui s'écoule entre la présentation d'un forçage à l'entrée et l'apparition d'une réponse à la sortie peut être assimilé au temps de réponse tel que défini en hydrologie, c'est-à-dire le temps qui s'écoule entre le centre de gravité du hétérogramme et le centre de gravité de l'hydrogramme (UNESCO – OMM, 1992).

Mangin (1981) montre que si les forçages d'entrée sont considérés comme des processus aléatoires (c'est le cas avec les pluies, en particulier dans le cas de situations fortement convectives), un corrélogramme croisé entrée-sortie représente la réponse impulsionnelle du système, ou tout du moins pour un système non-linéaire la réponse impulsionnelle correspondant au point de fonctionnement (l'état actuel du système).

Les fenêtres temporelles d'entrée, notées  $w$ , permettent de fournir simultanément les  $w$  dernières informations de la variable d'entrée considérée. On comprend que si on limite  $w$  au temps de réponse associé aux pluies de chaque pluviomètre, on limite le volume de données au strict nécessaire pour le fonctionnement du modèle. Des fenêtres temporelles trop longues détérioreraient la parcimonie du modèle et fourniraient des informations supplémentaires de nature à desservir la clarté de la relation entre entrée et sortie.

Ce procédé employé et décrit par Mangin (1984) ou encore Kong A Siou et al. (2011) sur des systèmes karstiques peut aisément être transposé aux systèmes hydrologiques rapides de surface.

Dans le cas qui nous intéresse, une analyse corrélatrice a été menée pour chaque pluviomètre en relation avec le débit observé à l'exutoire. Sur le corrélogramme croisé, le premier maximum a été interprété comme le temps de réponse. Seuls les corrélogrammes montrant un maximum bien identifiable ont été conservés. Par exemple, on peut comprendre que pour l'évènement n°230 présenté dans la figure 40, le corrélogramme est difficilement exploitable. Par ailleurs, durant certains évènements, il arrive qu'un pluviomètre ne collecte que peu voire pas de précipitations. Il est donc exclu du processus puisque son effet n'est que marginal sur le débit par rapport aux autres pluviomètres bien plus arrosés.

De plus, autant parce que le modèle défini a vocation à s'appliquer aux évènements intenses, que parce que les évènements trop faibles ne donnent pas lieu à des corrélogrammes exploitables, nous avons sélectionné les évènements pour estimer le temps de réponse. Deux critères ont été utilisés :

- les évènements présentant des intensités pluvieuses maximales supérieures à la médiane des intensités pluvieuses maximales des évènements pour le pluviomètre visé,
- les évènements présentant des cumuls supérieurs à la médiane des cumuls des évènements conservés pour le pluviomètre visé.

L'utilisation des deux méthodes de sélection ici décrites permet d'évaluer la pertinence des résultats. De fait, l'une et l'autre ont donné des résultats quasiment identiques. Des histogrammes de fréquence d'apparition des temps de réponse par pluviomètre ont été tracés et ont permis d'obtenir les intervalles de temps de réponse présentés dans le tableau 6.

La figure 47 présente l'histogramme obtenu pour Barre des Cévennes. Un pic apparaît pour la classe 7 à 9. De façon plus précise, les plus fréquentes apparitions de temps de réponse se situent autour de la classe 9. L'investigation de la dimension des fenêtres temporelles d'entrée a donc été menée de 7 à 11 pas de temps (soit 3,5 à 5,5 heures).

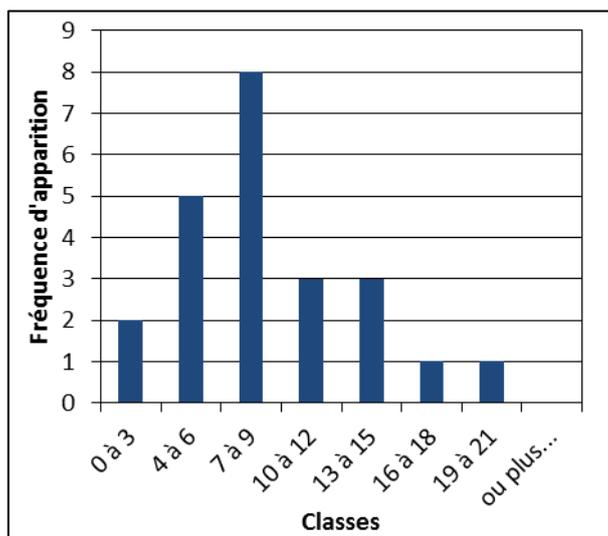


Figure 47: histogramme de fréquence d'apparition de différents temps de réponse entre les précipitations enregistrées à Barre-des-Cévennes et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h).

La figure 48 présente l'histogramme obtenu pour Saint-Roman-de-Tousque. Un pic apparaît pour la classe 4 à 6. De façon plus précise, les plus fréquentes apparitions de temps de réponse se situent autour de la classe 6. L'investigation de la dimension des fenêtres temporelles d'entrée a donc été menée de 4 à 9 pas de temps (soit 2 à 4,5 heures).

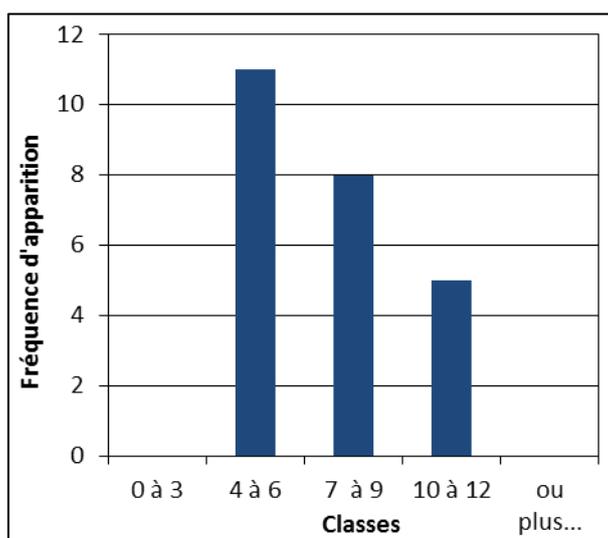


Figure 48: histogramme de fréquence d'apparition de différents temps de réponse entre précipitations enregistrées à Saint-Roman-de-Tousque et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h).

Enfin, la figure 49 présente l’histogramme obtenu pour Mialet. Un pic apparait pour la classe 4 à 6. De façon plus précise, les plus fréquentes apparitions de temps de réponse se situent autour des classes 4 et 6, très peu au delà. L’investigation de la dimension des fenêtres temporelles d’entrée a donc été menée de 2 à 7 pas de temps (soit 1 à 3,5 heures).

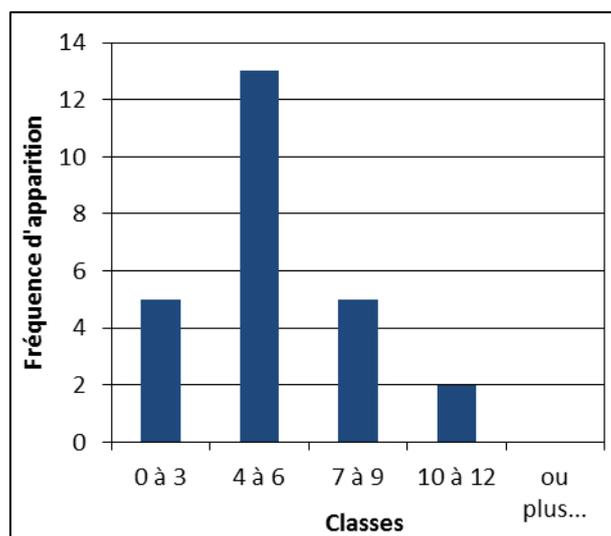


Figure 49: histogramme de fréquence d’apparition de différents temps de réponse entre précipitations enregistrées à Mialet et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h).

Dans le tableau 6, on donne le temps de réponse de la classe la plus représentée de l’histogramme. Les intervalles d’investigation pour les optimisations (pour les méthodes empiriques qui vont suivre) représentent les temps de réponse les plus représentés (plus de 66% des temps de réponse obtenus).

Tableau 6: temps de réponse et intervalles d’investigation pour les optimisations, par pluviomètre.

Pluviomètre	Mialet	Saint-Roman-de-Tousque	Barre-des-Cévennes
Temps de réponse (*0,5h)	4	6	9
Intervalle d’investigation (*0,5h)	2 à 7	4 à 9	7 à 11

Suivant cette méthode de définition des fenêtres temporelles, les fenêtres à appliquer pour les variables de précipitation des trois pluviomètres sont donc voisines de celles données par les temps de réponse du tableau. Les temps de réponse seraient donc, en moyenne, de deux heures pour Mialet, trois heures pour Saint-Roman-de-Tousque et quatre heures trente pour Barre-des-Cévennes. Une logique géographique est respectée : les valeurs des pluviomètres les plus amont sont les valeurs générant une corrélation plus tardive avec le débit à l’exutoire. Cependant comme l’analyse corrélatoire ne prend pas en compte le caractère non linéaire du processus étudié, nous avons souhaité valider ce premier dimensionnement à l’aide d’une autre méthode qui intègre la part non linéaire du processus.

#### 4.3.1.2 Méthode empirique de sélection de variables

Dans le chapitre 3 nous avons présenté la validation croisée comme une méthode empirique de sélection de variables qui a l’avantage de pouvoir opérer en l’absence de connaissance sur le phénomène modélisé. Ainsi la validation croisée va nous permettre de sélectionner aussi bien la

dimension optimale des fenêtres temporelles de précipitation, que celle des débits observés précédents ou celle de l'ordre du modèle, selon que l'on considère le modèle non récurrent ou récurrent.

La validation croisée faisant intervenir l'apprentissage du réseau, il convient dans un premier temps de définir quels vont être les sous-ensembles d'apprentissage, d'arrêt et de test que nous avons définis au chapitre 3.

De manière simple, la base d'apprentissage comporte tous les évènements moins ceux utilisés pour l'arrêt précoce et le test. 56 évènements étaient donc disponibles pour l'apprentissage.

Dans son esprit, l'approche de validation croisée doit être appliquée à l'ensemble de la base. Pourtant, il est possible de procéder en n'utilisant la validation croisée que sur des évènements qui présentent des caractéristiques intéressantes pour la problématique explorée. Parmi tous les modèles conçus (avec les différentes initialisations des paramètres, les différentes fenêtres temporelles, etc.), la validation croisée peut donc être utilisée comme une méthode de sélection de modèles. Dans notre étude, elle permet de sélectionner les modèles les plus performants sur les évènements intenses ( $>1\text{m}^3\cdot\text{s}^{-1}\cdot\text{km}^2$ ) qui sont connus pour être les plus dangereux mais aussi les plus difficiles à modéliser, et donc à prévoir.

Cette utilisation de la validation croisée est appelée « validation croisée partielle » (Toukourou, 2009) ; elle a par ailleurs l'avantage d'opérer sur moins d'évènements et donc de diminuer le temps de calcul de cette méthode consommatrice d'unités de calcul.

De plus, pour permettre au modèle de répondre rapidement et précisément à une impulsion de pluie, nous avons retenu pour effectuer la validation croisée partielle les évènements qui présentent dans tout ou partie de leur déroulement une forme impulsionnelle, c'est-à-dire ceux dans lesquels il existe une relation indubitable entre un stimulus pluvieux puissant et identifié et sa réponse hydrologique. Un exemple d'un tel type d'évènement est présenté en figure 50.

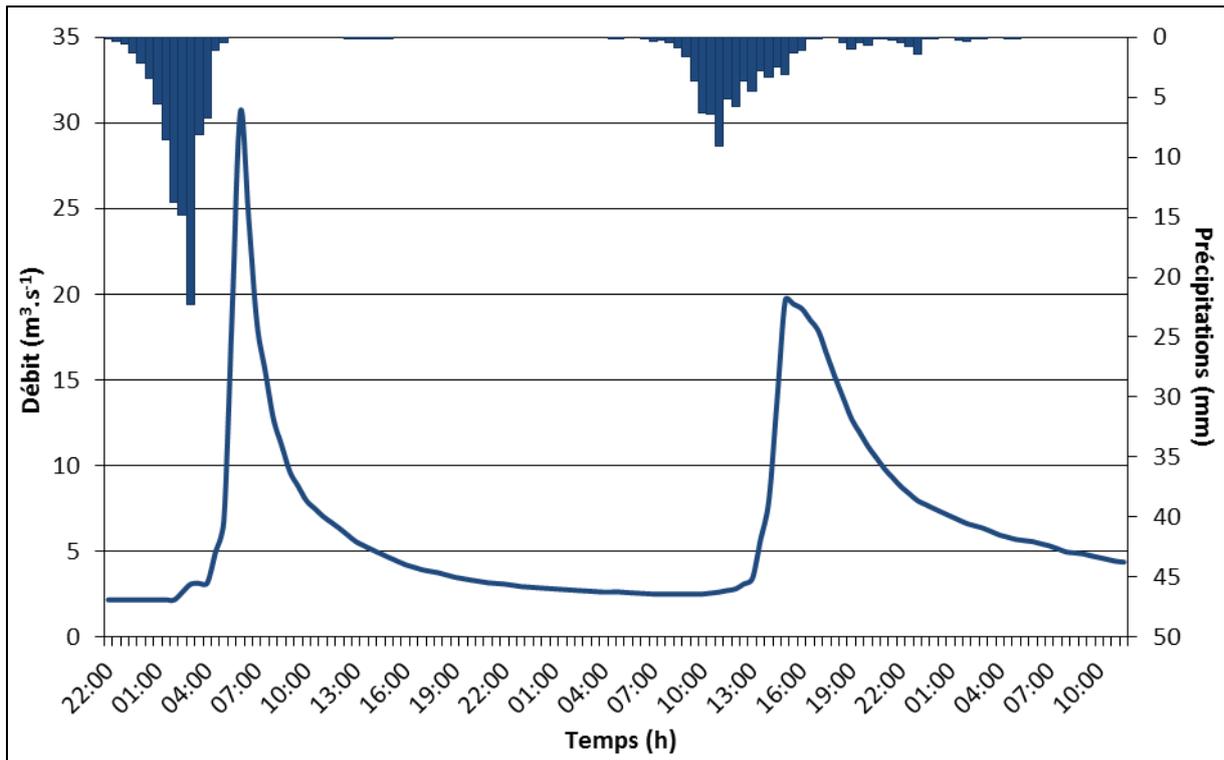


Figure 50: évènement n°208 présentant une pluie clairement reliée à une réponse hydrologique, comme une réponse à une impulsion.

Cela nous a conduits à sélectionner 18 évènements auxquels il faut le cas échéant, retrancher l'évènement utilisé en test.

L'évènement choisi pour effectuer le test pourrait être n'importe lequel dans la série 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).

Chaque ensemble de test différent a donc donné lieu à un ensemble d'apprentissage différent et à un ensemble d'évènements intenses ou impulsions (pour la validation croisée) éventuellement différent (selon que l'évènement de test choisi a les caractéristiques des évènements de validation croisée ou non). Les changements d'ensemble pour tester les différents évènements ont été opérés en maintenant les règles définies ci-dessus.

Le tableau 7 précise quelques caractéristiques des évènements choisis pour le test.

Tableau 7 : 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.

N°	Date	Pics	$Q_x$ ( $m^3.s^{-1}$ )	$Q_m$ ( $m^3.s^{-1}$ )	$RR$ (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

Une fois les ensembles d'apprentissage et de test définis, la validation croisée partielle est mise en œuvre en faisant varier :

- L'initialisation aléatoire des paramètres du réseau. En effet, du fait de la présence de minimum locaux lors de la minimisation de l'erreur quadratique il est intéressant d'initialiser le modèle dans différentes valeurs initiales afin de ne conserver, après apprentissage, que la meilleure des initialisations.
- Les dimensions des fenêtres temporelles des pluviomètres.
- Les dimensions des fenêtres temporelles des débits observé ou estimés, selon que le modèle est statique ou récurrent.

Pour chacun des modèles ainsi créé, la validation croisée partielle est réalisée et le score de validation croisée partielle est calculé tel que présenté en section 3.3.3.2. Le meilleur score de validation croisée partielle permet de déterminer quelle est la fenêtre temporelle adéquate.

Concernant les optimisations spécifiques des initialisations des paramètres, il a pu être constaté une très faible sensibilité aux différentes suites aléatoires de paramètres utilisées, dans les deux types de modèles (récurrent non-dirigé ou statique dirigé). Sur 50 suites d'initialisation aléatoires pour une même architecture, l'écart-type des scores de validation croisée ne dépassait jamais 5% de la moyenne.

Les intervalles d'investigation définis par la méthode des corrélations croisées ont été utilisés pour l'optimisation des dimensions des fenêtres temporelles. Lorsque l'optimisation a donné un résultat correspondant à une borne de l'intervalle d'investigation, ce dernier a été élargi pour s'assurer que l'optimum ne dérogeait pas aux conjectures émises dans la définition de l'intervalle. Cela n'a jamais été le cas. Les investigations ont été menées pour chaque entrée, y compris pour les doublons.

Pour les autres intervalles, une investigation plus large couplée à une réflexion sur la fonction des entrées dans le modèle ont été menées :

- Pour le cumul moyen, dans chaque partie du modèle, l'on a supposé que les valeurs devaient être plutôt faibles puisqu'il est important de connaître la valeur absolue du cumul pour obtenir l'état du bassin et puisque les informations des trois pluviomètres renseignent sur l'évolution de cet état. Nous avons mené ces investigations sur un intervalle de 1 à 7 pas de temps.
- Pour l'entrée récurrente, de la même manière que pour le débit observé, il peut être intéressant d'avoir la valeur estimée ou observée et une autre valeur précédente au moins pour connaître l'évolution du débit, qu'il soit estimé ou observé. Les investigations ont donc été menées de 1 à 5 pas de temps.

Dans le même temps, il a fallu dimensionner la complexité du réseau à proprement parler, c'est-à-dire le nombre de neurones cachés. Ceci a donc été réalisé en même temps que les fenêtres

temporelles avec la validation croisée partielle, en faisant varier le nombre de neurones cachés de 1 à 5. Toute l'étape d'optimisation a donc été conduite en minimisant le score de validation croisée obtenu sur chacune des nombreuses combinaisons possibles.

#### 4.3.2 Architecture optimale

Les architectures optimales obtenues pour les deux types de modèles en prévision de  $k+1$  à  $k+4$  sont présentées dans le tableau 8. L'horizon de prévision a été limité à deux heures en raison du temps de réponse estimé à Mialet, qui est de deux heures.

Tableau 8: architectures optimales pour les deux types de modèles utilisés. Les valeurs sont indiquées en nombre de pas de temps (soit en 0,5h).

Élément 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-Cevennes / $w_3$	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-Cevennes / $w_7$	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

Au niveau des entrées, nous avons précisé qu'aucune fenêtre temporelle n'était sortie des intervalles d'investigation définis par les corrélations croisées. Mieux, l'ordre géographique des pluviomètres est respecté puisqu'aucune fenêtre temporelle n'est plus grande que celle correspondant au pluviomètre plus amont. Sans encore avoir apporté de résultats, les modèles ont déjà identifié une architecture correspondant à la configuration du bassin et des sites de mesure.

La complexité reste assez limitée (peu de neurones sur la couche cachée), mais il apparait que la part non-linéaire du modèle améliore les prévisions par rapport à la seule contribution linéaire.

L'évolution des cumuls moyens semble prendre plus d'importance dans la partie linéaire du modèle puisque les optimisations ont demandé plusieurs pas de temps de recul. Ceci peut être compris comme une utilisation plus importante du cumul par la partie linéaire, suggérant que le mécanisme qui conduit le bassin à être plus réactif lorsque le cumul augmente est assez peu complexe et peut être décrit par une relation linéaire. Ceci-dit, l'évolution du cumul sur une certaine durée est nécessaire aux meilleures configurations de fonctionnement.

Enfin, l'ordre et le nombre de débits observés précédents (suivant le type de modèle) sont minimaux pour  $k+1$  ce qui laisse penser que pour un horizon de prévision faible, l'évolution des débits précédents, estimés ou mesurés, est moins importante que leur grandeur. En revanche, cette évolution semble devoir être prise en compte pour des horizons plus lointains, ce qui ne semble, là aussi, pas dénué de sens vis-à-vis du fonctionnement du système hydrologique : pour se projeter à des horizons de prévision plus importants, le modèle a besoin d'un historique un peu plus profond et donc de prendre en compte la tendance à l'augmentation ou à la diminution sur une période plus importante.

### 4.3.3 Régularisation par l'arrêt précoce

L'arrêt précoce a été présenté en section 3.3.3.3. Il consiste à arrêter la phase d'apprentissage à partir du moment où la capacité de généralisation commence à se dégrader. Pour cela il faut choisir un ensemble de données complémentaire : l'ensemble d'arrêt.

Le principal inconvénient de cette méthode est que le choix de l'ensemble d'arrêt peut influencer grandement le modèle obtenu si la base de données est réduite et que l'ensemble d'arrêt ne représente pas complètement la base d'apprentissage. Ce cas correspond à notre cadre de travail puisque les événements sont en nombre réduit, et très différents les uns des autres. Toukourou et al. (2011) ont proposé une solution élégante pour déterminer l'ensemble d'arrêt dans une problématique similaire : afin qu'il représente au mieux la base d'apprentissage, on choisit comme ensemble d'arrêt l'événement le mieux prédit de toute la base. Ceci signifie que c'est celui qui est le plus en cohérence avec le reste de l'ensemble des événements.

Nous avons donc procédé à une succession d'apprentissages (autant que d'évènement dans la base) afin d'estimer l'erreur de chaque événement en situation de validation. Pour cela les fenêtres temporelles ont été dimensionnées avec les temps de réponse (cf tableau 6). L'évènement qui a donné le meilleur score en validation est l'évènement n°222 de janvier 1998 (figure 51).

Il s'agit d'un évènement plutôt simple, intense au sens que nous avons défini dans la section 4.1, marqué par des précipitations qui s'intensifient progressivement pour s'arrêter plus brusquement, ce qui est assez typique des épisodes cévenols simples. La réaction assez importante est surtout liée aux pluies plus abondantes et plus intenses sur l'amont, également probablement à la saison (le mois de janvier présente généralement des taux d'écoulement importants). Le fait que cet évènement simple et intense soit le mieux prédit par le modèle est satisfaisant. Il aurait en effet été inquiétant que le modèle soit le plus performant sur un évènement faible et lent.

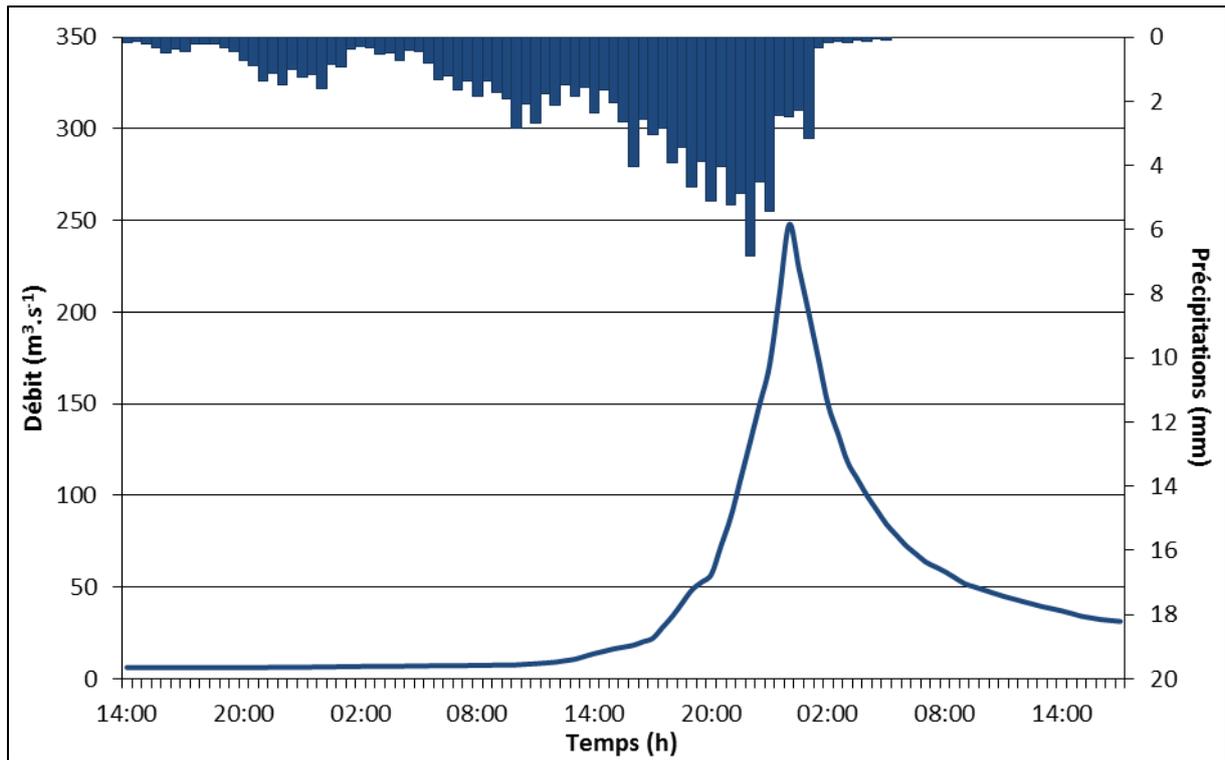


Figure 51: évènement n°222 de janvier 1998 utilisé comme ensemble d'arrêt.

#### 4.4 Résultats en prévision en l'absence de prévisions de pluie

Dans tous les cas qui seront évoqués ici, la performance a été évaluée par plusieurs moyens :

- le critère de Nash,
- le critère de persistance (utile seulement pour le modèle statique dirigé, puisqu'il existe un risque de prévision naïve de la part du modèle),
- le *SPPD* et le *PPD*,
- le décalage temporel, en pas de temps, du pic prévu par rapport au pic observé,
- le dernier, plus qualitatif, correspond à la comparaison entre le niveau de vigilance qui aurait été déclenché si l'on avait suivi le modèle et celui qui correspond aux mesures de débit effectuées pour le cours d'eau de Mialet. Les seuils de vigilance sont définis par le service de prévision des crues, selon les règles détaillées dans le chapitre 1. Pour le bassin versant du Gardon à Mialet, ces seuils sont présentés dans le tableau 9.

Tableau 9 : Seuils de débit approximatifs correspondant aux niveaux de vigilance.

Niveau	Seuil de débit
Vert	< 105 m <sup>3</sup> .s <sup>-1</sup>
Jaune	> 105 m <sup>3</sup> .s <sup>-1</sup>
Orange	> 370 m <sup>3</sup> .s <sup>-1</sup>
Rouge	> 600 m <sup>3</sup> .s <sup>-1</sup>

Il convient de remarquer ici que la décision de diffuser un niveau de vigilance n'est pas prise de manière automatique uniquement en fonction du débit. Le prévisionniste doit aussi identifier les augmentations plus ou moins locales ou plus ou moins ponctuelles des enjeux à proximité du cours d'eau.

Notons encore avant de débiter l'exposé des résultats que toutes les prévisions fournies le sont sans aucune prévision de la pluie au-delà de l'instant de prévision. En d'autres termes, la pluie observée s'arrête à l'instant de prévision  $k$  et n'est en aucun cas prolongée par une pluie prévue (nulle ou constante) ou par une pluie « parfaite » entre  $k$  et  $k+l$  (où  $l$  est l'horizon de prévision). De même, aucun débit mesuré au-delà de l'instant de prévision n'est utilisé. Aucun débit estimé qui n'aurait pas encore été estimé par le modèle n'est utilisé. La configuration est celle des conditions de prévision opérationnelles.

#### 4.4.1 Modèle statique dirigé : cas du bassin bien jaugé

Le modèle statique dit « dirigé », c'est-à-dire dirigé par les débits observés aux pas de temps précédents et jusqu'à l'observation, tant en apprentissage qu'en utilisation (test), donne d'excellents résultats en prévision pour les quatre événements et les quatre horizons de prévision sélectionnés. Cela confirme, si besoin en était, que la disponibilité des données est un des facteurs clé de notre problématique.

Le tableau 10 présente les résultats obtenus pour chaque événement de test et chaque horizon de prévision.

Tableau 10: Récapitulatif des résultats obtenus avec le modèle statique dirigé (Evt. correspond au numéro de l'évènement,  $l$  à l'horizon de prévision, Pers. au critère de persistance,  $V_p$  à la vigilance prévue par le modèle,  $V_o$  à la vigilance correspondant au débit observé à  $k+l$  et D au décalage temporel).

<b>Evt.</b>	<b><math>l</math> (0.5 h)</b>	<b>Nash</b>	<b>Pers.</b>	<b>PPD</b>	<b>SPPD</b>	<b><math>V_p</math></b>	<b><math>V_o</math></b>	<b>D(0.5 h)</b>
13	1	0.99	0.73	98%	94%	O	O	1
13	2	0.98	0.92	96%	95%	O	O	1
13	3	0.97	0.96	96%	88%	O	O	1
13	4	0.95	0.95	95%	94%	O	O	1
19	1	0.93	0.42	86%	71%	R	R	-1
19	2	0.94	0.65	89%	89%	R	R	0
19	3	0.94	0.7	94%	94%	R	R	0
19	4	0.78	0.67	87%	82%	R	R	1
8	1	0.97	0.64	99%	99%	J	J	0
8	2	0.94	0.75	90%	85%	J	J	1
8	3	0.91	0.67	88%	79%	J	J	1
8	4	0.88	0.70	83%	77%	J	J	1
302	1	0.97	0.65	91%	88%	J	J	2
302	2	0.95	0.8	87%	82%	J	J	1
302	3	0.93	0.84	83%	82%	J	J	2
302	4	0.89	0.79	81%	76%	J	J	1
<b>Moyenne</b>		0.93	0.74	90%	86%	-	-	24min
<b>Ecart-type</b>		0.05	0.14	6%	8%	-	-	24min

Le modèle statique dirigé est capable de fournir des résultats suffisamment précis pour les besoins de vigicrues d'un horizon de  $k+1$  à  $k+4$ , c'est-à-dire d'une demi-heure à deux heures.

Les critères de Nash sont souvent proches de 0,9 et perdent en valeur quand l'horizon de prévision augmente. Les critères de persistance sont globalement bons, éloignés de la prévision naïve, sauf généralement à  $k+1$ , ce qui est compréhensible car, comme le débit varie assez peu d'une période à la période suivante, il est difficile pour le modèle d'effectuer une prévision

meilleure que la prévision naïve. Pour l'évènement le plus intense, le n°19 de septembre 2002, la prévision devient moins efficace à deux heures mais le niveau de vigilance est toujours bien appréhendé.

La plupart du temps, le pic des prévisions est légèrement en retard sur l'observation mais le *SPPD* reste très élevé, ce qui atténue ce problème. Nous ne présenterons pas les 16 hydrogrammes ici mais nous présentons dans la figure 52 les évènements n°13 et n°19 pour un horizon de prévision de  $k+3$ . Tous les hydrogrammes sont présentés en annexe 4.

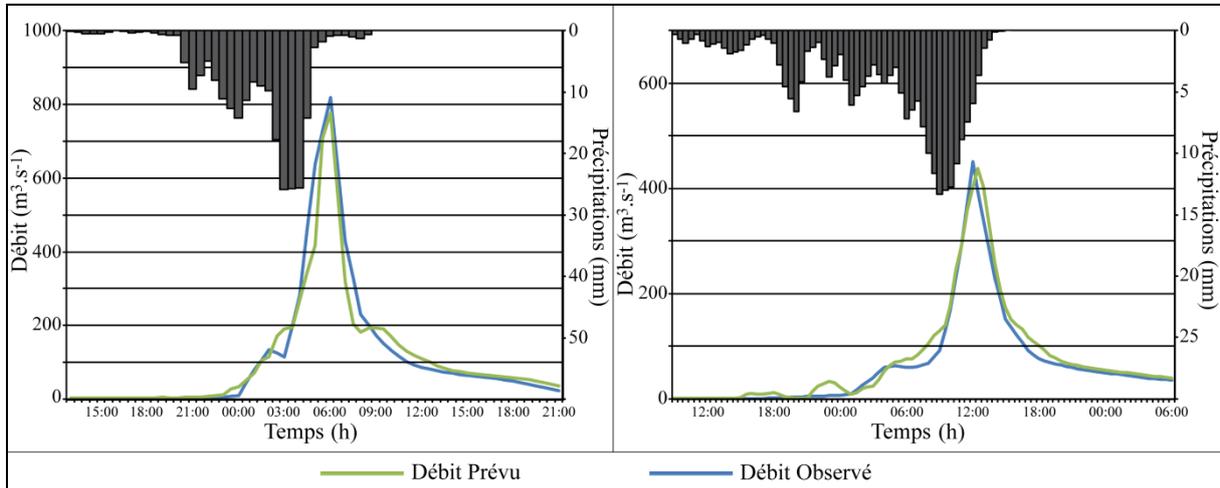


Figure 52: hydrogrammes d'observation et de prévision à  $k+3$  des évènements 19 (à gauche) et 13 (à droite) avec le modèle statique dirigé.

Nous pouvons noter la bonne qualité des prévisions même pour l'évènement de septembre 2002, qui est nettement plus intense que n'importe quel autre évènement de la base d'apprentissage (environ 50% plus intense).

Il faut aussi remarquer qu'en règle générale, ce type de modèle est connu pour utiliser beaucoup l'information des débits précédemment mesurés, ce qui en fait, en pratique, un modèle bien éloigné d'un vrai modèle pluie-débit. Ce constat est à l'origine du fait que le modèle ne simule en réalité que des incréments de débit par rapport au débit observé. Ainsi pour observer au mieux la performance de ce modèle, nous pouvons ajouter, en plus des *SPPD* les pourcentages d'anticipation des seuils de vigilance pour chaque évènement.

Le tableau 11 les présente et montre que dans la plupart des cas, les pourcentages d'anticipation des seuils de vigilance sont supérieurs à 70%. En d'autres termes, à l'horizon de prévision concerné, dans la plupart des cas, plus de 70% du débit des seuils de vigilance hydrologique est prévu. Seule une dégradation notable est observée : celle des prévisions aux horizons 1h30 et 2h pour l'évènement 19, probablement en relation avec le temps de réponse plus faible sur cet évènement. En effet, l'aval du bassin étant beaucoup plus arrosé que l'amont, la réponse du bassin est plus rapide que pour la majorité des évènements de la base d'apprentissage et induit donc un retard dans l'accroissement des débits préjudiciable à l'anticipation des seuils.

Tableau 11 : pourcentages d'anticipation des seuils de vigilance crue pour chaque évènement et chaque horizon de prévision. L'évènement n°8 étant à deux pics, on distingue l'anticipation pour chacun d'eux.

<b>Evènement 8</b>					<b>Moyenne</b>
Horizon de prévision (h)	0.5	1	1.5	2	4 horizons
jaune 1	90%	85%	76%	73%	81%
jaune 2	94%	103%	92%	83%	93%
<b>Evènement 13</b>					
Horizon de prévision (h)	0.5	1	1.5	2	4 horizons
orange	103%	95%	95%	97%	97%
<b>Evènement 19</b>					
Horizon de prévision (h)	0.5	1	1.5	2	4 horizons
orange	93%	82%	82%	66%	81%
rouge	97%	91%	69%	63%	80%
<b>Evènement 302</b>					
Horizon de prévision (h)	0.5	1	1.5	2	4 horizons
jaune	91%	96%	106%	117%	103%
Moyenne	95%	92%	87%	83%	89%

Pour compléter ces informations, la figure 53 présente les incréments de débit prévus en fonction des incréments de débit observés pour les quatre évènements testés et les quatre horizons de prévision explorés.

La grande majorité des débits prévus se situe dans les limites de +/-30% de la valeur observée. De fait, si l'on s'intéresse aux incréments de +/- 10 m<sup>3</sup>.s<sup>-1</sup>, représentant les fortes hausses ou fortes diminutions du débit, plus importantes dans le cadre de notre problématique, la moyenne du rapport entre débit prévu et débit observé est de 71%.

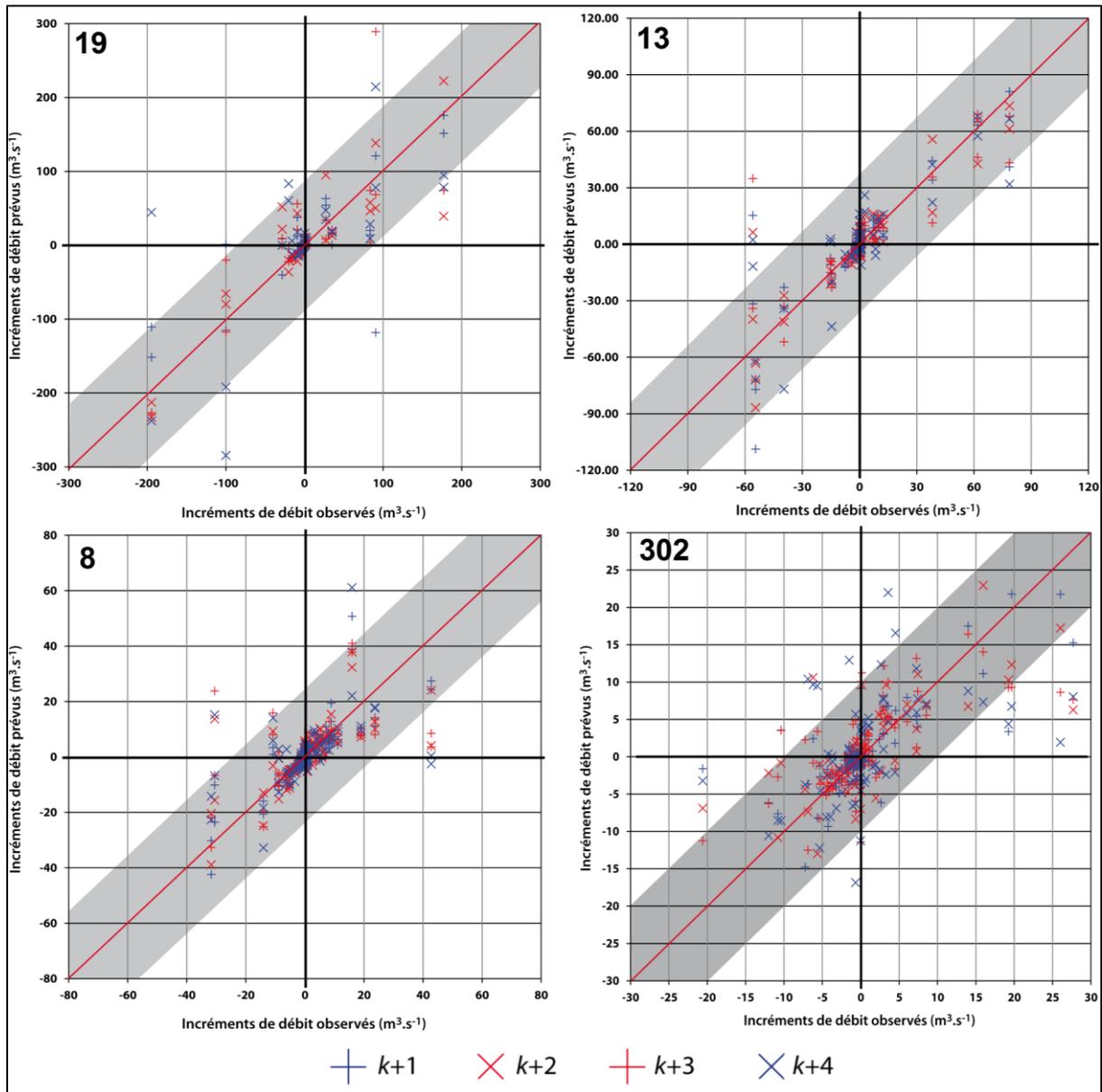


Figure 53 : représentation incréments de débit prévus en fonction des incréments de débit observés pour les quatre évènements aux quatre horizons de prévision.

En restant dans l'objectif de prévisions opérationnelles sur un bassin bien jaugé, l'utilisation de ce modèle semble tout à fait pertinente.

Par ailleurs, l'aspect des hydrogrammes, les critères de persistance et la représentation des incréments (figure 53) semblent indiquer que le modèle a su faire un usage pondéré des informations de pluie et de débit observé pour produire des prévisions généralement fiables.

#### 4.4.2 Modèle récurrent non dirigé : cas du bassin mal jaugé

Afin de pouvoir effectuer des prévisions sur un cours d'eau dont les mesures durant l'évènement ne sont pas fiables, mais qui après ré-analyse peuvent cependant être utilisées pour alimenter une base de données (cf. 1.6.) nous envisageons dans cette section la mise en œuvre d'un modèle qui n'utilise en aucun cas les débits observés précédents mais ceux estimés par le modèle. Il s'agit d'un modèle récurrent dit « non dirigé », c'est-à-dire non dirigé par les débits observés aux pas de temps précédents ou futurs, tant en apprentissage qu'en utilisation (test). Ce modèle a été

mis en œuvre en mode déterministe, sans prise en compte d'incertitudes sur les précipitations et en mode ensembliste, en essayant de prendre en compte l'incertitude sur les pluies observées, tel que décrit dans la partie 4.4.2.2.

#### 4.4.2.1 Prévisions déterministes

Les résultats de ces prévisions sont présentés dans le tableau 12, de la même manière que pour le modèle statique dirigé. Même si cela n'est pas directement visible sur les résultats, le modèle récurrent est plus instable que le modèle statique dirigé. En effet, contrairement à ce dernier, il ne dispose pas d'une information de débit observé pour rattraper une éventuelle errance. Par ailleurs, lors de la manipulation, le modèle récurrent est plus long à converger que le modèle statique. Ceci-dit, dans les deux cas, le temps de calcul reste dans des délais inférieurs à la minute.

Tableau 12: récapitulatif des résultats obtenus avec le modèle récurrent non dirigé (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).

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

Comme l'on pouvait s'y attendre, les résultats sont moins bons que pour la version statique dirigée ce qui accrédite l'utilisation, lorsque c'est possible, des débits observés comme variables d'état. Les résultats présentés ici confirment que cette information est importante dans le fonctionnement du modèle. Cependant, ces résultats montrent qu'en moyenne, le modèle récurrent non dirigé est capable de fournir des prévisions alternatives lorsqu'un problème survient avec l'obtention de données de débit (matériel endommagé, problèmes de transmission).

En effet, 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 inférieur à celui du modèle statique : le modèle récurrent représente mieux la dynamique du bassin.

Là encore, nous ne présenterons pas tous les hydrogrammes mais quatre d'entre eux : pour l'évènement 8 à  $k+1$ , l'évènement 13 à  $k+2$ , l'évènement 19 à  $k+3$  et l'évènement 302 à  $k+4$ . Les autres hydrogrammes sont présentés en annexe 5.

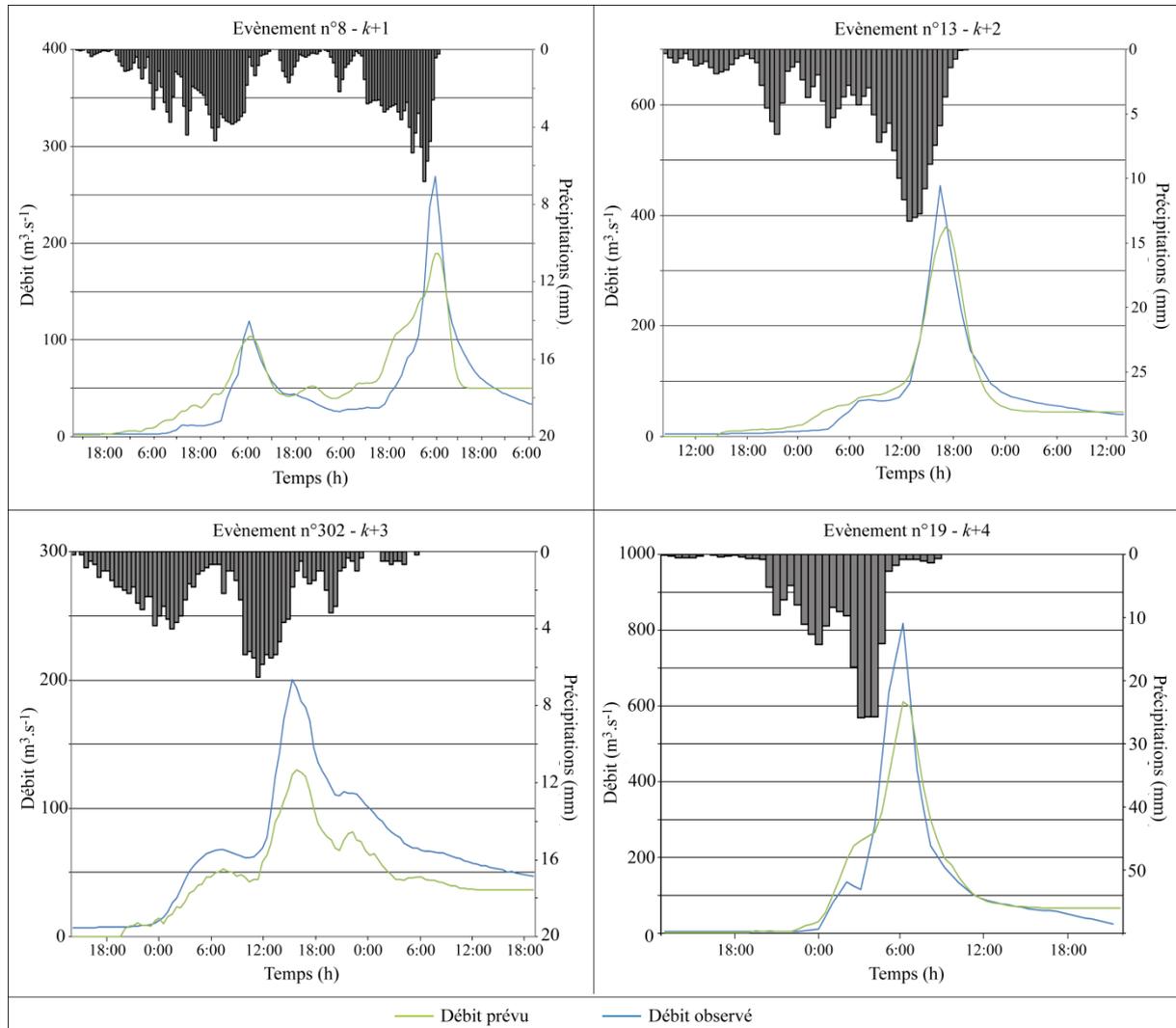


Figure 54: hydrogrammes observés et prévus par le modèle récurrent pour les évènements 8 à  $k+1$ , 13 à  $k+2$ , 302 à  $k+3$  et 19 à  $k+4$ .

#### 4.4.2.2 Prévisions ensemblistes

Compte tenu de la grande incertitude liée aux données pluviométriques soulignée au chapitre 1, nous avons souhaité évaluer la sensibilité des modèles obtenus en introduisant un facteur de hasard représentant l'incertitude dans les pluies observées. Nous avons estimé ces incertitudes à plus ou moins 20% ce qui est cohérent avec les conclusions de Marchandise (2007).

En conséquence, une modification aléatoire des pluies observées a été réalisée : chaque valeur de précipitation de l'ensemble de test a été multipliée par un nombre pris au hasard entre 0,8 et 1,2. Ce processus a été répété 20 fois pour donner 20 vecteurs de pluies différents pour chaque pluviomètre. Aucun biais vers des valeurs plus élevées ou plus faibles n'a été introduit puisque rien ne permet vraiment d'affirmer que les pluies mesurées par les pluviomètres sous-estiment ou à l'inverse surestiment les pluies réellement tombées sur le bassin.

Pour comparer les résultats obtenus par ces prévisions ensemblistes, face aux prévisions déterministes, nous avons choisi de nous baser sur le *SPPD* qui s'avère être le critère le plus représentatif du succès d'une prévision dans la problématique qui nous intéresse.

Nous avons donc présenté dans la figure 55 les valeurs maximum, minimum et moyennes atteintes par ce critère vis-à-vis des celles obtenues en mode déterministe.

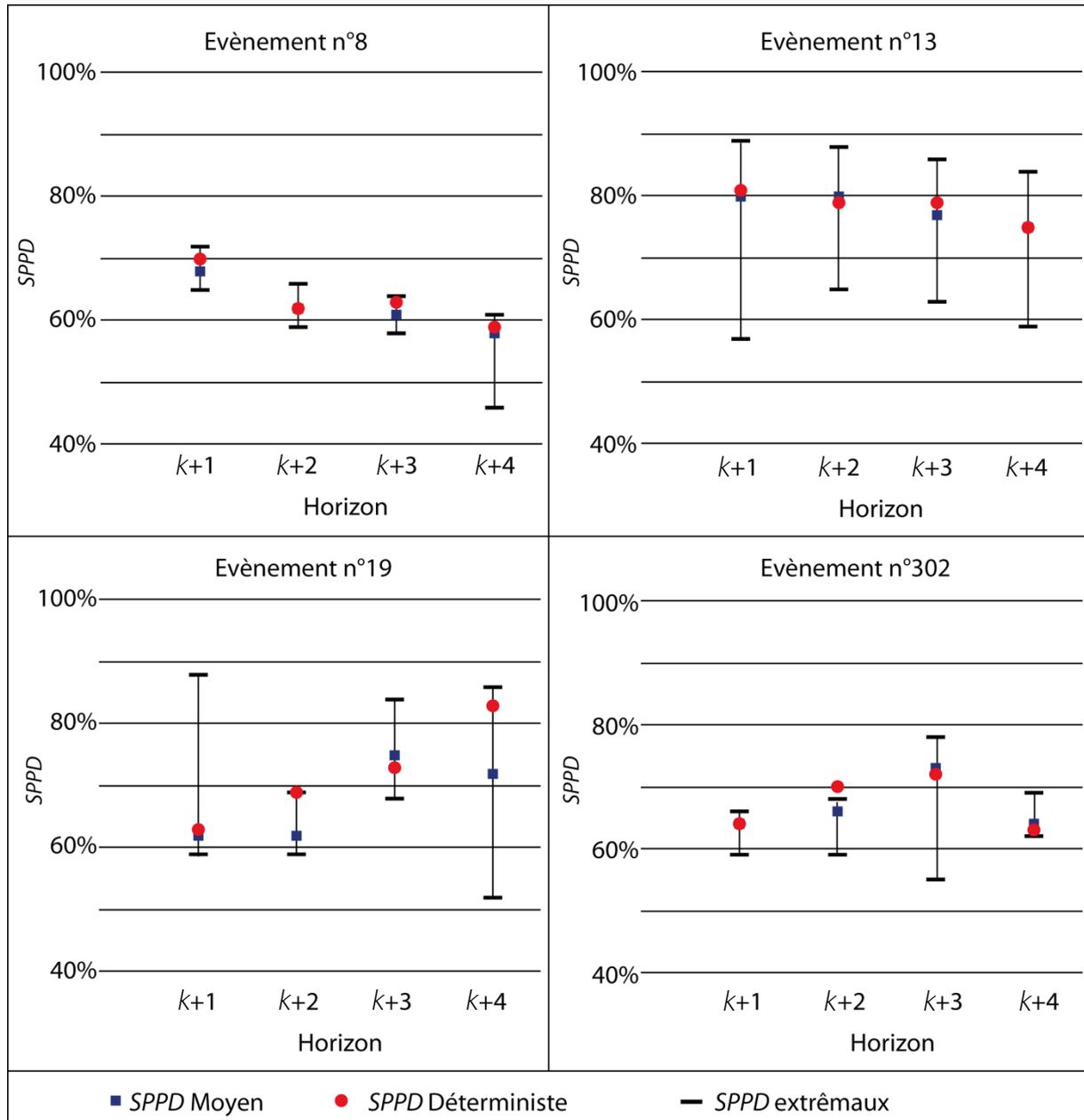


Figure 55: comparaison des résultats moyens et extrêmes du modèle ensembliste avec les résultats du modèle déterministe pour les quatre évènements à quatre horizons de prévision.

Aucune amélioration n'est obtenue en utilisant cette méthode. Dans la majorité des cas, les résultats moyens de sont pas meilleurs (et parfois pires) que ceux précédemment exposés. Dans certains cas, la valeur maximale donne d'excellents résultats mais rien ne permet au prévisionniste de choisir ce scénario plutôt qu'un autre à l'opérationnel. Il est en effet couramment admis que des prévisions ensemblistes apportent une information complémentaire utile lorsque la moyenne des scénarios montre une amélioration par rapport aux prévisions déterministes. Ce n'est pas le cas ici.

Cette étape nous apprend toutefois que le modèle n'amplifie globalement pas les erreurs présentes dans les entrées ce qui est un point très positif pour la robustesse du modèle, compte tenu de toutes les incertitudes et du bruit qui sont reliés aux mesures hydrométéorologiques.

#### 4.4.3 Utilisation de l'humidité issue de la couche racinaire du modèle SIM

Nous avons vu que, au travers du débit observé précédemment ou le cumul durant l'épisode, qui doivent fournir une information sur l'état du système, on cherche surtout à obtenir une information sur l'humidité des terrains. Cette humidité conditionne beaucoup la réaction hydrologique comme de nombreux auteurs ont pu le montrer et comme on peut le constater au regard des événements de fin d'été présents dans la base. Le couplage d'informations sur l'humidité des sols avec un réseau de neurones a déjà été pratiqué par Anctil et al. (2004) sur des bassins bien plus lents que les bassins cévenols, avec des résultats montrant que ces indices sont bien corrélés aux conditions réelles sur le bassin, conditions qui vont fortement influencer le débit observé. Le réseau de neurones peut donc avantageusement utiliser l'information.

La chaîne de modèles Safran-Isba-Modcou (Habets et al., 2008) que nous appellerons SIM délivre l'estimation d'une de ses variables d'état qui est l'indice d'humidité du sol. Ces indices ont aussi été fournis dans le cadre du projet BVNE pour le Gardon d'Anduze. La maille du modèle est de huit kilomètres sur toute la France avec une sortie toutes les 24 heures, à six heures UTC.

Ces indices présentent l'avantage de fournir une information directe sur l'état du terrain. En revanche, dans le cas des crues les plus rapides, le fait que l'humidité soit fournie toutes les 24h, à six heures UTC, représente un inconvénient majeur. En 24h, l'évolution de l'humidité peut avoir le temps de favoriser (ou non) une réaction hydrologique forte. On peut aussi faire démarrer la simulation au moment où sort la valeur d'humidité mais l'évolution durant l'épisode pourra de toute façon être plus rapide.

Nous avons testé l'utilisation de cet indice en remplacement du cumul observé dans le modèle récurrent déterministe. Les résultats sont mis en parallèle avec ceux du modèle récurrent déterministe utilisant le cumul des précipitations à la manière de ceux de la section 1.4.2 dans la figure 56.

L'impact de l'utilisation de cet indice n'est pas net. Sur les événements n°302 et 13, l'amélioration est claire et parfois importante. Sur les événements n°19 et 8, on n'observe généralement pas d'amélioration. On peut penser que l'évènement n°19, bref et intense, est moins bien appréhendé avec SIM du fait du pas de temps du modèle et de la chronologie de l'évènement (pic de crue peu de temps après l'actualisation de l'indice SIM) mais cela n'explique rien pour l'évènement n°8. Ce dernier aurait dû être mieux géré puisqu'avec un double pic, l'état d'humidité entre les deux pics a forcément de l'importance, pourtant il n'en a rien été.

L'utilisation de cet indice apporte parfois une amélioration mais rend le modèle hydrologique ainsi élaboré dépendant d'un autre modèle hydrométéorologique. Il est difficile d'évaluer l'incertitude liée à SIM et donc difficile de savoir quel est son impact sur notre modèle.

Nous avons essayé de fournir au modèle à la fois SIM et le cumul sur l'épisode. Les résultats ne se sont pas améliorés la plupart du temps ; ils se sont même parfois détériorés. L'introduction de la double information est forcément moins parcimonieuse, voire contradictoire, et a pu nuire au bon fonctionnement du modèle : le cumul peut augmenter alors que SIM reste stable pendant 48 pas de temps.

Pour pallier le problème du pas de temps de sortie du champ d'humidité de SIM, nous avons envisagé d'en faire une désagrégation temporelle tenant compte des pluies observées (pour qu'elle ne nécessite pas la prochaine valeur de SIM) mais, en quelque sorte, cela revenait à modéliser un modèle, ce qui nous éloigne de la réalité du terrain et surtout, qui peut être long et complexe. Nous nous serions ainsi détachés de notre objectif de parcimonie et de simplicité, indispensable pour un travail à l'opérationnel.

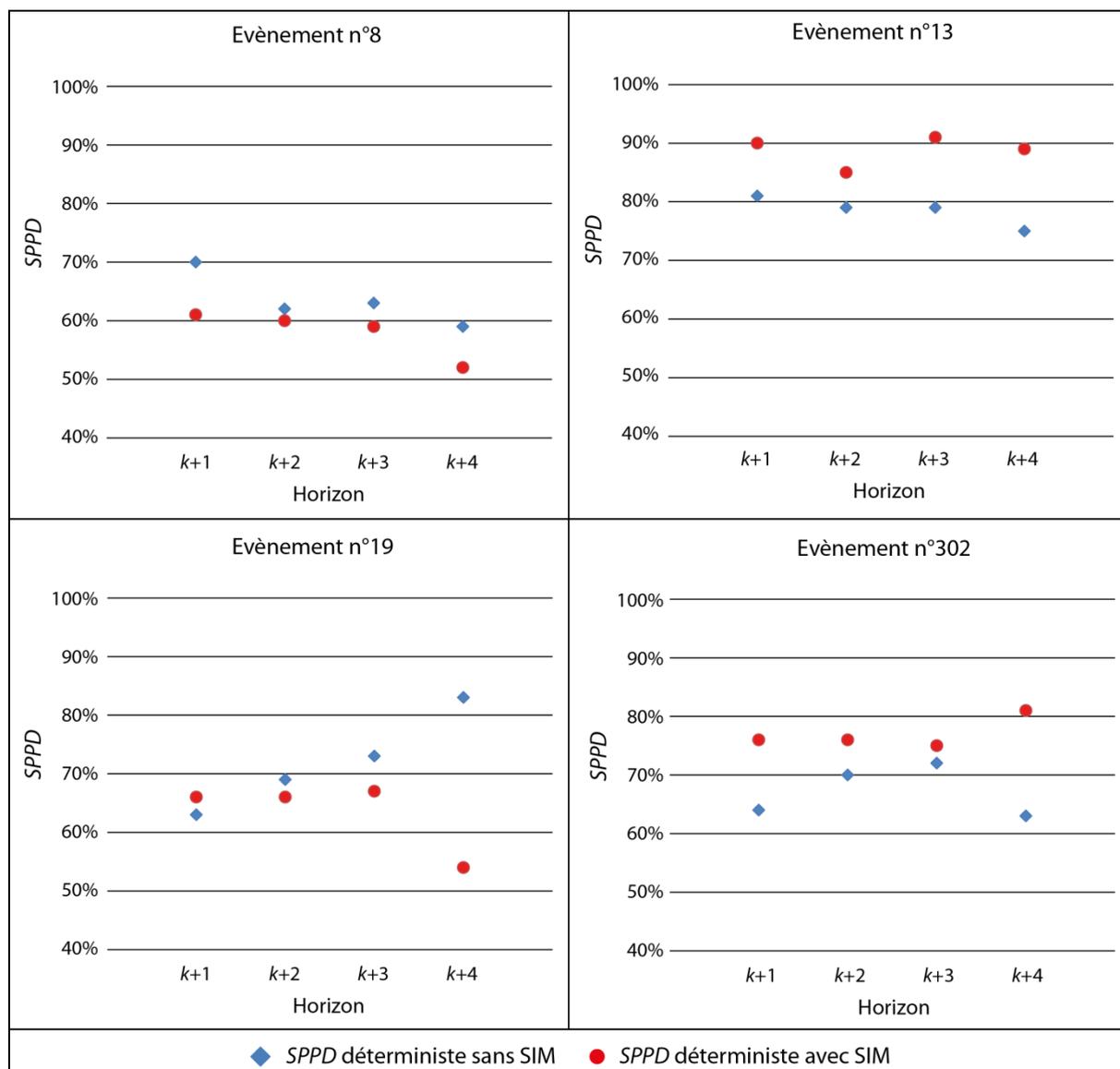


Figure 56: comparaison des résultats du modèle récurrent déterministe non dirigé sans ou avec SIM.

## 4.5 Conclusions

Ce chapitre s'est attaché à présenter la mise en œuvre de l'apprentissage statistique pour la prévision des crues du Gardon de Mialet à Mialet jusqu'à un horizon de prévision de deux heures et en l'absence de toute information sur les pluies futures. Pour cela une première base de données a été extraite et présentée.

Les spécificités de la conception d'un modèle neuronal de prévision ont ensuite été évoquées puis étudiées sur la base de données précédemment présentée. Trois architectures ont tout

d'abord été présentées et mises en œuvre : l'architecture purement linéaire, purement non linéaire (utilisant un perceptron multicouche) et enfin l'architecture juxtaposant dans un même réseau les modèles linéaires et non linéaires. La dernière architecture ayant fournis les meilleurs résultats, trois types de modèles ont à nouveau été développés à partir de celle-ci; les deux premiers sont qualifiés de déterministes puisqu'ils n'intègrent pas de bruit ajouté. Tous deux effectuent les prévisions de débits avec des architectures différentes, l'une non récurrente qui utilise en prévision les mesures de débit obtenues en temps réel (contexte bien jaugé), l'autre qui est utile en mode dégradé, lorsque les mesures de débit ne sont plus disponibles, soit par défaut de l'instrumentation, soit par défaut de la télétransmission (bassin mal jaugé). Nous avons pu montrer dans ce chapitre que ces deux types de modèle à réseaux de neurones (statique et récurrent) étaient tout à fait aptes à soutenir une utilisation opérationnelle dans un contexte de crues rapides. Ce propos est bien vérifié pour un contexte bien jaugé, dans lequel il est possible d'utiliser les données de débit observées précédemment (cas du modèle statique ou dirigé). Dans un contexte de bassin mal jaugé (cas du modèle récurrent ou non-dirigé), il convient de le nuancer puisque les résultats sont de moins bonne qualité. Toutefois, ces prévisions en mode dégradé, peu sensibles aux incertitudes de la mesure pluviométrique, peuvent constituer un solide apport à la prévision opérationnelle dans la mesure où elles sont stables ; il est en particulier remarquable que les prévisions effectuées avec le modèle récurrent se détériorent nettement moins que celles du modèle statique lorsque l'horizon de prévision augmente. Le modèle récurrent est donc plus à même de représenter la dynamique du processus de génération des crues rapides, on peut imaginer qu'il serait une très bonne solution si l'on disposait de prévisions fiables de précipitations.

Nous avons également pu présenter que les améliorations recherchées par le biais de la prévision non déterministe, ou ensembliste, n'ont pas apporté de progrès, ni de détérioration, dans la performance en configuration mal jaugée. Elles ont néanmoins de ce fait pu démontrer la robustesse du modèle neuronal face aux incertitudes.

En revanche, l'utilisation de l'humidité des sols issue de SIM apporte quelques améliorations, même si la mise à jour de cette information est réalisée à une période sans relation avec la dynamique des phénomènes étudiés ici. Si la donnée d'humidité de SIM est disponible, il est donc possible de l'utiliser pour la réalisation de cette prévision en mode dégradé.

Ces résultats extrêmement positifs, en particulier pour la prévision des crues avec le modèle récurrent (ou non dirigé) vont nous permettre d'envisager de transposer les prédicteurs ainsi obtenus à des bassins versants pour lesquels on ne dispose pas du tout de mesure de débit (non jaugés). Il peut paraître étrange, voire vain, de vouloir appliquer une méthode d'apprentissage statistique à des bassins sans mesure aucune, mais nous allons voir dans la suite comment une méthode de généralisation adaptée peut permettre d'envisager la réalisation de cet objectif.



## **5. Généralisation aux bassins non jaugés**

Les résultats obtenus en prévision sur le Gardon à Mialet avec le modèle récurrent sont ici généralisés à des bassins voisins de la bordure cévenole.

Dans un premier temps, nous présenterons les grandes lignes de la démarche envisagée qui consiste en un procédé de correction des sorties en fonction des caractéristiques du bassin visé. Il s'agit d'appliquer le modèle récurrent déterministe à une sélection de bassins et d'analyser les erreurs pour mettre au point une méthode de correction efficace.

Cette sélection de bassins sera présentée tout comme les évènements qui ont été sélectionnés sur ceux-ci. Nous verrons qu'une intéressante variété de bassins et d'épisodes de la bordure cévenole est disponible pour effectuer les premiers essais de cette méthode, depuis le très petit bassin du Lyonnais de quatre kilomètres carrés près d'Alès, aux bassins plus vastes du Chassezac ou du Gardon d'Anduze de plus de 500 km<sup>2</sup>.

La mise en œuvre de la méthode ainsi élaborée sera présentée. Nous verrons que toutes les corrections nécessaires ne peuvent être appliquées à l'aide des simples caractéristiques du bassin visé, mais qu'une bonne partie du chemin vers une correction efficace des débits prévus en fonction du bassin ciblé peut tout de même être parcourue.

Enfin, une application de la méthode aux bassins sélectionnés sera présentée. Les résultats seront détaillés, bassin par bassin, en mettant en perspective leur exploitabilité dans un contexte opérationnel ainsi que les pistes de recherche qu'ils ouvrent pour leur amélioration.

Nous tirerons des conclusions plus générales sur le niveau d'efficacité, de fiabilité et de robustesse de la méthode employée qui fournit des résultats remarquables sur la qualité de la dynamique des réponses. Seront définis les points forts, les points à améliorer, en particulier pour les petits bassins et la question des temps de réponse.

## 5.1 Démarche proposée

A l'aide d'une base de données suffisamment étoffée nous avons pu réaliser des prévisions relativement efficaces dans un contexte bien jaugé (modèle statique dirigé à Mialet) et dans un contexte mal jaugé (modèle récurrent non dirigé à Mialet). Conformément aux objectifs définis dans le chapitre 1, il nous faut désormais trouver un moyen de réaliser des prévisions dans un contexte non jaugé, c'est-à-dire sans donnée de débit, ni pour l'apprentissage, ni pour la validation (où le débit peut être utilisé comme information d'état).

Nous avons songé à créer un bassin « normalisé » qui représente un bassin de caractéristiques moyennes pour la bordure cévenole et ce à partir d'un bassin pilote, celui du Gardon à Mialet. Un apprentissage aurait été effectué sur le bassin normalisé, avec des variables normalisées, et il aurait suffi de « dénormaliser » les variables de sortie que l'on obtienne le débit prévu sur un bassin dont les caractéristiques entraînent dans le domaine de validité de la méthode. Une telle normalisation était assez complexe, notamment au niveau temporel (nécessité de l'introduction de pas de temps variables), et, dans le cas particulier de cette étude où un seul bassin pilote était utilisé, pouvait être considérée comme superflue.

Le Gardon à Mialet, suivant sa présentation réalisée dans le chapitre 4, nous semble représentatif d'un bassin cévenol. C'est, avec l'abondance des données disponibles, la raison qui nous pousse à le choisir dans cette démarche comme *bassin pilote*, c'est à dire le bassin que nous choisissons pour réaliser l'apprentissage et qui sert de base pour la généralisation des prévisions à des bassins non-jaugés, bassins que nous appellerons des *bassins cibles*.

La méthode retenue procède à partir du bassin pilote : à l'issue de l'apprentissage effectué sur ce bassin, on fait directement fonctionner le modèle récurrent calé en validation sur un bassin cible.

Dans la plupart des cas, les résultats ne sont pas satisfaisants, ce à quoi l'on pouvait s'attendre. En effet, les bassins pilote et cible sont généralement trop différents pour que le modèle représente correctement leurs fonctionnements respectifs. Quoi qu'il en soit, les résultats obtenus sont analysés, en particulier les erreurs temporelles et d'amplitude, dont la réduction est une question centrale dans la prévision des crues. Andréassian et al. (2010) nous enseignent en effet que l'analyse des erreurs et des échecs, du moins apparents, peuvent contribuer à faire avancer nombre de problématiques.

L'opération est répétée sur tous les bassins cibles visés par la généralisation. On essaie alors de trouver une relation empirique entre les erreurs (temporelles et d'amplitude) et certaines des caractéristiques simples du bassin versant (topographie notamment). Ces relations doivent permettre de corriger les sorties du modèle prédictif de telle sorte qu'elles correspondent mieux aux sorties observées du bassin cible. Ce processus est décrit dans la figure 57.

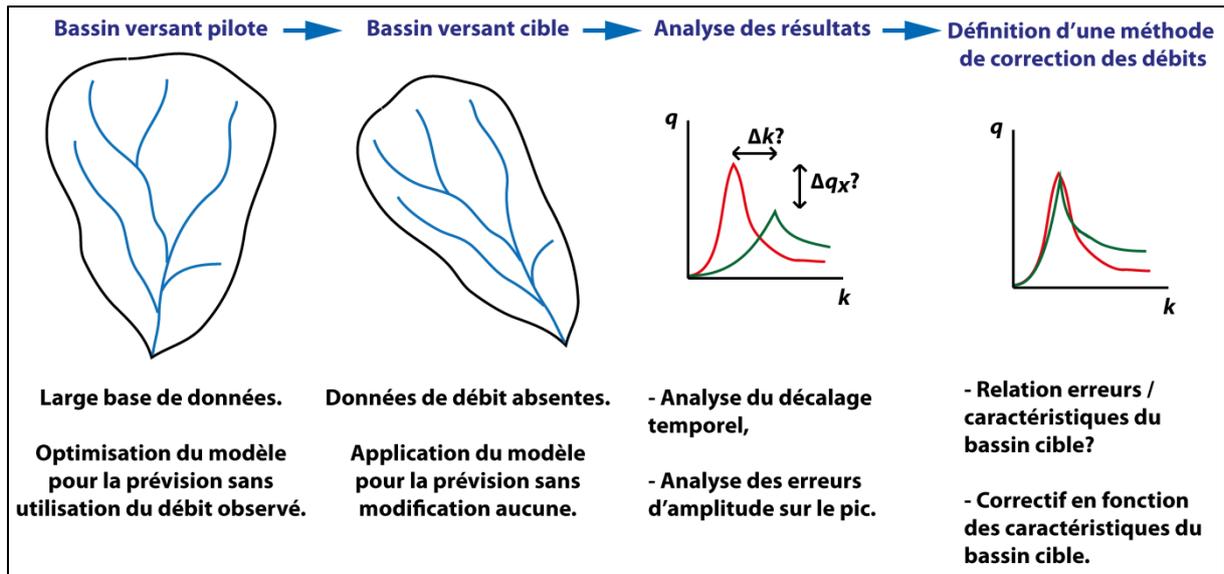


Figure 57: schématisation du raisonnement conduisant à la généralisation aux bassins non jaugés.

Le modèle est en quelque sorte régionalisé puisque ses paramètres sont considérés comme communs à de multiples bassins cibles d'une même région. Ceci étant dit, les paramètres ne sont pas les seuls éléments du modèle à rester identiques : la dimension des vecteurs de variables de forçage est elle aussi commune à tous les bassins, de même que la complexité du modèle. C'est donc uniquement un post-traitement qui va permettre d'adapter les résultats du modèle à leur cible et à ses caractéristiques propres qui, elles-mêmes, pilotent la dynamique hydrologique.

## 5.2 Méthodologie développée

### 5.2.1 Rappel des principales contraintes liées à la problématique

Les contraintes imposées par la pratique de l'hydrologie opérationnelle sont les principales mais s'y ajoutent celles spécifiques aux bassins non jaugés. Il faudra tenir compte de l'ensemble de ces contraintes pour proposer des résultats utilisables en conditions opérationnelles. Notre cahier des charges est donc le suivant :

- Le pas de temps : la période d'échantillonnage est choisie d'une demi-heure, correspondant aux contraintes opérationnelles des agents des services de prévision des crues, et comme dans les travaux précédemment exposés dans ce mémoire. Cependant, pour recalibrer temporellement les sorties, il pourra être nécessaire de procéder à des variations du pas de temps. Cet aspect sera développé dans ce qui suit.
- Les variables d'entrée : comme dans le chapitre précédent, compte tenu du fait que nous ne disposons pas de prévision des précipitations aux échelles de temps et d'espace considérées, nous n'utiliserons que des valeurs de pluie observées jusqu'à l'instant de prévision.
- La défaillance d'un ou plusieurs pluviomètres sur le bassin : il est nécessaire de pouvoir pallier ce problème. La donnée de pluie est la seule information de terrain disponible dans le cas des bassins non-jaugés. Son absence est donc particulièrement préjudiciable et doit être solutionnée au mieux.
- Les chroniques de débit du bassin pilote : il est impossible de pratiquer un apprentissage statistique sur des bassins non jaugés. Le bassin pilote, sur lequel l'apprentissage est réalisé, ne peut fournir les données nécessaires au calage des paramètres pour le bassin cible. Cela implique aussi, par exemple, que les corrélations croisées qui ont permis

d'estimer les temps de réponse du bassin pilote (chapitre 4) pour chaque pluviomètre ne seront plus disponibles ; les longueurs des fenêtres temporelles ne seront donc pas non plus ajustées de façon optimale.

- Les données de débit en temps réel du bassin cible : elles ne sont pas non plus disponibles dans un bassin non-jaugé, évidemment. Nous avons vu dans le chapitre 4 qu'il était possible de fournir des prévisions relativement fiables sans utiliser de données de débit en temps réel du bassin de calage. Nous aurons l'occasion d'observer sur plusieurs bassins en quoi l'absence de ces données affecte les prévisions.
- La robustesse : les prévisions doivent être testées sur plusieurs bassins avec autant d'évènements que possible pour permettre d'évaluer la robustesse de la méthode développée dans ces travaux. Idéalement, avec un nombre suffisant de bassins et d'évènements, il faudrait définir un domaine de validité pour la généralisation (limites de la méthode employée).
- La parcimonie : elle permet, comme nous l'avons vu, de réduire la complexité du modèle et par là de limiter l'impact des erreurs de mesure sur son fonctionnement (dilemme biais-variance).

Ce cahier des charges posé, nous allons détailler les étapes de la méthode qui a été utilisée pour mener à bien la généralisation aux bassins non-jaugés et qui a été synthétiquement présentée dans la section 5.1.

### 5.2.2 Analyse des données et des prévisions brutes : vers une méthode de correction des débits prévus

La première étape ayant été réalisée dans le chapitre 4, nous allons, dans un premier temps, appliquer le modèle récurrent déterministe existant à une série de bassins cibles sur lesquels on dispose, respectivement, d'un ou plusieurs évènements. L'échantillon de bassins et d'évènements sera l'échantillon de validation de la méthode.

#### *5.2.2.1 Sélection et aperçu des bassins cibles*

Ces bassins ont été sélectionnés en fonction de leur localisation, dans la zone d'étude définie dans le chapitre 1, mais aussi en fonction de l'existence de données de débit observé ou estimé par retour d'expérience. En effet, si l'on s'interdit d'utiliser les débits observés ou reconstitués pendant le fonctionnement du modèle, ils sont en revanche indispensables pour évaluer les performances du modèle et, éventuellement, crédibiliser son utilisation opérationnelle sur des bassins réellement non jaugés. Ont été éliminés les bassins qui comportent un barrage écrêteur de crues (Cèze et Gardon d'Alès). Cependant, à n'en pas douter, au regard des résultats du chapitre 4, l'impact d'un barrage passif pourrait être pris en charge par l'apprentissage d'un réseau de neurones, dans un contexte d'abondance des données. Dans le cadre d'une généralisation effectuée par un modèle ne comportant pas de bassin avec barrage dans sa base d'apprentissage, la difficulté serait trop importante. La numérotation et les principales caractéristiques des bassins sont présentés en annexe 3.

Les bassins pour lesquels nous avons pu réunir les informations sont donc au nombre de 14 et sont localisés dans la figure 58.

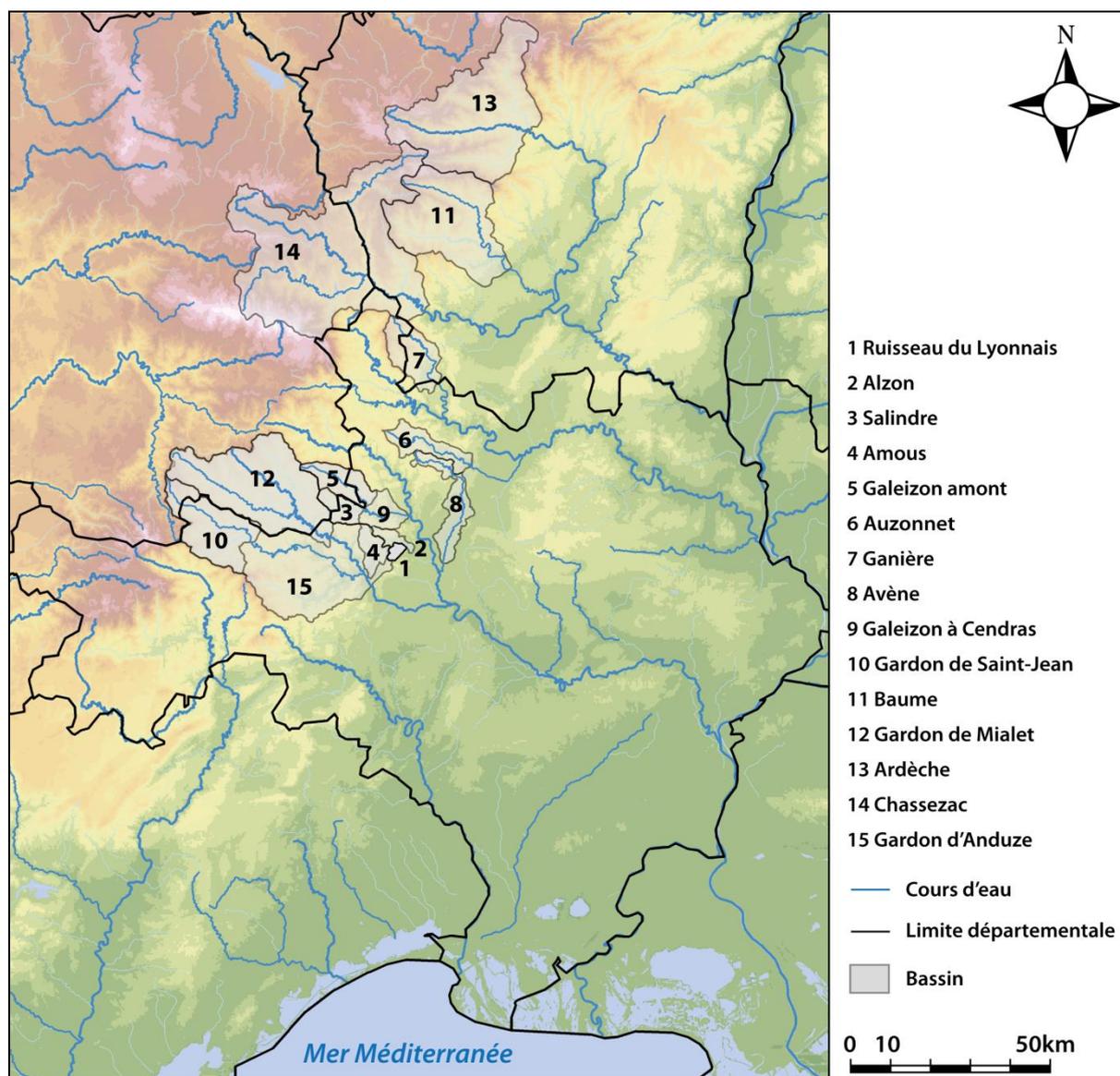


Figure 58: localisation des bassins cibles ainsi que du bassin du Gardon de Mialet à Mialet, le bassin pilote (numéros par ordre de surface croissante).

Les bassins cibles présentent des caractéristiques assez variées sur le plan topographique (dimensions, pentes, dénivelées...). Pour les caractériser, nous avons pu déterminer les valeurs présentées dans le tableau 13 : surface, pente du drain, longueur du drain, dénivelée du drain, pente des versants ou encore dénivelée du bassin. Cependant, leurs caractéristiques plus générales restent assez semblables en termes d'occupation des sols, de géologie (pour la plupart), de régime hydrologique ou encore de climatologie ; c'est également en fonction de ces propriétés communes qu'ils ont été choisis. Dans la suite, chaque présentation des résultats en validation sur un bassin cible sera précédée d'une description de celui-ci.

Tableau 13: principales caractéristiques topographiques des bassins sélectionnés (les statistiques en bas de tableau sont calculées, Gardon de Mialet exclu, seulement présenté ici comme référence ; les cellules en rouge représentent les valeurs minimales et les vertes les valeurs maximales).

Numéro attribué	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	729	36%	881
11	Baume	210	3,80%	30	1132	35%	1352
12	Gardon à Mialet	220	3,00%	29	864	33%	993
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	923	32%	1084
<b>Moyenne</b>		141	3,30%	22.8	657	27,93%	803
<b>Médiane</b>		56	3,20%	18.5	649	29,50%	731
<b>Maximum</b>		545	6,30%	54.7	1263	36,00%	1505
<b>Minimum</b>		4	2,30%	2	56	12,00%	203
<b>Ecart-type</b>		180	1,09%	17.9	362	6,84%	373

### 5.2.2.2 Sélection et description des événements retenus

Avant toute chose, il faut noter que parmi les événements conservés, une sur-représentation de l'évènement extrême de septembre 2002 apparait clairement. De fait, sur les 35 événements de la base de données cible, 11 d'entre eux concernent cet épisode. Etant donné les cumuls et les débits spécifiques extrêmes observés durant cet événement assez exceptionnel, il est nécessaire de signaler ce biais avant de livrer des statistiques et analyses sur la base de données.

Notons qu'un événement se définit comme une période durant laquelle une pluie a pu être reliée à un débit spécifique de plus de  $1 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  sur un bassin donné. La même période peut donc correspondre à plusieurs événements puisqu'il y a pu avoir plusieurs bassins touchés, comme c'est le cas, par exemple, pour l'évènement de septembre 2002.

Le tableau 14 présente le nombre d'évènements récoltés par bassin cible ainsi que la source et le type de données de débit obtenues.

Tableau 14: nombre d'épisodes par bassin, type de données et source de ces données (Rex 2002 : retour d'expérience de l'épisode de septembre 2002 ; SIEE / Syndicat du Galeizon : données issues des simulations de l'épisode de septembre 2002 par SIEE commandées par le syndicat du Galeizon ; OHMCV : données issues de la base de données Sevnol de l'Observatoire Hydrométéorologique Cévennes Vivarais ; BVNE : données fournies dans le cadre du projet BVNE).

Numéro attribué	Bassin cible	Nombre d'épisodes	Type	Source
1	Ruisseau du Lyonnais	1	Estimation	Rex 2002
2	Alzon	1	Estimation	Rex 2002
3	Salandre	1	Simulation et estimation	SIEE / syndicat du Galeizon
4	Amous	1	Estimation	Rex 2002
5	Auzonnet	3	Hydrogramme	OHMCV
6	Galeizon amont	1	Simulation et estimation	SIEE / syndicat du Galeizon
7	Ganière	11	Hydrogramme	OHMCV
8	Avène	1	Estimation	Rex 2002
9	Galeizon	1	Simulation et estimation	SIEE / syndicat du Galeizon
10	Gardon à Saumane	3	Hydrogramme	BVNE
11	Baume	3	Hydrogramme	OHMCV
13	Ardèche	3	Hydrogramme	OHMCV
14	Chassezac	3	Hydrogramme	OHMCV
15	Gardon à Anduze	2	Hydrogramme	BVNE

Comme dans le chapitre 4, on considère qu'un évènement débute à la première goutte de pluie qui touche l'un des pluviomètres et se termine avec la fin de la réaction hydrologique significative. En revanche, les épisodes ont été sélectionnés sur le débit afin de ne tester que les évènements significatifs en généralisation. De plus, une grande partie d'entre eux est issue de données limitées dans le temps, qui n'existent que parce qu'un épisode hydrologique important est survenu (c'est le cas des estimations *a posteriori*, qui ne sont généralement pas pratiquées pour des épisodes moyens).

L'histogramme de la figure 59 présente la répartition des débits spécifiques en fonction de trois classes. On constate sans surprise qu'une très grande majorité (près de 70%) d'évènements possède un débit spécifique inférieur à  $3 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  alors que les débits spécifiques élevés et extrêmes sont nettement moins représentés.

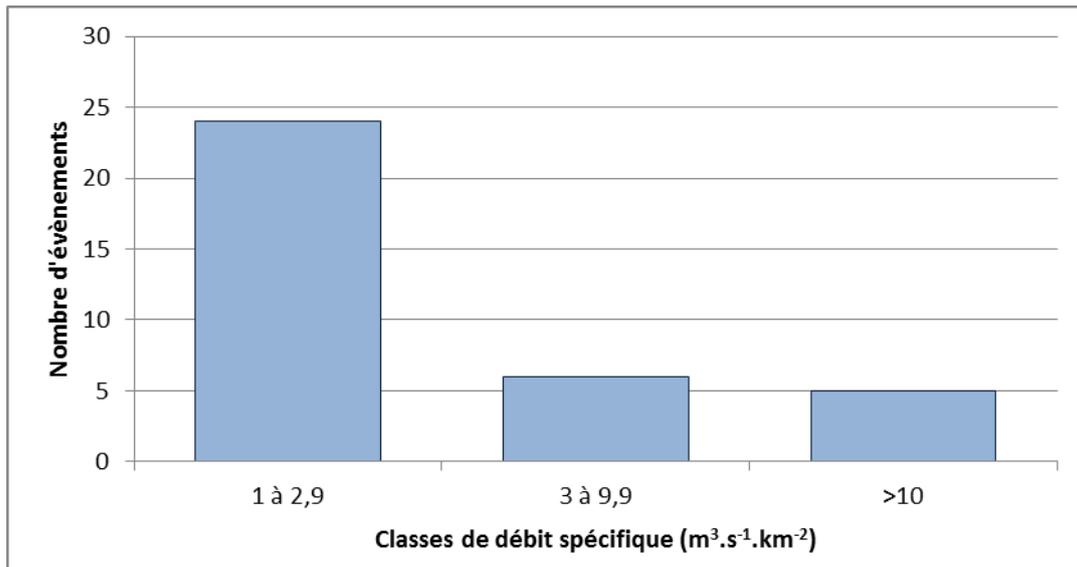


Figure 59: nombre d'évènements par classes de débits spécifiques sur un total de 35 évènements.

Ces valeurs de débit spécifique s'échelonnent de 1,01 à 30  $m^3.s^{-1}.km^{-2}$ . Notons que la totalité des épisodes des deux classes supérieures se sont produits en septembre 2002.

Concernant les cumuls moyens de précipitations observés, ils s'échelonnent de 84 à 529mm. La figure 60 les représente par classes de 100mm.

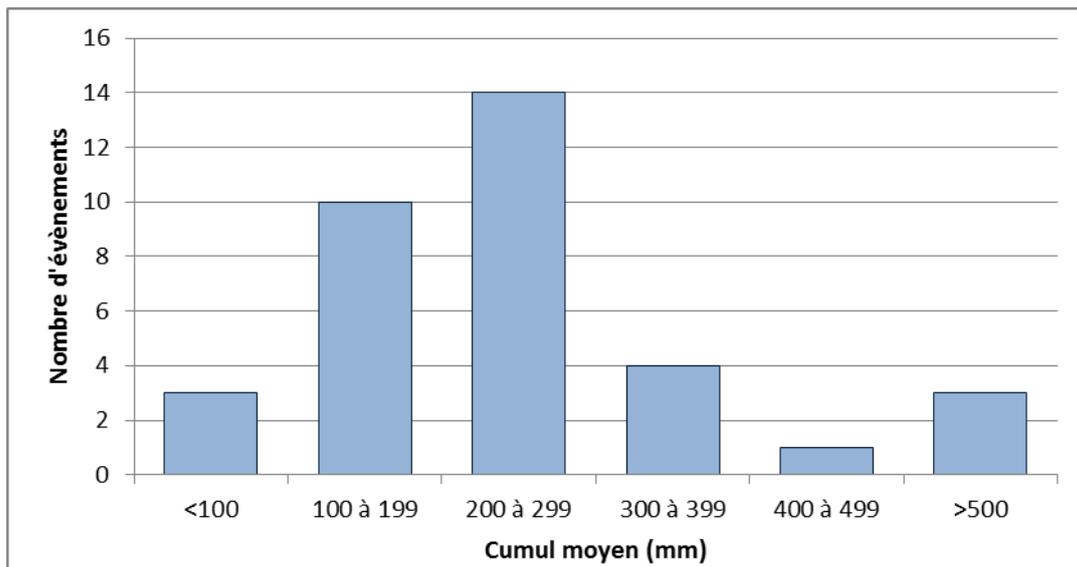


Figure 60: nombre d'épisodes par classes de cumul moyen.

La totalité des valeurs supérieures à 400mm sont issues de l'épisode de septembre 2002. Une grande majorité d'évènements a donné un cumul pluviométrique moyen de 100 à 300mm.

Les épisodes durent de 27 à 162 heures, en moyenne 76 heures avec un écart-type important de 42 heures. La grande majorité des valeurs courtes (inférieures à la moyenne) sont issues de l'épisode de septembre 2002.

Pour synthétiser les informations apportées sur cette base de données, malgré une forte représentation d'un seul épisode marquant et très intense sur une grande partie des bassins cibles sélectionnés, une intéressante variété d'évènements est disponible et va pouvoir être exploitée pour évaluer la méthode de généralisation développée ci-après.

### *5.2.2.3 De l'importance de la dimension temporelle et des amplitudes*

Le modèle récurrent déterministe sans estimation de l'humidité des sols développé dans le chapitre 4 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 la section 5.2.2.1. 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.

Concernant tout d'abord le décalage temporel observé en prévision entre le pic observé et le pic estimé, nous avons calculé sa moyenne par bassin versant cible, pour les quatre horizons de prévisions, au besoin sur l'ensemble des événements testés, et avons essayé de le mettre en relation avec les caractéristiques des bassins. Rappelons que, comme dans le chapitre 4, un décalage temporel positif signifie que le modèle est en retard par rapport à l'observation.

Dans toutes les figures qui vont suivre, en nuage de points, les marqueurs rouges représentent les bassins cibles alors que le marqueur bleu place le bassin pilote du Gardon à Mialet comme référence à 0 de décalage temporel théorique.

En premier lieu, le décalage a été représenté en fonction de la surface (figure 61). Si pour certains bassins, on peut trouver une logique dans les erreurs réalisées (grands bassins : modèle en avance ; petits bassins : modèle en retard), il n'en reste pas moins que certains décalages sont difficilement explicables et que l'ensemble ne permet pas de dégager une relation claire entre le décalage produit par le modèle et la surface du bassin.

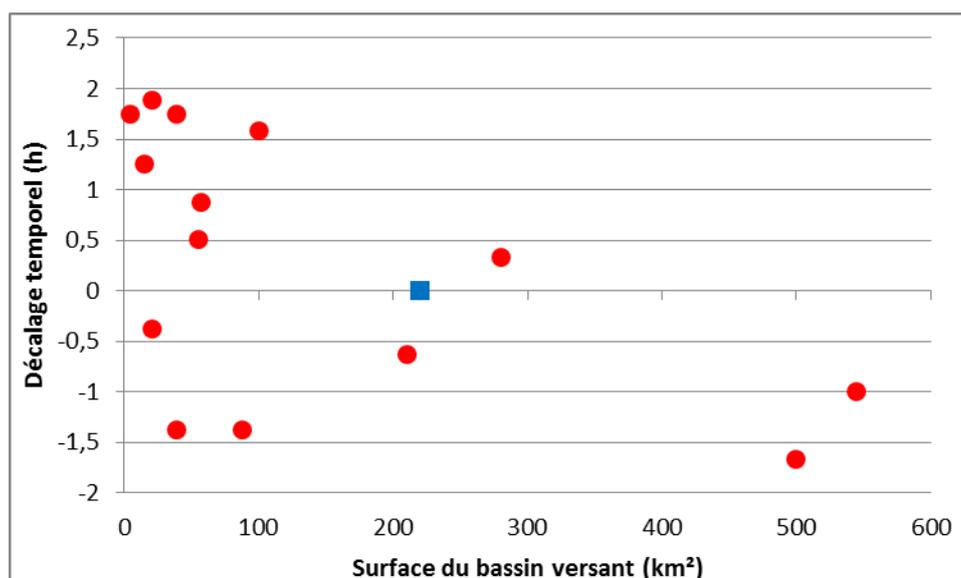


Figure 61: décalage temporel moyen en fonction de la surface des bassins considérés. Le marqueur bleu symbolise le décalage théorique du bassin pilote.

Le décalage a ensuite été représenté en fonction de la pente du drain (figure 62). On y note des valeurs de décalage temporel très hétéroclites pour des pentes relativement proches les unes des autres.

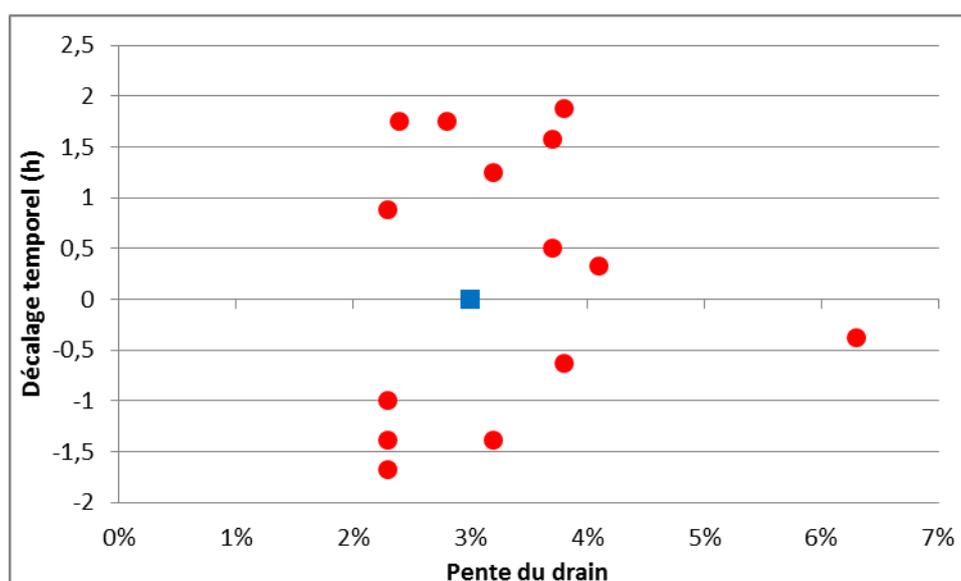


Figure 62: décalage temporel moyen en fonction de la pente du drain des bassins considérés. Le marqueur bleu symbolise le décalage théorique du bassin pilote.

Le décalage temporel est également présenté en fonction de la longueur du drain (figure 63). Une tendance à une avance du modèle pour les drains les plus longs et à un retard de celui-ci pour les drains les plus courts semble se dessiner, sans pour autant qu'une relation puisse être définie clairement.

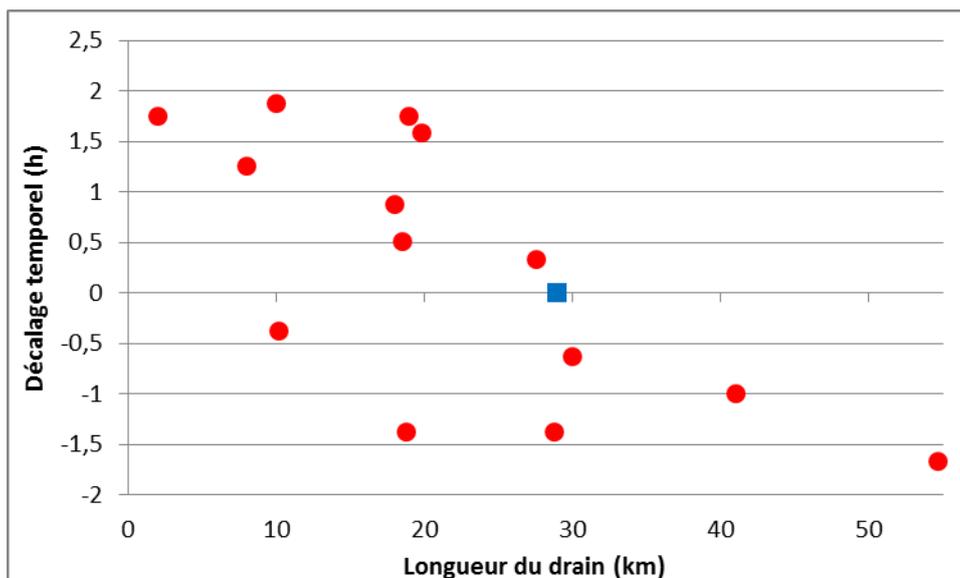


Figure 63: décalage temporel moyen en fonction de la longueur du drain du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote.

Avec la dénivelée du drain, le décalage temporel reste à peu près dans la même configuration qu'avec la longueur du drain (figure 64). Une légère tendance se dessine, mais il n'est pas possible d'en tirer une relation claire.

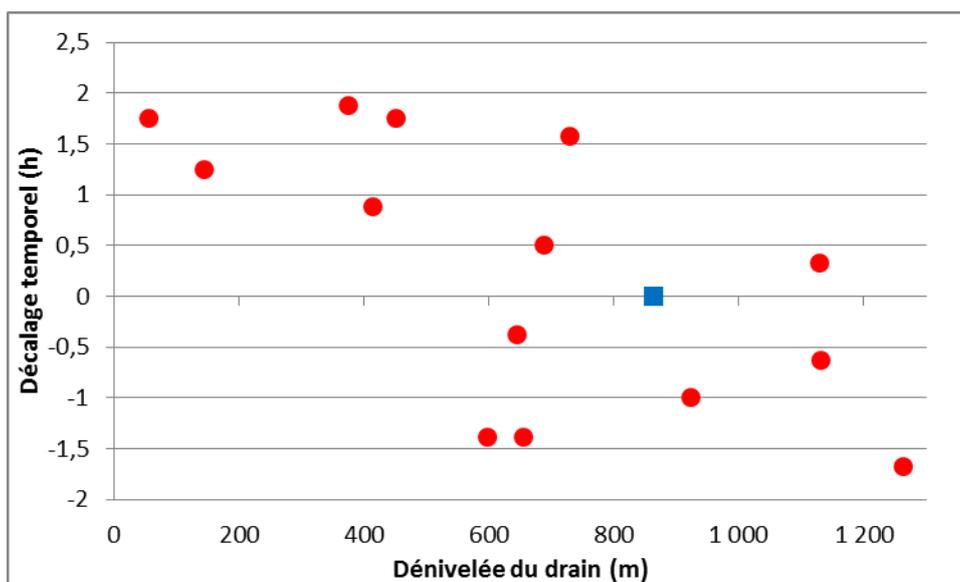


Figure 64: décalage temporel en fonction de la longueur du drain du bassin considéré.

Nous poursuivons avec la représentation du décalage temporel en fonction de la pente des versants (figure 65). Là encore, aucune tendance claire ne se dessine.

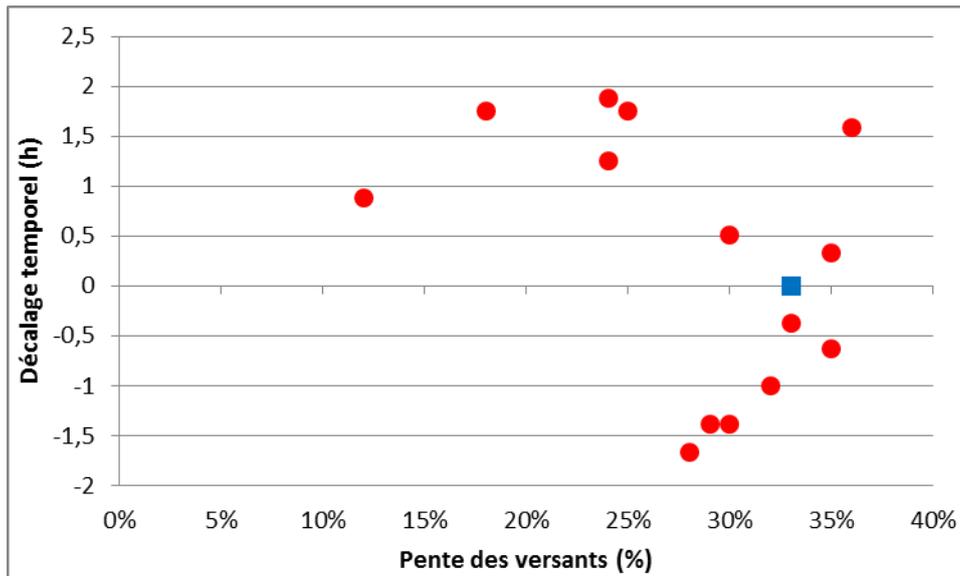


Figure 65: décalage temporel en fonction de la pente des versants du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote.

En dernier lieu, le décalage temporel moyen a été représenté en fonction de la dénivelée du bassin versant (figure 66). Dans ce cas aussi, une légère tendance se dessine mais il apparaît toujours impossible de définir une relation claire et fiable. Par exemple, les valeurs situées aux alentours de 700 mètres de dénivelée conduisent à des décalages séparés de trois heures.

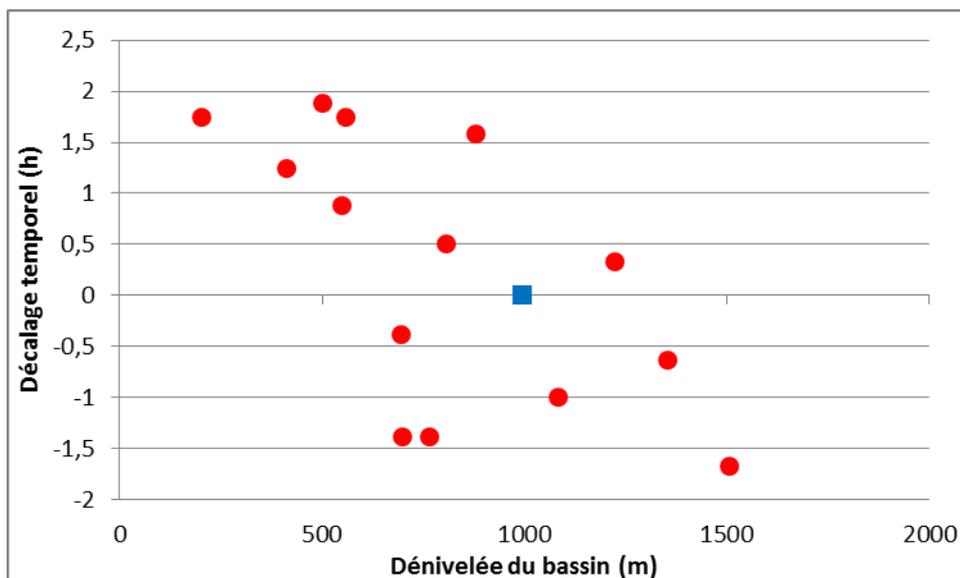


Figure 66: décalage temporel en fonction de la dénivelée du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote.

On voit donc qu'il n'est pas possible de corriger directement les décalages temporels à l'aide des caractéristiques topographiques générales définies. Dans un premier temps, nous avons l'idée d'utiliser une telle relation pour mettre en place une variation du pas de temps en fonction du temps de réponse du bassin (contraction ou dilatation temporelle). Bien que séduisante, cette idée ne peut pas être mise en œuvre avec ces bassins et ces évènements.

Par ailleurs, si les corrélations croisées effectuées pour le bassin du Gardon à Mialet nous ont amené à définir des intervalles d'investigation cohérents pour l'optimisation par validation croisée, elles ont aussi fourni un certain nombre de valeurs extrêmes qui laissent penser que le

temps de réponse d'un bassin donné est loin d'être une constante et que de nombreux facteurs peuvent l'influencer (intensités et succession des averses, état d'humidité initial, évolution et vitesse d'évolution de cet état, topographie et morphologie du bassin et du drain, occupation des sols, géologie, pédologie...). De ce fait, la sur-représentation d'un épisode particulier (septembre 2002) entraîne une distorsion qui limite toute analyse plus poussée.

Ainsi, nous ne pouvons déterminer une relation empirique qui serait probablement basée sur de très nombreuses variables avec une population statistique (nombre de bassins cibles et d'évènements) aussi réduite.

Quoi qu'il en soit, si l'on s'intéresse de façon plus détaillée à ces décalages temporels, on constate que la majorité d'entre eux sont faibles et que seules quelques valeurs sur des évènements spécifiques sont élevées. Ce biais ne pourra pas être corrigé dans ces travaux mais il reste relativement peu important et nous allons pouvoir examiner si l'importance de ces décalages peut être atténuée par de bonnes corrections de l'amplitude, qui pourraient conduire à des *SPPD* de bonne qualité.

Nous avons donc, de la même manière, moyenné les amplitudes sur les quatre horizons de prévision par bassin et nous les avons représentées en fonction des différentes caractéristiques du bassin cible. La représentation de la figure 67 montre le *PPD* (*percentage of the peak discharge* défini dans la section 2.2.1.4) moyen obtenu avec le modèle « tel quel » en fonction de la surface du bassin cible.

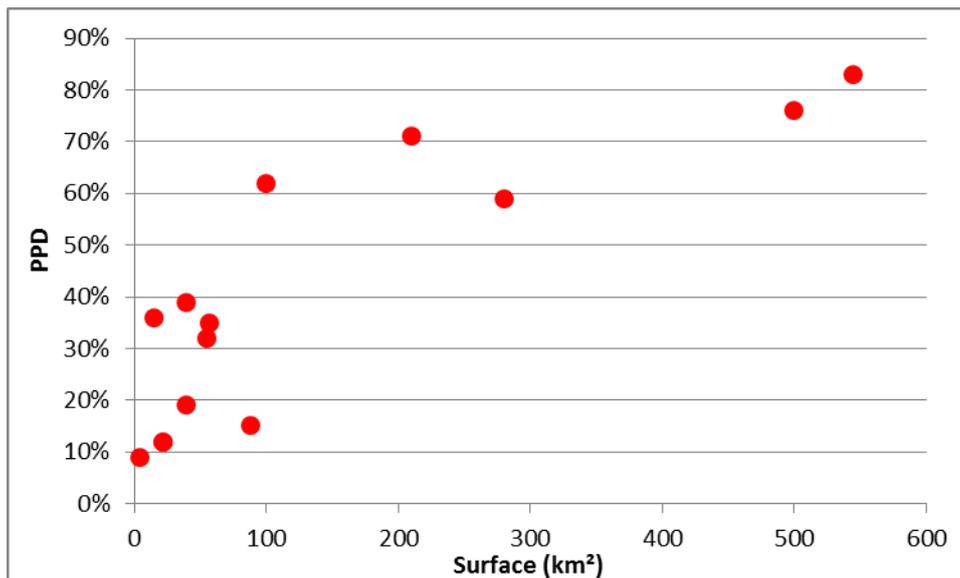


Figure 67: PPD moyen en fonction de la surface du bassin considéré.

On distingue une relation assez claire (coefficient de corrélation linéaire de 0,69) entre la surface du bassin cible et le *PPD* du modèle déterministe calé sur le bassin pilote.

Clairement, le *PPD* est d'autant plus important que le bassin est grand. De même, la relation se définit de manière intéressante entre les autres caractéristiques dimensionnelles du bassin qui sont généralement corrélées les unes aux autres et à la surface (longueur du drain : coefficient de corrélation linéaire de 0,77 ; dénivelée du drain : coefficient de corrélation linéaire de 0,74 ; dénivelée du bassin : coefficient de corrélation linéaire de 0,80).

Nous ne représenterons pas ici tous les diagrammes et allons-nous intéresser de façon plus pointue à la surface des bassins cibles. C'est en effet une caractéristique qui a été beaucoup étudiée pour la prédétermination parcimonieuse des débits extrêmes.

#### *5.2.2.4 Synthèse*

Nous avons vu qu'il était très difficile de définir une relation claire entre les décalages temporels produits par le modèle et les principales caractéristiques topographiques du bassin considéré. Nous laisserons ce problème en suspens, sa résolution étant compromise en l'état des moyens que nous utilisons.

En revanche, il est beaucoup plus aisé de trouver une relation entre le *PPD* et d'autres éléments physiques du bassin. Nous nous attarderons préférentiellement sur la surface puisque l'hydrologie a déjà fait beaucoup en ce sens et qu'une relation nette apparaît entre les erreurs d'estimation du *PPD* et la superficie du bassin.

Ce dernier point nous mène à développer le raisonnement présenté dans la section suivante.

### 5.2.3 Méthode de correction des débits prévus

Nous pouvons facilement comprendre que le débit spécifique maximum atteignable à Mialet est supérieur à  $3.75 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ , valeur maximale entre 1992 et 2008, atteinte lors de l'épisode de septembre 2002.

Pour s'en convaincre, il convient de regarder les débits spécifiques extrêmes atteints sur des bassins voisins : presque  $5 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  à Anduze en 2002 ( $545 \text{ km}^2$ ) et probablement un peu plus en 1958, plus de  $5,3 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  en 1958 à Alès ( $315 \text{ km}^2$ ) ou encore plus de  $6 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  à Ners en septembre 2002 ( $1100 \text{ km}^2$ ). Tous ces bassins ont connu des débits spécifiques maximaux sur de longues périodes (>50ans) supérieurs à celui septembre 2002 sur le Gardon de Mialet alors que ces bassins sont de 1,5 à 5 fois plus grands.

Sachant que les débits spécifiques de pointe tendent à diminuer avec l'augmentation de la surface du bassin, il paraît peu probable que la valeur maximale sur 1992-2008 à Mialet soit réellement extrême. De plus, des repères de crue anciens (1900 à Saint-Croix-Vallée-Française notamment) existent en amont de Mialet et témoignent de crues autrement plus importantes que celle de septembre 2002, même en tenant compte de la propagation vers l'aval.

Pour essayer de définir quel pourrait être le débit spécifique maximal théorique du Gardon à Mialet, il existe des méthodes d'évaluation à partir de critères simples, qui ne nécessitent pas de faire intervenir des séries pluviométriques.

#### *5.2.3.1 Formules empiriques de prédétermination des débits spécifiques de pointe*

Réménieras (1986) note que l'absence d'observations hydrométriques de longue durée a conduit à l'établissement de très nombreuses formules pour la prédétermination des débits de pointe probables du bassin, parfois en différentes périodes de retour.

Nombre d'entre elles n'utilisent que l'aire du bassin comme variable. Les paramètres sont des coefficients qui ont généralement une portée régionale. Pardé (2004) souligne d'ailleurs l'inexistence d'une formule universelle, la variabilité des territoires et des climats étant beaucoup trop importante.

Comme Rodier (1964) entre autres, Réménieras (1986) note que les formules de prédétermination du débit sont régionales et sont de la forme :

$$Q=C.A^\alpha$$

Où  $Q$  est le débit de pointe pour une fréquence donnée (en  $m^3.s^{-1}$ ),  $A$  la superficie du bassin ( $km^2$ ),  $C$  est un coefficient appelé « cote de Myer » dépendant des caractéristiques du versant (la formule dérive de formules de type Myer) et notamment de la pente moyenne des versants (Réméniéras, 1986) et  $\alpha$  est en exposant, inférieur à 1, généralement entre 0,4 et 0,8. On remarque que si l'on veut obtenir la dimension d'un débit spécifique, le coefficient  $\alpha$  devra être négatif.

Pour ce dernier exposant, la valeur est souvent considérée proche de  $\frac{1}{2}$ , conduisant à placer l'aire du bassin en racine en première approximation. Pardé (2004) note que d'un point à l'autre d'un vaste bassin versant, on aura peine à considérer que les valeurs de  $\alpha$  et de  $C$  sont stables. Ceci dit, pour des bassins voisins, de climats très proches et de caractéristiques semblables, il est possible de considérer que ces paramètres restent identiques. Tricart et al. (1960) nous apprennent d'ailleurs que les coefficients mis en jeu sont régionaux, dépendent notamment de la pluviosité et ne sont pas indépendants entre eux.

Concernant le débit spécifique, des formules ont aussi été créées, notamment par le Service Hydrographique Italien à partir d'assez longues chroniques de débits (Réméniéras, 1986) :

Pour des bassins de moins de  $1000km^2$ , la formule de Scimemi (*in* Réméniéras, 1986) s'écrit :

$$q = \frac{600}{(A+10)} + 1$$

Où  $q$  est le débit spécifique de la crue maximum ( $m^3.s^{-1}.km^{-2}$ ) et  $A$ , l'aire du bassin ( $km^2$ ).

Pour des bassins de 20 à  $1000km^2$ , avec les mêmes notations, Pagliaro (*in* Réméniéras, 1986) établit :

$$q = \frac{2900}{(A+90)}$$

Forti (*in* Réméniéras, 1986) fait intervenir deux formules pour les bassins montagneux de moins de  $1000km^2$  :

- Pour des pluies de l'ordre de 400mm/24h au maximum :

$$q = 3,25 \cdot \frac{500}{(A+125)} + 0,5$$

- Pour des pluies de l'ordre de 200mm/24h au maximum :

$$q = 2,35 \cdot \frac{500}{(A+125)} + 0,5$$

Naturellement, chaque région hydrologique doit avoir son jeu de formules pour la prédétermination de différents critères.

Dans le cas qui nous intéresse, c'est-à-dire les crues importantes sur la bordure cévenole, en pays languedocien, la détermination des paramètres imposerait l'examen de longues chroniques d'observation, dont nous ne disposons pas. Il est possible d'en faire une détermination

statistique en utilisant des débits spécifiques de pointe ayant été observés dans la région. En effet, les caractéristiques languedociennes sont assez homogènes dans le sens de l'étirement de la plaine et des chaînes du sud du Massif-Central : climat, occupation des sols et géologie sont assez semblables. Il est donc pertinent de considérer qu'en première approximation, un seul jeu de paramètres pourra correspondre à la configuration des versants méditerranéens de la chaîne.

### 5.2.3.2 Etablissement d'une règle de prédétermination des débits spécifiques importants pour les Cévennes

Pour établir une règle de prédétermination des débits spécifiques maximum cévenols, nous avons utilisé 35 débits spécifiques de pointe observés en Languedoc. Ces débits, issus des crues de référence de vigicrues, de la base de données BVNE, du retour d'expérience de l'évènement de septembre 2002 ou de novembre 1999, correspondent tous à des épisodes marquants sur le plan pluviométrique comme hydrologique. Ils ont été représentés en fonction de la surface des bassins sur lesquels ils ont été observés ou estimés (figure 68).

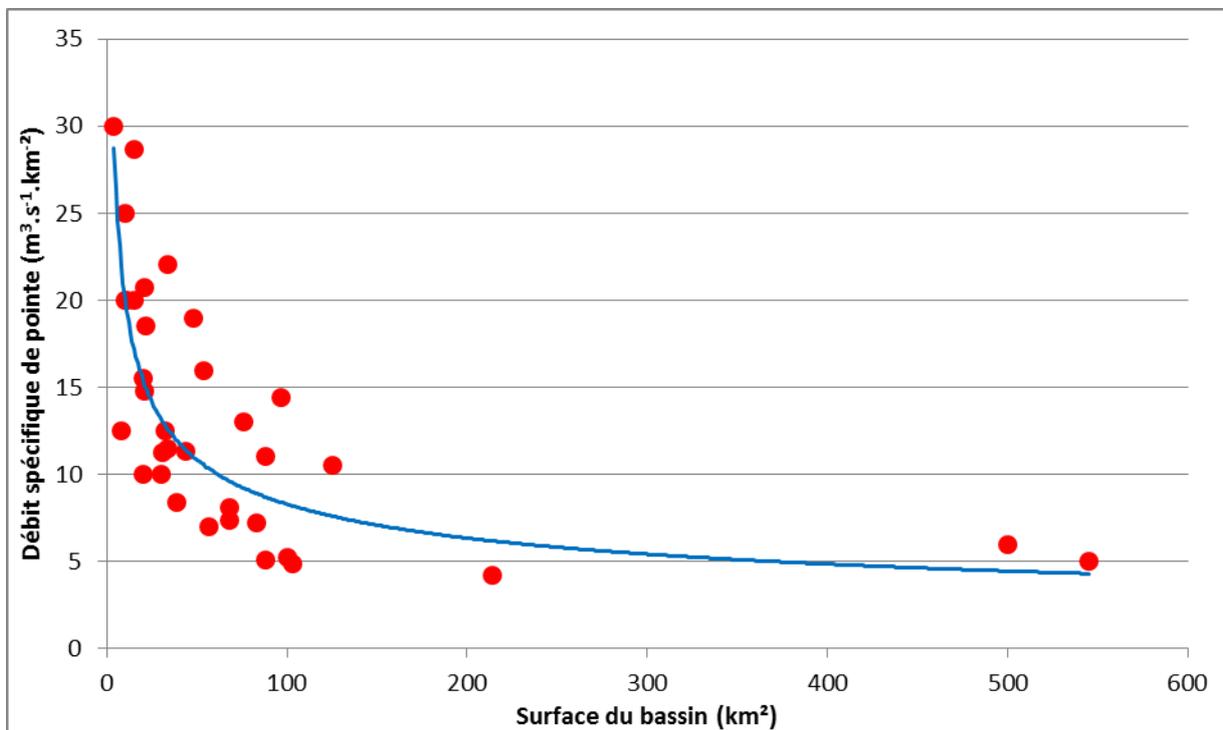


Figure 68: représentation du débit spécifique de pointe en fonction de la surface des bassins sur lesquels il a été observé (points rouges) et courbe de régression de type puissance (courbe bleue).

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.

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.

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. Ainsi, à l'aide de cette formule empirique et de la nouvelle valeur établie, nous allons pouvoir définir un coefficient correctif pour les débits sortants du modèle.

### 5.2.3.3 Etablissement d'une règle de détermination de coefficient correctif des débits prévus

Pour établir ce coefficient, nous devons rendre adimensionnelle la relation établie dans la section précédente. Pour ce faire, 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 :

$$b = 1,6331.S^{-0,386}$$

Où  $b$  est la valeur du coefficient correctif pour un rapport de la surface du bassin cible à la surface du bassin pilote  $S$  donnée. Cette relation est tracée dans la figure 69.

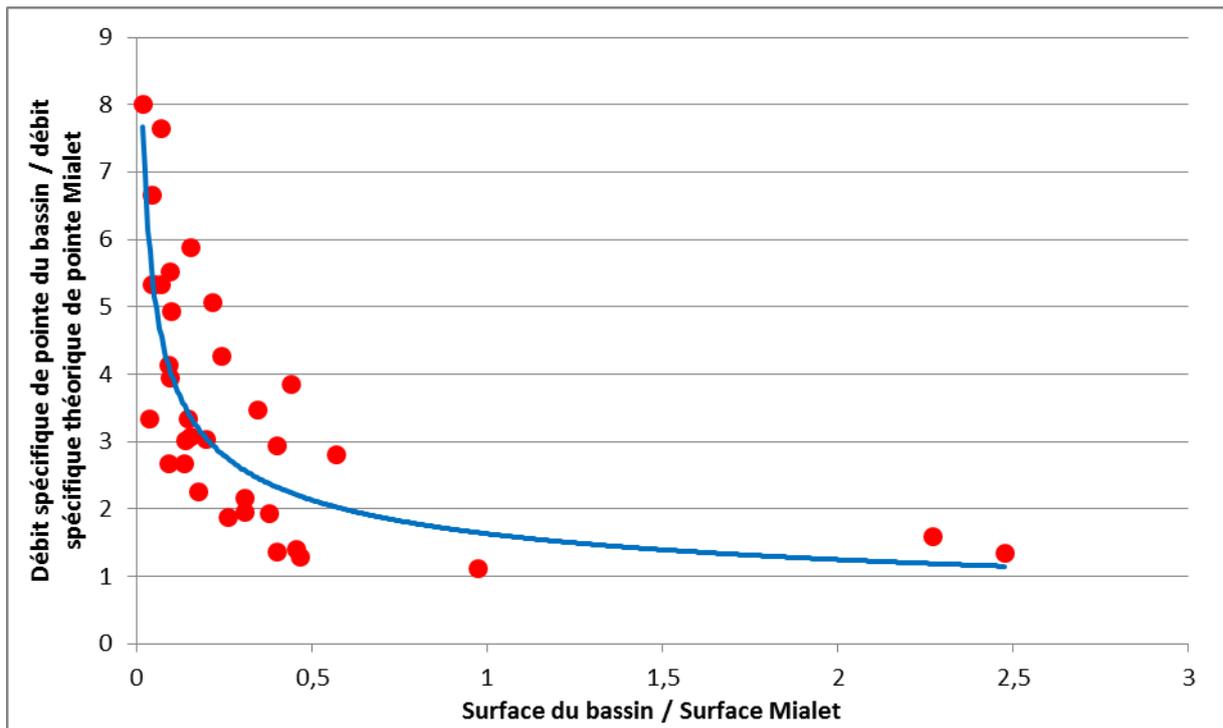


Figure 69: représentation des rapports des débits spécifiques de pointe et des surfaces des bassins aux valeurs respectives du bassin de Mialet.

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  $b$  du débit sortant du modèle ayant réalisé son apprentissage sur le bassin pilote.

Ainsi, pour les bassins visés, les coefficients correctifs sont fournis dans le tableau 15.

Tableau 15: coefficient correctif correspondant à chaque bassin visé par l'étude.

Bassin	Surface (km <sup>2</sup> )	<i>b</i>
Ruisseau du Lyonnais	4	7,67
Alzon	15	4,60
Salandre	21	4,04
Amous	21	4,04
Galeizon amont	39	3,18
Auzonnet	49	2,92
Ganière	55	2,79
Avène	57	2,75
Galeizon	88	2,33
Gardon à Saumane	100	2,21
Baume	210	1,66
Ardèche	280	1,49
Chassezac	500	1,19
Gardon à Anduze	545	1,15

Pour prendre en compte la correction envisagée à la section 5.1, il suffira d'appliquer ces coefficients correctifs aux débits spécifiques prévus par le modèle sans autre modification.

Dans les faits, on constate que la méthode revient à introduire une composante physique pour compenser l'incapacité, bien naturelle, du modèle à réseau de neurones à :

- à s'adapter à un bassin de surface différente de celle pour laquelle il a été conçu (implicitement, le modèle fonctionne pour le bassin pilote),
- à extrapoler à des débits spécifiques nettement plus importants que ceux présents dans la base d'apprentissage du bassin pilote.

Il s'agit en définitive d'unir la puissance d'identification d'un processus global apportée par l'apprentissage statistique à la force de réalisme apportée par la prise en compte d'éléments tangibles du bassin.

Nous allons appliquer la méthode ainsi élaborée aux bassins cibles dans la section suivante.

### **5.3 Application aux bassins cibles**

Pour chaque bassin, tous les événements 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 dans le chapitre 4, 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 en annexe 6. 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 entrées comme s'il provenait de trois pluviomètres différents ayant donné la même valeur.

Les résultats sont présentés bassin par bassin, par ordre croissant de surface.

### 5.3.1 Le ruisseau du Lyonnais

#### 5.3.1.1 *Présentation rapide du bassin*

Le bassin, de très petite taille (tout juste 4km<sup>2</sup>) présente une forme légèrement allongée dans une direction sud-ouest / nord-ouest peu commune dans la région. C'est dans cette direction, le long de la faille des Cévennes, que coule le drain principal qui est un cours d'eau intermittent : le ruisseau du Lyonnais. Ce dernier se jette dans l'Alzon à Saint-Jean-du-Pin, lui-même tributaire du Gardon d'Alès. La figure 70 représente ce bassin.

Le bassin est assez peu pentu (18%) ce qui s'explique assez aisément par la présence d'un petit plateau au centre du bassin (La Plaine, près du hameau de Blanas). Le drain l'est un peu plus, relativement à la moyenne des bassins cible (2.8%) mais reste moins pentu que les cours d'eau cévenols plus amont.

L'occupation des sols est presque exclusivement naturelle (forêts de chênes verts, landes et quelques terrains agricoles) puisque moins de 5% de la surface du bassin est couverte par des terrains artificiels. D'un point de vue géologique, les terrains marno-calcaires du secondaire (du lias au jurassique supérieur) sont omniprésents dans un contexte faillé de direction sub-cévenole. Quelques terrains karstiques d'étendue et de puissance limitées sont présents au sud-est.

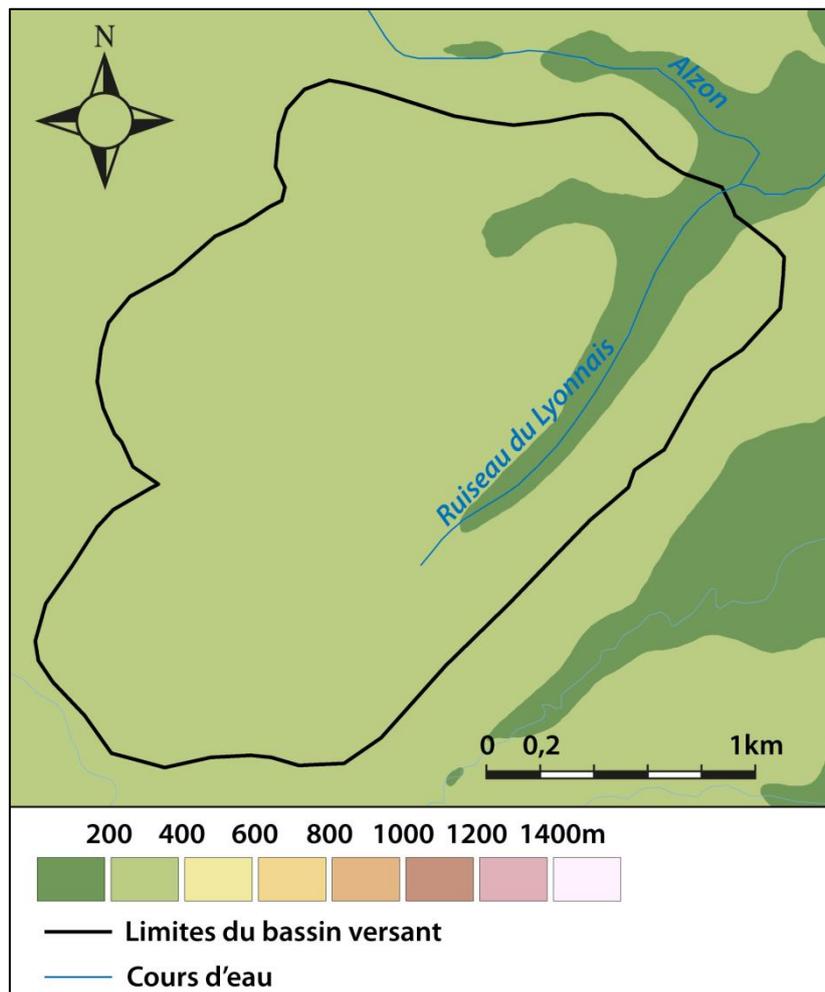


Figure 70: carte du bassin versant du ruisseau du Lyonnais.

### 5.3.1.2 Evènements et résultats

Un seul évènement est disponible sur le ruisseau du Lyonnais. Il s'agit de l'évènement extrême des 8 et 9 septembre 2002. Seul le débit de pointe évalué lors du retour d'expérience de l'épisode est disponible de sorte qu'aucun hydrogramme ne peut être tracé : une évaluation à  $120 \text{ m}^3 \cdot \text{s}^{-1}$  (soit  $30 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ) vers 6h30 du matin le 9 septembre est disponible. Elle est basée sur des témoignages et des observations de terrain. La valeur avancée est forcément entachée d'une forte incertitude qu'il est difficile d'évaluer de façon fiable.

Le cumul de pluie pour cet épisode de est de 524mm avec des intensités horaires et semi-horaires fortes (jusqu'à  $53 \text{ mm}/0.5 \text{ h}$  et  $90 \text{ mm}/1 \text{ h}$ ). Ces valeurs sont le résultat d'une estimation radar à l'exutoire qui est utilisée pour les trois entrées, étant donnée l'absence de pluviomètre et la petite taille du bassin.

La figure 71 présente les résultats obtenus en prévision à  $k+3$  (soit 1h30 d'horizon de prévision).

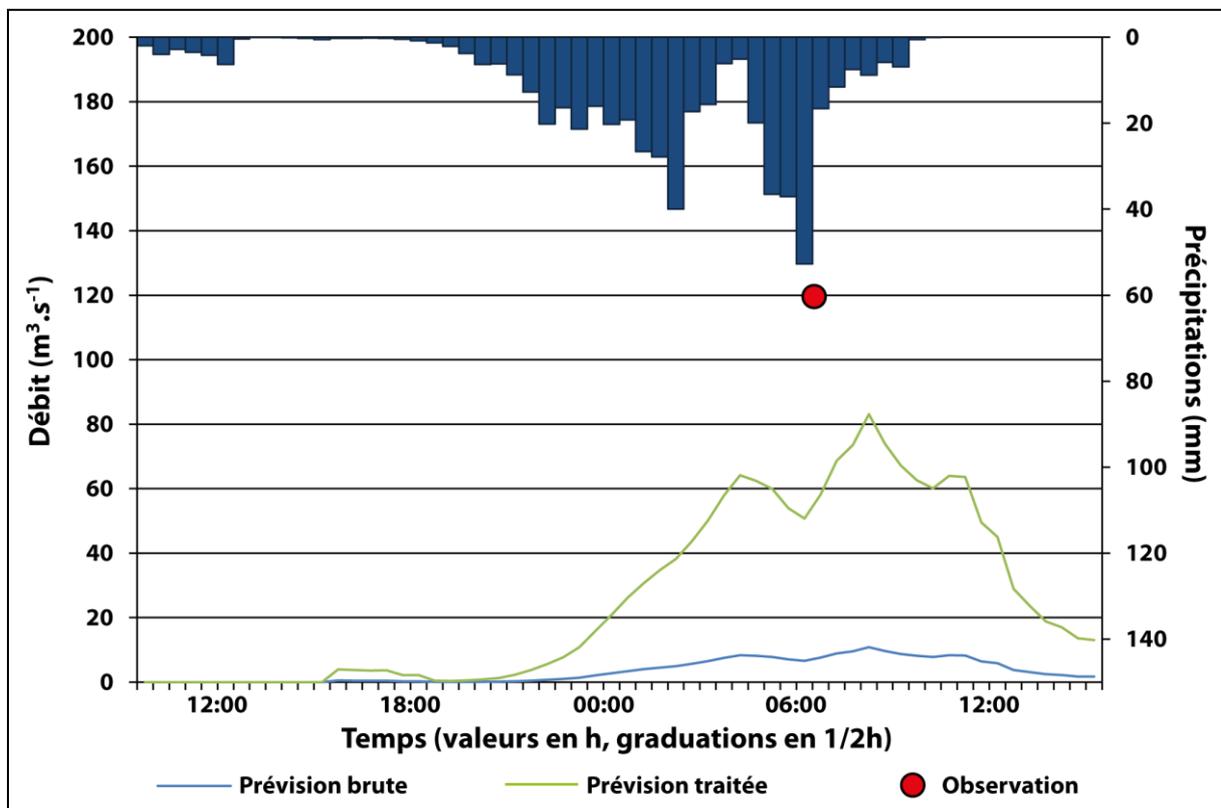


Figure 71: prévision de l'évènement des 8 et 9 septembre 2002 sur le ruisseau du Lyonnais à  $k+3$ .

Le tableau 16 détaille les résultats selon les critères prédéfinis pour les quatre horizons de prévision.

Tableau 16: résultats détaillés en prévision de l'évènement de septembre 2002 sur le ruisseau du Lyonnais.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	Sans objet	Sans objet	46%	62%	3,00
$k+2$	Sans objet	Sans objet	45%	63%	1,00
$k+3$	Sans objet	Sans objet	49%	69%	1,50
$k+4$	Sans objet	Sans objet	31%	73%	1,50

### 5.3.1.3 Analyse des résultats obtenus

Malgré un coefficient correctif très fort, le modèle ne parvient pas à approcher la valeur estimée par le retour d'expérience. On peut expliquer ce résultat d'une part par l'incertitude vraisemblablement très forte sur la valeur déterminée par le retour d'expérience, et d'autre part par l'incertitude également très importante sur la quantification des précipitations qui sont ici les pluies observées par le radar. A ce sujet, les récents travaux de Visserot (2012) au laboratoire ont montré que la pluviométrie radar pouvait être de à 20 à 60% inférieure aux valeurs des pluviomètres sur une dizaine d'évènements différents sur le Gardon à Anduze. Etant donné que le modèle a été conçu avec des données de pluviomètres, le manque de volume en entrée peut-être important.

Ceci étant dit, la valeur de débit estimée par le retour d'expérience peut aussi être sous-estimée et la sous-estimation du radar peut n'être que faible, ou dans les fourchettes les plus basses. Rien ne permet de le vérifier en l'absence d'instrumentation sur le bassin. De plus, ces explications ne peuvent pas masquer le retard, non corrigé par notre méthode et relativement important au regard des *SPPD*.

Ce seul test sur un très petit bassin pourrait laisser penser que le modèle n'est pas adapté aux bassins de taille très réduite capables de débits spécifiques extrêmes et de temps de réponse très brefs. Mais on peut aussi penser que les forçages comme le débit estimé ne sont pas quantifiés correctement. En tout état de cause, comme un seul évènement est disponible, il est difficile de tirer des conclusions fiables sur cet aspect mais nous pourrons revenir sur ce point lorsque nous considérerons les autres bassins cible.

## 5.3.2 L'Alzon

### 5.3.2.1 Présentation rapide du bassin

Le bassin est là-aussi de très petite taille (seulement 15km<sup>2</sup>) et présente une forme globalement allongée vers le sud-est, direction plus traditionnellement cévenole que pour le ruisseau du Lyonnais, affluent précité. Ce bassin est plus pentu, autant pour ce qui est des versants (24%) que pour ce qui est du drain (3.2%). Ceci-dit, ces valeurs restent inférieures à la moyenne des bassins cible dans l'étude.

L'occupation des sols est encore une fois très largement naturelle, composée pour l'essentiel de forêts de chênes verts, de pins maritimes et de châtaigniers. D'un point de vue géologique, l'extrême aval se situe sur le bloc toit d'un bras de la faille des Cévennes, bloc érodé jusqu'aux calcaires liasiques. En amont, dans le bloc mur moins érodé, on remonte le cours d'eau en même temps que les séries (jurassique supérieur, moyen, inférieur, lias, trias) jusqu'au contact avec les schistes des Cévennes et leurs filons granitiques et lamprophyriques dans le bassin moyen. Ces derniers sont à nouveau remplacés en amont par le lias et le trias.

La figure 72 présente la topographie du bassin de l'Alzon.

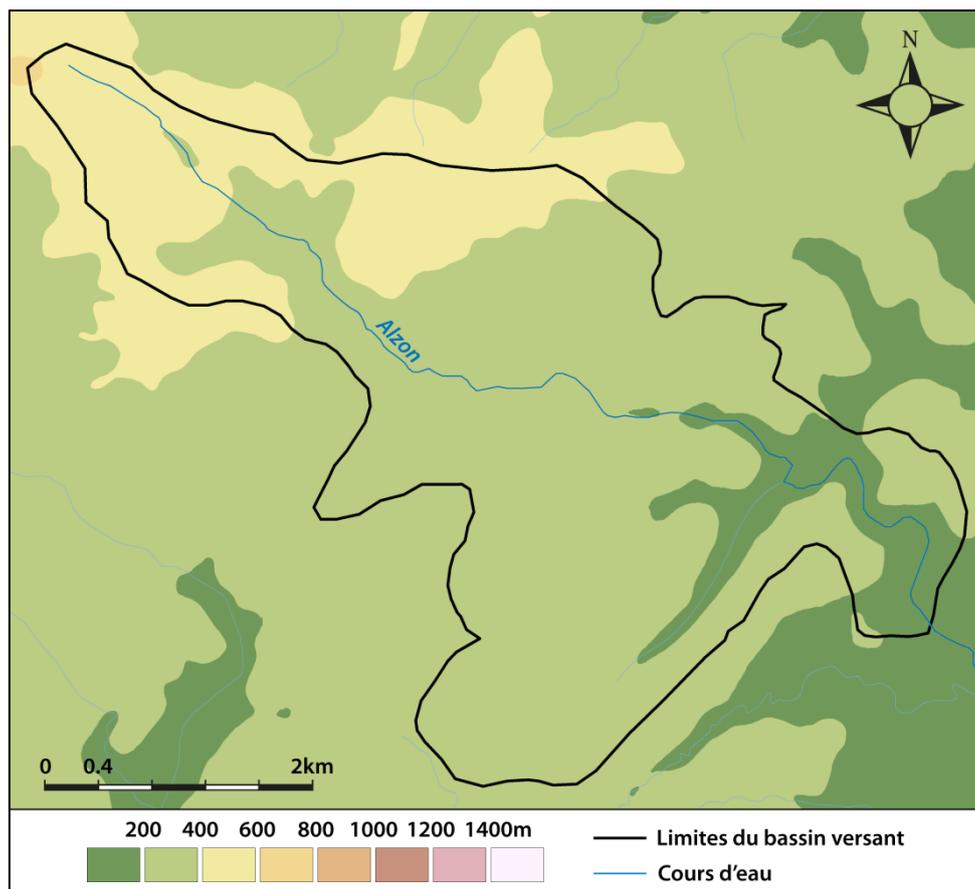


Figure 72: carte du bassin versant de l'Alzon à Saint-Christol-lez-Alès.

### 5.3.2.2 Evènements et résultats

Là encore, un seul évènement est disponible : celui des 8 et 9 septembre 2002. La pluie utilisée est la même que pour le ruisseau du Lyonnais : une évaluation radar à l'exutoire de ce dernier ruisseau fournie aux trois entrées du modèle. De même, seule une évaluation du débit à l'exutoire est disponible. Elle fait état d'une pointe à  $430 \text{ m}^3 \cdot \text{s}^{-1}$  à 7h le 9 septembre, avec toutes les incertitudes que ce type d'évaluation *a posteriori* implique.

La figure 73 présente les résultats obtenus en prévision à  $k+2$  (soit 1h d'horizon de prévision).

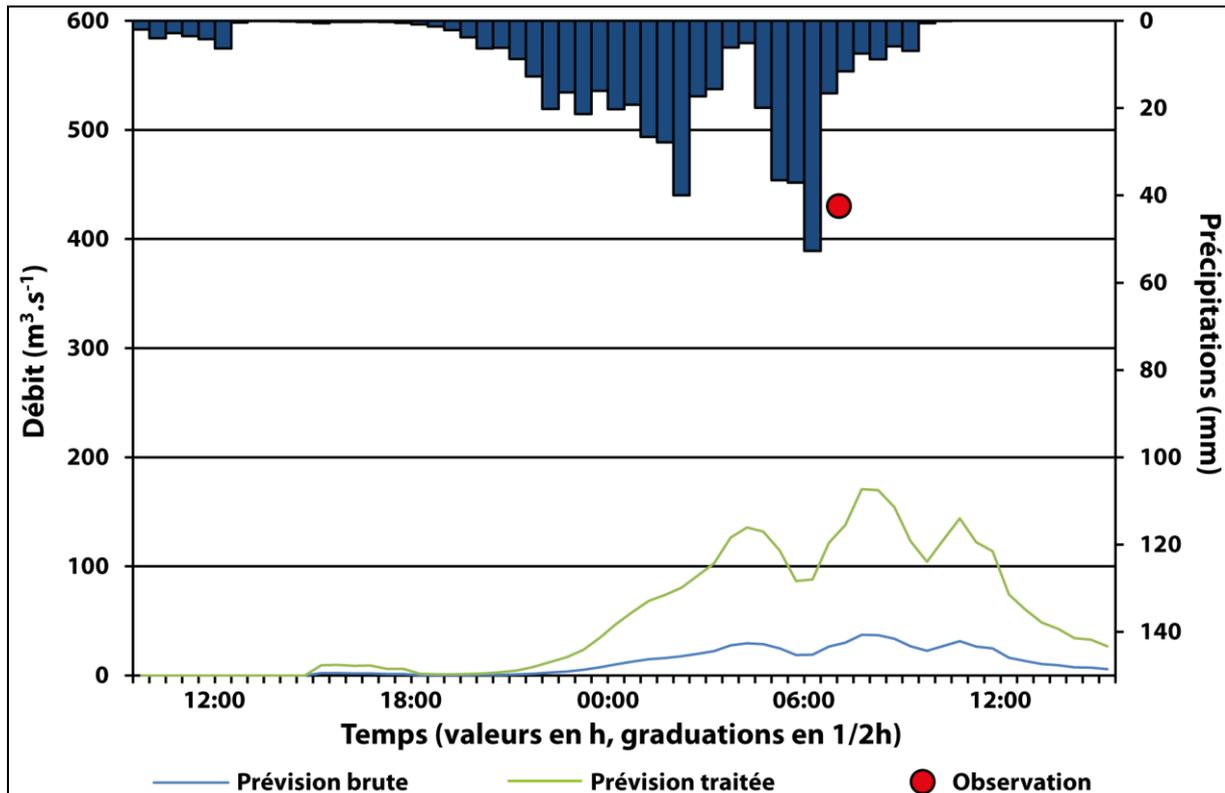


Figure 73: prévision de l'évènement des 8 et 9 septembre 2002 sur l'Alzon à  $k+2$ .

Le tableau 17 détaille les résultats selon les critères prédéfinis aux quatre horizons de prévision.

Tableau 17: résultats détaillés en prévision de l'évènement de septembre 2002 sur l'Alzon.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	Sans objet	Sans objet	39%	31%	2,50
$k+2$	Sans objet	Sans objet	40%	32%	0,50
$k+3$	Sans objet	Sans objet	43%	36%	1,00
$k+4$	Sans objet	Sans objet	46%	27%	1,00

### 5.3.2.3 Analyse des résultats obtenus

Avec des pluies strictement identiques et ayant donc les mêmes incertitudes que pour le ruisseau du Lyonnais, la seule différence avec ce dernier ne peut tenir qu'au coefficient correctif et au décalage temporel. Malheureusement, avec un coefficient plus faible (4,60 contre 7,47) et un débit spécifique de pointe de presque  $29 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ , l'on ne pouvait attendre de meilleurs résultats. Seule une amélioration d'une demi-heure sur le décalage du pic est à noter. Elle est due à un temps de réponse probablement légèrement supérieur, pour ce bassin qui est plus grand et plus sinueux en aval (courtes gorges), et qui a pu permettre à la crue d'atteindre son pic une demi-heure plus tard. Ceci étant dit, le temps de réponse ne peut être estimé que grossièrement compte tenu des incertitudes temporelles inhérentes à ce type d'estimation de débit. D'une manière générale, les conclusions incertaines tirées sur le ruisseau du Lyonnais sont identiques à celles que l'on peut tirer ici.

### 5.3.3 La Salandre

#### 5.3.3.1 *Présentation rapide du bassin*

La Salandre est un affluent du Galeizon dont le bassin mesure 21km<sup>2</sup>. Le cours d'eau a une pente très forte (6,3%), la plus forte rencontrée parmi les bassins visés. Les versants du bassin ne sont pas en reste puisqu'ils présentent une pente de 33%, supérieure à la moyenne des bassins ici considérés. La Salandre s'écoule vers le sud-est, dans une direction typiquement cévenole.

Justement, au niveau géologique, les schistes des Cévennes sont omniprésents, ponctués de gneiss et de filons de granite. Seule une très fine frange de trias occupe certaines crêtes méridionales. L'occupation des sols est presque exclusivement naturelle, composée majoritairement de forêts de chênes verts et de châtaigniers.

La figure 74 présente une carte du bassin de la Salandre.

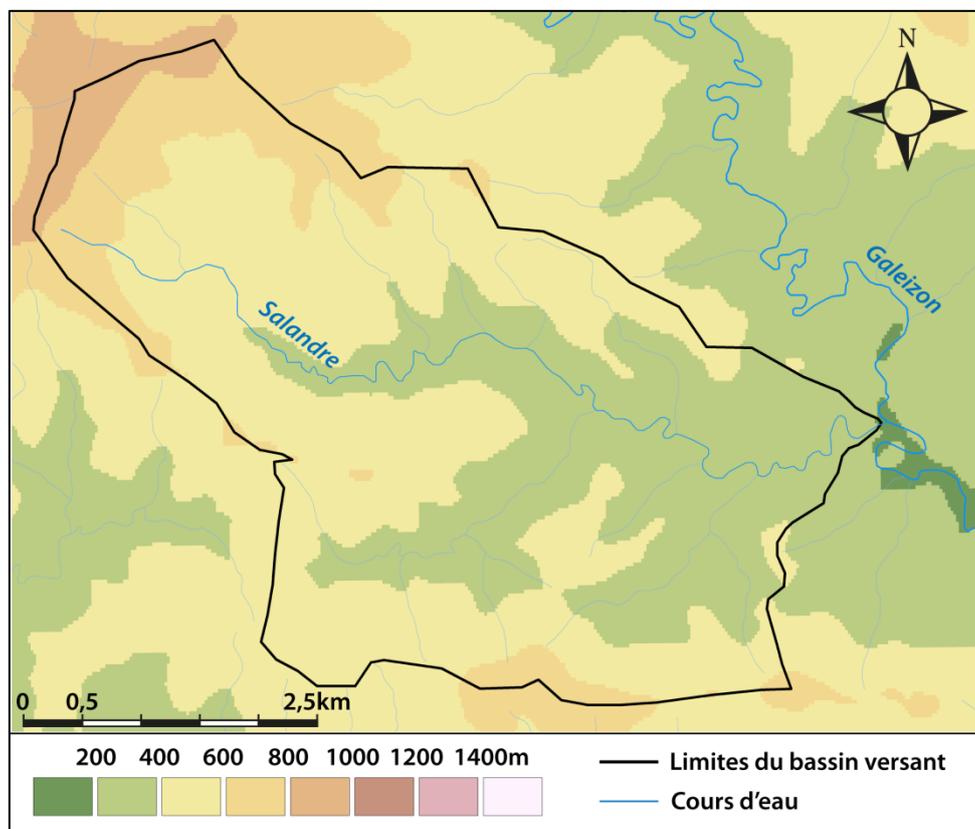


Figure 74: carte du bassin versant de la Salandre.

#### 5.3.3.2 *Evènements et résultats*

Grâce au retour d'expérience et aux simulations effectuées par SIEE pour le Syndicat Mixte du Galeizon suite à l'évènement de septembre 2002 sur le Galeizon et affluents, nous disposons d'hydrogrammes simulés calés sur les estimations de débit de pointe en différents points du bassin du Galeizon.

Nous disposons donc, sur la Salandre, d'un hydrogramme dont la forme présente une incertitude importante mais dont le pic peut être considéré comme moins imprécis que dans le cas précédent, surtout compte tenu de la forme de la vallée, assez encaissée, qui facilite l'évaluation des débits à partir d'une hauteur d'eau.

Les pluies sont issues de pluies radar moyennes sur l'ensemble du bassin du Galeizon. Elles sont utilisées pour les trois entrées qui sont donc identiques. Il convient de noter qu'une partie nord-est du bassin a subi peu de pluie en septembre 2002, conférant à la moyenne une valeur certainement inférieure à ce qui aurait pu être observé par un pluviomètre situé sur la partie sud-ouest du bassin. Malheureusement, aucun pluviomètre représentatif n'est disponible dans le secteur. Les fortes intensités, traditionnellement génératrices de crues rapides et volumineuses, sont donc lissées et ne peuvent plus participer à un forçage important du modèle.

Quoi qu'il en soit, le pic de crue est estimé à  $310 \text{ m}^3 \cdot \text{s}^{-1}$  (presque  $15 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ ), avec un cumul de précipitations de 232mm en moyenne sur le bassin et des intensités maximales de 29mm en 0,5h et 59mm en 1h, ce qui est donc plutôt faible au regard de la crue générée et des autres intensités pluvieuses observées sur la bordure cévenole lors de cet épisode. C'est bien l'effet du lissage des données par la moyenne sur le bassin qui se fait sentir.

La figure 75 présente l'hydrogramme obtenu à  $k+2$ , soit un horizon de prévision d'une heure.

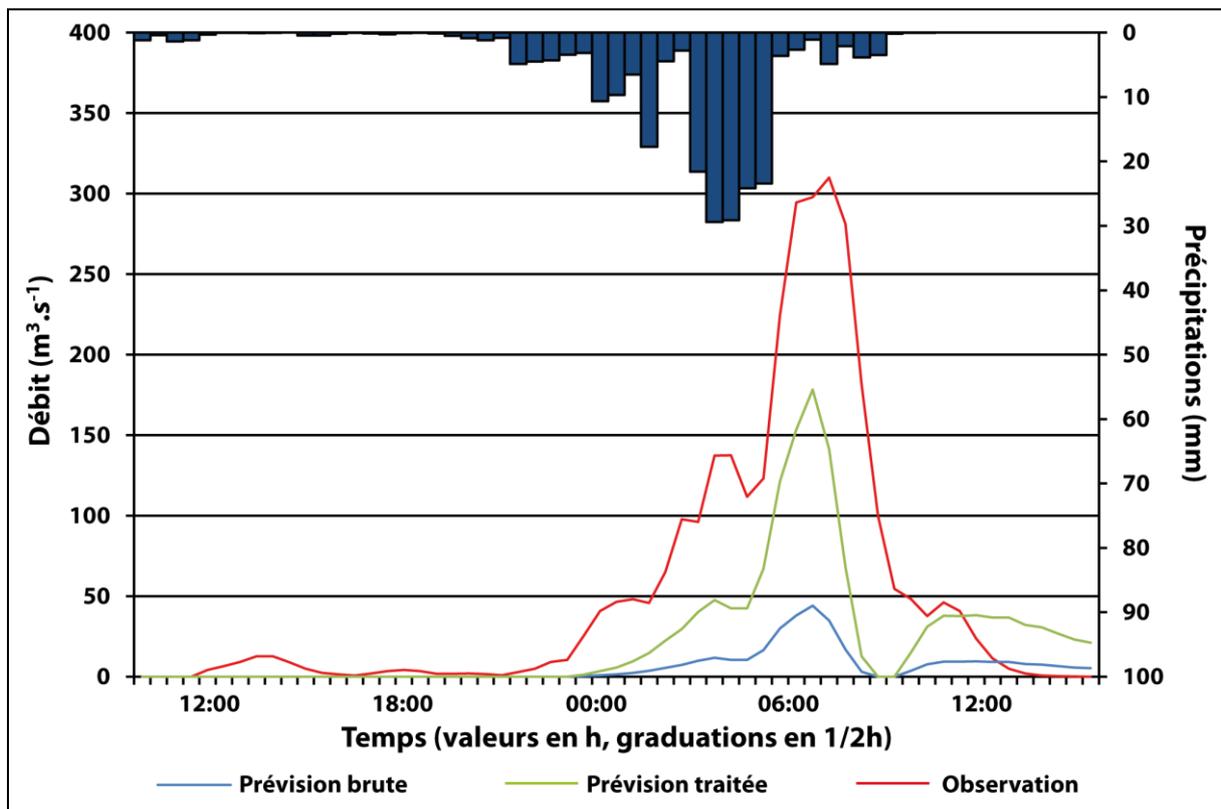


Figure 75: prévision de l'évènement des 8 et 9 septembre 2002 sur la Salandre à  $k+2$ .

Le tableau 18 présente quant à lui les résultats plus détaillés pour chaque horizon de prévision.

Tableau 18: résultats détaillés en prévision de l'évènement de septembre 2002 sur la Salandre.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	0,56	-3,29	41%	44%	-0,50
$k+2$	0,47	-0,62	46%	58%	-0,50
$k+3$	0,48	0,11	47%	53%	-0,50
$k+4$	0,40	0,29	44%	44%	0,00

### 5.3.3.3 Analyse des résultats obtenus

Les résultats obtenus sont de mauvaise qualité. En moyenne, moins de la moitié du pic est prévue à l'instant où il se réalise. Hors de toute considération sur le fonctionnement du modèle, plusieurs éléments sont à souligner pour tenter d'expliquer ces problèmes :

- La faiblesse des pluies comparativement au débit spécifique de pointe : même sur le bassin le plus réactif qui soit, il est difficile d'imaginer que des cumuls de 239mm avec des intensités maximales de moins de 60mm/h soient en mesure de faire réagir un bassin jusqu'à des débits spécifiques aussi élevés. Ces cumuls restent en effet assez courants pour la région. De plus, ici encore, le modèle est prévu pour et calé par des données de pluviomètres ; il n'est donc pas optimal de travailler avec des données radar, en particulier si celles-ci sont moyennées sur un bassin entier.
- La qualité de l'estimation ne semble pas devoir être remise en cause dans de larges proportions mais elle présente forcément son lot d'incertitudes, préjudiciables à la démarche.
- L'intensité extrême de la crue générée : elle est, peut-être, même avec le coefficient correctif appliqué en sortie, hors du domaine de validité du modèle, comme cela a possiblement été le cas pour l'Alzon et le Ruisseau du Lyonnais. Seuls des tests plus nombreux pourraient permettre de valider cette hypothèse.

Plus surprenants sont les décalages temporels sur le pic, très faibles voire nuls. En comparant le pic de précipitations et le pic de débit, on peut raisonnablement estimer le temps de réponse sur l'évènement à trois heures ce qui est plutôt élevé pour un bassin de si faibles dimensions et somme toute assez proche de ce qu'on pourrait attendre du bassin du Gardon à Mialet. Nous gardons en mémoire cet élément pour vérifier s'il peut se retrouver sur d'autres bassins et éventuellement en tirer des conclusions robustes, malgré la faible relation entre décalages et dimensions du bassin (section 5.2.2.3).

## 5.3.4 L'Amous

### 5.3.4.1 Présentation rapide du bassin

L'Amous est un affluent du Gardon d'Anduze dont le bassin mesure 21km<sup>2</sup>. Si la pente du drain principal est supérieure à la moyenne (3,8%), la pente des versants est, elle, inférieure (24%). Le cours d'eau coule du nord au sud alimenté par un bassin relativement allongé.

L'occupation des sols y est essentiellement naturelle, hormis des terrains miniers à l'est qui ne représentent finalement que peu de surface et quelques cœurs de villages ou hameaux peu étendus. On trouve donc essentiellement des forêts d'essences comparables à celles de la Salandre. Géologiquement, le bassin de l'Amous est en revanche bien différent : dans le faisceau de failles des Cévennes, à l'aval, des chicots de terrains essentiellement carbonatés du jurassique inférieur, du lias et du trias occupent les sous-sols alors qu'en amont, les séries liasiques et triasiques plus vastes et non compartimentées sont très majoritaires. Les terrains karstiques sont minoritaires avec des réseaux assez peu développés.

La figure 76 présente une carte du bassin versant.

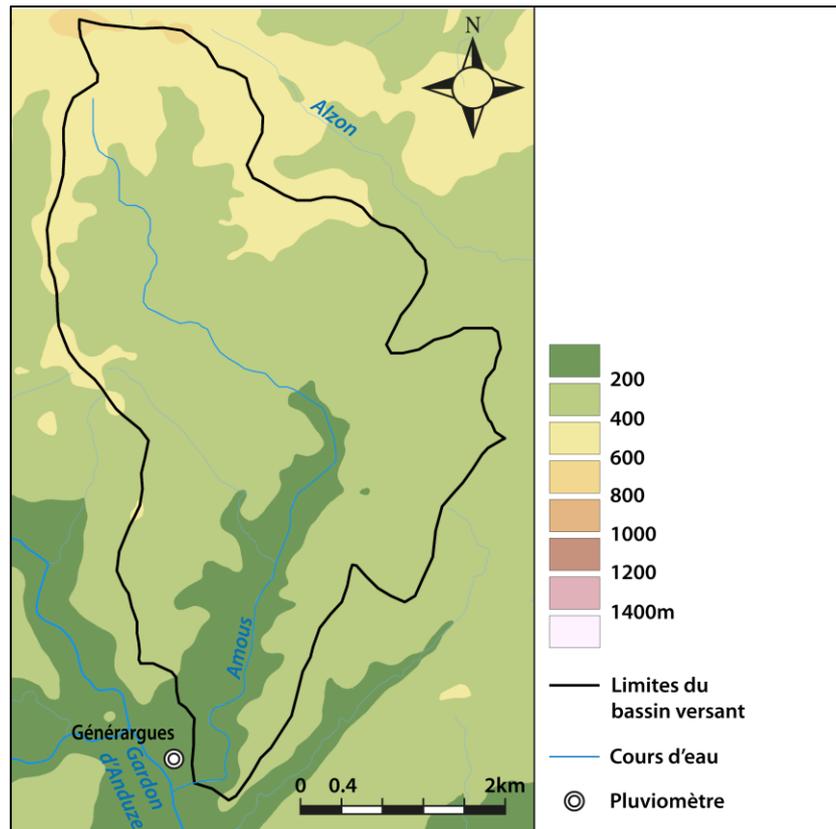


Figure 76: carte du bassin versant de l'Amous.

#### 5.3.4.2 Evènements et résultats

Ici encore, c'est le retour d'expérience de septembre 2002 qui nous fournit une valeur estimée du débit de pointe. Elle est de  $435\text{m}^3.\text{s}^{-1}$  vers 6h du matin le 9, soit près de  $21\text{m}^3.\text{s}^{-1}.\text{km}^{-2}$ . Un pluviomètre Météo France est présent à la Bambouseraie de Prafrance à l'aval et donne des valeurs horaires pour l'épisode. Une estimation radar a par ailleurs été faite en amont du bassin, conférant à l'évènement un deuxième hyétogramme. La troisième entrée du modèle sera la moyenne des deux autres. Cette valeur par défaut est choisie en raison de l'absence de certitudes sur le poids de chacune des données disponibles. Générargues remplacera Mialet et l'estimation radar en amont remplacera Barre-des-Cévennes dans le modèle. Le cumul moyen s'élève à 472mm (587mm en aval et 358mm en amont) et les intensités maximales atteignent 108mm/1h et 67mm/0,5h à Générargues, jusqu'à 72mm/1h et 44mm/0,5h en moyenne sur le bassin.

La figure 77 présente le résultat obtenu à  $k+3$ .

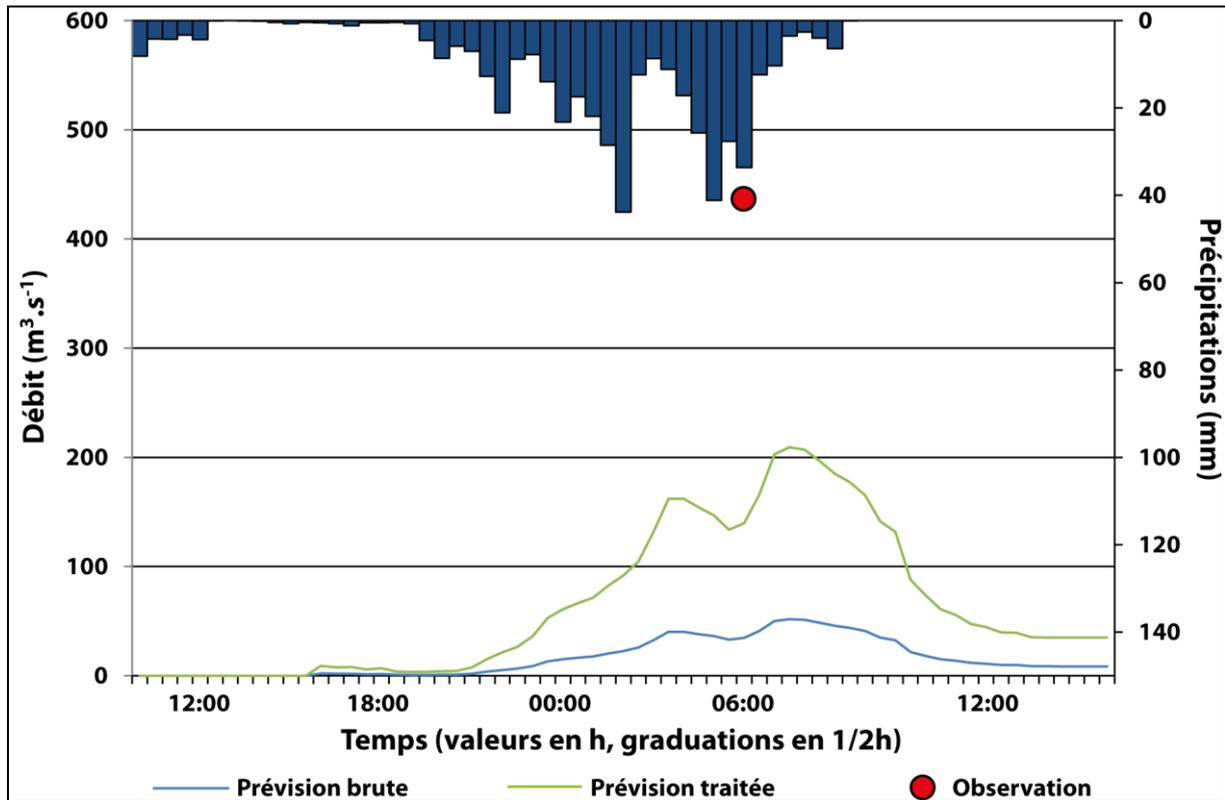


Figure 77: prévision de l'événement de septembre 2002 sur l'Amous à  $k+3$ .

Le tableau 19 donne les résultats détaillés pour les quatre horizons de prévision.

Tableau 19: résultats détaillés en prévision de l'événement de septembre 2002 sur l'Amous.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	Sans objet	Sans objet	34%	48%	2,50
$k+2$	Sans objet	Sans objet	30%	46%	1,50
$k+3$	Sans objet	Sans objet	32%	48%	1,50
$k+4$	Sans objet	Sans objet	22%	51%	2,00

#### 5.3.4.3 Analyse des résultats obtenus

Les résultats ne sont pas de bonne qualité. De fait, les pourcentages du pic restent faibles et la prévision est assez largement en retard sur tous les horizons de prévision. Contrairement au bassin précédent, ce bassin semble réagir assez nettement plus vite. Pour ce qui est de l'amplitude, le débit spécifique semble, malgré le coefficient correctif, être encore trop extrême pour le modèle ce qui ajoute à la conjecture selon laquelle le modèle serait mal adapté aux trop petits bassins générant des crues aux débits spécifiques extrêmes. Cependant, encore une fois, les forçages étant partagés entre des données pluviométriques (réputées fiables) et des données radar (réputées peu fiables et inférieures aux pluviomètres), et le débit estimé étant très incertain, il est toujours aussi difficile de tirer une conclusion robuste.

### 5.3.5 L'Auzonnet

#### 5.3.5.1 Présentation rapide du bassin

L'Auzonnet est un affluent de l'Auzon, lui-même tributaire de la Cèze en son bassin moyen. Ce bassin mesure  $39\text{km}^2$  aux Mages. Le drain principal a une pente plutôt faible (2,4%), tout comme les versants (25%). S'écoulant vers l'est-sud-est, il présente une forme très allongée.

Le bassin est essentiellement occupé par des milieux naturels. Même si les remblais et terrils miniers abandonnés y sont relativement nombreux, le reboisement naturel progressif de ces terrains limite leur impact sur l'hydrologie et les rend très proches des occupations des sols naturelles. Géologiquement, le bassin est assez complexe. Près de l'exutoire, le faisceau de failles des Cévennes découpe les séries carbonatées jurassiques en blocs assez ténus. Le fond de la vallée, d'aval en amont, voit une forte représentation des terrains carbonifères (argiles, grès, charbons...). Le reste du bassin est occupé en amont par des gneiss, plus en aval par le lias et le trias que l'on retrouve sur toute la bordure cévenole. La figure 78 présente une carte du bassin.

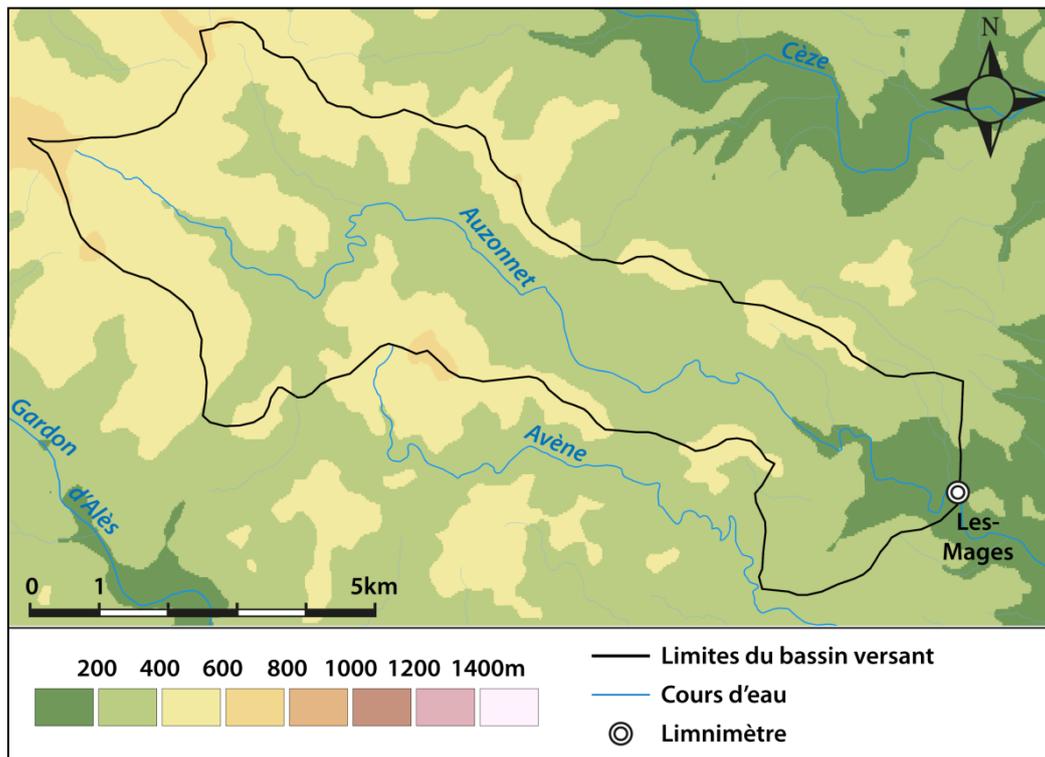


Figure 78: carte du bassin versant de l'Auzonnet.

#### 5.3.5.2 Evènements et résultats

Contrairement aux bassins traités précédemment, trois évènements sont disponibles sur le bassin de l'Auzonnet. Les données de débit disponibles dans la base de données Sevnol de l'Observatoire Hydrométéorologique Cévennes Vivarais (OHMCV) permettent de tracer les hydrogrammes pour les trois évènements à la station des Mages.

En revanche, l'éloignement et la disponibilité aléatoire des données issues des pluviomètres voisins nous ont conduits à privilégier une estimation radar ponctuelle à l'exutoire.

Les trois évènements sont numérotés 921, 922 et 923 et sont décrits dans le tableau 20.

Tableau 20: description synthétique des trois évènements disponibles pour l'Auzonnet. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
921	déc-00	135	9	5	58	1,49	radar	1 sur 1
922	oct-01	98	23	14	68	1,74	radar	1 sur 1
923	sept-02	378	67	38	169	4,33	radar	1 sur 1

Les résultats obtenus sur l'évènement le plus intense, le numéro 923, de septembre 2002, à un horizon de prévision de  $k+2$  sont présentés dans la figure 79.

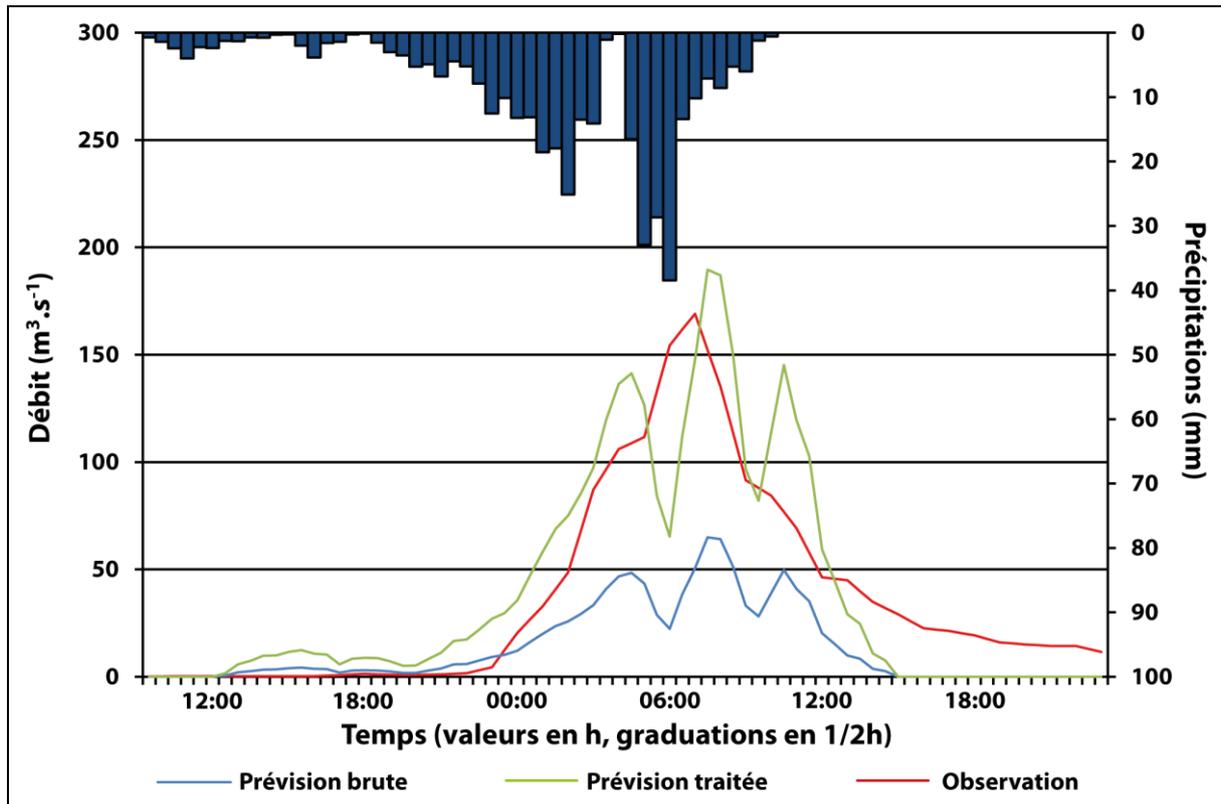


Figure 79: prévision de l'évènement de septembre 2002 sur l'Auzonnet à  $k+2$ .

Le détail des critères de qualité de la prévision pour chaque évènement est présenté dans le tableau 21.

Tableau 21: résultats détaillés en prévision des évènements 921, 922 et 923 sur l'Auzonnet.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
921	k+1	0,67	-37,03	90%	101%	2,00
921	k+2	0,61	-10,43	95%	103%	2,00
921	k+3	0,64	-3,97	93%	102%	2,50
921	k+4	0,61	-2,05	96%	105%	1,50
922	k+1	-0,13	-9,36	91%	106%	-2,00
922	k+2	-0,06	-2,07	92%	104%	-2,00
922	k+3	-0,08	-0,74	87%	98%	-2,00
922	k+4	-0,14	-0,27	93%	114%	-2,00
923	k+1	0,61	-16,78	84%	118%	2,50
923	k+2	0,77	-1,96	87%	112%	0,50
923	k+3	0,65	-1,23	103%	129%	1,00
923	k+4	0,68	-0,28	81%	140%	1,00

### 5.3.5.3 Analyse des résultats obtenus

D'une manière générale, les résultats obtenus sur l'Auzonnet sont de bonne qualité. Plusieurs points sont à souligner :

- Des décalages temporels significatifs sont observés entre les pics estimés et mesurés (jusqu'à deux heures). Cette observation est à pondérer par les SPPD de très bonne qualité obtenus sur l'ensemble des évènements.
- A SPPD convenables, peuvent correspondre, via un décalage temporel important, des PPD de mauvaise qualité. Hormis sur l'évènement 923 à k+3 et k+4, les surestimations du PPD restent modérées.
- Les montées de pic, en particulier vers la base, sont systématiquement en avance sur l'observation ce qui est un point positif pour la prévision opérationnelle.

Ainsi, sur ce bassin de 39km<sup>2</sup>, avec des données de pluie réduites à une estimation radar à l'exutoire, deux évènements moyens et un évènement fort, le modèle développé est utilisable en prévision jusqu'à k+4. L'information qu'il fournit peut être un outil de choix pour la prévision sur ce petit bassin. Il s'agit du premier bassin de l'échantillon sur lequel une prévision exploitable est obtenue. On remarque que sur les trois épisodes de ce bassin, la donnée radar n'a pas posé les problèmes que nous avons envisagés sur les premiers bassins traités.

## 5.3.6 Le Galeizon amont

### 5.3.6.1 Présentation rapide du bassin

Le bassin versant du Galeizon en amont du Pont-de-Roubarbel (c'est-à-dire là où la Salandre conflue) mesure 39km<sup>2</sup>. Le drain est moyennement pentu au regard des bassins concernés par l'étude (3.2%) alors que les versants sont plus pentus que la moyenne avec 29%. De forme allongée, le bassin s'étire du nord-ouest au sud-est et est assez sinueux à l'aval où l'on rencontre une morphologie encaissée.

Géologiquement, le bassin est exclusivement composé de schistes et de gneiss cévenols. L'occupation des sols est, quant à elle, très largement naturelle puisque seules les rares habitations et les étroites routes représentent les occupations des sols artificielles.

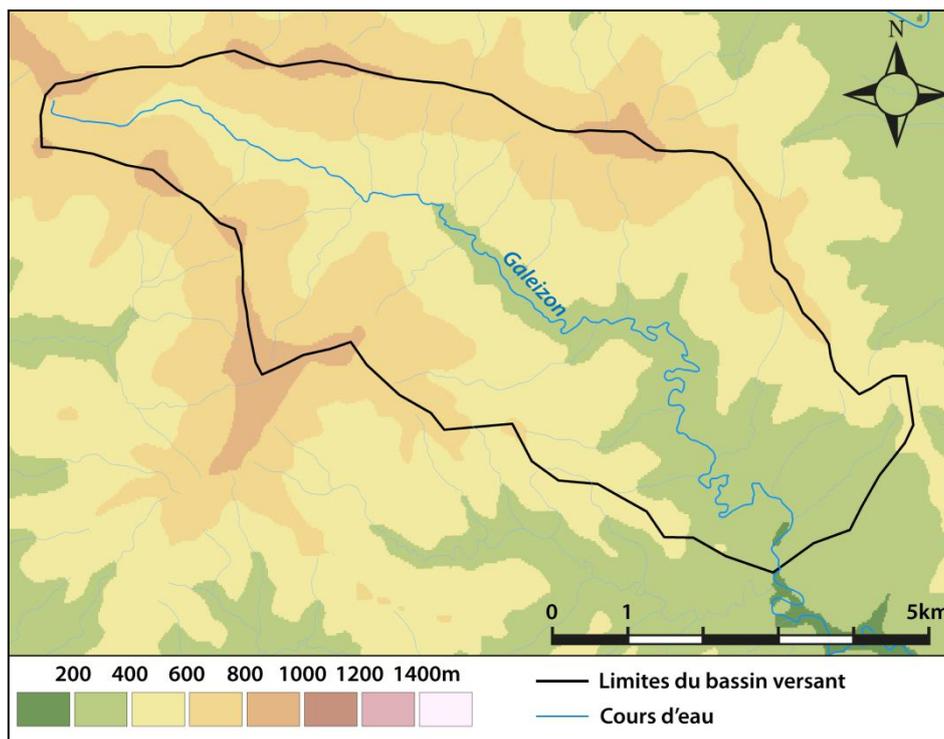


Figure 80: carte du bassin du Galeizon amont.

#### 5.3.6.2 Evènements et résultats

Les données disponibles concernent l'évènement de septembre 2002 et sont, en entrée, identiques à celles utilisées sur la Salandre. En revanche, une estimation particulière du débit de pointe et un hydrogramme simulé sont disponibles. Dans ce cas encore, seul le coefficient correctif pourra changer quelque chose aux prévisions obtenues sur la Salandre.

Le débit de pointe de l'épisode considéré est de  $330 \text{ m}^3 \cdot \text{s}^{-1}$  soit environ  $8,5 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ .

La figure 81 représente les résultats obtenus sur cet évènement en prévision à  $k+2$ .

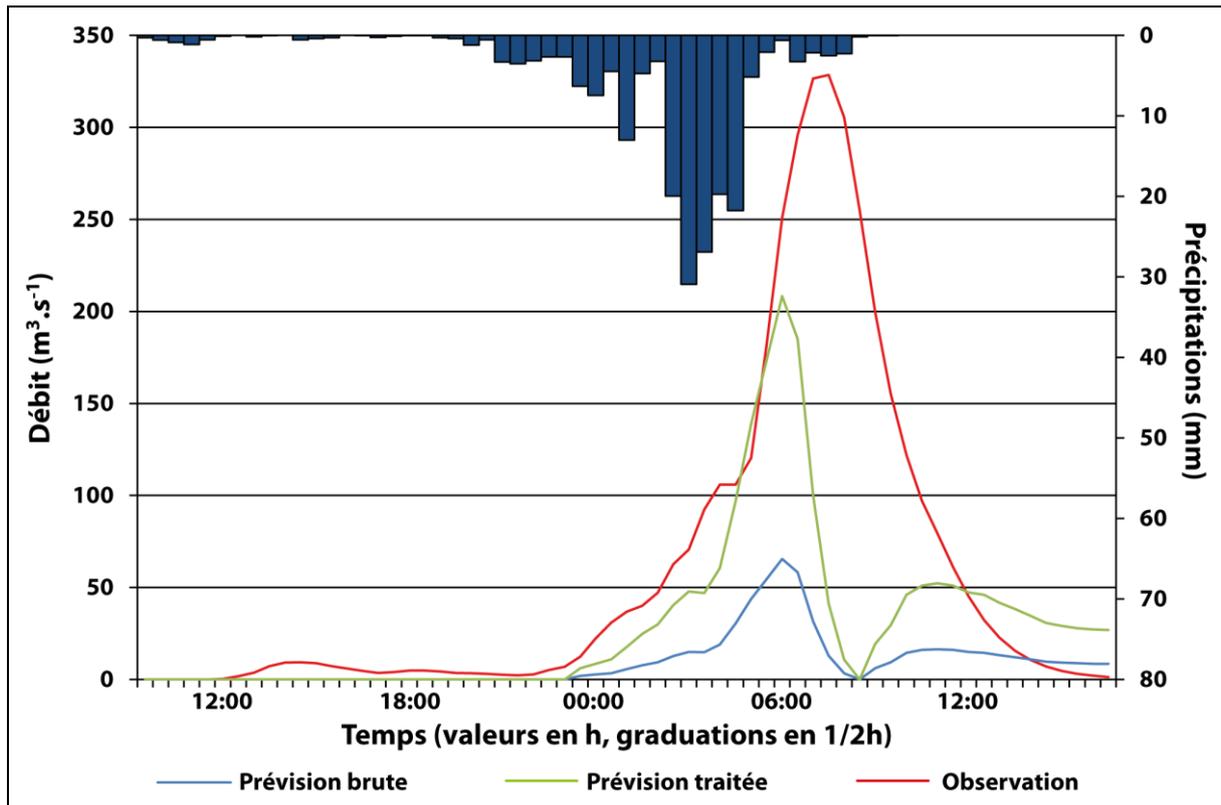


Figure 81: prévision de l'événement de septembre 2002 sur le Galeizon amont à  $k+2$ .

Le tableau 22 représente quant à lui les résultats détaillés de l'événement selon plusieurs critères de qualité.

Tableau 22: résultats détaillés en prévision de l'événement de septembre 2002 sur le Galeizon amont.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	0,74	-4,48	54%	56%	-1,50
$k+2$	0,22	-3,42	7%	69%	-1,50
$k+3$	0,27	-1,00	13%	63%	-1,50
$k+4$	0,20	-0,38	4%	54%	-1,00

### 5.3.6.3 Analyse des résultats obtenus

De même que les résultats, la conclusion est presque identique à celle que l'on pouvait tirer de l'événement sur la Salandre, à ceci près que le débit spécifique n'est pas aussi élevé, ce qui permet de mettre plus en avant les problèmes liés à la faiblesse de la pluie radar disponible que le caractère extrême de l'événement. On constate, là encore, que la grossière estimation que l'on peut faire du temps de réponse amène une valeur élevée pour un si petit bassin.

## 5.3.7 La Ganière

### 5.3.7.1 Présentation rapide du bassin

La Ganière est un affluent de la Cèze dont le bassin, étiré du nord-nord-ouest vers le sud-sud-est mesure  $55\text{km}^2$ . Le bassin est assez pentu et présente une dénivelée importante pour sa surface : le drain est en pente de 3.7%, les versants de 30% et la dénivelée est de plus de 800m, trois valeurs supérieures à (ou voisines de) la moyenne de l'échantillon de bassins. La figure 82 présente une carte du bassin.

Un tiers aval du bassin est occupé par des terrains carbonifères (grès, psammistes, schistes, charbons) alors que le reste du bassin est schisteux, granitique et gneissique. L'occupation des sols est là encore presque exclusivement naturelle.

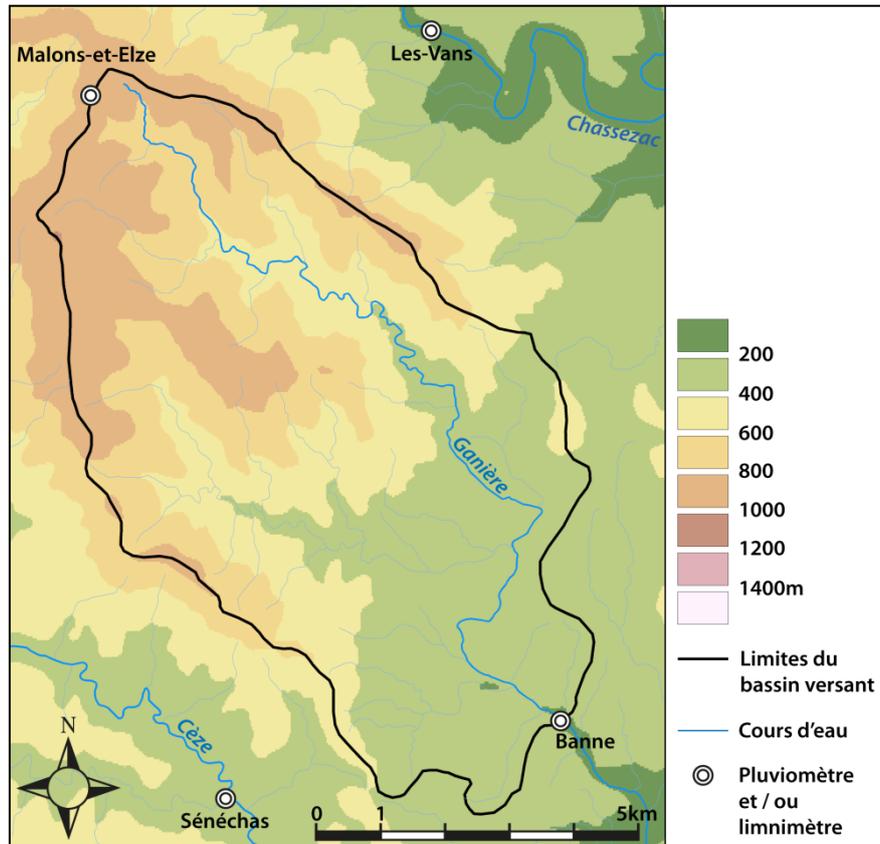


Figure 82: carte du bassin versant de la Ganière.

### 5.3.7.2 Evènements et résultats

La base de données de l'OHMCV ainsi que les données fournies par le projet BVNE ont permis d'extraire 11 évènements pour ce bassin. Le limnimètre se situe sur la commune de Banne et les données de débit ont été extraites directement de la base OHMCV. Trois pluviomètres dont deux hors-zone permettent d'obtenir un champ de pluie relativement convenable sur le bassin : Sénéchas au sud-est (données BVNE), Malons-et-Elze à l'extrême nord du bassin (données BVNE) et Les-Vans (OHMCV) au nord-est. Les données des pluviomètres ne sont pas systématiquement disponibles et les règles de substitution vues en début de section s'appliquent au gré des disponibilités. Par rapport au modèle élaboré à Mialet, le pluviomètre de Barre-des-Cévennes est représenté par Malons-et-Elze, le pluviomètre de Saint-Roman-de-Tousque par Les-Vans et celui de Mialet par Sénéchas. Ce choix peut paraître discutable mais comme rien ne permettait réellement de différencier les deux pluviomètres septentrionaux, des tests sur les évènements où les trois pluviomètres sont disponibles ont permis d'identifier la configuration optimale. La différence est toutefois restée minimale, ce qui atteste, encore une fois, d'une certaine robustesse du modèle face à ses entrées. Le tableau 23 fournit une description rapide des évènements.

De l'évènement 915 à l'évènement 920 ainsi que pour le 913, c'est le pluviomètre de Malons-et-Elze qui est indisponible. Sur l'évènement de septembre 2002, le 914, c'est le pluviomètre des Vans qui est le seul disponible.

Tableau 23: description synthétique des 11 évènements disponibles pour la Ganière. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
910	nov.-00	159	11,5	6,9	88	1,60	Pluvio.	3 sur 3
911	nov.-00	108	10	6,5	63	1,15	Pluvio.	3 sur 3
912	déc.-00	185	8,5	5	77	1,40	Pluvio.	3 sur 3
913	oct.-01	222	37	21	107	1,95	Pluvio.	2 sur 3
914	sept.-02	93	15	9	206	3,75	Pluvio.	1 sur 3
915	sept.-03	126	20	12	55	1,00	Pluvio.	2 sur 3
916	nov.-03	136	10	6	64	1,16	Pluvio.	2 sur 3
917	nov.-03	117	16	10	81	1,47	Pluvio.	2 sur 3
918	nov.-03	242	10	6	93	1,69	Pluvio.	2 sur 3
919	oct.-04	246	25	15	88	1,60	Pluvio.	2 sur 3
920	nov.-04	84	13	8	78	1,42	Pluvio.	2 sur 3

Les figures 83 et 84 présentent les résultats obtenus sur l'évènement n°919 (évènement moyen) et sur l'évènement 914 (intense), assez représentatifs de la problématique spécifique rencontrée sur ce bassin.

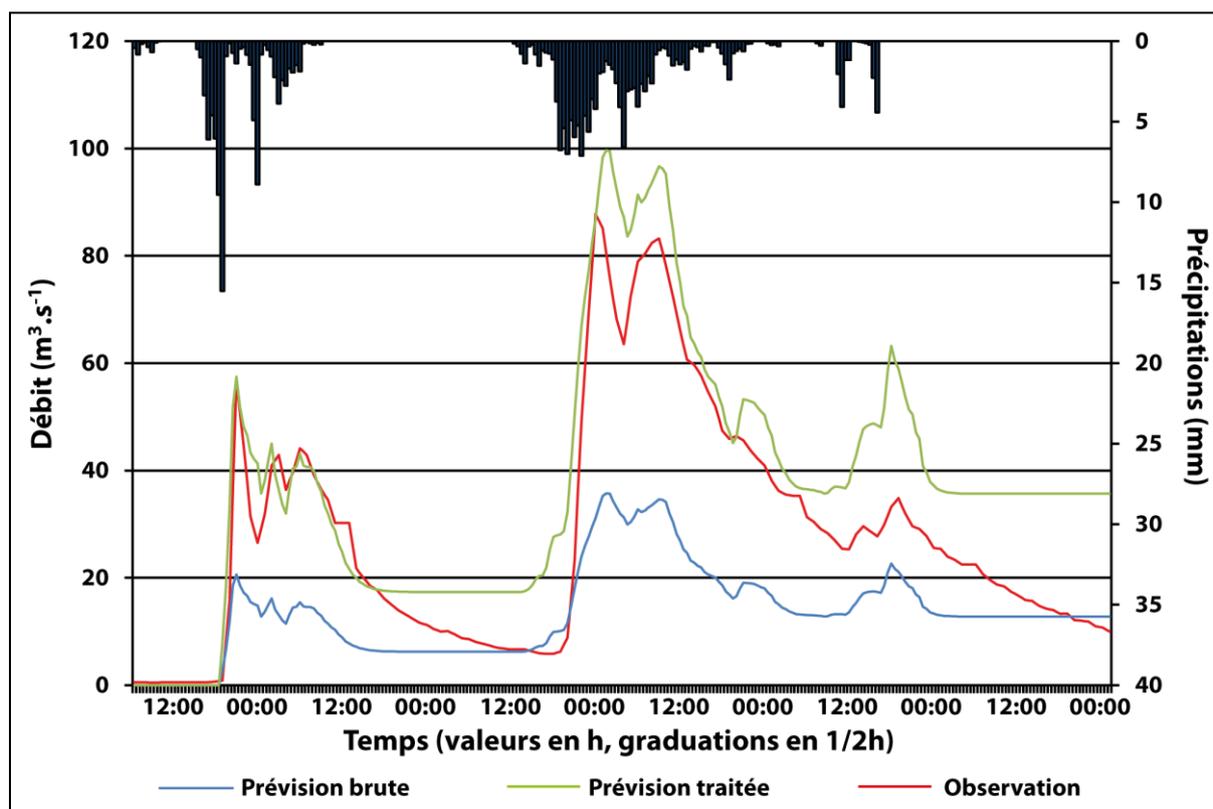


Figure 83: prévision de l'évènement d'octobre 2004 sur la Ganière à  $k+4$ .

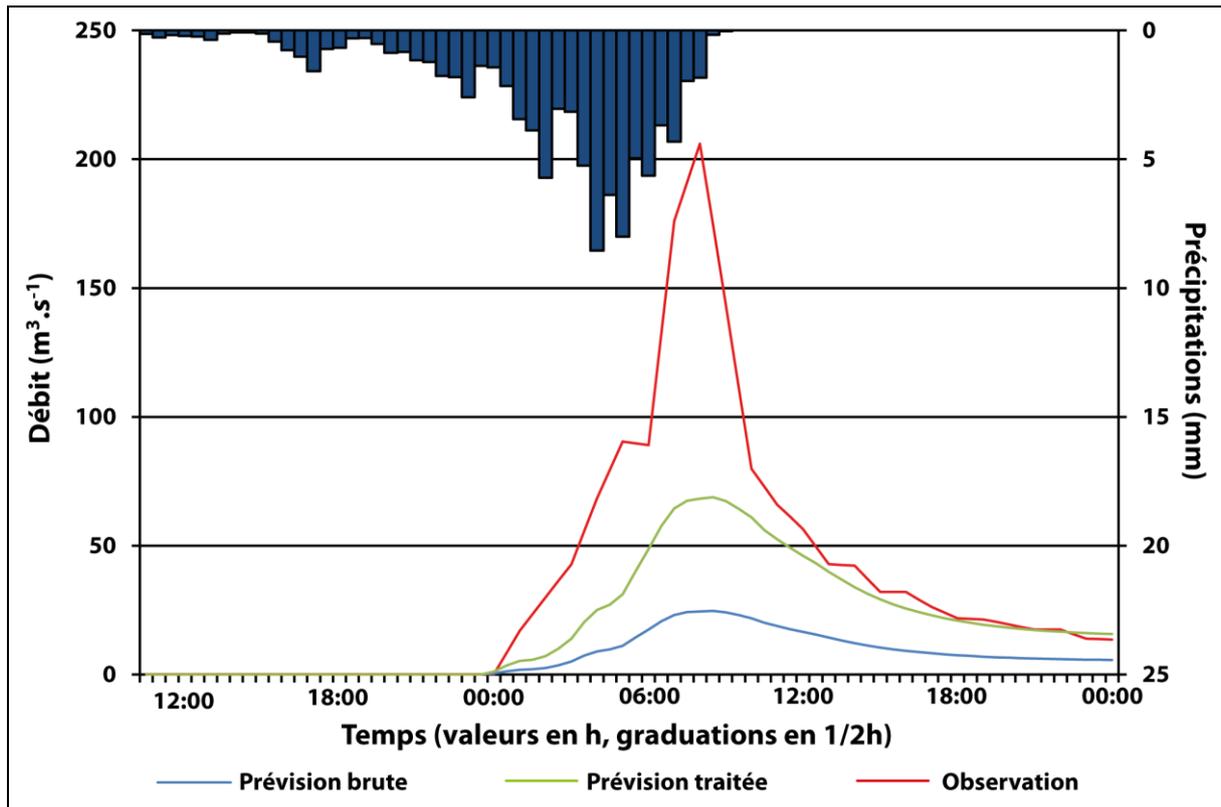


Figure 84: prévision de l'évènement de septembre 2002 sur la Ganière à k+2.

En annexe 7, un tableau détaille les résultats pour chaque évènement à chaque horizon de prévision selon les critères retenus dans cette partie.

### 5.3.7.3 Analyse des résultats

Les deux évènements présentés dans la section précédente sont assez représentatifs des cas rencontrés sur ce bassin. D'une manière générale, le modèle réagit bien, c'est-à-dire qu'il est dans les temps et dans les amplitudes adéquates, chaque fois que les variables d'entrées sont représentatives du champ de pluie réel.

En effet, pour les évènements où tous les pluviomètres sont disponibles, les résultats sont globalement satisfaisants. En revanche, les évènements où seuls un ou deux pluviomètres sont disponibles ne fournissent que des résultats mitigés, voire mauvais à l'image de l'évènement de septembre 2002 (figure 82). Sur ce dernier évènement en particulier, malgré un démarrage et un pic assez bien calés en temps, l'amplitude reste beaucoup trop faible. Seul le pluviomètre des Vans fonctionne et l'on sait que cette localité était excentrée du plus gros de l'évènement. Il est de toute façon improbable qu'une crue de  $3,75 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  ( $200 \text{ m}^3 \cdot \text{s}^{-1}$ ) se produise avec 93mm de cumul moyen en environ 18 heures sur le bassin... Ceci explique la forte sous-estimation observée. Le coefficient d'écoulement serait proche de 100%, ce qui ne s'est produit sur aucun évènement de ce bassin dans la base ni dans la zone d'étude lors de cet épisode. Comme on l'a vu dans la section 4.2.2.4, le modèle n'a pas appris de coefficients d'écoulement si forts, qui ne correspondent manifestement pas à des valeurs observables dans la zone d'étude.

Sur certains évènements à deux pluviomètres, le modèle propose de bons résultats (916 à 919) et sur d'autres d'assez larges sous-estimations ou surestimations. Dans ces cas-là et compte tenu du bon fonctionnement observé sur d'autres épisodes, l'on peut penser que le champ de pluie n'est pas bien représenté par deux des trois pluviomètres. Une estimation radar ponctuelle

serait alors nécessaire pour remplacer le pluviomètre manquant. Sur ce bassin, les cas où les résultats ne sont pas satisfaisants ne semblent donc pas liés au fonctionnement du modèle mais à des défauts de représentativité des forçages.

### 5.3.8 L'Avène

#### 5.3.8.1 Description rapide du bassin

L'Avène est un affluent du Gardon d'Alès dont le bassin mesure 57km<sup>2</sup> quasiment à la confluence, à Saint-Hilaire-de-Brethmas. Que ce soit pour le drain où les versants, c'est le bassin le moins pentu de ceux ici considérés. Cet état de fait est lié à la traversée du bassin d'Alès, quasiment plat. La figure 85 présente une carte de ce bassin.

De direction cévenole (nord-ouest au sud-est) sur un tiers amont, le cours d'eau et le bassin pivotent vers le sud-sud-ouest ce qui marque un fort impact de la tectonique sur le drainage. En effet, le changement de direction s'opère au passage de la faille des Cévennes qui entraîne aussi un changement de relief fort.

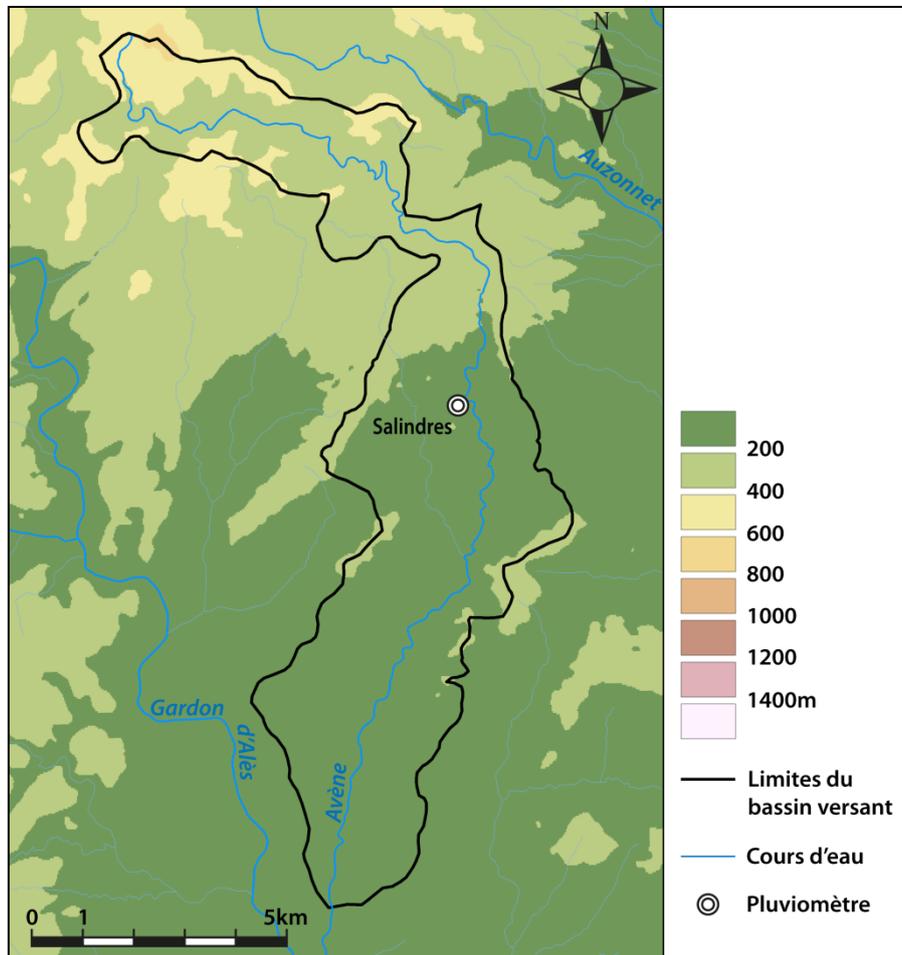


Figure 85: carte du bassin versant de l'Avène.

Une part minime de l'extrême amont est occupée par des schistes des Cévennes et des terrains carbonifères. Sur le reste de la partie située en amont de la faille des Cévennes, on retrouve des terrains secondaires liasiques, jurassiques et crétacés, plus ponctuellement au contact de la faille. La présence de terrains karstiques est très limitée en puissance et en surface. En aval de la faille des Cévennes, le fossé d'Alès présente ses marnes et ses grès à niveaux particuliers conglomératiques ou calcaires (calcaires en plaquettes). L'Avène serpente alors dans la plaine

entre quelques dômes surélevés, olistostromes (klippes sédimentaires) de hauterivien (calcaires argileux) avant de rentrer dans la plaine alluviale du Gardon.

L'occupation des sols est naturelle dans la section amont, agricole pour les trois quarts et urbanisée pour le reste dans la section aval. Entre les terrains géologiques et l'occupation des sols anthropique, l'aval constitue une zone assez peu perméable.

### 5.3.8.2 Evènements et résultats

Le seul évènement disponible sur ce bassin, qui n'est pas équipé d'un limnimètre, est issu d'une estimation du retour d'expérience de septembre 2002. On a estimé au point de l'exutoire ici considéré un débit de  $400 \text{ m}^3 \cdot \text{s}^{-1}$  soit environ  $7 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$ . Le pluviomètre Météo-France de Salindres a fourni la donnée pluviométrique pour cet épisode. Un cumul de 529mm avec des intensités maximales de 74mm/1h et 42mm/0.5h ont été relevés.

La figure 86 présente les résultats obtenus en prévision sur cet évènement à  $k+3$ .

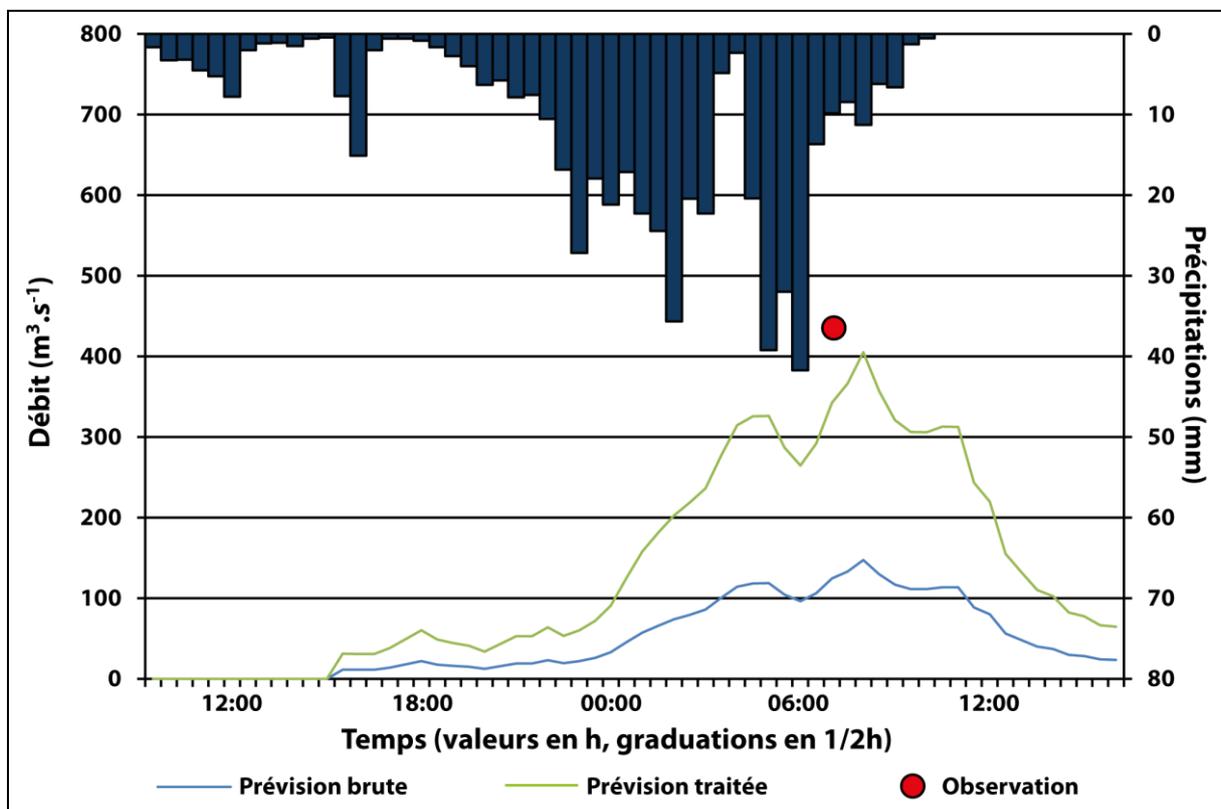


Figure 86: prévision de l'évènement de septembre 2002 sur l'Avène à  $k+3$ .

Le tableau 24 présente quant à lui les résultats détaillés selon les critères de qualité prédéterminés pour les quatre horizons de prévision.

Tableau 24: résultats détaillés en prévision de l'évènement de septembre 2002 sur l'Avène.

Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
$k+1$	Sans objet	Sans objet	79%	88%	1,00
$k+2$	Sans objet	Sans objet	80%	95%	0,50
$k+3$	Sans objet	Sans objet	86%	101%	1
$k+4$	Sans objet	Sans objet	63%	103%	1,00

### *5.3.8.3 Analyse des résultats*

Compte-tenu de l'importante incertitude associée à une estimation de débit sur un évènement important en zone peu accidentée, l'on peut considérer que les résultats sont de bonne qualité.

La quasi-totalité du débit de pointe est prévue alors que le décalage, modéré, est limité par des *SPPD* convenables. Même avec une taille relativement petite, la faiblesse des pentes a peut-être pu compenser l'écart temporel entre le bassin du Gardon à Mialet et celui de l'Avène. En tout cas, ces résultats montrent que le modèle aurait été exploitable en utilisation opérationnelle sur cet évènement et ce bassin, même avec des données de pluie réduites aux valeurs d'un seul pluviomètre. Il aurait été intéressant de pouvoir le tester sur d'autres évènements de typologie différente mais seules ces données étaient disponibles.

### **5.3.9 Le Galeizon à Cendras**

#### *5.3.9.1 Description rapide du bassin*

D'une surface de 88km<sup>2</sup>, le Galeizon est un affluent du Gardon d'Alès. Il s'écoule du nord-ouest au sud-est dans un bassin relativement allongé. Le drain fait partie des moins pentus de l'échantillon de bassins. Malgré des tronçons amont très pentus, l'aval est assez plat. En revanche, les moindres pentes des versants en aval ne parviennent pas à compenser les pentes très fortes de l'amont, pour une moyenne d'environ 30%. La figure 87 présente une carte de ce bassin.

La géologie du bassin est assez simple : un petit quart aval ajouté à une frange à l'extrême sud sont composés de terrains liasiques carbonatés et triasiques et le reste de schistes des Cévennes, ponctués de filons granitiques ou lamprophyriques, ou de bancs de quartzites.

L'occupation des sols est quant à elle presque exclusivement naturelle.

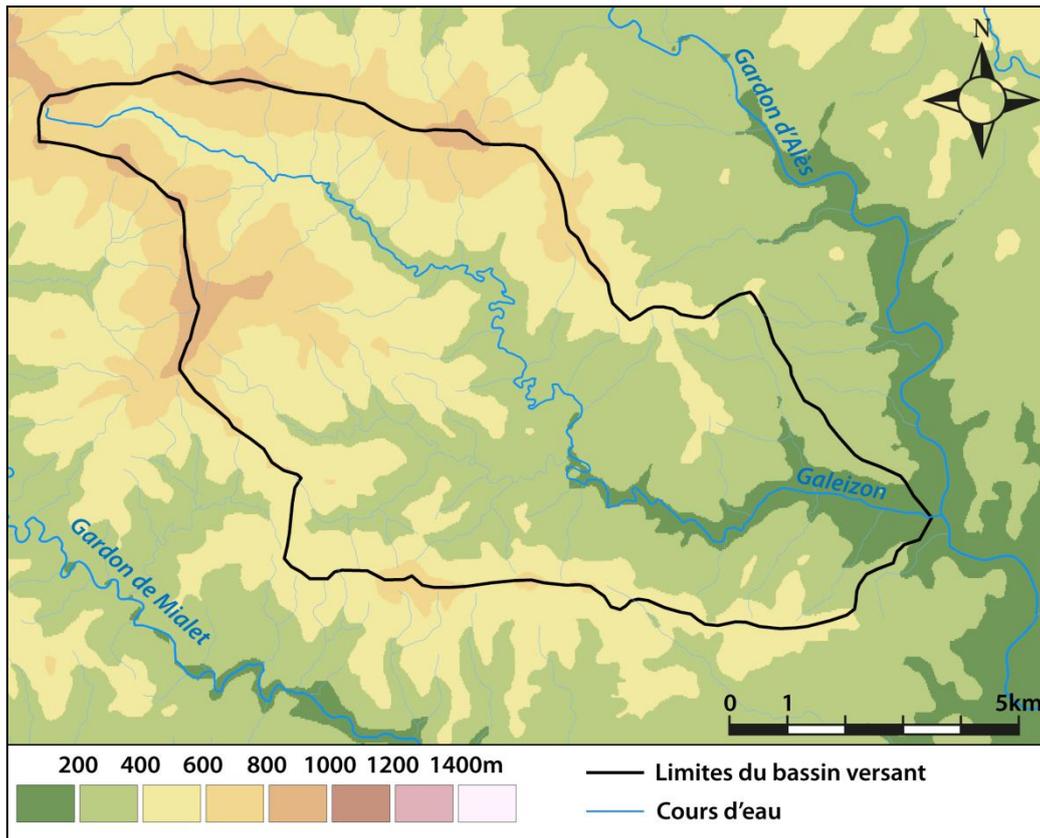


Figure 87: carte du bassin versant du Galeizon.

### 5.3.9.2 Evènements et résultats

La source des données est la même que pour les bassins de la Salandre et du Galeizon amont. Nous disposons d'un hydrogramme simulé calé sur une évaluation du débit de pointe. Les pluies sont strictement identiques à celles utilisées sur les deux bassins amont et donc probablement sousestimées. C'est donc le coefficient correctif qui devra faire la différence par rapport aux bassins de la Salandre et du Galeizon amont.

Le débit de pointe est estimé à  $970 \text{ m}^3 \cdot \text{s}^{-1}$  soit  $11 \text{ m}^3 \cdot \text{s}^{-1} \cdot \text{km}^{-2}$  ce qui est un débit spécifique très important pour un bassin de cette taille. Compte-tenu du débit brut proposé par le modèle sur les deux autres évènements alimentés par les mêmes pluies, le coefficient du Galeizon (complet) ne peut permettre d'atteindre des résultats exploitables. Le tableau 25 présente les résultats détaillés pour chaque horizon de prévision. Les hydrogrammes sont, eux, présentés en annexe 5.

Tableau 25: résultats détaillés en prévision de l'évènement de septembre 2002 sur le Galeizon.

Horizon	Nash	Persistence	SPPD	PPD	Décalage (h)
$k+1$	0.31	-27.24	33%	33%	-1.50
$k+2$	0.04	-9.15	11%	41%	-1.50
$k+3$	0.07	-3.58	13%	38%	-1.50
$k+4$	-0.02	-1.96	9%	32%	-1.00

### 5.3.9.3 Analyse des résultats

A résultats identiques, les causes de la mauvaise qualité des prévisions sont identiques, à quelques décalages temporels près. Ainsi, une pluie trop faible, si moyennée qu'elle ne représente plus les intensités fortes, et un débit spécifique extrême n'ont pas permis au modèle d'atteindre les résultats escomptés.

### 5.3.10 Le Gardon à Saumane

#### 5.3.10.1 Description rapide du bassin

Le Gardon de Saint-Jean à Saumane mesure 100km<sup>2</sup>. En aval de Saumane, après avoir traversé Saint-Jean-du-Gard et reçu l'apport de la Salindrenque, il conflue avec le Gardon de Mialet pour donner le Gardon d'Anduze, quelques kilomètres en amont de sa ville éponyme. Il s'étire du nord-ouest au sud-est dans des pentes fortes, de drain comme de versants. Cette dernière valeur, de 36%, est d'ailleurs la plus forte de l'échantillon de bassins. La figure 88 représente une carte de ce bassin.

La géologie du bassin est très simple : les schistes des Cévennes sont omniprésents à l'exception de quelques filons granitiques à l'extrême amont, qui jouxtent les terrains karstiques secondaires (d'étendue et de puissance très limitées) de la partie ouest de la *can* de l'Hospitalet.

L'occupation des sols n'est pas plus complexe : naturelle, avec des forêts de châtaigniers et des garrigues pour l'essentiel, elle ne laisse qu'une place tout à fait mineure aux occupations artificielles (Saumane surtout). En somme, ce bassin est non seulement voisin de celui du Gardon à Mialet, mais aussi assez analogue à la taille près.

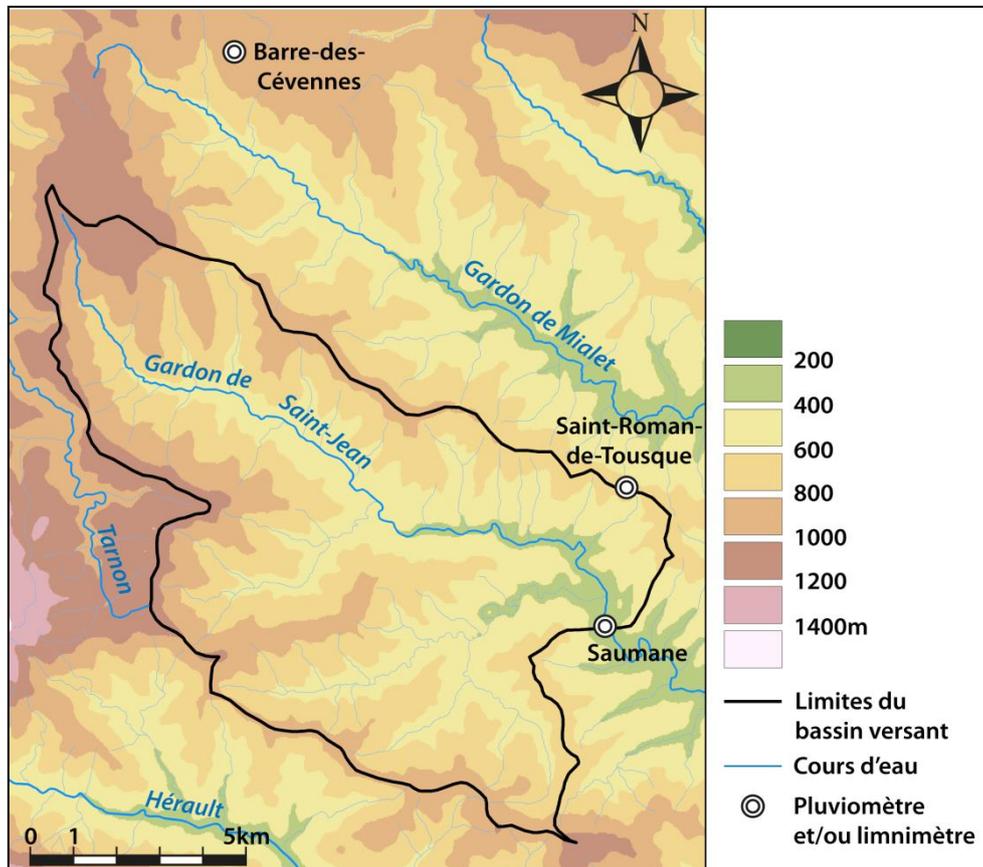


Figure 88: carte du bassin du Gardon à Saumane.

#### 5.3.10.2 Evènements et résultats

Trois évènements d'intensités différentes ont été extraits de la base de données BVNE. Le limnimètre couplé à un pluviomètre et disposé à Saumane a permis, grâce à la courbe de tarage correspondante d'obtenir les débits sur les évènements. Les pluviomètres utilisés sont les mêmes que pour le bassin de Mialet, voisin, à l'exception de celui de Saumane, qui remplace celui de Mialet dans le modèle initial. Seul le pluviomètre de Barre-des-Cévennes est à l'extérieur du

bassin. Il peut être considéré comme représentatif des zones amont du bassin du Gardon d'Anduze. Les données de pluie sont disponibles à 100% pour tous les évènements. Ces évènements sont synthétiquement présentés dans le tableau 26.

Tableau 26: description synthétique des trois évènements disponibles pour le Gardon à Saumane. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
931	sept-02	230	58	31	527	5,27	Pluvio.	3 sur 3
932	sept-00	243	32	16	212	2,12	Pluvio.	3 sur 3
933	nov-96	206	12	7	111	1,11	Pluvio.	3 sur 3

La figure 89 présente les résultats pour l'évènement de septembre 2002 sur ce bassin à  $k+4$ .

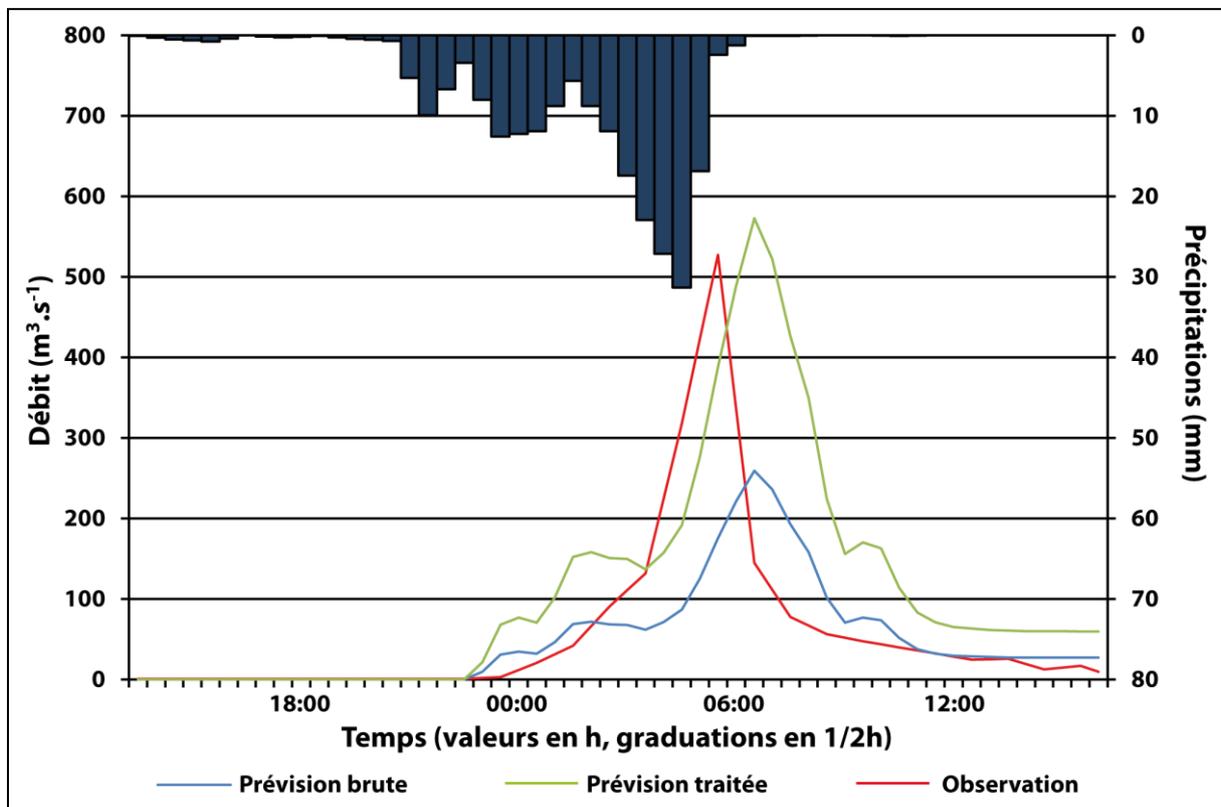


Figure 89: prévision de l'évènement de septembre 2002 sur le Gardon à Saumane à  $k+4$ .

Le tableau 27 présente les résultats détaillés pour chaque horizon de prévision et chaque évènement.

Tableau 27: résultats détaillés en prévision des événements 931, 932 et 933 sur le Gardon à Saumane.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
931	k+1	-0,35	-6,29	103%	120%	1,00
931	k+2	-0,68	-2,04	111%	131%	0,50
931	k+3	-0,84	-1,01	102%	142%	1,00
931	k+4	-0,22	0,01	74%	109%	1,00
932	k+1	-1,30	-37,05	82%	139%	2,50
932	k+2	-1,68	-11,82	84%	149%	2,00
932	k+3	-1,35	-4,71	79%	147%	2,00
932	k+4	-1,23	-2,53	74%	140%	2,50
933	k+1	0,31	-46,56	102%	130%	1,50
933	k+2	-0,14	-22,74	124%	144%	2,50
933	k+3	-0,19	-12,12	120%	141%	1,00
933	k+4	-0,17	-3,80	104%	143%	1,50

### 5.3.10.3 Analyse des résultats

Avec un champ de pluie presque commun à celui de Mialet et un bassin analogue mais plus petit, l'on pouvait s'attendre à ce que le modèle surestime les débits et soit en retard. C'est ici exactement ce qu'il se passe. Une certaine logique est respectée vis-à-vis de la configuration des deux bassins. La correction d'amplitude a joué un rôle trop important sur ce bassin. Sans elle, les résultats étaient largement sous-estimés. Avec, ils sont largement surestimés. Si les résultats ne sont pas parfaits et même s'ils présentent un léger risque de fausse alerte, ils permettent tout de même d'appréhender le débit à l'instant du pic à l'intérieur d'une marge d'erreur de 30%, ce qui est une incertitude couramment admise sur la prévision des débits.

## 5.3.11 La Baume

### 5.3.11.1 Description rapide du bassin

La Baume est un affluent de l'Ardèche dont le bassin mesure 210km<sup>2</sup>, très pentu avec 35% de pentes moyennes et 3,8% de pente du drain, et qui prend sa source sur les contreforts du Mont-Tanargue. La figure 90 présente une carte du bassin.

Assez typiquement cévenol, le bassin est limité au sud par un relais de la faille des Cévennes. Le quart aval du bassin est triasique (rudites, arénites, dolomies...) alors que le reste se divise entre schistes, orthogneiss, paragneiss et granites, ponctués de nombreux éléments filoniens.

L'occupation des sols est naturelle, à de rares exceptions agricoles et urbanisées.

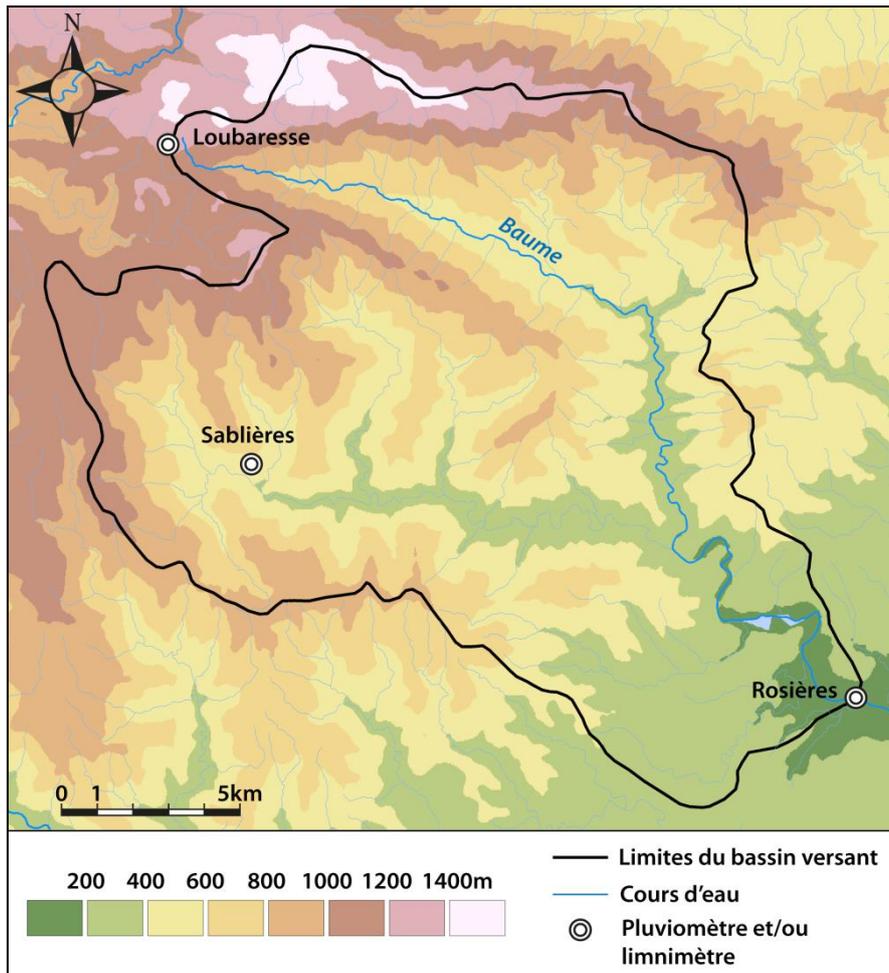


Figure 90: carte du bassin versant de la Baume.

### 5.3.11.2 Evènements et résultats

Trois évènements à Rosières issus de la base de données SevnoI de l’OHMCV sont disponibles pour ce bassin. En sus du limnimètre de Rosières, le bassin comprend trois pluviomètres assez bien répartis sur le bassin. Ils seront, d’amont en aval, utilisés dans le même ordre qu’à Mialet. La disponibilité des données est optimale pour les trois évènements.

Le tableau 28 présente une description synthétique de ces évènements, plutôt faibles pour un bassin de cette dimension.

Tableau 28: description synthétique des trois évènements disponibles pour la Baume. RR représente le cumul moyen,  $I_{xl}$  l’intensité maximale moyenne en 1h,  $I_{xc}$  l’intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d’appareils disponibles pour l’évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
904	nov-00	177	10	5	212	1,01	Pluvio.	3 sur 3
905	oct-01	254	26	16	330	1,57	Pluvio.	3 sur 3
906	oct-04	252	10	6	230	1,10	Pluvio.	3 sur 3

La figure 91 présente les résultats en prévision pour l’évènement d’octobre 2004 sur la Baume à  $k+4$  (soit 2h).

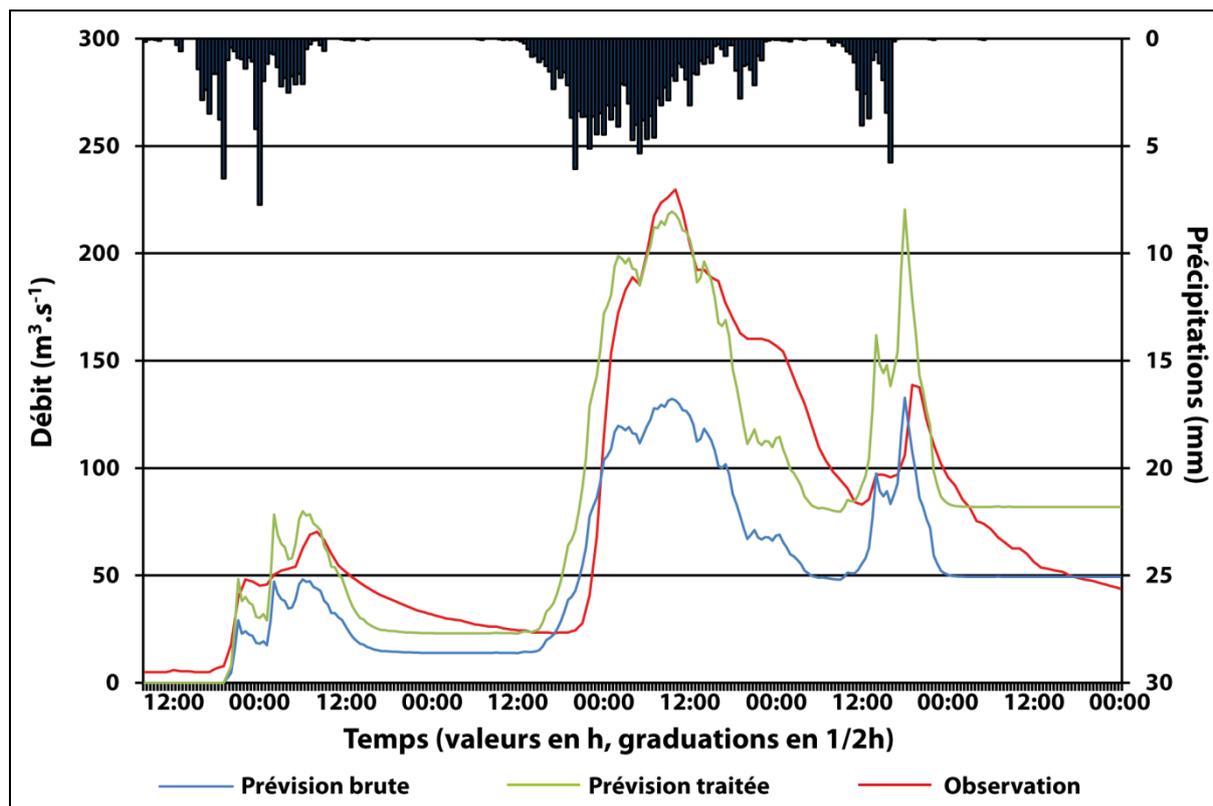


Figure 91: prévision de l'évènement d'octobre 2004 sur la Baume à Rosières à  $k+4$ .

Le tableau 29 présente quant à lui les résultats pour tous les évènements et horizons explorés.

Tableau 29: résultats détaillés en prévision des évènements 904, 905 et 906 sur la Baume.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
904	$k+1$	0,53	-49,60	93%	117%	-1,00
904	$k+2$	0,71	-7,29	95%	106%	-1,00
904	$k+3$	0,70	-2,91	97%	106%	-1,00
904	$k+4$	0,70	-1,35	104%	116%	-1,00
905	$k+1$	-0,29	-29,85	189%	194%	-0,50
905	$k+2$	0,57	-2,04	130%	130%	0,00
905	$k+3$	0,50	-0,87	128%	128%	0,00
905	$k+4$	0,42	-0,43	143%	143%	0,00
906	$k+1$	0,77	-44,62	96%	100%	-1,00
906	$k+2$	0,85	-7,28	89%	91%	-0,50
906	$k+3$	0,83	-3,35	91%	91%	-1,00
906	$k+4$	0,80	-1,96	95%	96%	-0,50

### 5.3.11.3 Analyse des résultats

Les résultats montrent des décalages temporels faibles, ce qui est conforme à ce que l'on pouvait attendre d'un bassin de surface proche de celle du Gardon de Mialet. En revanche, les résultats sont plus mitigés au niveau de l'amplitude. Si l'évènement 906 est très bien prévu jusqu'à  $k+4$ , les évènements 904 et 905 sont plus délicats : le premier souffre d'une sous-estimation du premier pic alors que le second pic (principal) est bien prévu par le modèle ; le second souffre d'une importante surestimation du second pic (principal).

Dans l'ensemble, la dynamique du bassin est relativement bien reproduite mais la prise en compte du cumul courant sur l'épisode conduit à des problèmes de sur ou sous-estimation sur les évènements à plusieurs pics. Ceci oblige à considérer les très bons résultats de l'évènement 906 comme peu robustes. Plus généralement, il est clair que l'échantillon d'évènements n'est pas suffisant pour conclure au bon ou au mauvais fonctionnement du modèle sur ce bassin.

### 5.3.12 L'Ardèche

#### 5.3.12.1 Description rapide du bassin

L'Ardèche est un long cours d'eau au vaste bassin tributaire du Rhône. Ici, nous considérerons son extrême amont, à la station de Pont-de-Labeaume, où le bassin ne fait que 280km<sup>2</sup>. Les pentes du drain (4,1%) et des versants (35%) sont très fortes. La rivière prend sa source au nord du Mont-Tanargue, entre Plateau-Ardéchois et Margeride. La figure 92 présente une carte du bassin.

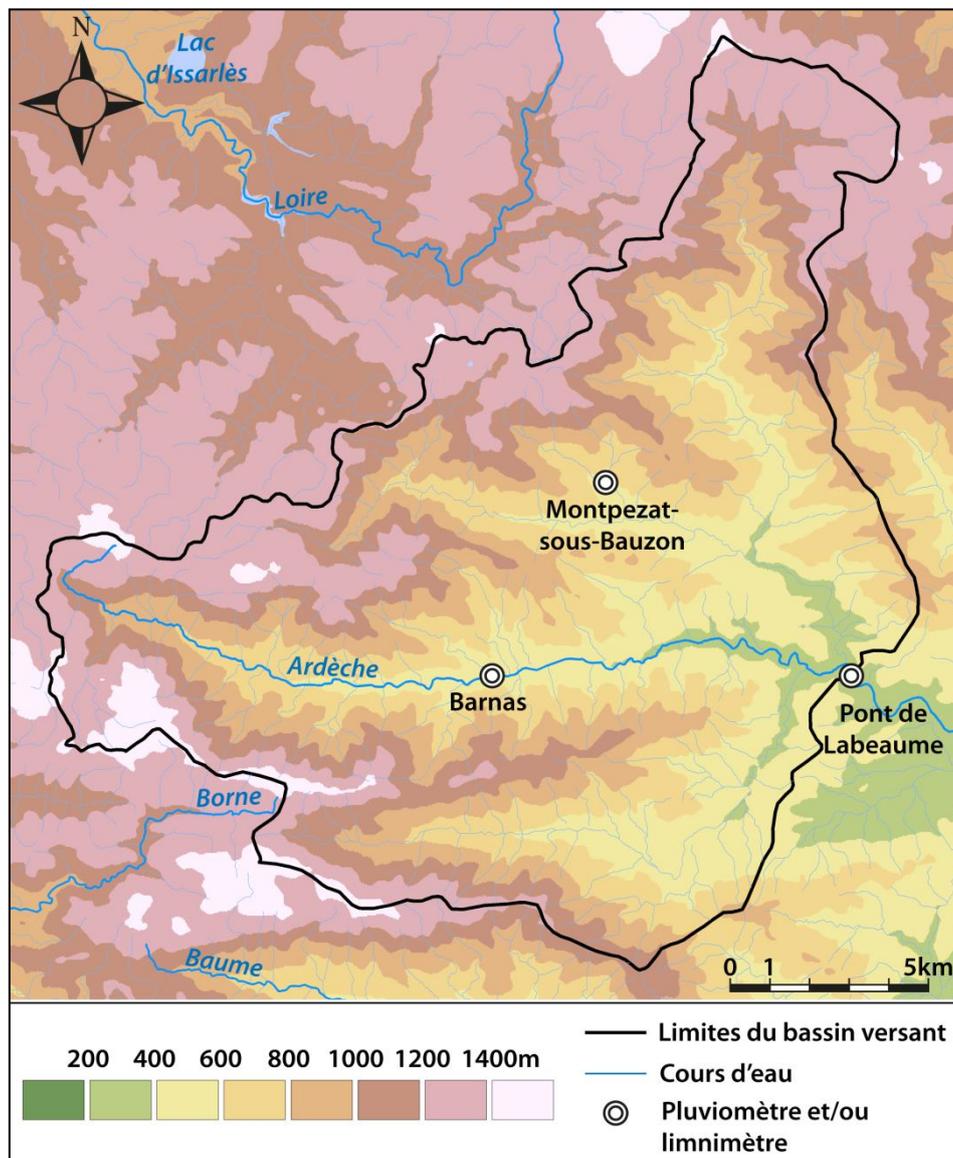


Figure 92: carte du bassin versant de l'Ardèche à Pont-de-Labeaume.

D'un point de vue géologique, le bassin est assez simple : à l'exception de quelques zones recouvertes de basaltes, les terrains sont métamorphiques (schistes, gneiss, leptynites,

nébulites, quartzites...) ou magmatiques (granites). L'occupation des sols est, ici aussi, naturelle en très grande majorité.

### 5.3.12.2 Evènements et résultats

Trois évènements ont été extraits de la base de données Sevnol de l'OHMCV pour ce bassin. Les débits sont fournis à l'exutoire, à Pont-de-Labeaume, de même que la pluie. Deux autres pluviomètres étaient disponibles à 100% sur les épisodes considérés : Barnas et Montpezat-sous-Bauzon. Barnas sera considéré comme le pluviomètre amont (Barre-des-Cévennes), Montpezat comme le pluviomètre central (Saint-Roman-de-Tousque) et Pont-de-Labeaume comme le pluviomètre aval (Mialet).

Le tableau 30 présente une description des évènements étudiés ici.

Tableau 30: description synthétique des trois évènements disponibles pour l'Ardèche. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
901	déc-00	230	14	8	434	1,55	Pluio.	3 sur 3
902	oct-01	274	37	21	494	1,76	Pluio.	3 sur 3
903	déc-03	322	15	9	484	1,73	Pluio.	3 sur 3

La figure 93 présente les résultats obtenus en prévision sur l'évènement n°902 à  $k+3$ .

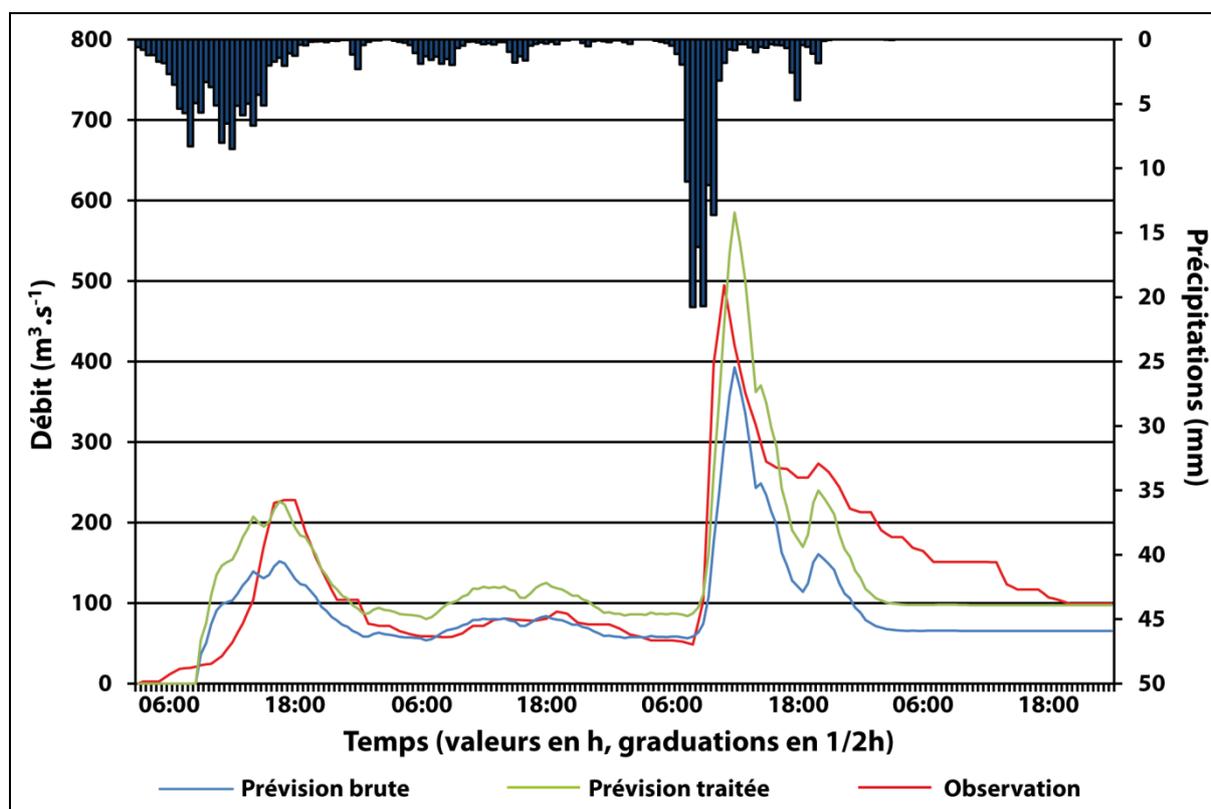


Figure 93: prévision de l'évènement d'octobre 2001 à un horizon de  $k+3$

Le tableau 31 présente quant à lui les résultats détaillés pour chaque évènement compte tenu des critères prédéfinis.

Tableau 31 : résultats détaillés en prévision des évènements 901, 902 et 903 sur l'Ardèche.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
901	k+1	-0,27	-171,66	78%	86%	-1,00
901	k+2	-0,32	-51,65	65%	66%	1,50
901	k+3	-0,23	-24,31	67%	68%	-0,50
901	k+4	-0,22	-14,95	76%	76%	0,00
902	k+1	0,53	-11,00	155%	155%	0,00
902	k+2	0,73	-1,03	88%	113%	1,00
902	k+3	0,70	-0,14	90%	118%	1,00
902	k+4	0,60	0,01	98%	122%	1,00
903	k+1	0,26	-148,52	66%	67%	-1,00
903	k+2	0,20	-47,04	59%	60%	0,50
903	k+3	0,25	-22,56	61%	61%	1,00
903	k+4	0,25	-14,04	59%	60%	0,50

### 5.3.12.3 Analyse des résultats

Les résultats sur les évènements 901 et 903 ne sont pas de très bonne qualité alors qu'ils sont beaucoup plus exploitables pour l'évènement 902. On remarque que pour les deux évènements où la prévision est délicate, les débits initiaux observés sont de plusieurs dizaines de mètres cube par seconde ce qui est une caractéristique de l'Ardèche. Ce cours d'eau est en effet capable de soutenir des débits assez élevés pendant de longues périodes, permettant parfois à des évènements de débiter à près de 100 m<sup>3</sup>.s<sup>-1</sup>. Ce n'est que très rarement le cas sur le Gardon à Mialet qui se tarit beaucoup plus vite. La prise en compte de cette information est, on le sait dans le cadre des bassins non jaugés, impossible de façon directe.

Avec toutes les réserves nécessaires lorsque seuls trois évènements sont testés, on peut penser que ce type de régime hydrologique n'est pas bien appréhendé par le modèle. Comme pour le bassin de Saumane, ces prévisions peuvent cependant trouver leur utilité dans le cadre de la prévision sur des bassins non jaugés.

## 5.3.13 Le Chassezac

### 5.3.13.1 Présentation rapide du bassin

Le Chassezac est un bassin de 500km<sup>2</sup> tributaire de l'Ardèche qui prend sa source en Margeride. Grossi par les cours de la Borne qui prend sa source à proximité de la source de l'Ardèche et par l'Altier qui prend sa source non loin du sommet du Mont-Lozère, il s'écoule en direction du sud-est. Ces affluents donnent au bassin une forme peu commune de sorte qu'il fait le lien entre les massifs du Tanargue et du Mont-Lozère du sud-ouest au nord-est et entre les rudes plateaux lozériens et la clémente plaine de la basse Ardèche du nord-ouest au sud-est. Le drain est peu pentu et les versants conformes à la moyenne des bassins de l'échantillon. Cinq barrages, à la vocation essentiellement hydroélectrique sont situés sur le bassin : trois sur le Chassezac, un sur la Borne et un sur l'Altier. La capacité d'écrêtement de ces barrages est très limitée et selon le règlement d'information sur les crues de même que selon le site de la communauté de communes de Villefort, leur vocation secondaire (au-delà de l'hydroélectricité) tient plus au soutien d'étiage qu'à l'écrêtement des crues. La figure 94 présente une carte de ce bassin.

Géologiquement, on retrouve majoritairement des terrains cristallins : schistes des Cévennes et granites, ponctués de nombreux filons. Quelques terrains triasiques gréseux sont perchés sur les plateaux au nord-est de Villefort, mais restent marginaux tant en surface qu'en puissance.

L'occupation des sols est, encore une fois, presque exclusivement forestière ou agricole avec des pâturages et des prairies, très proches de l'état naturel.

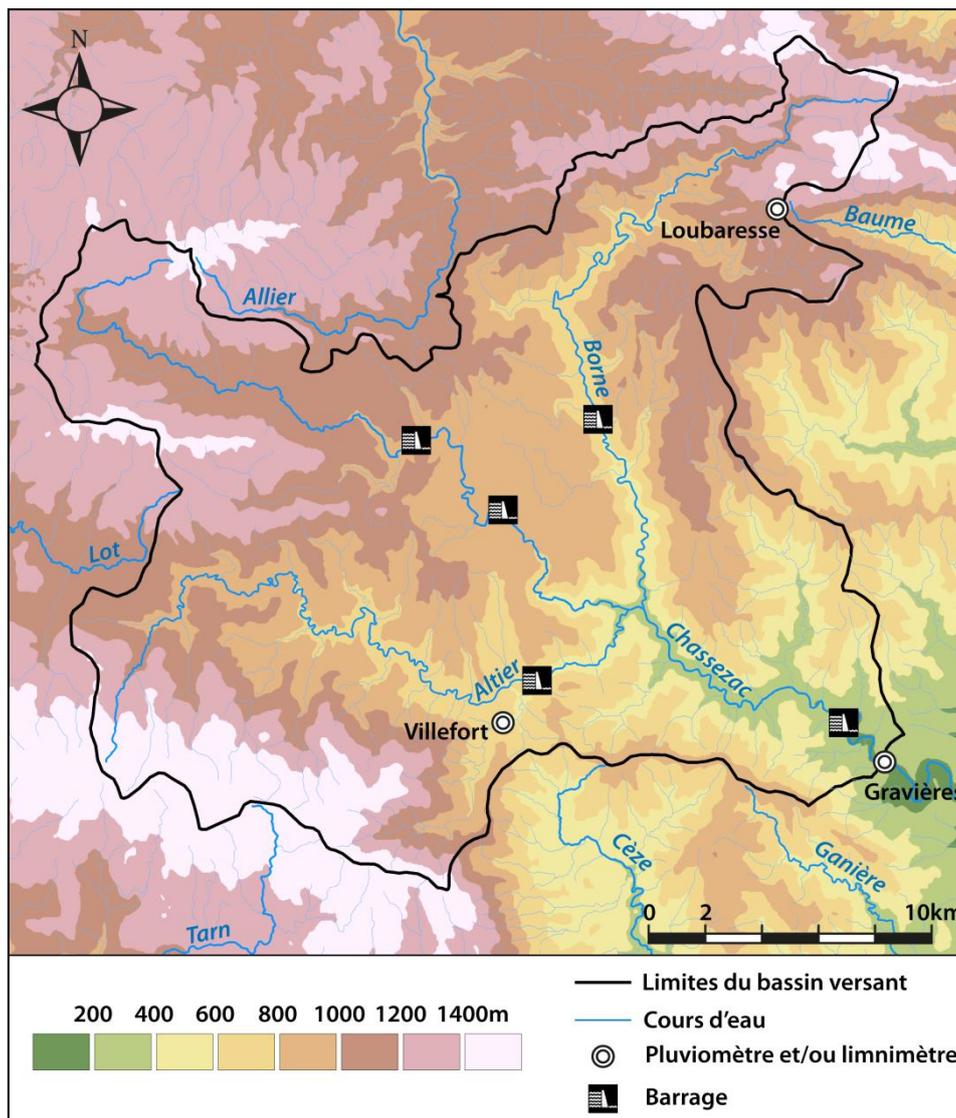


Figure 94 : carte du bassin versant du Chassezac.

### 5.3.13.2 Evènements et résultats

Trois évènements ont été extraits de la base de données SevnoI de l'OHMCV pour ce bassin. Les données de débit sont fournies pour la station limnimétrique de Gravières. Les pluies sont disponibles à 100% pour Gravières (correspondant à Mialet), pour Villefort (correspondant à Saint-Roman-de-Tousque) et pour Loubaresse (correspondant à Barre-des-Cévennes).

Il faut noter que sur ce bassin, qui englobe une part des versants nord du Mont-Lozère et de la Montagne-du-Goulet, ainsi que l'extrémité méridionale de la Margeride, la fréquence de chutes de neige abondantes sur les hauteurs peut être assez élevée suivant les hivers. Etant donné que les pluviomètres se situent tous trois dans des zones peu impactées par ces éventuelles chutes de neige en flux méditerranéen et que les pluies méditerranéennes ne dépassent guère le bassin supérieur, on peut considérer l'influence de la neige comme négligeable en première approximation. D'ailleurs, les épisodes présentés dans le tableau 32 se produisent à des dates peu favorables aux épisodes neigeux.

Tableau 32 : description synthétique des trois évènements disponibles pour le Chassezac. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
907	sept-00	189	17	10	564	1,13	Pluvio.	3 sur 3
908	oct-01	313	25	15	635	1,27	Pluvio.	3 sur 3
909	nov-03	230	14	9	728	1,46	Pluvio.	3 sur 3

La figure 95 présente un hydrogramme en prévision de l'évènement 907 à  $k+4$ .

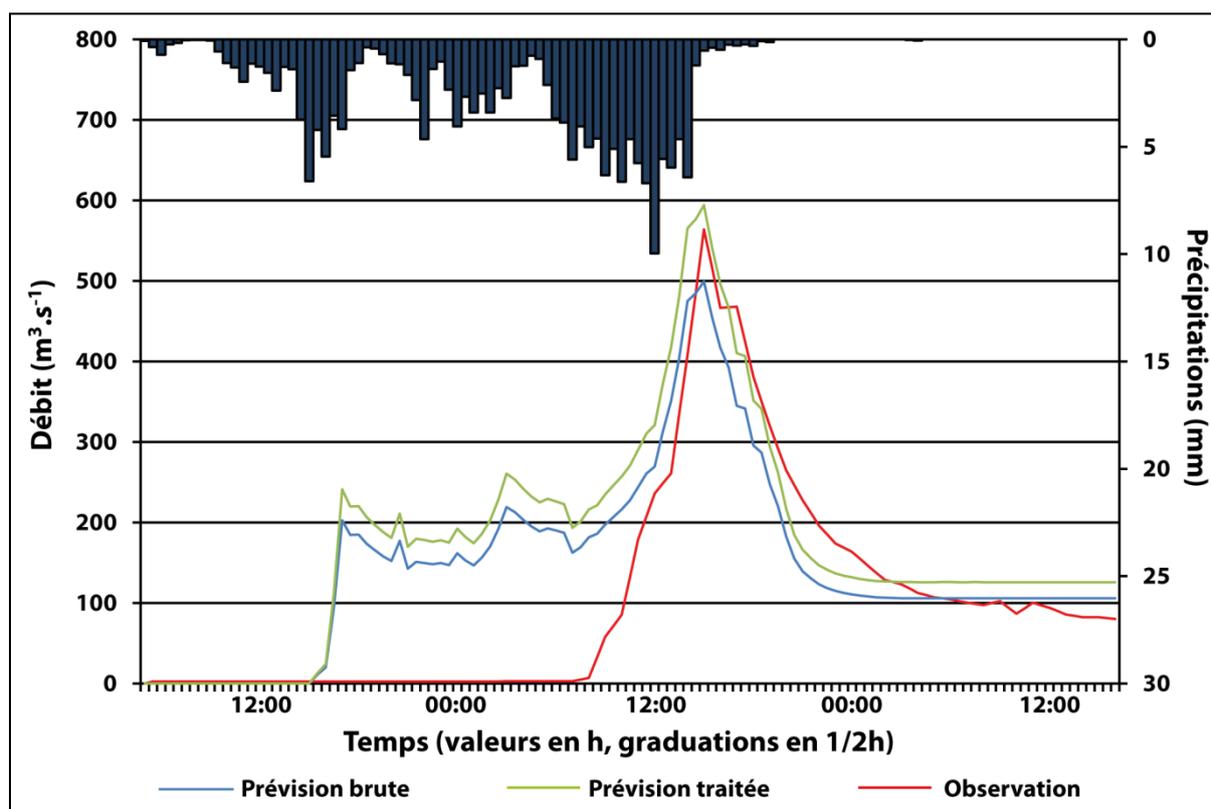


Figure 95: prévision de l'évènement de septembre 2000 à  $k+4$ .

Le tableau 33 présente quant à lui les résultats détaillés pour chaque évènement compte tenu des critères prédéfinis.

Tableau 33: résultats détaillés en prévision des évènements 907, 908 et 909 sur le Chassezac.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
907	k+1	0,51	-22,42	104%	105%	-0,50
907	k+2	0,38	-7,68	90%	93%	-0,50
907	k+3	0,40	-3,39	93%	96%	-0,50
907	k+4	0,24	-2,49	105%	105%	0,00
908	k+1	-0,19	-70,61	78%	149%	-5,00
908	k+2	-0,02	-18,32	78%	105%	-4,00
908	k+3	0,00	-9,63	80%	106%	-4,00
908	k+4	-0,09	-6,88	74%	121%	-4,00
909	k+1	0,59	-79,46	53%	55%	-1,00
909	k+2	0,62	-19,19	50%	50%	-0,50
909	k+3	0,63	-8,24	52%	52%	0,00
909	k+4	0,60	-5,13	53%	53%	0,00

### 5.3.13.3 Analyse des résultats

Au sens de la prévision du pic les évènements 907 et 908 ont été bien prévus dans l'ensemble. En revanche, dans chacun des cas, un problème évident de prise en compte des conditions initiales se présente : l'humidité des sols n'est pas bien appréhendée. La hausse du débit est trop rapide du côté du modèle. Ceci tient peut-être au fait que les barrages, même de capacité modérée, peuvent significativement retarder la survenance de la base du pic. Quant à l'évènement 909, bien qu'étant le plus intense, il ne comporte pas d'intensité forte et un cumul assez bas. Par manque de forçages structurés, le modèle ne parvient pas à obtenir une amplitude suffisante. Les résultats sont en l'état difficilement exploitables en conditions opérationnelles mais peuvent être utiles sous réserve de confirmer par des tests sur d'autres évènements que les barrages affectent bel et bien le début de l'hydrogramme et que le pic est assez systématiquement bien prévu. Il pourrait aussi être bon, si la tendance au stockage se confirmait, de comparer la différence entre les volumes réellement écoulés et les volumes écoulés prévus avec la capacité de stockage éventuelle de ces barrages.

### 5.3.14 Le Gardon à Anduze

#### 5.3.14.1 Description rapide du bassin

Le Gardon à Anduze possède un bassin de 545km<sup>2</sup> et est issu de la confluence entre le Gardon de Mialet et le Gardon de Saint-Jean (coté Saumane). Lui-même conflue avec le Gardon d'Alès pour donner, à Ners, le Gard ou Gardon. Peu pentu pour ce qui est du drain, il l'est plus que la moyenne pour ses versants, ce qui est essentiellement dû à sa partie amont. De forme relativement équilibrée, l'écoulement général s'y fait du nord-ouest au sud-est, comme pour la plupart des cours d'eau cévenols dont on peut considérer qu'il est un archétype. La figure 96 présente une carte de ce bassin.

Géologiquement, le bassin à Anduze est limité au sud-est par la faille des Cévennes. Un quart aval du bassin est pour l'essentiel composé de roches carbonatées du secondaire, très compartimentées par le faisceau de failles des Cévennes à l'aval. Certaines parties peu étendues du jurassique supérieur et plus étendues du lias sont karstiques (karst de Mialet) mais les séries concernées sont découpées par les vallées et souvent perchées de sorte que leur influence sur l'hydrologie du bassin dans sa globalité semble assez limitée. Le reste du bassin, à l'exception des petits causses à l'extrême amont (vus dans les bassins à Saumane et Mialet) est occupé par

des schistes des Cévennes (micaschistes pour la plupart) ponctués de filons, des gneiss et des granites. L'occupation des sols est très majoritairement naturelle avec une importante proportion de forêts.

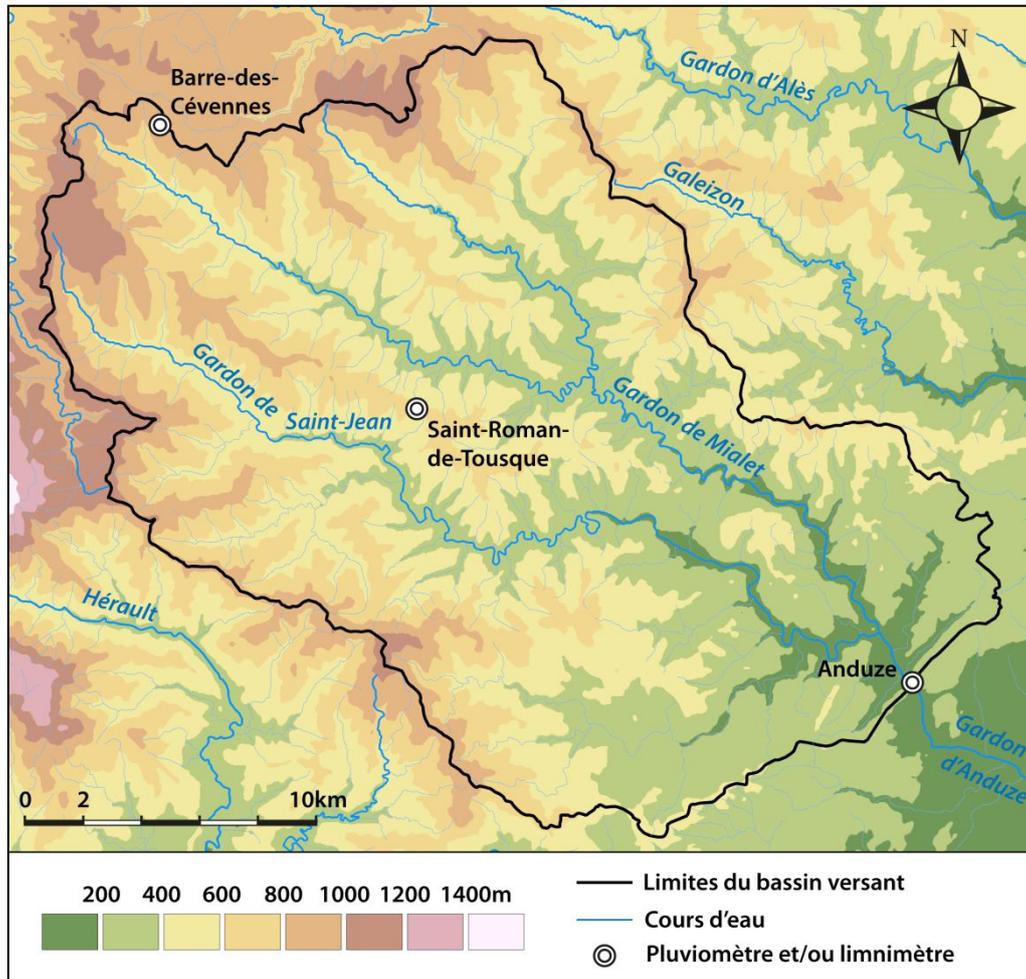


Figure 96: carte du bassin versant du Gardon d'Anduze.

#### 5.3.14.2 Evènements et résultats

Deux évènements issus de la base de données BVNE ont été utilisés pour ce bassin. Le limnimètre d'Anduze et la courbe de tarage associée permettent de donner les débits. Les pluviomètres utilisés sont ceux d'Anduze, de Saint-Roman-de-Tousque et de Barre-des-Cévennes. Par rapport au modèle initial, seul Mialet change pour être remplacé par Anduze. Cette configuration doit pouvoir permettre de tirer parti de la relation causale partielle qui lie le débit à Mialet au débit à Anduze tout en prenant en compte trois zones dans le bassin : une zone aval, une zone centrale et une zone amont. Cette vision n'est pas dénuée de sens dans la mesure où la majorité des épisodes de type cévenol montre une organisation sud-ouest / nord-est en bandes souvent relativement étroites. Elle permet en tout cas de limiter le nombre de pluviomètres (trois autres sont disponibles sur le bassin). Les évènements traités sont synthétiquement décrits dans le tableau 34.

Tableau 34: description synthétique des deux évènements disponibles pour le Gardon à Anduze. RR représente le cumul moyen,  $I_{xl}$  l'intensité maximale moyenne en 1h,  $I_{xc}$  l'intensité maximale moyenne en 0.5h,  $Q_x$  le débit de pointe,  $q_x$  le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement.

n°	Date	RR (mm)	$I_{xl}$ (mm/h)	$I_{xc}$ (mm/0,5h)	$Q_x$ ( $m^3.s^{-1}$ )	$q_x$ ( $m^3.s^{-1}.km^{-2}$ )	Pluie	Dispo.
934	sept-02	339	46	23	2742	5,03	Pluvio.	3 sur 3
935	sept-00	198	22	11	801	1,47	Pluvio.	3 sur 3

La figure 97 présente l'hydrogramme de prévision de l'évènement 934 à  $k+3$ .

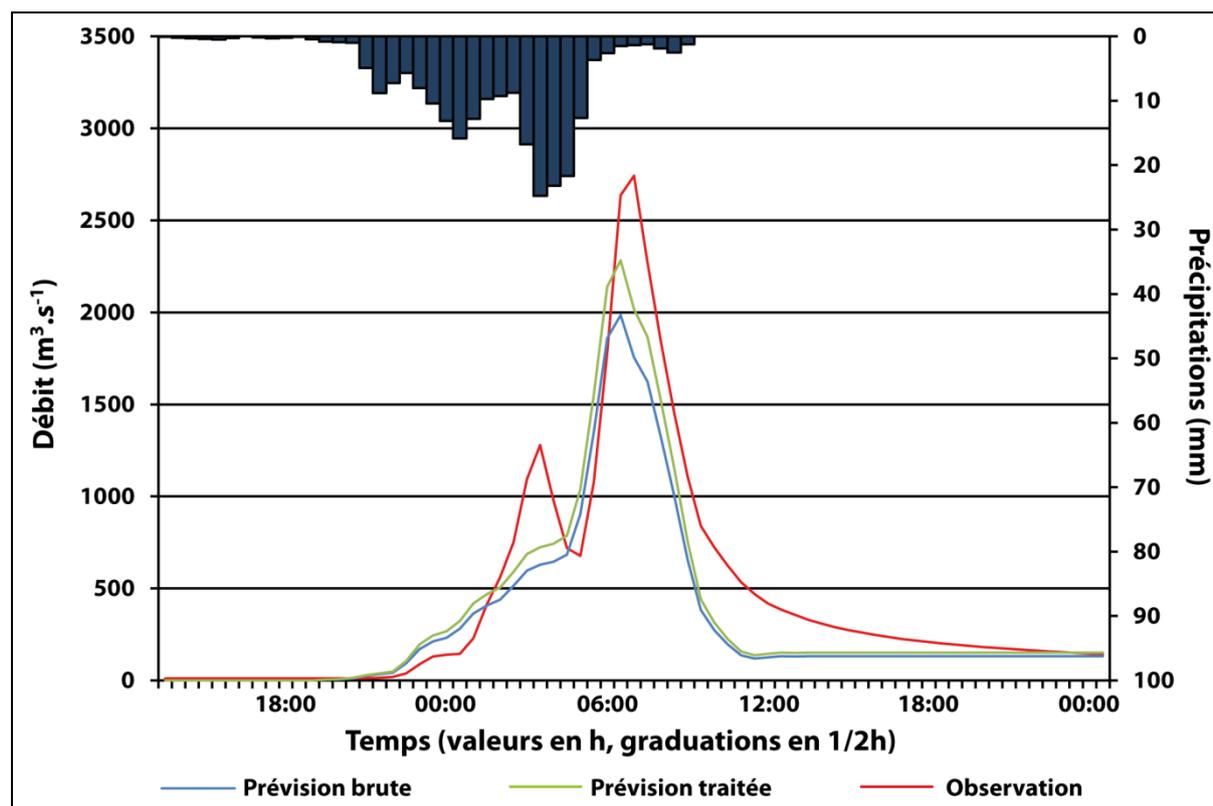


Figure 97: prévision de l'évènement de septembre 2002 à  $k+3$ .

De même, le tableau 35 présente les résultats détaillés des deux évènements pour les quatre horizons de prévision selon les critères prédéfinis.

Tableau 35: résultats détaillés en prévision des évènements 934 et 935 sur le Gardon à Anduze.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
934	$k+1$	0,45	-4,52	58%	108%	-1,50
934	$k+2$	0,83	0,49	67%	72%	-1,00
934	$k+3$	0,87	0,79	74%	83%	-0,50
934	$k+4$	0,78	0,73	57%	60%	-0,50
935	$k+1$	0,72	-8,72	106%	123%	-1,00
935	$k+2$	0,80	-0,83	83%	111%	-1,50
935	$k+3$	0,79	0,06	85%	111%	-1,00
935	$k+4$	0,75	0,33	80%	94%	-1,00

### 5.3.14.3 Analyse des résultats

Avec cette configuration semblable à celle observée à Mialet ainsi que la relation qui lie le débit à Mialet au débit à Anduze, l'on pouvait s'attendre à observer les résultats que l'on observe ici à savoir des prévisions systématiquement en avance et assez raisonnablement sous-estimées au moment du pic.

Ces prévisions issues d'une vision très simpliste du bassin du Gardon à Anduze mériteraient d'être confirmées sur plusieurs autres évènements mais, avec des biais relativement constants, elles pourraient s'avérer intéressantes en conditions opérationnelles.

## 5.4 Conclusions

Nous avons pu voir dans cette partie qu'il était réaliste de généraliser les prévisions issues de l'apprentissage statistique à des bassins non jaugés de caractéristiques assez diverses. Dans cette conclusion, après une analyse générale, nous détaillerons les apports et les limites de cette méthode.

### 5.4.1 Synthèse des résultats

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 subjective 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, en ne perdant pas de vue que le prévisionniste, en conditions opérationnelles, ne dispose à aucun moment du débit observé.

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 objective, comme dans le chapitre 4, nous privilégions le *SPPD* pour fournir une estimation de la qualité des résultats. Le tableau 36 présente d'une part les *SPPD* moyens et médians et d'autre part le taux d'évènement 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 évènements 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 évènements donnent lieu à des prévisions de 70 à 130% du pic au moment où il survient.

Tableau 36 : 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.

<b>SPPD sur tous les évènements</b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
SPPD moyen	81%	76%	77%	76%	78%
SPPD 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
SPPD moyen	86%	80%	82%	82%	83%
SPPD médian	84%	83%	84%	81%	83%
Taux d'évènements entre 70 et 130%	67%	64%	70%	64%	66%
<b>SPPD sur tous les évènements avec données de pluie représentatives</b>					
Horizon de prévision	0h30	1h	1h30	2h	Moyenne 4 horizons de prévision
SPPD moyen	88%	81%	82%	80%	83%
SPPD 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
SPPD moyen	94%	87%	87%	87%	89%
SPPD médian	87%	84%	86%	86%	86%
Taux d'évènements entre 70 et 130%	80%	80%	88%	80%	82%

Etant entendu que le succès d'une prévision réside dans sa capacité à permettre aux autorités (gestion de crise) et aux populations (information préventive) d'anticiper une inondation potentiellement préjudiciable à la sécurité des personnes et des biens, l'on peut considérer que les prévisions sont, dans la majorité des cas, de bonne qualité. Elles permettent d'apporter une information prédictive en moyenne assez efficace au prévisionniste, à l'heure où celui-ci n'en dispose pas.

#### 5.4.2 Analyse générale

Dans la grande majorité des cas, les résultats montrent que le choix du bassin du Gardon à Mialet comme bassin pilote était judicieux. En effet, sur la quasi-totalité des évènements, même ceux avec plusieurs pics, la dynamique hydrologique est respectée. Si quelques excès ou déficits d'amplitude peuvent être observés, il n'en reste pas moins que la relation causale entre pluie et débit apprise sur le Gardon à Mialet est utilisée à bon escient. Ainsi, la représentativité de ce bassin apparaît adéquate dans l'ensemble.

Les coefficients correctifs élaborés avec la méthode de prédétermination des débits spécifiques de pointe ont apporté une amélioration significative. Ils ont permis de corriger :

- d'une part la relative faiblesse du débit spécifique le plus important observé à Mialet durant la période qui a servi à établir la base de données (1992-2008),
- d'autre part, les différences de caractéristiques topographiques et en particulier de surface des bassins cibles.

Le tableau 37 présente l'augmentation en valeur des critères de Nash, *PPD* et *SPPD* obtenue entre la prévision brute et la prévision traitée.

Tableau 37 : augmentation en valeur des critères de Nash, du *PPD* et du *SPPD* pour l'ensemble des évènements testés avec le traitement appliqué.

Horizon	Nash	<i>PPD</i>	<i>SPPD</i>
<b>0h30</b>	0,15	47%	41%
<b>1h</b>	0,22	46%	39%
<b>1h30</b>	0,16	47%	40%
<b>2h</b>	0,14	48%	39%

Les difficultés à extrapoler aux évènements très intenses sont plutôt limitées aux petits bassins. Ces bassins (moins de 30km<sup>2</sup>) donnent de moins bons résultats que les autres (tableau 36). Dans le cas de la Salandre (tout comme sur le Galeizon amont et complet), on sait que les forçages sont trop faibles (lissés par une moyenne géographique du radar en zone accidentée). Mais dans le cas du ruisseau du Lyonnais, de l'Amous et de l'Alzon, trois hypothèses subsistent :

- La première est que les estimations radar ponctuelles seraient sous-estimées et que le modèle, ayant appris sur des pluviomètres, ne peut pas atteindre de forts débits avec des forçages plus faibles que ceux qui doivent entraîner une telle réaction avec des pluviomètres. Si l'idée est séduisante puisqu'appuyée par des vérifications statistiques (Visserot, 2012, entre autres), il n'en reste pas moins que sur l'Auzonnet, les résultats obtenus avec une estimation radar ponctuelle du même type sont de bonne qualité ;
- La seconde serait que, les estimations de débit du retour d'expérience étant incertaines, le débit prévu est potentiellement plus proche de l'observation réelle que ne le laisse penser cette valeur estimée. Cet argument est à double tranchant puisque la valeur réelle peut aussi être supérieure à ce qui a été estimé par le retour d'expérience ;
- La troisième serait que les très petits bassins capables de débits spécifiques extrêmes sont hors du domaine de validité du modèle. Vue l'incertitude qui existe sur la régression ayant donné les coefficients correctifs au niveau des surfaces faibles, il est possible que la relation ne soit pas adaptée dans ces cas-là.

Il faut aussi tempérer cette observation par la faible importance de la population d'évènements sur ces bassins. De fait, les estimations existent suite à l'évènement extrême de septembre 2002 et souvent, sur ces bassins, de telles estimations ne sont pratiquées que pour ce type d'évènements, peu fréquents. Par ailleurs, leur petite taille réduit les chances qu'un pluviomètre y soit implanté. Il faudra donc certainement se contenter, au moins pour quelques années, d'estimations forcément incertaines et de données radar différentes de celles des pluviomètres.

D'un point de vue temporel, les décalages sont assez systématiquement compensés par des *SPPD* convenables. L'impact reste donc faible sur l'échantillon cible. On peut émettre l'hypothèse d'une faible variabilité du temps de réponse, en particulier en terrains schisteux, typiquement cévenol. Il se peut en effet que ces terrains aient un temps de production élevé et à peu-près équivalent sur l'ensemble du massif et un temps de transfert uniquement dépendant des caractéristiques des drains. De ce fait, la variabilité des temps de réponse serait beaucoup plus faible et

expliquerait la pauvreté des variations observées en fonction des caractéristiques topographiques.

### 5.4.3 Apports et limites

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

- Bien qu'important, le nombre d'évènements utilisé 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.
- Encore une fois, bien qu'important, le nombre d'évènements, de bassins, et d'évènements par bassin (sauf, à la limite, pour la Ganière) utilisés pour valider la méthode est un peu faible pour tirer des conclusions robustes sur le fonctionnement du modèle. Cependant, il ne peut relever du hasard que le fonctionnement du modèle soit aussi performant sur autant de cas différents.
- 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.
- 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.

Ces limites ne semblent pas insurmontables et nous détaillerons les perspectives de recherche qui viseront à les repousser, mais auparavant, les apports de ce modèle couplé à cette méthode de correction sont présentés :

- Les fondements scientifiques du modèle ont pu être justifiés : les hypothèses qui sous-tendent son fonctionnement, de même que celui de la méthode de correction, sont vérifiables.
- A l'exception du bassin pilote, le volume de données nécessaire est faible et la difficulté pour les acquérir l'est aussi. Les données nécessaires sont donc bien en accord avec leur disponibilité.
- Les travaux présentés dans le chapitre 4 ont montré à quel point la recherche d'un point d'équilibre entre complexité et flexibilité avait été privilégiée. De fait, la complexité est minimale pour le bon fonctionnement du modèle.
- Si la méthode de correction nécessite d'autres validations pour prouver sa robustesse, le modèle a quant à lui démontré à quel point il était robuste et transposable. Sa validation est assurée par tous les procédés de validation croisée ainsi que par les tests effectués.
- Même si le modèle est de type boîte noire, son fonctionnement reste assez aisément compréhensible pour un utilisateur quelconque, avec un peu d'expérience d'utilisation.

Nous voyons donc que le modèle répond aux critères de qualité généralement attendus pour des modèles hydrologiques et qui ont été énoncés dans la section 2.2.1.3.

En sus, l'apport majeur consiste en la réalisation de prévisions relativement fiables – au sens de l'anticipation des inondations préjudiciables à la sécurité des personnes et des biens - sur plusieurs bassins non jaugés et plusieurs types d'évènements dans une région réputée particulièrement délicate pour la prévision des crues, y-compris sur des bassins jaugés. Ces prévisions se font sans aucune prévision de pluie ni utilisation de débits observés à l'opérationnel. Le modèle développé est donc réellement prédictif sur des bassins non jaugés.

Il peut apporter une information importante pour la protection des personnes et des biens en temps réel. Grâce au fait qu'il fonctionne sur des bassins de faible taille, il peut aussi aider à quantifier les apports en crue aux drains principaux. Par exemple, sur la Cèze, de bonnes prévisions sur l'Auzonnet et la Ganière peuvent aider à appréhender de façon plus fiable le débit à Tharoux, à l'entrée des gorges de la Cèze.

Pour ce qui est du domaine de validité géographique, la question mériterait d'être approfondie par de nombreuses autres validations mais plusieurs pistes se dégagent déjà :

- Les caractéristiques topographiques de la zone d'étude sont très spécifiques : pentes fortes et densités de drainage élevées en relation avec des cumuls de pluie annuels très importants (>2000mm sur les crêtes) et une grande fréquence des épisodes pluvieux intenses. Il paraît pertinent de limiter les travaux à des bassins qui sont alimentés par ces zones.
- Les bassins de la plaine languedocienne sont très fréquemment impactés par des perturbations des écoulements de surface d'origine karstique. Il paraît difficilement envisageable d'utiliser ce modèle et cette méthode sur ces bassins.

L'on peut donc penser qu'entre faille des Cévennes et ligne de partage des eaux Atlantique-Méditerranée, le modèle couplé à cette méthode fournit des performances optimales.

Cela pourrait être confirmé par des validations sur des bassins plus septentrionaux (Eyrieux, Doux, par exemple) ou plus occidentaux (Orb par exemple, en évitant la Lergue, l'Arre ou la Vis, aux bassins très influencés par des karsts).

Cependant, les bassins atlantiques en leurs extrêmes amonts présentent encore un certain nombre de caractéristiques cévenoles. L'on peut penser que des validations sur les bassins du Tarn amont, du Tarnon ou encore de la Mimente, par exemple, peuvent être pertinentes pour peut-être étendre légèrement le domaine de validité.

Enfin, dans d'autres régions, il s'agirait de trouver des zones aux caractéristiques relativement proches. Ce pourrait être le cas de certains bassins des Alpes-Maritimes (mais de nombreux barrages hydroélectriques sont présents), des Pyrénées-Orientales (amont) et de la Corse orientale (mais la quantité de données disponibles y est assez limitée).

Quoi qu'il en soit, pour des régions de caractéristiques différentes, rien n'empêche d'appliquer la même méthode avec un autre bassin pilote et des coefficients correctifs spécifiques, y compris sur des régimes hydrologiques différents. La sélection des variables (utilisation d'informations complémentaires sur l'état initial, sur la fonte des neiges, pas de temps différents...) devra en revanche être adaptée aux systèmes hydrologiques visés.

Pour rendre le modèle pleinement opérationnel sur la zone d'étude, une période de mise à l'épreuve sur des bassins jaugés comme la Ganière à Banne, l'Auzonnet aux Mages ou le Gardon à Saumane combinée à la collecte d'autres données sur les bassins cibles pour effectuer des validations complémentaires seraient nécessaires. Ces opérations permettraient d'analyser le comportement du modèle sur d'autres événements, anciens et actuels, et par conséquent, une prise en main optimale par les prévisionnistes.

## **Conclusion générale**

L'objectif de ces travaux était de parvenir à fournir des prévisions de crues éclair permettant une information préventive des populations et une aide à la décision en situation critique pour les autorités locales dans différents contextes de mesures hydrométriques : celui des bassins bien jaugés, là où les données de débit sont disponibles en temps réel aussi bien que sur de longues périodes passées ; celui des bassins mal jaugés, là où des chroniques de débit ont été acquises mais où les données ne sont pas disponibles ou fiables en temps réel et enfin celui des bassins non jaugés, là où aucune donnée de débit ou de hauteur d'eau n'est disponible. Les enjeux concernent la protection des populations et des biens des zones impactées par ce type de crues qui font régulièrement des victimes et des dégâts matériels considérables.

L'idée directrice des travaux a été de limiter la complexité de toutes nos démarches au strict nécessaire, en préservant la possibilité d'une utilisation finale des résultats en conditions opérationnelles. Ainsi, l'expertise sur les phénomènes connus a naturellement laissé sa place aux travaux statistiques dès lors que les processus en jeu sont apparus difficilement identifiables ou quantifiables aux échelles de temps et d'espace nécessaires.

Nous avons commencé par caractériser précisément la zone d'étude, ses problématiques hydrométéorologiques spécifiques et les mesures existantes mises en place pour se prémunir des risques d'inondation. Cette approche pluridisciplinaire nous a permis de conclure à une grande importance de la prévision de l'aléa hydrologique et de poser clairement la problématique de ces travaux, de sorte qu'elle soit adaptée au contexte de la zone.

A travers une identification et une définition des processus hydrologiques intervenant sur ces crues, nous avons pu mesurer la complexité et la non-linéarité de la relation entre pluie et débit. Ce constat, mis en perspective avec un tour d'horizon des modèles hydrologiques existants, nous a poussés à considérer, en l'état actuel des connaissances, la solution de l'apprentissage statistique par réseaux de neurones comme pertinente pour la résolution de notre problématique. Sans perdre de vue l'importance capitale de l'identification et de la quantification des processus hydrologiques, dans les délais impartis, la modélisation de type « boîte noire » est apparue adaptée à la situation.

Cette famille de modèles a donc été présentée. Ses capacités et propriétés ont été décrites et ses biais ont été caractérisés, de même que les méthodes élaborées pour les mitiger. Forts de ces informations nous avons appliqué ces modèles au Gardon à Mialet, dans le but d'élaborer un modèle capable de fournir des prévisions en contexte bien jaugé et mal jaugé. L'application rigoureuse des méthodes de régularisation et d'optimisation de la capacité de généralisation des modèles a permis d'atteindre une anticipation de 2 heures sans utilisation de pluies prévues sur ce petit bassin. Ces résultats étant suffisamment robustes, nous avons pu envisager une utilisation opérationnelle sur ce bassin ainsi que leur généralisation à des bassins non jaugés.

C'est donc ce qui a été effectué dans le chapitre 5. Nous y avons ainsi défini un post-traitement pour généraliser les prévisions ainsi disponibles à 35 événements sur 14 bassins utilisés comme des bassins non jaugés. Les résultats de la méthode ainsi exposée valident notre démarche : ils fournissent des temps d'anticipation exploitables en temps sur la majorité des cas testés avec la prévision, en moyenne, de 80% du pic observé (et plus encore sur les bassins de plus de 30km<sup>2</sup> disposant de données de pluie représentatives).

Si les très petits bassins aux réactions extrêmes posent encore un certain nombre de problèmes et si les incertitudes liées à la qualité des forçages ou au besoin d'un plus grand nombre d'évènements et de bassins en validation sont encore présentes, il n'en reste pas moins qu'à partir de 30km<sup>2</sup> environ et jusqu'à plus de 500km<sup>2</sup>, la dynamique des bassins est bien appréhendée avec une erreur moyenne sur l'amplitude prévue inférieure à 20%.

Le défi qui consistait à réaliser l'apprentissage d'un prédicteur neuronal dédié à un bassin versant dépourvu de base de données de débit a donc été relevé avec succès. La réalisation d'un prédicteur statistique pour les bassins versants non jaugés est, de plus, riche en enseignements et en démarches nouvelles de recherches.

En premier lieu, comme c'est le cas avec toutes les démarches statistiques, la quantité de données pour le calage est primordiale. De ce fait, on peut conjecturer que plus le temps passera, meilleurs seront les résultats. Au fil des épisodes, l'on pourra affiner les modèles jusqu'à ce que des prévisions déterministes comme probabilistes atteignent un haut niveau de fiabilité. De même, la correction des débits en fonction du bassin cible en généralisation ne pourra que s'améliorer par l'apport d'autres évènements et bassins. A terme, il est même possible d'envisager l'identification de processus (boîtes blanches) à l'aide de l'apprentissage statistique, comme c'est déjà le cas sur des systèmes karstiques très complexes (Kong-A-Siou, 2011).

La quantité, la qualité intrinsèque et la qualité de la sélection et du traitement des données est en définitive tout l'objet de ce mémoire.

Plus spécifiquement, la méthode de correction gagnerait assurément en efficacité avec trois avancées :

- L'augmentation du nombre de débits spécifiques de pointe qui ont servi à la définition des coefficients correctifs du post-traitement. Avec de nombreuses analyses *a posteriori* et la remontée d'anciennes et de nouvelles informations sur de plus nombreux bassins, la relation pourrait s'affiner. Il pourrait alors être possible de définir plusieurs régressions, pour les crues moyennes ou les plus fortes, pour certaines tailles de bassins, certaines pentes ou encore différents types d'épisodes pluvieux... Les domaines de validité seraient nettement plus robustes et permettraient donc d'appliquer la méthode à bon escient en toutes circonstances.
- Le choix d'autres bassins pilotes pourrait apporter la possibilité de prévisions d'ensemble et donc, une éventuelle amélioration des résultats moyens. Cependant, la quantité de données nécessaire à un apprentissage correct sur un bassin pilote est assez limitante dans la zone d'étude. L'on peut aussi imaginer la création, un temps envisagée, d'un bassin normalisé basé sur plusieurs bassins pilotes bien cernés.
- Une recherche élargie sur les temps de réponse, basée là encore sur de nombreuses données, pourrait permettre d'apporter une information temporelle plus fine et reliée aux caractéristiques du bassin, à la structure et à l'intensité des pluies ou encore à l'état initial. Cette information affinée pourrait alors être utilisée pour mener à bien les corrections temporelles qui n'ont pas pu être traitées ici.

Naturellement, la validation de la méthode proposée sur d'autres bassins et d'autres évènements permettrait de la rendre plus robuste et d'en analyser encore mieux les faiblesses et les forces. En particulier, des validations en périphérie de la zone d'étude permettraient de définir plus clairement son domaine de validité.

De même, l'utilisation d'un bassin pilote alimenté par des pluies radar pourrait permettre de savoir si la différence entre le signal issu des pluviomètres et celui issu d'une estimation radar ponctuelle explique les résultats médiocres obtenus sur les très petits bassins. Cependant, le nombre d'évènements disponibles avec le radar est assez faible, repoussant probablement cette échéance à plusieurs années.

Enfin, l'apport de prévisions de pluies fiables sur  $n$  pas de temps futurs pourrait repousser l'horizon de prévision d'autant de pas de temps. Avec l'utilisation des images radar, le cumul de la prochaine demi-heure peut souvent être estimé de façon relativement fiable par un prévisionniste aguerri et ainsi, augmenter légèrement l'horizon de prévision.

L'élaboration et l'application de cette méthode ont non seulement montré une efficacité intéressante en conditions opérationnelles, mais elles ont, en plus, ouvert des perspectives de recherche passionnantes pour la compréhension des phénomènes hydrologiques.



## Références bibliographiques

- Abrahart, R.J., See, L.M., 2007, Neural network modeling of non-linear hydrological relationships, *Hydrol. Earth. Syst. Sci.*, 11, pp1563-1579.
- Abrahart, R.J., Anctil, F., Coulibaly, P., Dawson, C.W., Mount, N.J., See, L.M., Shamseldin, A.Y., Solomatine, D.P., Toth, E., Wilby, R.L., 2012, Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting, *Progress in physical geography*, 36-4, pp 481-513.
- Alfieri, L., Smith, P.J., Thielen-del-Pozo, J., Beven, K.J., 2011, A staggered approach to flash flood forecasting – case study in the Cévennes region, *Advances in Geosciences*, 29, pp13-20.
- Ambroise, B., 1998, Genèse des débits dans les petits bassins versants ruraux en milieu tempéré : processus et facteur, *Revue des sciences de l'eau*, pp. 471-495.
- Anctil, F., Michel, C., Perrin, C., Andréassian, V., 2004, A soil moisture index as an auxiliary input for stream flow forecasting, *Journal of Hydrology* 286(2004), pp. 155-167.
- Andréassian, V., Perrin, C., Parent, E., Bardossy, A., 2010, The court of miracles in hydrology : can failure stories contribute to hydrological science ?, *Hydrological sciences journal*, 55(6), pp849-856.
- Araghinejad, S., Azmi, M., Kholghi, M., 2011, Application of artificial neural network ensembles in probabilistic hydrological forecasting, *Journal of Hydrology*, 407(2011), pp. 94-104.
- 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, doi : 10.1109/NaBIC.2011.6089612.
- 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.
- Ayral, P.A., 2005, Contribution à la spatialisation du modèle du modèle opérationnel de prévision des crues éclair ALTHAIR, thèse de l'Université de Provence, 311p.
- Bachoc, A., Garçon, R., Daly, F., Gautheron, A., Gaume, E., Perrin, C., Obled, C., Lepelletier, T., 2011, Besoins opérationnels et progrès actuels en matière de connaissance des phénomènes physiques dans l'hydrologie des crues, *La Houille Blanche*, n°2011-1, pp 14-21.
- Barron, A.R., 1993, Universal Approximation Bounds for Superpositions of a Sigmoidal Function, *IEEE Transactions on Information Theory* IT-39, pp. 930-945.
- Bartholmes, J., Todini, E., 2005, Coupling meteorological and hydrological models for flood forecasting, *Hydrology and Earth System Sciences*, 9(4), pp. 333-346.
- Beaudeau, P., Leboulanger, T., Hanneton, S., 1994, Prévision des crues turbides : apport des méthodes neuronales aux démarches classiques sur l'exemple du forage d' Yport (Seine-Maritime), *TSM. Techniques sciences méthodes, génie urbain génie rural*, n° 2, pp. 63-93 (16 ref.), pp. 97-109.

- Bender, E., 1996, *Mathematical Methods in Artificial Intelligence*, California, IEEE Computer Society Press, 636p.
- Bergström, S., 1991, Principles and confidence in hydrological modeling, *Nordic Hydrology*, 22, pp. 123-136.
- Bergström, S., 1995, The HBV-IWS model, *in Computer Models of Watershed Hydrology* (éd. by V.P. SINGH), pp. 443-476, Water Resources Publications, Littleton, Colorado, USA.
- Bergström, S., Forsman, A., 1973, Development of a conceptual deterministic rainfall-runoff model, *Nordic Hydrology* 4, pp. 147-170.
- Beven, K.J., Kirkby, M.J., 1979, a physically based variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24(1), pp. 43-69.
- Beven, K.J., 1989, Changing ideas in hydrology, the case of physically-based models, *Journal of hydrology*, 105, pp. 157-172.
- Beven, K.J., Lamb, R., Quinn, P.F., Romanowicz, R., Freer, J., 1995, TOPMODEL, in V P Singh (Ed). *Computer Models of Watershed Hydrology*, Water Resources Publications, pp. 627-668.
- Beven, K.J., 2004, Infiltration excess at the Horton hydrology Laboratory (or not ?), *Journal of Hydrology*, 293(2004), pp219-234.
- Borrell, V., 2004, *Vers une modélisation hydrologique adaptée à la prévision opérationnelle des crues éclair, application à de petits bassins versants du sud de la France*, thèse de l'INP Toulouse, 254p.
- Borrell, V., Chorda, J., Dartus, D., 2005, *Prévision des crues éclair*, C.R. Géoscience 337, pp. 1109-1119.
- Borrell, V., Dartus, D., Alquier, M., 2002, Forecast of flash floods in ungauged basins with satellite data, PUB, Kick-off Workshop of the IAHS DECADE ON PREDICTION IN UNGAGED BASINS, Hydrological Sciences on Mission, Communications.
- Boudevillain, B., Argence, S., Claud, C., Ducrocq, V., Joly, B., Joly, A., Lambert, D., Nuissier, O., Plu, M., Ricard, D., Arbogast, P., Berne, A., Chaboureau, J.P., Chapon, B., Crépin, F., Delrieu, G., Doerflinger, E., Funatsu, B., Kirstetter, P.E., Masson, F., Maynard, K., Richard, E., Sanchez, E., Terray, L., Walpersdorf, A., 2009, Cyclogenèses et précipitations intenses en région méditerranéenne : origines et caractéristiques, *La Météorologie*, 66, pp 18-28.
- Campbell, E., P., 2006, Physical-statistical models for predictions in ungauged basins, PUB, *International perspectives on the state of the art and pathways forward*, IAHS Publ. 301, pp. 291-298.
- Cappus, P., 1960, Bassin expérimental d'Alrance, Etude des lois de l'écoulement, application au calcul et à la prevision des debits, *La Houille Blanche*, Juillet-Aout 1960 - N°A, pp. 493-520.
- Chneiweis, A., 2008, Introduction de « Catastrophes naturelles, crise financière : sommes-nous entrés dans l'ère des évènements extrêmes ? », Dossier de la Revue Risques n° 76/décembre 2008 : Rubrique Risques et solutions - Les évènements extrêmes.
- Corzo, G., Solomatine, D., 2007, Knowledge-based modularization and global optimization of artificial neural network models in hydrological forecasting, *Neural Networks* 20 (2007) 528-536.
- Cosandey, C., 1994, Formation des crues "cévenoles" dans des bassins élémentaires du Mont-Lozère, *Revue des sciences de l'eau*, pp. 377-393.

- Cosandey, C., Robinson, M., 2000, Hydrologie continentale, Armand Colin, Collection Universitaire, Géographie, Paris, 360p.
- Cosandey, C., Bigot, S., Dacharry, M., Gille, E., Laganier, R., Salavador, P.G., 2003, Les eaux courantes, Géographie et environnement, éd. Belin sup, 240p.
- Cosandey, C., Andréassian, V., Martin, C., Didon-Lescot, J.F., Lavabre, J., Folton, N., Mathys, N., Richard, D., 2005, The hydrological impact of the mediterranean forest: a review of French research, *Journal of Hydrology* 301(2005), pp. 235-249.
- Coulibaly, P., Anctil, F., Bobee, B., 2000, Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, *Journal of Hydrology*, Volume 230, Issues 3-4, 8 pp. 244-257.
- Coustau, M., 2011, Contribution à la prévision des crues sur le bassin du Lez : modélisation de la relation pluie-débit en zone karstique et impact de l'assimilation des débits, Thèse de l'Université Montpellier 2, 230p.
- Cybenko, G., 1989, Approximation by superpositions of a sigmoidal function, *Mathematics of Control, Signals and Systems*, vol 2, pp. 303-314, 1989.
- Dreyfus, G., Martinez, J.M., Samuelides, M., 2008, Apprentissage statistique, EYROLLES, Algorithmes, 417p.
- Dolciné, L., Andrieu, H., Sempere-Torres, D., Creutin, D., 2001, Flash flood forecasting with coupled precipitation model in mountainous Mediterranean basin, *Journal of Hydrologic Engineering*, 6-1 pp 1-10.
- Dunne, T., Black, R.D., 1970, Partial area contributions to storm runoff in a small new England watershed, *Water Resources Research*, Vol.6 No5, pp 1296-1311.
- Emmanuel, I., Andrieu, H., Tabary, P., 2012, Evaluation of the new French operational weather radar product for the field of urban hydrology, *Atmospheric Research*, Article in press.
- Faure, M., 1995 Late orogenic carboniferous extensions in the Variscan French Massif Central Tectonics, 14 (1995), pp. 132-153.
- FFSA, 2011, Inondations du Var, le bilan chiffré un an après, [http://www.ffsa.fr/sites/jcms/p1\\_429021/inondations-du-var-le-bilan-chiffre-un-an-apres?cc=fp\\_8801](http://www.ffsa.fr/sites/jcms/p1_429021/inondations-du-var-le-bilan-chiffre-un-an-apres?cc=fp_8801), consulté le 01/02/2012.
- Fiorentino, M., Iacobellis, V., 2005, Towards reduction of uncertainty on MEDCLUB (MEDiterranean CLimate Ungauged Basins), PUB, International perspectives on the state of the art and pathways forward, *IAHS Publ.* 301, pp. 275-291.
- Funahashi, K., 1989, On the approximate realization of continuous mappings by neural networks, *Neural Networks*, vol 2, pp. 183-192.
- Gallinari, P., Cibas, T., 1999, Practical complexity control in multilayer perceptrons, *Signal Processing*, vol 74, pp. 29-46.
- Gaume, E., 2002, *Eléments d'analyse des crues éclair*, thèse de doctorat ENGREF, INRS, 320p.
- Gaume E., 2003, Hydrologie, Cours de DEA Sciences et Techniques de l'Environnement de l'Ecole Nationale des Ponts et Chaussées, Chapitres 2 et 3 Hydrologie de versants et de bassins versants et modélisation pluie-débit, 93p.

- Geman, S., Bienenstock, E., Doursat, R., 1992, Neural networks and the bias / variance dilemma, *Neural Computation*, vol 4, pp. 1-58.
- Gibbs, M.S., Maier, H.R., Dandy, G.C., 2011, A generic framework for regression regionalization in ungauged catchments, *Environmental Modelling & Software*, vol. 27-28, pp. 1-14.
- Goswami, M., O'Connor, K.M., 2006, Flow simulation in an ungauged basin: an alternative approach to parametrization of a conceptual model using regional data, Large sample basin experiments for hydrological parametrization: results of the model parameter experiment – MOPEX, IAHS Publ. 307, pp. 149-158.
- Habets F., Boone, A., Champeaux, J.L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Quintana Segui, P., Rousset-Regimbeau, F., Viennot, P., 2008, The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France, *Journal of Geophysical Research D : Atmospheres* 113, D06113 (2008) 18.
- Hewlett, J.D., Hibbert, A.R., 1963, Moisture and energy condition within a sloping soil mass during drainage, *Journal of geophysical research*, 68(4), pp. 1081-1087.
- Hewlett, J.D., Hibbert, A.R., 1967, Factors affecting the response of small watersheds to precipitation in humid areas, *in* Sopper, W.E., Lull, H.W., editors, *Forest Hydrology*, New York: Pergamon press, pp. 275-290.
- Hornik, K., Stinchcombe, M., White, H., 1989, Multilayer feedforward networks are universal approximators, *Neural Networks*, vol 2, pp. 359-366.
- Horton, R. E., 1933, The role of infiltration in the hydrologic cycle, *Eos Trans. AGU*, 14, pp446-460, 1933.
- Huet, P., Martin, X., Prime, J.L., Foin, P., Laurain, C., Cannard, P., 2003, Retour d'expérience des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches-du-Rhône, de l'Ardèche et de la Drôme, Rapport consolidé après phase contradictoire, IGA, CGPC, CCGREF, IGE.
- Hundecha, Y., ZEHE, E., BARDOSSY, A., 2002, Regional parameter estimation from catchment properties for the prediction of ungauged basins, PUB, Kick-off Workshop of the IAHS DECADE ON PREDICTION IN UNGAGED BASINS, Hydrological Sciences on Mission, Communications.
- Institut Géographique National, 2006, BD CARTHAGE 3.0, Descriptif de livraison Format MIF/MID, pp 17-18.
- Jacq, V., 1996, Phénomènes remarquables n°3, Inventaire des situations à précipitations diluviennes en Languedoc-Roussillon / PACA / Corse, Période 1958-1994, Service Central d'Exploitation de la Météorologie, BCIR Sud-Est et Centres Départementaux, Météo France, 190p.
- Kessler, J., Chambraud, A., 1990, *Météo de la France, tous les climats localité par localité*, Editions JC Lattès, 391p.
- Kitadinis, P.K., Bras, R.L., 1980, Real-Time forecasting with a conceptual hydrologic model, applications and results, *Water Resources Research*, vol 16, n° 6, pp. 1034-1044.
- Klemeš, V., 1986, Operational testing of hydrological simulation models, *Hydrological Sciences Journal* 31(1), pp. 13-24.
- Kong A Siou, L., Johannet, A., Borell Estupina, V., Pistre, S., 2012, Optimization of the generalization capability for rainfall-runoff modeling by neural networks: The case of the *Lez*

aquifer (southern France), *Environmental Earth Sciences*, Volume 65, Number 8, Pages 2365-2375.

Kong-A-Siou, L., 2011, Modélisation des crues de bassins karstiques par réseaux de neurones, cas de l'aquifère du Lez (Montpellier), thèse de l'Université Montpellier 2, 230p.

Lavabre, J., Grégoris, Y., 2005, AIGA : un dispositif d'alerte des crues sur l'ensemble du réseau hydrographique. *Ingénieries EAT*, n° 44, 3-12.

Ledoux Consultants, 2003, MEDD SDPRM, Estimation quantitative et qualitative des dommages économiques dans le Gard à la suite des inondations des 8 au 10 septembre 2002. Cas des entreprises et des habitats des particuliers, 49p.

Ledoux, B., 2006, La gestion du risque inondation, Ed. Lavoisier, 769p.

Le-Lay, M., Saulnier, G.M., 2007, Exploring the signature of climate and landscape spatial variabilities in flash flood events: case of the 8-9 September 2002 Cévennes-Vivarais catastrophic event, *Geophys. Res. Lett.* 34:L13401.

Leone, F., 2007, Caractérisation des vulnérabilités aux catastrophes « naturelles » : contribution à une évaluation géographique multirisque (mouvements de terrain, séismes, tsunamis, éruptions volcaniques, cyclones), Mémoire d'Habilitation à Diriger des Recherches (HDR), Section 23 (Géographie), Université Paul Valéry, Montpellier III, Laboratoire GESTER, Montpellier, 245 p. + Annexes.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. Bergström, S., 1997, Development and test of the distributed HBV-96 hydrological model, *Journal of Hydrology* 201, pp. 272-288.

Linsley, R.K., 1982, Rainfall-runoff models - an overview, in V. P. SINGH (Editor), *Proceedings of the international symposium on rainfall-runoff modeling*, Water Resources Publications, Littleton, CO, USA, pp. 3-22.

Mckay, D. J. C., 1992, A practical Bayesian framework for backpropagation networks, *Neural Computation*, vol 4, pp. 448-472.

Mangin, A., Apports des analyses corrélatoire et spectrale croisées dans la connaissance des systèmes hydrologiques, note de Alain Mangin transmise par Michel Durand-Delga, *C.R Acad. Sc. Paris*, 293, Série II, pp. 1011-1014.

Mangin, A., 1984, Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale, *Journal of Hydrology*, 67(1984), pp 25-43.

Marchandise, A., 2007, Modélisation hydrologique distribuée sur le Gardon d'Anduze ; étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et tests d'hypothèses sur les processus hydrologiques, thèse de l'Université Montpellier II, 215p.

Martin, C., Didon-Lescot, J.F., Jolivet, J., 2008, Réflexion sur le rôle hydrologique des aménagements anciens des versants et des thalwegs : l'exemple des Cévennes, *Travaux du Laboratoire de Géographie Physique Appliquée* n°26, pp. 31-40.

Mathevet, T., 2005, Quels modèles Pluie-Débit globaux au pas de temps horaire ? Développements empiriques et comparaison de modèles sur un large échantillon de bassins versants, thèse, ENGREF, 463p.

Mackay, D., 1992, Information based objective functions for active data selection, *Neural Computation* 4, pp 590-604.

MEDDTL, 2012, [http://www.vigicrues.gouv.fr/som\\_aide.php](http://www.vigicrues.gouv.fr/som_aide.php).

Merz, R., Blöschl, G., Parajka, J., 2006 Regionalization methods in rainfall-runoff modelling using large catchment samples, Large sample basin experiments for hydrological parametrization: results of the model parameter experiment – MOPEX, IAHS Publ. 307, pp. 117-125.

Météo France, 2012, Pluies Extrêmes en France Métropolitaine, <http://pluiesextremes.meteo.fr/index.php>, Dir. Pub. Jacq, F., consulté en janvier 2012.

Meunier, M., 1996, Couvert forestier et crues sur les petits bassins versants de montagne, Unasylva n°185 - Influences forestières.

Miquel, P., 2001, Excès climatiques sur la Montagne Languedocienne et conséquences catastrophiques, 244p.

Montz, B.E., Grunfest, E., 2002, Flash flood mitigation: recommendations for research and applications, Environmental Hazards, 4(1) pp. 15-22.

Moussa, R., 2010, When monstrosity can be beautiful while normality can be ugly: assessing the performance of event-based flood models, Hydrological Sciences Journal, 55(6)2010, Special Issue : The court of miracles of hydrology, pp. 1074-1084.

Moussa, R., Chahinian, N., Bocquillon, C., 2007, Distributed hydrological modelling of a Mediterranean mountainous catchment – Model construction and multi-site validation, Journal of Hydrology (2007) 337, pp. 35-51.

Musy, A., Higy, C., 2009, Hydrologie : 1 Une science de la nature, Presses polytechniques universitaires romandes, Ingénierie de l'Environnement, 314p.

Nash, J., Sutcliffe, J., 1970, River Flow Forecasting through Conceptual Models, part I – A Discussion of Principles, Journal of Hydrology, 10, pp. 282-290.

Obled, C., Saulnier, J.M., 1999, Le versant et le bassin versant : quelques aperçus sur la formation des crues et la production du ruissellement, La Houille Blanche, Revue Internationale de l'Eau, n°7-8, pp. 42-53.

Oudin, L., Kay, A., Andréassian, V., Perin, C., 2010, Are seemingly physically similar catchments truly hydrologically similar?, Water Resources Research, 46 W11558 15p.

Pardé, M., 2004, Sur la puissance des crues en diverses parties du monde, Ed. Géocarrefour, 373p.

Pereira Filho, A.J., Dos Santos, C.C., 2006, Modeling a densely urbanized watershed with an artificial neural network, weather radar and telemetric data, Volume 317, Issues 1–2, Pages 31–48.

Piotrowsky, A., Napiorkowski, J.J., Rowinski, P.M., 2006, Flash Flood Forecasting by Means of Neural Networks and Nearest Neighbour Approach - A Comparative Study, Nonlin. Processes Geophys. 13, 443-48.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992, Numerical recipes in C, Cambridge University Press.

Rasmussen T.-C., Baldwin R.H., Dowd R.H., Williams A.G., 2000, Tracer vs. Pressure wave velocities through unsaturated saprolites, *Soil Science Society of American Journal*, 64, pp. 75-85.

Réménieras, 1986, Hydrologie de l'ingénieur, Collection de la Direction des Etudes et Recherches d'Electricité de France, Ed. Eyrolles, 456p.

Rode, S., 2012, Le chêne ou le roseau : quelles stratégies de gestion du risque d'inondation en France ?, *Cybergeo : European Journal of Geography [Online], Regional and Urban Planning*, article 603, DOI : 10.4000/cybergeo.25299

Rodier, J., 1964, Principes généraux de l'étude des crues, Conférence présentée à l'Université de Gand les 12 et 13 février 1964, ORSTOM Fonds Documentaire n°32928, Cote B.

Rodier, J.A., Roche, M., 1984, Répertoire mondial des crues maximales observées, AISH, Wallingford (UK), n°143.

Roustan, M., 2007, Synthèse des travaux du Groupe d'Etudes sur les Inondations, Assemblée Nationale, XII<sup>e</sup> législature, 192p.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986, Learning Internal Representations by error backpropagation, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, pp. 318-362, MIT Press.

Salomon, J.N., 1997, L'homme face aux crues et aux inondations, Presses Universitaires de Bordeaux, 132p.

Sahoo, G.B., Ray, C., De Carlo, E.H., 2006, Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii, *Journal of Hydrology* 317, 63-80.

Schmitz, G.H., Cullmann, J., 2008, PAI-OFF: A new proposal for online flood forecasting in flash flood prone catchments, *Journal of Hydrology* 360, 1-14.

Sjöberg, J., Zhang, Q., Ljung, L., 1995, Nonlinear black-box modeling in system identification: a unified overview, *Automatica*, vol 31, pp. 1691-1724.

Soltner D., 1990. Les bases de la production végétale : le sol, le climat, la plante. Tome 1, le sol, Sciences et techniques agricoles.

Stone, M., 1974, Cross-validators choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, B36, pp. 111-147.

Sudheer, K. P., Gosain, A. K., Saheb, S. M., Rangan, D. M., 2002, Modelling evaporation using artificial neural network algorithm, *Hydrological Processes*, 16(16): 3189-3202.

Toth, E., 2009, Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting, *Hydrol. Earth Syst. Sci.*, 13, pp. 1555-1566.

Toukourou, M.S., 2009, Application de l'apprentissage artificiel à la prévision des crues éclair, thèse de l'ENSMP, 180p.

Toukourou, M.S., Johannet, A. Dreyfus, G., 2009, Flash flood forecasting by statistical learning in the absence of rainfall forecast: a case study, *Engineering Applications of Neural Networks Proceedings*, pp. 98-107.

Toukourou, M.S., Johannet, A., Dreyfus, G. and Ayrat P.-A., 2011, Rainfall-runoff modeling of flash floods in the absence of rainfall forecasts: the case of "Cévenol flash floods", *Journal of Applied Intelligence* vol. 35, 2, pp. 1078-189.

Tramblay, Y., Bouvier, C., Martin, C., Didon-Lescot, C., Todorovik, D., Domergue, J.M., 2010, Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling, *Journal of Hydrology*, 387, 176-187. <http://dx.doi.org/10.1016/j.jhydrol.2010.04.006>.

Tricart, J., Hirsch, F., 1960, Relation entre le débit et la superficie des bassins fluviaux, *Annales de Géographie*, 1960, 69, n°375, pp. 449-461.

UNESCO, OMM, 1992, Glossaire international d'hydrologie, en ligne sur : <http://webworld.unesco.org/water/ihp/db/glossary/glu/aglu.htm>.

Vinet, F., 2008, Geographical analysis of damage due to flash floods in southern France: The cases of 12–13 November 1999 and 8–9 September 2002, *Applied Geography* 28 (2008) 323–336.

Vinet, F., 2010, Le risque inondation, diagnostic et gestion, Coll. Sciences du Risque et du Danger, Lavoisier, 316p.

Visserot, Y., 2012, Intégration de l'information hydrométéorologique spatialisée pour la modélisation et la prévision des crues éclair à l'aide de réseaux de neurones, rapport de fin Master 2 Gestion des Catastrophes et des Risques Naturels, Université Montpellier III, 108p.

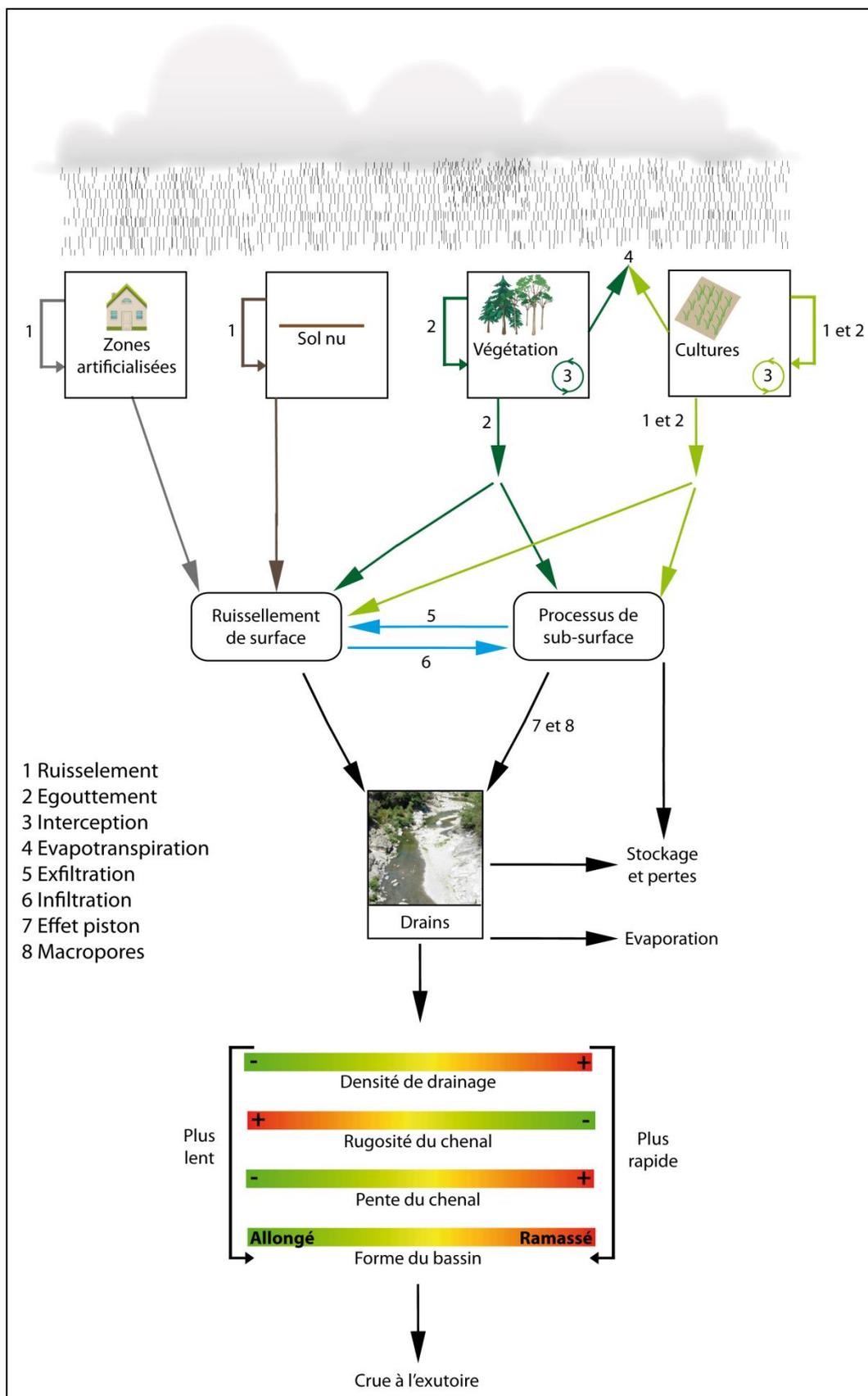
Yacoub, M., Bennani, Y., 2001, Une mesure de pertinence pour la sélection de variables dans les perceptrons multicouches, *Revue d'intelligence artificielle*, vol.15, pp. 393-410.

Yadav, M., Wagener, T., Gupta, H., 2007, Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, *Advances in Water Resources* 30 (2007) pp. 1756-1774.

Younis, J., Anquetin, S. and Thielen, J., 2008, The benefit of high-resolution operational weather forecasts for flash flood warning, *Hydrology and Earth System Sciences*, 12, pp. 1039-1051.

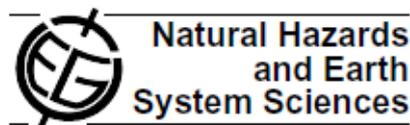
## Annexes

### ANNEXE 1 : schématisation du processus de transformation de la pluie en écoulement en crue.



**ANNEXE 2** : 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.

Nat. Hazards Earth Syst. Sci., 12, 3307–3324, 2012  
www.nat-hazards-earth-syst-sci.net/12/3307/2012/  
doi:10.5194/nhess-12-3307-2012  
© Author(s) 2012. CC Attribution 3.0 License.



## Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon de Mialet basin (southern France)

G. Artigue<sup>1</sup>, A. Johannet<sup>1</sup>, V. Borrell<sup>2</sup>, and S. Pistre<sup>2</sup>

<sup>1</sup>Ecole des Mines d'Alès, Centre des Matériaux de Grande Diffusion, Alès, France

<sup>2</sup>Université Montpellier II, Hydrosiences Montpellier, Montpellier, France

Correspondence to: A. Johannet (ame.johannet@mines-ales.fr)

Received: 8 February 2012 – Revised: 3 July 2012 – Accepted: 17 September 2012 – Published: 12 November 2012

**Abstract.** In southern France, flash flood episodes frequently cause fatalities and severe damage. In order to inform and warn populations, the French flood forecasting service (SCHAPI, *Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations*) initiated the BVNE (*Bassin Versant Numérique Expérimental*, or Experimental Digital Basin) project in an effort to enhance flash flood predictability. The target area for this study is the *Gardon d'Anduze* basin, located in the heart of the *Cévennes* range. In this Mediterranean mountainous setting, rainfall intensity can be very high, resulting in flash flooding. Discharge and rainfall gauges are often exposed to extreme weather conditions, which undermines measurement accuracy and continuity. Moreover, the processes governing rainfall-discharge relations are not well understood for these steeply-sloped and heterogeneous basins. In this context of inadequate information on both the forcing variables and process knowledge, neural networks are investigated due to their universal approximation and parsimony properties. We demonstrate herein that thanks to a rigorous variable and complexity selection, efficient forecasting of up to two-hour durations, without requiring rainfall forecasting as input, can be derived using the measured discharges available from a feedforward model. In the case of discharge gauge malfunction, in degraded mode, forecasting may result using a recurrent neural network model. We also observe that neural network models exhibit low sensitivity to uncertainty in rainfall measurements since producing ensemble forecasting does not significantly affect forecasting quality. In providing good results, this study suggests close consideration of our main purpose: generating forecasting on ungauged basins.

### 1 Introduction

In Mediterranean regions, flash floods due to heavy rainfall events frequently occur, causing a large number of fatalities and costly damage. In the last few years, these episodes have had serious consequences in southern France, including the *Gard* Department, where in 2002 a total of 23 people died with damage amounting to €1.2 billion (Gaume et al., 2009). The cost associated with this type of episode can, in rural areas, exceed €15 000 per resident (Vinet, 2008). More recently, on 15 June 2010, the *Var* Department experienced extreme flash flooding that caused 23 deaths (Rouzeau et al., 2010).

Population growth, especially in these attractive areas, and a sustainable land management policy imply the need to enhance prevention measures and early warning alerts transmitted to authorities and local populations.

Fully aware of this issue, the French Flood Forecasting Center (SCHAPI, *Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations*) initiated in 2006 the Experimental Digital Basin project (BVNE, *Bassin Versant Numérique Expérimental*) as a mean of improving flash flood forecasting by providing input to the flood warning map in the form of high-water alarm levels (known as “vigicrues”, [www.vigicrues.gouv.fr](http://www.vigicrues.gouv.fr)), through the use of new models, which are more efficient than previous ones relative to rapid hydrological responses.

The BVNE project is focused on the *Cévennes* range, located in the southern part of France, and more precisely on the *Gardon d'Anduze* basin (Fig. 1).

One of the main characteristics of the observed rainfall data during these flood events is their tremendous

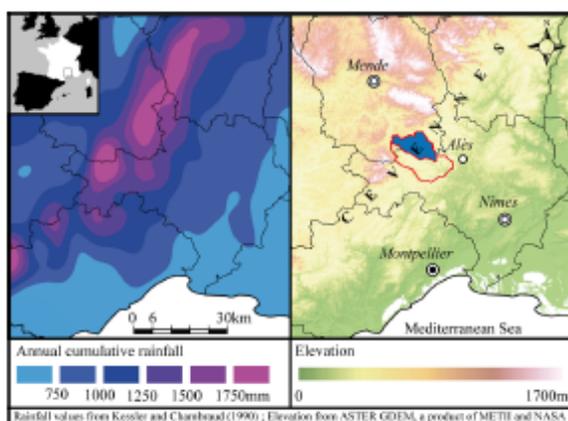


Fig. 1. Average annual cumulative rainfall (left); elevation and the Gardon d'Anduze Basin shown in the red frame. The blue zone indicates the Gardon de Mialet sub-basin (right) in the Cévennes area.

heterogeneity in both time and in space. For example, during the October 2008 event on the Gardon de Mialet Basin, rainfall reached a peak of 485 mm in Mialet over a 24-h period, and 17 mm within the same duration in Barre-des-Cévennes, a mere 15 km away. For this reason, reliable rainfall forecasts are not currently available for such a small area at the time scale considered useful (a few hours and a few kilometres).

The location and quantity of gauges do not, unfortunately, yield sufficiently accurate observations on the targeted basin. On the one hand, rain gauges require frequent maintenance and are not always operable under extreme conditions, while radar images do not provide sufficient accuracy in areas of relief. On the other hand, discharge gauges may be damaged or even destroyed by extreme floods, as was the case in Anduze during the great flood of September 2002 and in the Var Department (southern France) during June 2010. Moreover, these discharge gauges are not positioned at the outlet of each risky basin, thus, leaving rivers or creeks ungauged even though they may be hazardous in the event of flooding.

The human and economic issues related to flash floods have prompted researchers to define a procedure for performing discharge forecasts without having to rely on rainfall forecasts in poorly gauged or even ungauged basins. This specific aim is the topic of the present research.

In this paper, we will begin in Sect. 2 by presenting the Cévennes range and the basin chosen for this particular study, i.e., the Gardon de Mialet basin. In addition, a summary of previous studies on flash flood forecasting will be introduced, concerning hydrological physical modelling as well as neural networks modelling. Section 3 will present the design and optimisation of a feedforward model and of a recurrent model. One will be assessed in providing forecasting when discharge measurements are available (as state variable) and

the other will be tested on the same events when no discharge value is available in operational use (degraded mode in the case of a “poorly gauged basin”). Ensemble forecasting will also be tested in order to determine the eventual level of improvement for the recurrent model. After presenting the set of results, this paper will close with a discussion section and conclusion.

## 2 Target area: introduction and previous research

### 2.1 Background on the Cévennes Range

The Cévennes Range extends from the Hérault to Ardèche departments in a curved shape. Its distance to the Mediterranean Sea varies from 50 km in the Hérault to more than 100 km in the Ardèche. The elevations of this range can be as low as 100–200 m in the foothills area and rise to 1700 m, e.g., on Mont-Lozère crests. Most summits lie between elevations of 500 and 1000 m. The transition between Languedoc plains and Cévennes crests is very abrupt, thus, inducing steep slopes capable of exceeding 30 % in spots.

Elevation strongly affects the climate of this mountain range. Despite being dominated by Mediterranean influences due to the proximity of the sea, elevation still plays an important role, especially relative to rainfall. This aspect is responsible for the high spatial variability (Fig. 1).

Moreover, as highlighted by Moussa et al. (2007), rainfall is highly variable over time. On an annual basis, values can vary by up to 100 %, while major events can be quite common during several months and then be completely absent for a few years.

For the subject addressed in this paper, the most important characteristic of local climate is the high frequency of heavy rainfall events. The name of the range is synonymous with Cévenol episodes that often roughly correspond with the end of long summer droughts, intensified by high temperatures observed on southern slopes in the range's lower parts.

According to France's weather agency (Météo France, 2011), within a radius of 25 km around Saint-Jean-du-Gard (located in the heart of the Gardon d'Anduze Basin), during the period 1958–2010, over 200 mm in daily climatic events (from 06:00 to 06:00 in UTC) occurred 53 times, i.e., once a year on average. This figure is likely to be considerably less than that of events over 200 mm measured over a 24-h sliding period.

This weather pattern exerts a strong influence on local hydrology. For example, the first rainfall event at the end of summer may have fewer hydrological consequences than subsequent events, which occur on wet soils. At the end of a drought period, however, rivers and creeks may be totally dry, whereas they can reach into the thousands of cubic metres per second for the largest basins and  $30 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$  for the smallest ones (Gaume, 2002) for a rainfall event of sufficient intensity.

The local geology is dominated by crystalline rocks: granite over the upper part (Mont Aigoual, Mont Lozère, Mont Tanargue), schist over the middle part, and limestone and karst in spots over the lower part. These bedrocks, steep slopes, frequent droughts and intense downpours give rise to thin and rocky soils covered by sclerophyll vegetation over the lower part, while in contrast hardwoods and conifers cover the upper part.

## 2.2 The Gardon de Mialet basin

The Gardon de Mialet basin is representative of the Cévennes territory; it measures 220 km<sup>2</sup> and its elevation ranges from 147 m in Mialet to 1170 m, with steep slopes encountered (36% average slope). Its soils are rocky, whereas the substratum is composed of 95% mica-schists (Ayrat, 2005). Land use is homogeneous: 92% of the basin is covered by natural landscapes (principally chestnut trees, conifers, mixed forests and bush) (Ayrat, 2005) whereas 8% of the basin is covered by rocks and urban areas.

Three rain gauges located in Barre-des-Cévennes (elevation 930 m), Mialet (elevation 147 m) and Saint-Roman-de-Tousque (elevation 630 m) and one discharge gauge (in Mialet, at the outlet) are available on the basin (Fig. 2).

The characteristics of this basin, which regularly experiences very heavy localised rainfalls (at times exceeding 500 mm during just 24 h; to be compared with the 600-mm average annual rainfall in Paris, France), generating rapid and high-volume hydrological responses. The response and rise times observed during flood events are very short, leaving almost no reaction time to evacuate residents to shelters.

Snow accumulation on top of the Mialet basin rarely impacts flash floods hydrology. The importance of snowmelt in the hydrological system cannot be eliminated for an event occurring during January or February. Nevertheless it appears as highly improbable as warm air masses come from south-east (Mediterranean Sea) during most of the events. This meteorological configuration strongly limits the risk of snow on the summits of the basin. This influence can be considered as negligible.

The hydrological regime of the basin is essentially rain-fed. Consequently, very low discharges are observed in summer when droughts occur, whereas higher discharges are observed following rainfalls in the other seasons, particularly in autumn when the most intense rainfall episodes occur.

## 2.3 Flash flood forecasting

### 2.3.1 Physically based models

Flash floods are typically nonlinear phenomena. In order to characterise such an intense rainfall-discharge relation, physical hydrological models consider many parameters such as initial and constant soil moisture, soil thickness and slopes.

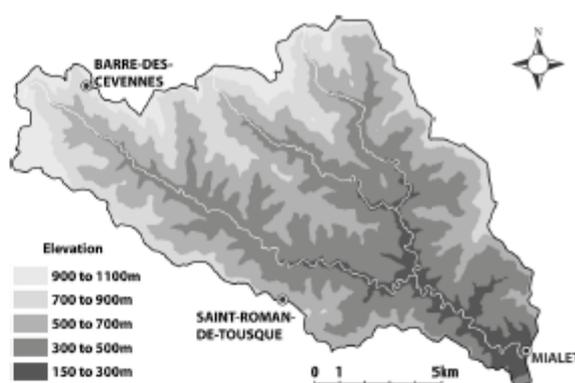


Fig. 2. Schematic topographic and hydrographic map of the Gardon-de-Mialet basin, including gauge locations (elevations derived from ASTER GDEM, a product of METII and NASA).

The majority of physical models have not been created for this specific purpose; consequently, some authors have tried to adapt them to flash floods of the Cévennes area, beginning this effort by performing simulations (i.e., modelling the behaviour of the basin, without forecasts).

As an example, Trambly et al. (2010) performed simulations using the Soil Conservation Service Curve Number model; they revealed the importance of initial soil moisture condition, a parameter difficult to assess. Gaume (2002) used a variant of this model to reconstruct past floods and improve their understanding. TOPMODEL (Beven and Kirkby, 1979; Beven et al., 1995) is based on a distribution of topographic parameters that impact the hydrological response. Using TOPMODEL, Le-Lay and Saulnier (2007) showed that the most influential variable to take into consideration for flash flooding was the spatial distribution of rainfalls, which underscores the issue of rainfall data quantity and quality. Moussa et al. (2007) used MODSPA on the Gardon-d'Anduze basin with a variable time step that allowed modelling both long- and short-term processes. With its multi-site calibration, MODSPA has yielded some interesting simulations.

Borrell et al. (2005) used MARINE, a model based on spatially distributed land characteristics (topography, land cover, etc.). These authors also performed high-quality real-time simulations, though these were not optimally suited for forecasting. In addition, MARINE is very sensitive to initialisation steps involving natural variables. These values are quite difficult to obtain in real-time, for event-based forecasts; hence, the initial soil moisture and values from piezometers have been used (Coustau et al., 2011).

Marchandise (2007) assessed various types of rainfall-discharge models on the Gardon d'Anduze basin. This author concluded that in general, model parameters do not represent very well the physical significance they were intended for. He

also noted difficulties when taking into account soil moisture and thickness as well as the model space and time variability.

Hydrological models have basically encountered two types of issues. First, they are not easily able to cope with the absence of rainfall forecasts. Second, initial conditions prove very difficult to adjust in real-time use; moreover, obtaining previous or real-time measurements of physical quantities in precise locations does not necessarily guarantee the representativeness of conditions over the entire basin (Marchandise, 2007).

The problem of the unavailability of rainfall forecasts for flash floods forecasting is pointed out from several years. Coupling weather forecasts with hydrological forecasts provides thus a good alternative to the zero-rainfall hypothesis. Some interesting results (more than 80% of the peak discharge nine days before the event) have been provided this way on large basins producing slower floods than those studied herein (Bartholmes and Todini, 2005). In the case of flash floods, the small size of the affected basins and the high rainfall intensities of localised convective cells do not allow for precise forecasting even though studies show that benefits would be obtained from such a process (Younis et al., 2008). Dolciné et al. (2001) showed on the Gardon d'Anduze basin, and for few (not very intense) events, that the basin operates filtering on the rainfall signal significantly affecting sensitivity to the rainfall forecasts. It was shown that quality of rainfalls forecasts are not critical for very short discharge lead time forecasts, whereas for greater lead-times, it could be really efficient to use rainfall forecasts coupled to hydrological forecasts. Such conclusions could be generalised to smaller basin as the Mialet basin and to stronger events as those addressed in the present study. More recently, Alfieri et al. (2011) illustrated coupling meteorological and hydrological modelling in the frame of the European Flood Alert System (EFAS), as well on Gardon d'Anduze basin, and exhausted the interest of such coupling for decision makers.

Concerning the use of RADAR data for rainfall estimation, only a few years of data have been recorded for extracting several events, though this would not be sufficient for machine learning. Raw RADAR information was not considered to be reliable enough in mountainous regions, or in case of localised rainfall (Sun et al., 2000), and would consequently be calibrated with rain gauges. Hence, RADAR information is continuously evolving; such rainfall information was, thus, not calculated using a stationary process, preventing then to capitalise on ancient events. Météo France is, thus, currently working on a project of reanalysis of raw RADAR information in order to provide uniformly treated data for France (Tabary et al., 2012) at sampling rate of one hour. Such more reliable information would increase the ability of physically-based models, but will necessitate treatments to be useful for machine learning models, considering the very high number of inputs provided by RADAR images.

### 2.3.2 Neural networks models

The issue regarding physically based models leads to many paths of investigation, for which neural networks seem to offer an interesting alternative paradigm. As "black-box" models, these networks do not presume any a priori behaviour, given that the model construction is data-driven and the parameters are devoid of physical significance. Thus, due to both their universal approximation (Hornik et al., 1989) and parsimony (Barron, 1993) properties, neural networks have become increasingly prevalent in the field of hydrology. Maier and Dandy (2000) provided an overview of neural networks use in hydrology and enhanced their application through forecasting. Nevertheless, the excellent capabilities that neural networks prove for training must be counterbalanced by their ability to reliably generalise to unknown dataset. This trap is well known in machine learning and was formalised as the bias-variance tradeoff (Geman et al., 1992).

To better cope with the bias variance tradeoff, regularisation methods must be used. For this purpose Coulibaly et al. (2000) used early stopping (Sjöberg and Ljung, 1992) and Sudheer et al. (2002) used cross-validation (Stone, 1974). Several studies, however, show a performance assessment not based on an independent dataset, which could be the cause of result quality overestimation. Corzo and Solomatina (2007) sought to optimize neural networks models for the purpose of flood forecasting. They relied upon modular modelling, which allows enhancing forecasting and better represents the processes of a complex natural system. Pereira Filho and Dos Santos (2006) modelled a highly urbanised basin using neural networks, assisted by data derived from remote sensing. Strong sensitivity to both data quality and the number of events in the training set can be observed, with the tremendous capability of this model to be run without process-related certainties being highlighted.

Concerning flash floods, Sahoo et al. (2006) proceeded with flood and water quality forecasting on very small basins. These authors did not use regularisation methods and obtained high quality results, which was likely due to the absence of approximations on hydrometeorological measurements (on a very small basin). Without resorting to regularisation methods, Piotrowski et al. (2006) obtained a low-quality generalisation for the selected test set. To solve the problem of generalisation with respect to intense events, Schmitz and Cullmann (2008) used the most intense event in the training set.

Toukourou et al. (2009) applied neural networks to the Gardon d'Anduze basin, Kong A Siou et al. (2012) to the Lez basin, and both showed that early stopping was providing more parsimonious and more reliable models than weight decay method. As well on the Gardon basin Toukourou et al. (2011) proposed an adaptation of the cross validation to specialise the model on intense events. So-called partial cross validation showed that feedforward models were able to forecast reliably water level up to the response time of the basin,

thereby providing a useful tool for population warning purposes.

It is worth noting that, the feedforward multilayer perceptron, used in the previous cited studies, is fed by rainfall and previous observed discharge (or water level) acting as state variable and playing the role of data assimilation. It is currently considered as the best neural predictor in the field of hydrology. Nevertheless, this attractive quality has limitations in real time forecasting, which is the purpose of the present paper, because real time data transmissions can be disturbed during the rain event, or instrumentation deleted, as it occurred in Gardon d'Anduze basin during the event of September 2002. Discharge measurements can, thus, no longer be available in real time conditions. For this reason it is necessary to focus studies on a model able to provide reliable forecasts without information about previous discharges. This situation is often referred as poorly gauged situation in the literature and has the prospective interest to be transferred to really ungauged basins. Investigate the ability to perform discharge forecasts without previous discharges observation is, thus, the aim of this study.

### 3 Flash flood forecasting on the Gardon de Mialet basin

#### 3.1 Available dataset

The SCHAPI provided 17 yr (from 1992 to 2008) of rainfall and discharge data generated from the discharge gauge and the three rain gauges on the Mialet Basin.

Since the purpose of this study is flood-oriented, we have selected a set of flood events within the database in order to focus the model on flood behaviour. Such a selection of events was based on rainfall so as to capture all types of behaviours, including events occurring at the end of summer when soils are dry and no discharge response to rainfall is observed and those occurring when the response is significant. The rainfall detection threshold was chosen at 100 mm during a 48-h sliding period for at least one of the three rain gauges. This value is not particularly high and typically induces a low hydrological response possible: finding low-intensity events to run statistical learning might prove to be very important in describing the wide variety of behaviour. This extraction method provided us with 58 complete and reliable rainfall/discharge events, without missing gaps.

The majority of selected events had occurred during the autumn (September, October and November accounting for 52 % of all events). Spring and winter were also well represented, whereas only one event occurred during summer (end of August). This distribution is very typical of a Mediterranean climate and can be found on almost all basins affected by flash floods in the Cévennes region.

Event durations range from 26 to 143 h and are equally distributed between three categories: less than 48 h, 48 to 71 h, and greater than 72 h. The beginning of an event is defined

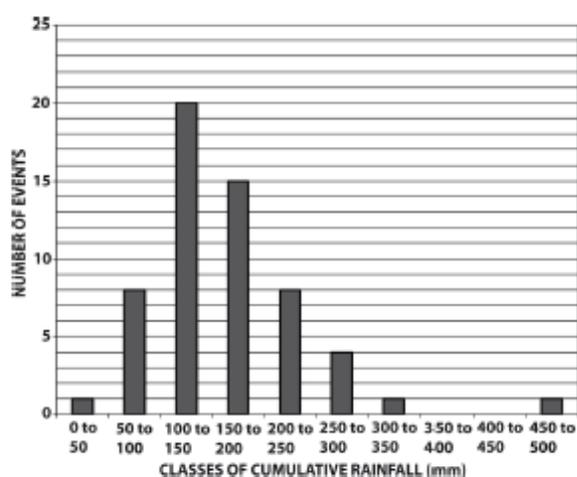


Fig. 3. Distribution of average cumulative rainfall (ACR) in events extracted from the database vs. 50-mm class.

when the first drop reaches one of the rain gauges and we consider that the end happens once the significant part of the hydrological response is complete.

In order to characterise the database, the average cumulative rainfall (ACR) was calculated for each event making a simple average of the cumulative rainfalls of the three rain gauges, without using elevation or location weighting. As each event is particular, it is difficult to state that one calculation method would be more reliable than the others. Consequently, in order to avoid adding another bias to data, the authors considered a simple average, without any hypothesis about rainfall episodes average structures.

The ACR varied from 44 to 462 mm during the selected events. Most events lie in the 100 to 200-mm range (Fig. 3). It can be noticed that ACR can be lower than the threshold of events selection (100 mm on at least one of the three rain gauges) because of the spatial heterogeneousness of rainfalls.

Even on this small basin, as indicated in Sect. 1, rainfall heterogeneity can be very significant. On the three rain gauges, the mean standard deviation of cumulative rainfall per event equals 33 % of the average cumulative rainfall and for some events can exceed 100 %. It would be easy to imagine even higher values when considering hourly rainfall, insofar as the average annual cumulative rainfall presented in Fig. 1 already shows great heterogeneity.

The hourly rainfall rate was also studied and results reveal a tendency favouring lower intensities in Barre-des-Cévennes and higher intensities in Mialet and Saint-Roman-de-Tousque. This observation is consistent with the calculations in Ceresetti (2011) concerning the probability of occurrence of extreme rainfall intensities, which increases when receding from areas of relief in the Cévennes region.

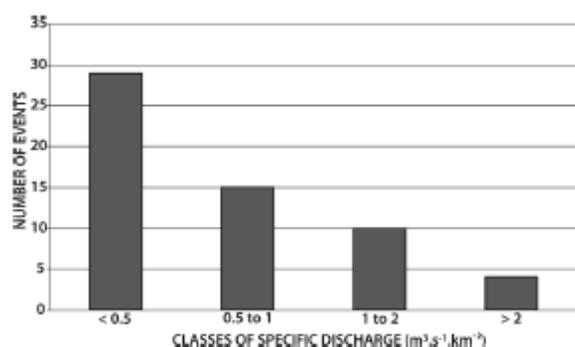


Fig. 4. Distribution of specific discharges in events extracted from the database vs. class.

Among these 58 events, roughly one-fourth can be qualified as intense events, i.e., in our opinion, the specific discharge reaches at least  $1 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$  (Fig. 4). The ultimate purpose of this study is focused on intense events, which are known to be the most hazardous for personal injury and the most destructive for infrastructure (MEDDTL, 2011). Nevertheless, the database must include both types of events (intense and non-intense) in order to effectively assess the capability to avoid false warnings.

In order to compare the rainfall and runoff volumes, the runoff coefficient was calculated for each event (Fig. 5). It can be noticed that an inherent inaccuracy exists in this calculus due to the arbitrary choice of the end of the event.

Runoff coefficients are often quite low suggesting that delayed runoff is important in the hydrological behaviour of the basin. Autumn marks an increase of runoff coefficient while soils are getting wetter. In November, the average coefficient reaches 45% on average and stays high throughout winter and spring. Nearly no event is available in summer but the only event in August and the events of September suggest that the runoff coefficient is very low during this period. This is consistent with the climatic observations for summer period (droughts and high temperatures).

It is worth noting that the signals recorded during flash floods contain a large proportion of noise and inaccuracy. Rainfall is measured by rain gauges, which represent the most accurate sensors currently available, yet the accuracy is estimated at roughly 10% or 20% (Marchandise, 2007). These signals, however, do provide local information, given that rainfall heterogeneity plays a major role in flash floods. To obtain a more representative depiction, radar acquisitions of rainfall with a  $1\text{-km}^2$  definition have been undertaken since 2002, though the number of flood events monitored within the basin under investigation was still too small to allow for reliable use in a machine learning approach.

The sampling period for discharge during flood events contained in the database was 1 h before 2002 and 5 min since 2002. In the present work, a 30-min sampling period

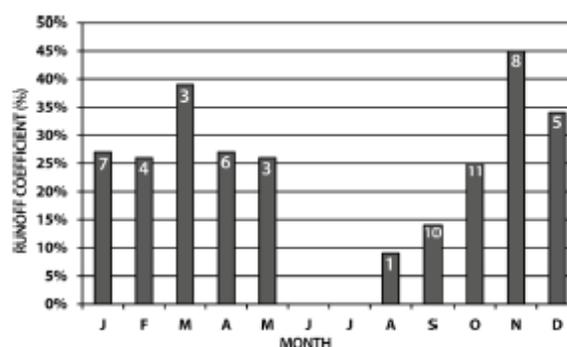


Fig. 5. Average runoff coefficient per month (number of events used to calculate the average is indicated inside bars).

was chosen, which is appropriate considering both the organisational constraints of SCHAPI and this kind of basin response time. For events occurring prior to 2002, re-sampling was performed by means of linear interpolation, which obviously does not provide the missing information, but still allows managing events without differentiation.

### 3.2 Requirements

As previously presented in Sect. 2.3, flash floods occur on small basin having a few hours response time. Rainfall forecasts are not actually available at so small scale of time. Faced with this major drawback, numerous works focused on rainfall forecasting, ensemble forecasting or consequences on discharge forecasting of such lack of information. In the present study, a different option was chosen which consists in considering that discharge forecasts must be carried out without any information about future rainfalls. Accordingly, no hypothesis like null rainfall, or constant rainfall would be assumed. The maximum forecasting lead time would, thus, be the concentration time of the basin. Considering that: (i) Dolciné et al. (2001) estimated the response time of the Gardon d'Anduze was around 3 h, (ii) Toukourou et al. (2011) studied forecasting on Gardon d'Anduze basin with the same kind of approach as the present paper and lead times up to 5 h, (iii) the Gardon d'Anduze and Gardon de Mialet basins are nested, the first one (the larger) being more complex in terms of geology, land uses, slopes and hydrographic network, the second one (the smaller) having a shorter response time, thus, it would be adequate to study forecasting up to 2 h ahead without any hypothesis about future rainfall.

As the sampling period is 30-min, the experiments were conducted for forecast lead times from  $k+1$  to  $k+4$  (extending from 30 min to 2 h) and tested on four major events described in Sect. 3.4.3.

### 3.3 Performance assessment

In order to assess the quality of the resultant forecasting, several methods have been developed.

#### 3.3.1 Level of vigilance

The Vigicrues map displayed on the [www.vigicrues.gouv.fr](http://www.vigicrues.gouv.fr) site offers four watch levels:

- Green: no specific vigilance required.
- Yellow: risk of flood or rapidly rising water level, though not causing any significant damage and only requiring special vigilance for exposed activities.
- Orange: a flood causing significant overflows likely to exert a major impact on community life, property and personal safety.
- Red: major flood risk. Direct and general threat to property and personal safety.

For each area covered by the SCHAPI service, a Flood Information Rule is published. The Cevennes Range is monitored by the Grand Delta flood forecasting service; its Flood Information Rule (Préfecture du Gard, 2010) provides approximate discharge values for the various levels of vigilance in each basin (Table 4). This information will help us in assessing the relevance of forecasts with respect to practice of the local flood forecasting service.

The model performance will be assessed with respect to its ability to indicate to the forecaster the appropriate maximum level within the range of the forecast lead time.

However, the decision to broadcast a level of vigilance is made not only by monitoring the predicted discharge. Forecasters must also identify local issues that can be correlated with a specific period during the year (e.g., campsite filling rate, popular events), thus, allowing them to modulate their decision criteria. In summary, the information provided by the level of vigilance forecasting will not be adequately thorough, prompting us to introduce other criteria that measure efficiency while evolving continuously.

#### 3.3.2 Nash and the persistence criteria

The Nash criterion (Nash and Sutcliffe, 1970) is often used in the field of hydrology; it corresponds to the  $R^2$  determination coefficient, i.e.:

$$R^2 = 1 - \frac{\sum_{k=1}^n (s_k^p - s_k)^2}{\sum_{k=1}^n (s_k^p - \bar{s}^p)^2}$$

where:

- $s_k^p$  is the observed value at time  $k$

- $s_k$  is the simulated value at time  $k$
- $n$  is the number of observed couples/simulated values targeted by the simulation
- $\bar{s}^p$  is the average observed value on the  $n$ -sized sample.

This criterion must be close to one, which means that the predicted discharge is close to the observed discharge. A 0 value represents an average discharge equivalent forecasting, whereas a negative value indicates that the forecasting provided is even worse than the simple average of the observed value during the event. Generally speaking, for flash floods purposes, a Nash criterion value greater than 0.8 is considered satisfactory. However, especially when using a feedforward model, a risk for the model to provide a naive forecasting (when the model provides the same value at the forecast lead time as the one observed at the instant of forecasting) exists. That kind of result generally induces, for short lead times, a good value of the Nash criterion, whereas the model does not bring any information. In order to assess the forecasting provided compared to the naive forecasting, the persistence criterion (Kitadimis and Bras, 1980) has been defined:

$$C_p = 1 - \frac{\sum_{k=1}^n (s_{k+l}^p - s_{k+l})^2}{\sum_{k=1}^n (s_{k+l}^p - s_k^p)^2}$$

where:

- $l$  is the forecast lead time,
- the other forms are similar to Nash criterion ones.

This criterion must be close to one too. A 0 value represents the score of the naive forecasting and a negative value means that the forecasting is even worse than the naive forecasting.

Since the quality of the peak discharge forecasting proves most important in terms of safety, other criteria more sensitive to this quality have been added.

#### 3.3.3 Peak discharge

In addition to the Nash and persistence criteria, we introduced complementary indicators focusing on the quality of peak forecasting:

- The percentage of peak discharge (PPD): this ratio compares the estimated and observed maximum peak discharges and is expressed as follows:

$$PPD = \frac{s_{\max}^p}{s_m^p}$$

where:

$s_{\max}^p$  is the estimated peak discharge and  $s_m^p$  the observed peak discharge.

- The synchronous percentage of the peak discharge (SPPD): this ratio compares the estimated discharge and observed discharge at the time of the observed peak discharge; it can be written as:

$$SPPD = \frac{s_m}{s_m^p}$$

where:

$s_m$  is the estimated discharge at the time of observed peak discharge and  $s_m^p$  the observed peak discharge.

- Lag: this value indicates the time delay between observed and measured peak discharges. It is expressed in step times, which for this study is in half-hour intervals. If this value is positive, then the forecasting is late; on the other hand, if it is negative, the forecasting is early.

Ultimately, it is easily understood that the SPPD criterion is the most critical in terms of safety. The other indicators, however, ensure that the global information on performance is as complete and accurate as possible.

### 3.4 Neural network model design

#### 3.4.1 Generic architecture

As often shown in literature, the standard multilayer perceptron seems to be the best candidate to provide flash flood forecasting. Nevertheless it was pointed out that such a model requires previous discharge information, which can be unavailable in a real-time situation. It is, thus, necessary to dispose of another model able to identify dynamic nonlinear behaviour. Such a model exists in the framework of nonlinear system theory and capitalises on estimated previous discharges as state variable. This model is called recurrent because the estimated output is fed back as state variable. This loop provides a dynamical behaviour to the model as the output can evolve continuously even when faced with constant exogenous input variables (in this study: the rainfalls). Generally the recurrent model was proved providing less accurate results than the feedforward one (Johannet, 2010). This study will be a good way of comparing performances of both models achieved with feedforward and recurrent design.

Moreover, as has been performed in numerous modelling approaches, the linear and nonlinear parts of the process can be identified separately. We have, thus, chosen to combine a linear model with a nonlinear one in order to allow the nonlinear part to focus exclusively on nonlinear relations. Indisputably, the ease with which such a separation can be achieved in a unique formalism offers one of the great advantages of neural networks. As this variant was proved useful in Artigue et al. (2011), specifically for recurrent model, we extended in the present work the assessment of such design to the sixth possible following generic architectures:

- Fully linear (feedforward or recurrent).

- Fully nonlinear: multilayer perceptron (feedforward or recurrent).

- Linear and nonlinear combined model (feedforward or recurrent).

We proceed in this section in a synthetic comparison between all these architectures and focus in a second stage (Sect. 3.4) on the detailed presentation of the chosen architecture ( $k, l, r, m, r_{nn}, m_{lin}, m_{nn}$  are positive integers).

The fully linear perceptron consists of only one output linear model. It can be noticed that this model is equivalent to a multiple linear regression model. The output can be expressed in the feedforward version as:

$$s(k+l) = g_{LIN}(q(k), q(k-1), \dots, q(k-r), u(k), u(k-1), \dots, u(k-m))$$

where  $s$  is the estimated discharge,  $k$  is the discrete time (sampled each half hour)  $q$  is the measured discharge,  $u$  is the vector of exogenous variables (rainfalls),  $r$  is the order of the model,  $m$  is the width of the sliding window of rainfalls information,  $l$  is the lead time of forecasts,  $g_{LIN}$  is the linear function implemented by a feedforward neural network.

In the recurrent linear model, one has (with the same notations):

$$s(k+l) = g_{LIN}(s(k+l-1), s(k+l-2), \dots, s(k+l-r), u(k), u(k-1), \dots, u(k-m))$$

In the standard feedforward multilayer perceptron (MLP), one has (with the same notations):

$$s(k+l) = g_{NN}(q(k), q(k-1), \dots, q(k-r), u(k), u(k-1), \dots, u(k-m))$$

where  $g_{NN}$  is the nonlinear function implemented by the feedforward neural network.

In the recurrent multilayer perceptron (MLP) one has (with the same notations):

$$s(k+l) = g_{NN}(s(k+l-1), s(k+l-2), \dots, s(k+l-r), u(k), u(k-1), \dots, u(k-m))$$

In the feedforward linear and nonlinear combined model one has (with the same notations):

$$s(k+l) = g_{LIN}(u(k), u(k-1), \dots, u(k-m_{lin})) + g_{NN}(q(k), q(k-1), \dots, q(k-r_{nn}), u(k), u(k-1), \dots, u(k-m_{nn}))$$

where the length of the sliding rainfalls window of the linear part is denoted  $m_{lin}$ , the order of the nonlinear part is noted  $r_{nn}$ , the length of the sliding rainfalls window of the nonlinear part is denoted  $m_{nn}$ .

**Table 1.** Mean performances of assessed models. Performances are synthesized as it will be shown in Sect. 3.5. Lag is the delay between the maximum observed discharge and maximum forecast discharge.

Mean performances		$N$	$C_p$	PPD	SPPD	Lag (0.5 h)
Pure linear, feedforward	Average	0.87	0.55	99 %	82 %	1
	Std. Dev.	0.11	0.11	5 %	12 %	0.55
Pure linear, recurrent	Average	0.68	-2.87	56 %	54 %	0.81
	Std. Dev.	0.12	6.57	5 %	5 %	1.17
Standard MLP, feedforward	Average	0.90	0.62	86 %	81 %	0.44
	Std. Dev.	0.11	0.20	15 %	13 %	0.51
Standard MLP, recurrent	Average	0.67	-2.9	55 %	52 %	1.44
	Std. Dev.	0.10	3.71	12 %	12 %	0.89
Combined, feedforward	Average	0.93	0.74	90 %	86 %	0.81
	Std. Dev.	0.05	0.14	6 %	8 %	0.8
Combined, recurrent	Average	0.81	-1.25	72 %	70 %	0.38
	Std. Dev.	0.14	3.6	7 %	8 %	0.81

In the recurrent linear and nonlinear combined model one has (with the same notations):

$$s(k+l) = g_{\text{LNN}}(\mathbf{u}(k), \mathbf{u}(k-1), \dots, \mathbf{u}(k-m_{\text{lin}})) + g_{\text{NN}}(s(k+l-1), s(k+l-2), \dots, s(k+l-r_{\text{nn}}), \mathbf{u}(k), \mathbf{u}(k-1), \dots, \mathbf{u}(k-m_{\text{nn}}))$$

In order to compare performances of all models in a synthetic way, Table 1 provides the mean performances of each model. Criteria were calculated by simple average over 4 various lead times and upon 4 intense events as it will be more explained in Sect. 3.4.2. Accurate architectures were adjusted for each model as presented in depth in the following section.

First it can be observed that recurrent models are not as efficient as their feedforward equivalent; this point is well known. Equally, it appeared that considering Nash criteria or persistency criteria, the combined model outperforms both linear and MLP models. The apparent good success of the linear model for the PPD and SPPD conceals a great overestimation of the flood peak. Regarding the lag time, it is interesting to note that it was better in the combined model for the recurrent version. This last point is very satisfying because it means that the combined neural model succeeds in catching both the amplitude and the dynamics of the hydrological process.

For these reasons, it is clear that the models implementing combination of linear and nonlinear parts outperform other ones. Such generic architecture is, thus, used in the following sections.

### 3.4.2 Variable selection and optimisation

#### Taking into account bias-variance tradeoff:

Working with neural networks involves an accurate consideration of model complexity. On the one hand, a very complex model adjusts to the underlying function added to the noise included in the data. This phenomenon is well known and labelled “overfitting”. In this particular case, the model is not

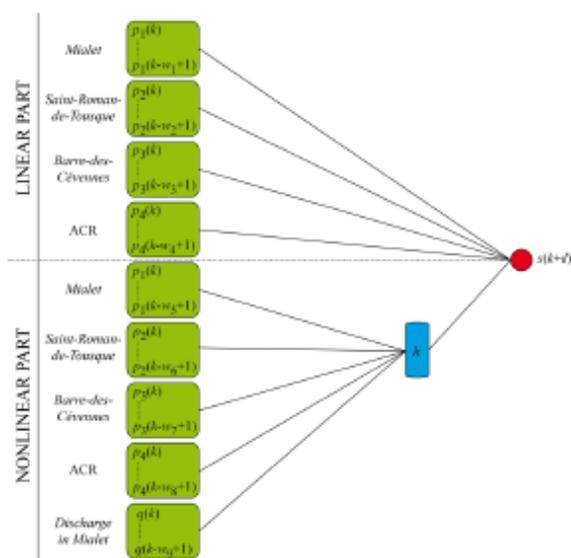
able to generalise to examples different from those used for training. On the other hand, a very simple model lacks the flexibility to adapt to the regression function (Dreyfus, 2005). In both cases, the results obtained are not reliable.

This problem is called the “bias-variance dilemma” (Ge-man et al., 1992). Bias decreases with greater complexity (the model becomes increasingly precise in adapting to the training data), yet variance is increasing at the same time (the model shows greater sensitivity to the training data details). Since bias and variance are positive terms, their sum indicates the minimum for a certain complexity that represents the best tradeoff.

Let us consider the model complexity being represented by the number of free parameters of the model. The variable selection allows for less model complexity in order to enhance the model’s generalisation abilities. Several methods yield a variable selection; these include knowledge-based methods, applied when the phenomenon to be identified is well known and statistics-based methods, employed when the phenomenon is not clearly understood or when the variables available are not mutually independent.

In this study, for example, as follows, the response time of a basin may be relevant information in defining the number of previous rainfall values necessary to produce a forecast (previously termed as window’s width). This information, which is simple and accurate, does not require a detailed understanding of the targeted basin, but merely a statistical analysis of the database; this analysis will be provided in following section.

Nevertheless, knowledge-based methods are necessary, and as shown in the state of the art (Sect. 2.3.1) the information of soil moisture is a necessary state variable, especially for the recurrent model, which does not dispose of this information (contrarily to the feedforward model which disposes of the discharge measurements, themselves influenced par soil moisture). It appeared interesting to input this kind of information to the model by the way of a sliding window



**Fig. 6.** Feedforward model. Green elements are rain gauge input variables or discharge in Mialet (noted  $q$ ), blue element represents hidden layer and red circle represent the linear output neuron. ACR represents the average cumulative rainfall;  $p_n$  refers to the rainfall (instantaneous or cumulative) value and  $w_i$  to the length of the respective sliding window.  $l$  is the forecast lead time. The sliding window of previous estimated discharges is  $r$ : the order of the model. The number of neurons on the hidden layer is  $h$ .

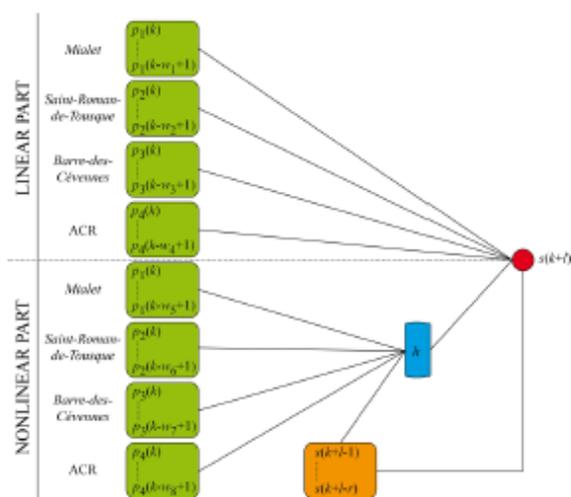
of the average cumulative rainfall defined as following:

$$ACR(k) = \frac{1}{n} \sum_{k_f=0}^k \sum_{n=1}^3 p_n(k_f)$$

Where  $k_f$  represents the time from the beginning of the event to the actual discrete time of forecasting  $k$ , and  $n$  is the index of available rain gauges in the basin; in the present study 3 rain gauges are available.

The  $ACR(k)$  variable obeys to the same definition than the ACR introduced in Sect. 3.1, nevertheless it was calculated in a causal way during the event. The aim is to provide to the model parsimonious information about the previous rainfalls, over a long time horizon, without increasing tremendously the number of variables, and consequently the complexity. Therefore, as the average cumulative rainfall  $ACR(k)$  includes the rainfall memory from the beginning of the event it was applied to the networks, as did for others variables, through a sliding window whose length must be appropriately selected. It can be noticed that if  $ACR(k)$  does not help the model to provide a better forecast it is removed from the architecture by the variable selection process described in next section.

Regarding the number of neurons, one can note that the pure linear model containing only one neuron was shown



**Fig. 7.** Recurrent model. Green elements are rain gauge input variables, orange elements are recurrent input variables (noted  $s$ ), blue element represents hidden layer and red circle represent the linear output neuron. ACR represents the average cumulative rainfall;  $p_n$  refers to the rainfall (instantaneous or cumulative) value and  $w_i$  to the length of the respective sliding window.  $l$  is the forecast lead time. The sliding window of previous estimated discharges is  $r$ : the order of the model. The number of neurons on the hidden layer is  $h$ .

unable to adjust adequately to the underlying function: it is too simple. For both recurrent and non-recurrent models, the separation of the linear and nonlinear behaviours expressed in the combined model helps to cleverly tune the level of complexity of the nonlinear part of the model.

Following these thoughts, the proposed generic architecture is presented in Fig. 6 for the feedforward mode and in Fig. 7 for the recurrent network.

After an accurate tuning of the complexity of the models (number of hidden neurons and widths of temporal windows) hereafter presented, the comparison of the performance of both feedforward and recurrent models will allow assessing the degradation in operational forecast conditions when measured discharges are no longer available.

### Complexity selection

The goal of complexity selection is to tune the number of hidden neurons and variables, for the studied models, namely the values of:  $h, w_1, \dots, w_9, r$ .

Considering the corresponding hydrological response times, the sliding time windows relative to rain gauge inputs were proportioned. The notion, as demonstrated by Kong A Siou et al. (2011), has been to estimate this characteristic time statistically in using cross-correlation between rainfall and discharge on all database events. On the resultant cross correlograms, the first maximum was interpreted as the

**Table 2.** Values of statistically estimated response times (based on cross-correlations), and indications on the range of sliding window lengths investigated using cross-validation (1 stands for Mialet, 2 for Saint-Roman-de-Toussque, and 3 for Barre-des-Cévennes).

Rain gauge	1	2	3
Average response time steps (0.5 h)	4	6	9
Response time steps range (0.5 h)	2–7	5–9	8–11

response time. Only correlograms showing an interpretable maximum have been retained. Among these results, we employed two techniques to select the events considered of high enough magnitude to represent a hydrological response:

- events whose maximum intensity exceeds the median,
- events whose average cumulative rainfall exceeds the median.

Use of these two selection methods also allows us to assess the relevance of results, which lastly were nearly similar for both methods. Cross correlograms were, therefore, plotted for all rain gauges and all selected events. Since each rainfall event does not always involve every rainfall gauge, some rainfall recordings were poorly correlated with the observed discharge. In such cases, no response time can be extracted from the correlogram; these configurations have been removed from the statistical processing.

From these response time estimates, the sliding rainfall window dimension was accurately adjusted for the three investigated rain gauges by means of cross-validation, as proposed by Kong A Siou et al. (2011). Training was performed using Levenberg-Marquardt second-order method in recognition of its strong performance. Cross validation is used jointly with early stopping.

It is important to recall that the model uses strictly no rainfall information after the instant of forecasting  $k$ . As well, strictly no observed discharge is used after the instant of forecasting  $k$ .

Table 2 presents the values of the statistically estimated response times as well as the range of sliding window lengths examined using cross-validation. It is reassuring to note that these response times are increasing when moving from the closest to the furthest rain gauge from the outlet. Equally, the same demarche was followed to select response times of others architectures presented in Sect. 3.4.1. It appeared that they were quasi-similar, thereby proving the robustness of the method. In the same spirit, time window of  $ACR(k)$  is adjusted using cross-validation inside the interval  $[0-7]$ . Selected values are given in Table 3 ( $w_4$  for the feedforward model and  $w_8$  for the recurrent one).

Once the variables have been selected, the optimisation routine relative to the hidden layer and loop or previous discharges (if the model is recurrent) is performed. This

**Table 3.** Optimal architecture selected using cross-validation for various forecast lead times.

Model element	$k+1$ (0.5 h)	$k+2$ (0.5 h)	$k+3$ (0.5 h)	$k+4$ (0.5 h)
$w_1 (p_1)$	2	4	4	5
$w_2 (p_2)$	9	8	8	8
$w_3 (p_3)$	9	9	9	8
$w_4 (p_4)$	5	3	3	3
$w_5 (p_1)$	6	4	5	4
$w_6 (p_2)$	7	7	6	7
$w_7 (p_3)$	10	10	10	9
$w_8 (p_4)$	1	1	1	1
$h$	3	2	2	2
$r/w_9$	1	2	2	2

optimisation step can also be conducted by minimising the cross-validation score obtained on each combination. The optimisations performed have specifically targeted the random initialisations of parameters; these outcomes indicate that the model is fairly insensitive to initial parameters.

It is important to realise that the underlying function to implement at specific  $t_1$  lead time is not the same that the one operating at another  $t_2$  lead time because the future rainfalls are not provided to the model. Considering that the neural “black box” model is not forced to realise strictly the physical process of the watershed, it is, thus, asked to provide implicit rainfall anticipation. In such a case, there is no way for the model to be the same for each lead time. Therefore, it is necessary to design as many models as requested lead time discharge forecasting.

4 models were, thus, adjusted for each lead time ( $k+1$ ,  $k+2$ ,  $k+3$ ,  $k+4$ ) and for both feedforward and recurrent models. The number of recurrent inputs (the order  $r$ ) has been tested from 1 to 5. The optimal architectures are presented in Table 3.

It can be noticed in Table 3 that except for the  $k+1$  model, the order best value was two, which suggests that the model does not only need the sole previous value, but also its variation to perform better. This variation allows the model to distinguish the rising of the peak and its decreasing.

### 3.4.3 Event selection for training, cross-validation, early stopping and testing

The training dataset was straightforward to create: all 58 database events, with the exception of those selected for testing and early stopping, were used. Thus, in all, 56 events were available for the training step.

Cross-validation was performed on the 58 events in order to select the early stopping event as proposed by Toukourou et al. (2011). The best cross-validation score among intense events (i.e., over  $1 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$ ) matches the chosen

**Table 4.** Discharge thresholds corresponding to the levels of vigilance.

Level	Discharge threshold
Green level	$< 105 \text{ m}^3 \text{ s}^{-1}$
Yellow level	$> 105 \text{ m}^3 \text{ s}^{-1}$
Orange level	$> 370 \text{ m}^3 \text{ s}^{-1}$
Red level	$> 600 \text{ m}^3 \text{ s}^{-1}$

early stopping event: this event occurred in January 1998 and reached an intensity of  $250 \text{ m}^3 \text{ s}^{-1}$ .

The cross-validation approach should be applied to the entire training database. Nevertheless, due to the small number of intense events and the desire to produce the best possible model for intense events, the use of cross-validation for selecting the architecture can only proceed on a selection of events, revealing useful characteristics for the purpose of this study. Such a practice is called “partial cross-validation” by Toukourou et al. (2011). Since we are focused here on flash floods, those most intense events are also the most hazardous, we have selected all of the intense events (specific discharge exceeding  $1 \text{ m}^3 \text{ s}^{-1} \text{ km}^{-2}$ ). Moreover, to allow the model to respond quickly and precisely to a rainfall impulse, we have retained the events displaying an “impulse shape”, meaning those events whose discharge response can be clearly correlated with a rainfall impulse. Consequently, this level of model specialisation, referred to as “partial cross-validation”, was applied to 17 events.

The test event could be any in the series, but we opted to test an extreme event, an average event, a minor event and a complex event (i.e., double-peak) in order to generate an overview of the developed model’s efficiency. The interchange of events between the test event and another dataset implies inevitably new training and new parameter’s initialisation selection.

Table 5 lists the characteristics of the selected events for the test step. From a simple examination of these characteristics, the system nonlinearity is quite noticeable: the rainfall-discharge relationship does not depend merely on the observed rainfall amount or hourly intensity.

## 3.5 Results

### 3.5.1 Feedforward model

The feedforward model displays excellent results on all four events tested. Table 6 presents these specific results. These results show that the feed-forward model is able to provide sufficiently accurate forecasting for the vigicrues requirements relative to  $k + 1$  to  $k + 4$ , i.e., for forecast lead times of 30 min to 2 h. The Nash criterion values often lie close to 0.9 and decrease as the forecast lead time increases. For the most intense event (September 2002, event no. 19), the forecasting becomes less accurate for  $k + 4$ , though the predicted

level of vigilance remains accurate. The persistence index is, generally speaking, showing that the model brings better information than a naive forecasting. On average, it is more than 0.7 and rarely under 0.5.

From a temporal standpoint, in most cases examined herein, lags indicate that the forecasting is slightly late, yet the limited degradation of SPPD vs. PPD serves to mitigate this concern.

While not all of the hydrographs will be presented herein, Fig. 12a and b displays the predicted hydrographs for a medium to strong event (no. 13) and the extreme event (no. 19) for the  $k + 3$  forecast lead time (90 min). Let us note the excellent quality of forecasting for this 90-min horizon, even on an extreme event, which is higher than any event in the training set.

The results are satisfactory, which often proves to be the case with this configuration, where previous discharges are inputted to the model. Moreover, positive persistence index show that the model combines discharge and rainfall in order to bring useful information.

In large lead times, it can be noticed that persistency increases. This phenomenon does not mean that the forecasting is better in an absolute way, but that it is easier to work better than the naive forecast for large lead times. This observation was highlighted in Kong A Siou et al. (2011).

### 3.5.2 Recurrent model

The recurrent model was run in a first step with real rainfall data (as measured by the rain gauges without correction) and in a second step with a 20-set range of modified rainfall data, thus, yielding 20 discharge forecasting, in order to enhance results. These results are presented in the current section.

#### Deterministic rainfall:

The recurrent model run with observed rainfall data produced the results shown in Table 6. As expected, this model yields poorer results than the feedforward model, which takes advantage of the measured discharges as state variables. As underscored by the results presented in this section, discharge information is quite important for the model; though the recurrent model is typically capable of providing reliable alternative forecasting when discharge measurement problems occur during floods (damaged devices, transmission problems, etc.). The persistence criteria are better for intense and simple events and, as previously, for large lead times.

Generally speaking, roughly 70 % of the peak discharge is accurately predicted, in many instances in terms of time as well (to within 30 min). The level of underestimation ranges from  $-41\%$  to  $-17\%$  relative to SPPD, with an average of  $-30\%$  and a standard deviation of 8 %. Despite this underestimation, the relative constancy could allow forecasters to adapt their decision during operational applications. For an

**Table 5.** Main characteristics of the events selected for the test step ( $Q_x$ : peak discharge,  $Q_a$ : average discharge,  $P$ : average rainfall during the event and  $I_x$ : maximum observed hourly intensity).

No.	Date	Peaks	$Q_x$ ( $\text{m}^3 \text{s}^{-1}$ )	$Q_a$ ( $\text{m}^3 \text{s}^{-1}$ )	$P$ (mm)	$I_x$ ( $\text{mmh}^{-1}$ )
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

**Table 6.** Results obtained by the feedforward model ( $V_p$ : level of vigilance predicted by the model,  $V_0$ : level of vigilance required given the observed discharge, Lag: delay between the maximum observed discharge and maximum forecast discharge).

Event	$l$ (0.5 h)	$R^2$	$C_p$	PPD	SPPD	$V_p$	$V_0$	Lag (0.5 h)
13	1	0.99	0.73	98%	94%	O	O	1
13	2	0.98	0.92	96%	95%	O	O	1
13	3	0.97	0.96	96%	88%	O	O	1
13	4	0.95	0.95	95%	94%	O	O	1
19	1	0.93	0.42	86%	71%	R	R	-1
19	2	0.94	0.65	89%	89%	R	R	0
19	3	0.94	0.7	94%	94%	R	R	0
19	4	0.78	0.67	87%	82%	R	R	1
8	1	0.97	0.64	99%	99%	Y	Y	0
8	2	0.94	0.75	90%	85%	Y	Y	1
8	3	0.91	0.67	88%	79%	Y	Y	1
8	4	0.88	0.70	83%	77%	Y	Y	1
302	1	0.97	0.65	91%	88%	Y	Y	2
302	2	0.95	0.8	87%	82%	Y	Y	1
302	3	0.93	0.84	83%	82%	Y	Y	2
302	4	0.89	0.79	81%	76%	Y	Y	1
Average		0.93	0.74	90%	86%	-	-	0.81
Std. Dev.		0.05	0.14	6%	8%	-	-	0.8

average event, the synchronization of peaks is better than that obtained by use of the feedforward model.

Here once again, we will not present all of the hydrographs produced, but limit our display to the following ones: event no. 8 for  $k+1$ , no. 13 for  $k+2$ , no. 302 for  $k+3$ , and no. 19 (the most intense of the entire database) for  $k+4$  (Fig. 10). The event 8 shows a good consideration of the hydrological dynamics by the model. Particularly, it is worth noting that the increase of discharge at the beginning of the peaks is anticipated by the model. However, the second peak is slightly underestimated. Even with a little underestimation, the event 13 is well forecast in any case. Its standard shape can be easily linked to the shape of the rainfall signal so that the model, trained on intense and “impulse shaped” events, responds well. The event 302 is the most underestimated, at any instant of the event. Here again, the model reproduces well the general dynamics of the event, but the lack of amplitude is important. Finally, the forecasting on event 19 is of excellent quality especially when recognising that it is the most intense event of the database (roughly two times greater

than the second most intense event) and when recalling that this event, as a test event, was not used in either the training or stopping sets. Such a result proves that neural networks, if used with rigorous regularisation methods, can generalise efficiently to extreme never encountered events.

Surprisingly (except for persistence), it can be observed on few events that performances tend to stabilise, or sometimes slightly increase, when the lead time increases. In order to confirm this observation over larger dataset, the cross-correlation score is shown in Table 8 for the various lead times. One can note that this score is slightly better for  $k+1$ , invalidating thus the previous observation in a general way. One can also note that the cross-validation score actually stabilises from  $k+2$  to  $k+4$  lead times. The last point would suggest that the recurrent model would actually be able to anticipate short term future rainfalls.

#### Ensemble forecasting:

In order to further enhance the results obtained, the uncertainty in rainfall measurements has been considered by

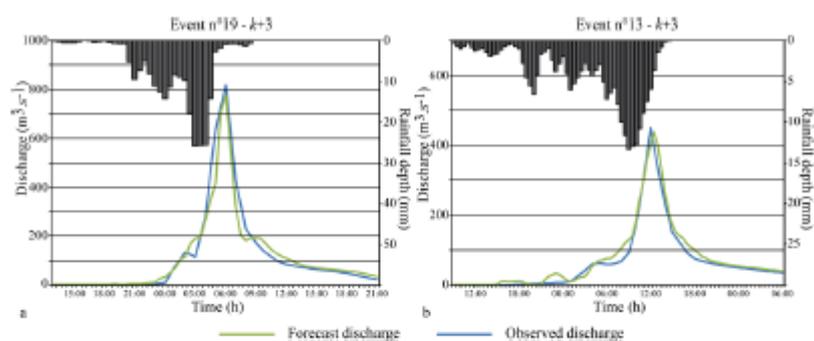


Fig. 8. (a) Predicted and observed hydrographs for the  $k+3$  horizon on event no. 19 using the feedforward model; (b) predicted and observed hydrographs for  $k+3$  horizon on event no. 13 with the feedforward model.

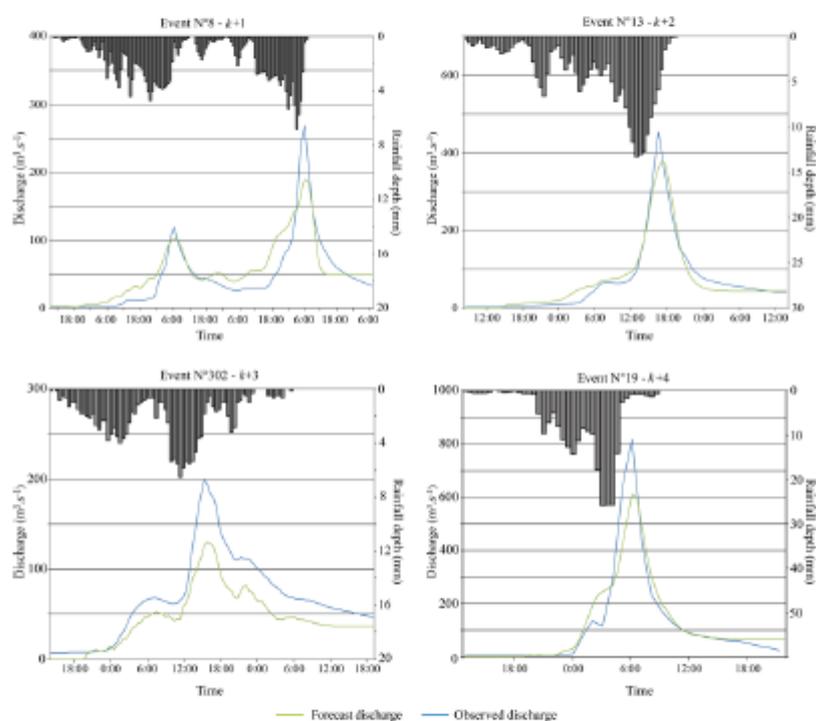


Fig. 9. Forecast and observed discharges for the four tested events with the recurrent model (the forecast lead times are indicated).

introducing a random component in the rainfall values. Future rainfall data were no longer applied to the model.

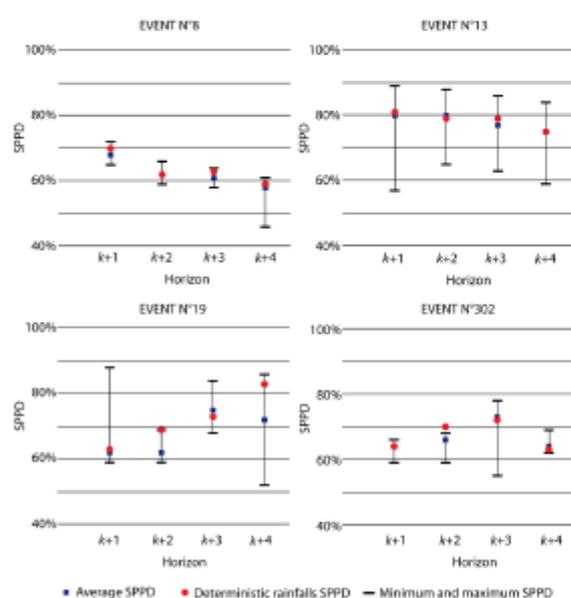
The two main sources of uncertainty are: (i) the rain gauge itself and its location; and (ii) the spatial variability of both rainfall (especially relative to convective rainfall, which is the case within the Cévennes range) and strong winds. In this study, these uncertainties have been estimated at around 20 %, which is consistent with the conclusions by Marchandise (2007).

Consequently, a random modification of the rainfall was performed: each test set value used was multiplied by a random number between 0.8 and 1.2. This process was repeated 20 times, providing 20 different rainfall inputs for the model. No bias was introduced due to the poor knowledge of the rainfall field on the eastern part of the basin, combined with the high rainfall heterogeneity.

In order to compare the results obtained using this model with the deterministic forecasting from Sect. 4.5.2., and as the most representative criterion for our purpose, SPPD was

**Table 7.** Results output by the recurrent model ( $V_p$ : level of vigilance predicted by the model,  $V_o$ : vigilance required given the observed discharge, Lag: delay between the maximum observed discharge and maximum predicted discharge).

Event	$l$ (0.5 h)	$R^2$	$C_p$	PPD	SPPD	$V_p$	$V_o$	Lag (0.5 h)
13	1	0.94	0.07	84%	81%	O	O	1
13	2	0.95	0.65	83%	79%	O	O	1
13	3	0.93	0.82	79%	79%	Y	O	1
13	4	0.93	0.86	77%	75%	Y	O	1
19	1	0.85	-0.25	67%	63%	O	R	-1
19	2	0.92	0.71	72%	69%	O	R	-1
19	3	0.91	0.84	74%	73%	R	R	-1
19	4	0.91	0.89	83%	83%	R	R	0
8	1	0.8	-4.52	70%	70%	Y	Y	0
8	2	0.78	-0.84	62%	62%	Y	Y	0
8	3	0.77	0.33	67%	63%	Y	Y	1
8	4	0.81	0.15	65%	59%	Y	Y	2
302	1	0.63	-13.3	64%	64%	Y	Y	0
302	2	0.64	-3.51	70%	70%	Y	Y	1
302	3	0.57	-2.31	72%	72%	Y	Y	0
302	4	0.57	-0.63	65%	63%	Y	Y	1
Average		0.81	-1.25	72%	70%	-	-	0.38
Std. Dev.		0.14	3.61	7%	8%	-	-	0.81



**Fig. 10.** Summary of results obtained from ensemble forecasting compared with deterministic forecasting output.

chosen as the descriptive parameter. Figure 10 presents the maximum, minimum and average SPPD values derived from the ensemble forecasting, along with the SPPD values output by deterministic rainfall forecasting.

**Table 8.** Evolution of the cross-validation score versus the forecast lead-time.

Lead time	1	2	3	4
Cross validation score ( $0.10^{-6}$ )	161	174	175	179

We have demonstrated herein that no improvement is gained by this method. The average results of ensemble forecasting are, as a general rule, either nearly the same or worse than the output of deterministic forecasting. The lag results, which have not been presented herein, also turn out to be similar. The maximum values can yield excellent results, yet in real-time use this information proves irrelevant. As a matter of fact, the question surrounding the choice of leading ensemble remains unsolved for the forecaster. It is currently recognised that when the average ensemble forecasting is similar to the deterministic forecasting, then the contribution of the ensemble forecasting becomes quite limited.

Nevertheless, this part of the study has sparked discussion about the robustness of the recurrent model. Generally speaking, events nos. 8 and 302 display a lower variability than the others. We analysed this result through a diminished impact of rainfall modification on lower intensities and cumulative rainfalls. In the case of low-intensity events, the model shows relative insensitivity to rainfall uncertainty. In the case of higher-intensity events, the conclusion is rather ambiguous, but with the exception of maximum and minimum values, a low dispersion can be noticed. The model used, therefore, shows limited sensitivity to noise in the test set input data.

If no strict forecasting enhancement is obtained, then model robustness is assessed and can be qualified as reasonable.

#### 4 Discussions

The results obtained prove again the neural networks ability to provide robust and efficient forecasting in complex nonlinear natural systems. Here is targeted a small basin regularly submitted to flash floods.

Previous work in the area was made by Toukourou et al. (2009, 2011) on the Gardon d'Anduze basin (545 km<sup>2</sup>). This basin is bigger than the Gardon de Mialet and not as homogeneous as the Gardon de Mialet basin is, even if they are nested. Consequently, the Gardon d'Anduze behaviour is more complex with lots of sub-basins, high slopes differences, geological heterogeneity... Thus, the maximum forecast lead time can be of the same order in both basins because the Gardon d'Anduze may be more sensitive to rainfall location, thereby reducing the average visibility. For these reasons, it would probably be difficult to build an efficient recurrent model for Anduze.

We showed, in the present study, that for a simple basin, a linear part combined to a nonlinear one in the model brings a significant enhancing to a MLP feedforward model. This behaviour may be linked to the basin behaviour. The transformation of rainfall into discharge is more direct for rapid basin: both signals (average rainfall and discharge) have the same shape so that a linear relation can bring enhancement to the modelling process, while the nonlinear part supports complex relations.

Concerning the shorter lead time  $k + 1$ , forecasts are sometimes worse on the intense events chosen compared to the other events. One can explain this lower quality by the hydrologic behaviour of the basin and by the bias variance tradeoff.

In order to explain this, we should observe Dolciné et al. (2001) work on the Gardon d'Anduze basin. They show that, very short-term rainfall forecast is not necessary to provide accurate discharge forecasting. Even if in the present work, no rainfall forecasting is used, the underlying phenomenon is the same. In both cases, the last rainfall values (forecast or observed) provided to the model are not useful to explain the short-term estimated discharge. Indeed, the rainfall recently observed (one or two step times before the instant of forecasting) is not yet participating to the hydrological response. In our study, this means that some recent values are unnecessary for short forecast lead times while they are still provided to the model. They increase complexity by increasing the number of parameters; taking into account the bias-variance tradeoff, this may degrade the forecasts obtained.

A point is common amongst all of the studies about neural network hydrological modelling: a rigorous application of regularisation methods is a reliable path to design a well-working neural network model. Then, if data are available,

the use of previous measured discharges is recommended for short-term forecasting; for longer terms forecasting it would be useful to take profit of the recurrent model, which seems to grasp the dynamics of the basin better.

Finally, ensemble predictions do not enhance the results obtained by the recurrent model. The advantage is very clear: the sensitivity of the model to rainfall uncertainty is low. On the other hand, no improvement is obtained. In these conditions, if a bias could have been defined (under or overestimating of the rain gauges), we could have been opening another path of inquiry, in which forecasting could have been enhanced considering better the reliability of the rainfall measurements.

#### 5 Conclusions

The potential human and financial losses related to flash floods have made their study and forecasting a very challenging concern. In this paper, we have evaluated six types of models based on machine learning: the first category belongs to feedforward models using previous observed discharge values, while the second represents recurrent models using previous estimated discharge values. In each category, linear, classical MLP and combined linear and MLP models were assessed, each one type was set forth in four versions corresponding to four forecast lead times. It appeared then that combined models were superior to the others and provided very good forecasts up to the response time of the basin without any assumption about future rainfalls. End users can benefit from these results in order to decrease the uncertainty inherent to flash flood forecasting.

A rigorous variable selection process (using basin response times) and an accurate application of regularisation methods (early stopping, cross-validation) have highlighted the ability of neural networks to model nonlinear recurrent systems such as rapid basins. Their parsimony is highly valued in the context of flash flood forecasting, as characterised by a poorly known hydrological context, missing data and partial data unavailability during real-time use.

As exhibited in the literature, the feedforward model is the more efficient. Applied to the Mialet Basin, it yields efficient forecasting on any kind of event tested up to a 2-h horizon. In the case of data transmission difficulties or discharge gauge damage, the recurrent model outputs forecasting in a degraded mode. Though efficiency is lower, still in most instances 70% of the peak discharge is predicted. In order to raise this ratio value, ensemble forecasting that take into account the great uncertainty on rainfall inputs were developed. Unfortunately, the average results of these forecasting were nearly equal to those obtained using the simple input version. The longer computation time these forecasting require is not justified by the improvement in results. Nevertheless, the low sensitivity of neural networks to the noise and inaccuracy found in the test set has been demonstrated. This

property constitutes another advantage of the neural network approach.

It can be highlighted that such conclusions are generic and can be applied to any other basin in which data would be available for a sufficient number of events. Next, by taking soil moisture into account using associated indices or previous long-term rainfall series, the model could be enhanced, especially when considering the great seasonal variations in initial conditions. Lastly, after enhancing these forecasting, a generalisation of recurrent model forecasting to ungauged basins remains the main purpose of our future research.

*Acknowledgements.* The authors would like to thank SCHAPI for the availability of their BVNE project, with special mention of Caroline Wittwer and Arthur Marchandise for supplying data and hosting constructive discussions on flash flood forecasting. Sincere thanks are also addressed to the Hydrosciences Montpellier and Centre des Matériaux de Grande Diffusion research teams. Our gratitude is extended to the Civil Engineering student body of the Ecole des Mines d'Alès School, particularly Guillaume Burg and Pierre Etchart whose work has contributed to these research projects. The constant efforts made by Dominique Bertin (with the Geonosis Company) to enhance and develop the neural network software RNF Pro are hereby acknowledged as well. And lastly, special thanks go to the French National Research Agency for its support of the FLASH project (ANR-09-SYSC.004), which has contributed to the present work.

Edited by: N. R. Dalezios

Reviewed by: two anonymous referees

## References

- Alfieri, L., Smith, P. J., Thielen-del Pozo, J., and Beven, K. J.: A staggered approach to flash flood forecasting – case study in the Cévennes region, *Adv. Geosci.*, 29, 13–20, doi:10.5194/adgeo-29-13-2011, 2011.
- Artigue, G., Johannet, A., Borrell, V., and Pistre, S.: Flash floods forecasting without rainfalls forecasts by recurrent neural networks, Case study on the Mialet basin (Southern France), Proceedings of the 3rd world congress on Nature and Biologically Inspired Computing, 19–21 Oct. 2011 in Salamanca, 303–310, 2011.
- Ayral, P. A.: Contribution à la spatialisation du modèle opérationnel de prévision des crues éclair ALHTAIR, approches spatiale et expérimentale, application au bassin-versant du Gardon d'Anduze, PhD Thesis, Université de Provence, 311 pp., 2005.
- Barron, A. R.: universal approximation bounds for superpositions of a sigmoidal function, *IEEE Transactions on Information Theory* IT-39, 930–945, 1993.
- Bartholmes, H. and Todini, E.: Coupling meteorological and hydrological models for flood forecasting, *Hydrol. Earth Syst. Sci.*, 9, 333–346, doi:10.5194/hess-9-333-2005, 2005.
- Beven, K. J. and Kirkby, M. J.: A physically based variable contributing area model of basin hydrology, *Hydrol. Sci. Bull.*, 24, 43–69, 1979.
- Beven, K. J., Quinn, P., Romanovicz, R., Freer, J., and Lamb, R.: TOPMODEL, in: *Computer models of watershed hydrology*, edited by: Singh, V. P., Water Resources Publications, Highlands Ranch, Colo., 627–668, 1995.
- Borrell, V., Chorda, J., and Dartus, D.: Prévision des crues éclair, *C. R. Geoscience*, 337, 1109–1119, 2005.
- Ceresetti, D.: Structure spatio-temporelle des fortes précipitations : application à la région Cévennes-Vivarais, PhD Thesis, Université de Grenoble, Laboratoire d'étude des Transferts en Hydrologie et Environnement, 283 pp., 2011.
- Corzo, G. and Solomatine, D.: Knowledge-based modularization and global optimisation of artificial neural network models in hydrological forecasting, *Neural Networks*, 20, 528–536, 2007.
- Coulibaly, P., Anctil, F., and Bobee, B.: Daily reservoir inflow forecasting using artificial neural networks with stopped training approach, *J. Hydrol.*, 230, 244–257, 2000.
- Coustau, M., Ricci, S., Borrell, V., Piacentini, A., Thual, O., Bouvier, C., and Marchandise, A.: Calibration of rainfall runoff hydrological model for flash flood forecasting using data assimilation, EGU, poster, European Geosciences Union General Assembly 2011 Vienna, Austria 3–8 April 2011, 2011.
- Dolciné, L., Andrieu, H., Sempere-Torres, D., and Creutin, D.: Flash flood forecasting with coupled precipitation model in mountainous Mediterranean basin, *J. Hydrol. Eng.*, 6, 1–10, 2001.
- Dreyfus, G.: *Neural networks, methodology and applications*, Springer, Berlin, 2005.
- Gaume, E.: *Éléments d'analyse sur les crues éclair*, PhD Thesis, Ecole Nationale du Génie Rural, des Eaux et Forêts and Institut National de la Recherche Scientifique, 359 pp., 2002.
- Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičová, L., Bloschl, G., Borga, M., Dumitrescu, A., Daliakopoulos, I., Garcia, J., Irimescu, A., Kohnova, S., Koutroulis, A., Marchi, L., Matreata, S., Medina, V., Preciso, E., Sempere-Torres, D., Stancalie, G., Szolgay, J., Tsanis, I., Velasco, D., and Viglione, A.: A compilation of data on European flash floods, *J. Hydrol.*, 367, 70–78, 2009.
- Geman, S., Bienenstock, E., and Doursat, R.: Neural networks and the bias/variance dilemma, *Neural Comput.*, 4, 1–58, 1992.
- Hornik, K., Stinchcombe, M., and White, H.: Multilayer Feedforward Networks Are Universal Approximators, *Neural Networks*, 2, 359–366, 1989.
- Johannet, A.: Artificial Neural Networks, *Environmental hydraulics 2, Mathematical Models*, 419–443, 2010.
- Kessler, J. and Chambraud, A.: *Météo de la France, tous les climats localité par localité*, J.C. Lattès, 391 pp., 1990.
- Kitadimis, P. K. and Bras, R.: Real time forecasting with a conceptual hydrologic model, applications and results, *Water Resour. Res.*, 16, 1034–1044, 1980.
- Kong A Siou, L., Johannet, A., Borrell, V., and Pistre, S.: Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France), *J. Hydrol.*, 403, 367–380, 2011.
- Kong A Siou, L., Johannet, A., Borrell, V., and Pistre, S.: Optimization of the generalisation capability for rainfall-runoff modelling by neural networks: the case of the Lez aquifer (southern France), *Environ. Earth Sci.*, 65, 2365–2375, 2012.
- Le-Lay, M. and Saulnier, G. M.: Exploring the signature of climate and landscape spatial variabilities in flash flood events: case of

- the 8–9 September 2002 Cévennes-Vivarais catastrophic event, *Geophys. Res. Lett.*, 34, L13401, doi:10.1029/2007GL029746, 2007.
- Maier, H. R. and Dandy, G. C.: Neural networks for the forecasting and forecasting of water resources variables: a review of modelling issues and applications, *Environ. Model. Softw.*, 15, 101–124, 2000.
- Marchandise, A.: Modélisation hydrologique distribuée sur le Gardon d'Anduze: étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et tests d'hypothèses sur le processus hydrologiques, PhD Thesis, Université des Sciences et Techniques du Languedoc (Montpellier), 2007.
- Météo France, Jacq, F.: Pluies Extrêmes en France Métropolitaine, available at: <http://pluiesextremes.meteo.fr/index.php>, last access: 28 July 2011.
- MEDDTL, Ministère de l'Ecologie, du Développement Durable, des Transports et du Logement, available at: <http://www.vigicrues.gouv.fr> (last access: October 2012), 2011.
- Moussa, R., Chahinian, N., and Bocquillon, C.: Distributed hydrological modelling of a Mediterranean mountainous catchment – model construction and multi-site validation, *J. Hydrol.*, 337, 35–51, 2007.
- Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- Pereira Filho, A. J. and Dos Santos, C. C.: Modeling a densely urbanized watershed with an artificial neural network, weather radar and telemetric data, *J. Hydrol.*, 317, 31–48, 2006.
- Piotrowski, A., Napiórkowski, J. J., and Rowiński, P. M.: Flash-flood forecasting by means of neural networks and nearest neighbour approach – a comparative study, *Nonlin. Processes Geophys.*, 13, 443–448, doi:10.5194/npg-13-443-2006, 2006.
- Préfecture du Gard: Direction Départementale de l'Équipement, Règlement de surveillance, de prévision et de transmission de l'information sur les crues, available at: [http://www.gard.equipement.gouv.fr/IMG/pdf/RIC\\_SPCGD\\_version\\_2011\\_cle7bd5c2.pdf](http://www.gard.equipement.gouv.fr/IMG/pdf/RIC_SPCGD_version_2011_cle7bd5c2.pdf) (last access: October 2012), 2006.
- Rouzeau, M., Martin, X., and Pauc, J. C.: Retour d'expérience des inondations survenues dans le département du Var les 15 et 16 Juin 2010, Rapport du Ministère de l'Ecologie, de l'Énergie, du Développement Durable et de la Mer, en Charge des Technologies Vertes et des Négociations sur le Climat, Ministère de l'Intérieur, de l'Outre-Mer et des Collectivités Territoriales, Secrétariat d'Etat Chargé de l'Ecologie, Inspection générale de l'Administration, Conseil Général de l'Environnement et du Développement Durable, 87 pp., 2010.
- Sahoo, G. B., Ray, C., and De Carlo, E. H.: Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii, *J. Hydrol.*, 317, 63–80, 2006.
- Schmitz, G. H. and Cullmann, J.: PAI-OFF: A new proposal for on-line flood forecasting in flash flood prone catchments, *J. Hydrol.*, 360, 1–14, 2008.
- Sjöberg, J. and Ljung, L.: Overtraining, regularisation and searching for minimum in neural networks, Preprint IFAC Symposium on Adaptive Systems in Control and Signal Processing, 1992.
- Stone, M.: Cross-validated choice and assessment of statistical forecasting, *J. Roy. Stat. Soc.*, B36, 111–147, 1974.
- Sudheer, K. P., Gosain, A. K., Rangan, D. M., and Saheb, S. M.: Modelling evaporation using an artificial neural network algorithm, *Hydrol. Process.*, 16, 3189–3202, 2002.
- Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F.: Flood estimation using radar and raingauge data, *J. Hydrol.*, 239, 4–18, 2000.
- Tabary, P., Dupuy, P., L'Henaff, G., Laurantin, O., Merlier, C., and Soubeyroux, J. M.: A 10-year (1997–2006) reanalysis of quantitative precipitation estimation over France: methodology and first results, *Hydrol. Sci. J.*, in press, 2012.
- Toukourou, M. S., Johannet, A., and Dreyfus, G.: Flash flood forecasting by statistical learning in the absence of rainfall forecast: a case study, *Engineering Applications of Neural Networks Proceedings*, 98–107, 2009.
- Toukourou, M., Johannet, A., Dreyfus, G., and Ayrat, P.-A.: Rainfall-runoff modelling of flash floods in the absence of rainfall forecasts: the case of “Cévenol flash floods”, *J. Appl. Intell.*, 35, 1078–1089, 2011.
- Tramblay, Y., Bouvier, C., Martin, C., Didon-Lescot, J. F., Todorovik, D., and Domergue, J. M.: Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling, *J. Hydrol.*, 387, 176–187, 2010.
- Vinet, F.: Geographical analysis of damage due to flash floods in southern France: The cases of 12–13 November 1999 and 8–9 September 2002, *Appl. Geogr.*, 28, 323–336, 2008.
- Younis, J., Anquetin, S., and Thielen, J.: The benefit of high-resolution operational weather forecasts for flash flood warning, *Hydrol. Earth Syst. Sci.*, 12, 1039–1051, doi:10.5194/hess-12-1039-2008, 2008.

### ANNEXE 3 :

Les évènements ont été numérotés de la manière suivante :

- <200 : évènements déjà disponibles dans la base de données du projet BVNE Gardons pour le Gardon de Mialet.
- 200 à 299 : évènements au pas de temps horaire convertis au pas de temps de la demi-heure pour le Gardon de Mialet.
- 300 à 399 : évènements au pas de temps de cinq minutes convertis au pas de temps de la demi-heure pour le Gardon de Mialet.
- 900 à 999 : évènements issus de diverses sources sur les 15 bassins visés par la généralisation et utilisés pour le travail de généralisation aux bassins non jaugés.

#### Évènements sur le Gardon de Mialet :

N°	Date et heure début	Durée (h)	Débit de pointe (m <sup>3</sup> .s <sup>-1</sup> )	Source
1	22/9/94 20:00	36	337	BVNE
3	3/10/95 23:00	45	531	BVNE
4	13/10/95 14:00	29	313	BVNE
8	10/11/96 10:00	82	269	BVNE
13	28/9/00 9:00	46	454	BVNE
19	8/9/02 13:00	33	819	BVNE
23	18/10/06 12:00	48	102	BVNE
25	20/11/07 22:00	67	53	BVNE
26	20/10/08 22:00	42	110	BVNE
27	31/10/08 11:00	75	528	BVNE
117	20/10/99 4:30	39	263	BVNE
120	12/11/00 2:00	71	80	BVNE
200	22/9/93 5:00	68	118	BVNE
201	10/10/93 6:00	119	54	BVNE
202	31/10/93 7:00	124	79	BVNE
203	3/2/94 4:00	67	69	BVNE
204	23/4/94 7:00	61	46	BVNE
205	19/10/94 8:00	53	377	BVNE
206	2/11/94 11:00	102	304	BVNE
207	17/1/95 10:00	50	53	BVNE
208	10/9/95 22:00	62	31	BVNE
209	24/11/95 17:00	94	132	BVNE
210	7/12/95 21:00	48	60	BVNE
211	15/12/95 12:00	79	109	BVNE
212	9/1/96 4:00	50	127	BVNE
213	11/1/96 22:00	86	193	BVNE
214	20/1/96 9:00	143	312	BVNE
215	25/2/96 17:00	71	60	BVNE
216	13/3/96 18:00	74	37	BVNE
217	22/4/96 14:00	54	20	BVNE

**Evènements sur le Gardon de Mialet (suite) :**

N°	Date et heure début	Durée (h)	Débit de pointe (m <sup>3</sup> .s <sup>-1</sup> )	Source
218	18/9/96 20:00	40	95	BVNE
219	13/10/96 0:00	80	43	BVNE
220	4/12/96 2:00	108	66	BVNE
221	27/8/97 13:00	35	11	BVNE
222	12/1/98 14:00	52	248	BVNE
223	14/4/98 0:00	45	102	BVNE
224	29/4/98 14:00	50	120	BVNE
225	27/5/98 14:00	40	44	BVNE
226	25/9/98 11:00	50	61	BVNE
227	28/12/98 6:00	133	143	BVNE
228	17/5/99 7:00	79	299	BVNE
229	11/11/99 18:00	94	72	BVNE
230	22/9/01 6:00	65	6	BVNE
300	14/3/02 0:00	26	63	BVNE
301	30/9/03 11:30	40	138	BVNE
302	15/11/03 14:00	54	200	BVNE
303	29/11/03 18:30	125	348	BVNE
304	20/2/04 19:30	66	66	BVNE
305	30/3/04 5:30	134	154	BVNE
306	29/4/04 0:00	60	217	BVNE
307	27/10/04 12:00	42	70	BVNE
308	5/9/05 21:30	84	74	BVNE
309	28/1/06 5:00	50	146	BVNE
310	13/9/06 10:30	42	145	BVNE
311	3/1/08 11:30	37	41	BVNE
312	3/2/08 12:30	37	118	BVNE
313	18/4/08 0:00	94	113	BVNE
314	13/12/08 0:00	60	89	BVNE

**Evènements sur les bassins non jaugés :**

Cours d'eau	Exutoire	Surface (km <sup>2</sup> )	Date début	Heure début	Débit de pointe (m <sup>3</sup> .s <sup>-1</sup> )	Source
Ardèche	Pont-de-Labeaume	280	05/12/2000	05:00	434	OHMCV - Sevnol
Ardèche	Pont-de-Labeaume	280	18/10/2001	03:00	494	OHMCV - Sevnol
Ardèche	Pont-de-Labeaume	280	29/11/2003	20:00	484	OHMCV - Sevnol
Baume	Rosières	210	12/11/2000	00:00	212	OHMCV - Sevnol
Baume	Rosières	210	17/10/2001	16:00	330	OHMCV - Sevnol
Baume	Rosières	210	25/10/2004	08:00	230	OHMCV - Sevnol

**Evènements sur les bassins non jaugés (suite) :**

Cours d'eau	Exutoire	Surface (km <sup>2</sup> )	Date début	Heure début	Débit de pointe (m <sup>3</sup> .s <sup>-1</sup> )	Source
Chassezac	Gravières	500	28/09/2000	05:00	564	OHMCV - Sevnol
Chassezac	Gravières	500	17/10/2001	15:00	635	OHMCV - Sevnol
Chassezac	Gravières	500	21/11/2003	09:00	728	OHMCV - Sevnol
Ganière	Banne	55	08/11/2000	00:00	88	OHMCV - Sevnol
Ganière	Banne	55	29/11/2000	00:00	63	OHMCV - Sevnol
Ganière	Banne	55	05/12/2000	13:00	77	OHMCV - Sevnol
Ganière	Banne	55	16/10/2001	00:00	107	OHMCV - Sevnol
Ganière	Banne	55	07/09/2002	12:00	206	OHMCV - Sevnol
Ganière	Banne	55	29/09/2003	00:00	55	OHMCV - Sevnol
Ganière	Banne	55	13/11/2003	01:00	64	OHMCV - Sevnol
Ganière	Banne	55	19/11/2003	18:00	81	OHMCV - Sevnol
Ganière	Banne	55	28/11/2003	02:00	93	OHMCV - Sevnol
Ganière	Banne	55	23/10/2004	01:00	88	OHMCV - Sevnol
Ganière	Banne	55	02/11/2004	06:00	78	OHMCV - Sevnol
Auzonnet	Les Mages	49	20/12/2000	01:00	58	OHMCV - Sevnol
Auzonnet	Les Mages	49	17/10/2001	01:00	68	OHMCV - Sevnol
Auzonnet	Les Mages	49	07/09/2002	13:00	169	OHMCV - Sevnol
Galeizon	Cendras	88	07/09/2002	13:00	970	Syndicat du Galeizon
Galeizon	Pt.-De-Roubarbel	39	07/09/2002	13:00	329	Syndicat du Galeizon
Salandre	Confl. Galeizon	21	07/09/2002	13:00	310	Syndicat du Galeizon
Avène	St-Hilaire-de-Brethmas	57	07/09/2002	13:00	400	Retour d'exp. 09/2002
Ruisseau du Lyonnais	St-Jean-du-Pin	4	07/09/2002	13:00	120	Retour d'exp. 09/2002
Amous	Généralgues	21	07/09/2002	13:00	435	Retour d'exp. 09/2002
Alzon	St-Christol-lez-Alès	15	07/09/2002	13:00	430	Retour d'exp. 09/2002
Gardon de St Jean	Saumane	100	08/09/2002	13:00	527	BVNE Gardons
Gardon de St Jean	Saumane	100	28/09/2000	09:00	212	BVNE Gardons
Gardon de St Jean	Saumane	100	10/11/1996	10:00	111	BVNE Gardons
Gardon d'Anduze	Anduze	545	08/09/2002	13:00	2742	BVNE Gardons
Gardon d'Anduze	Anduze	545	28/09/2000	09:00	801	BVNE Gardons

#### **ANNEXE 4 : Prévisions sur le bassin du Gardon à Mialet**

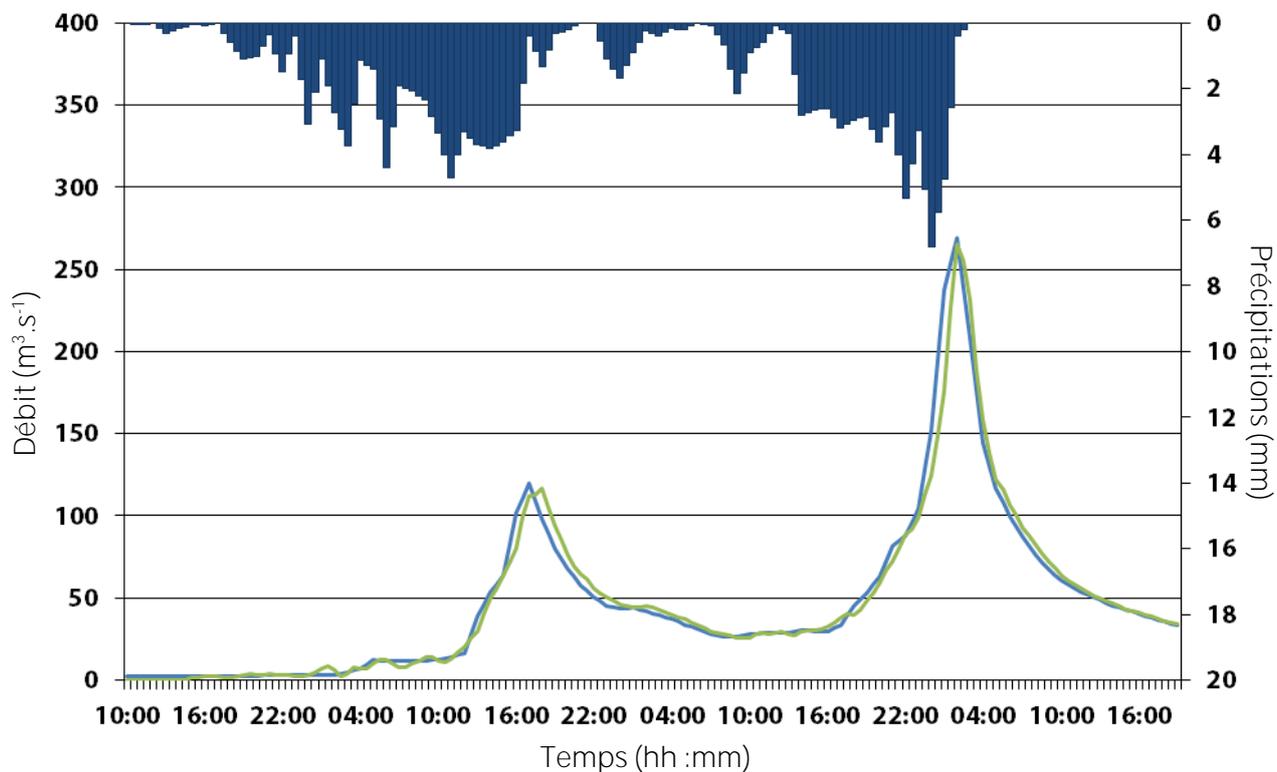
L'annexe représente les hydrogrammes en prévision avec le modèle statique (dirigé).

- Le débit observé est en bleu et les prévisions en vert.
- Les ordonnées normales indiquent le débit en mètres cube par seconde.
- Les ordonnées inverses indiquent le cumul de pluie en millimètres.
- Le temps est gradué en demi-heures.

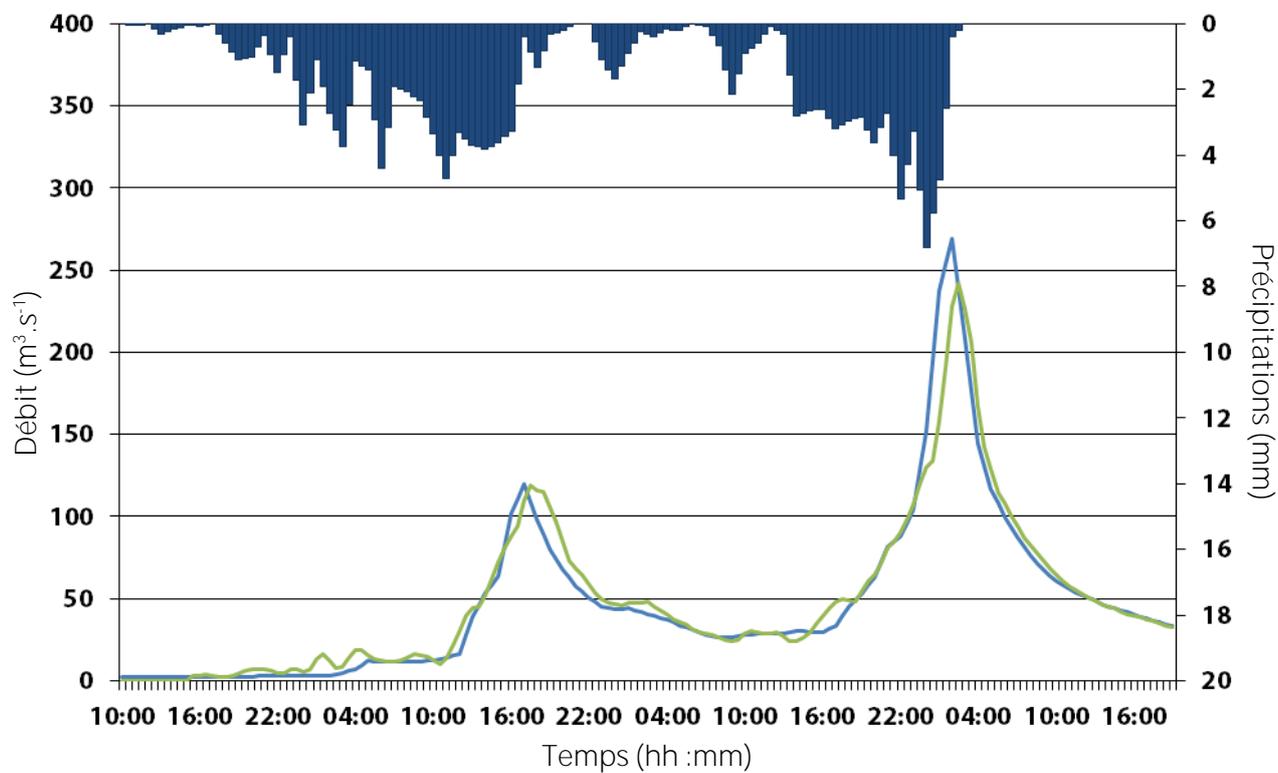
Il n'y a aucune prévision des pluies futures utilisée.

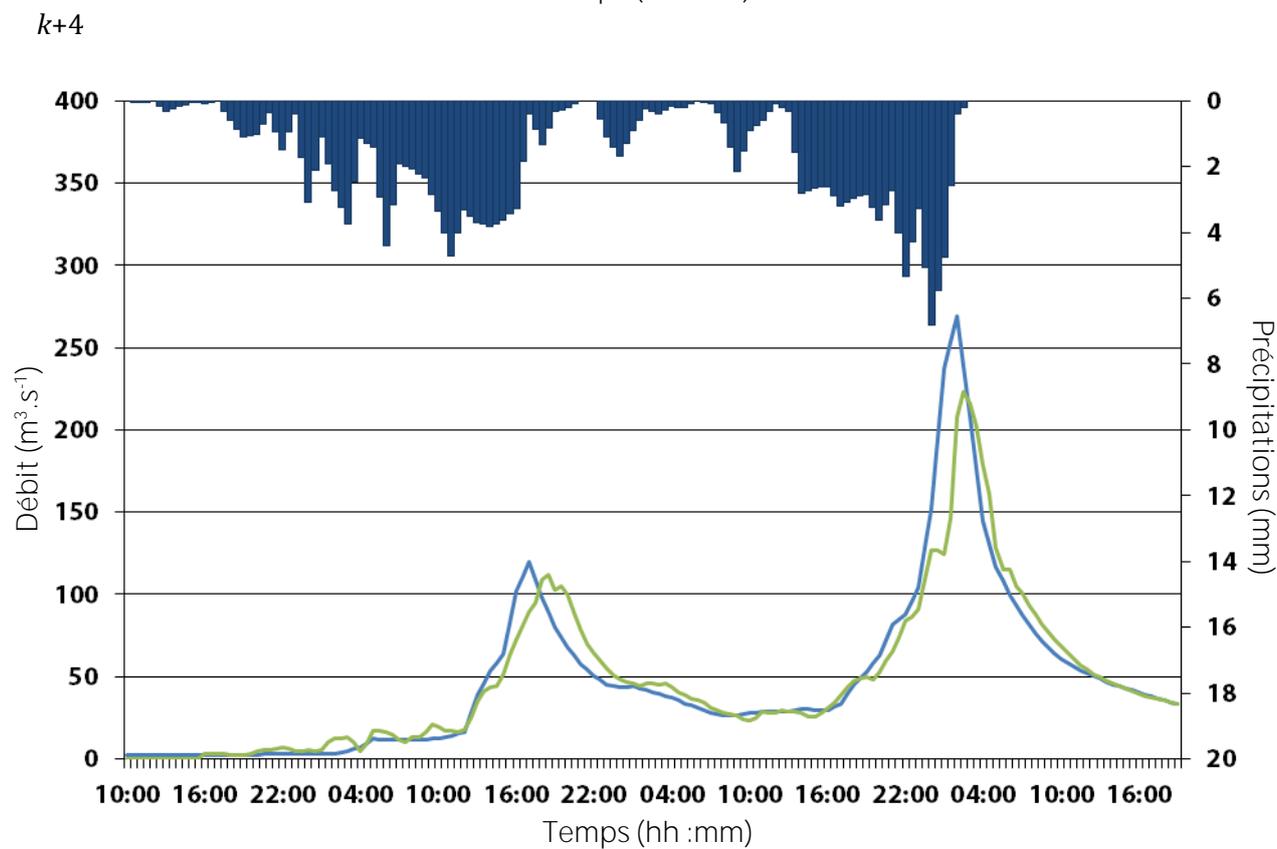
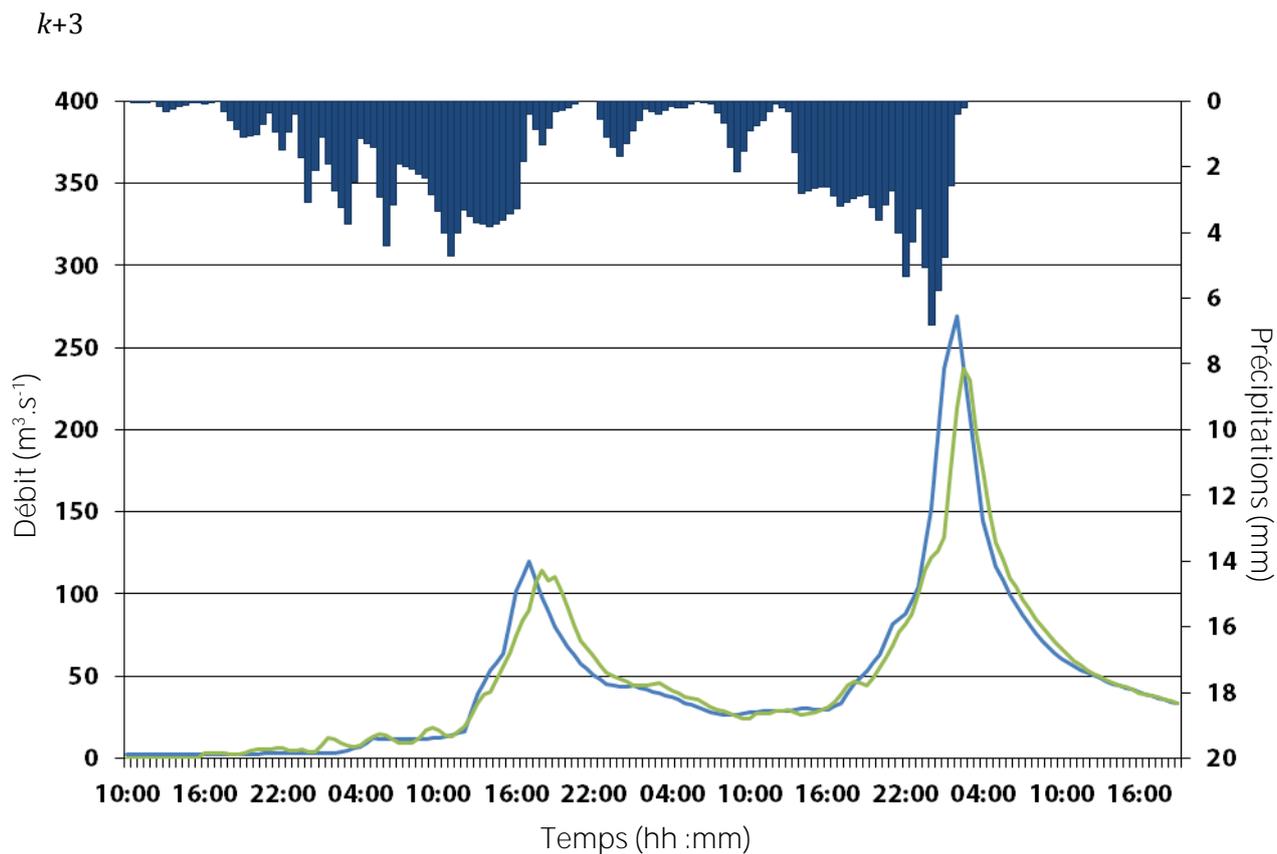
Evènement n°8 :

k+1



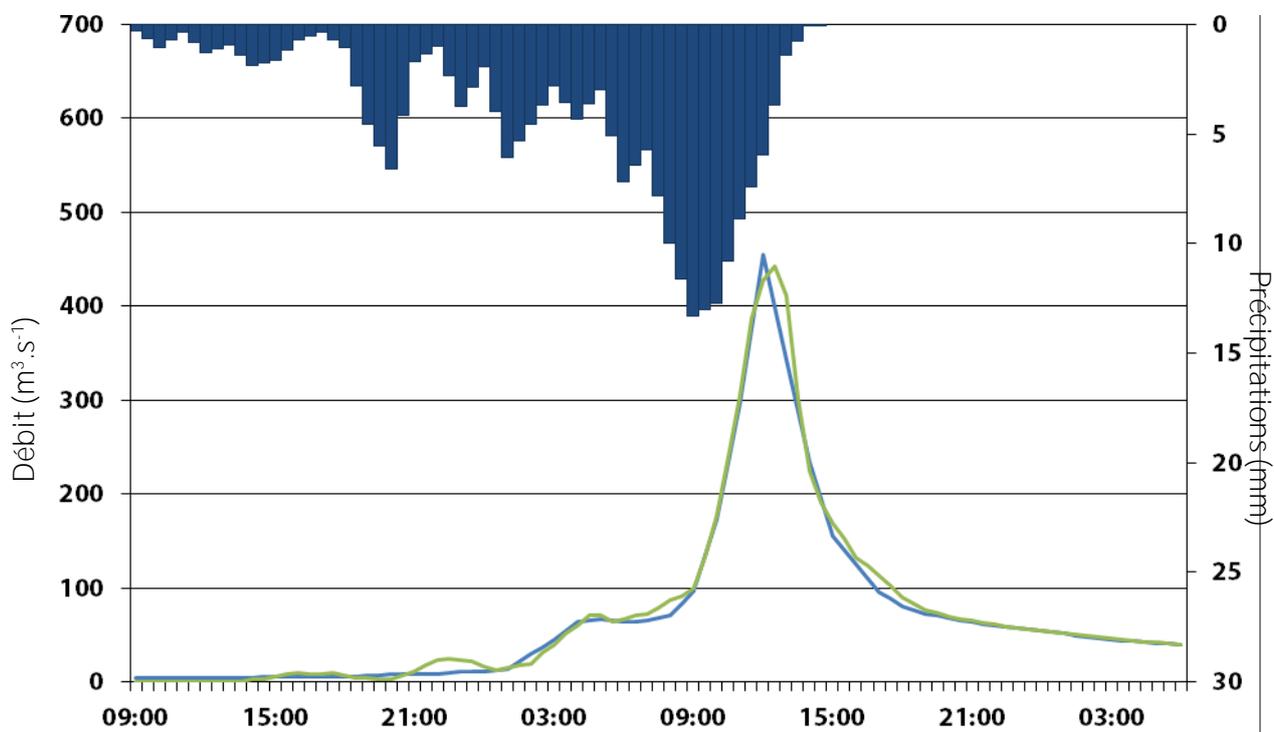
k+2



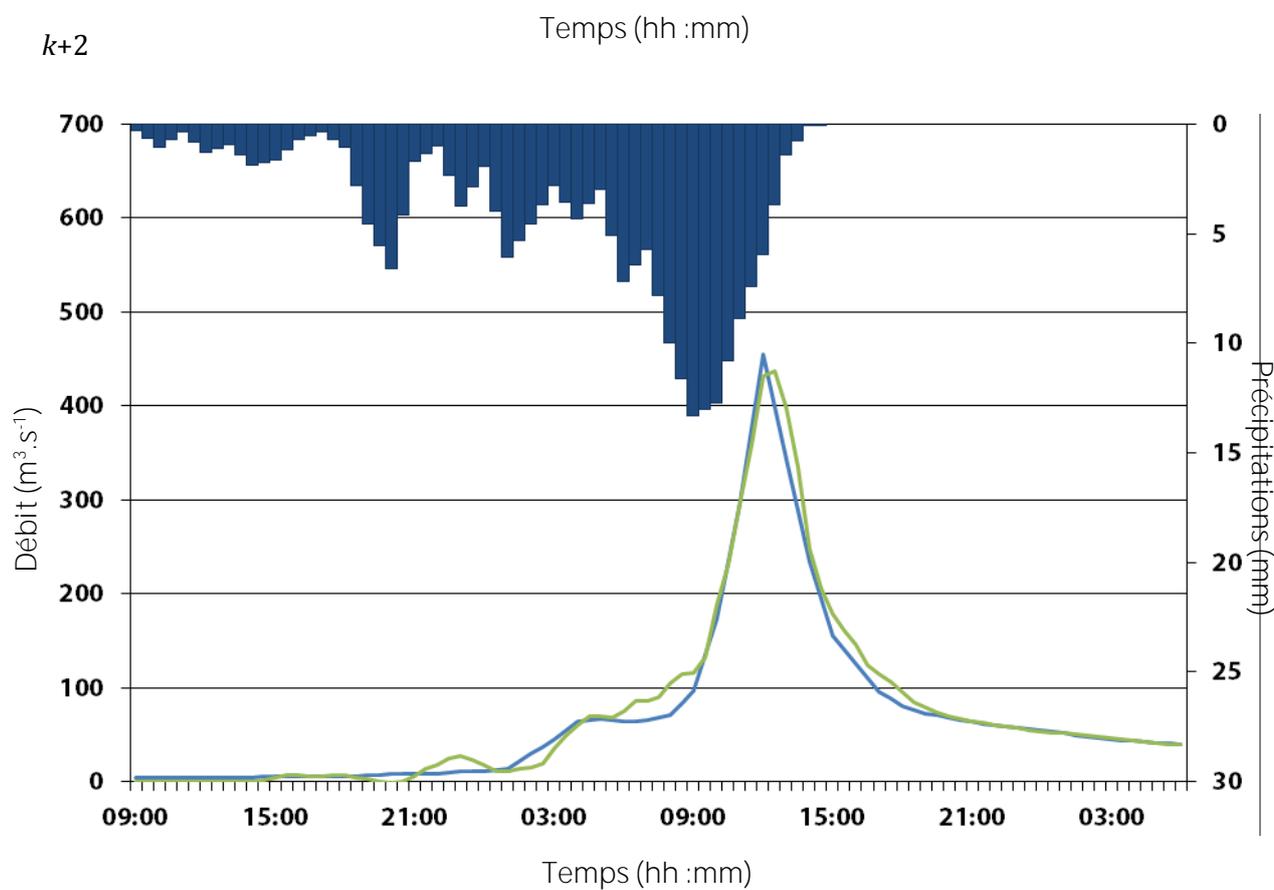


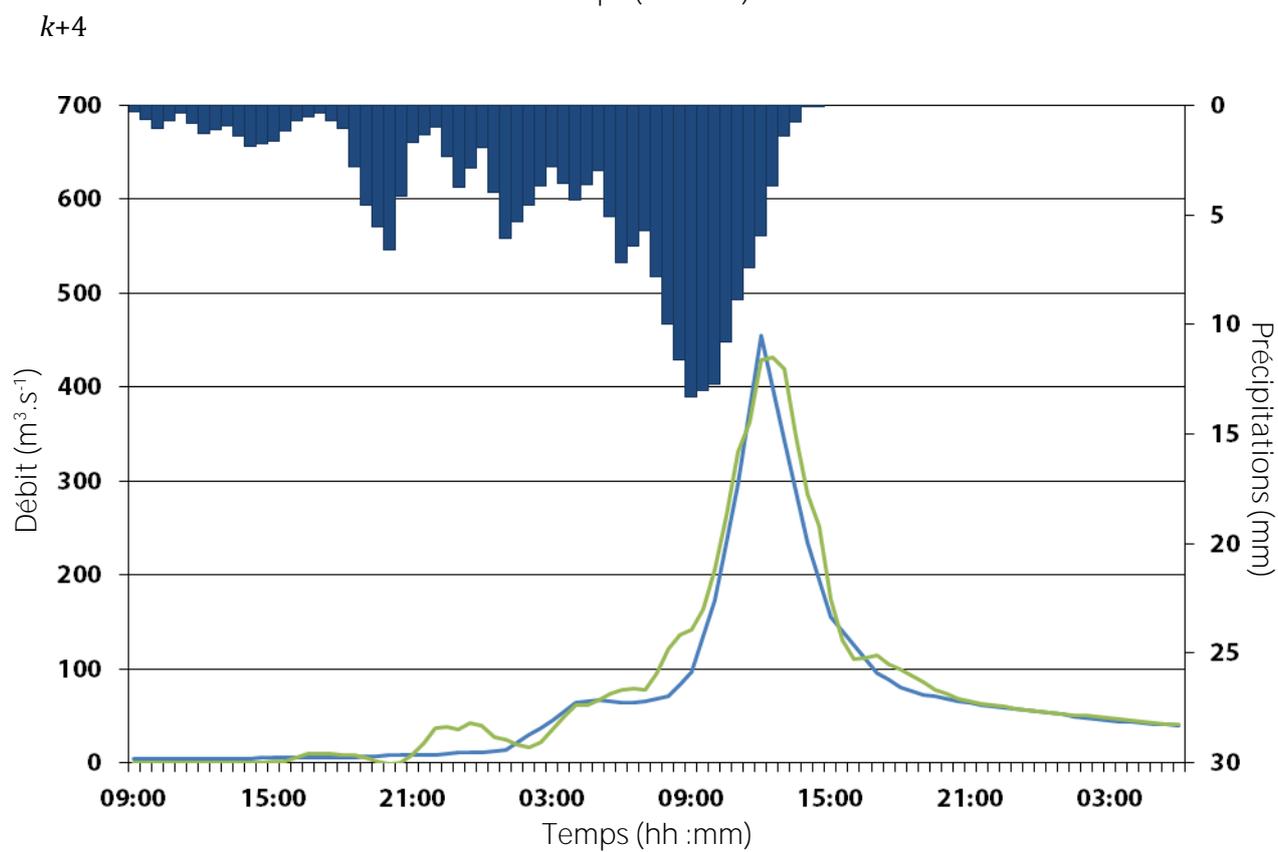
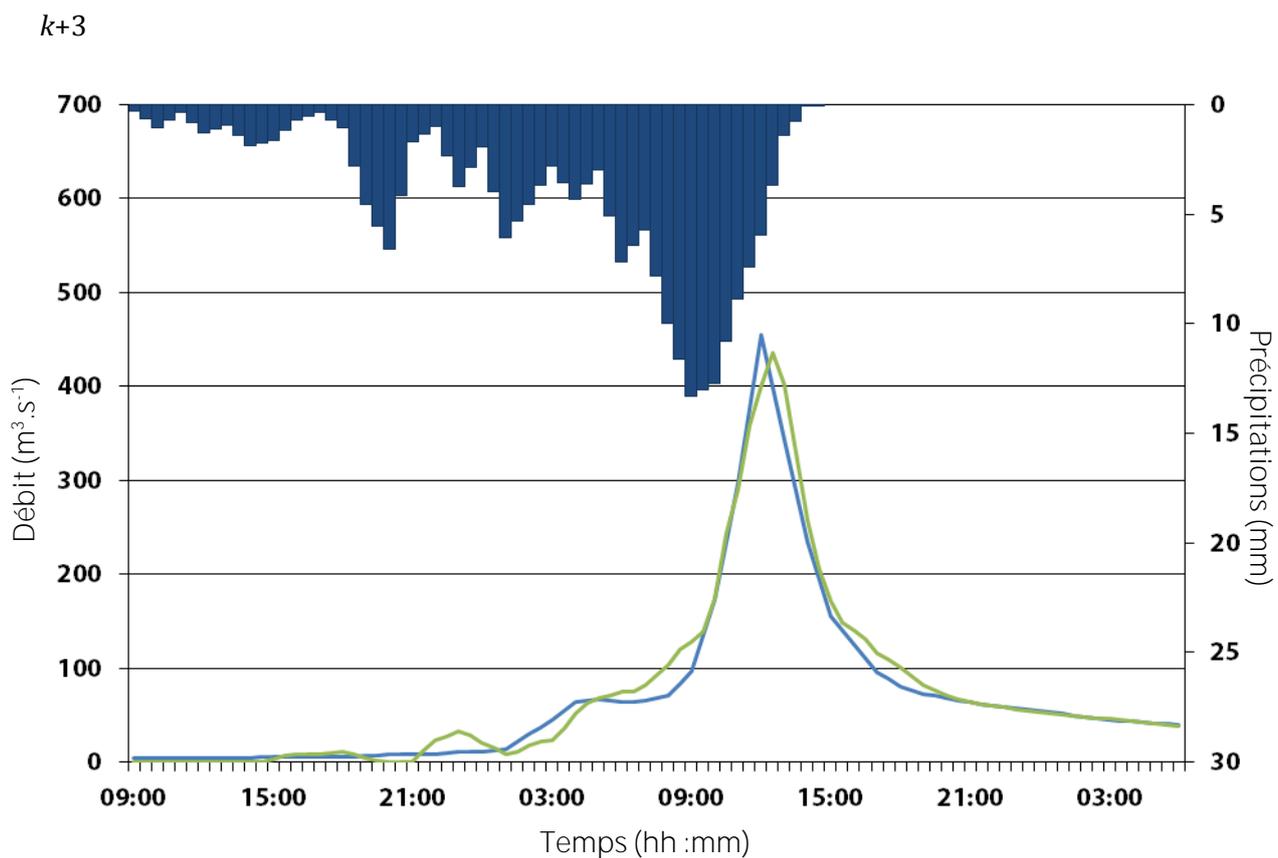
Evènement n°13 :

k+1



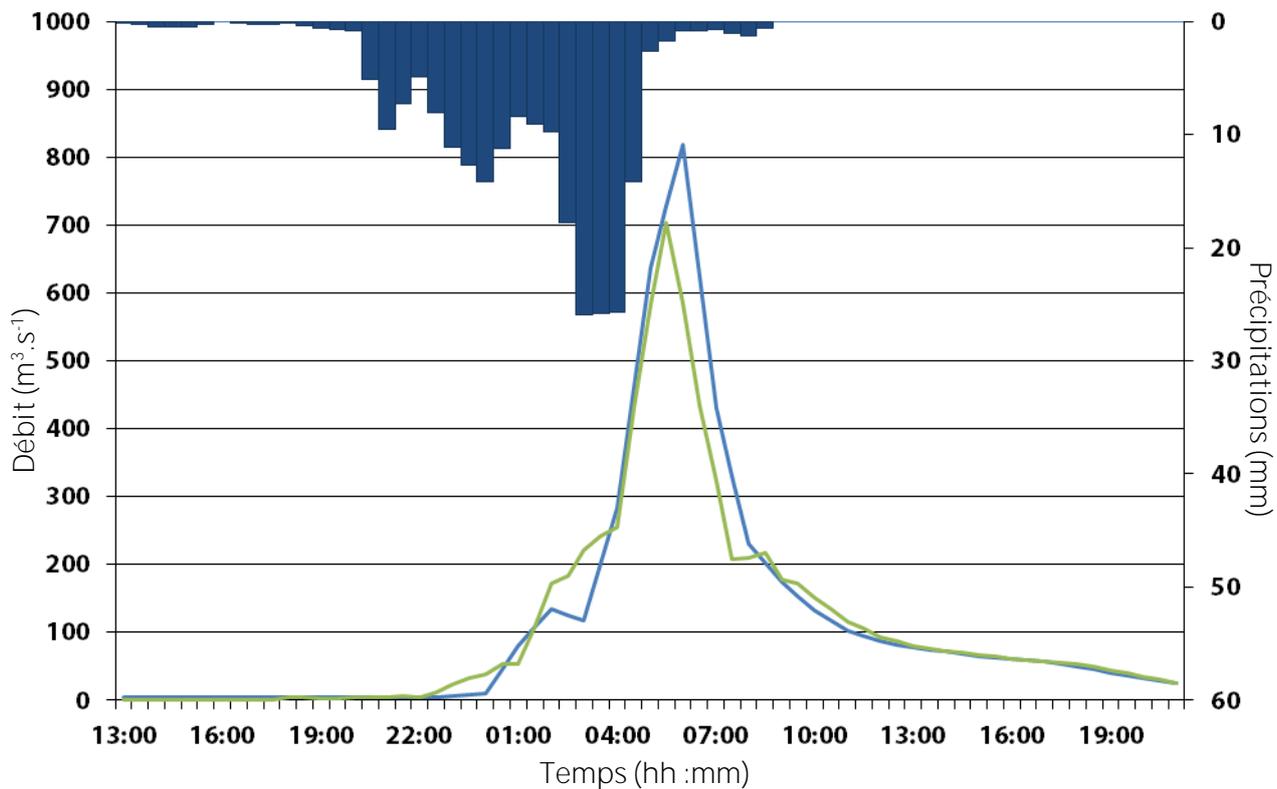
k+2



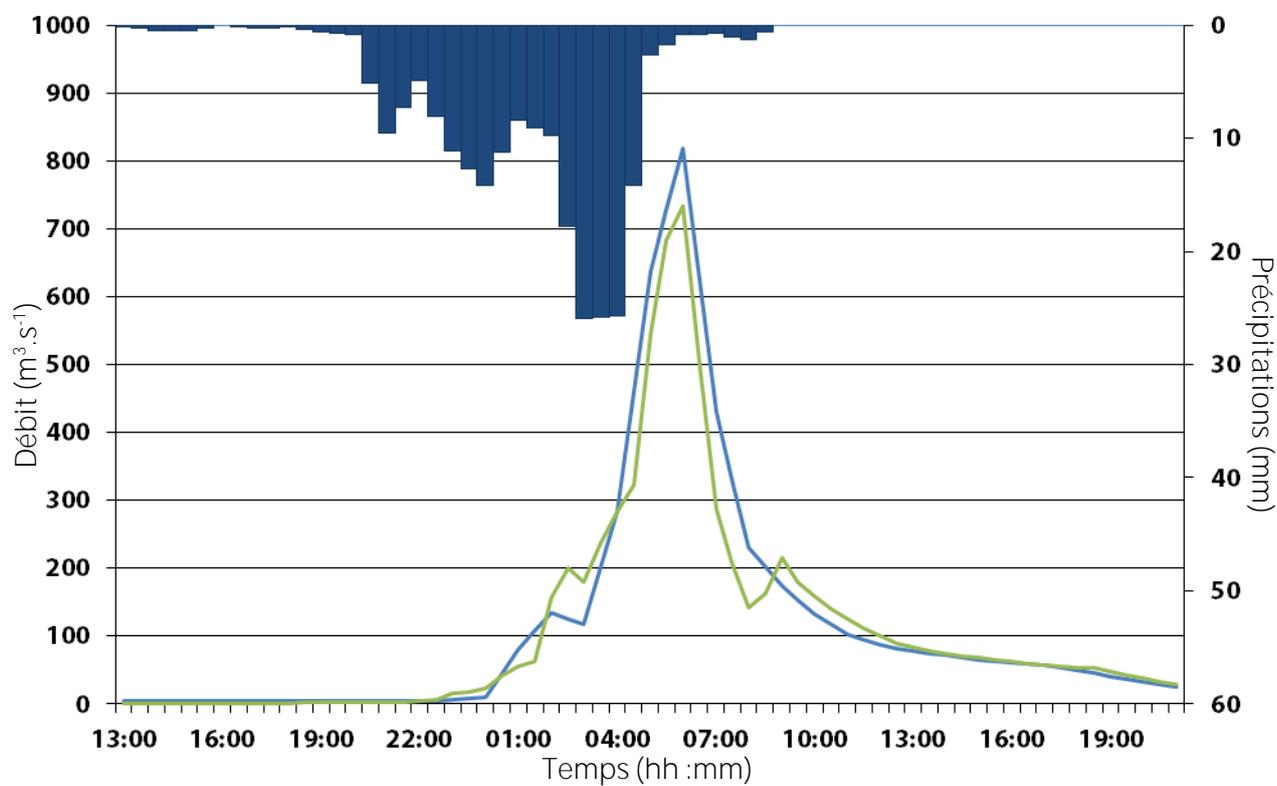


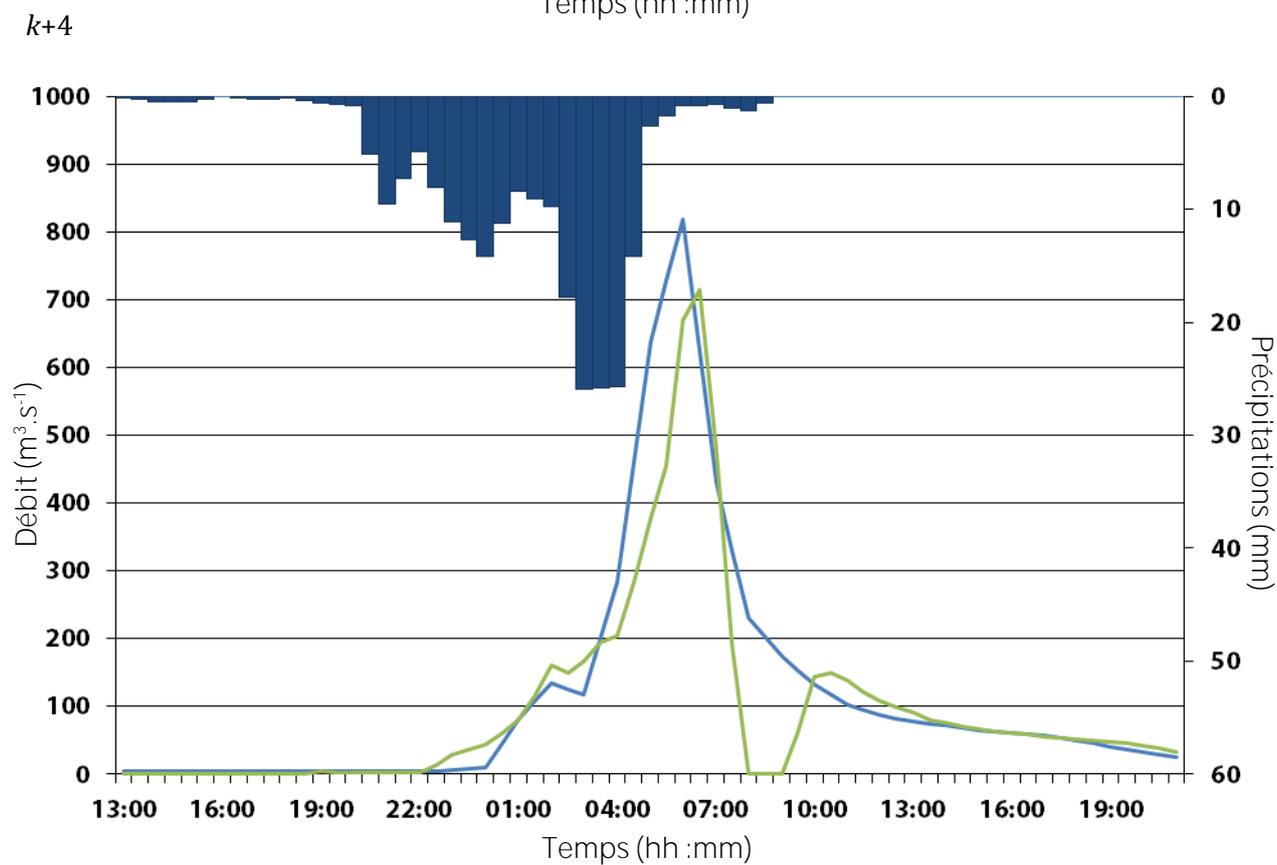
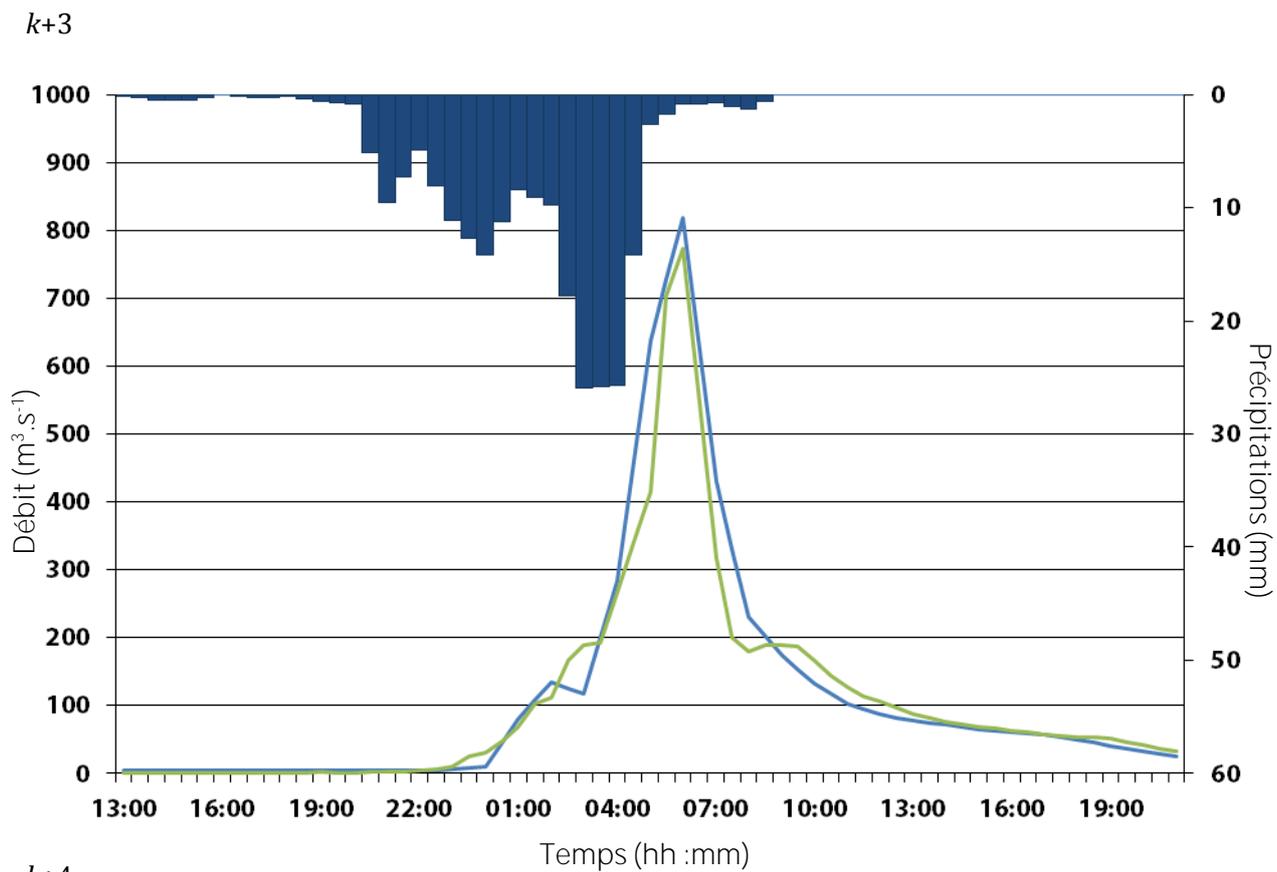
Evènement n°19 :

k+1



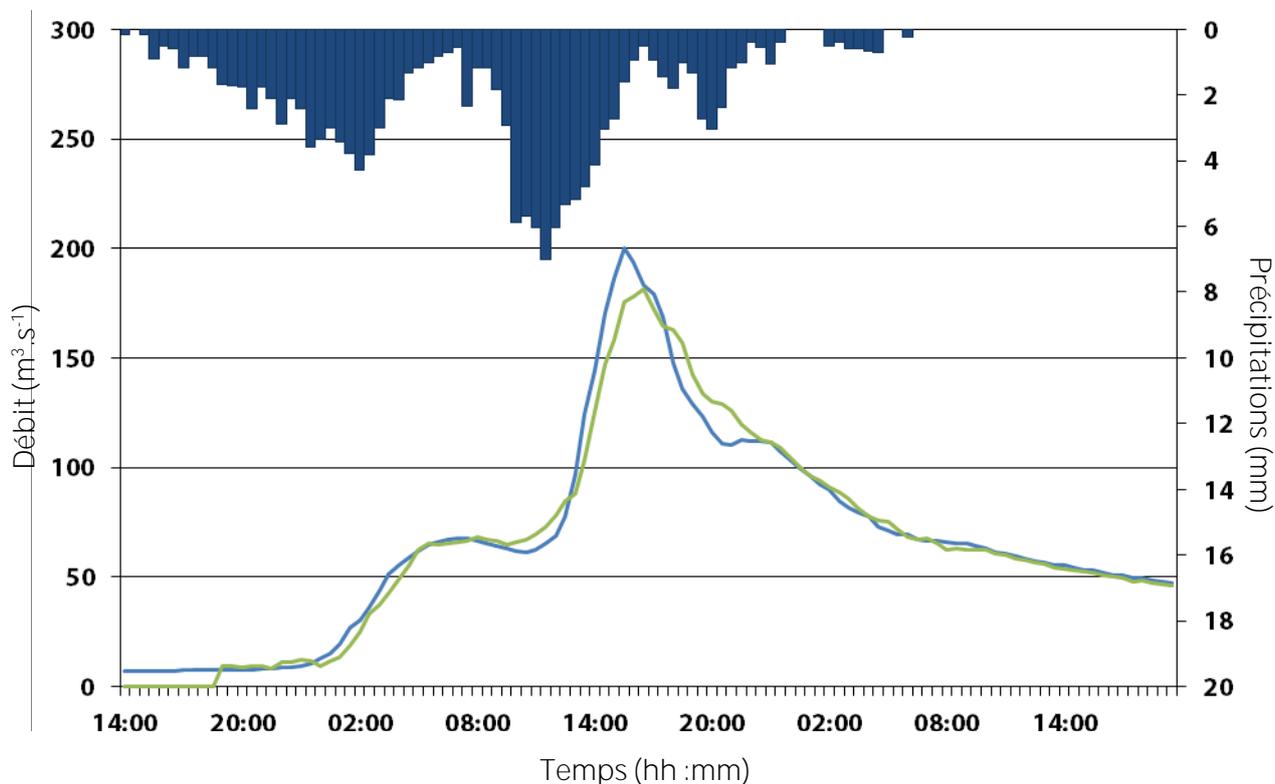
k+2



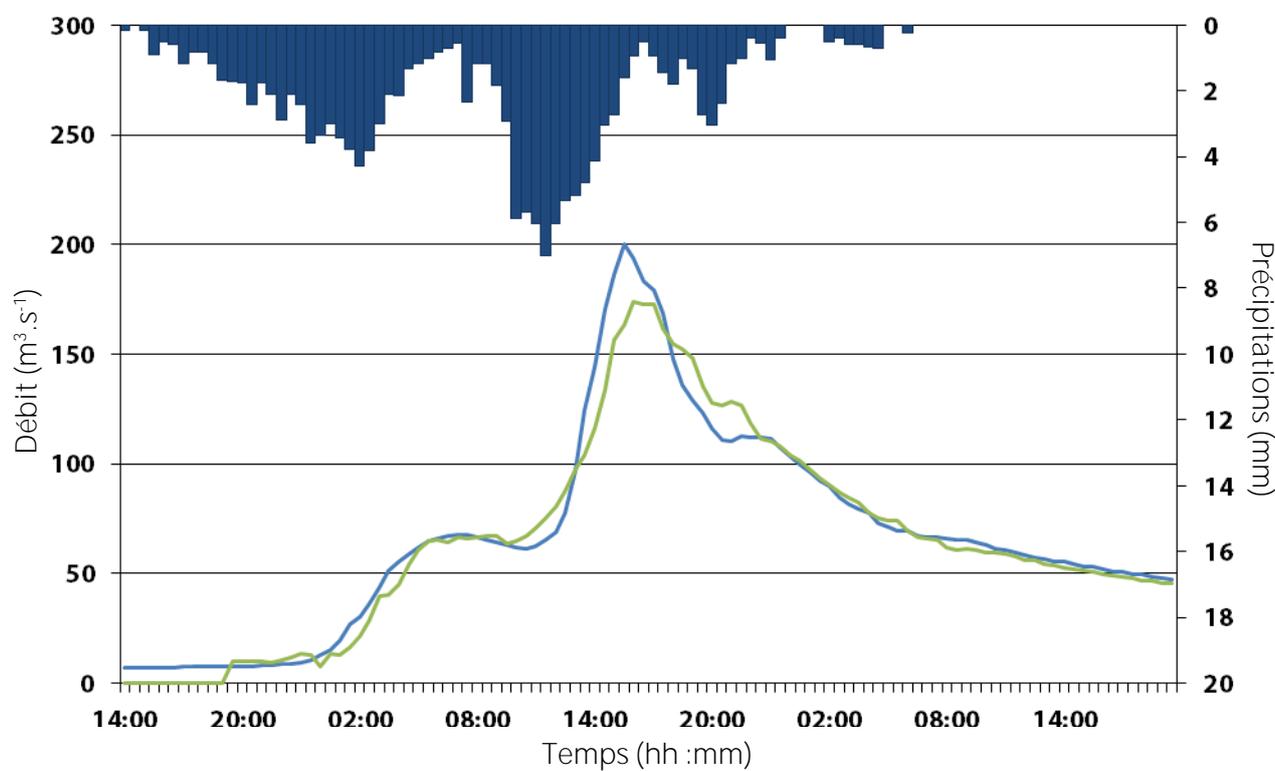


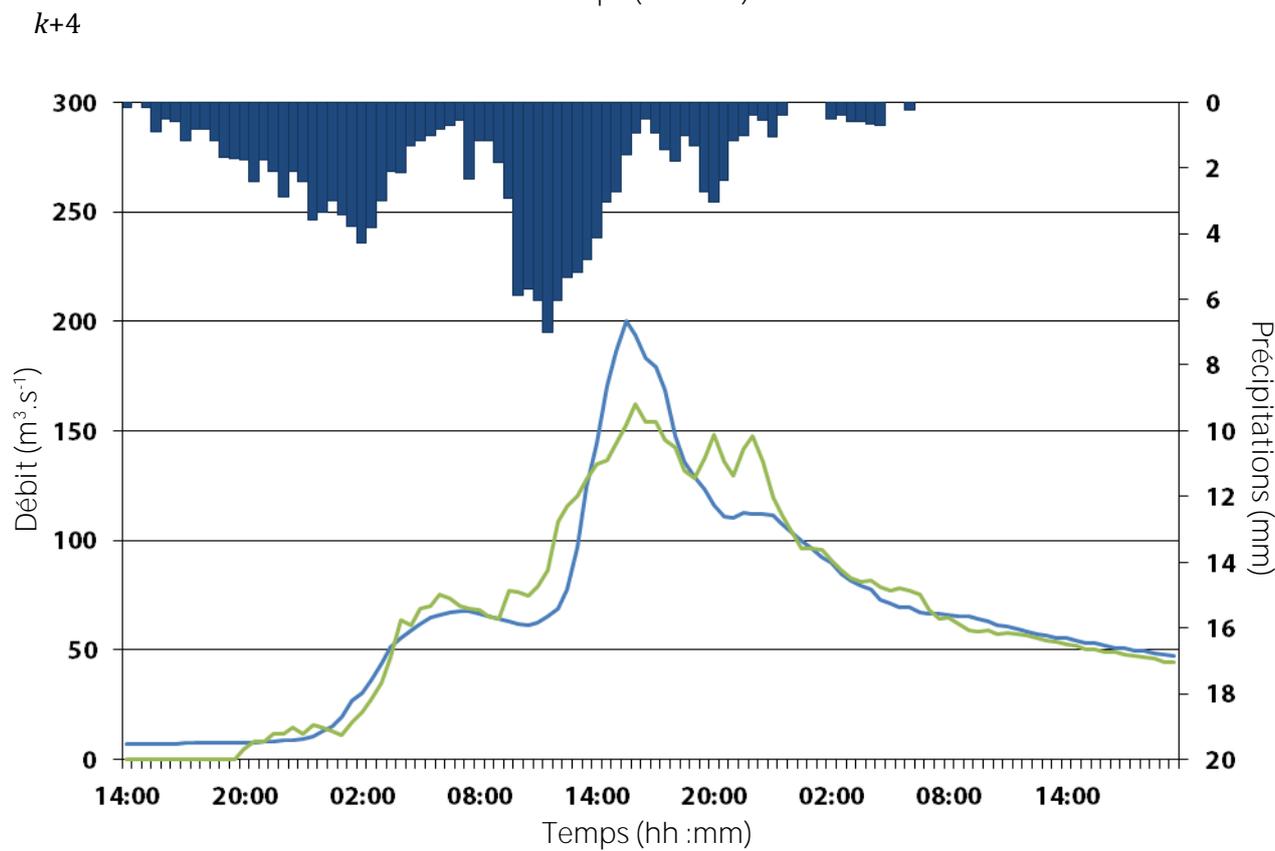
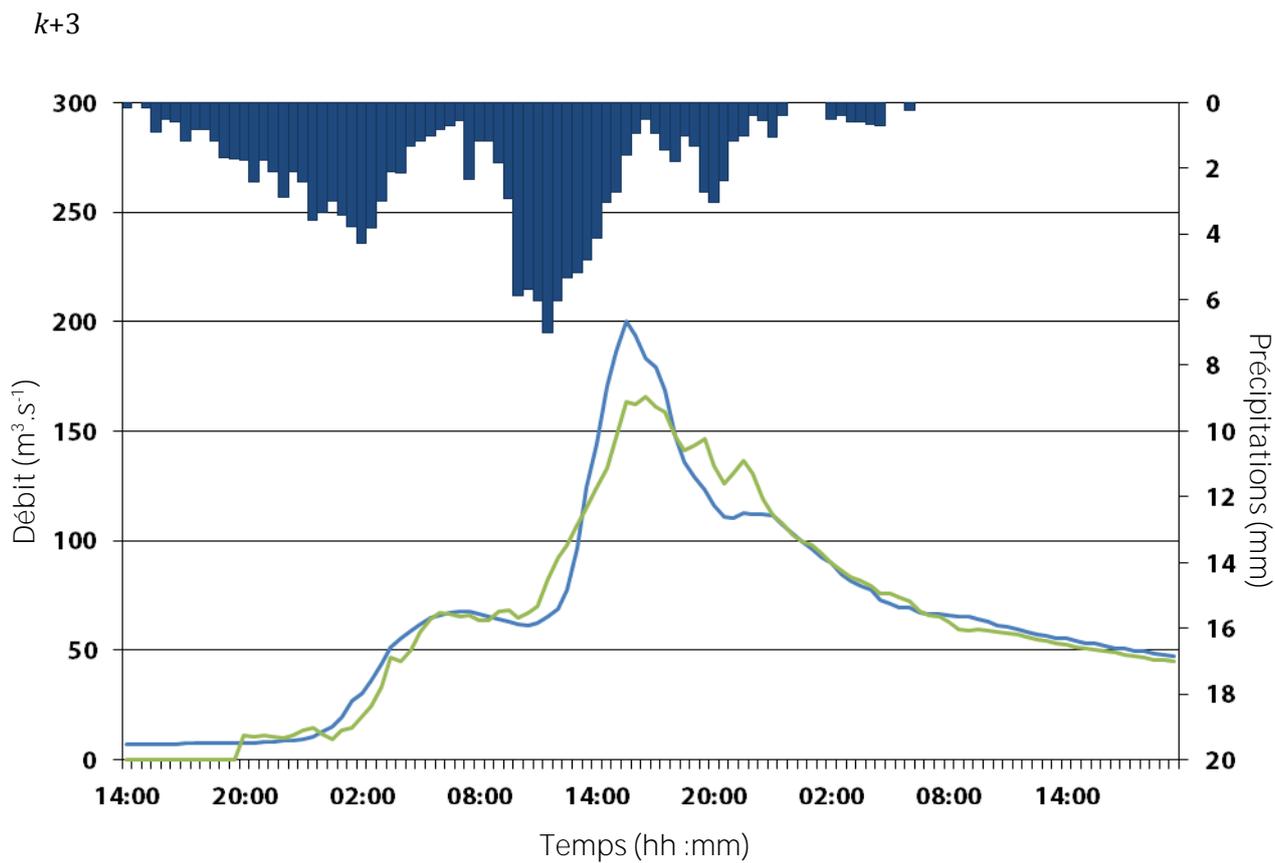
Evènement n°302 :

k+1



k+2





## **ANNEXE 5 : Prévisions sur le bassin du Gardon à Mialet**

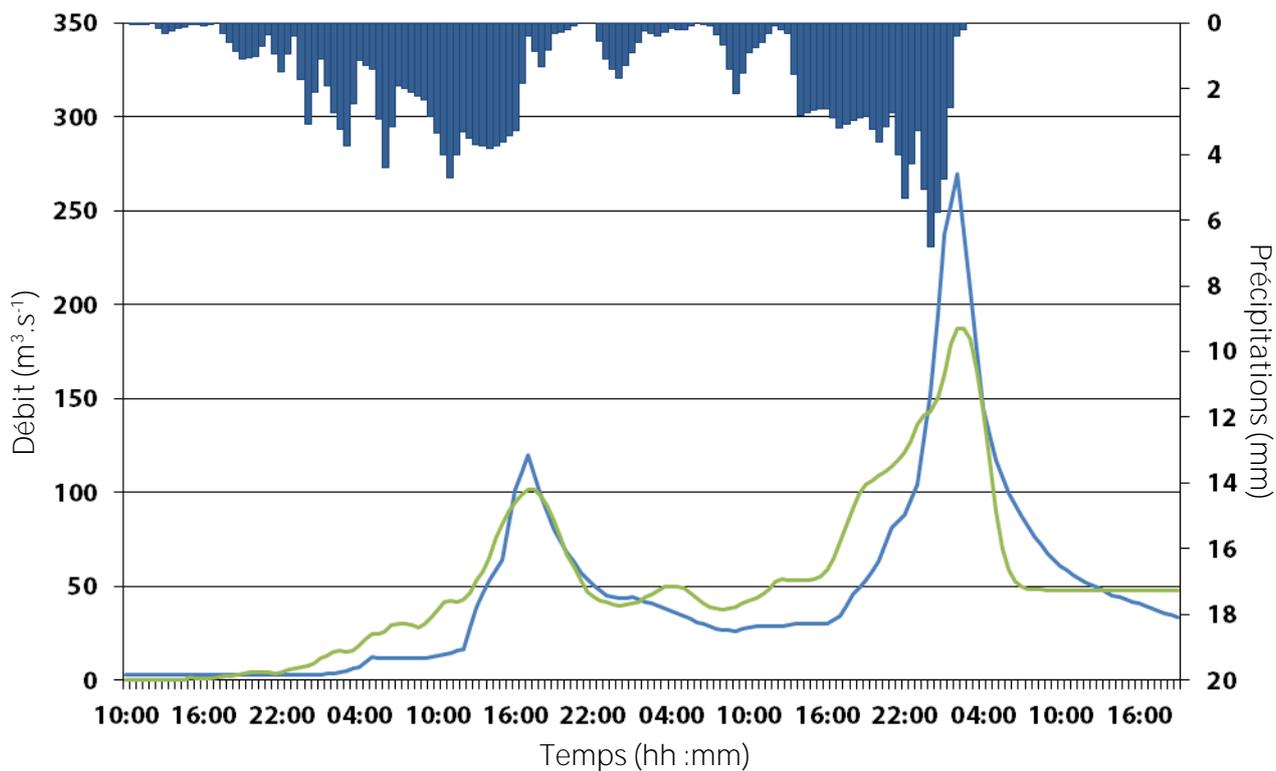
L'annexe présente les hydrogrammes en prévision avec le modèle récurrent déterministe.

- Le débit observé est en bleu et les prévisions en vert.
- Les ordonnées normales indiquent le débit en mètres cube par seconde.
- Les ordonnées inverses indiquent le cumul de pluie en millimètres.
- Le temps est gradué en demi-heures.

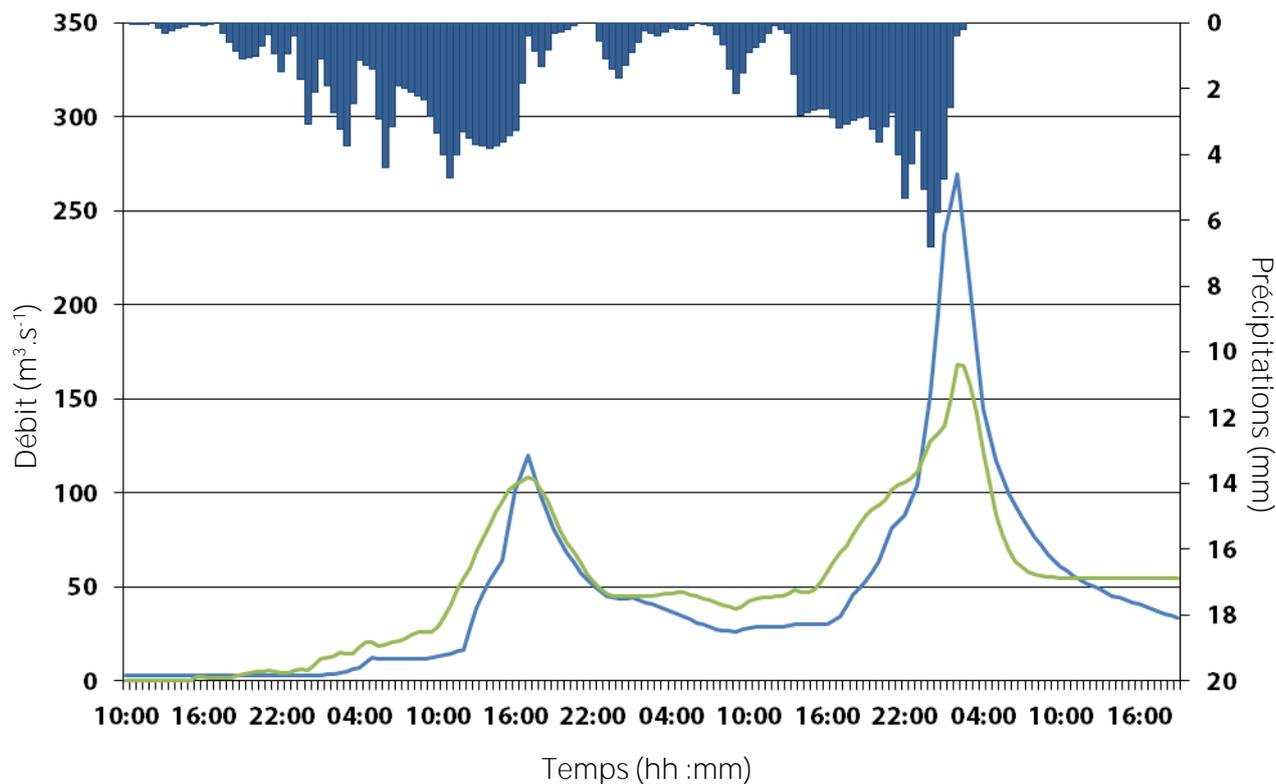
Il n'y a aucune prévision des pluies futures utilisée.

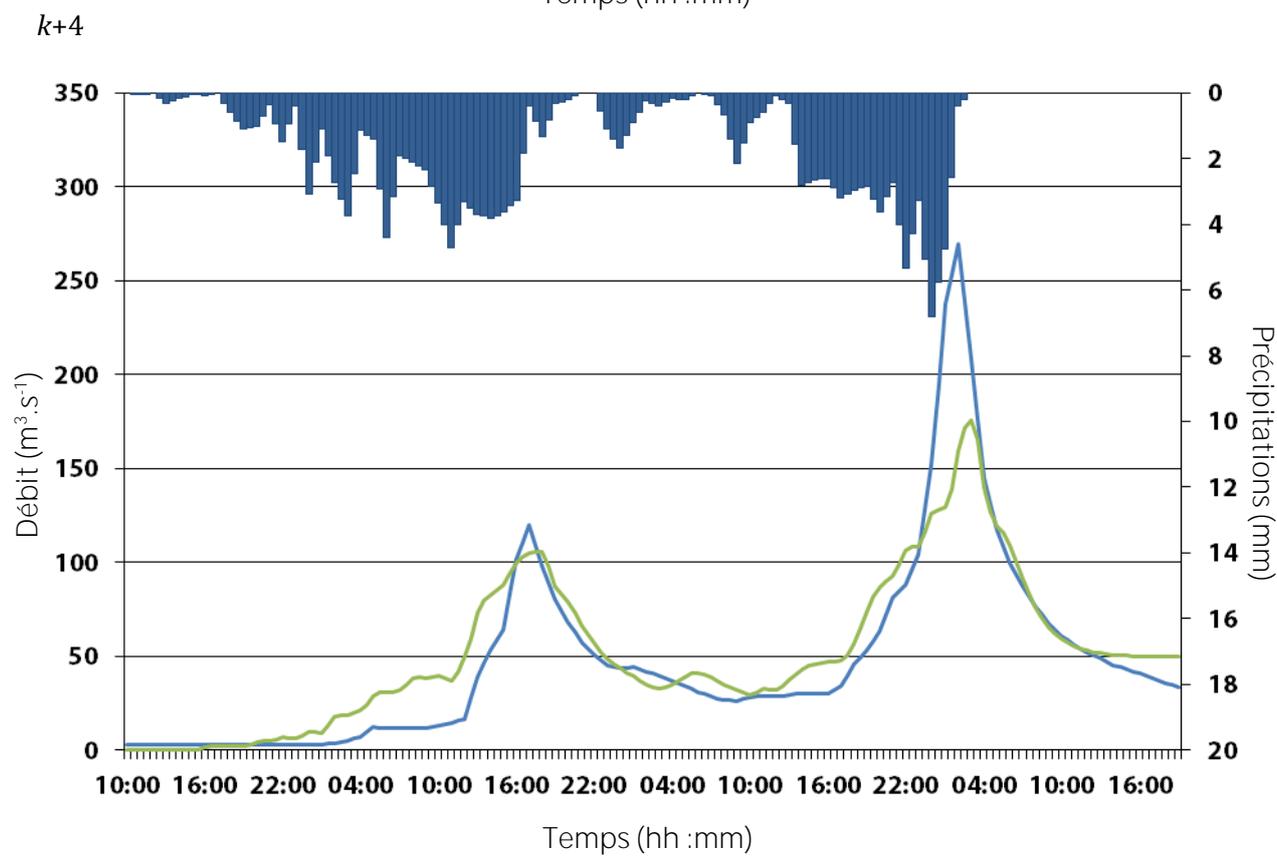
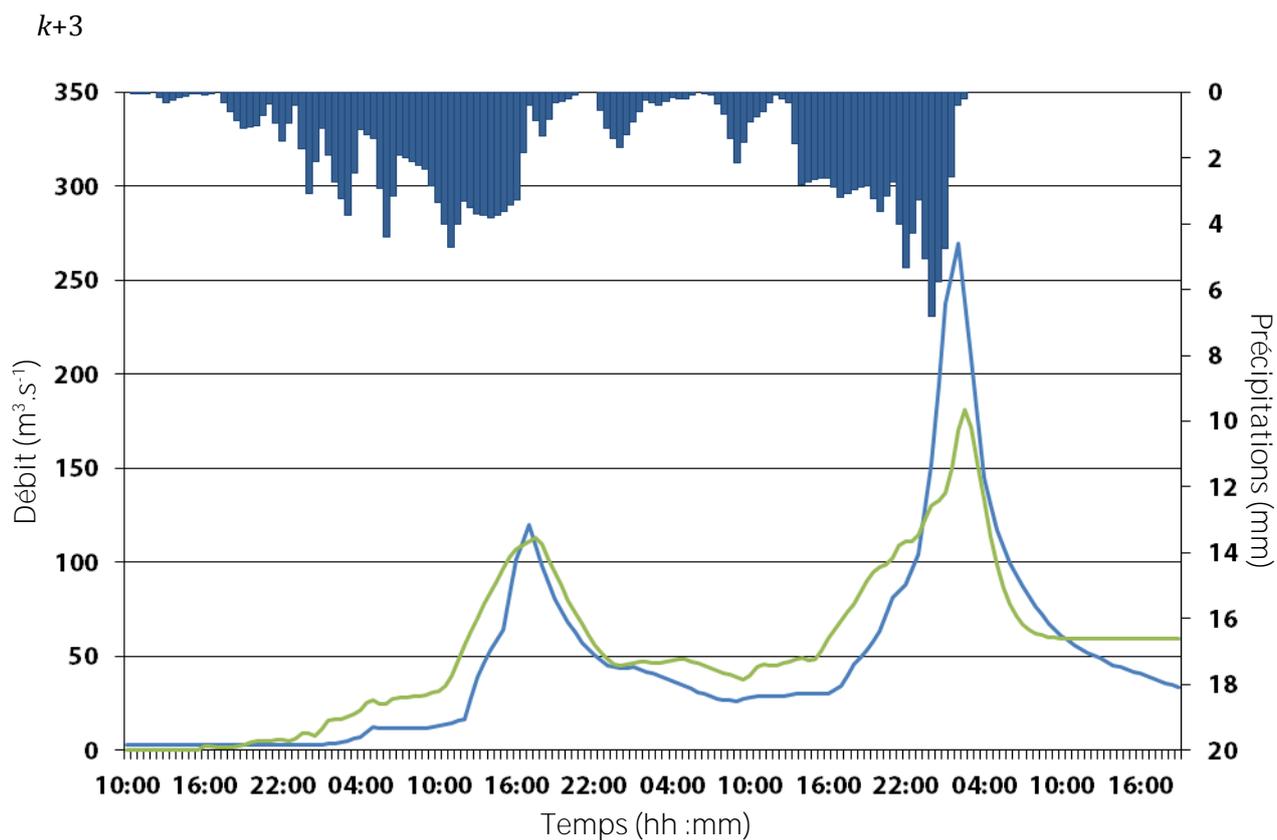
Evènement n°8 :

$k+1$



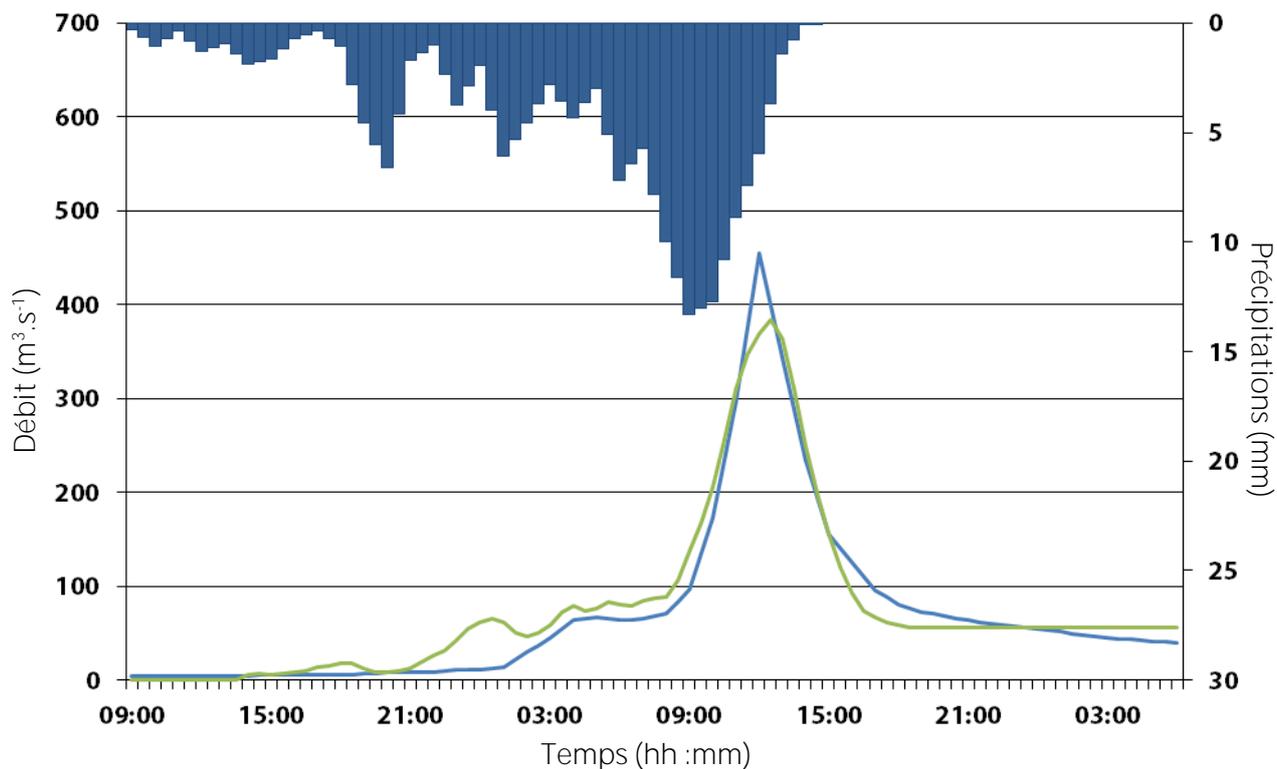
$k+2$



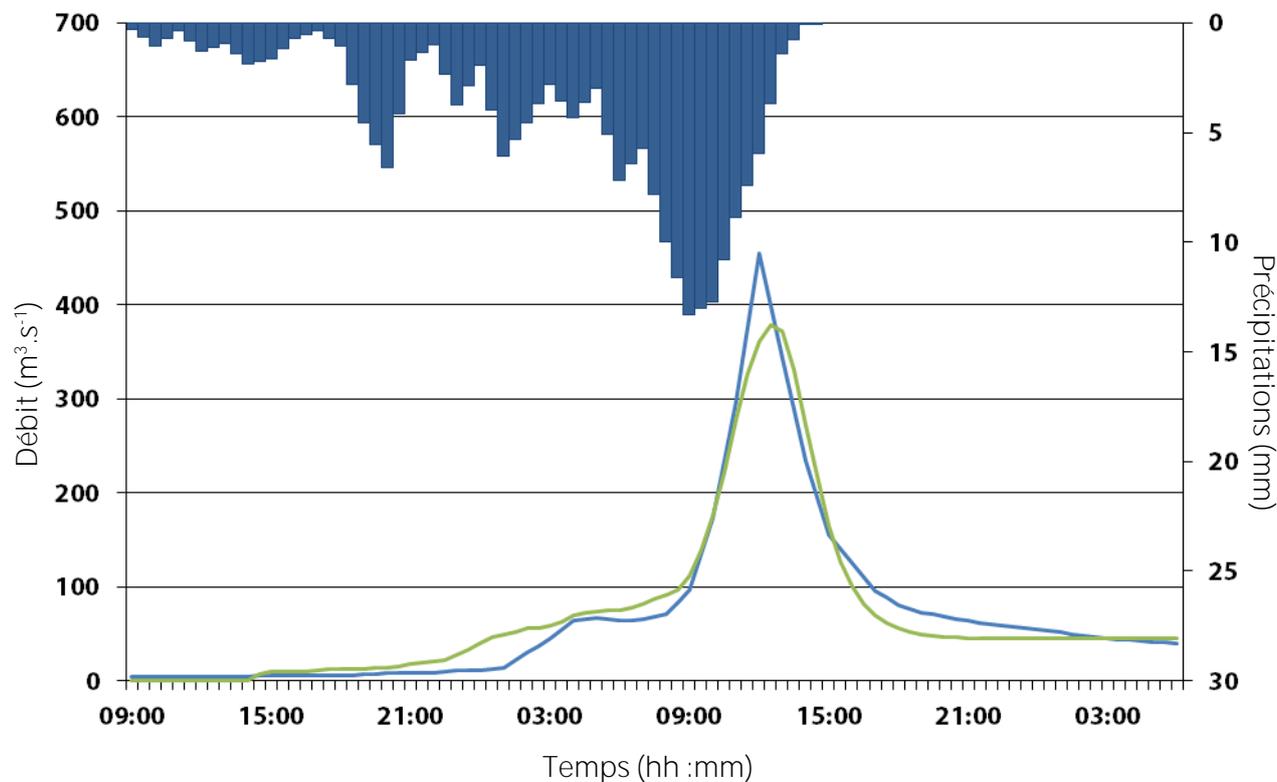


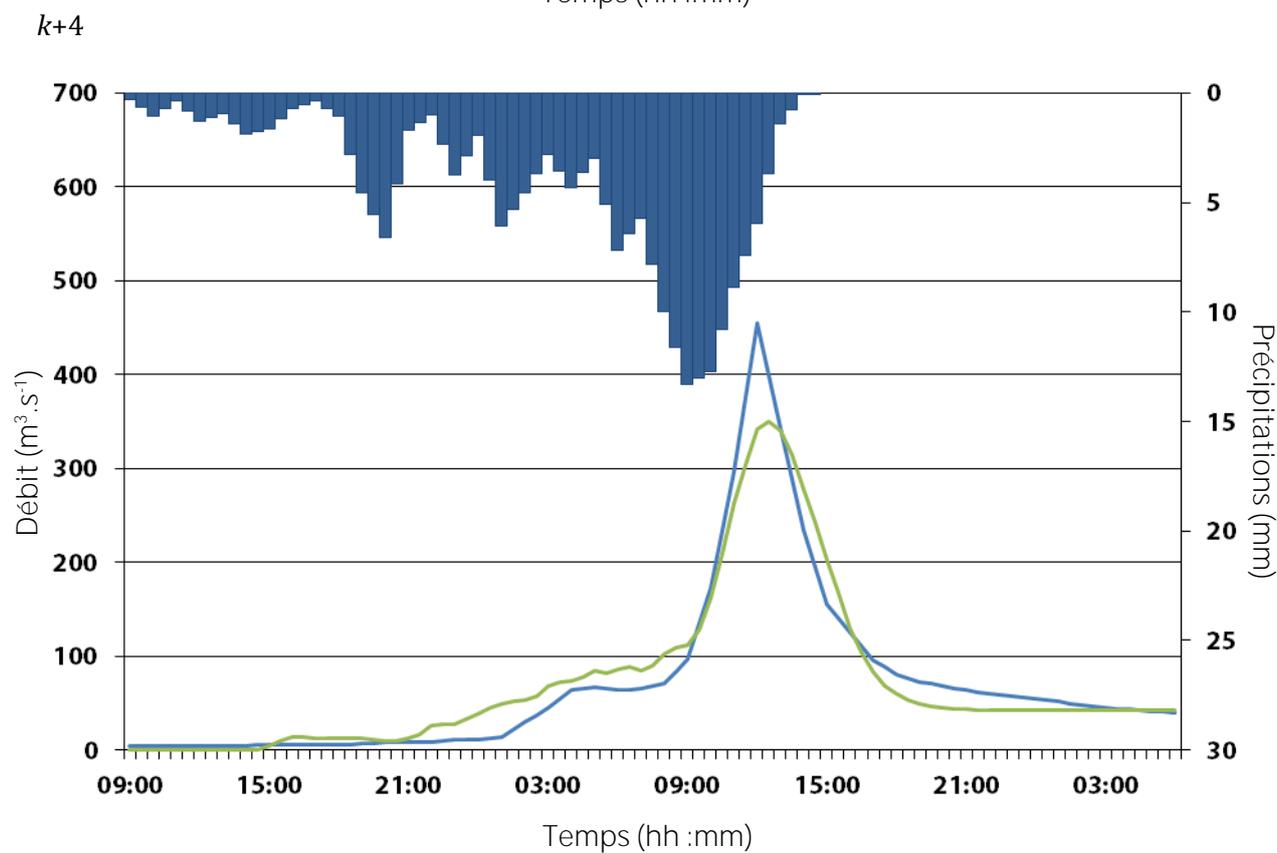
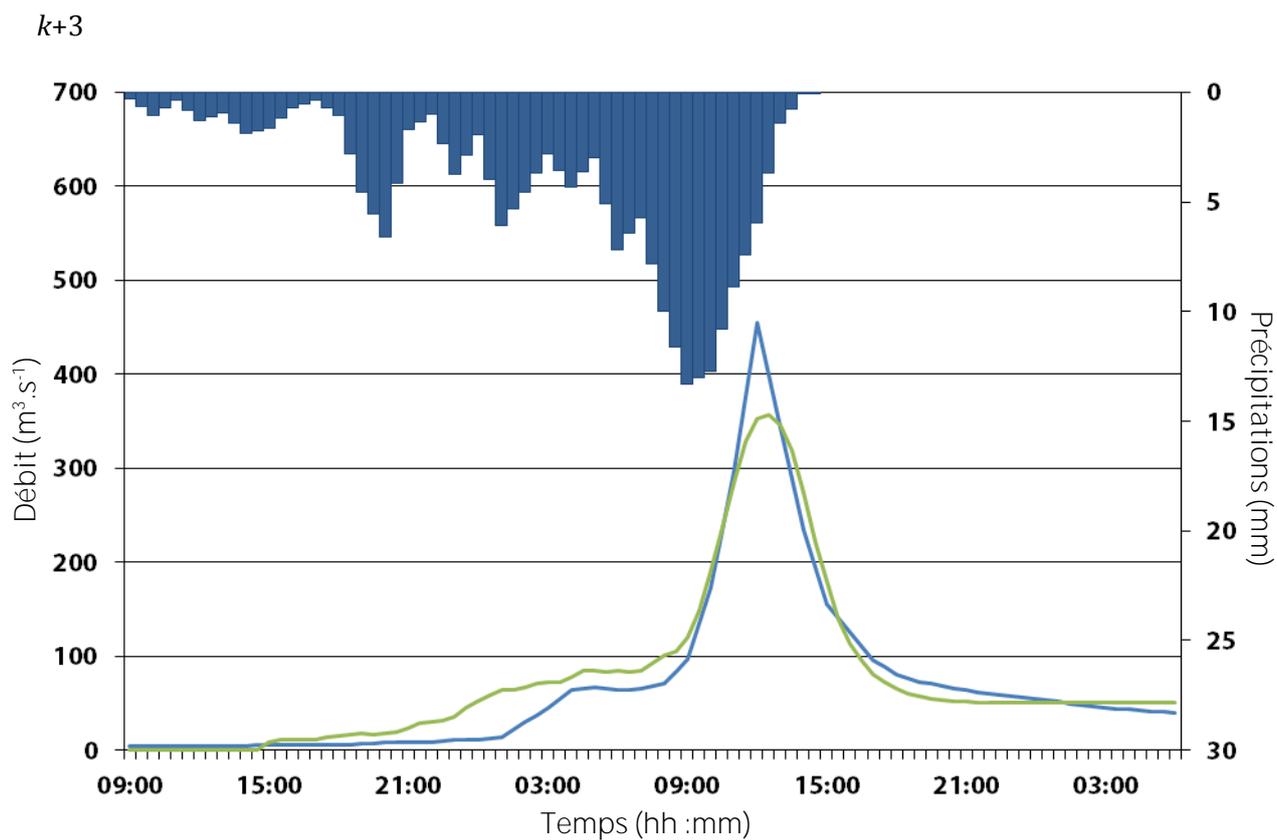
Evènement n°13 :

k+1



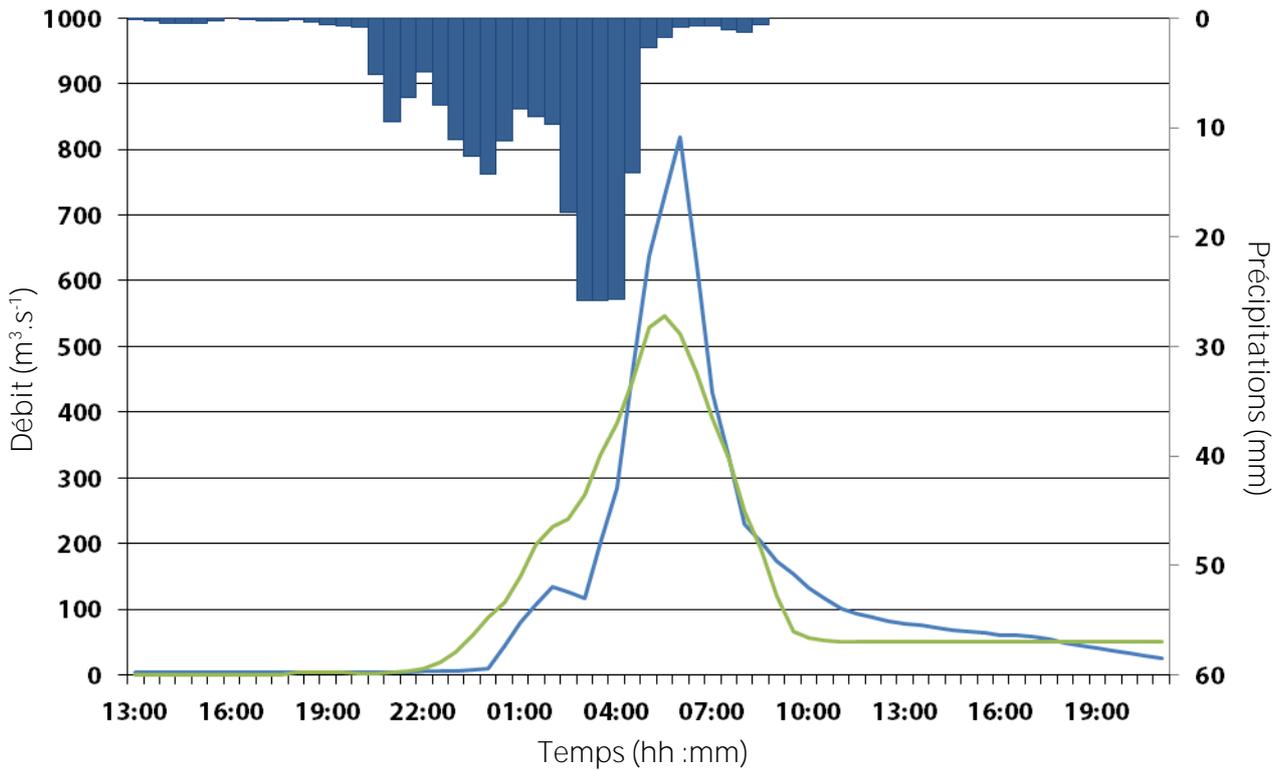
k+2



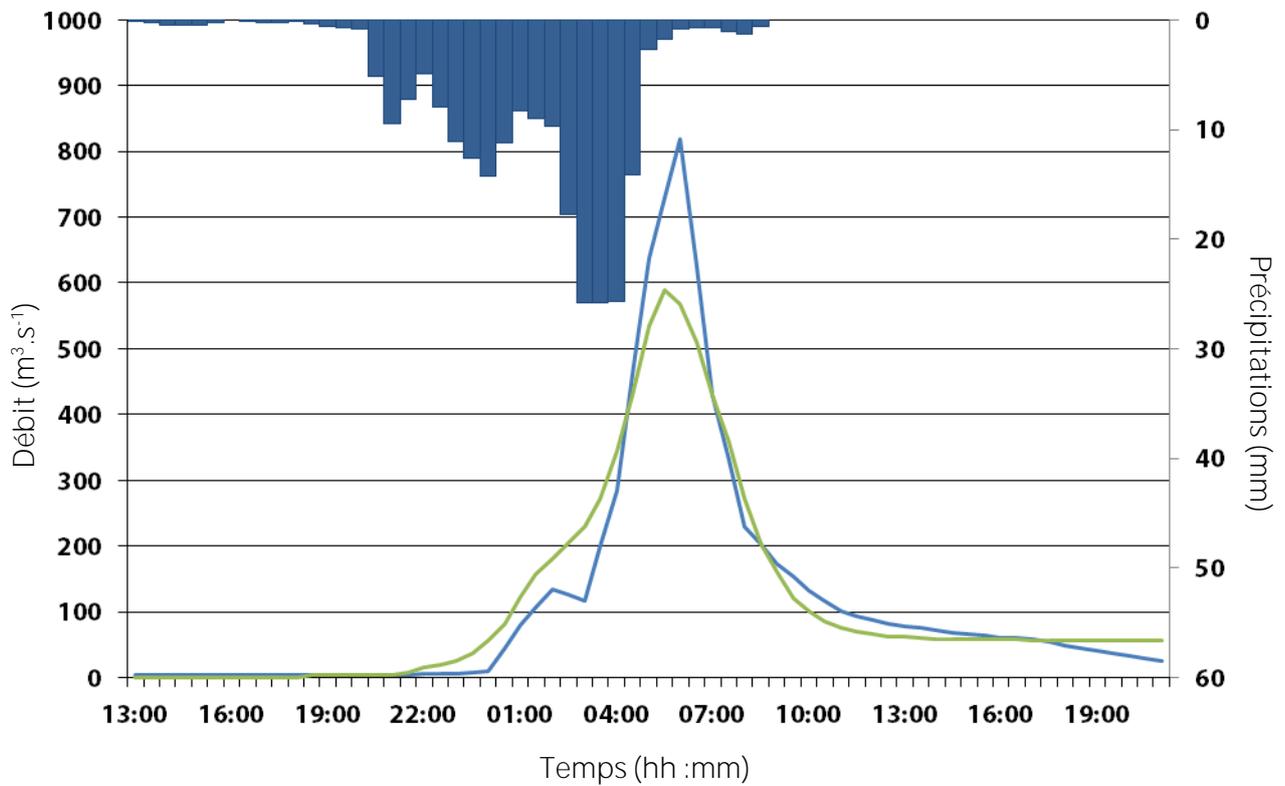


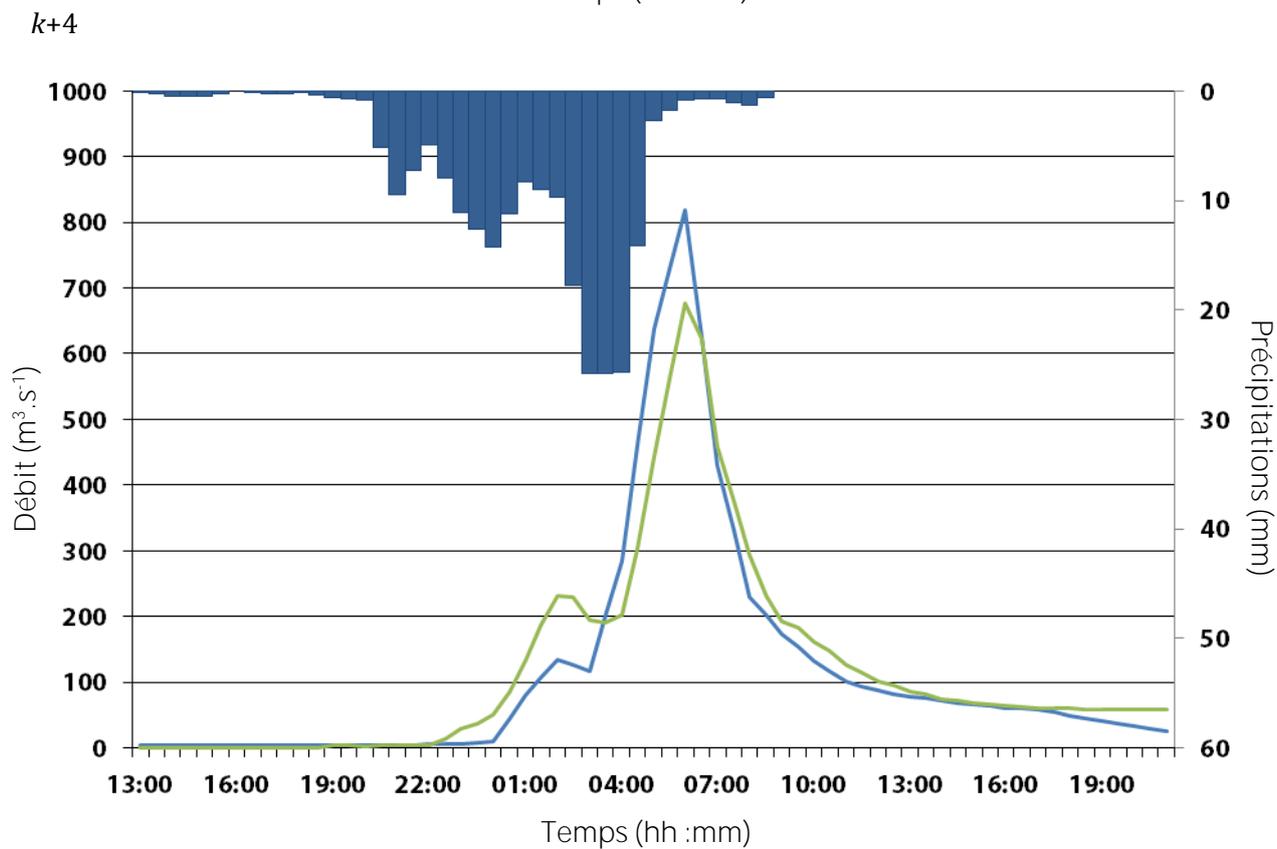
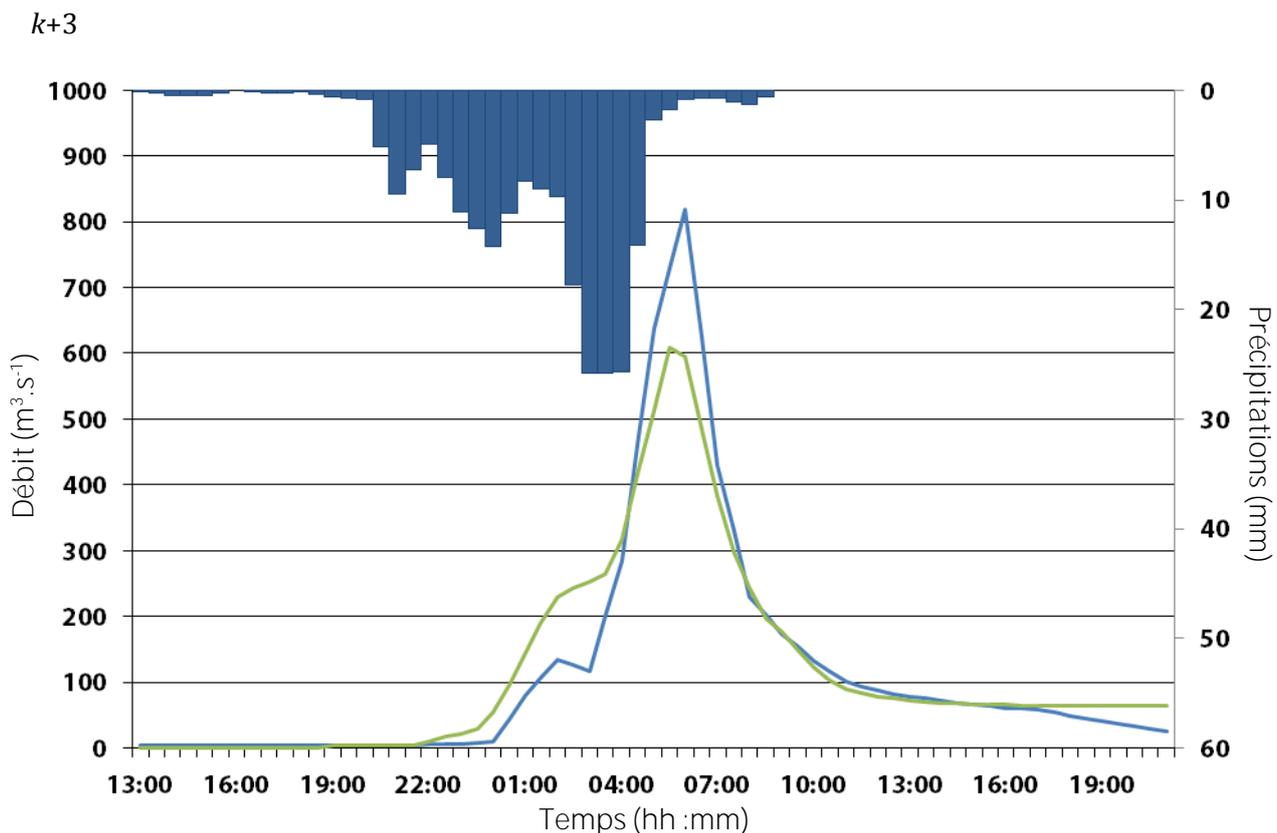
Evènement n°19 :

k+1



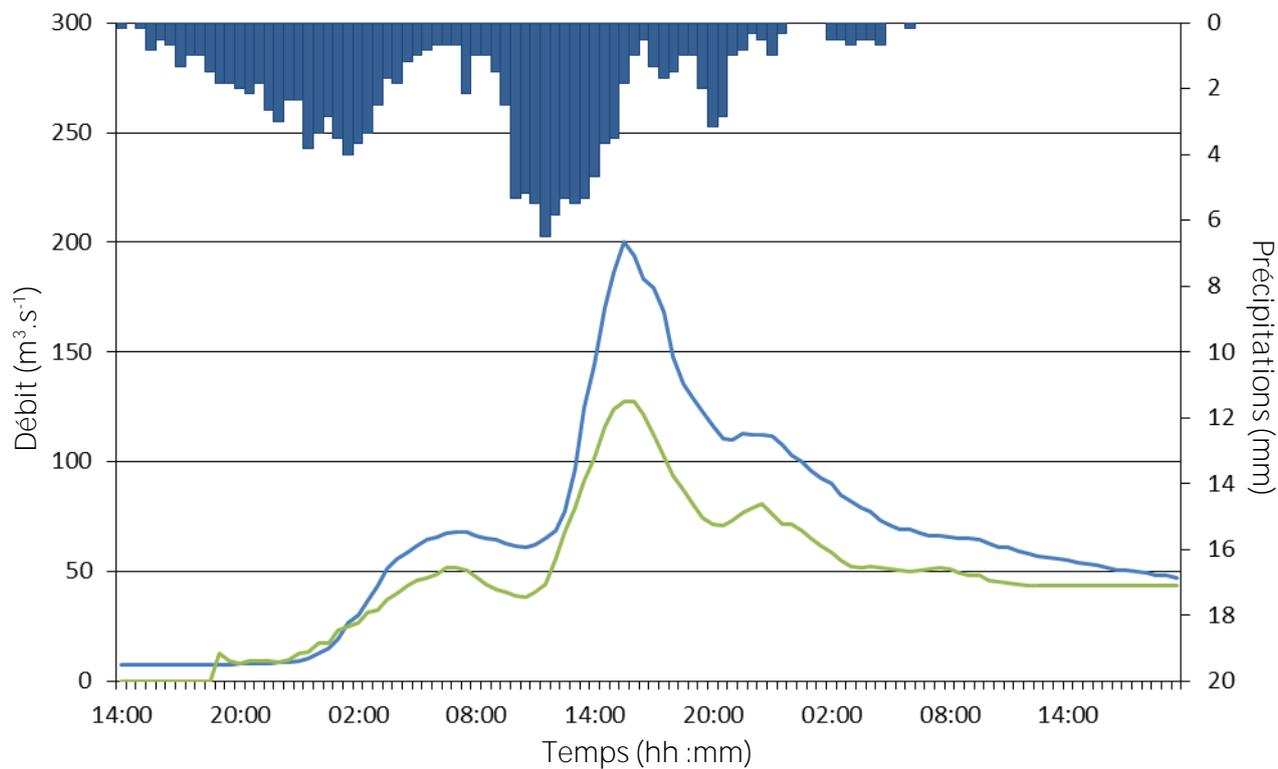
k+2



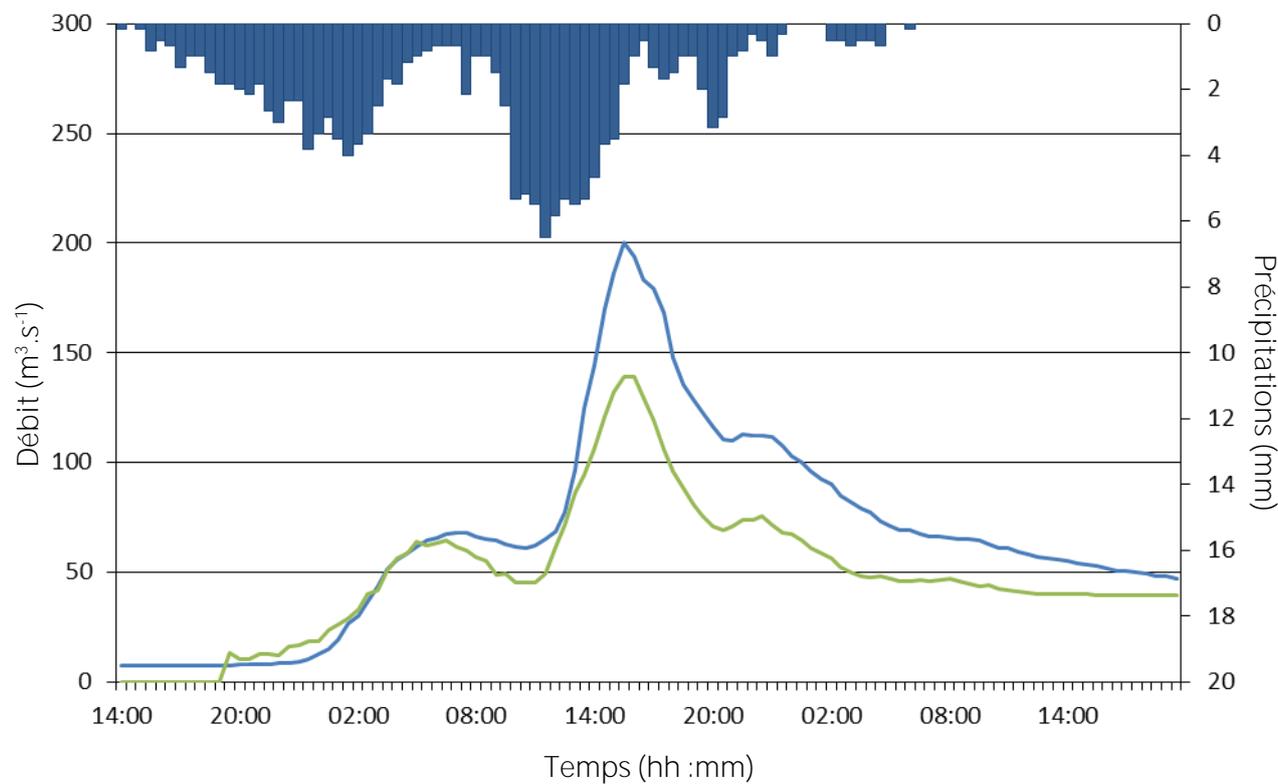


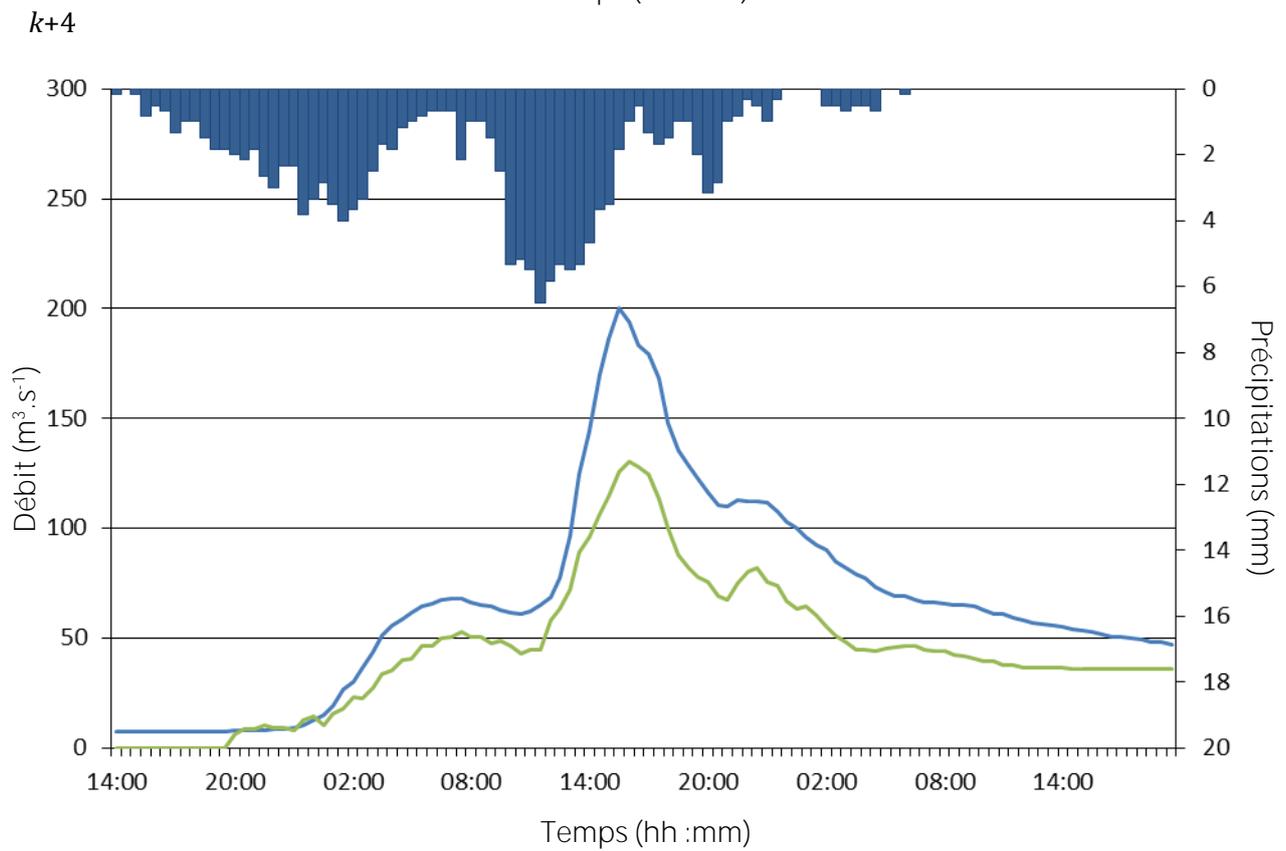
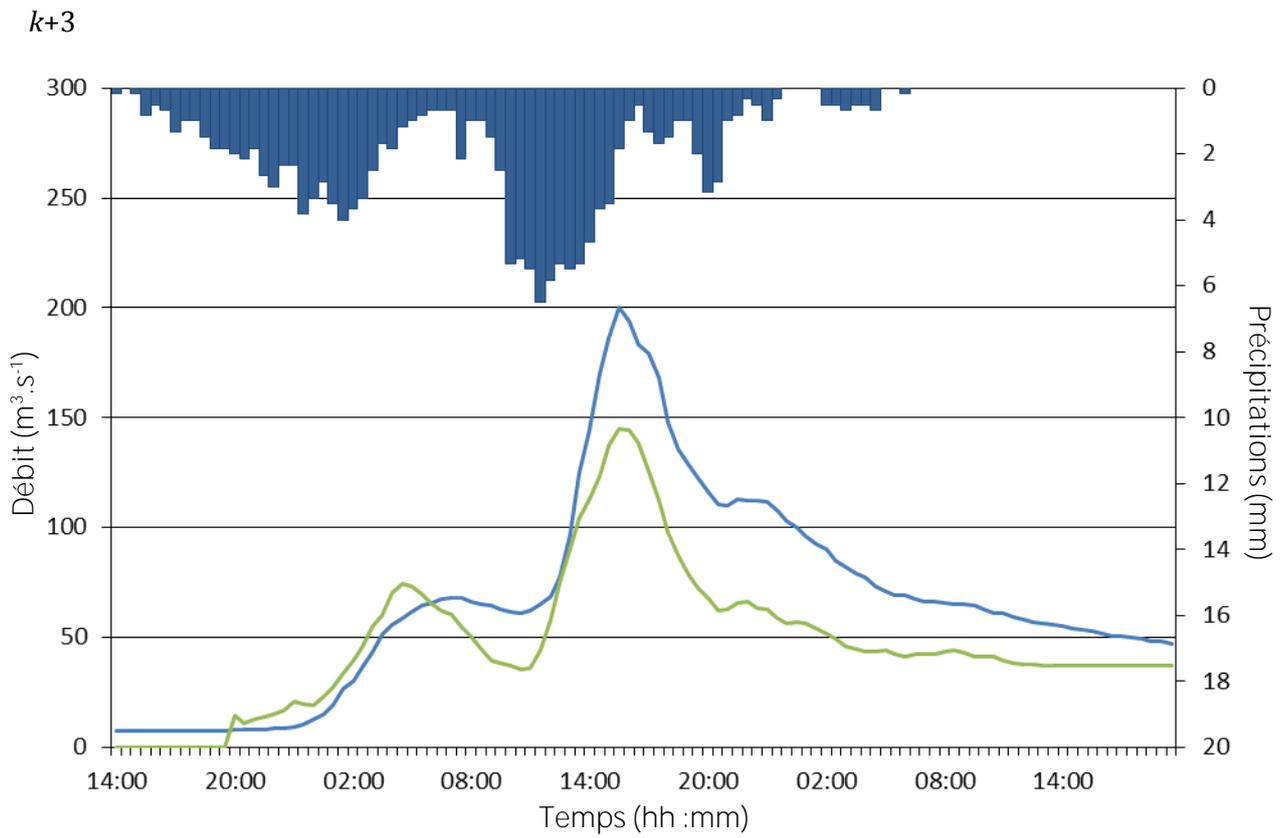
Evènement n°302 :

k+1



k+2





## **ANNEXE 6 : Prévisions sur les bassins non jaugés**

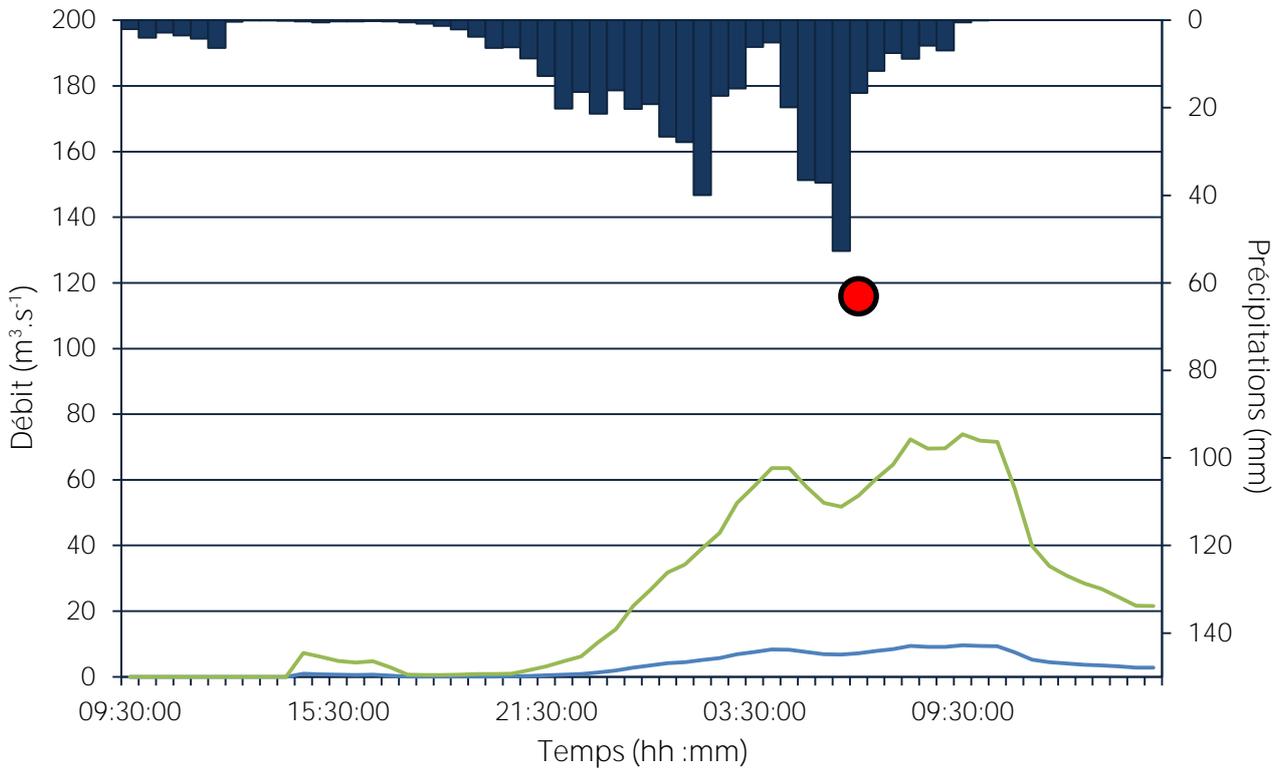
L'annexe présente les hydrogrammes en prévision sur les bassins cible.

- Le débit observé est en rouge. Lorsque la valeur maximale est reconstituée, elle est représentée par un cercle rouge à contour noir.
- Les prévisions brutes sont en bleu.
- Les prévisions corrigées sont en vert.
- Les ordonnées normales indiquent le débit en mètres cube par seconde.
- Les ordonnées inverses indiquent le cumul de pluie en millimètres.
- Le temps est gradué en demi-heures.

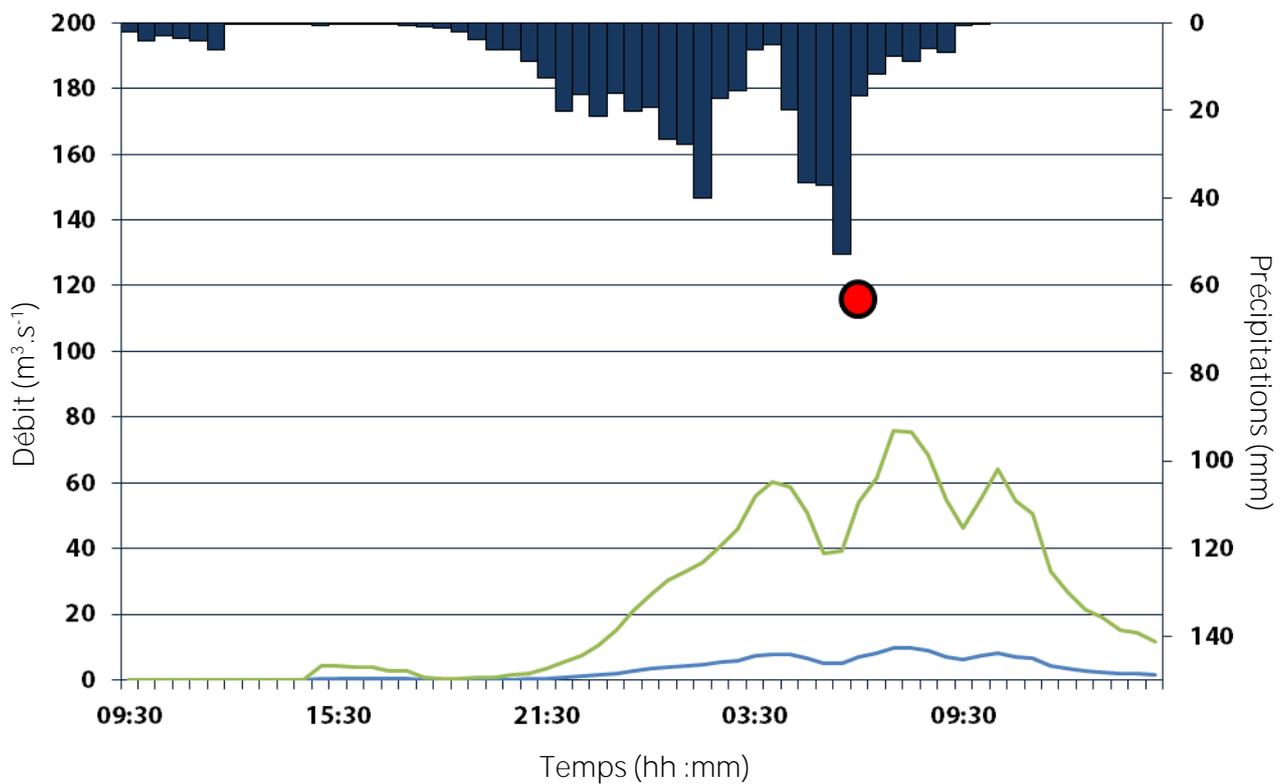
Il n'y a aucune prévision des pluies futures utilisée.

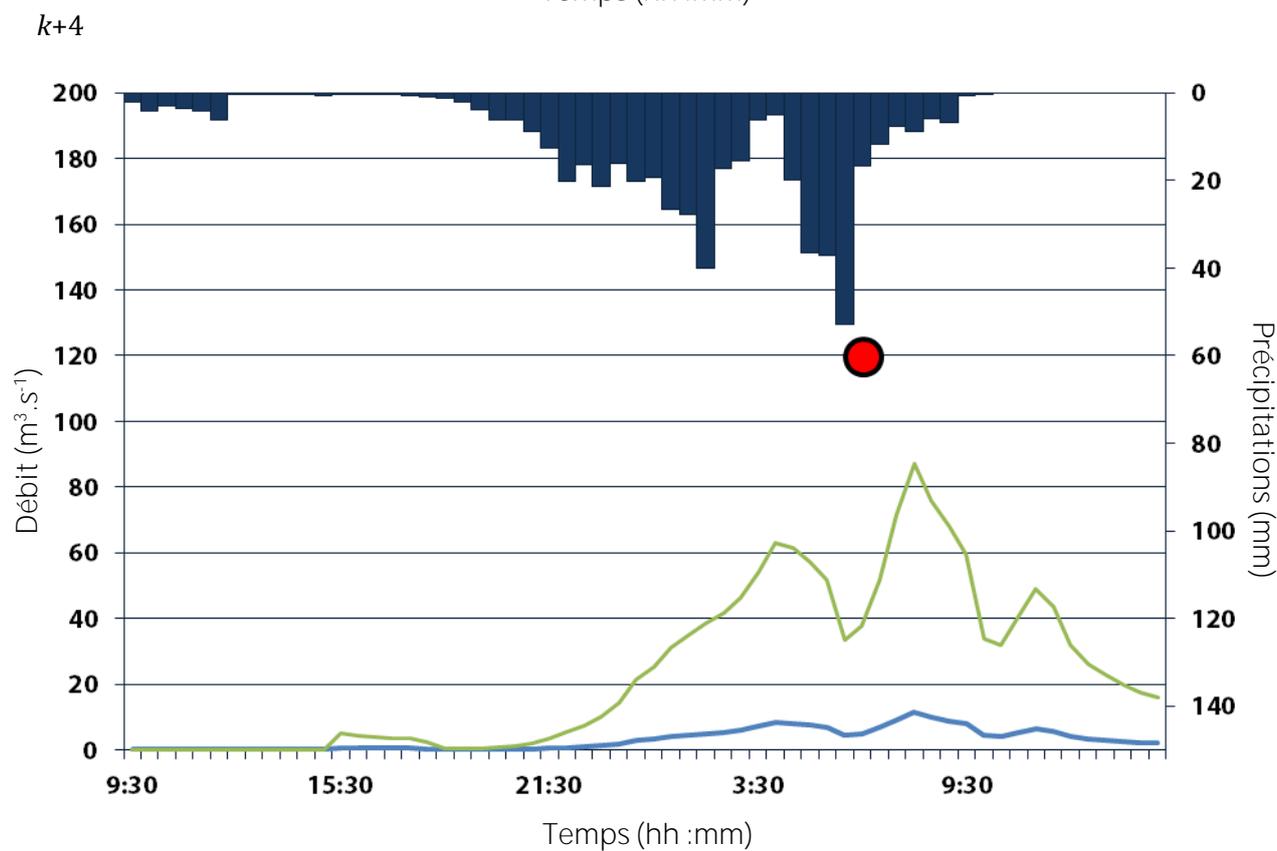
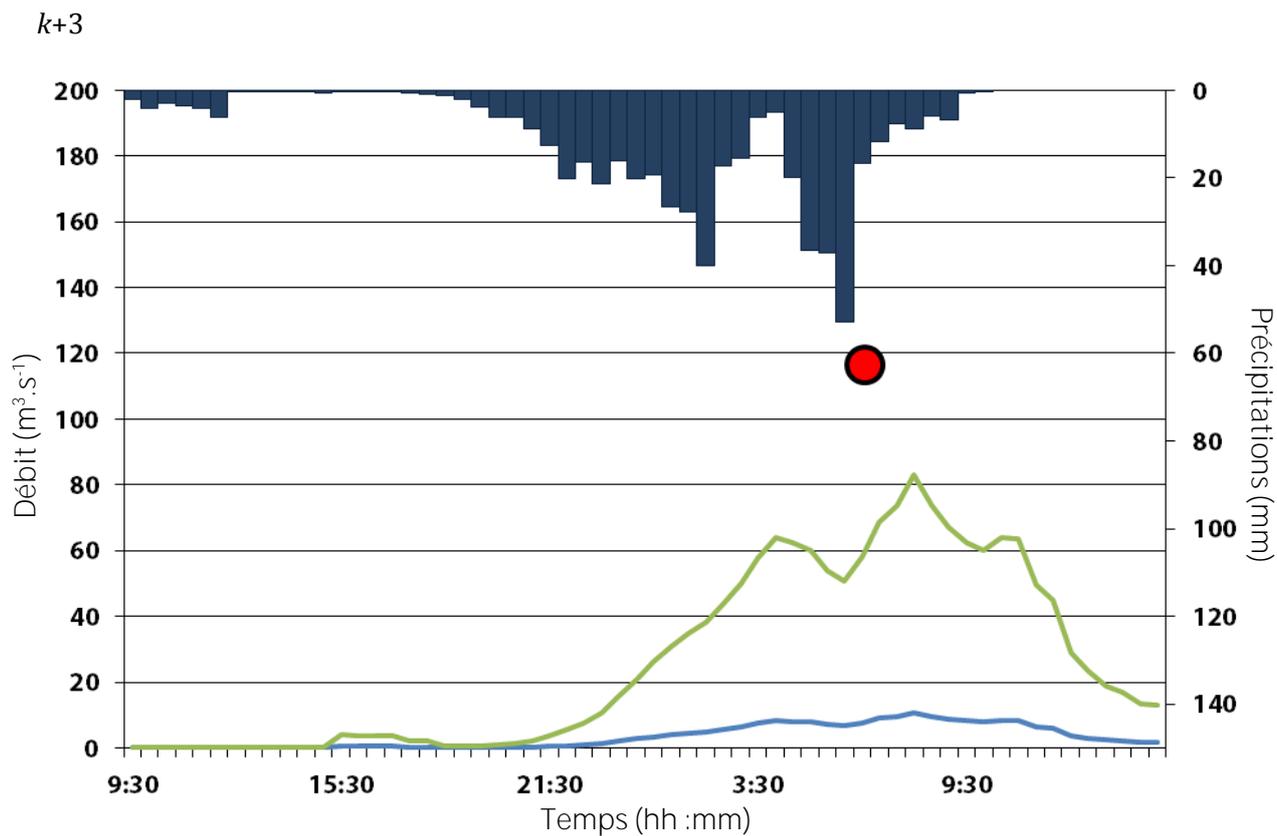
Bassin du ruisseau du Lyonnais, évènement n°928 :

$k+1$



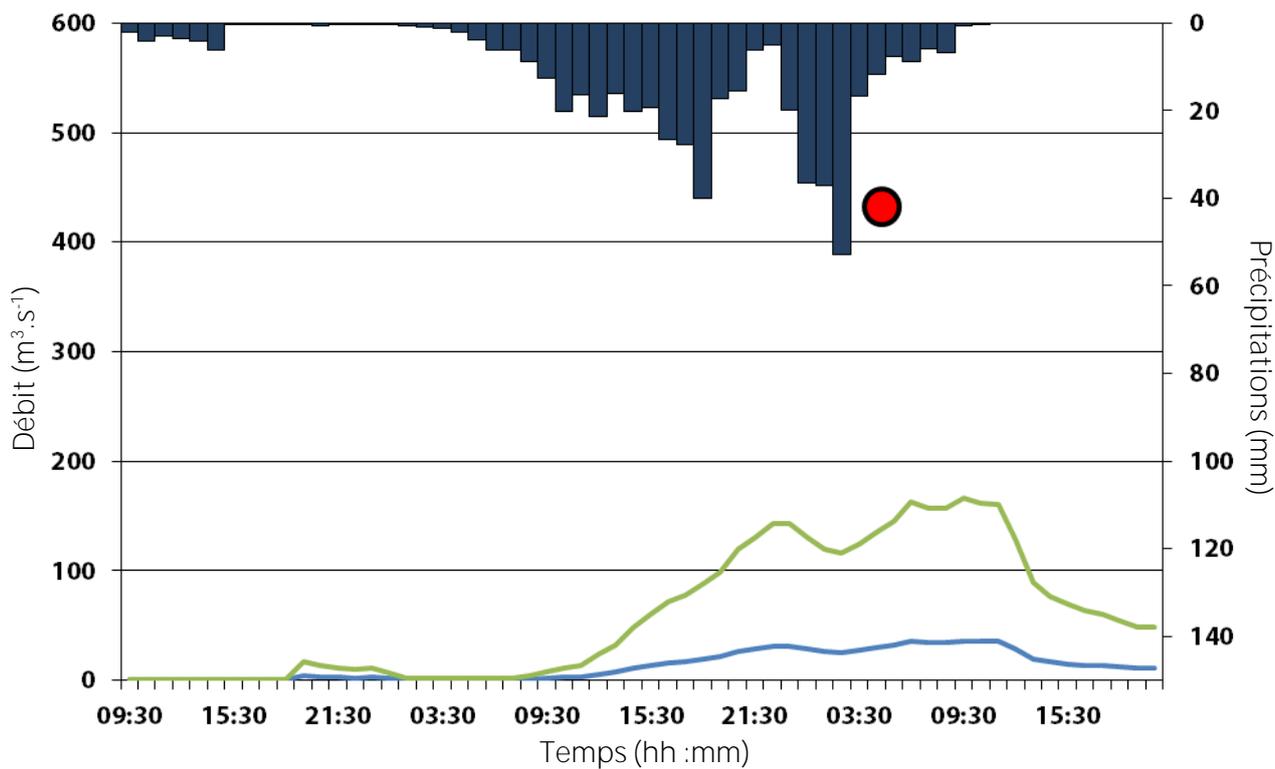
$k+2$



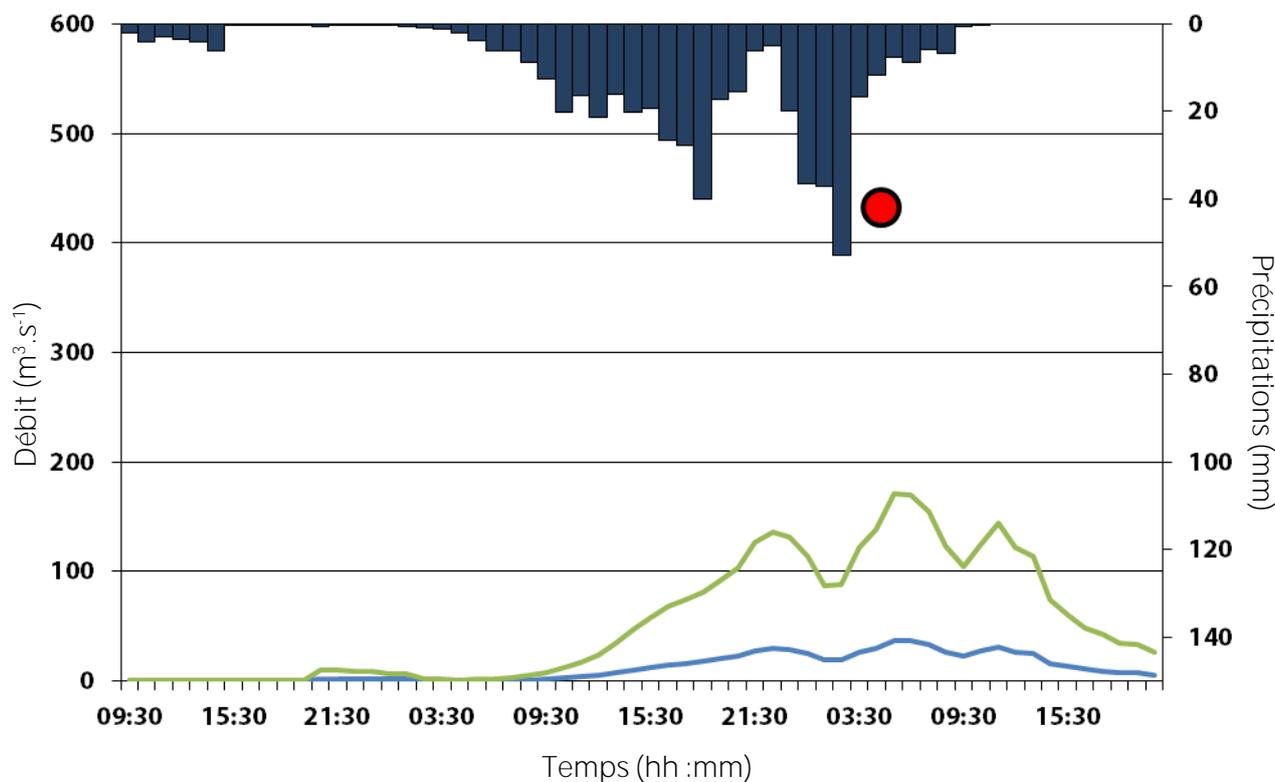


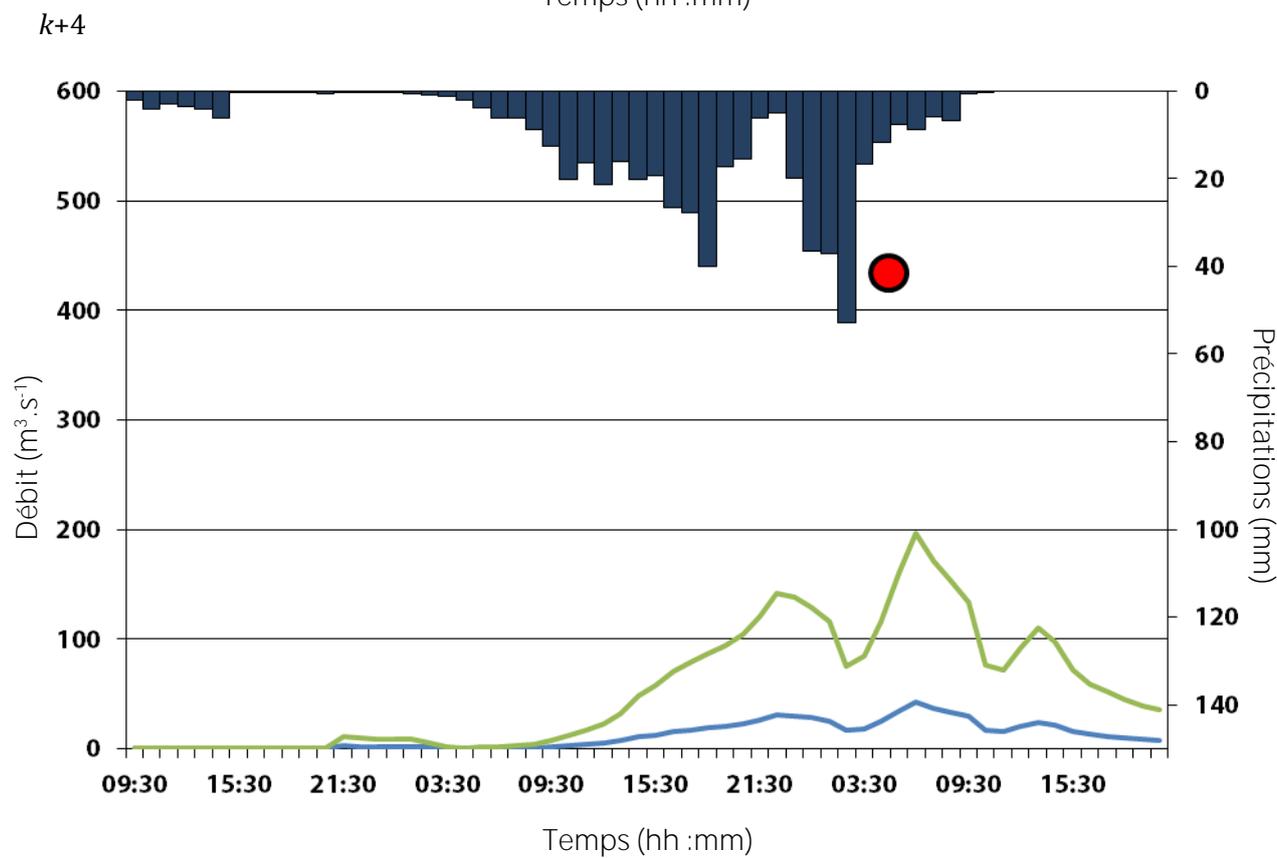
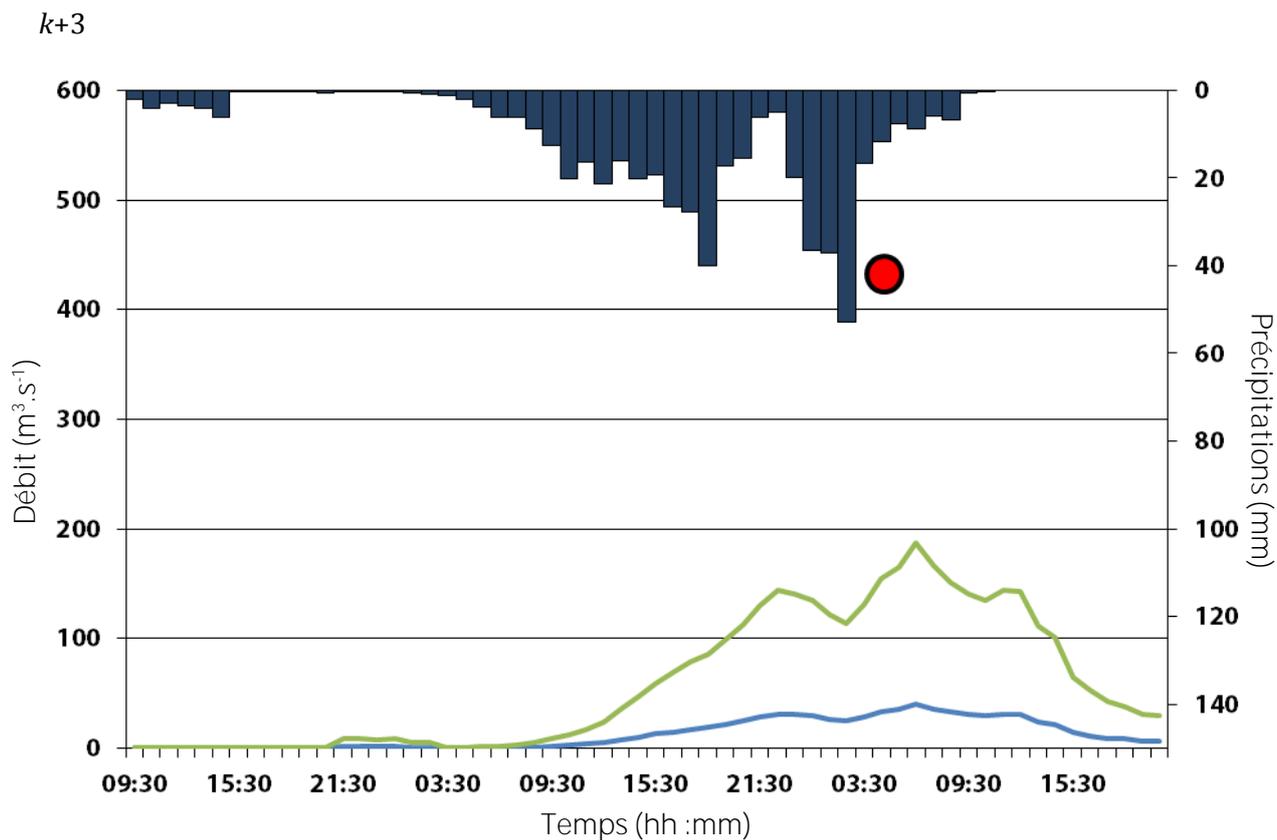
Bassin de l'Alzon, évènement n°930 :

k+1



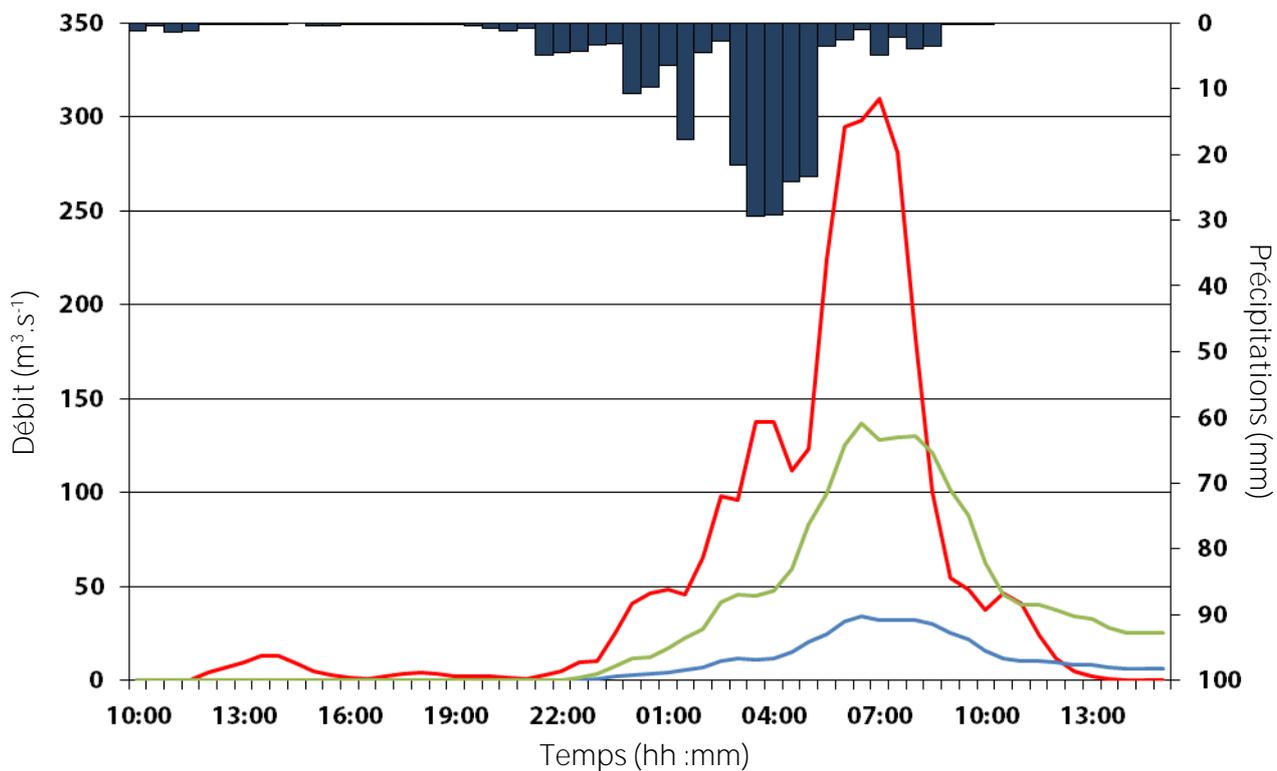
k+2



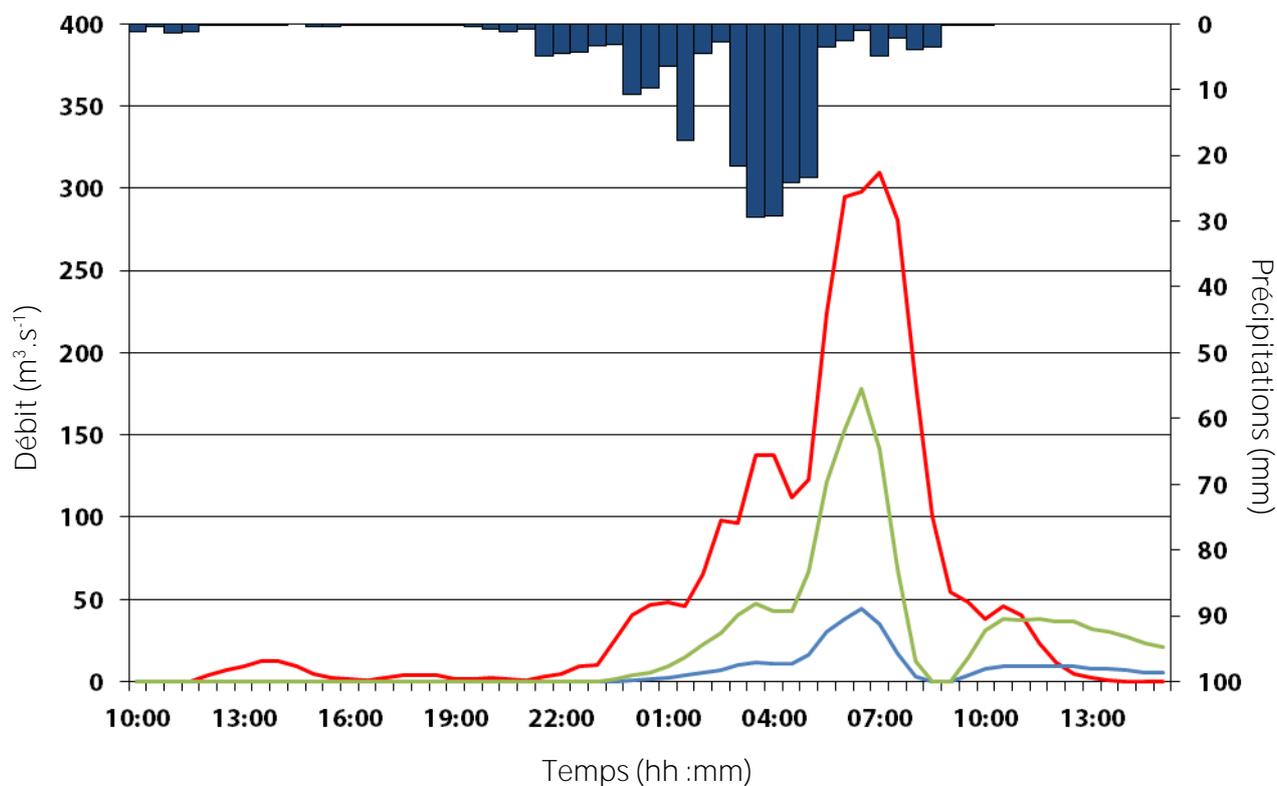


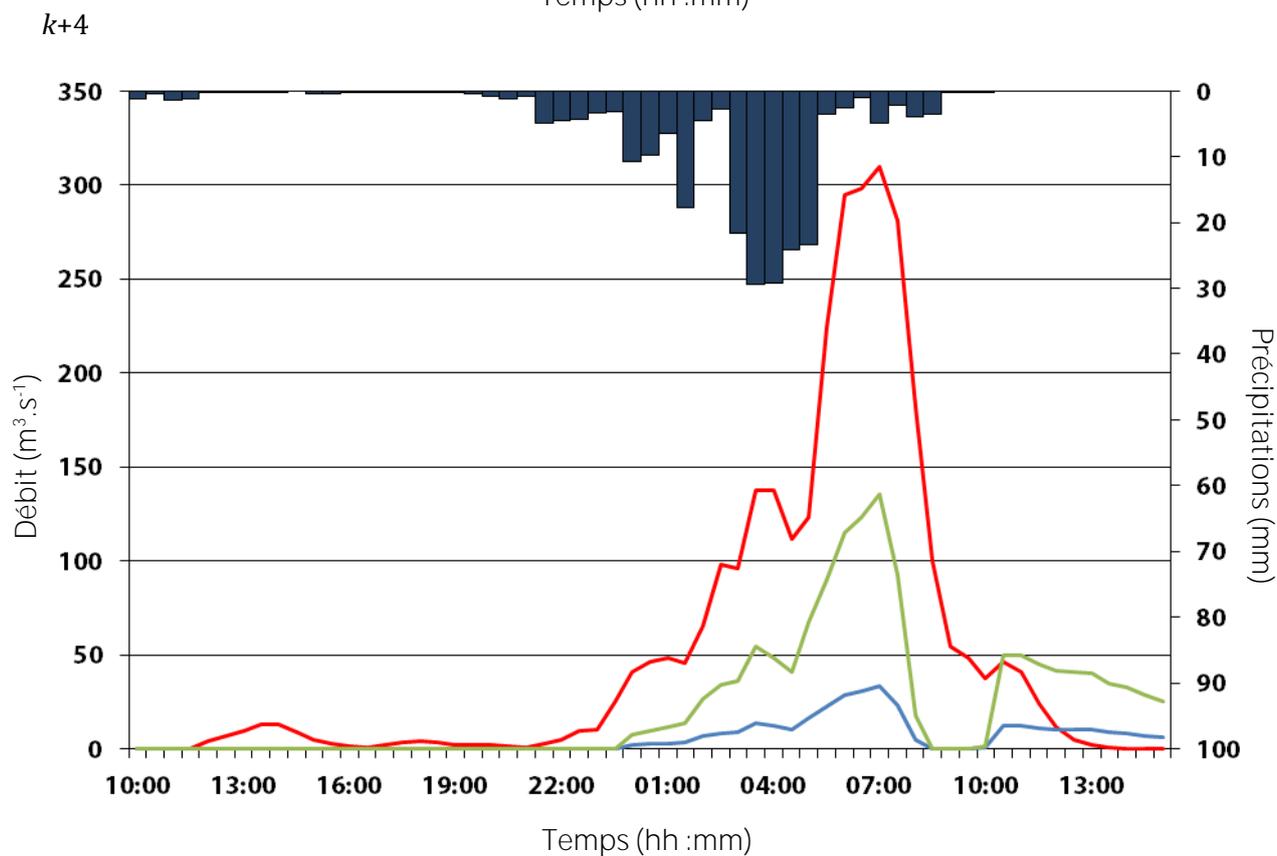
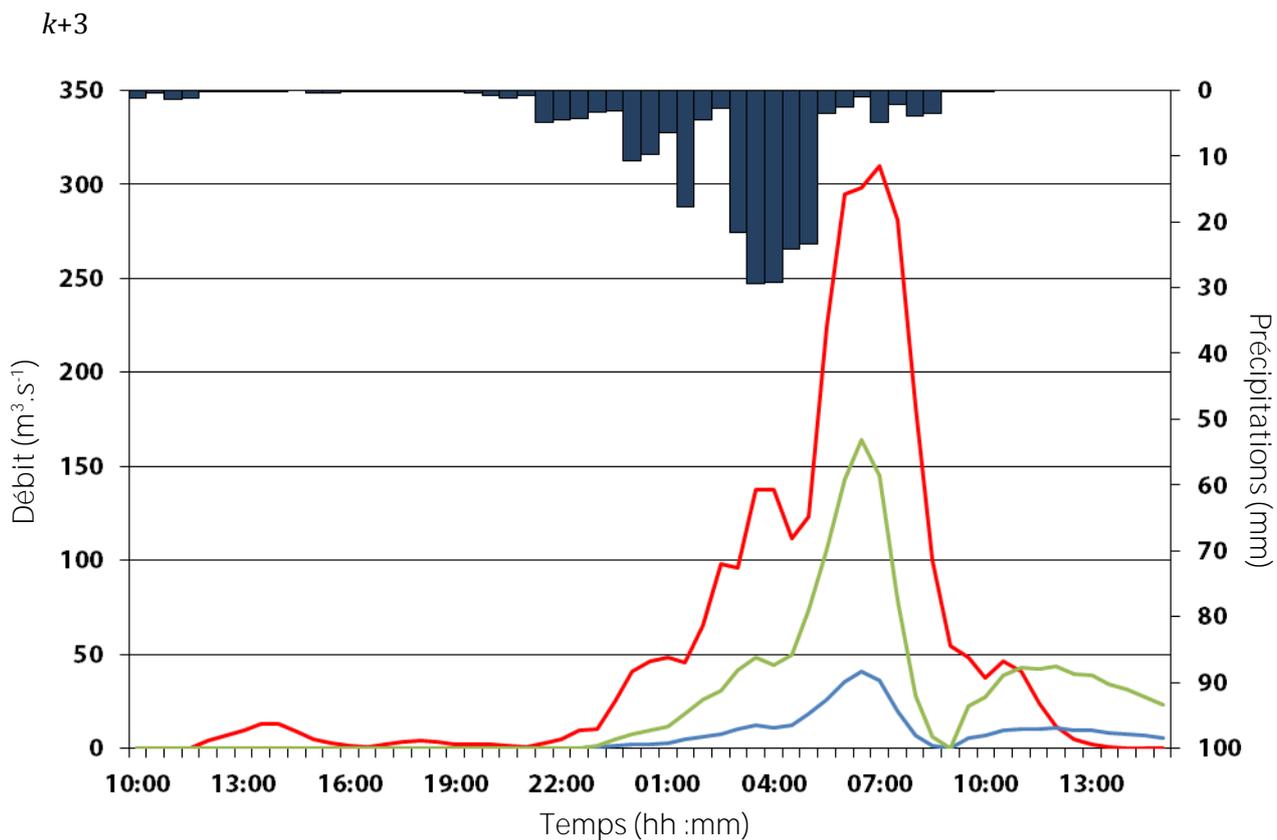
Bassin de la Salandre, évènement n°926 :

k+1



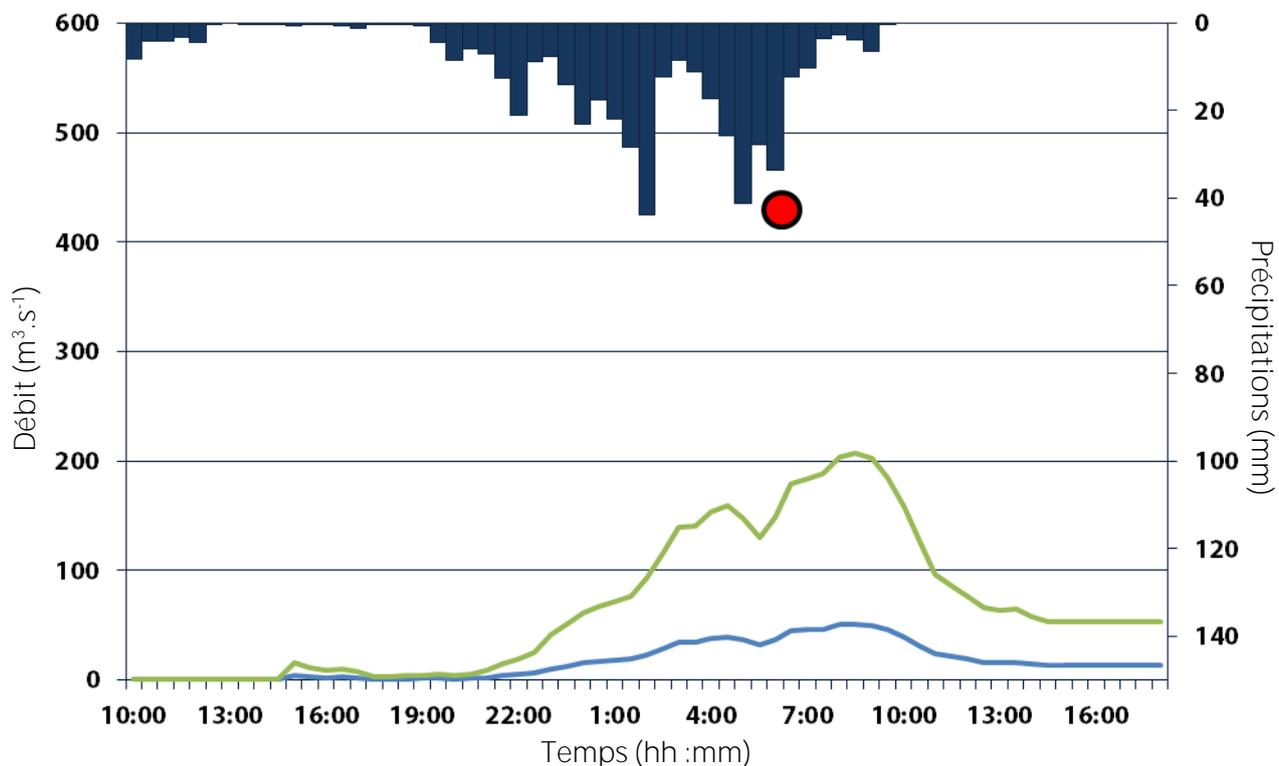
k+2



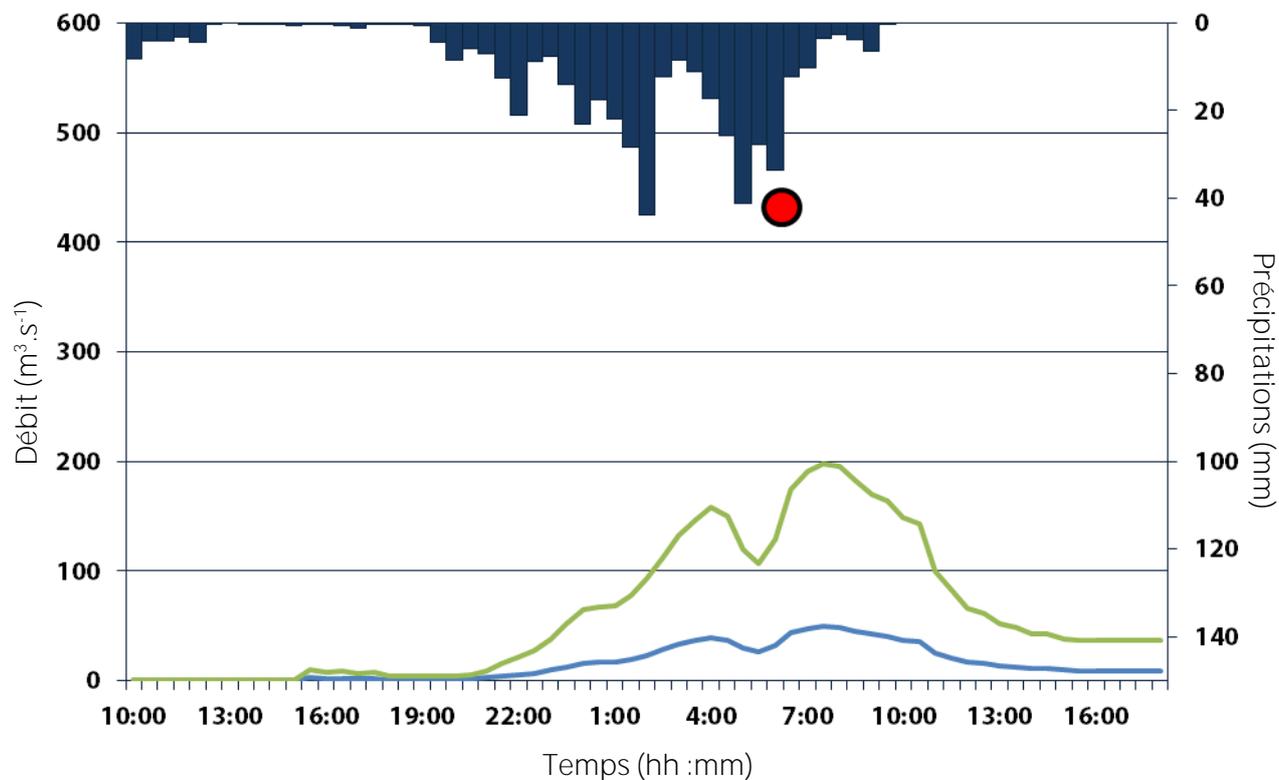


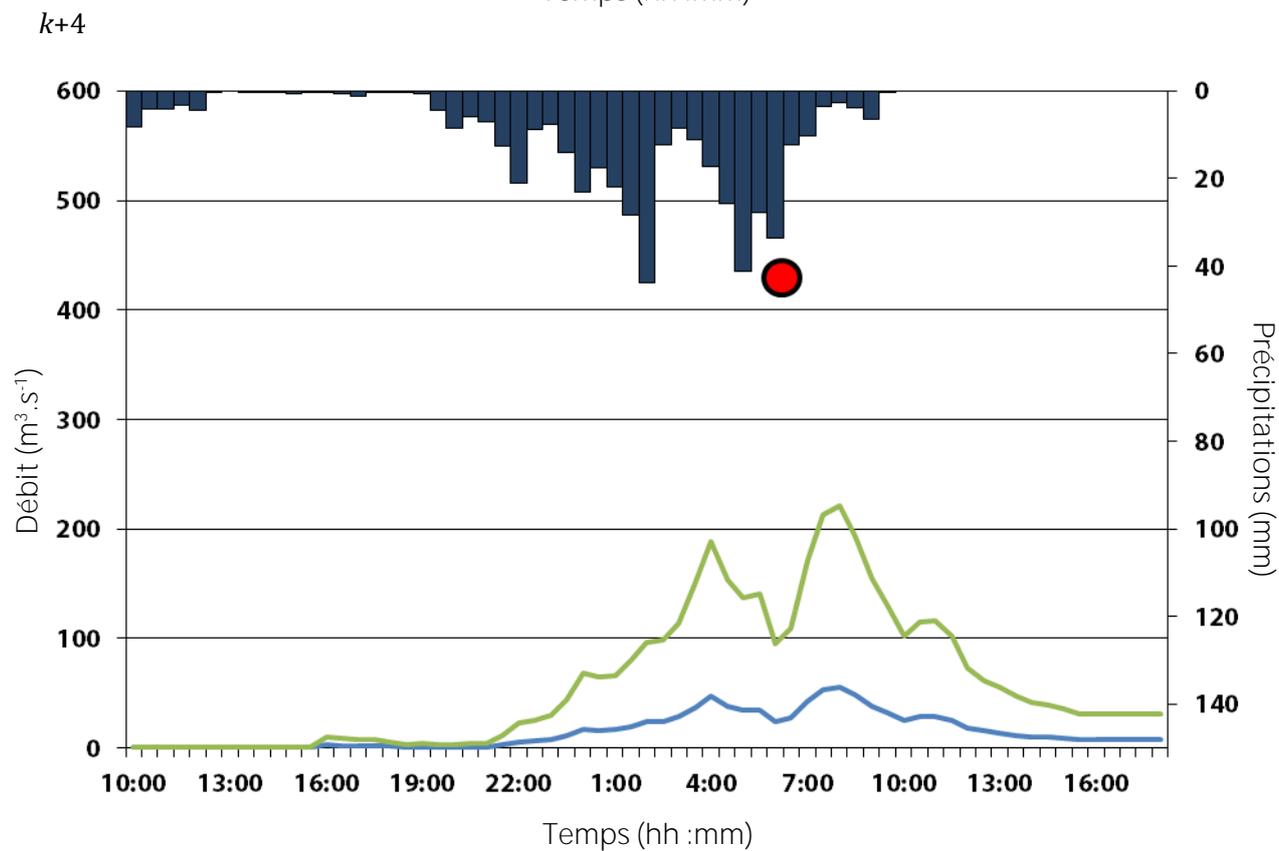
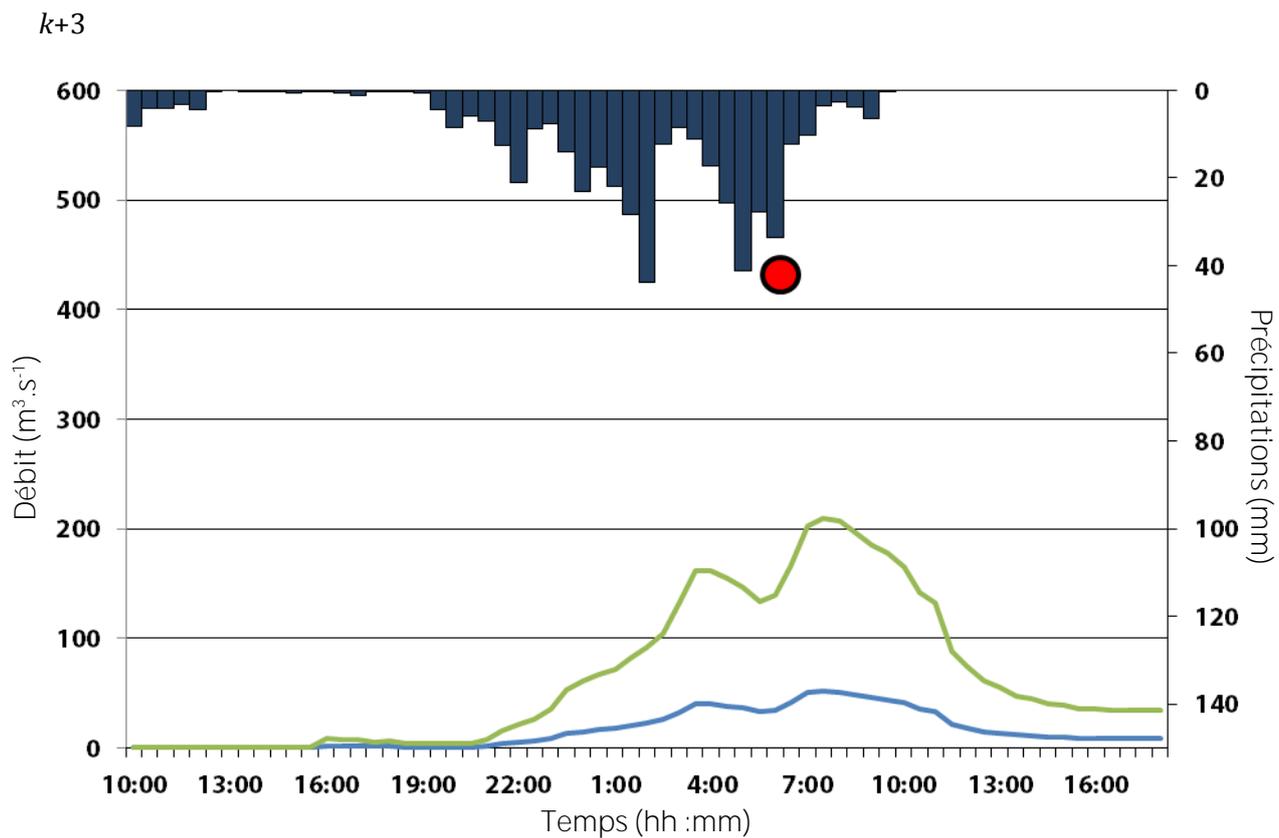
Bassin de l'Amous, évènement n°929 :

k+1



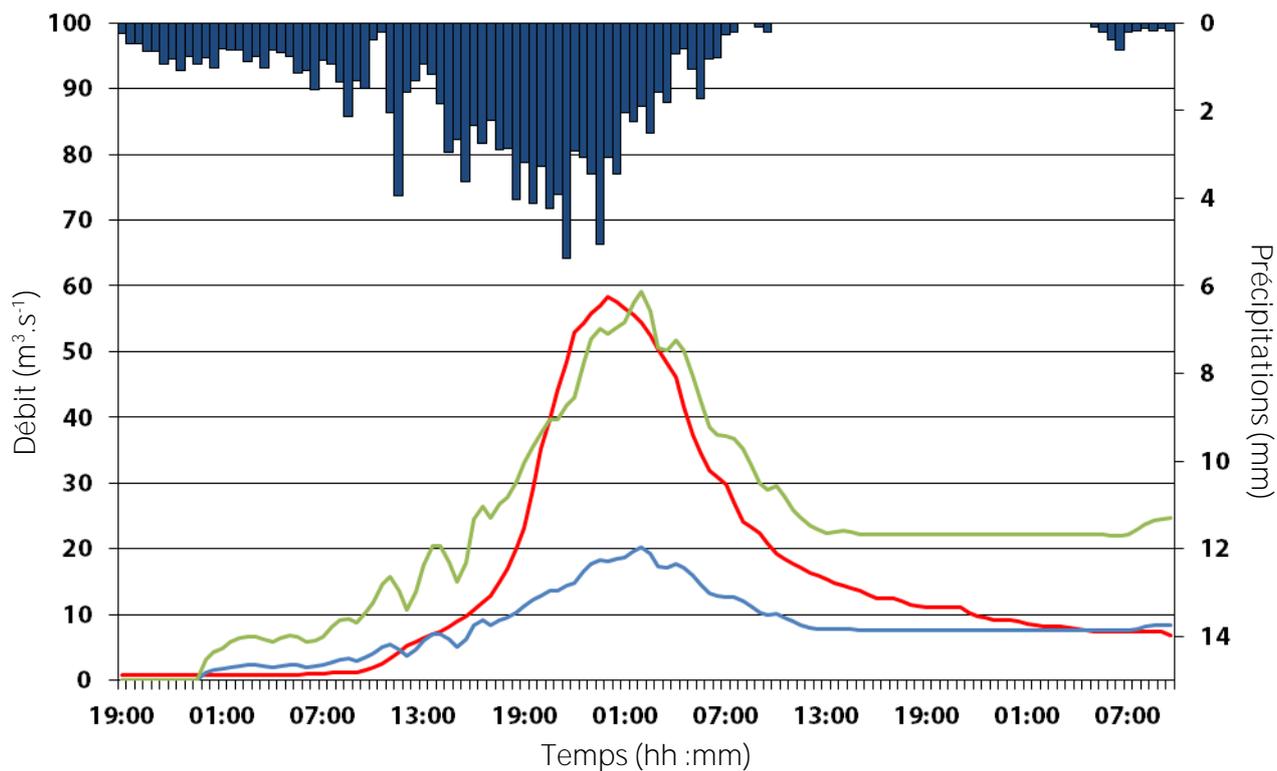
k+2



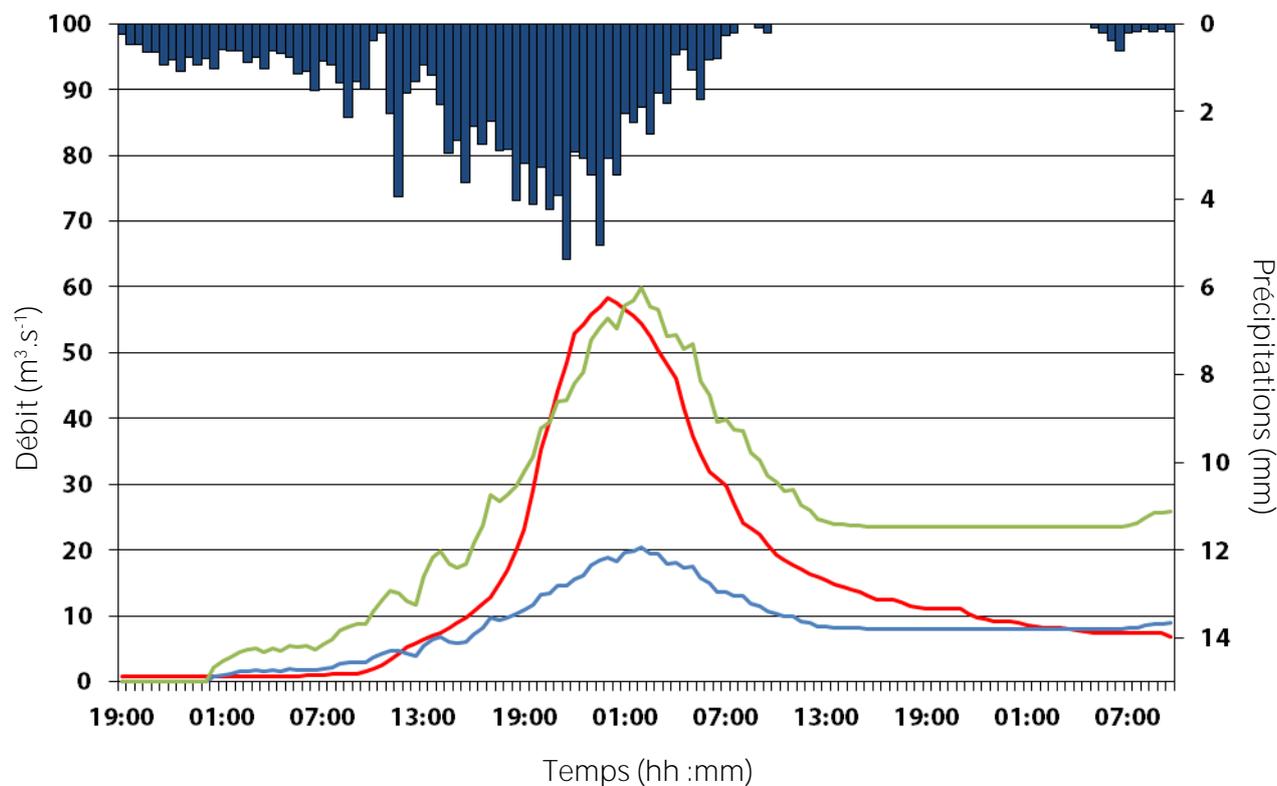


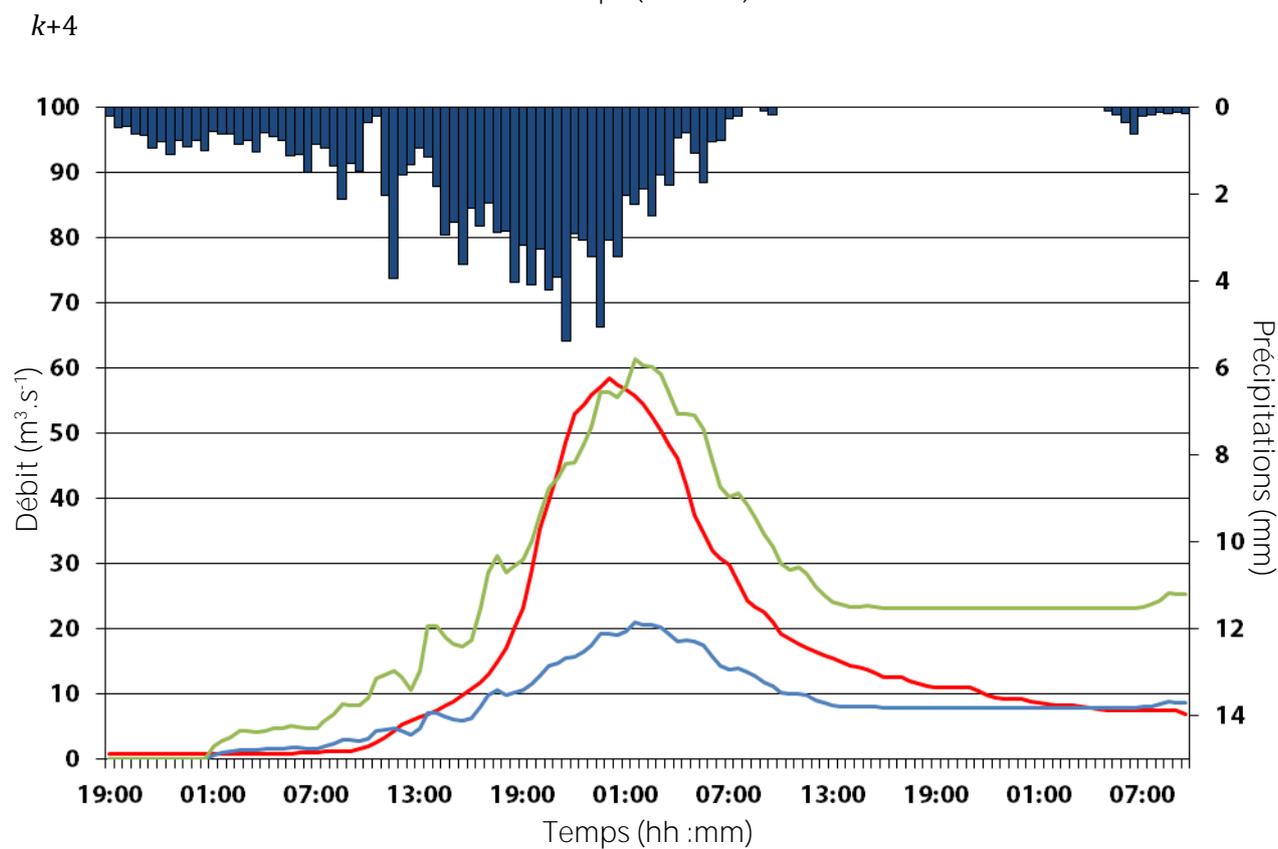
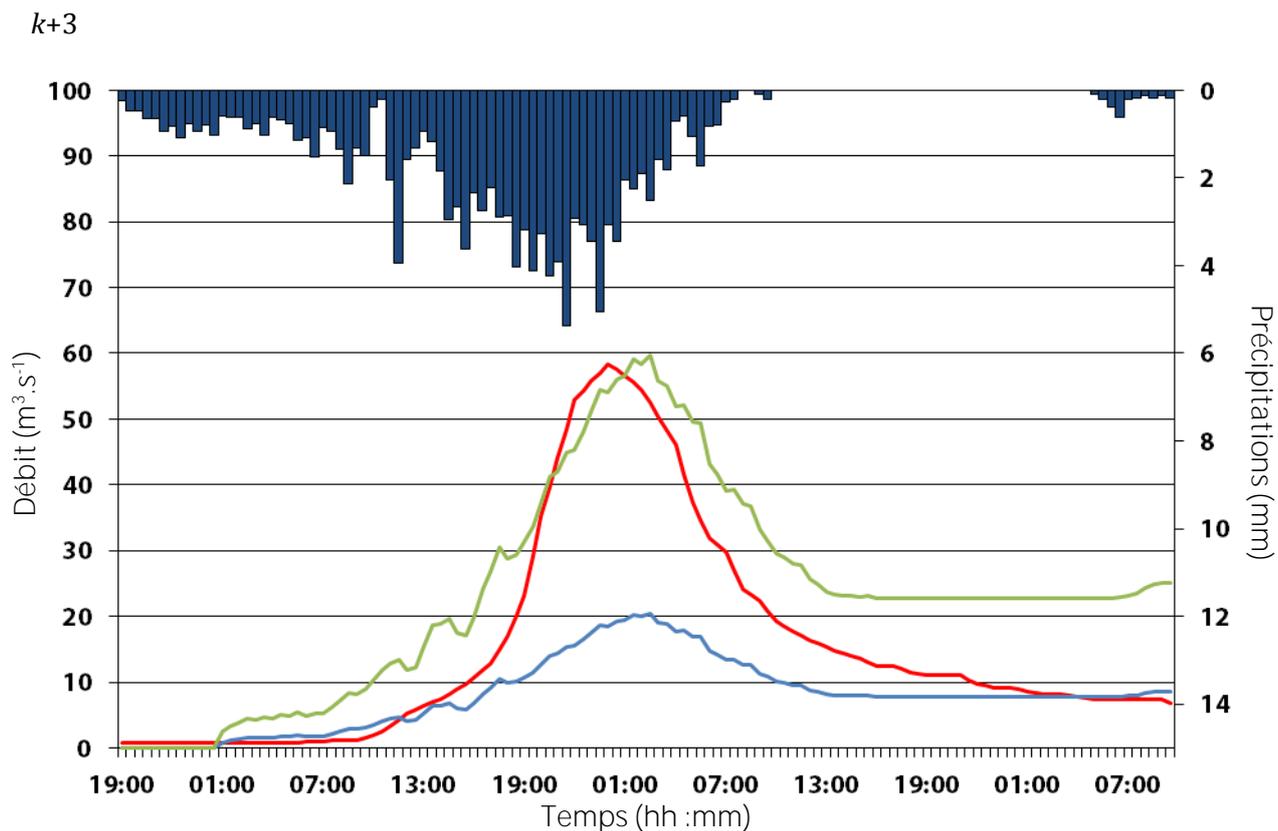
Bassin de l'Auzonnet, évènement n°921 :

$k+1$



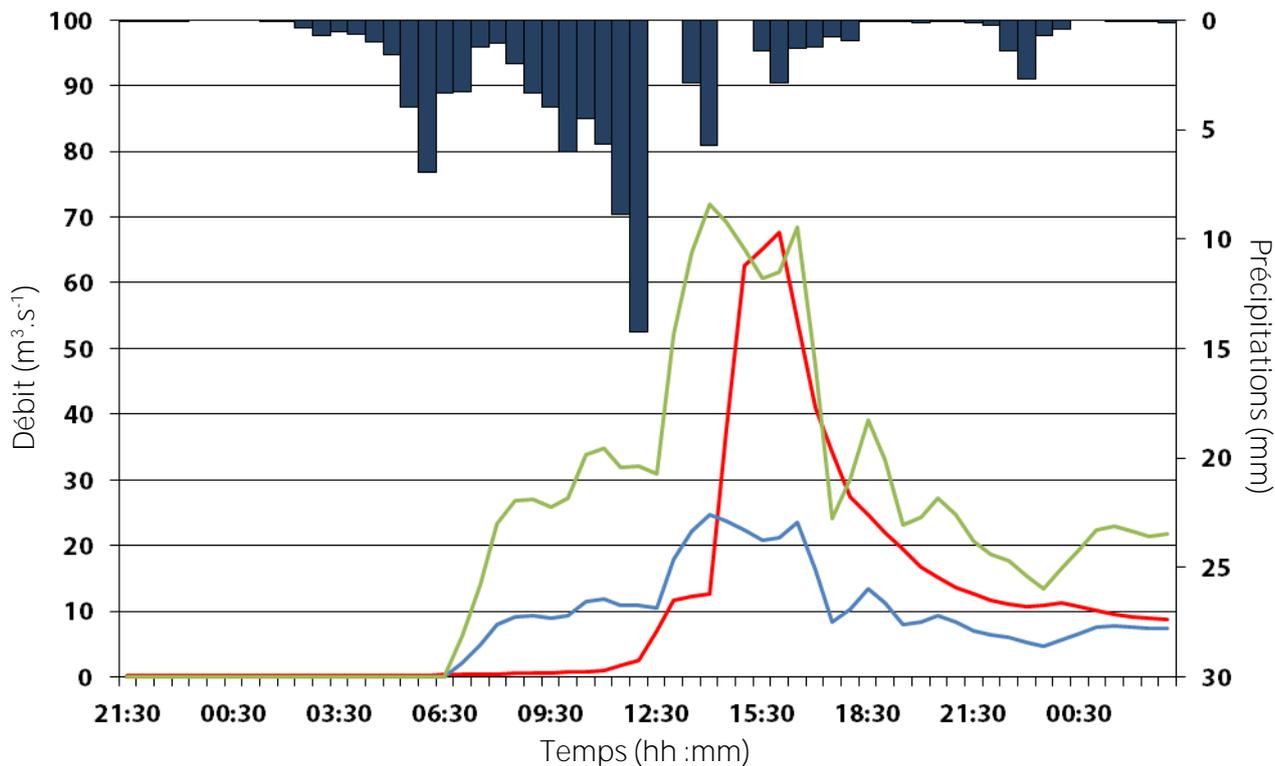
$k+2$



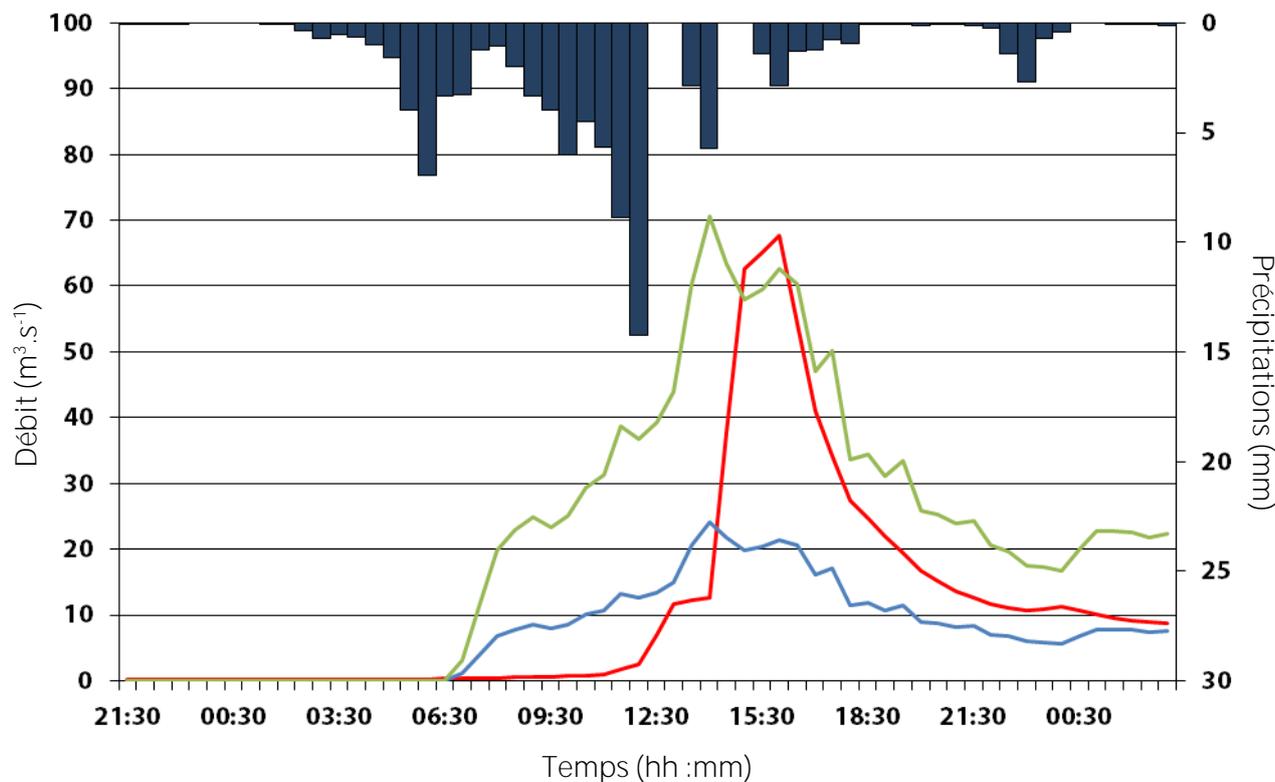


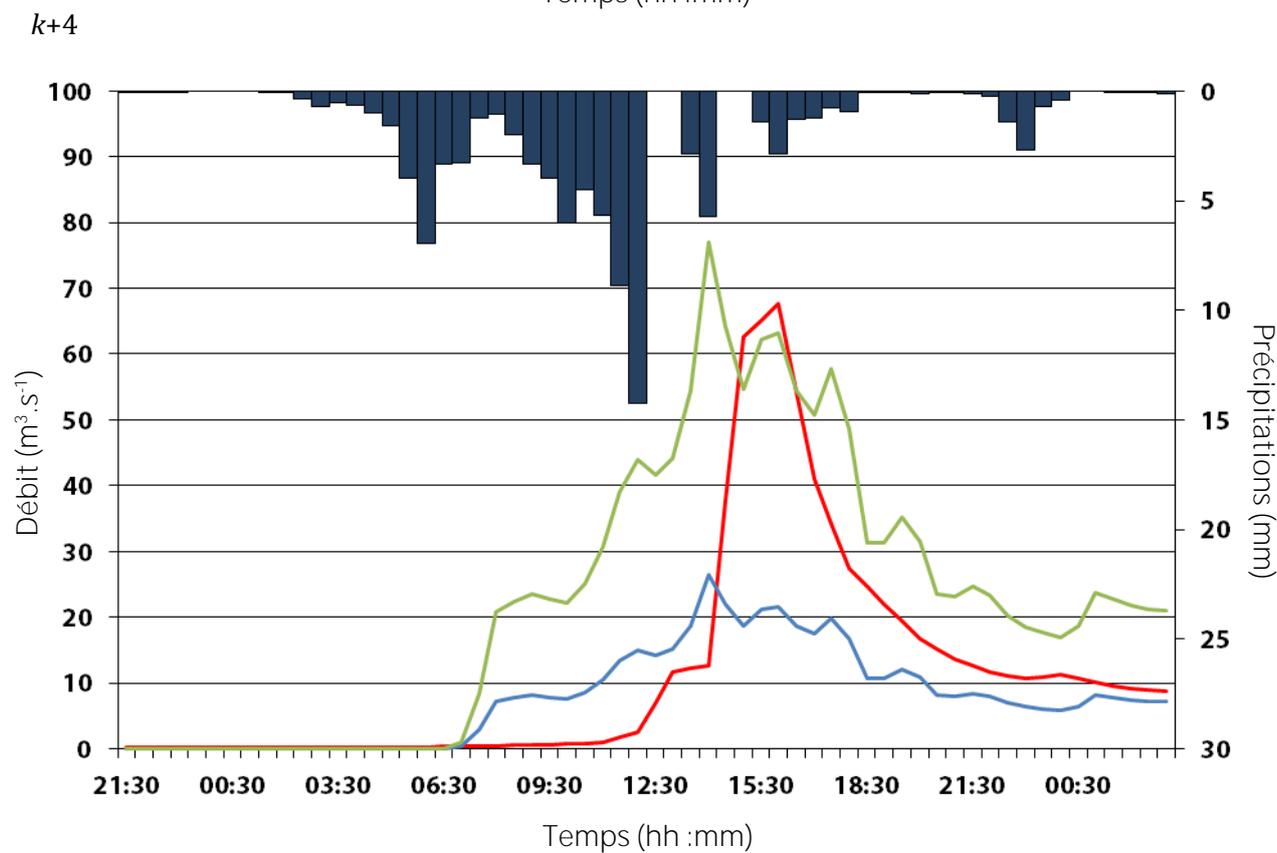
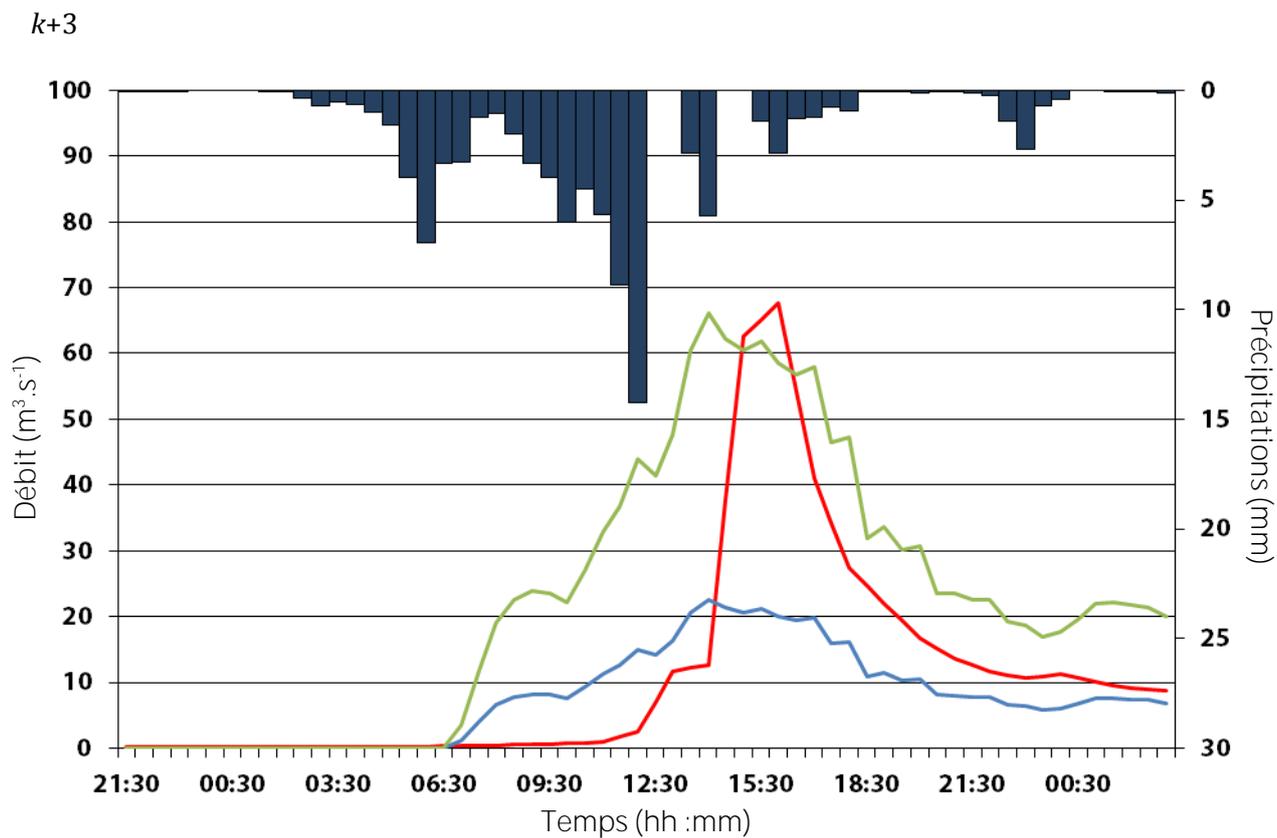
Bassin de l'Auzonnet, évènement n°922 :

k+1



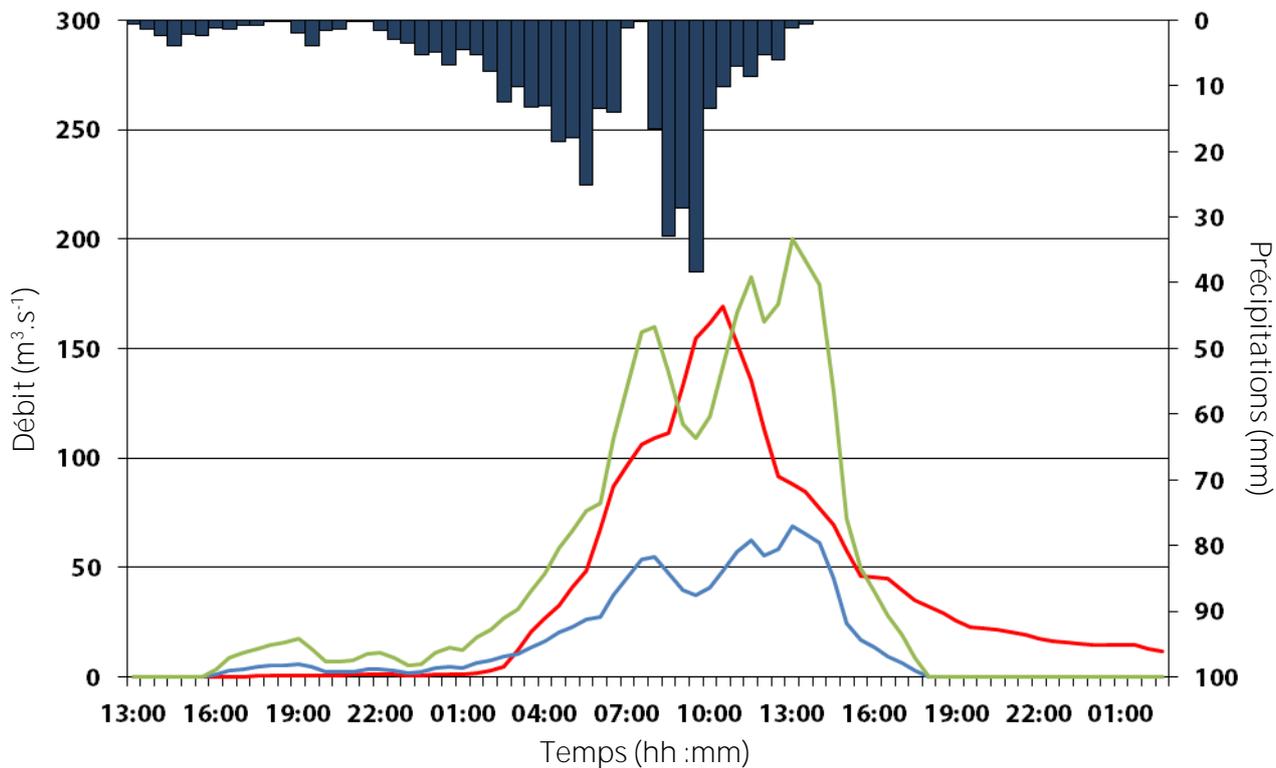
k+2



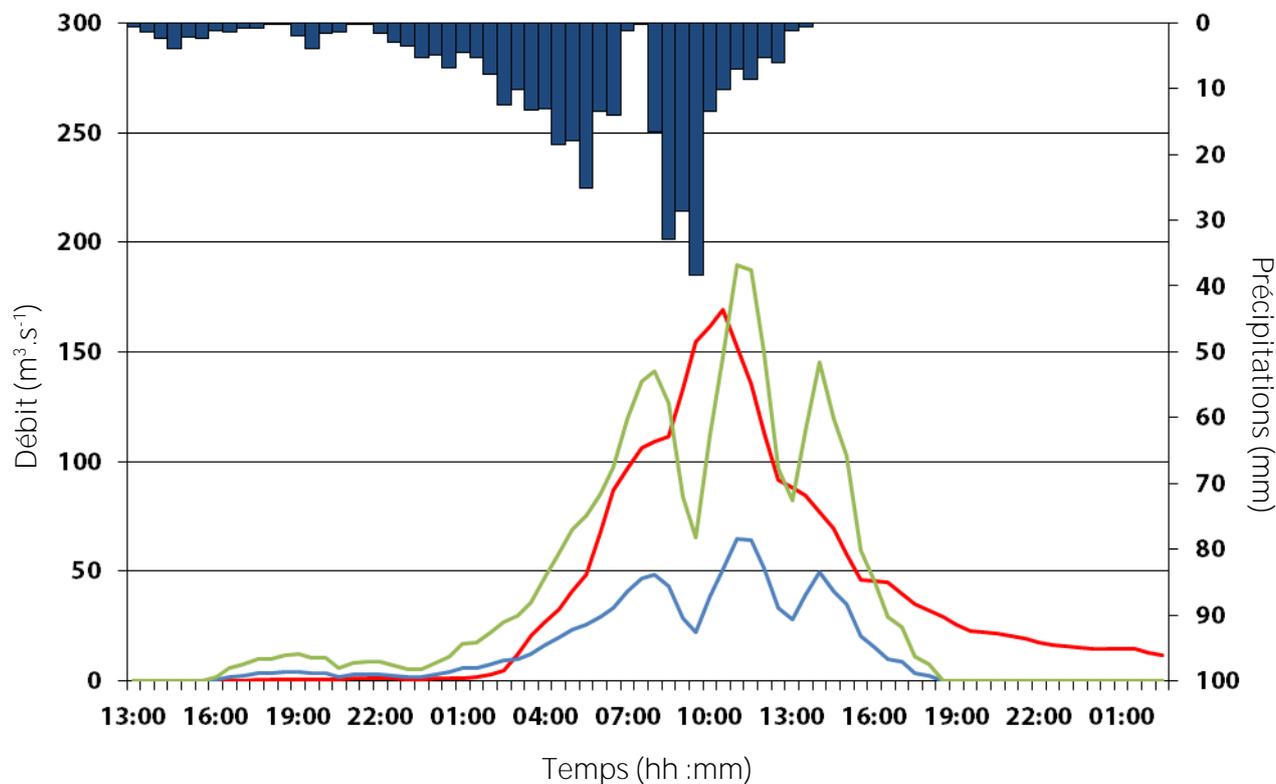


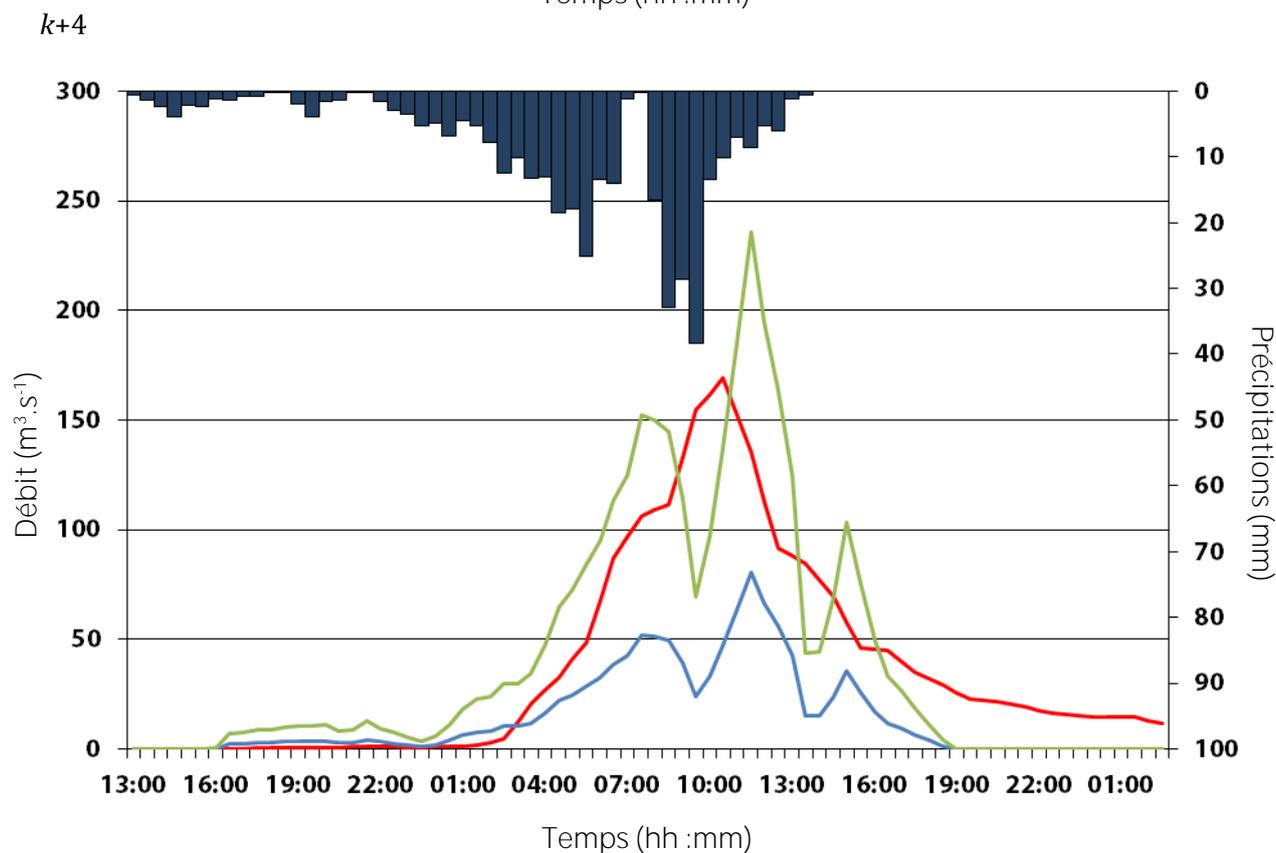
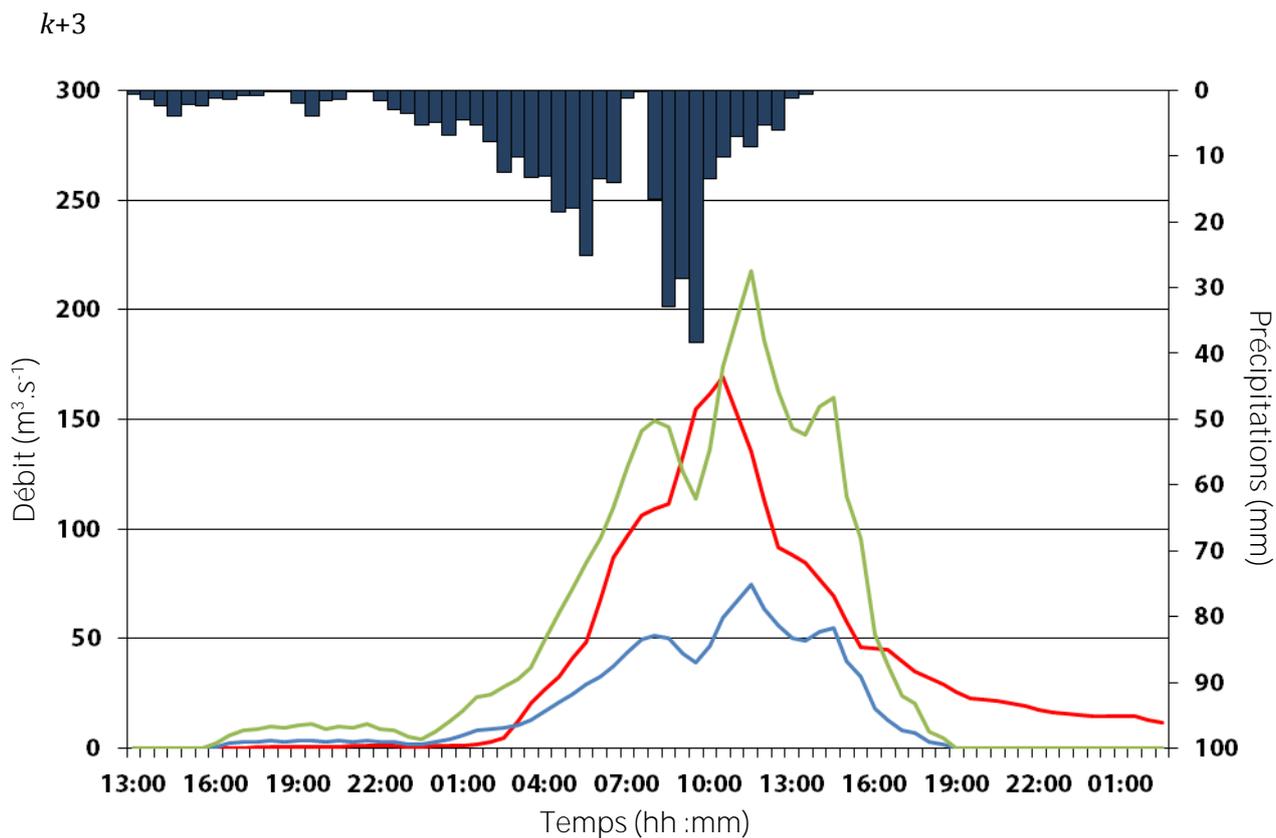
Bassin de l'Auzonnet, évènement n°923 :

k+1



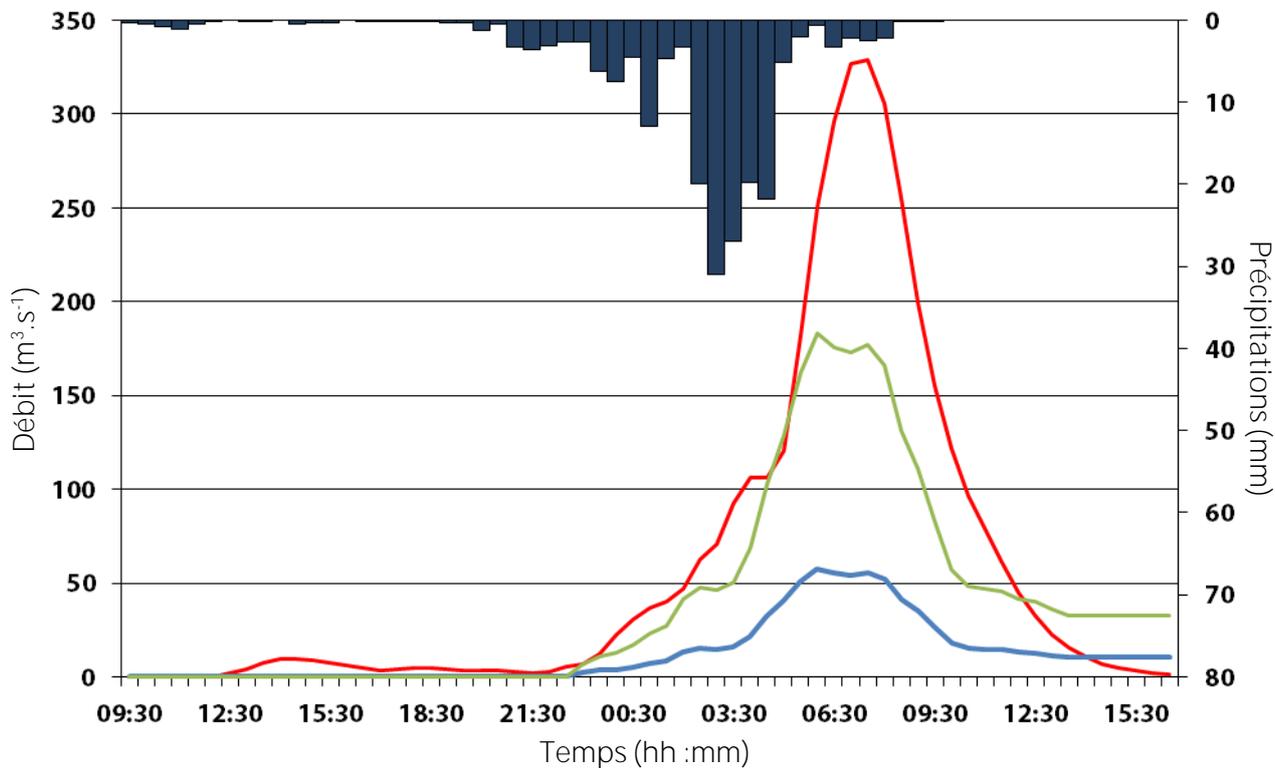
k+2



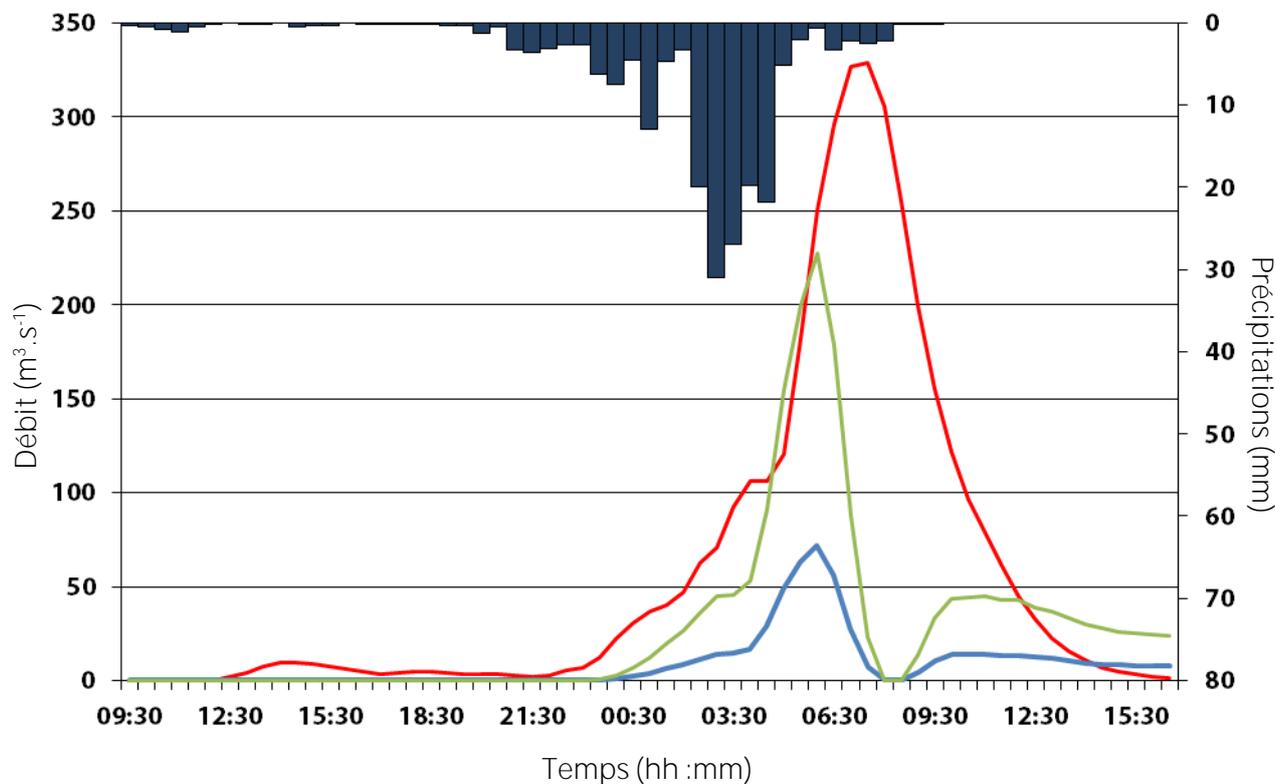


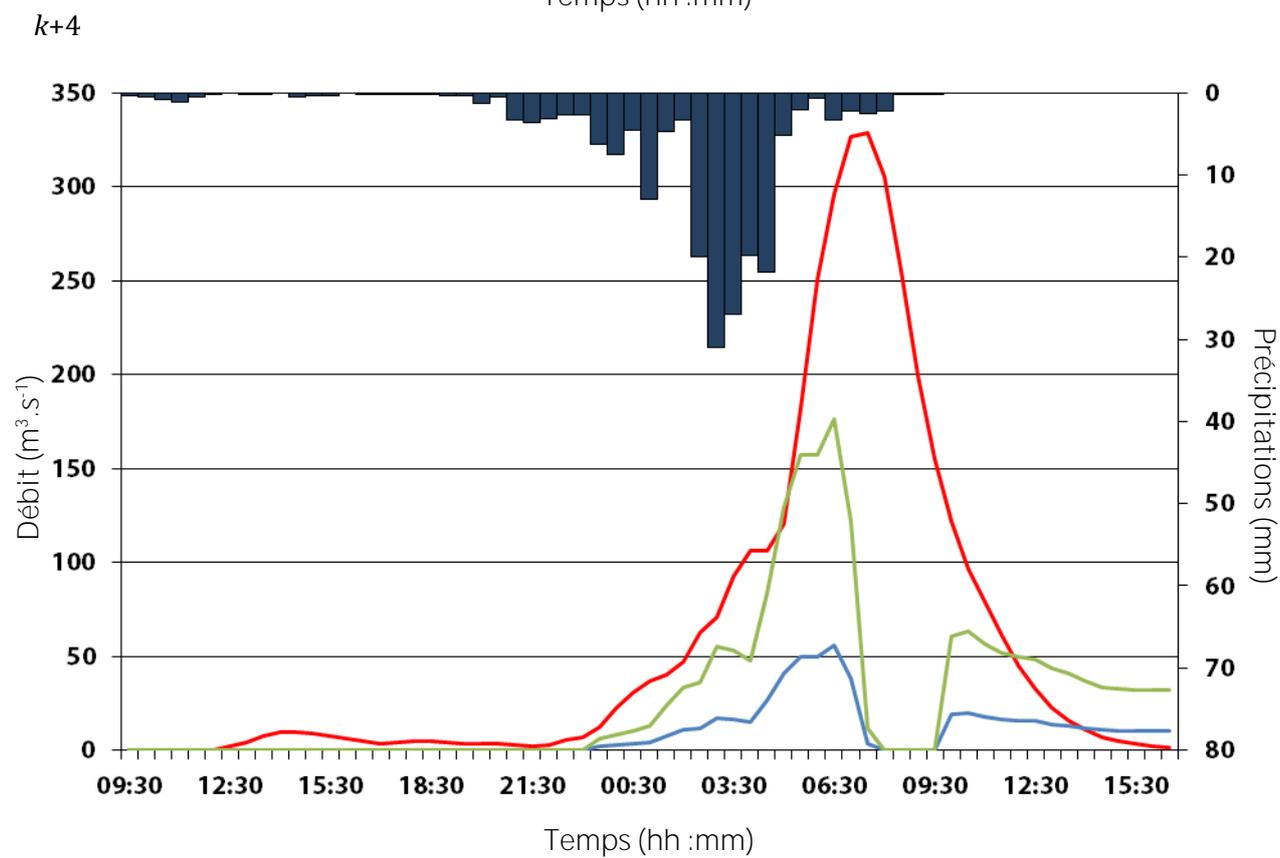
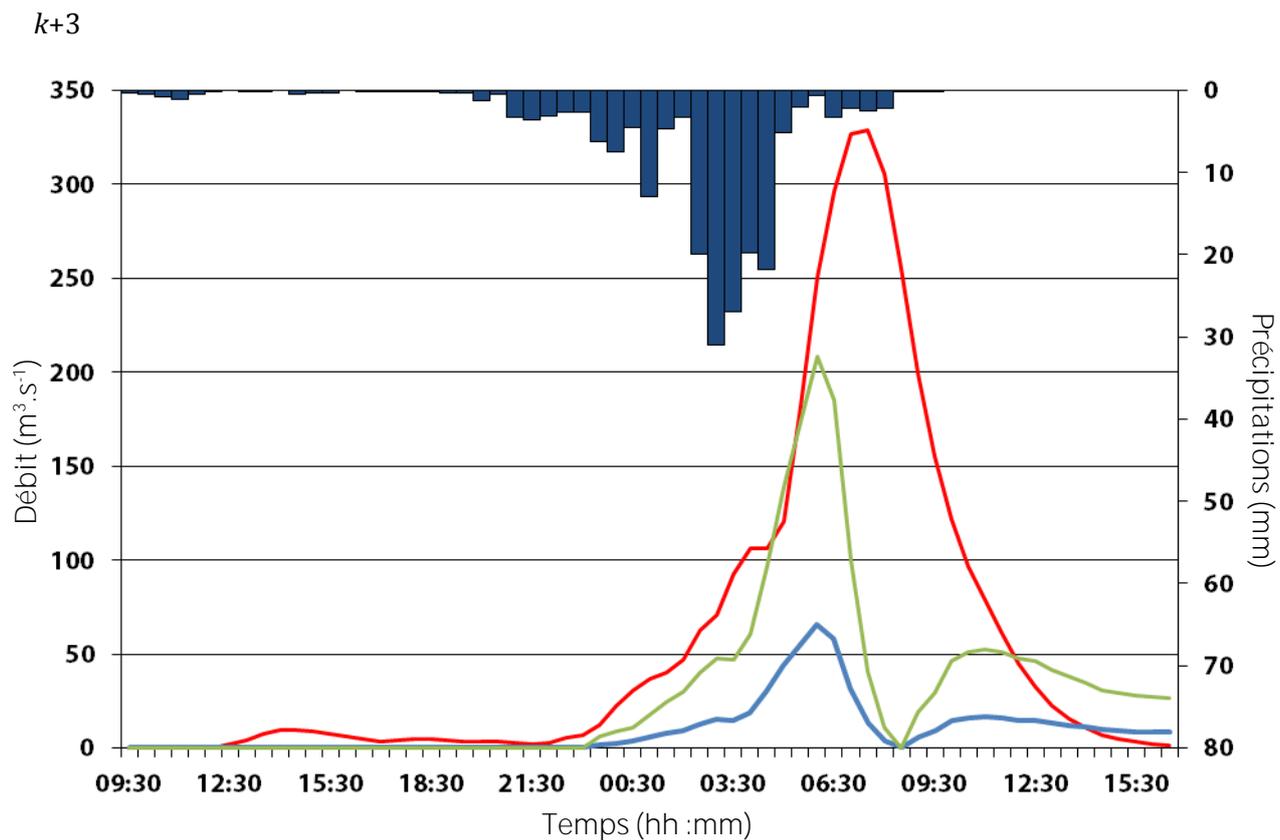
Bassin du Galeizon amont, évènement n°925 :

$k+1$



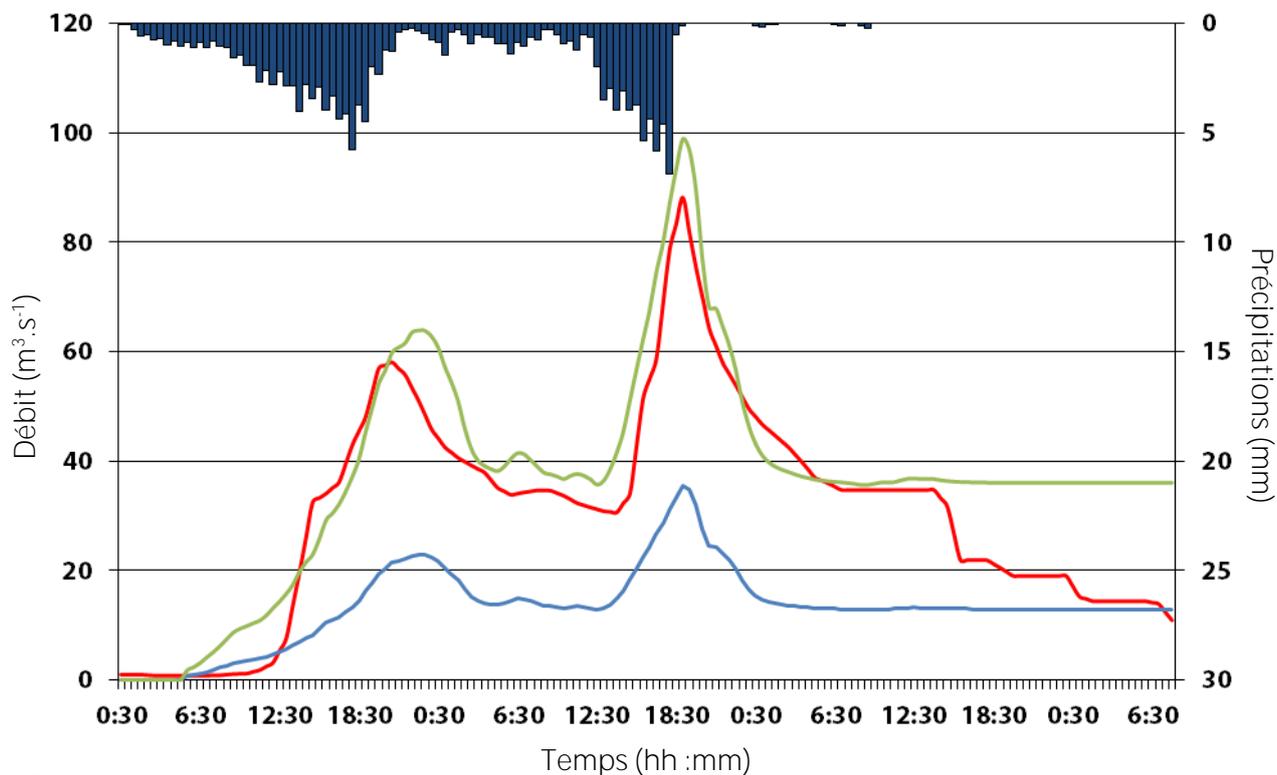
$k+2$



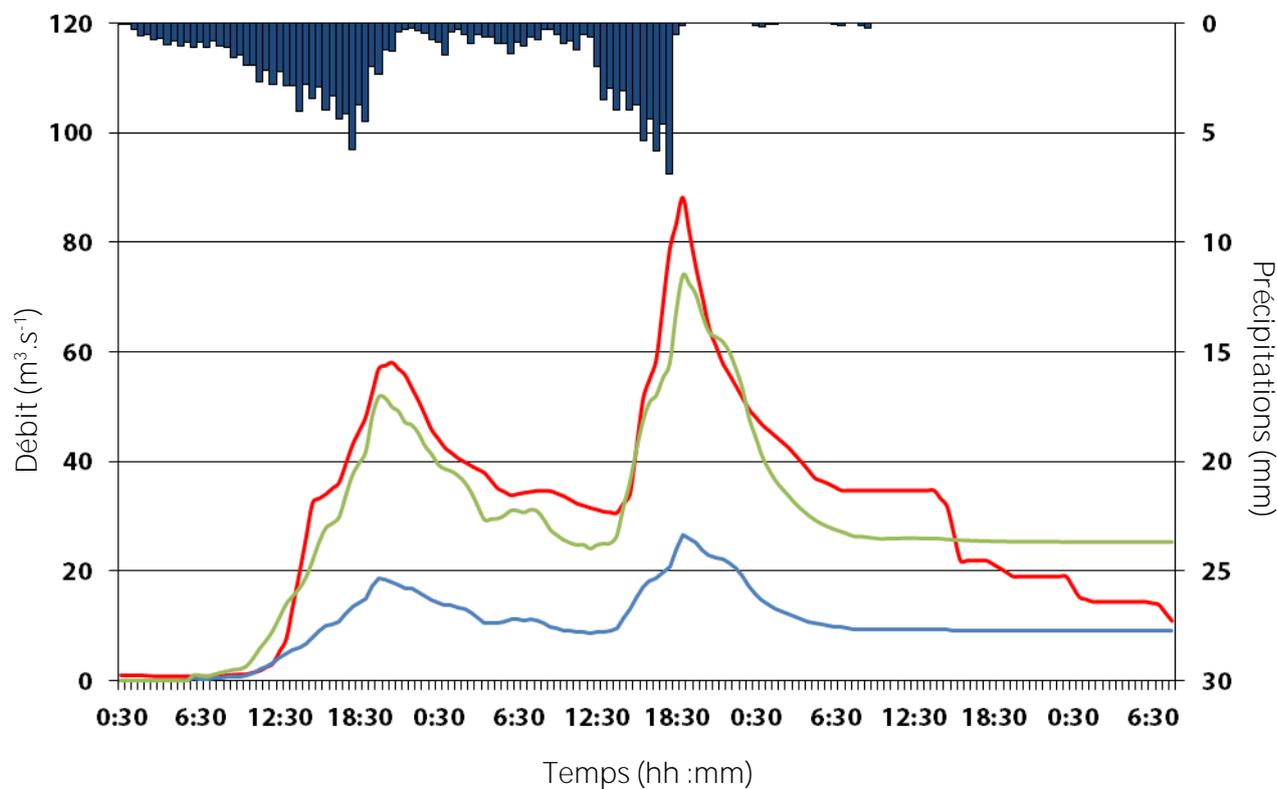


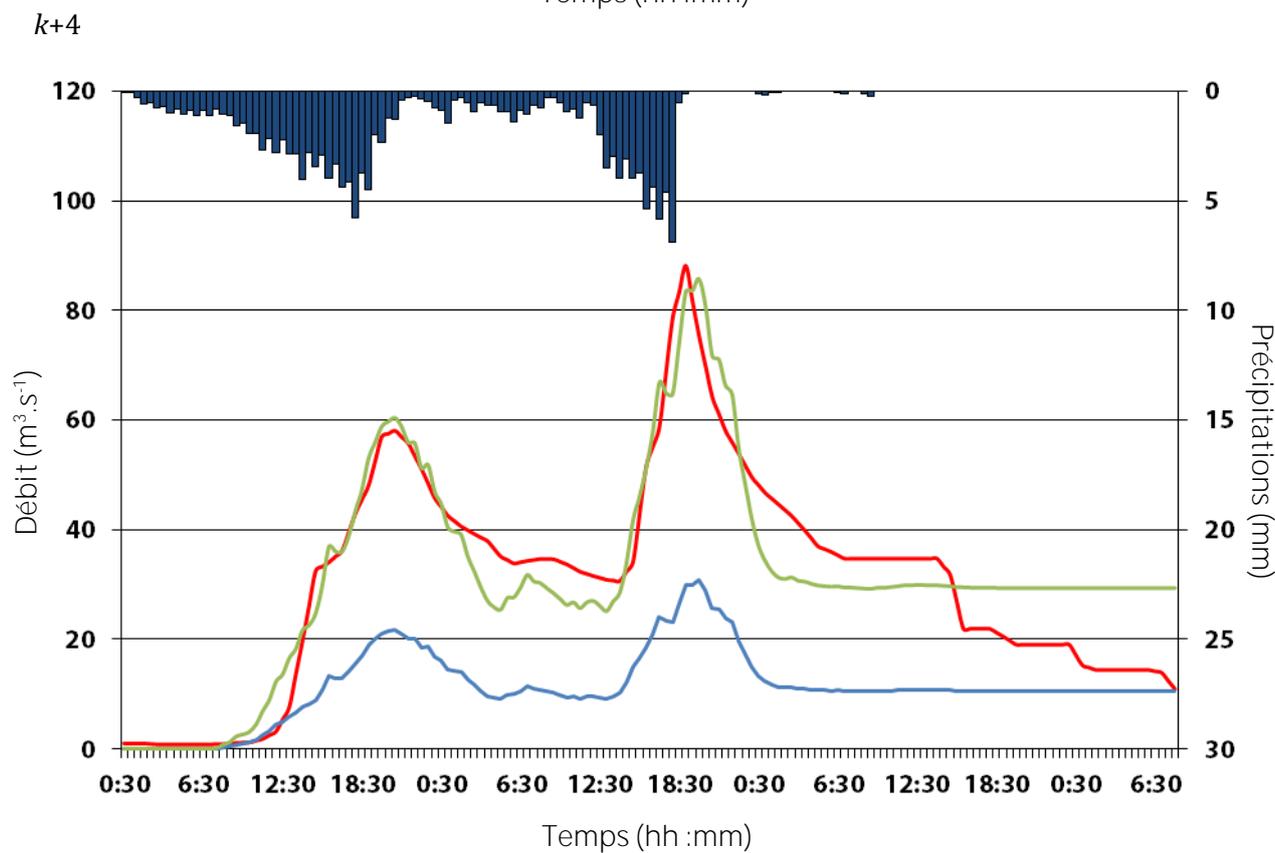
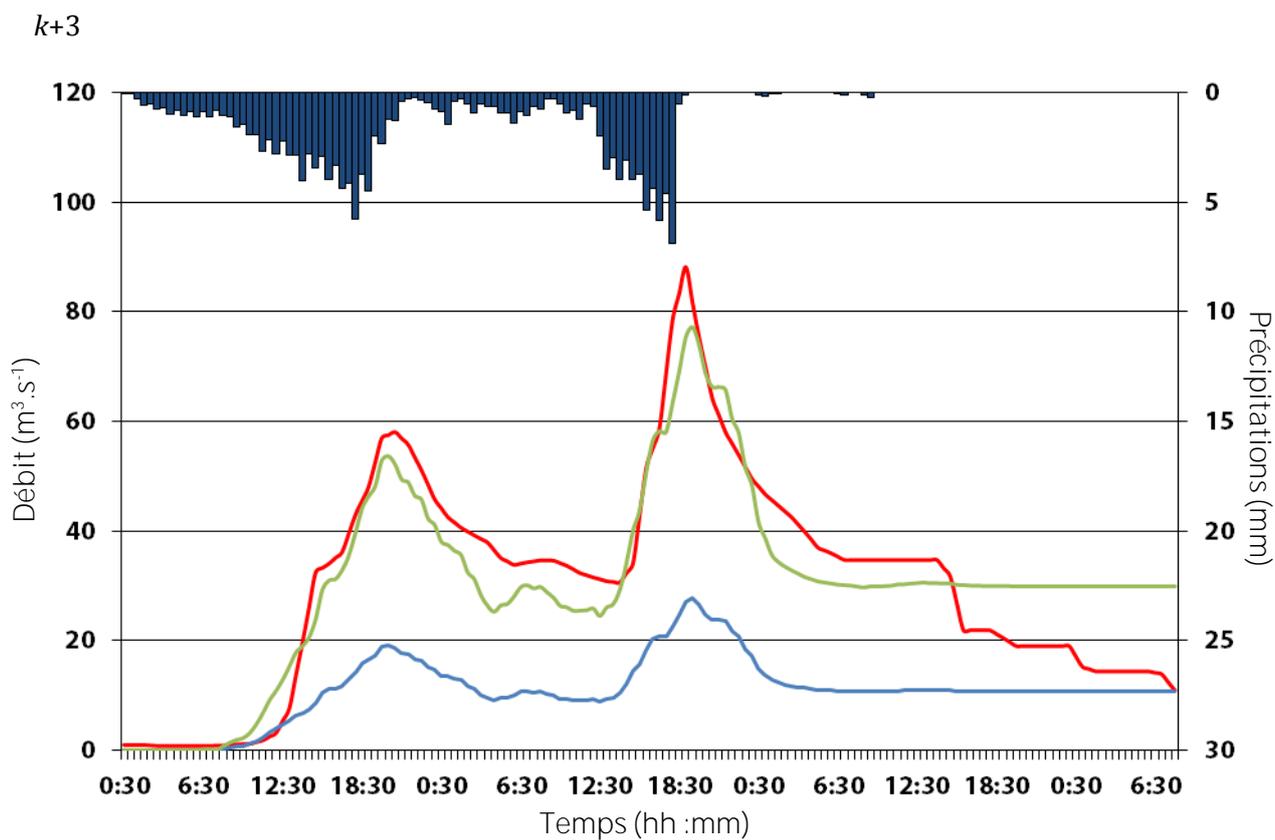
Bassin de la Ganière, évènement n°910 :

$k+1$



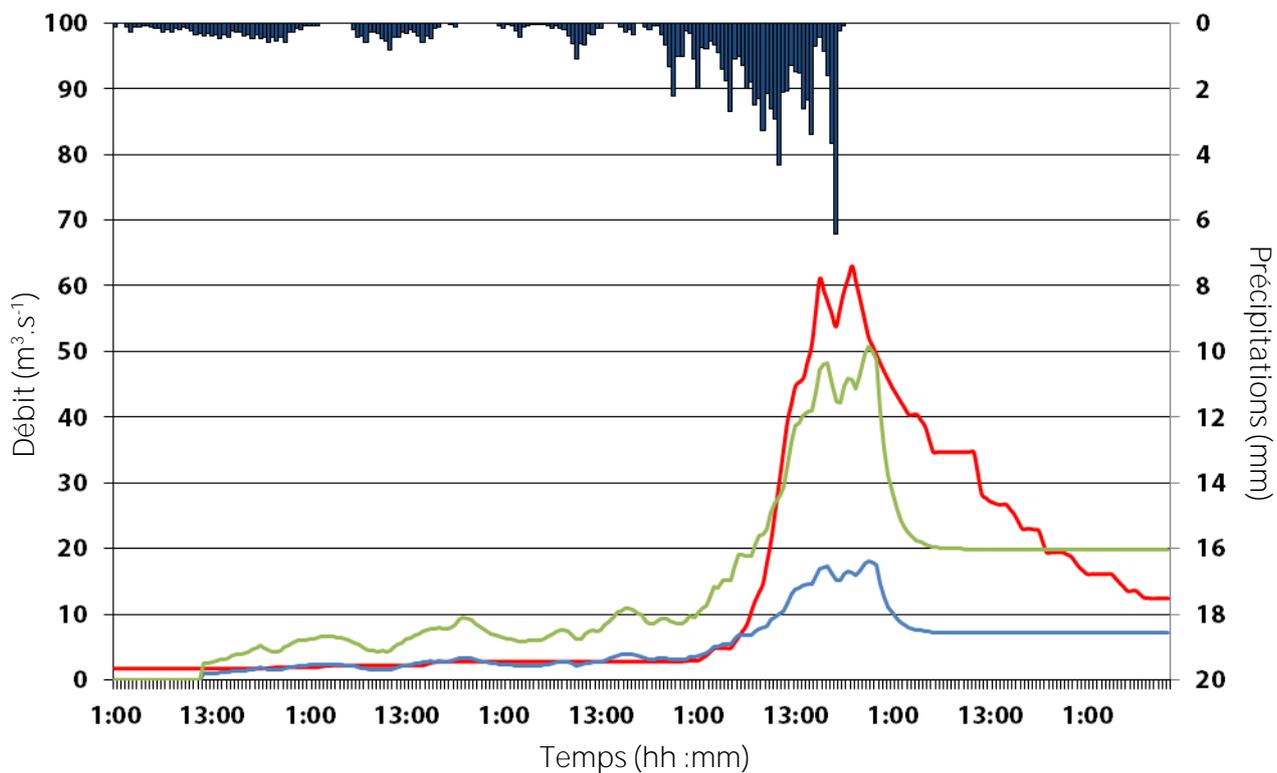
$k+2$



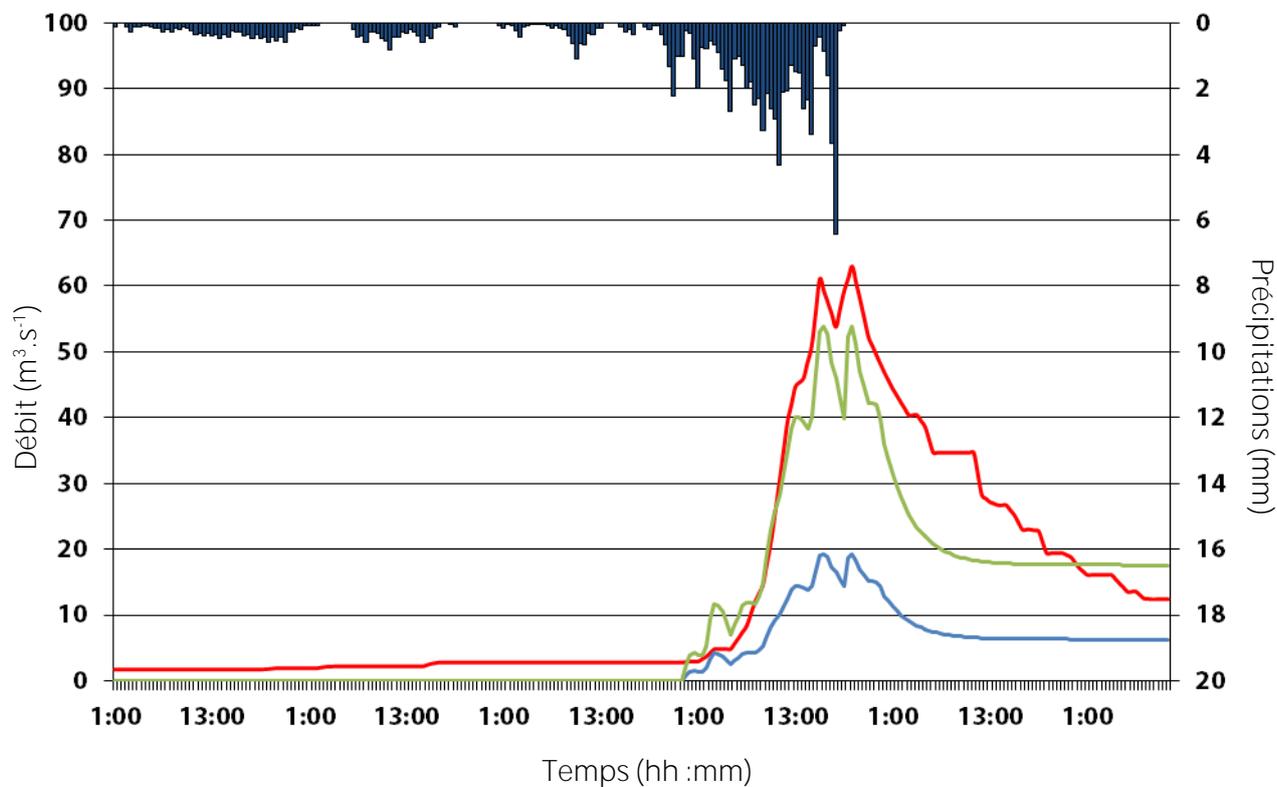


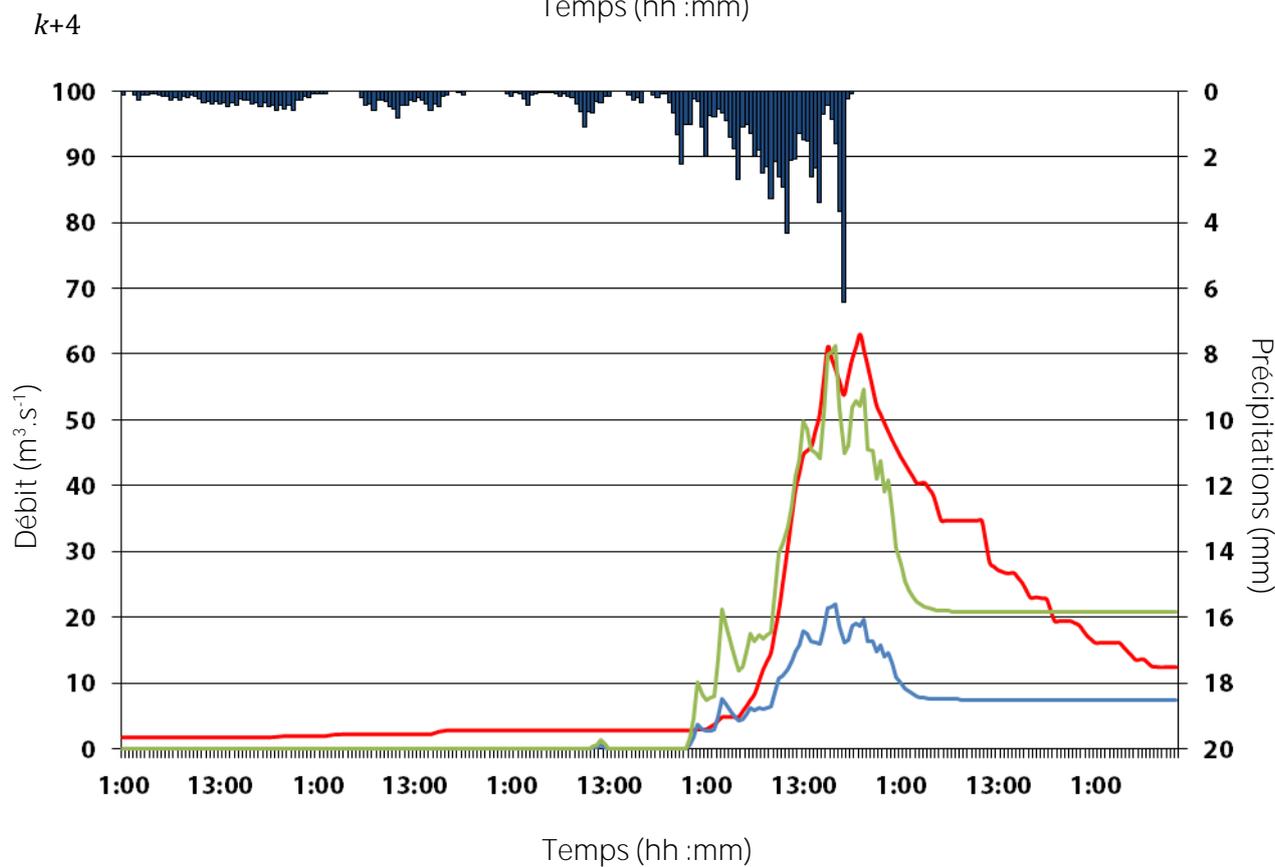
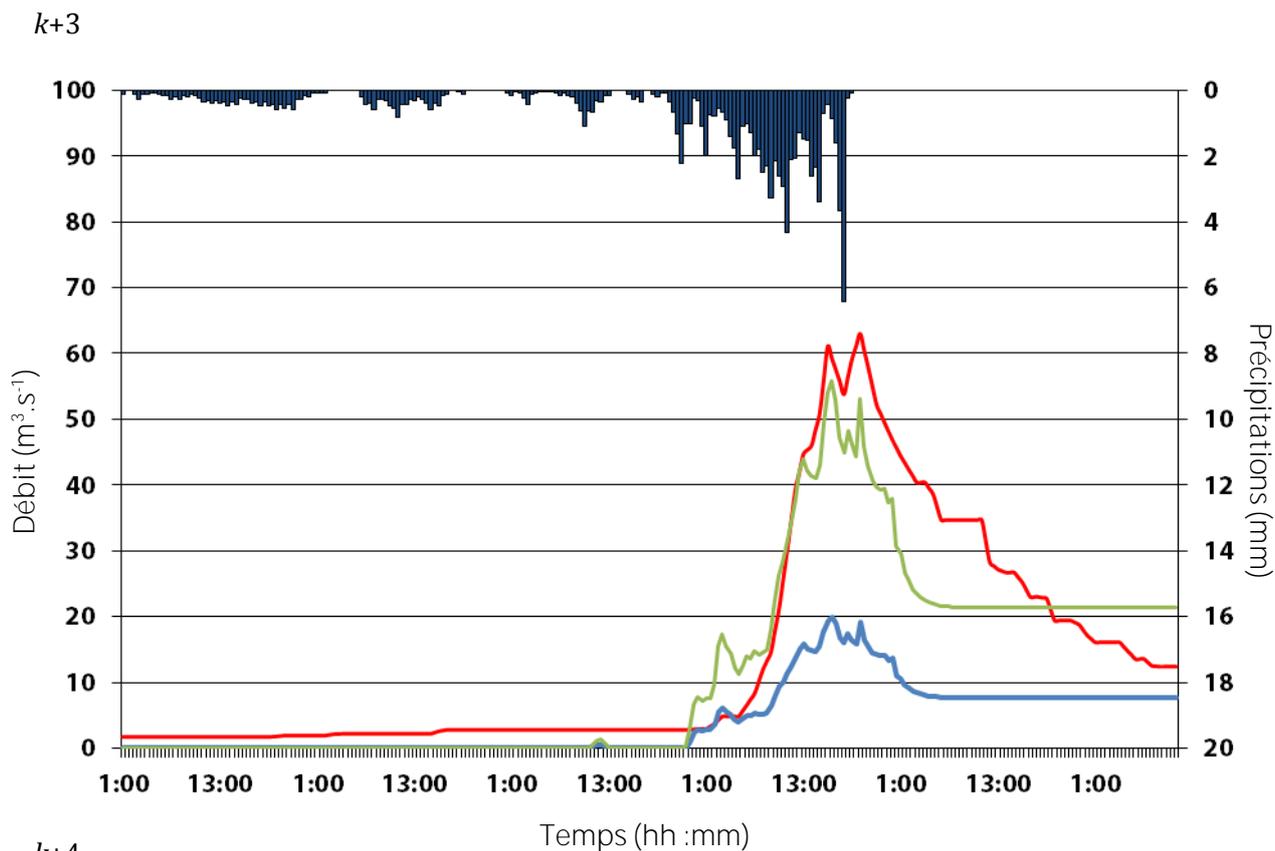
Bassin de la Ganière, évènement n°911 :

$k+1$



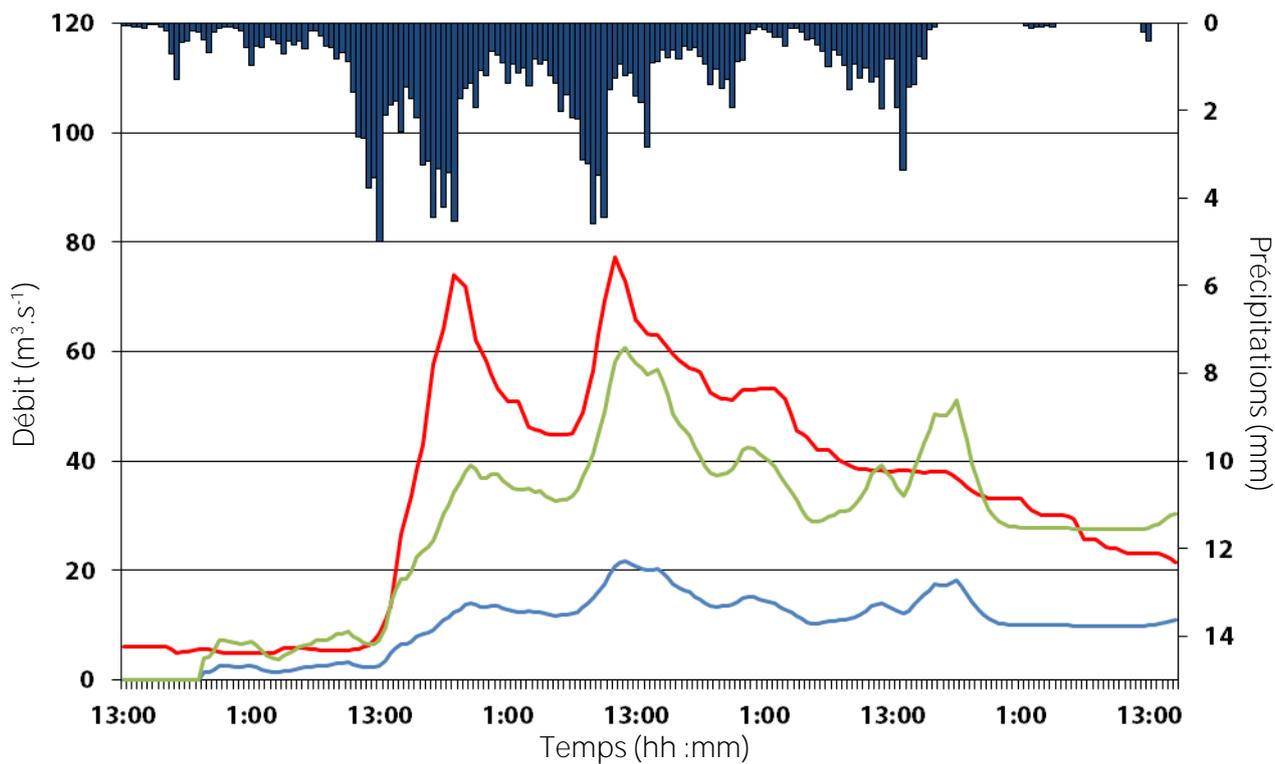
$k+2$



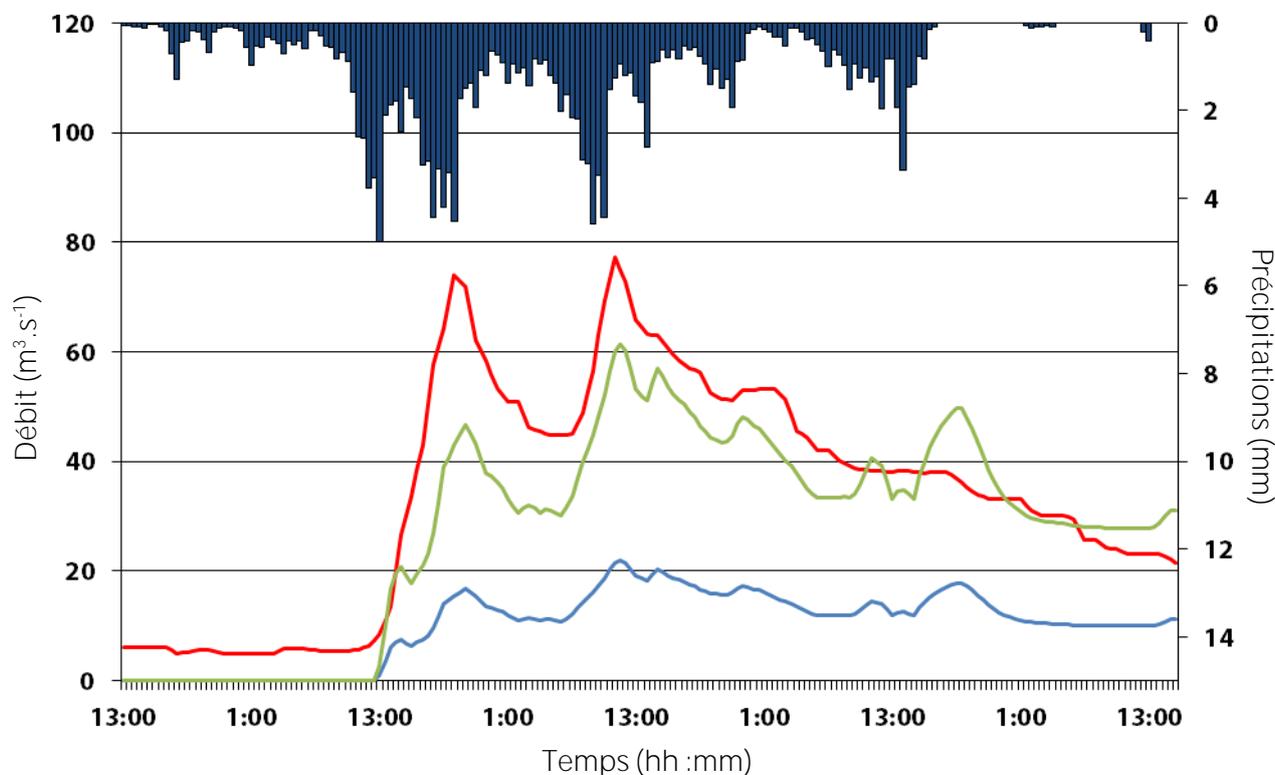


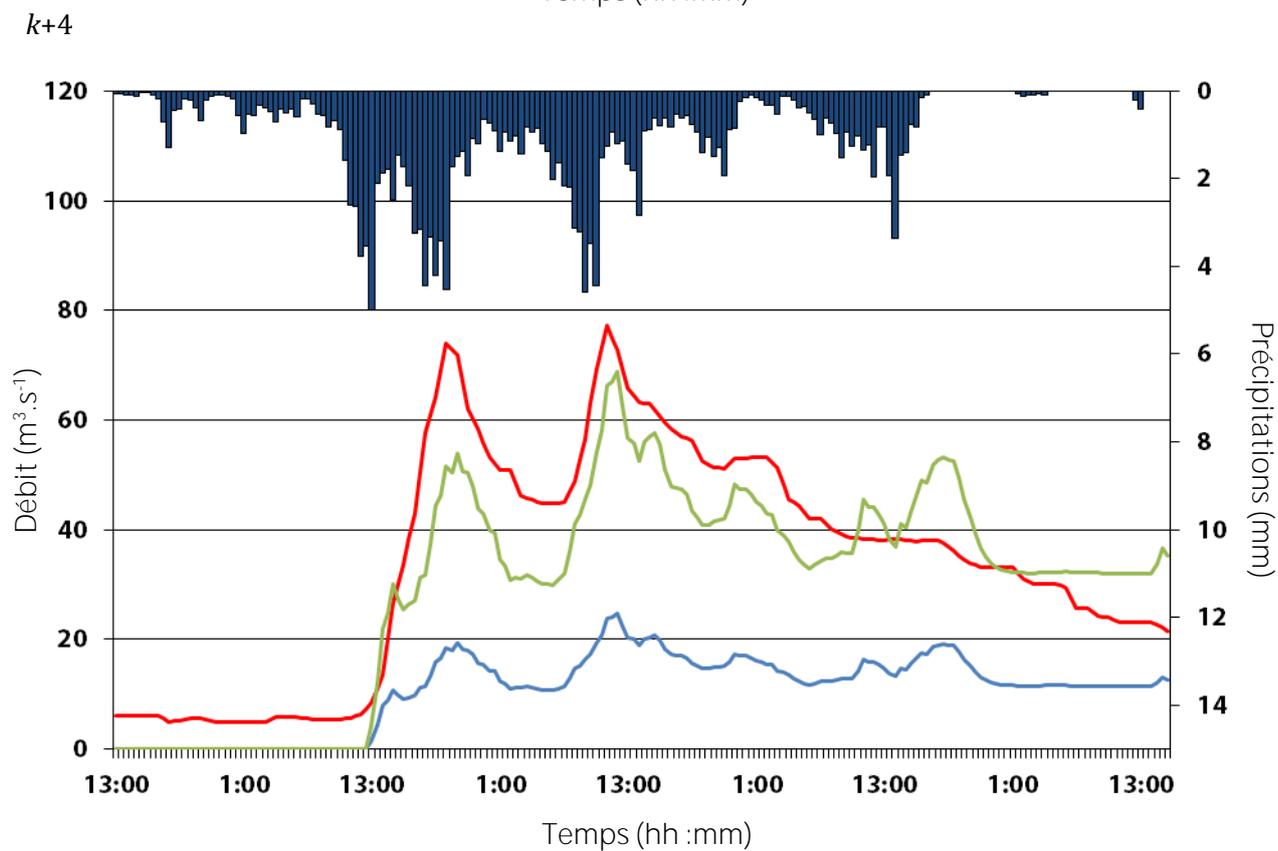
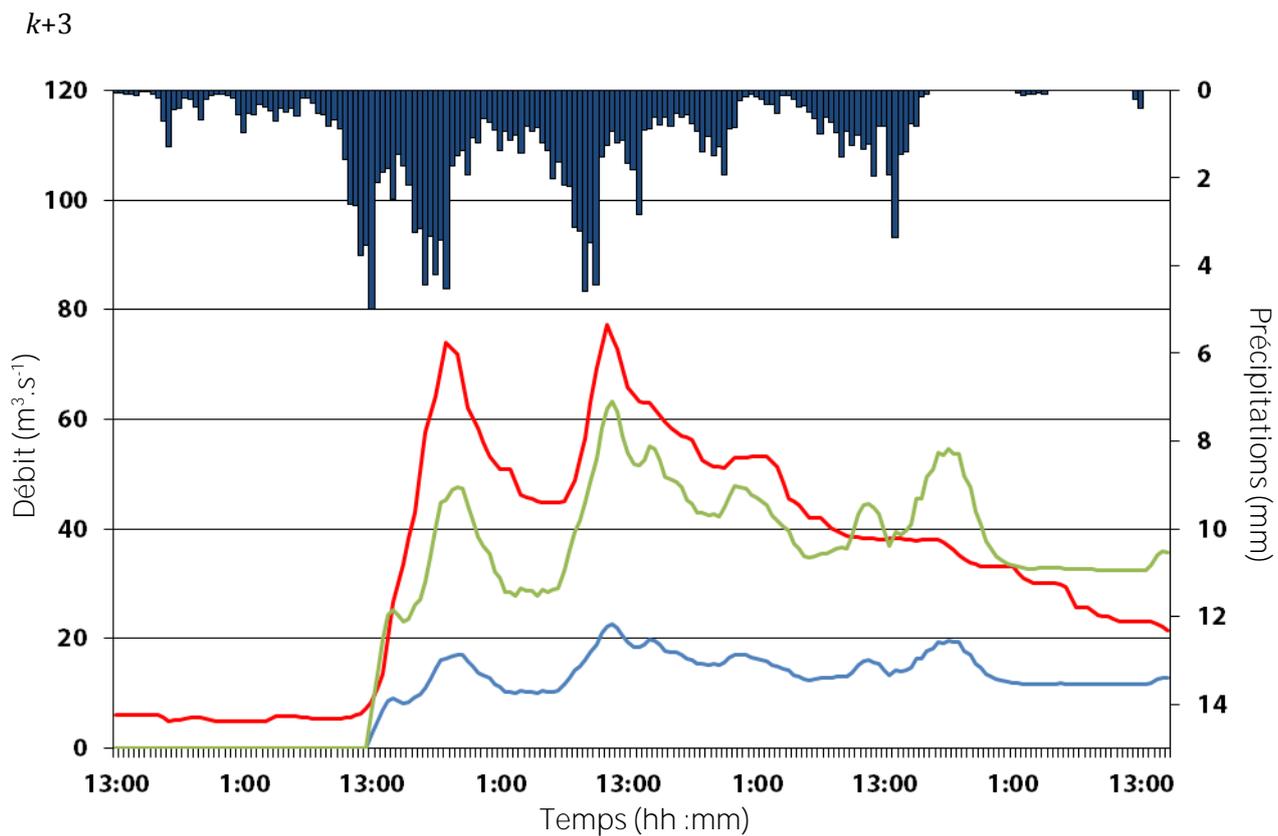
Bassin de la Ganière, évènement n°912 :

$k+1$



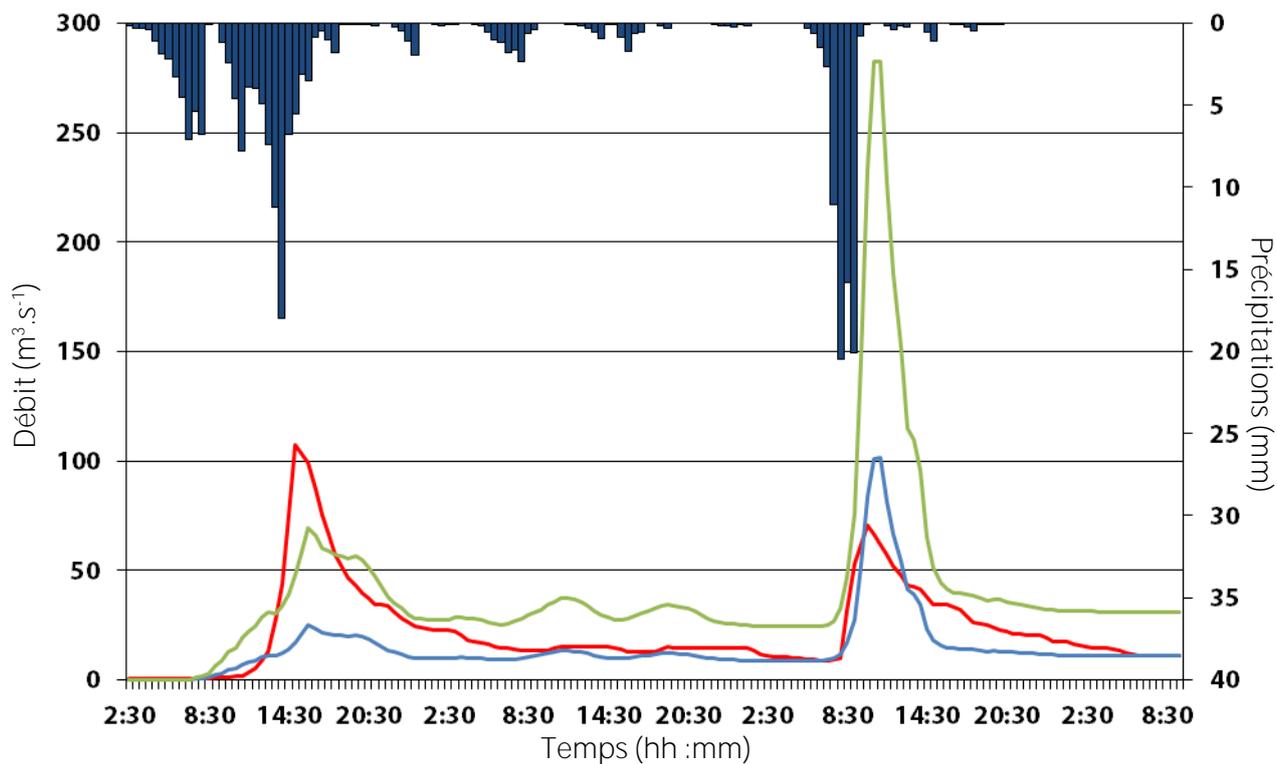
$k+2$



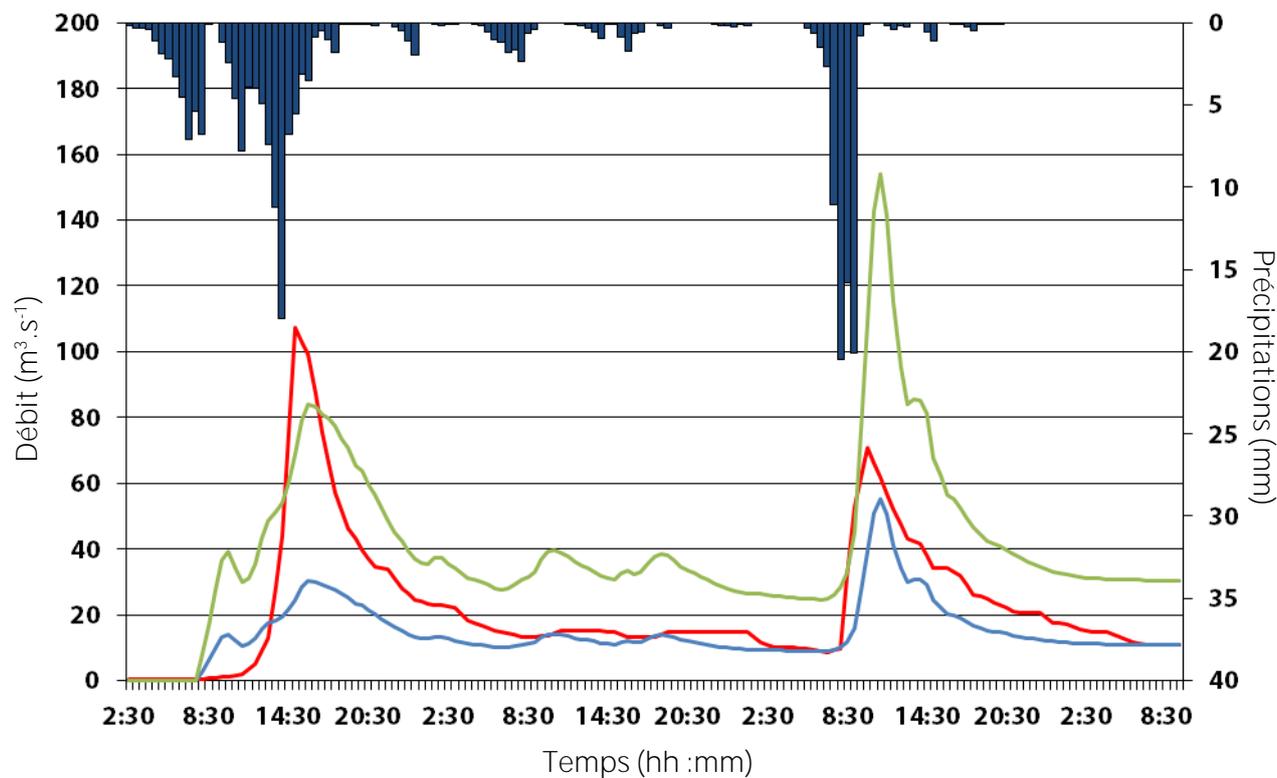


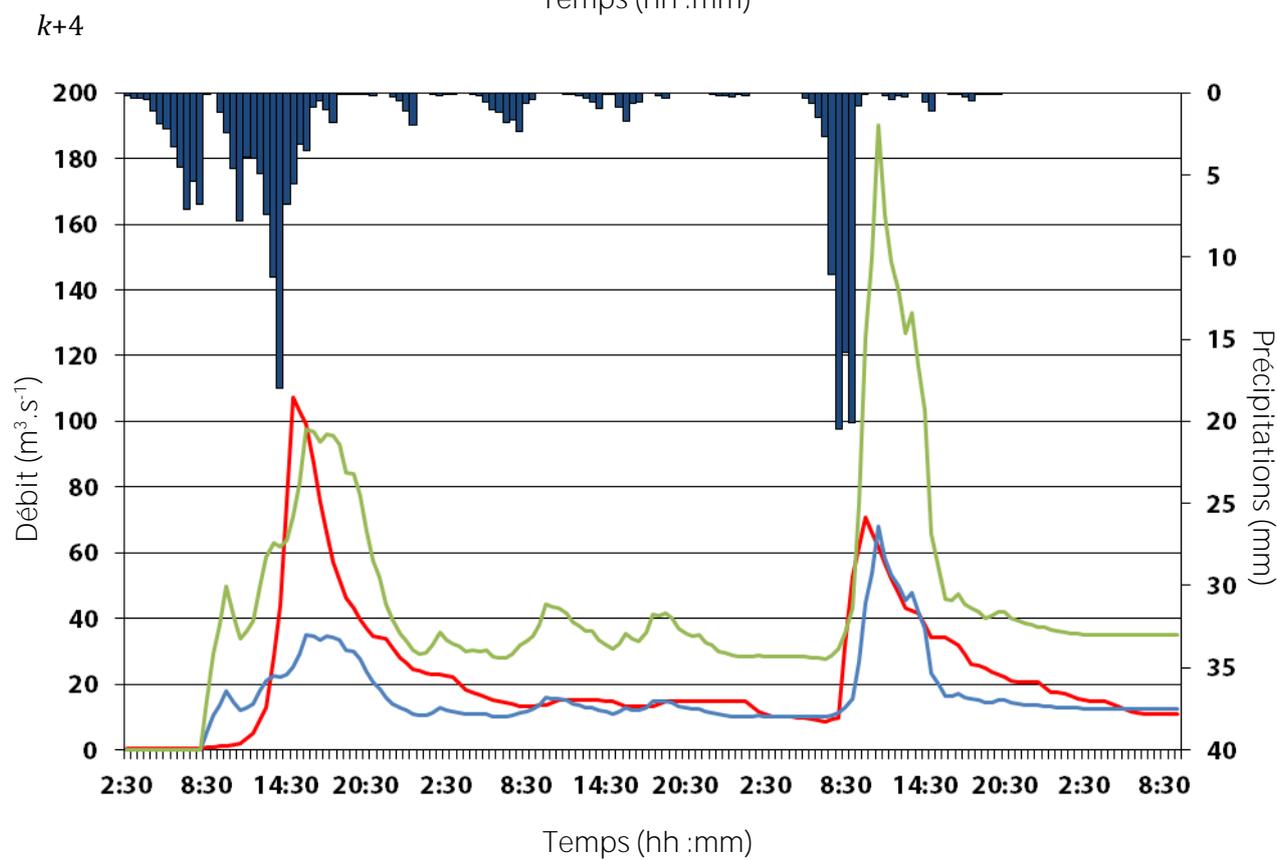
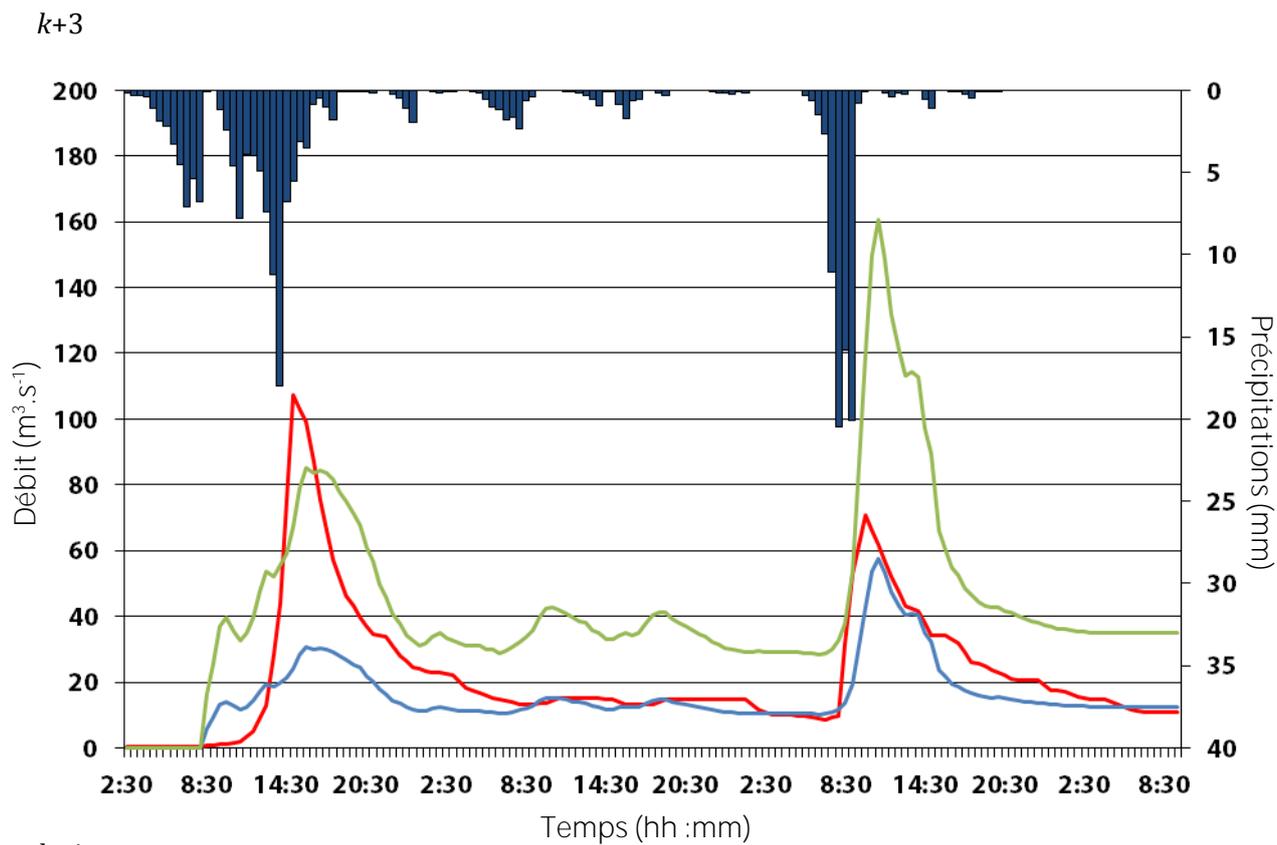
Bassin de la Ganière, évènement n°913 :

$k+1$



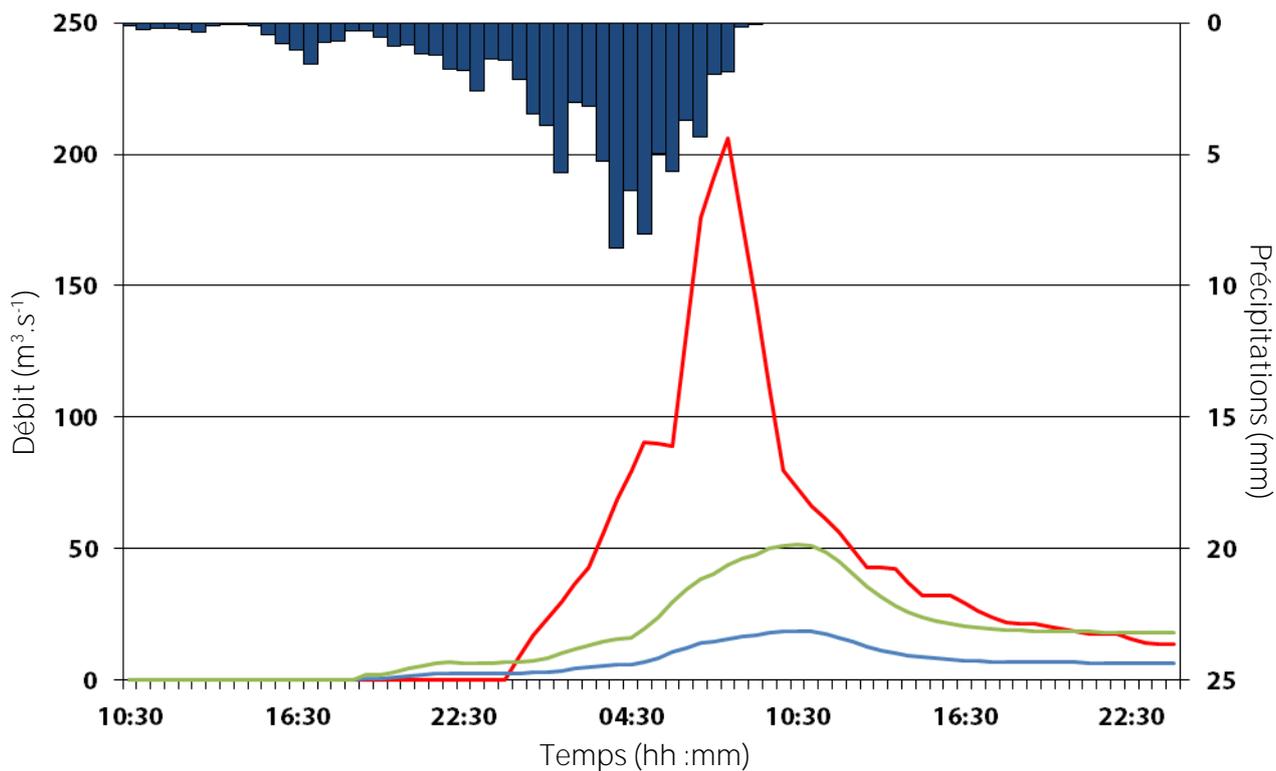
$k+2$



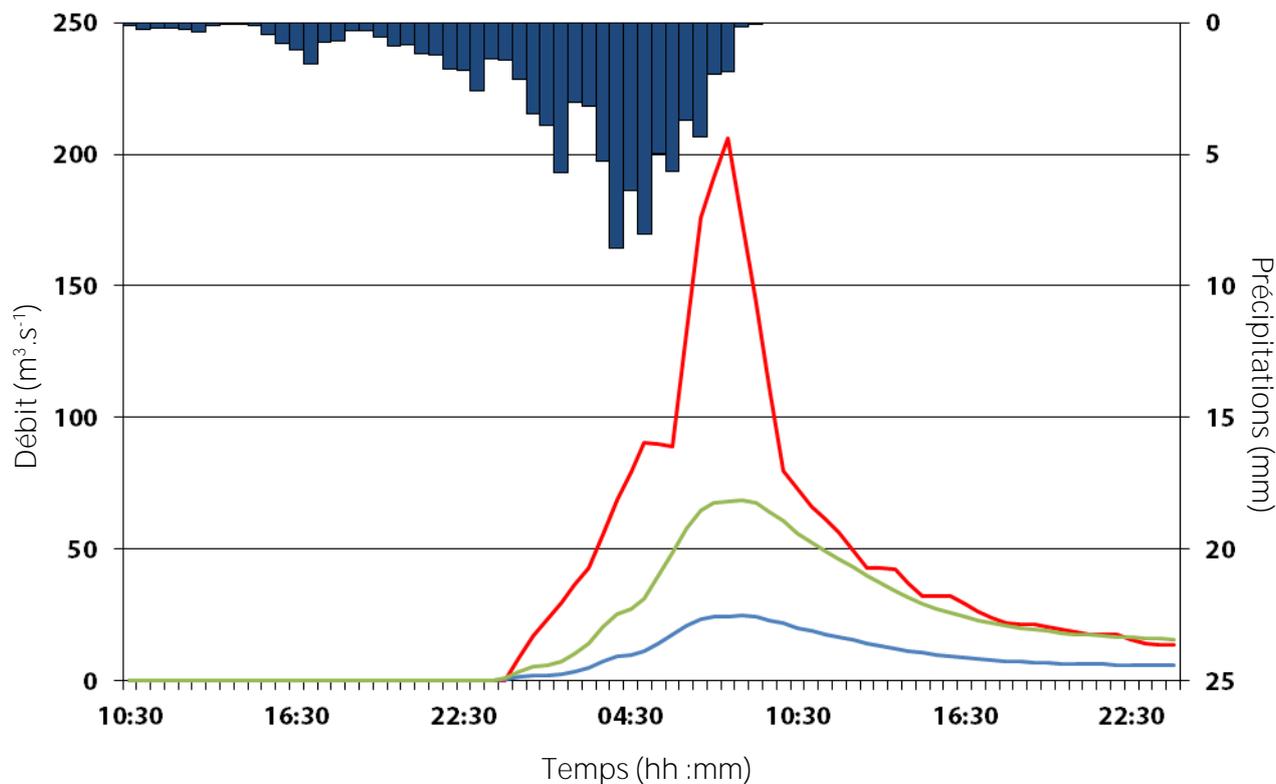


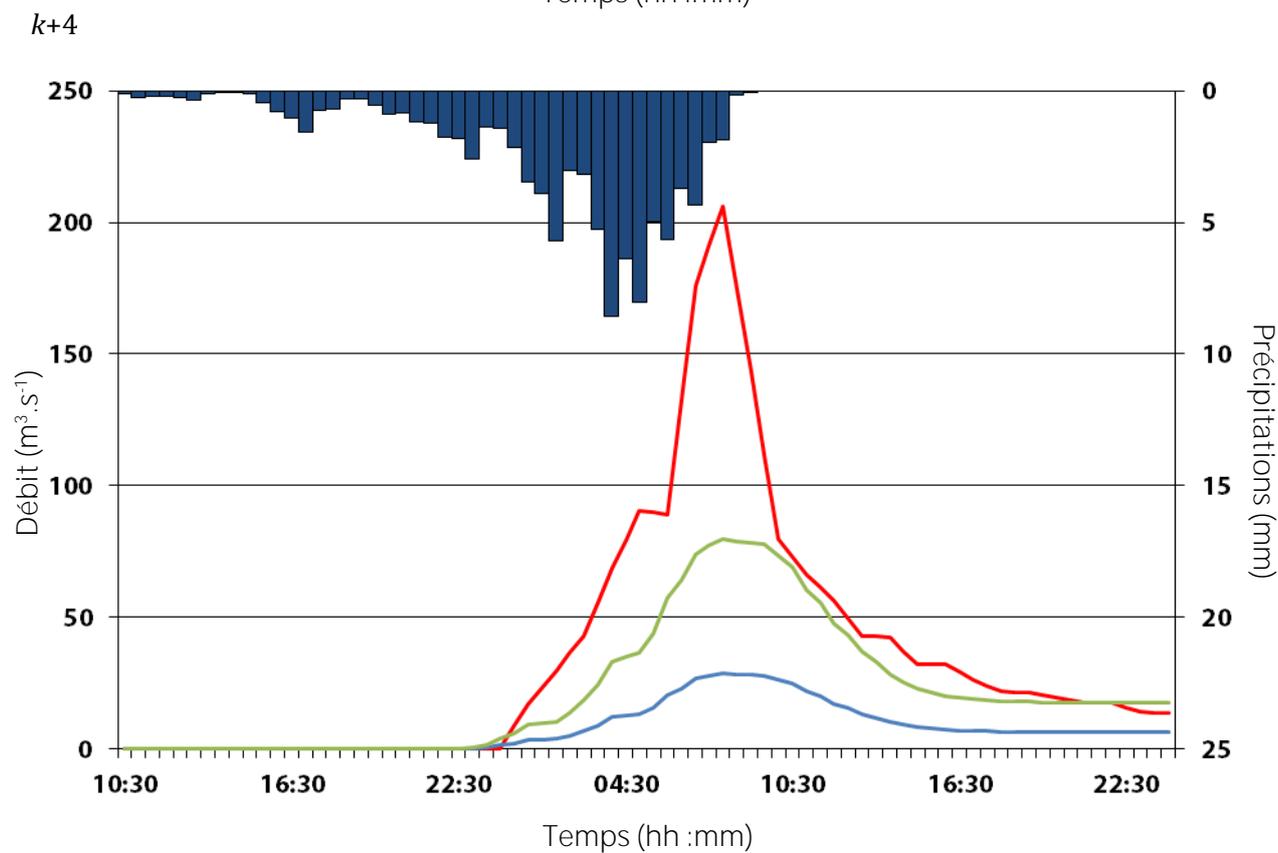
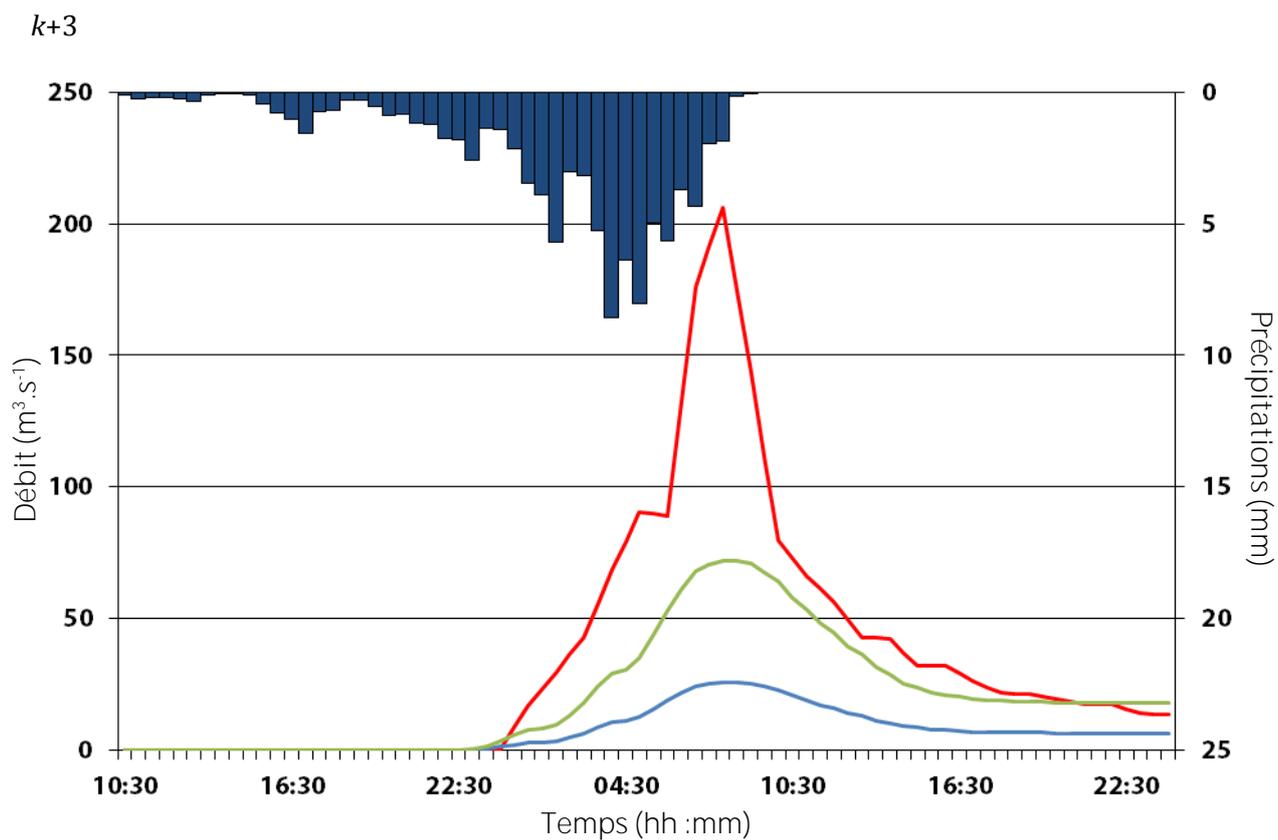
Bassin de la Ganière, évènement n°914 :

k+1



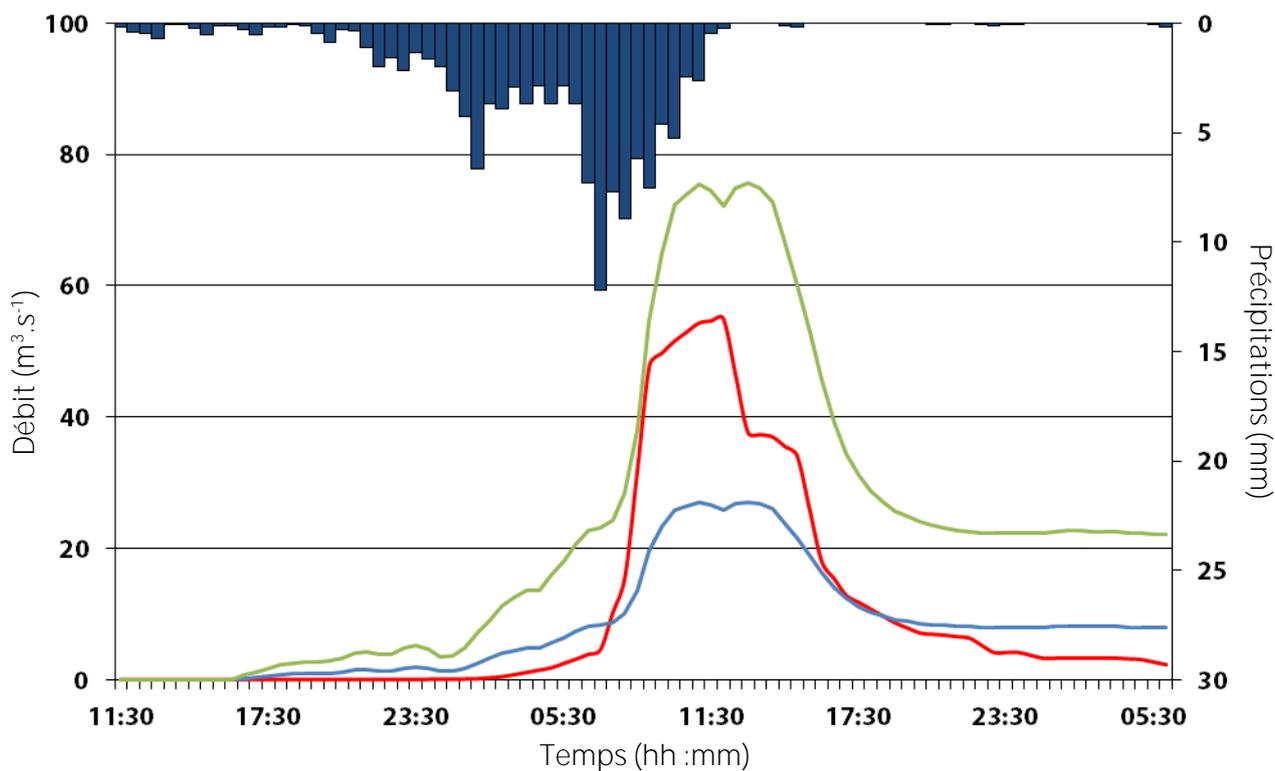
k+2



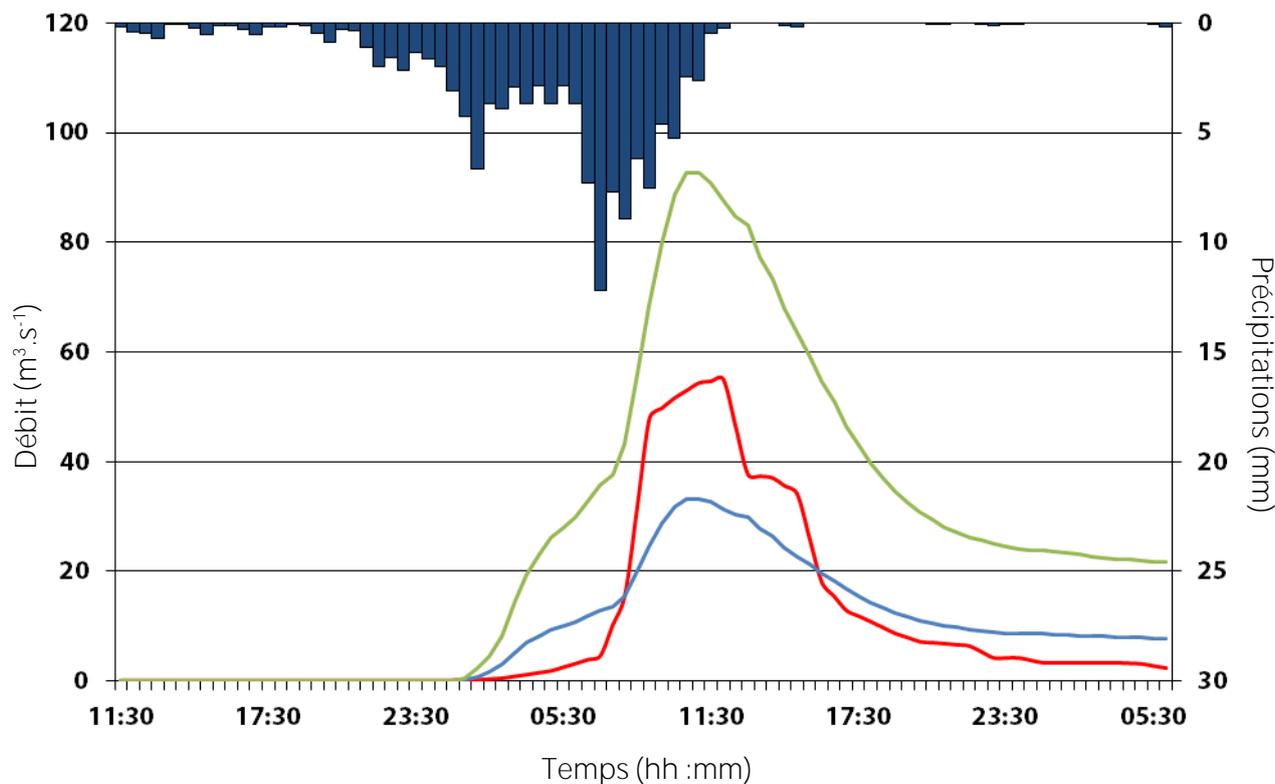


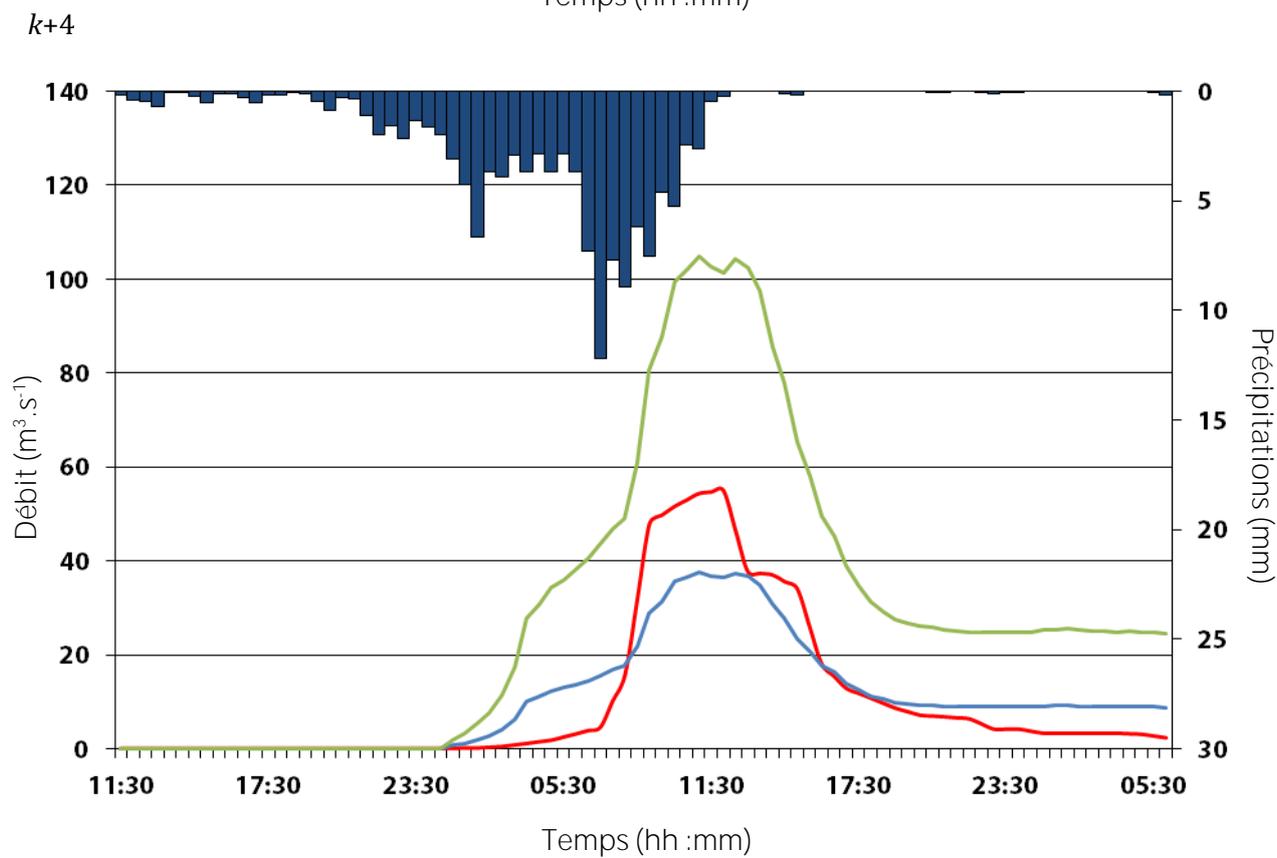
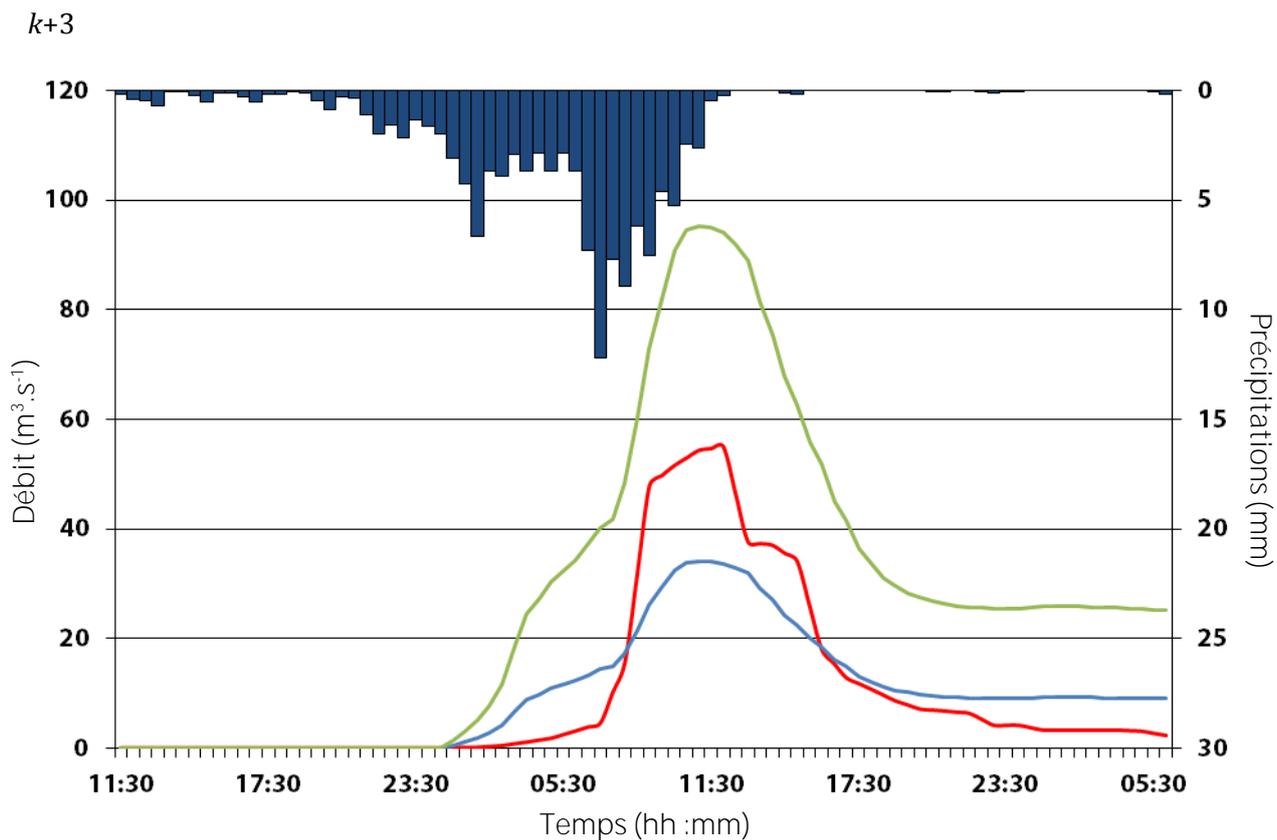
Bassin de la Ganière, évènement n°915 :

k+1



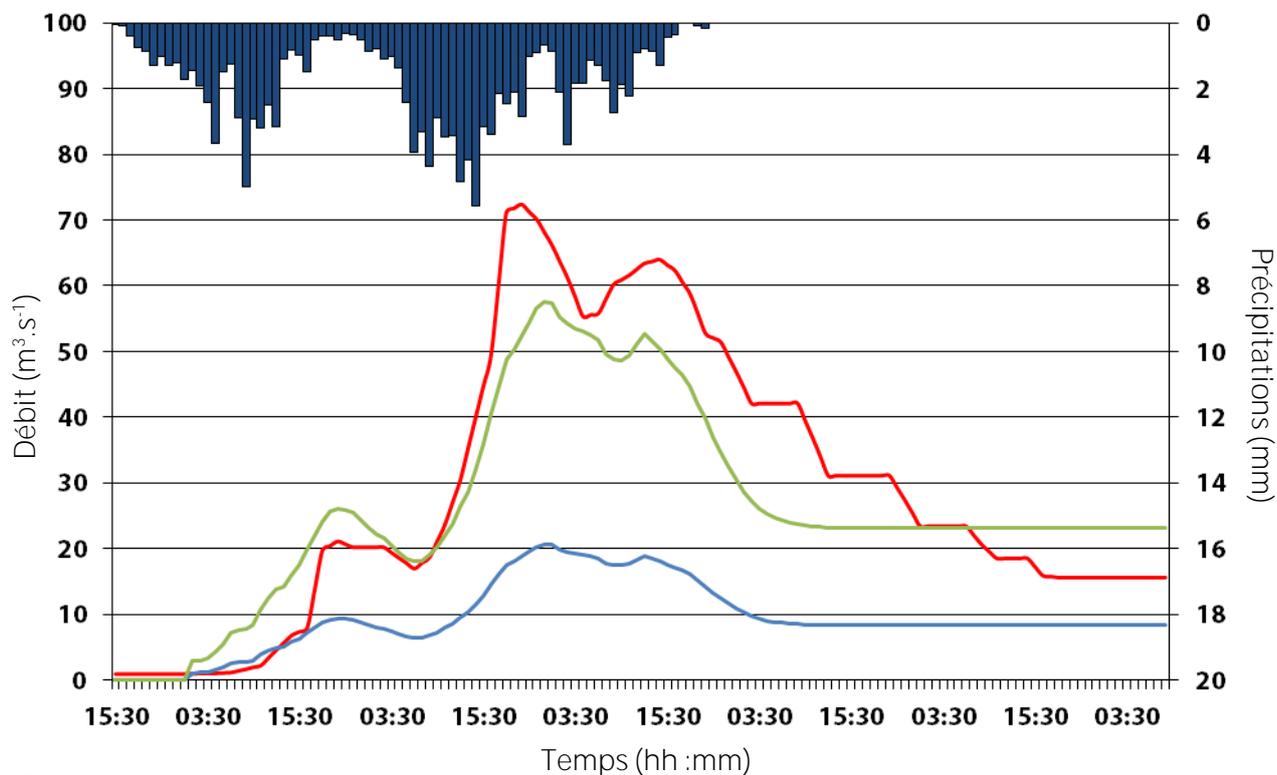
k+2



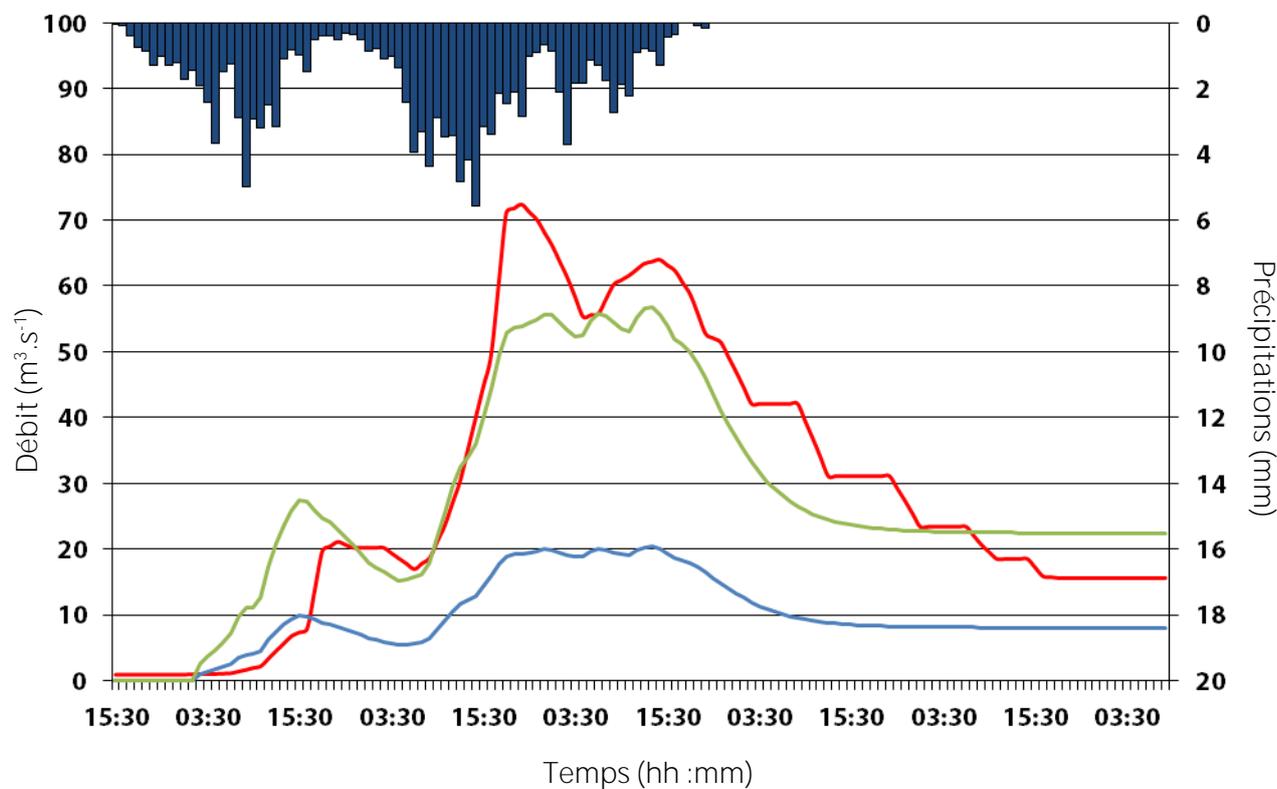


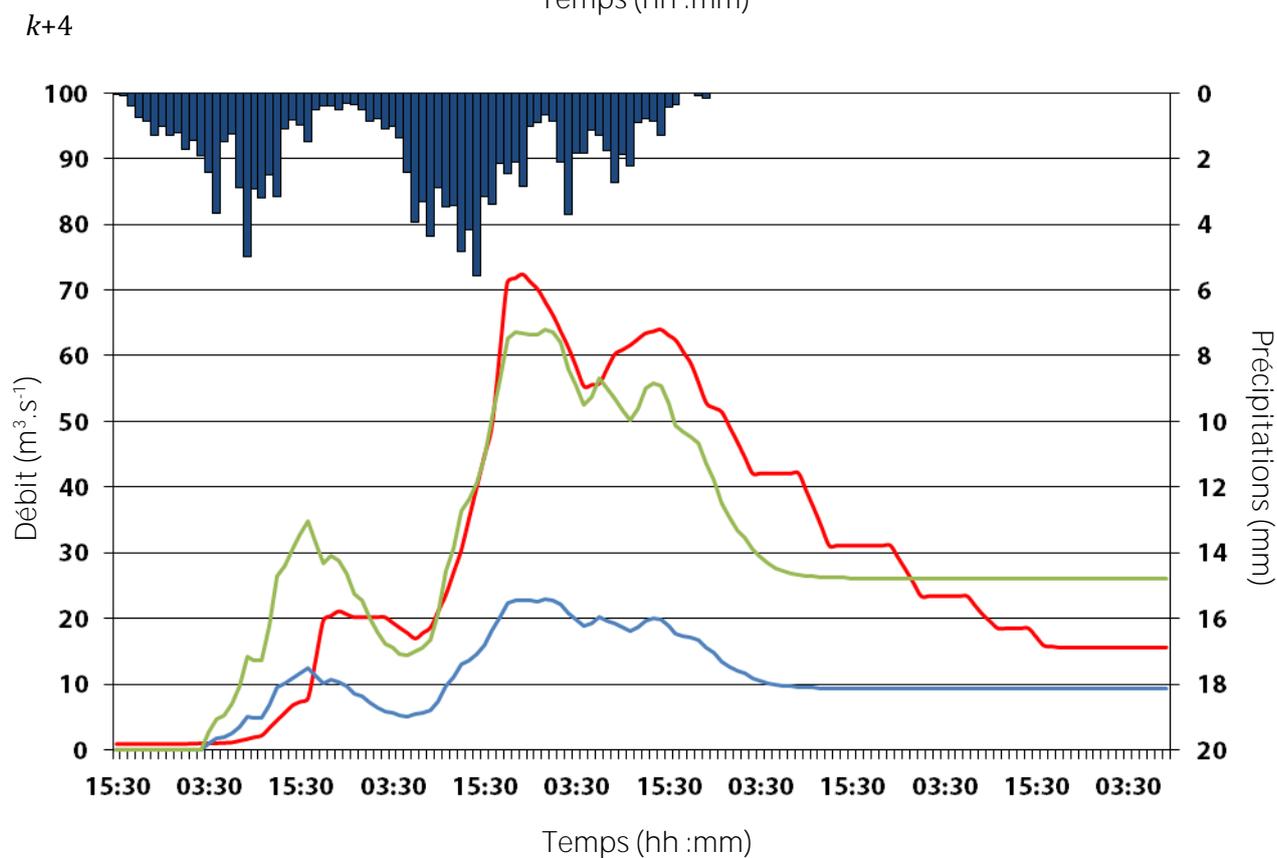
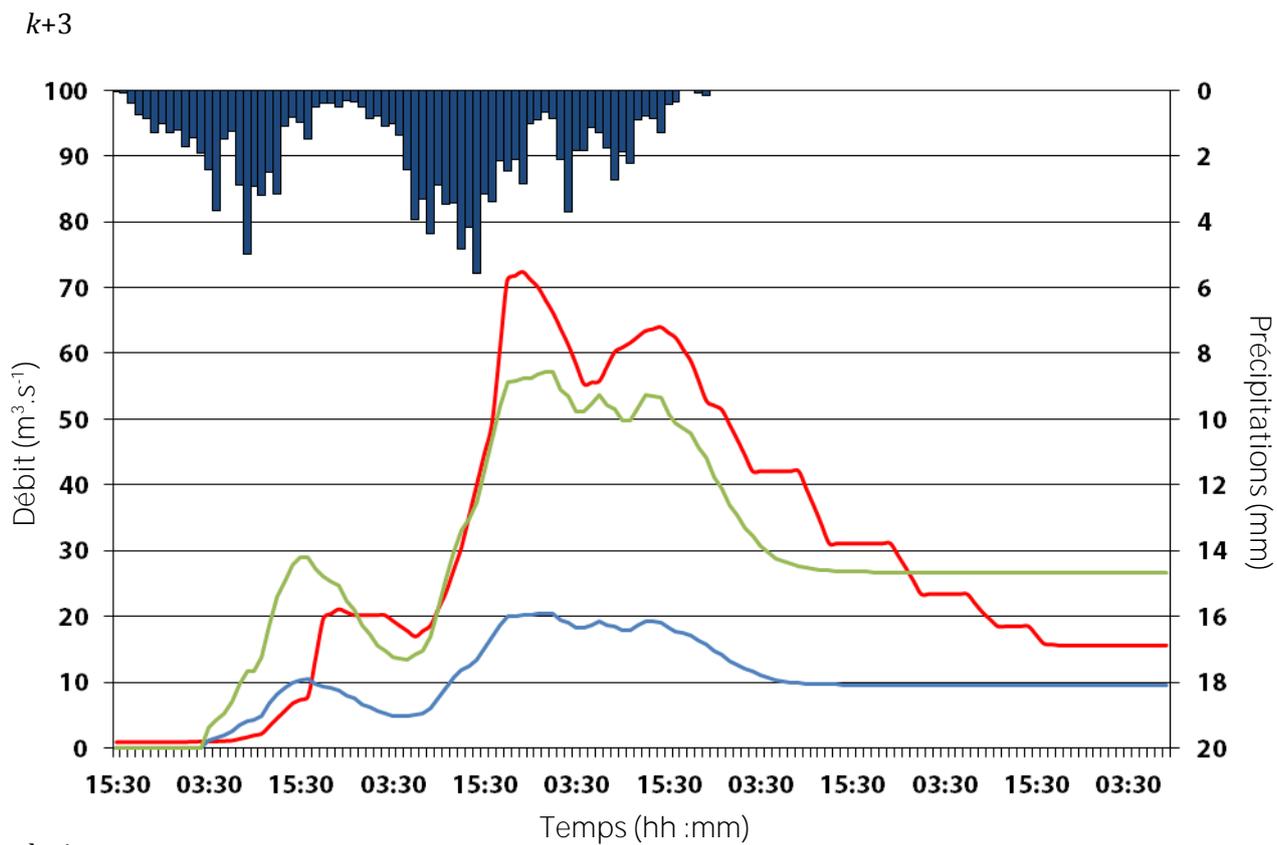
Bassin de la Ganière, évènement n°916 :

k+1



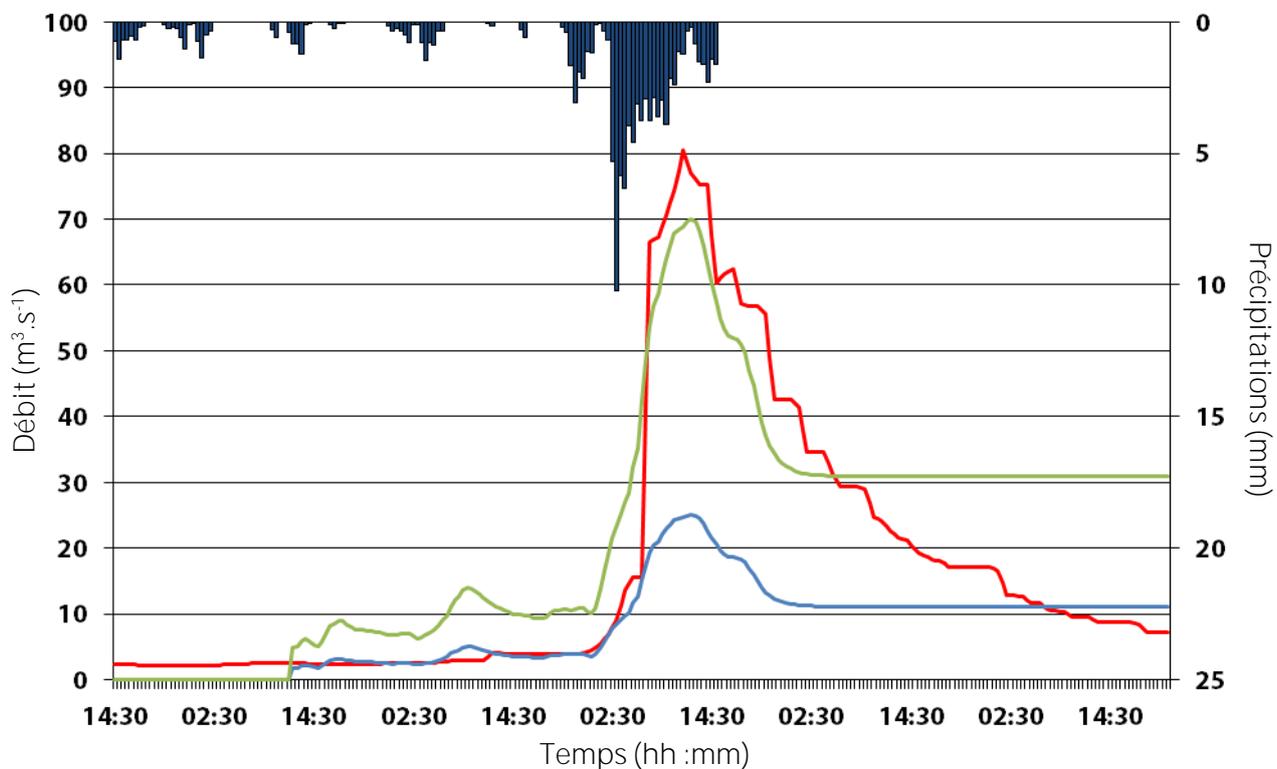
k+2



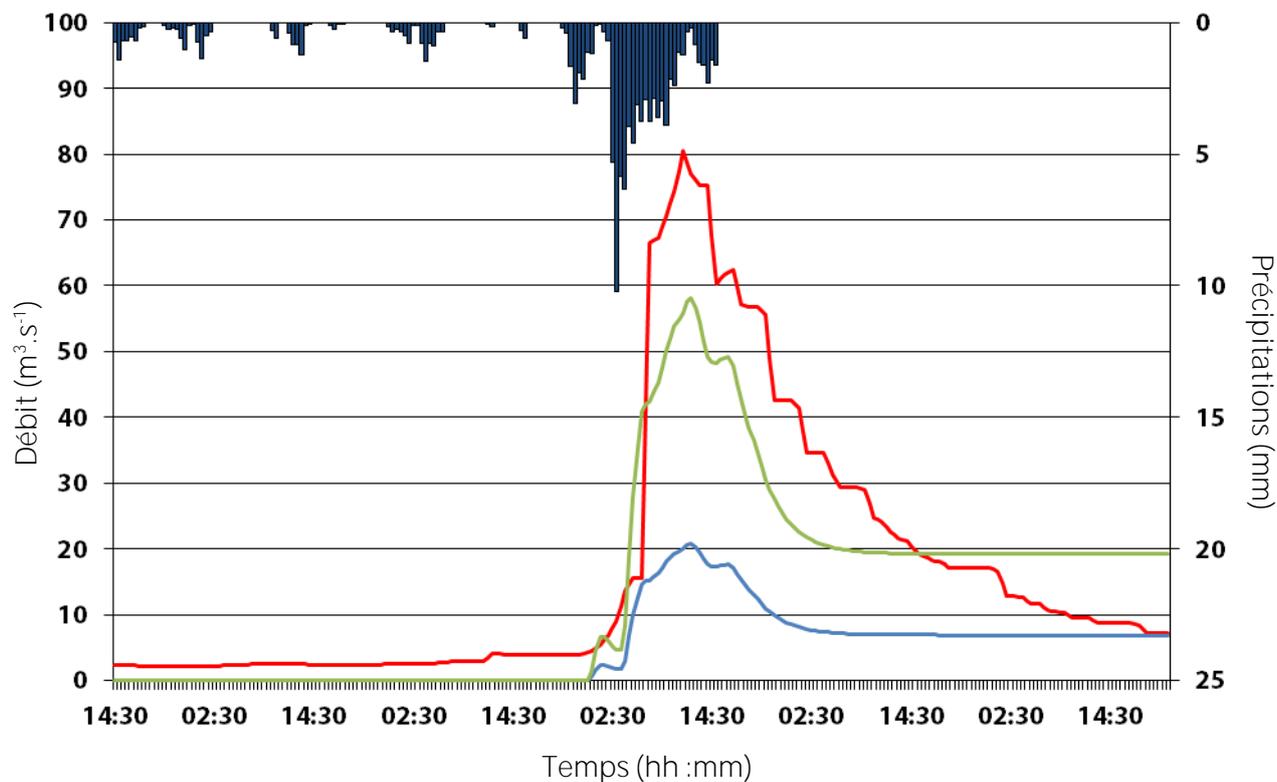


Bassin de la Ganière, évènement n°917 :

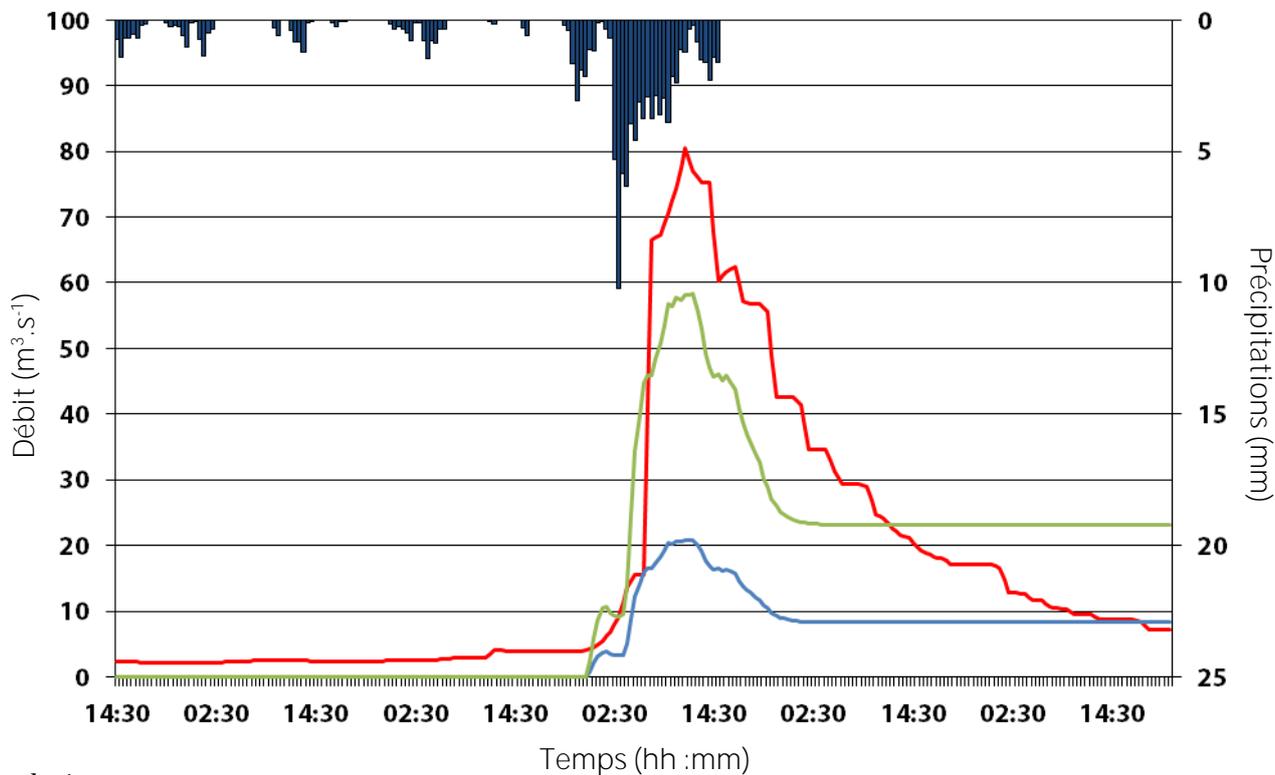
k+1



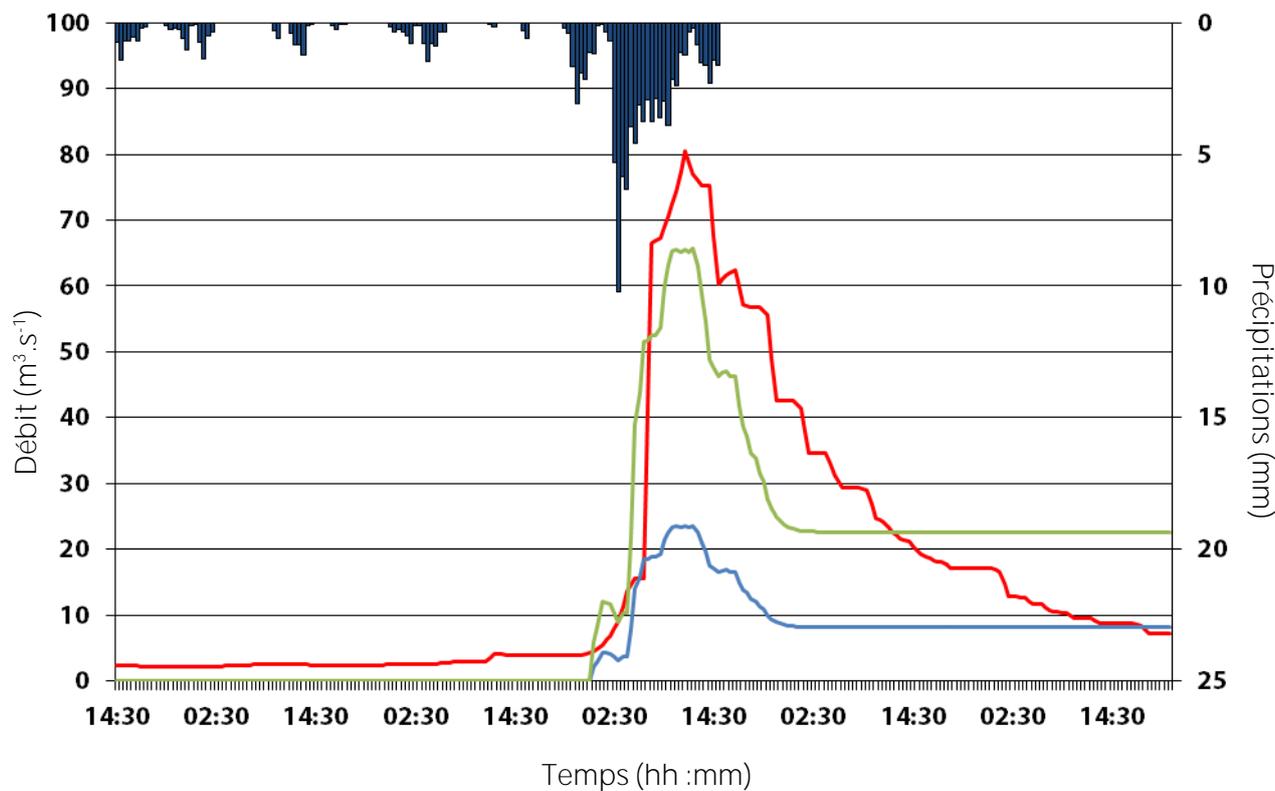
k+2



k+3

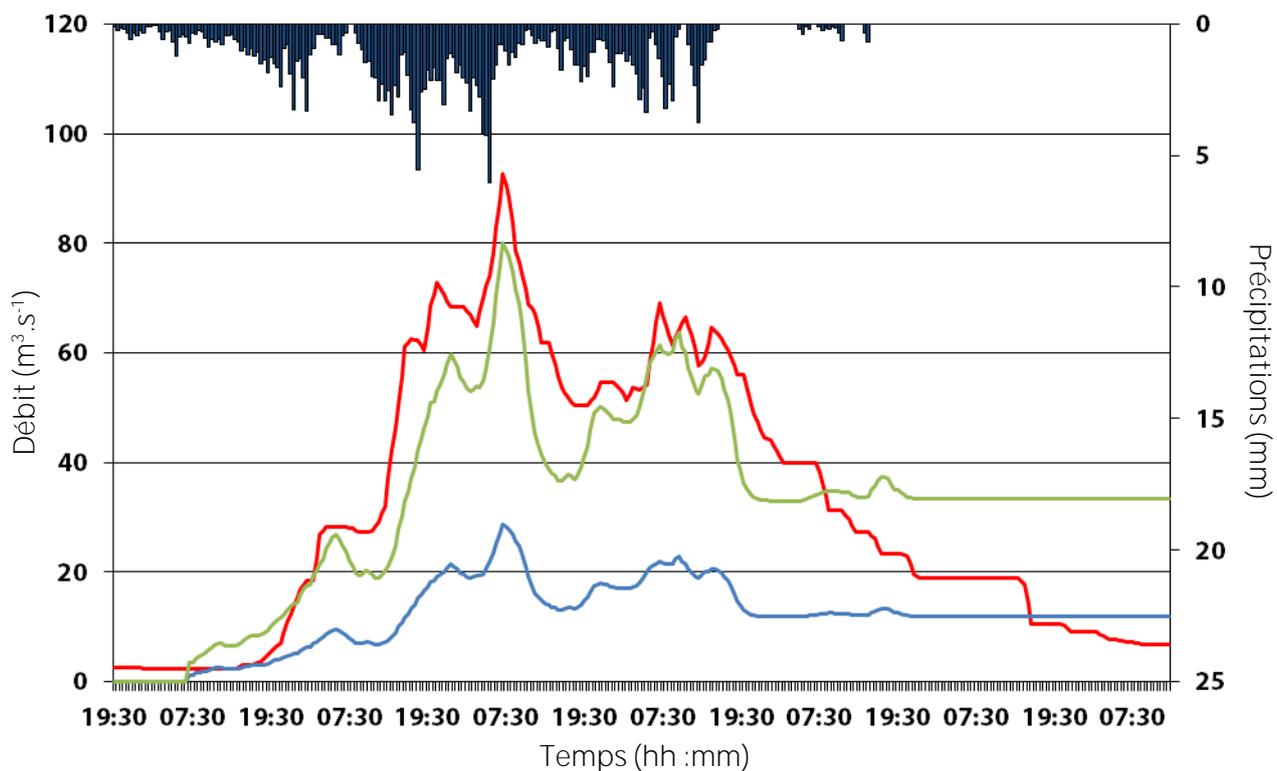


k+4

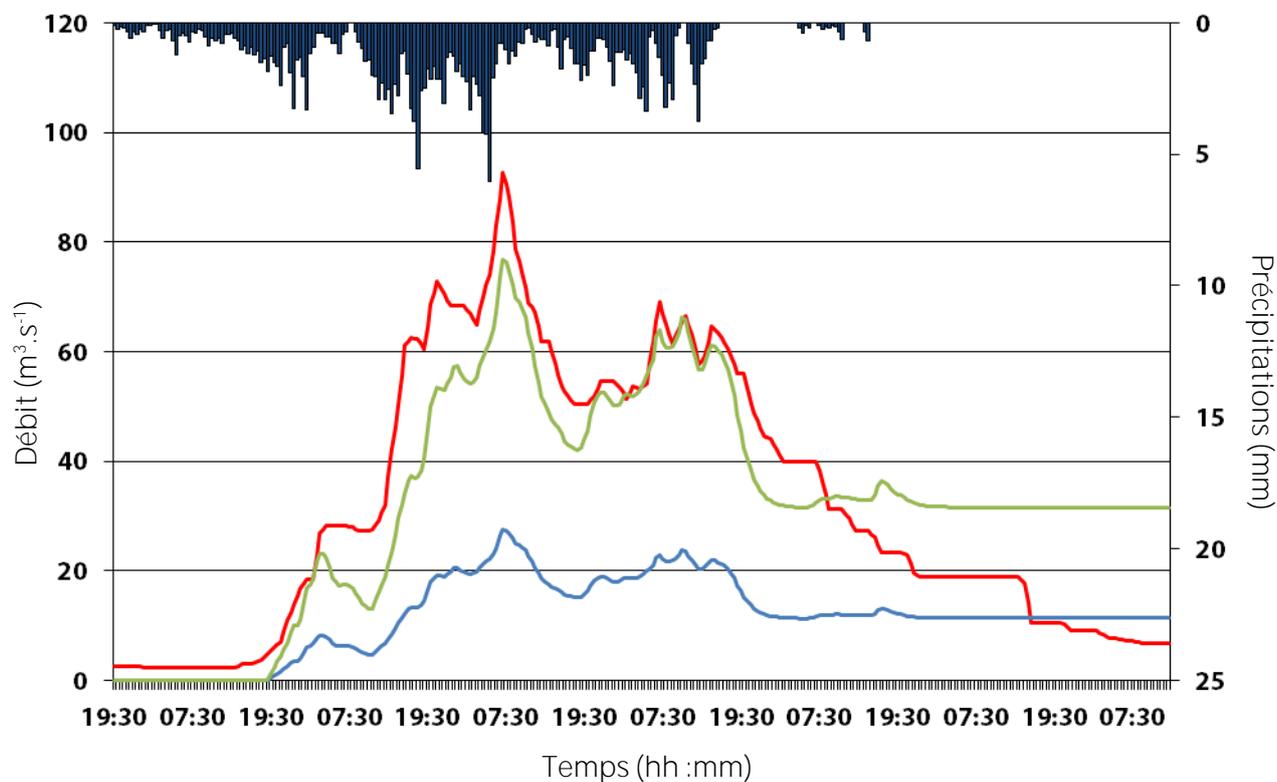


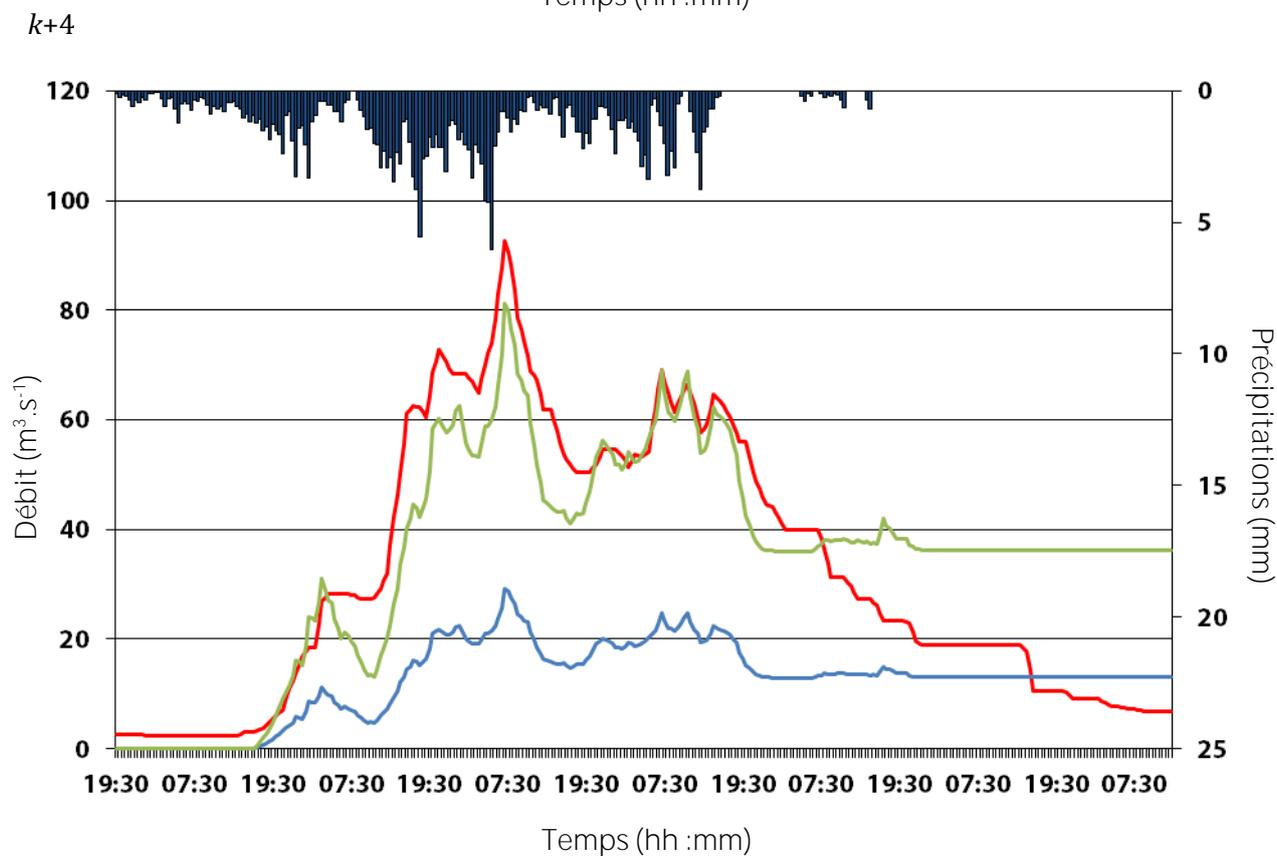
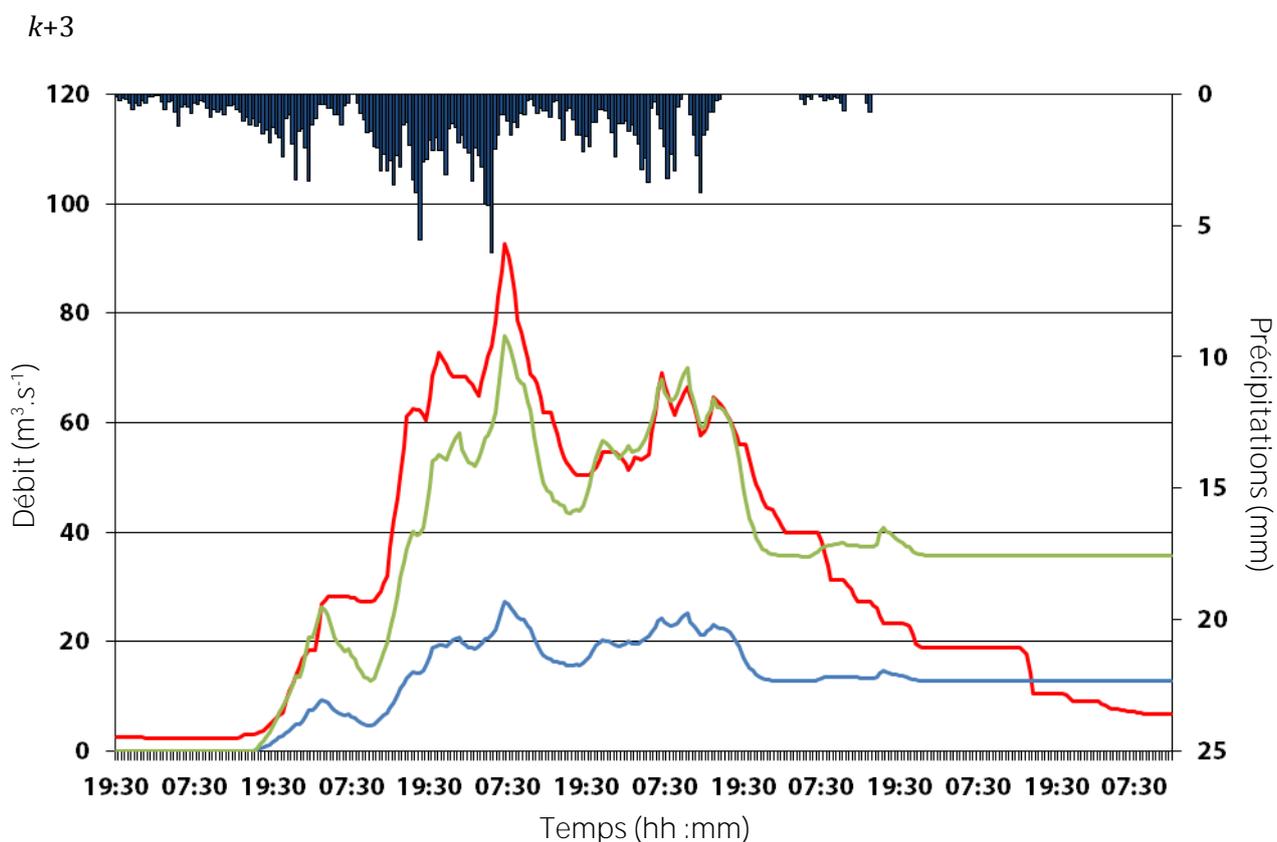
Bassin de la Ganière, évènement n°918 :

k+1



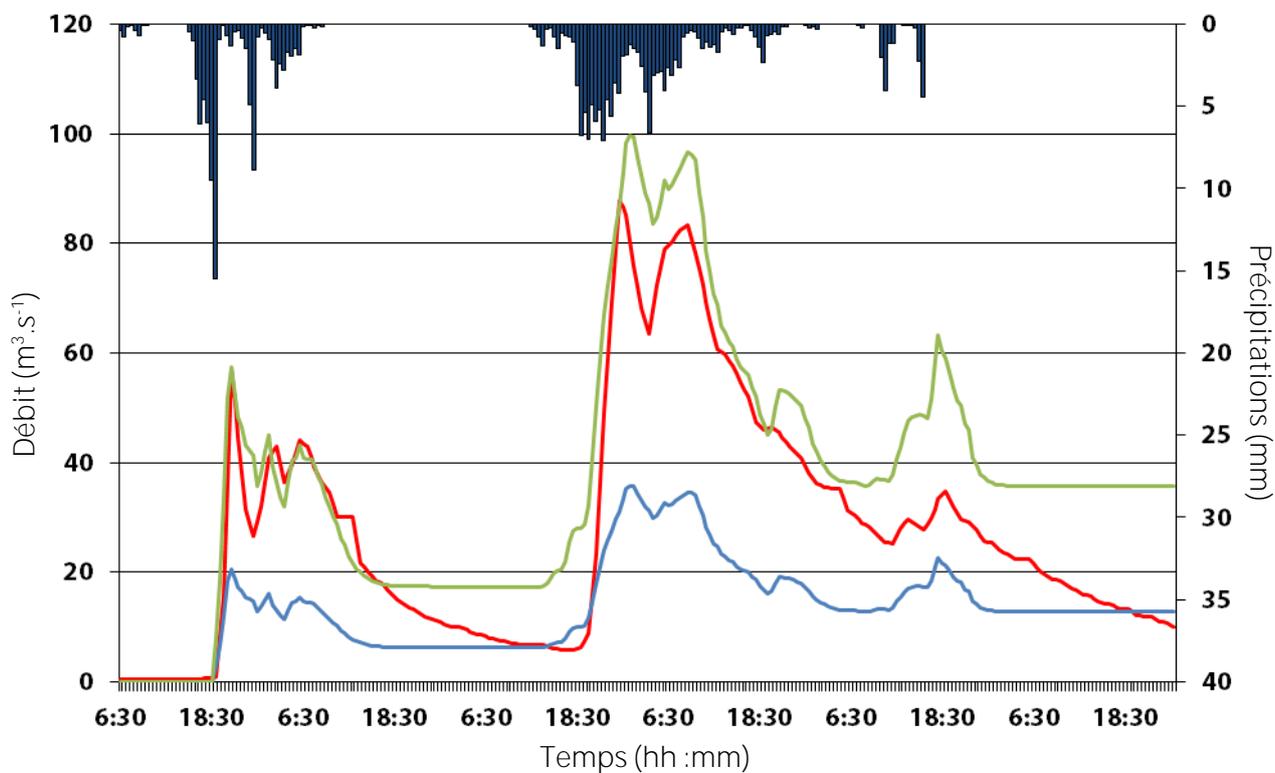
k+2



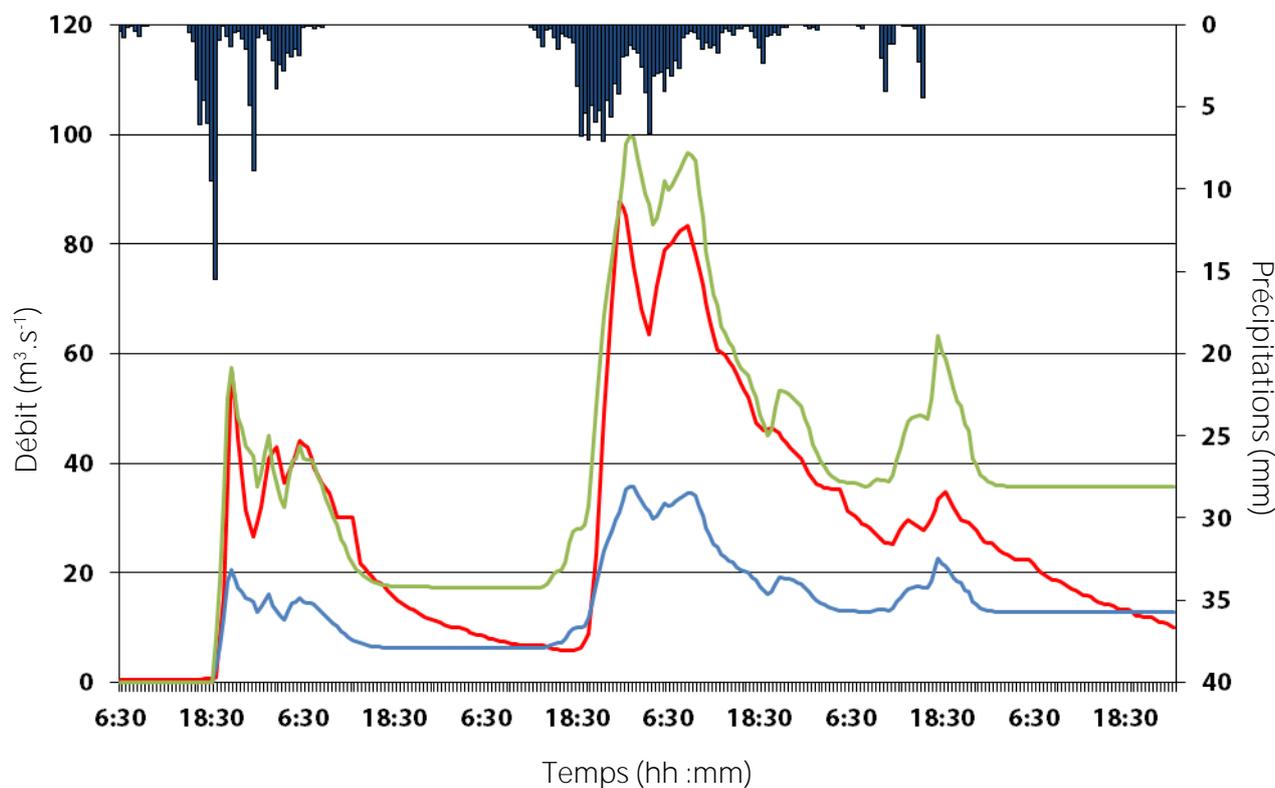


Bassin de la Ganière, évènement n°919 :

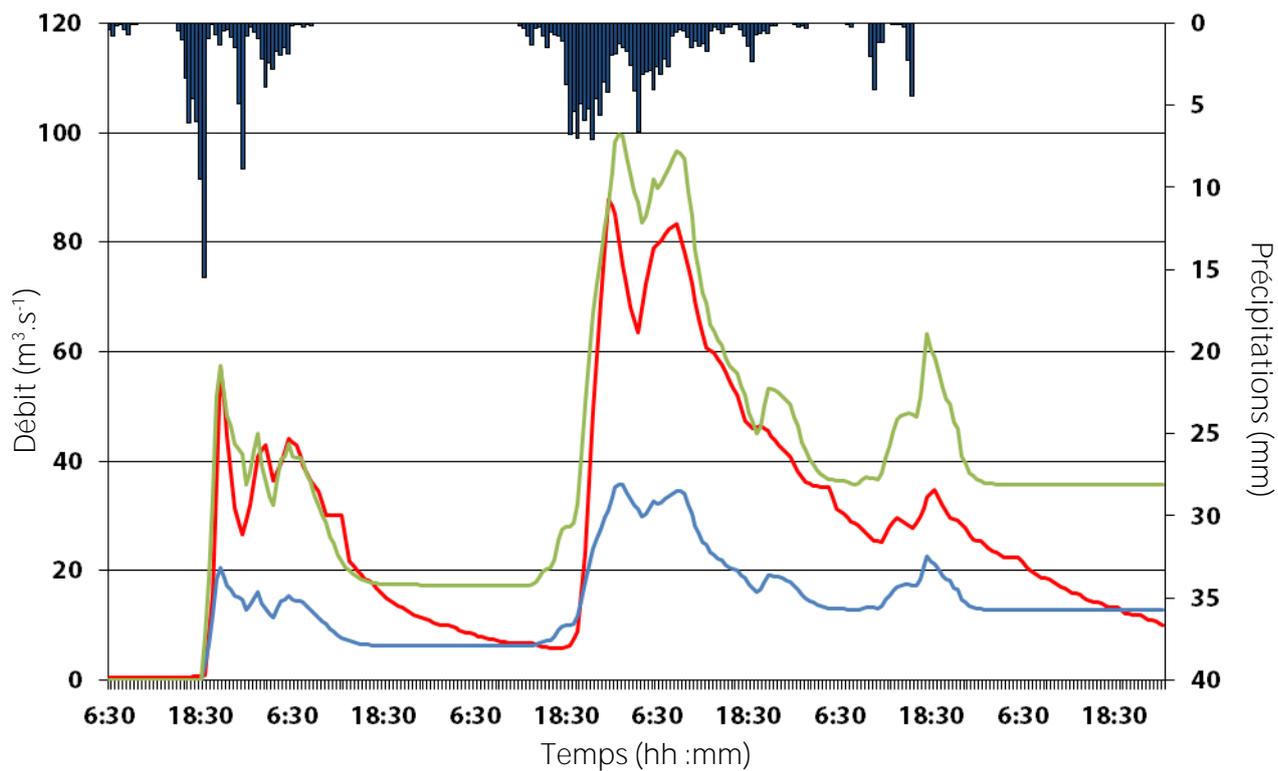
k+1



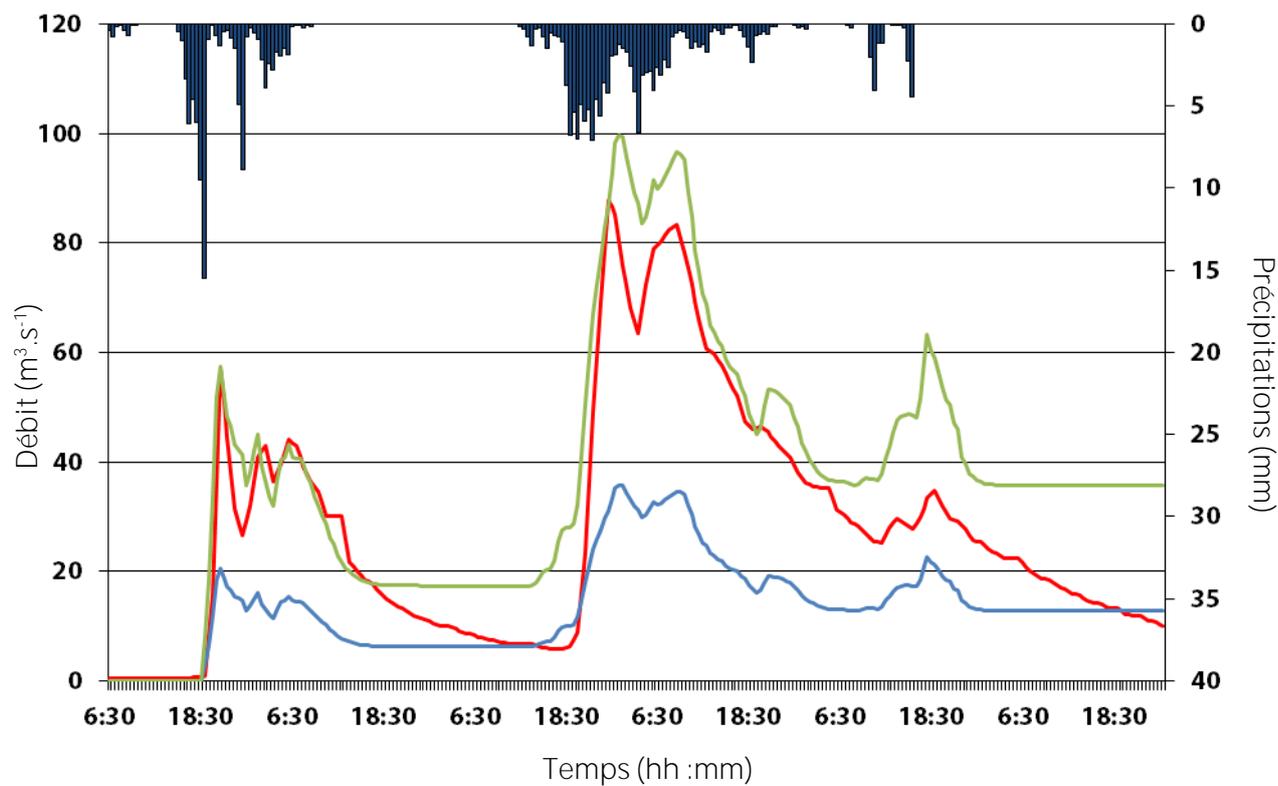
k+2



$k+3$

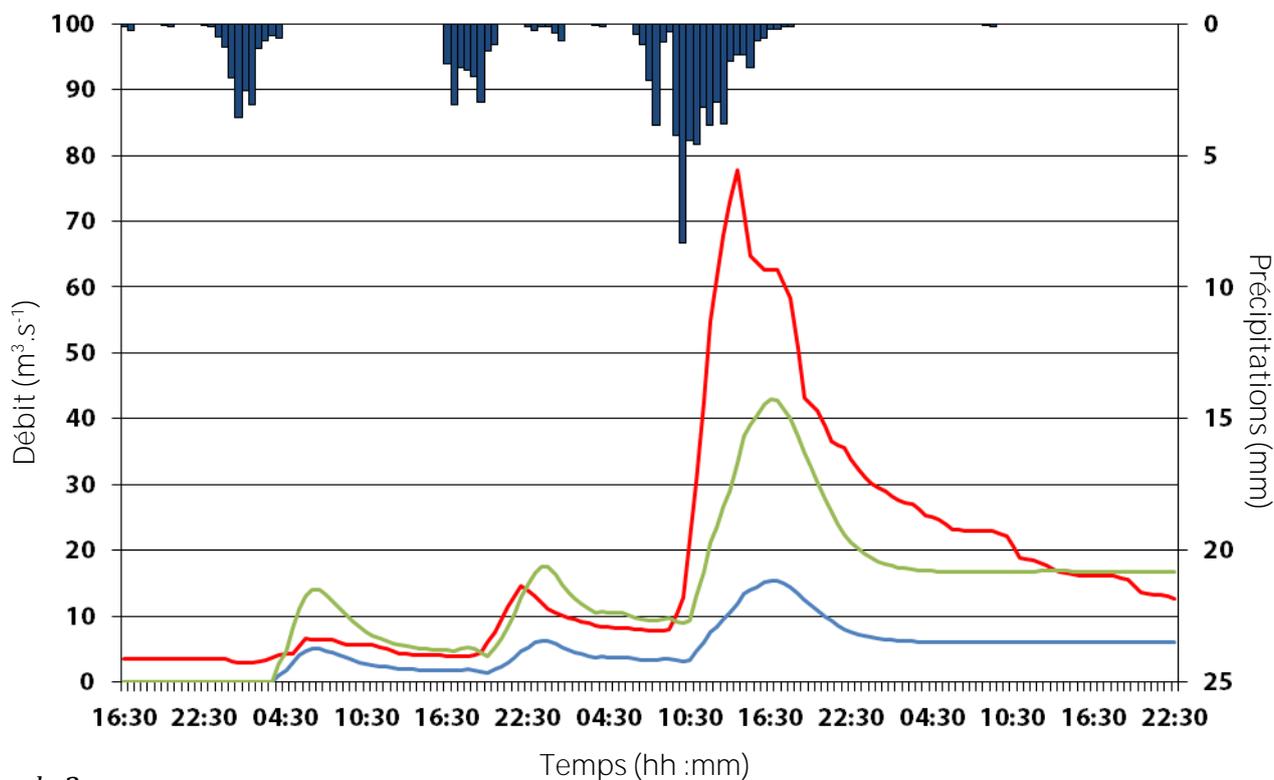


$k+4$

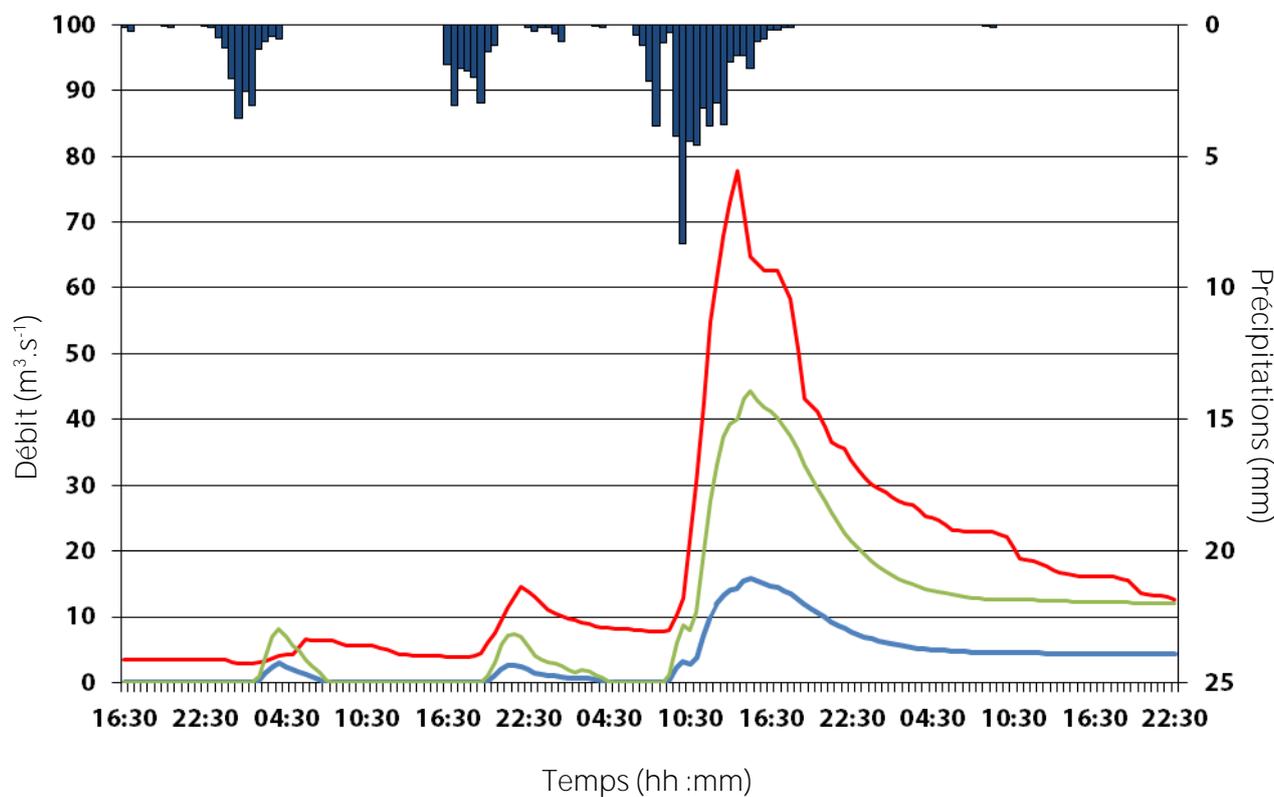


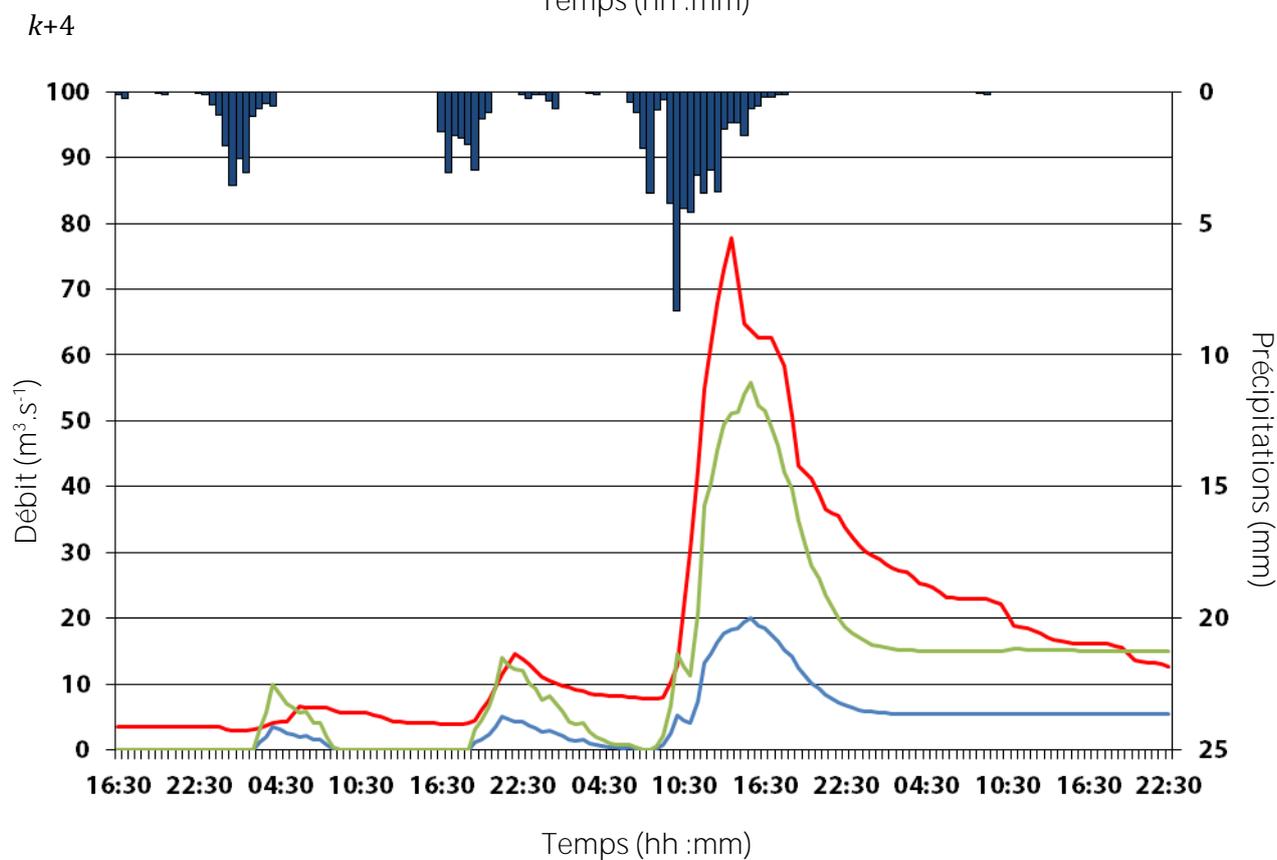
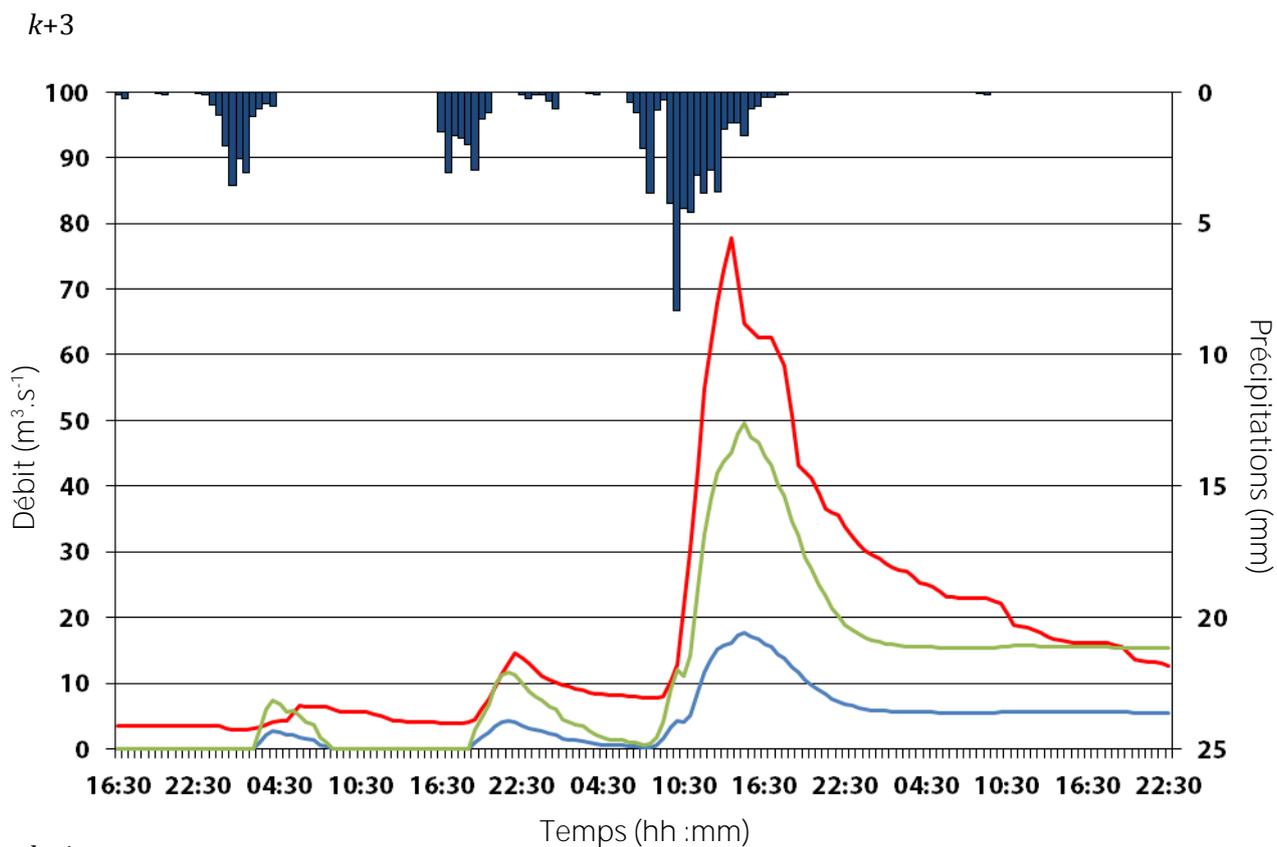
Bassin de la Ganière, évènement n°920 :

k+1



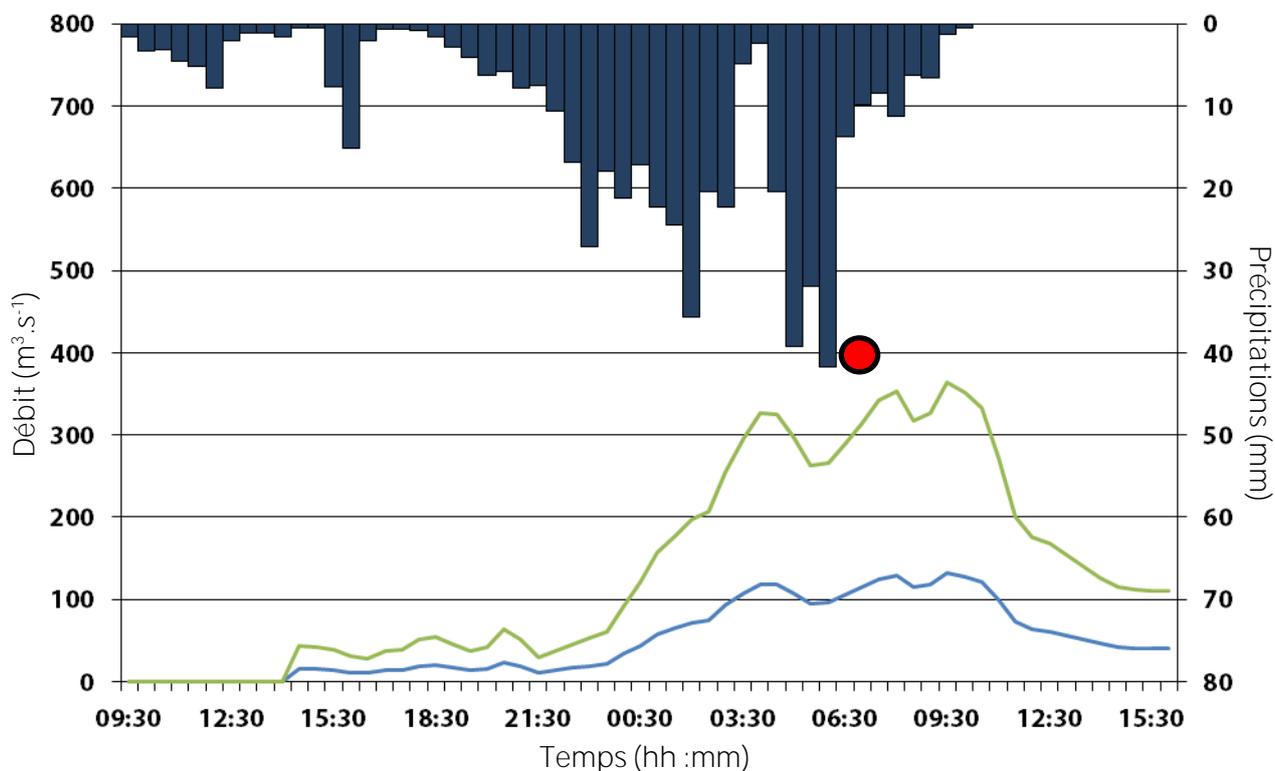
k+2



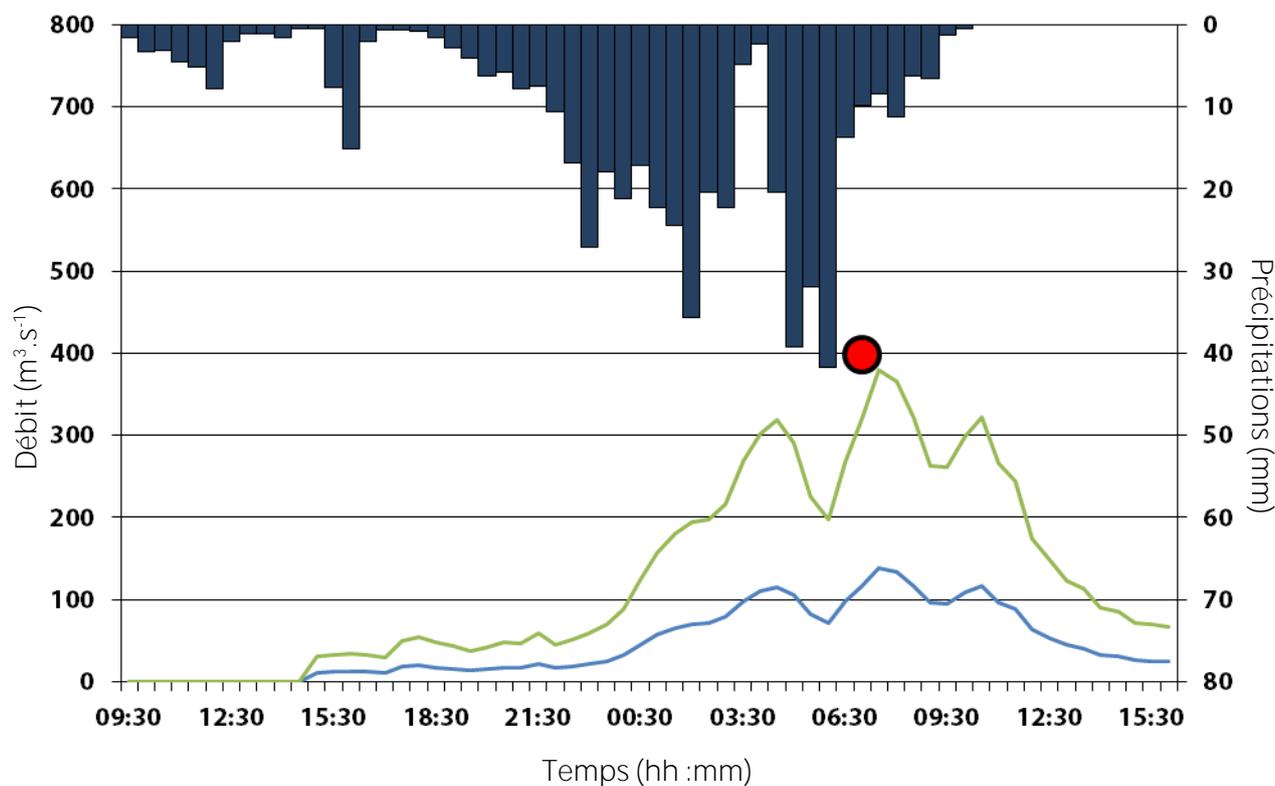


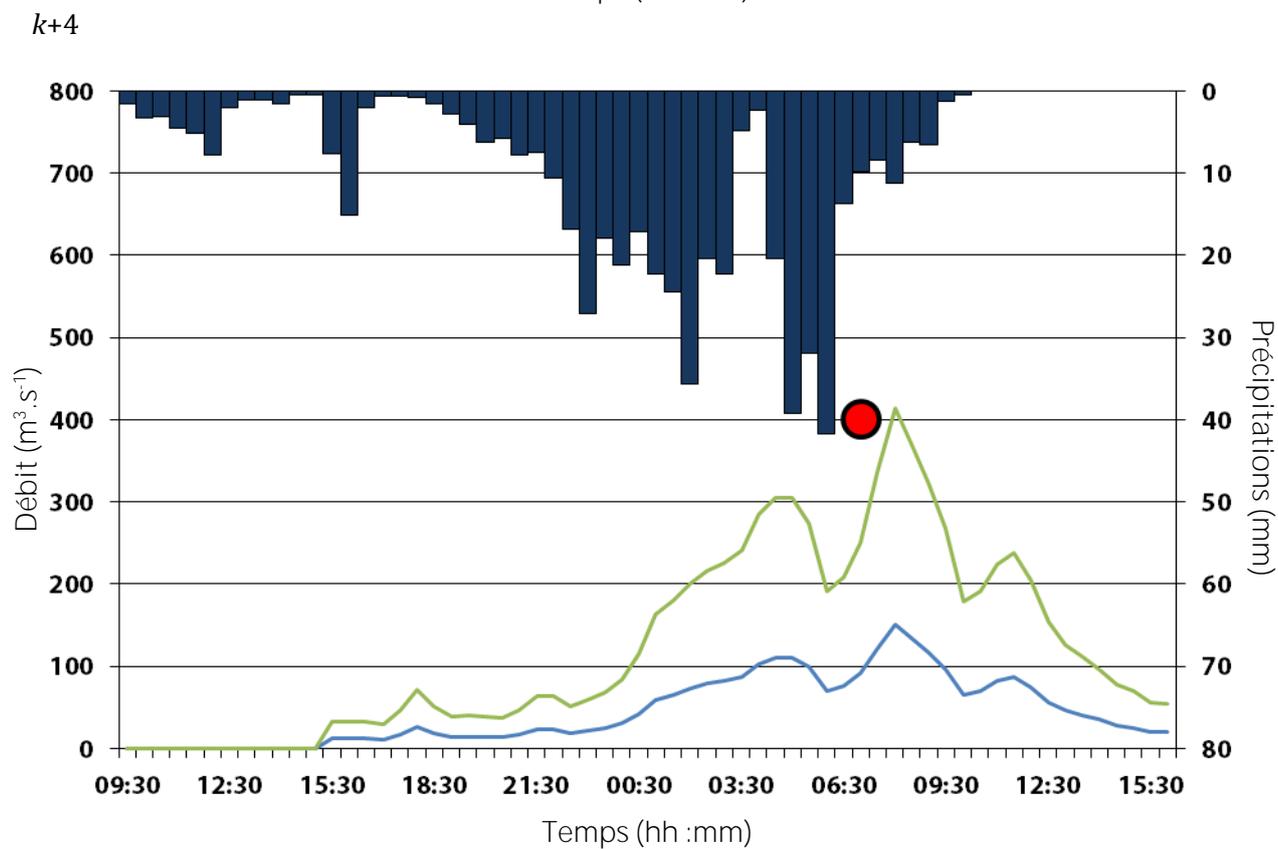
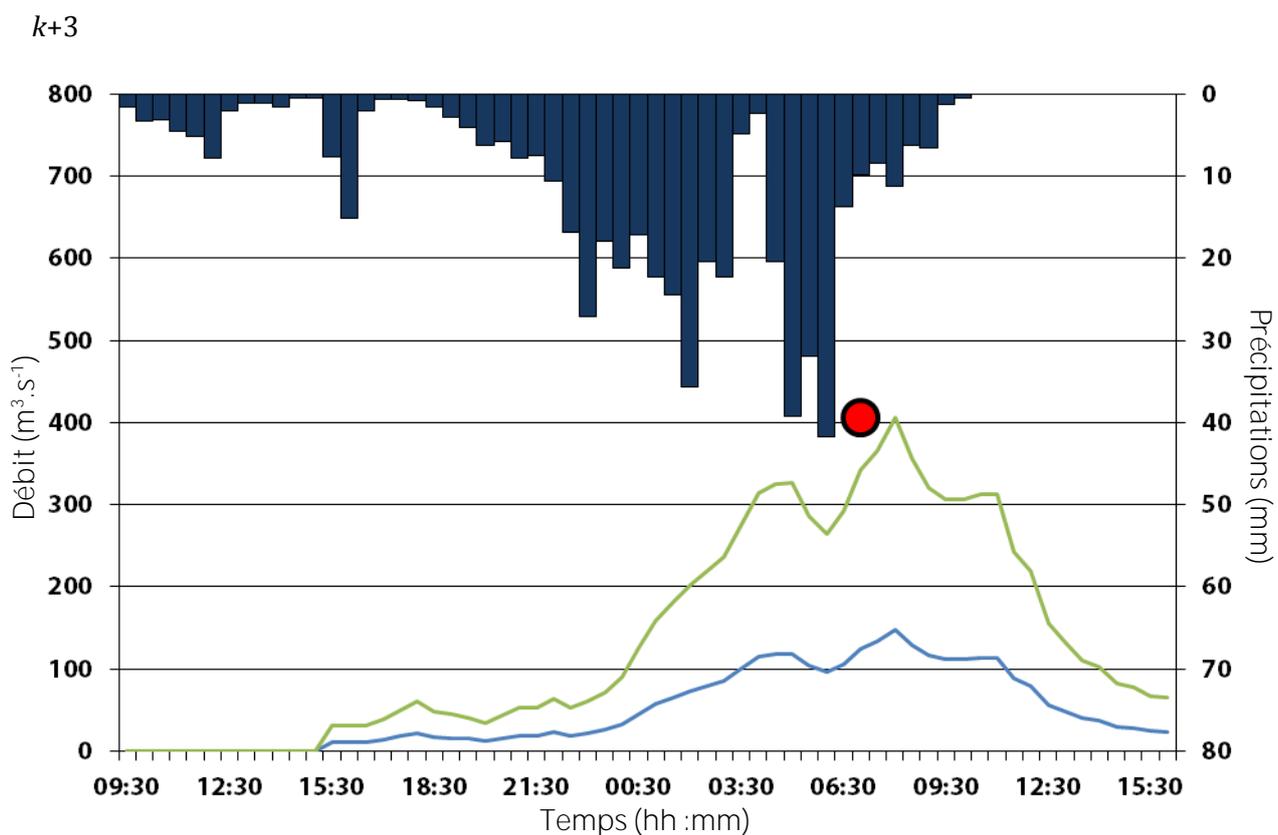
Bassin de l'Avène, évènement n°927 :

k+1



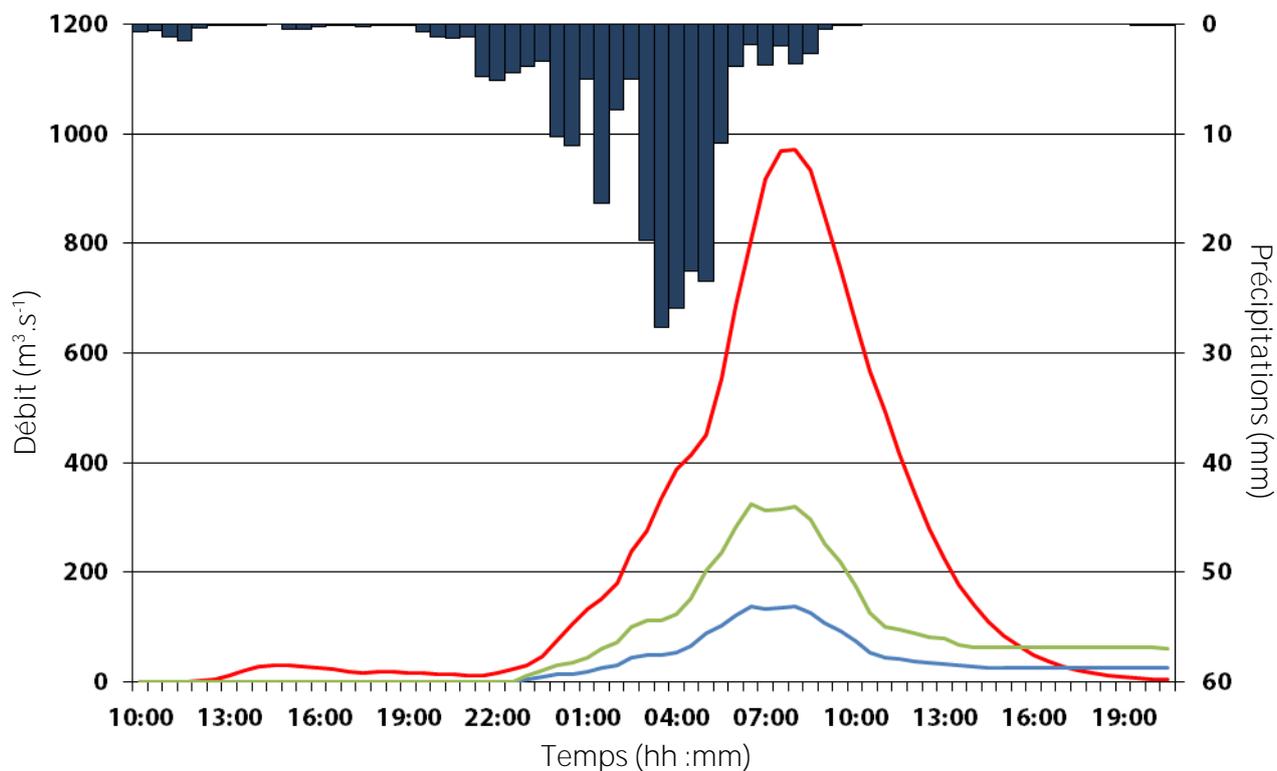
k+2



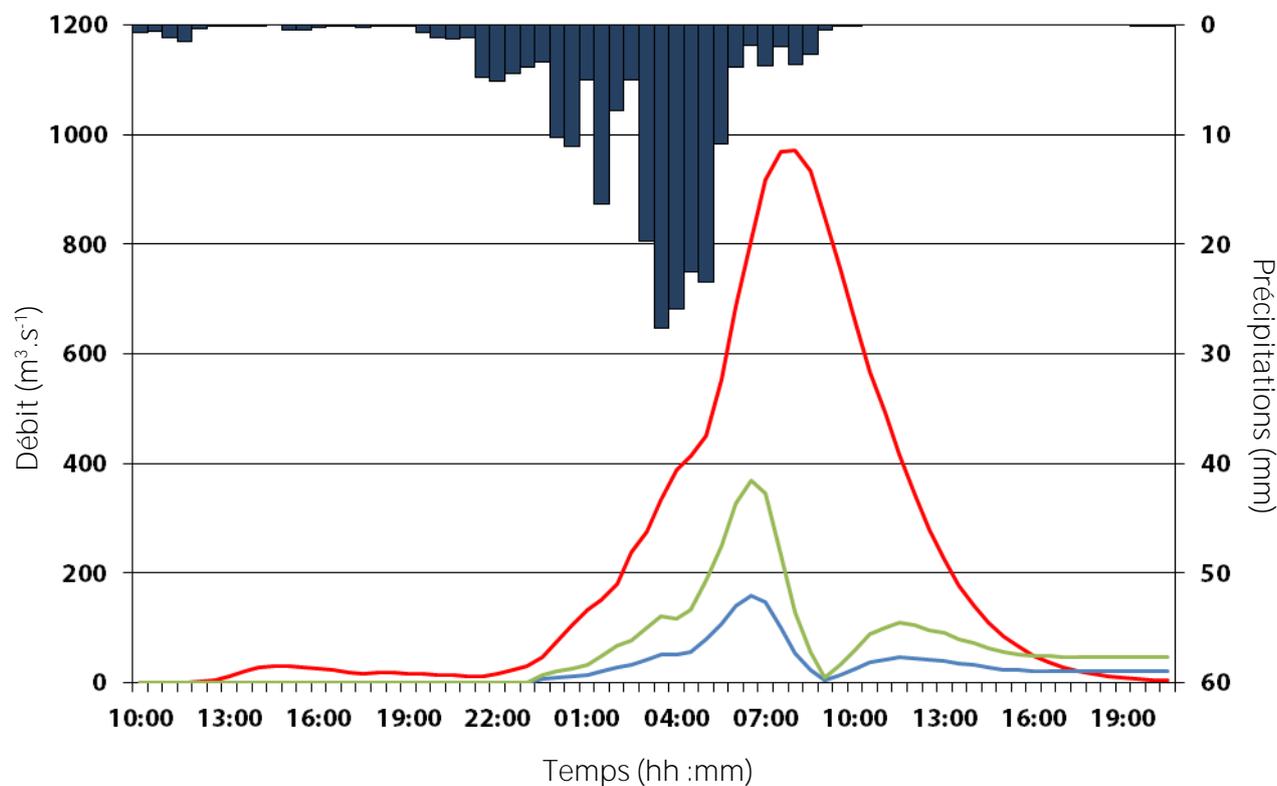


Bassin du Galeizon complet, évènement n°924 :

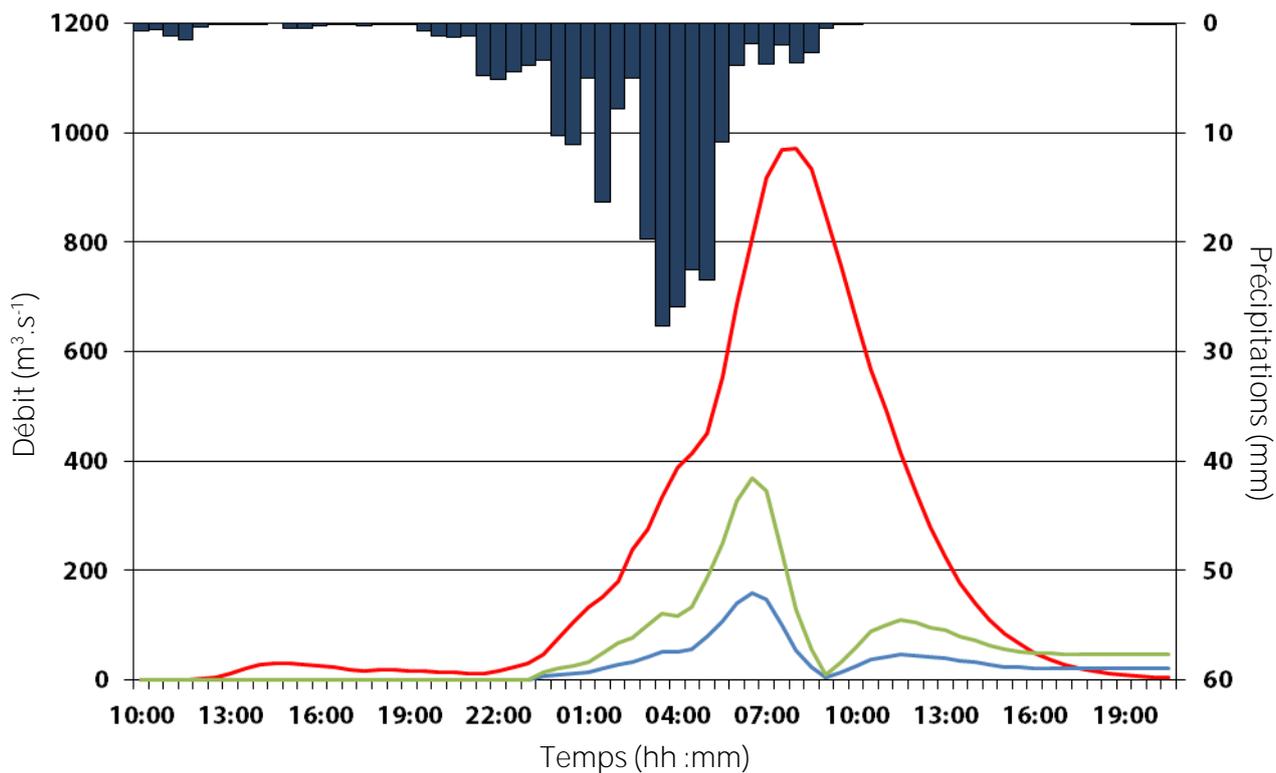
k+1



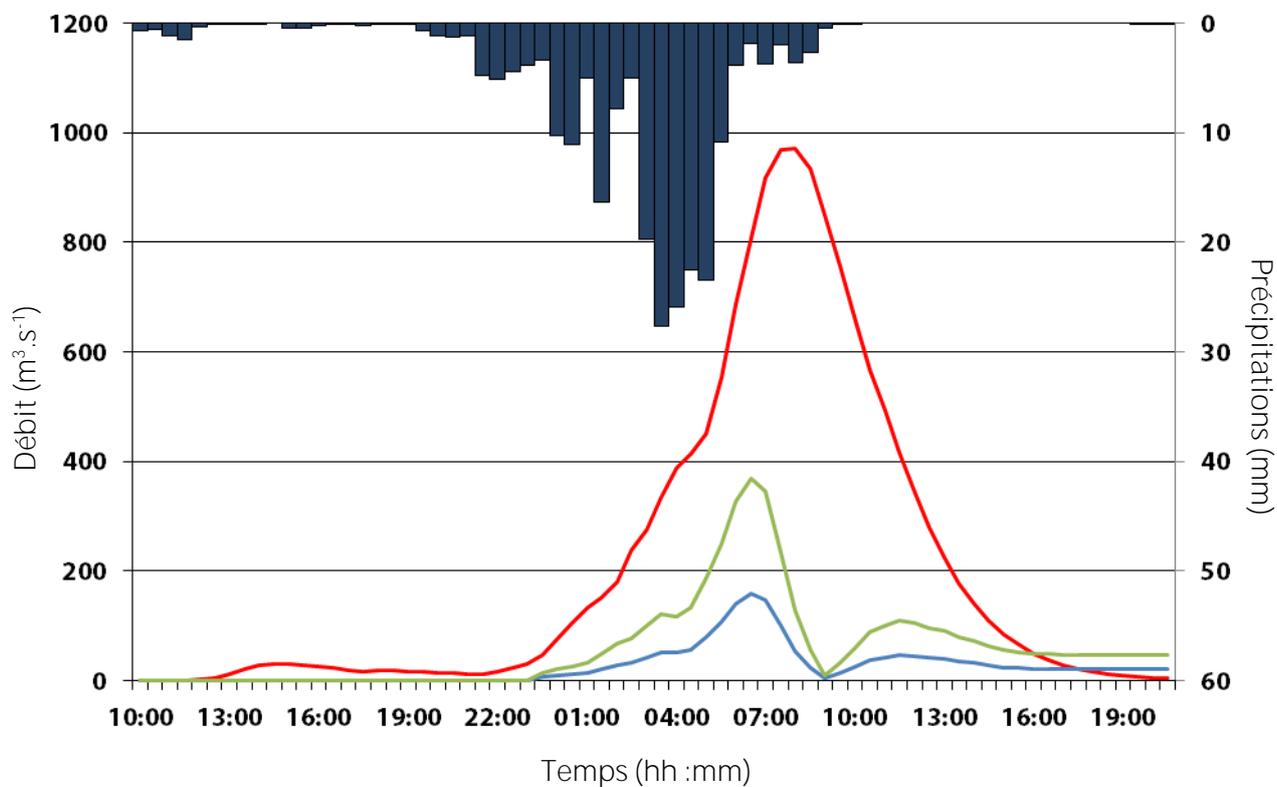
k+2



$k+3$

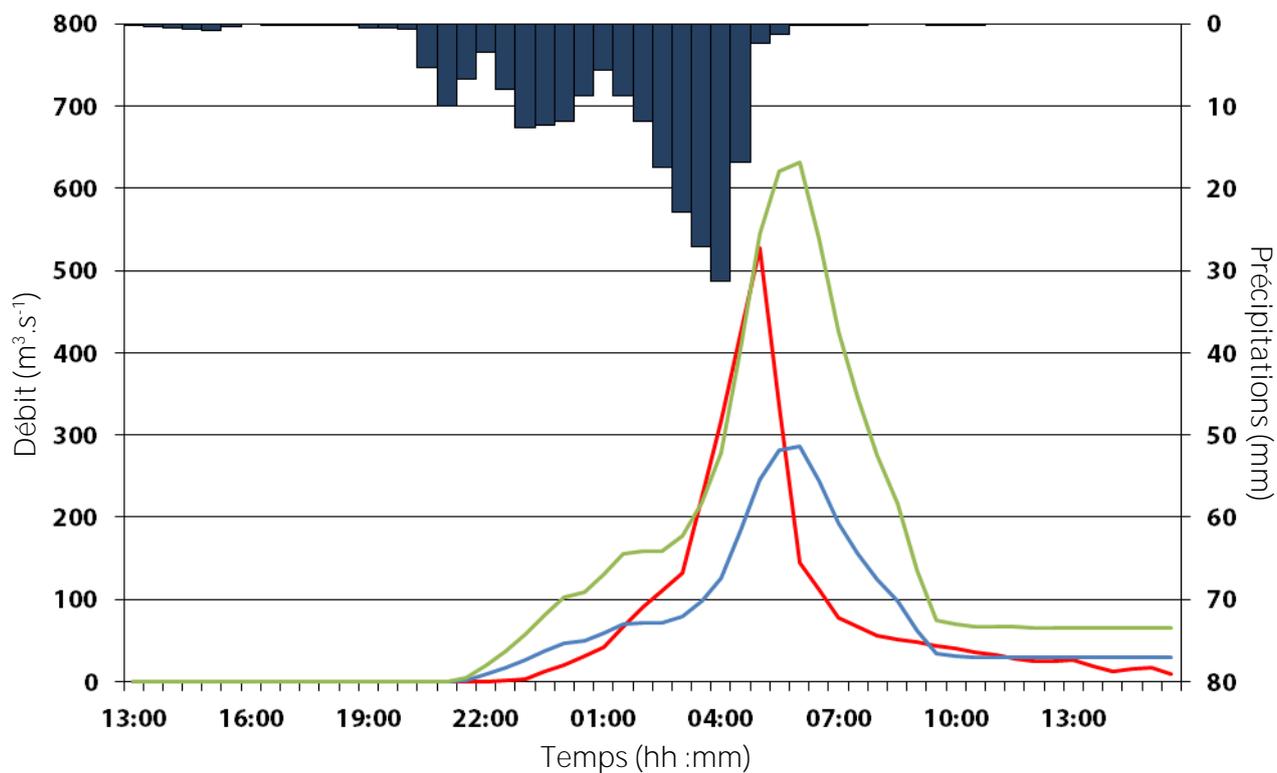


$k+4$

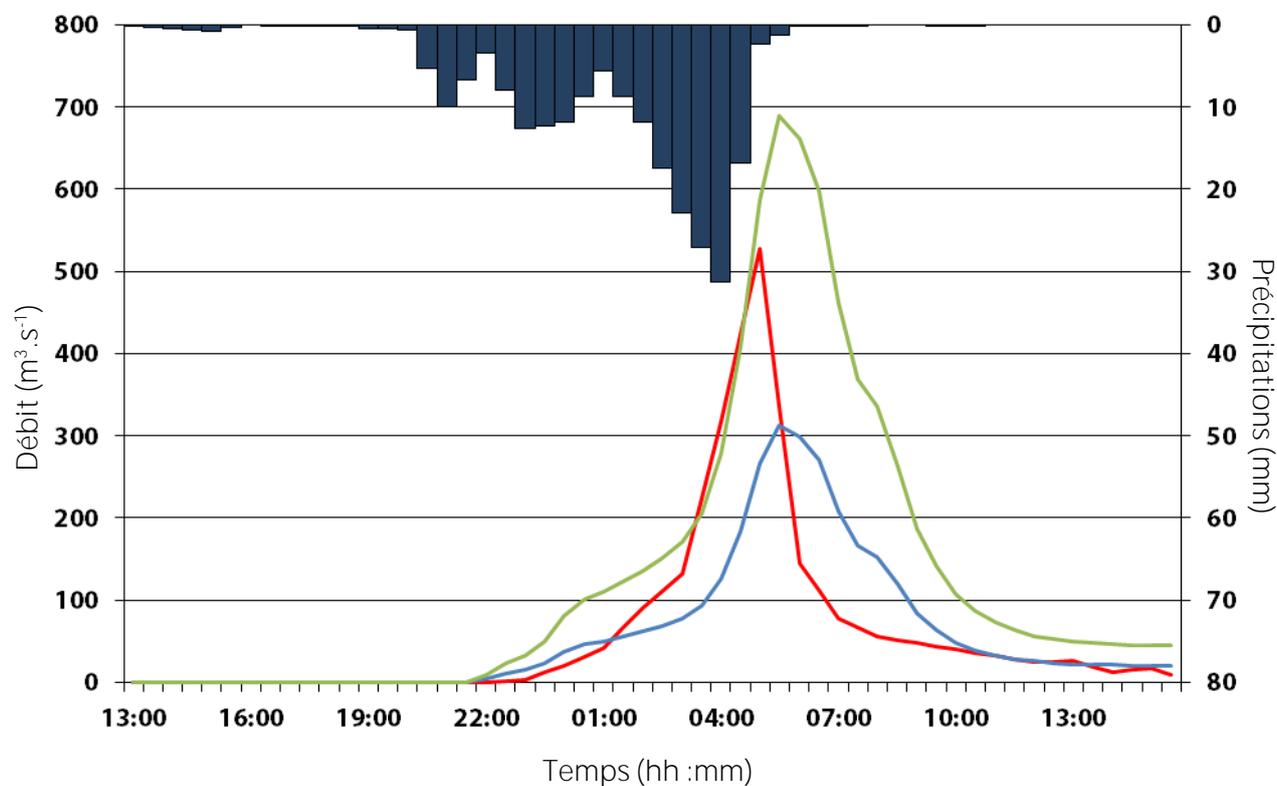


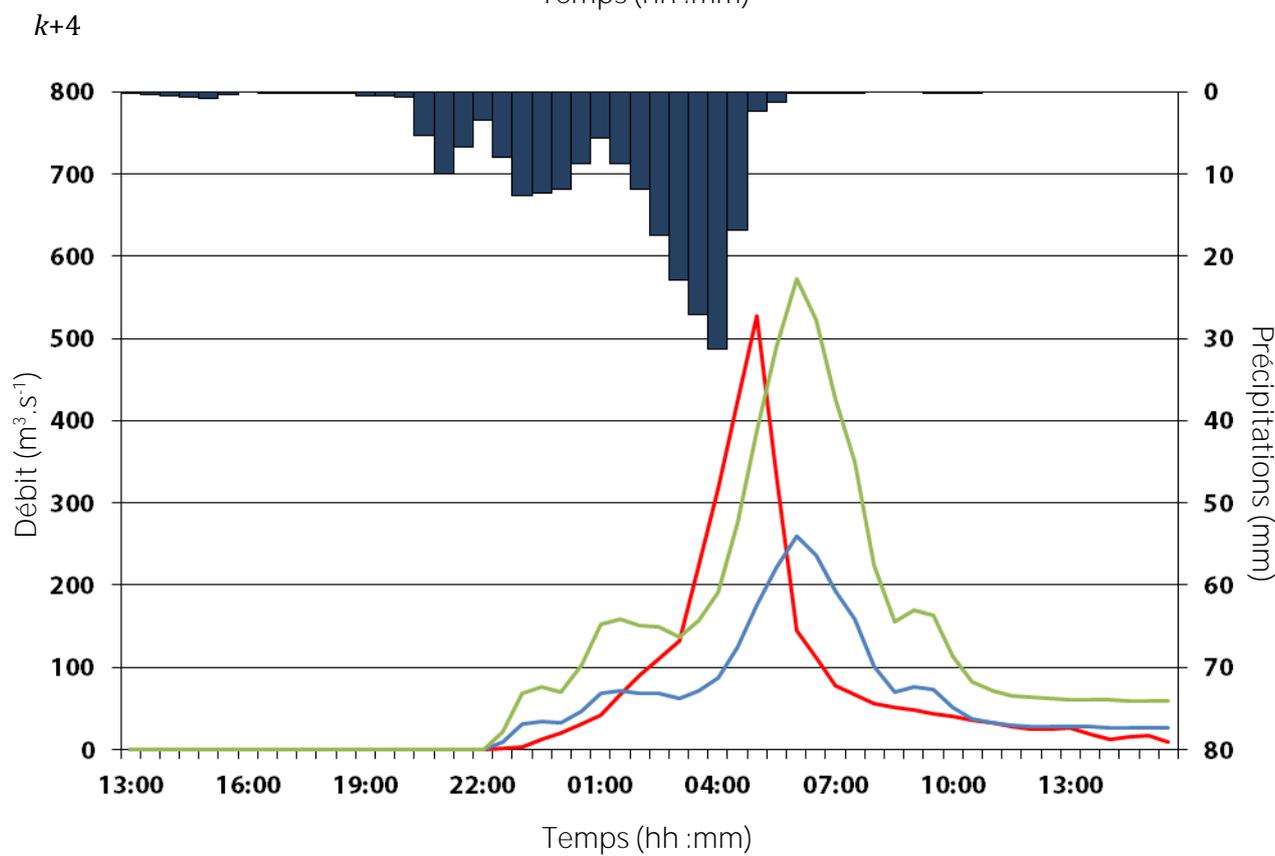
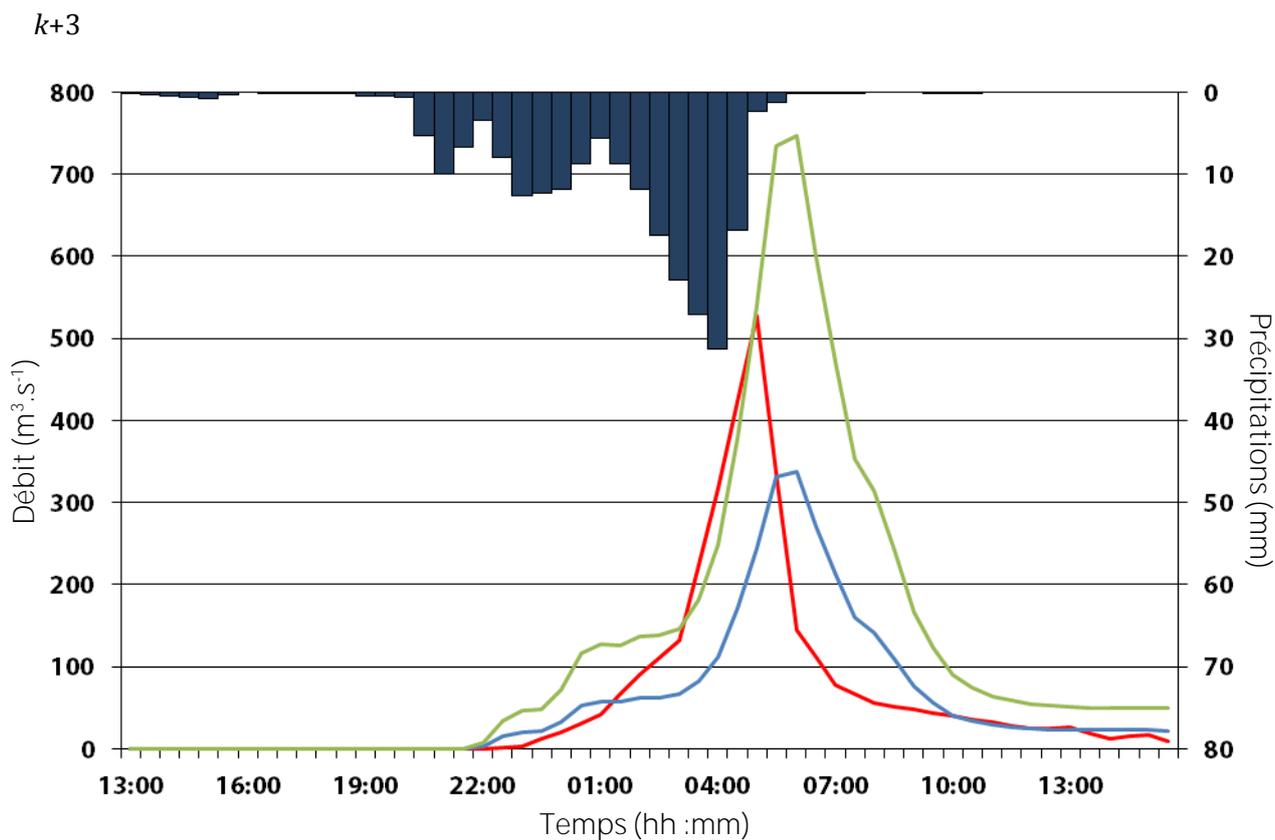
Bassin du Gardon à Saumane, évènement n°931 :

k+1



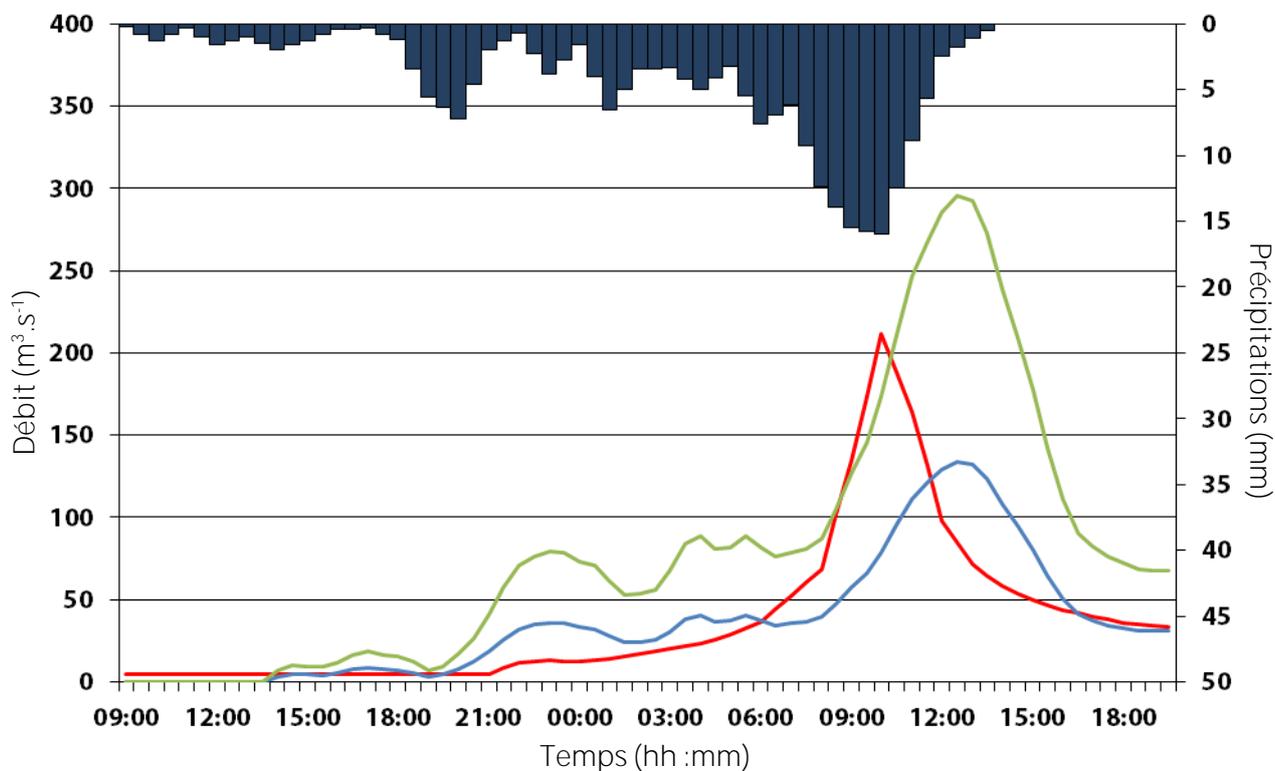
k+2



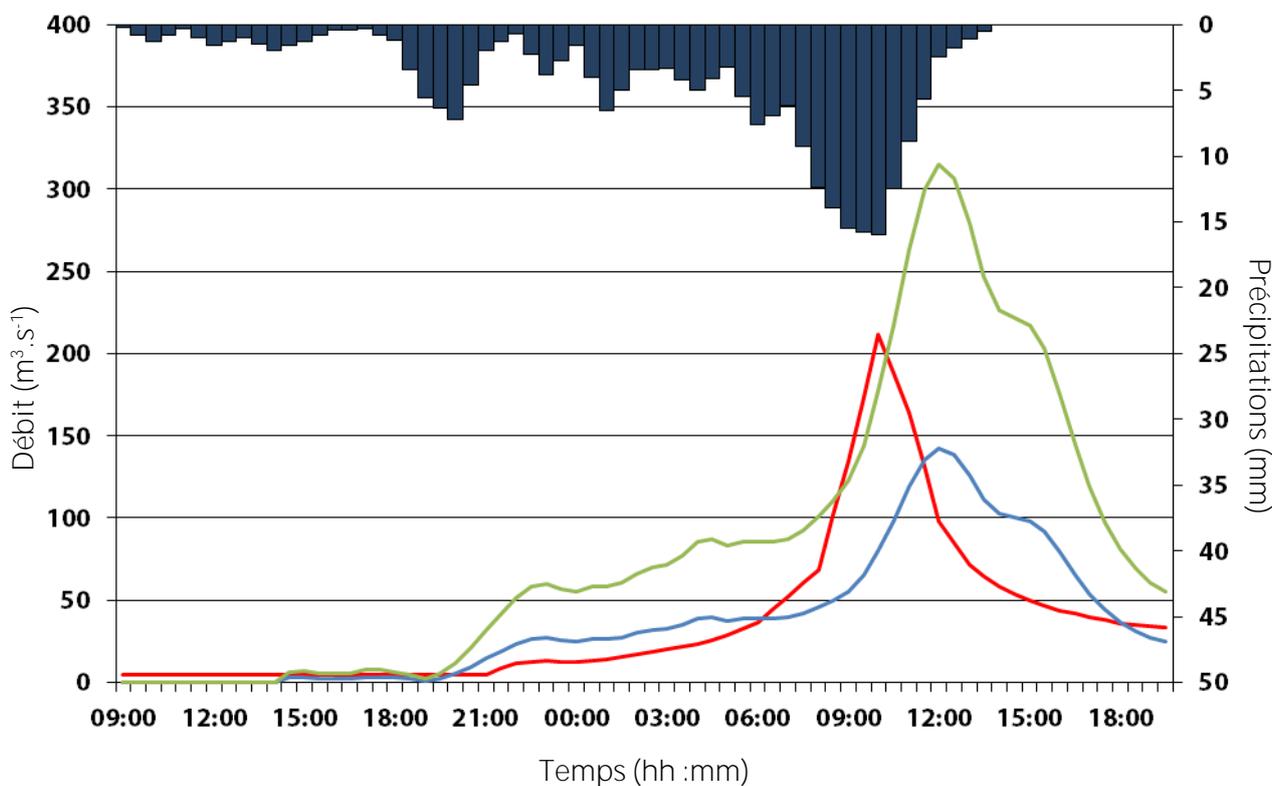


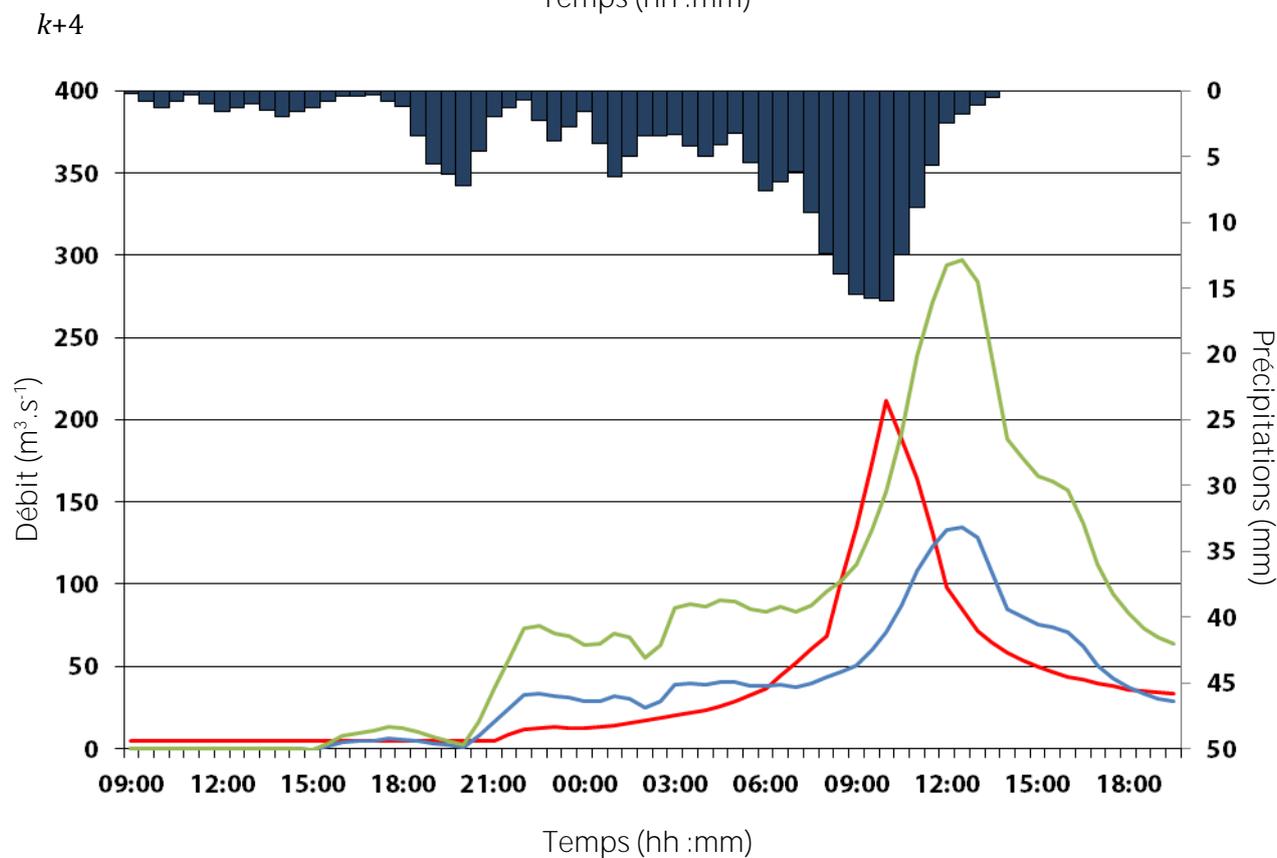
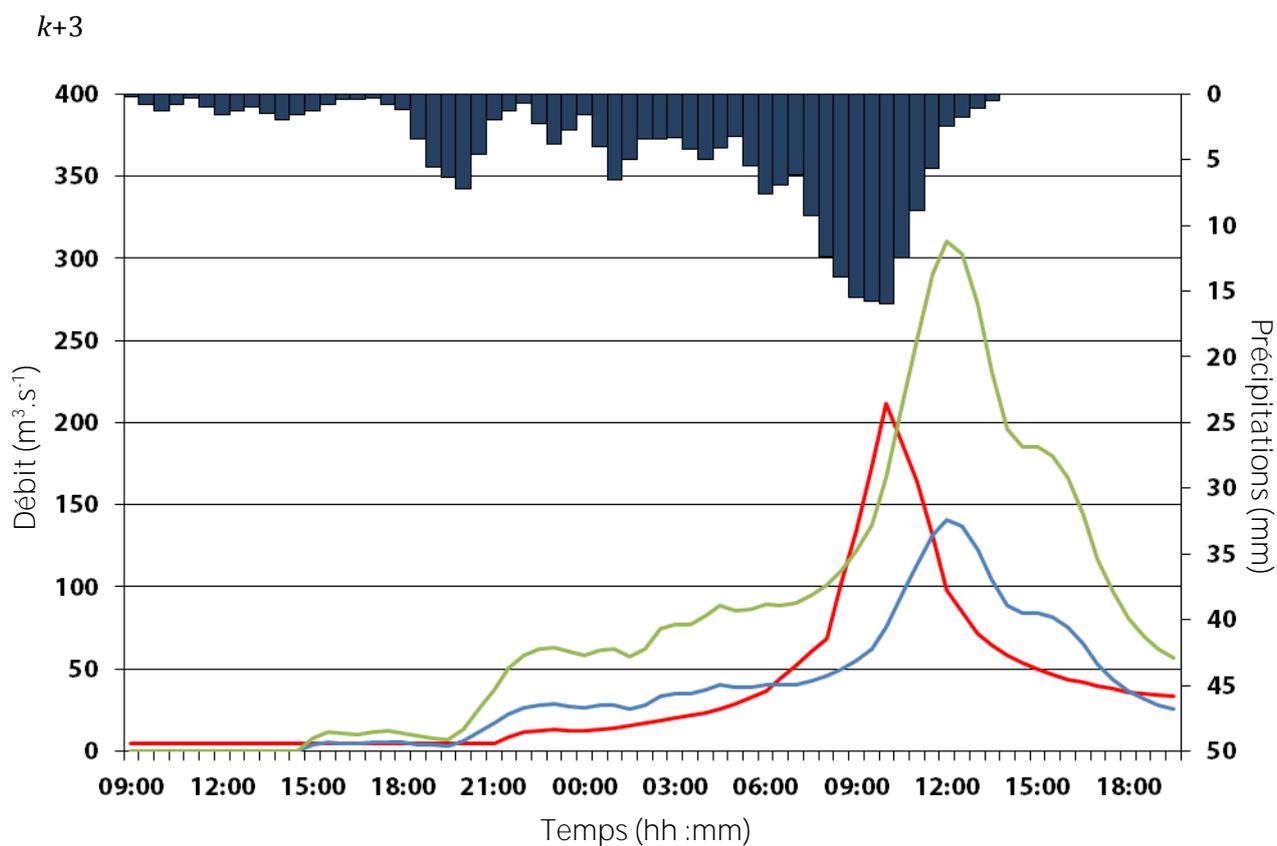
Bassin du Gardon à Saumane, évènement n°932 :

k+1



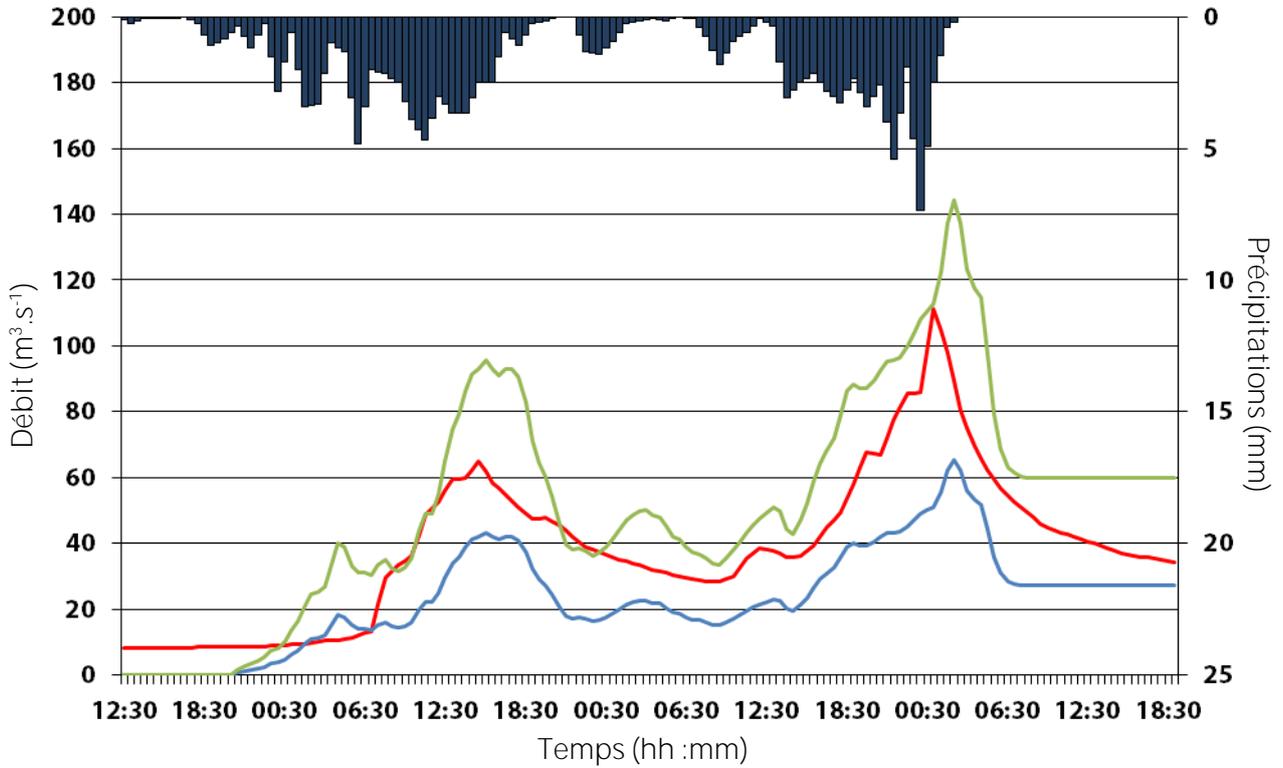
k+2



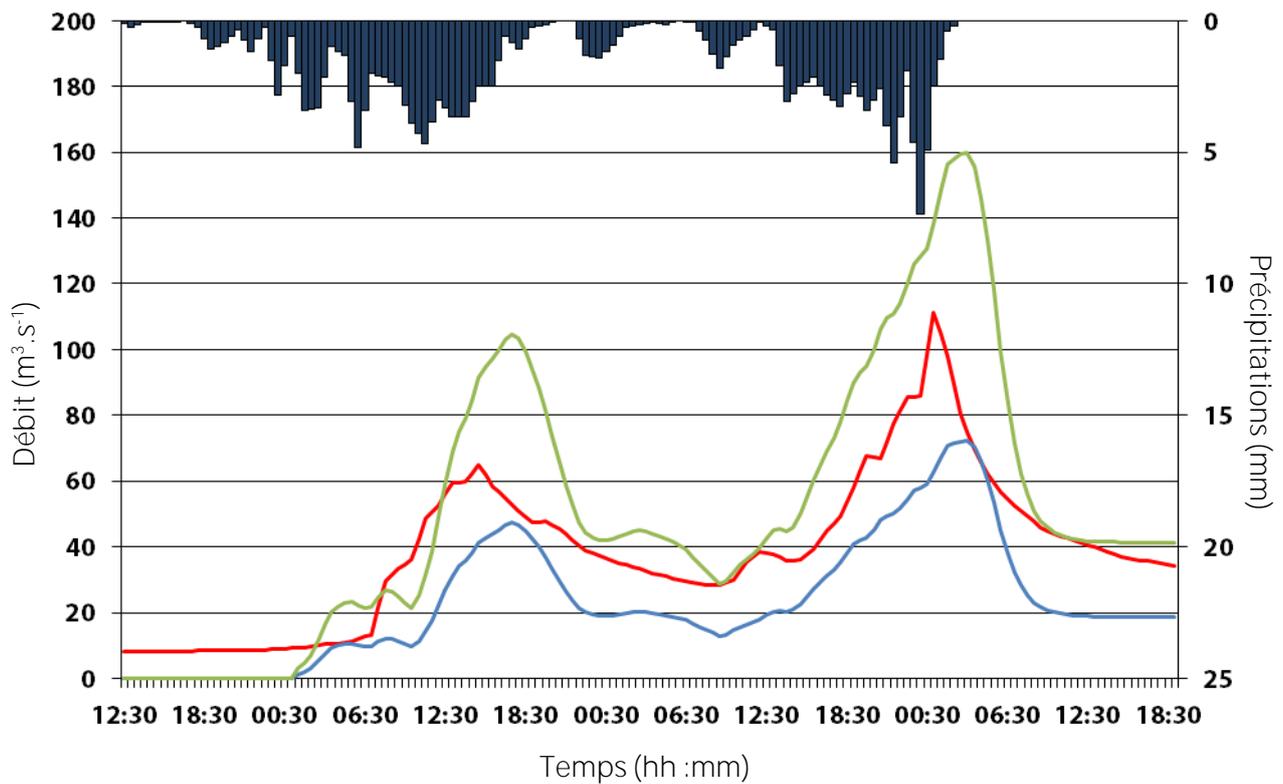


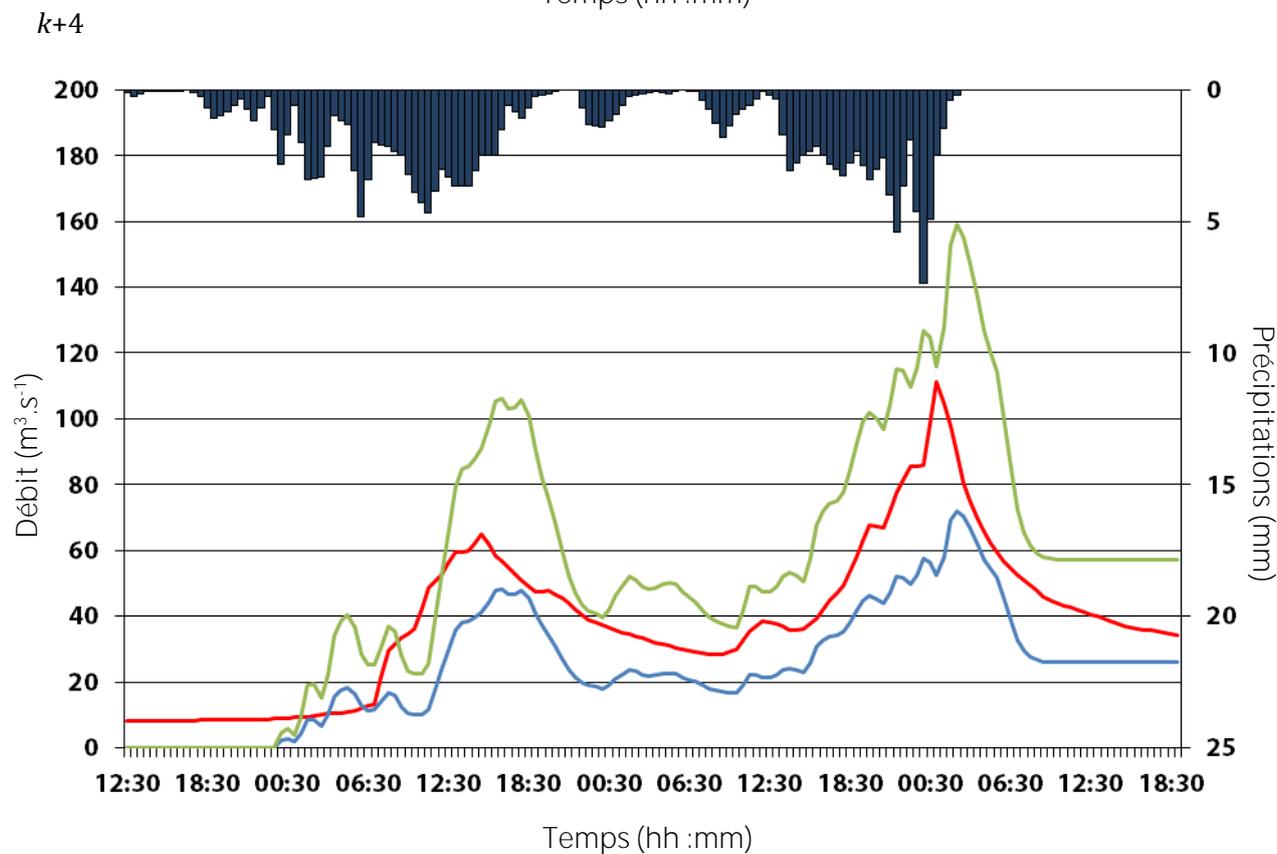
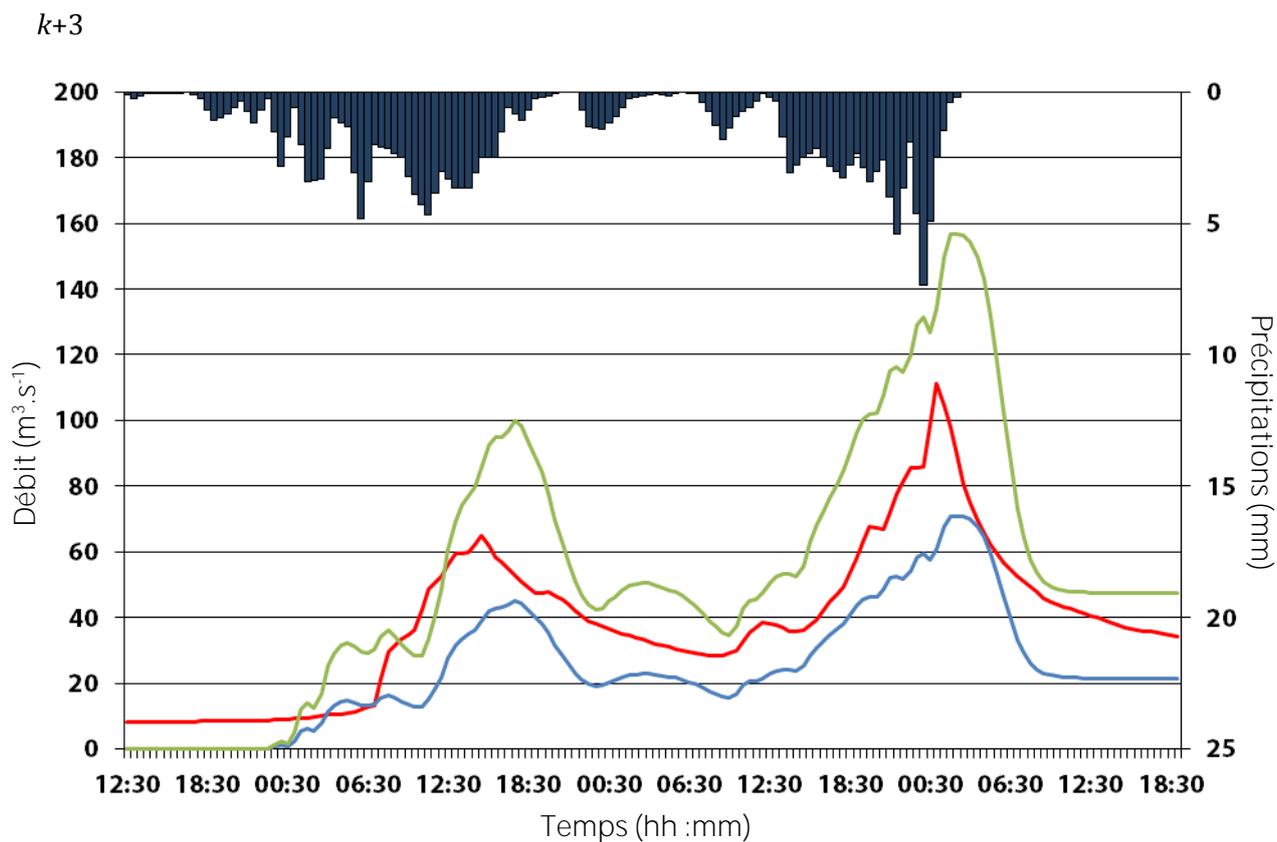
Bassin du Gardon à Saumane, évènement n°933 :

$k+1$



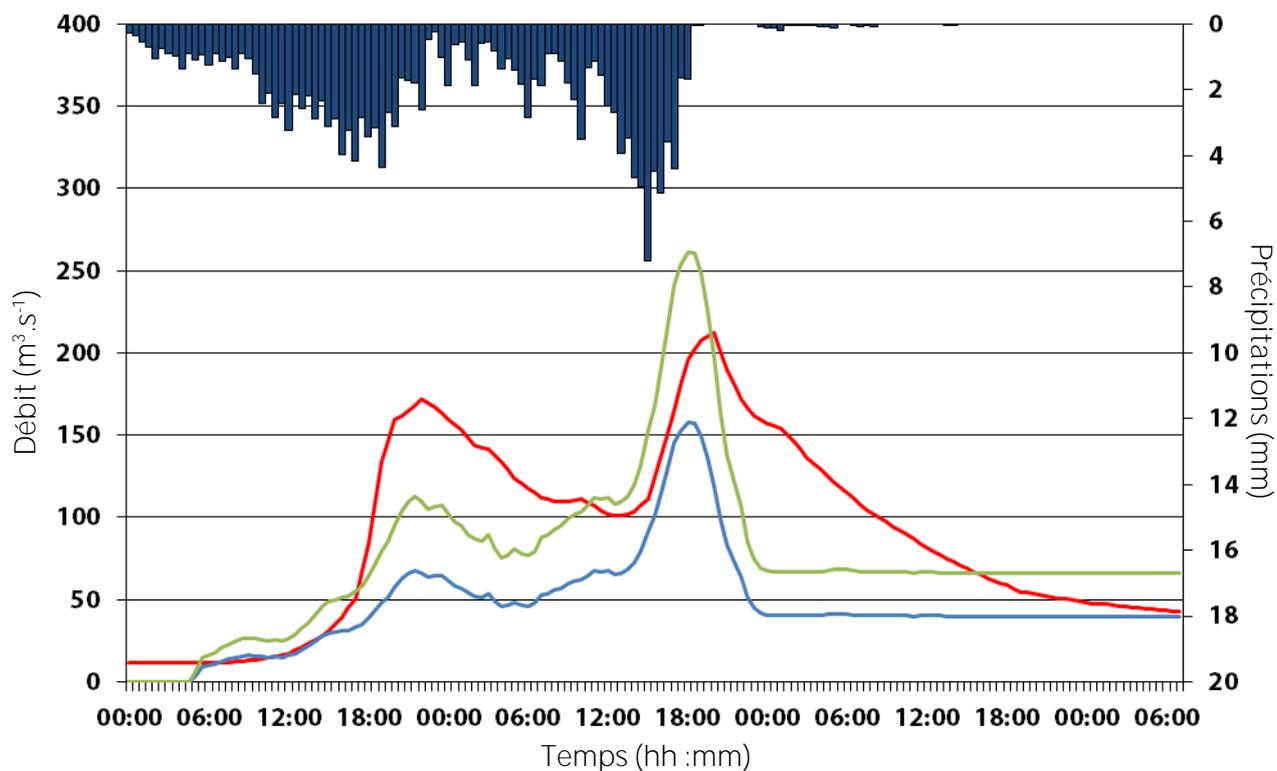
$k+2$



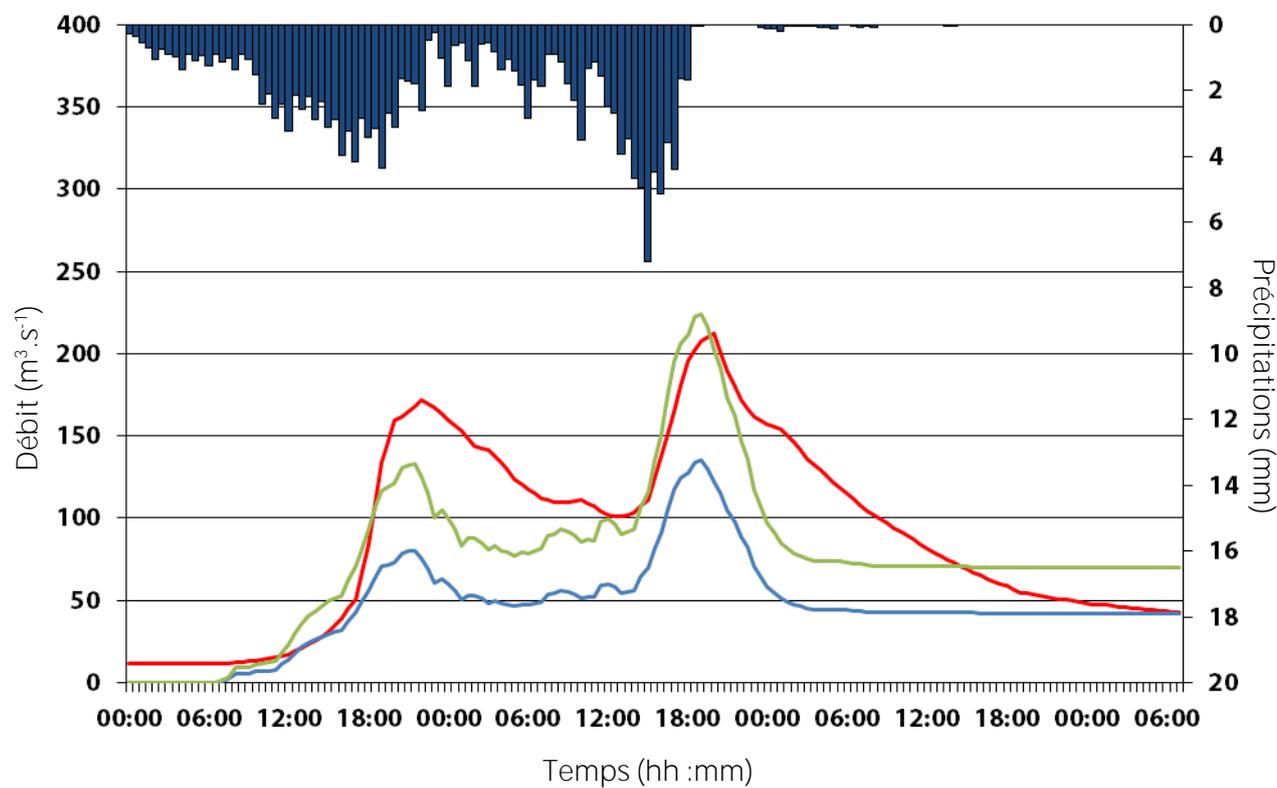


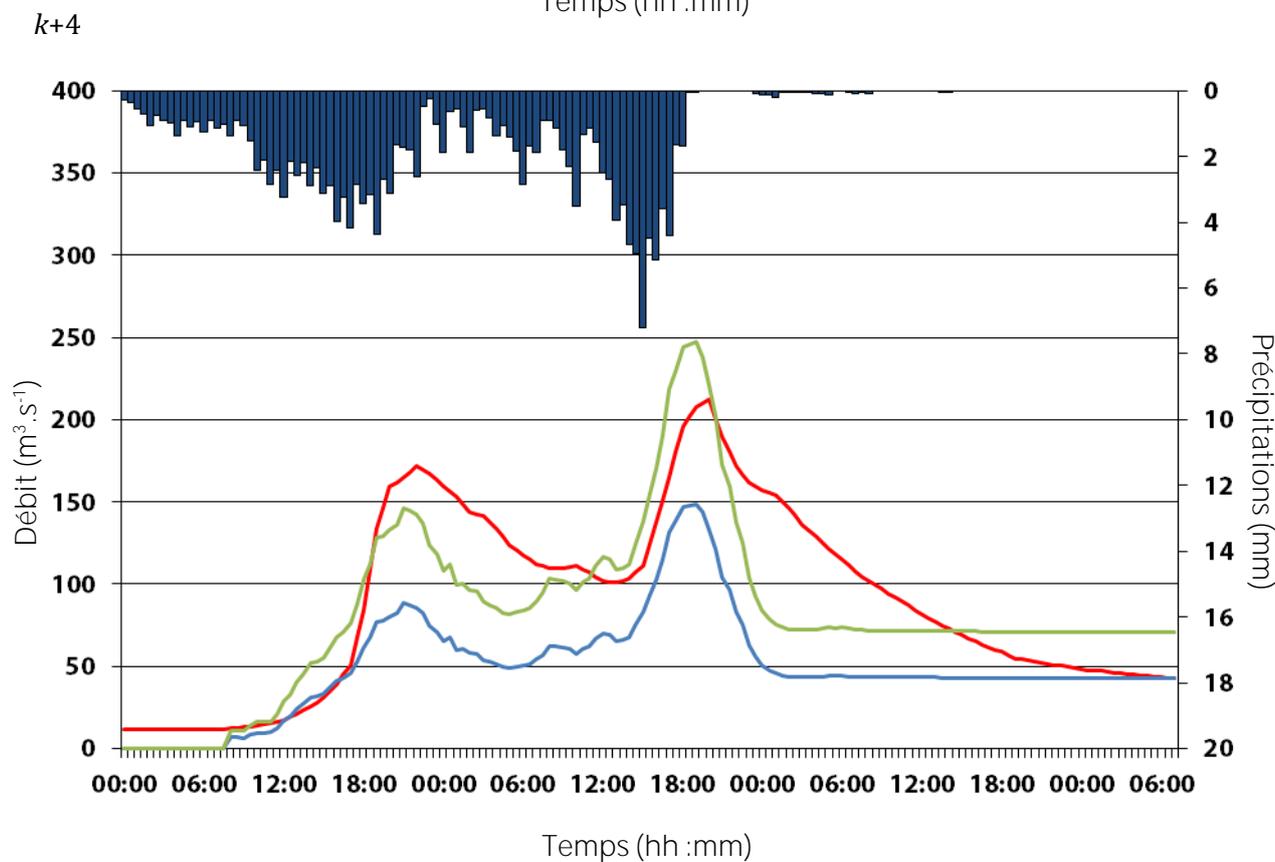
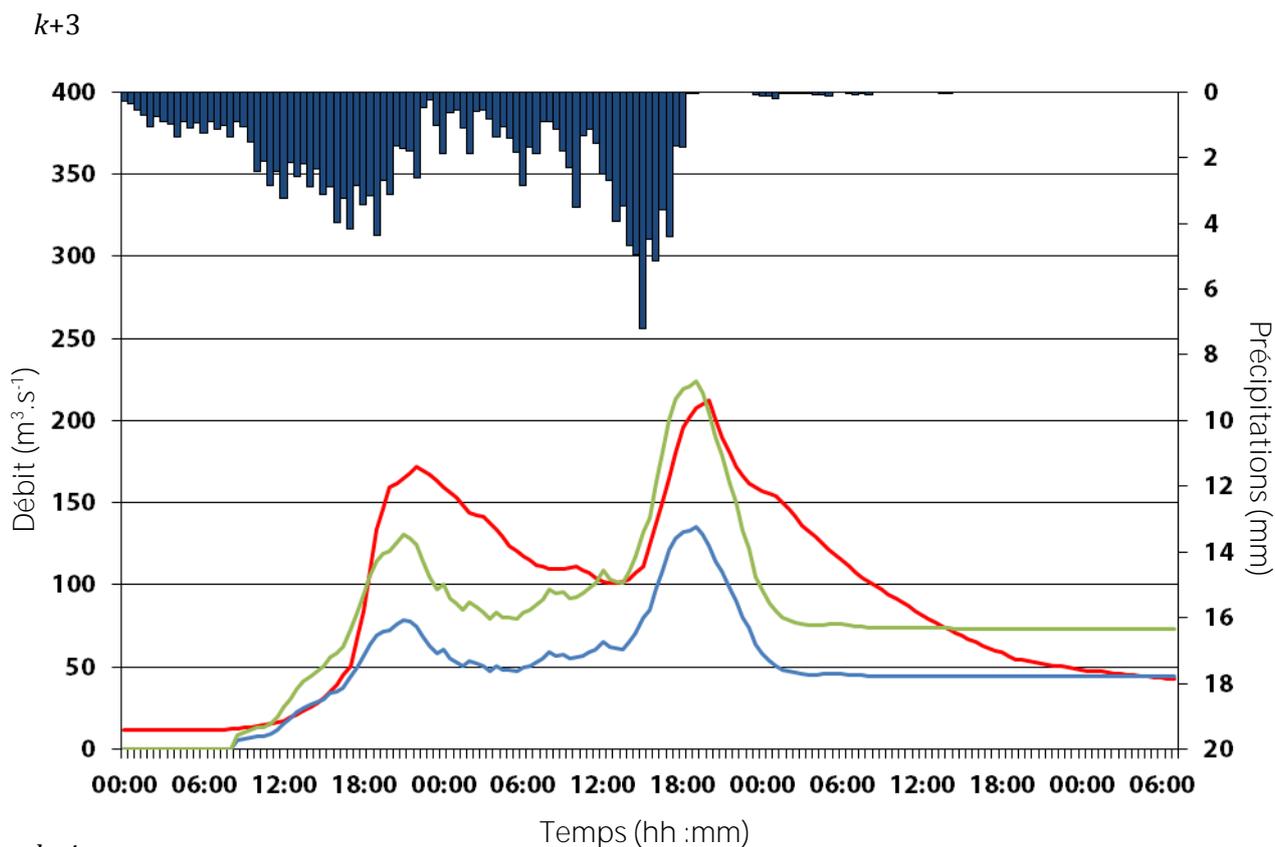
Bassin de la Baume, évènement n°904 :

k+1



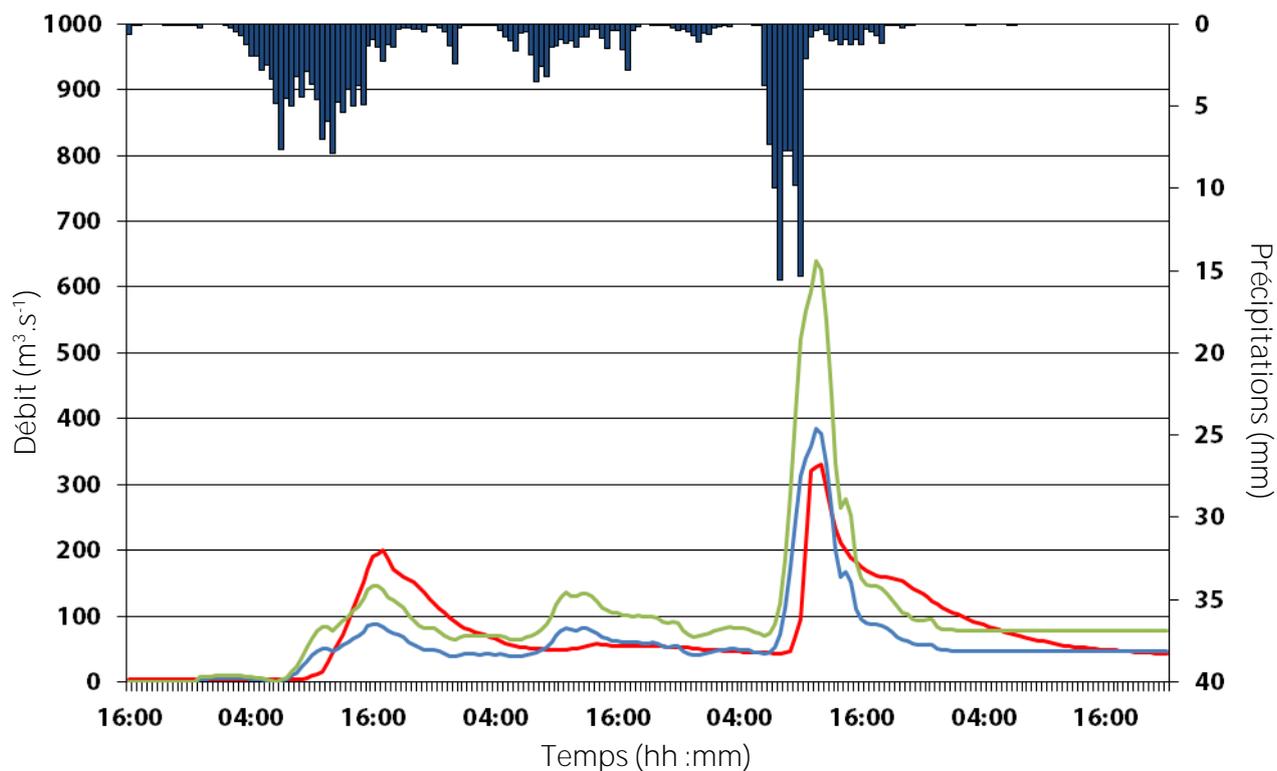
k+2



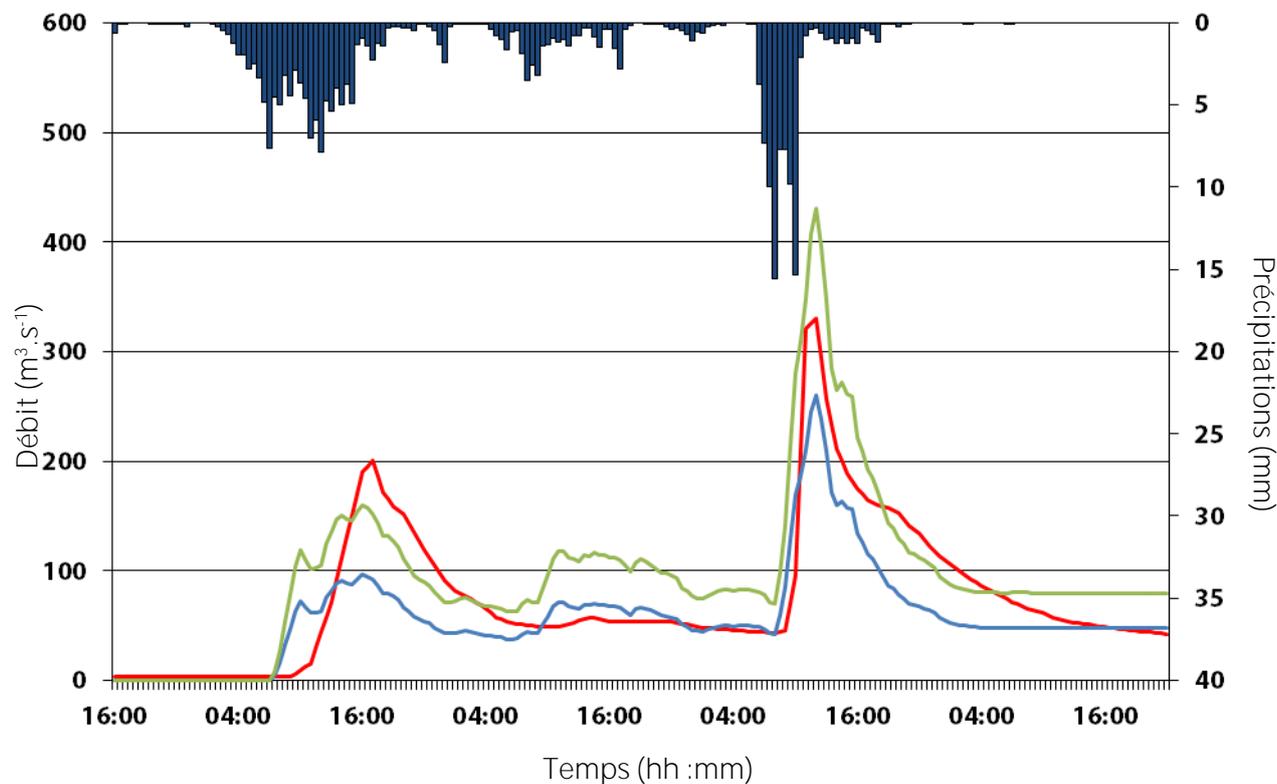


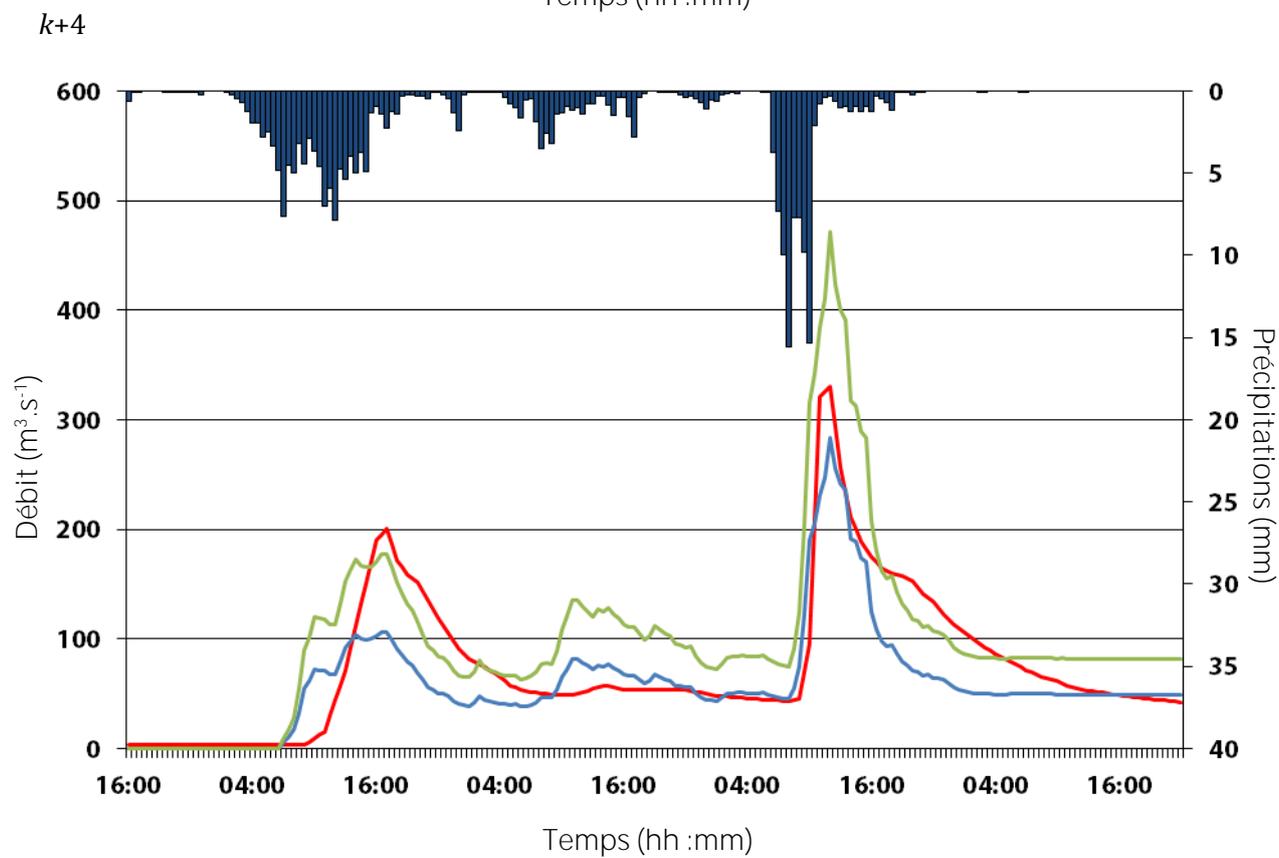
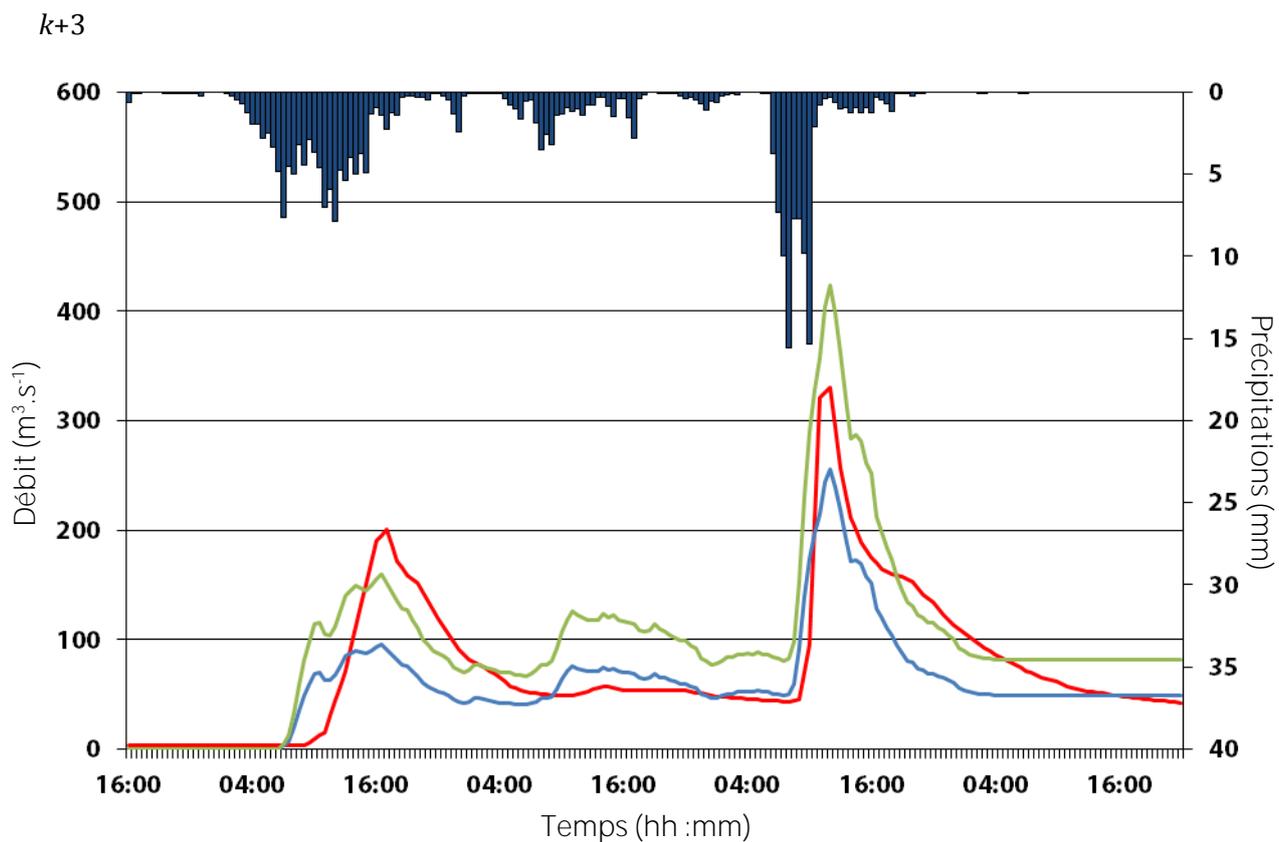
Bassin de la Baume, évènement n°905 :

k+1



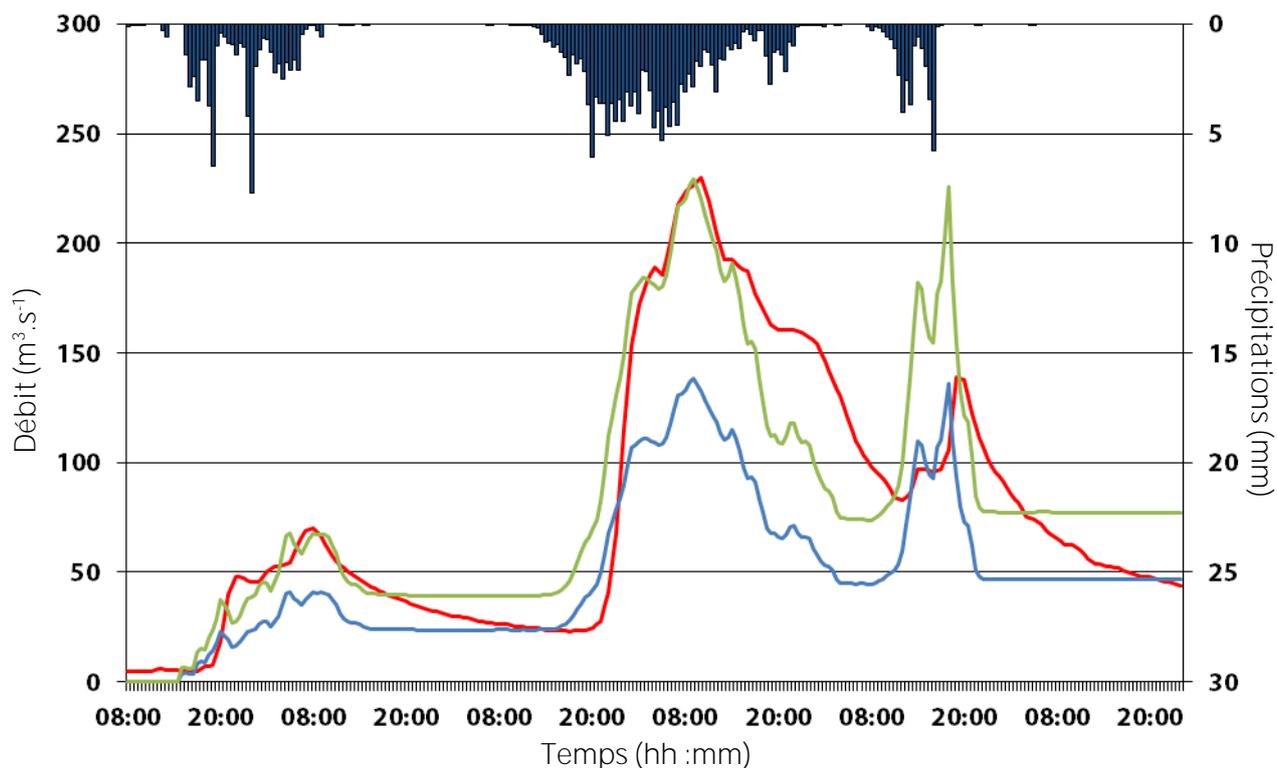
k+2



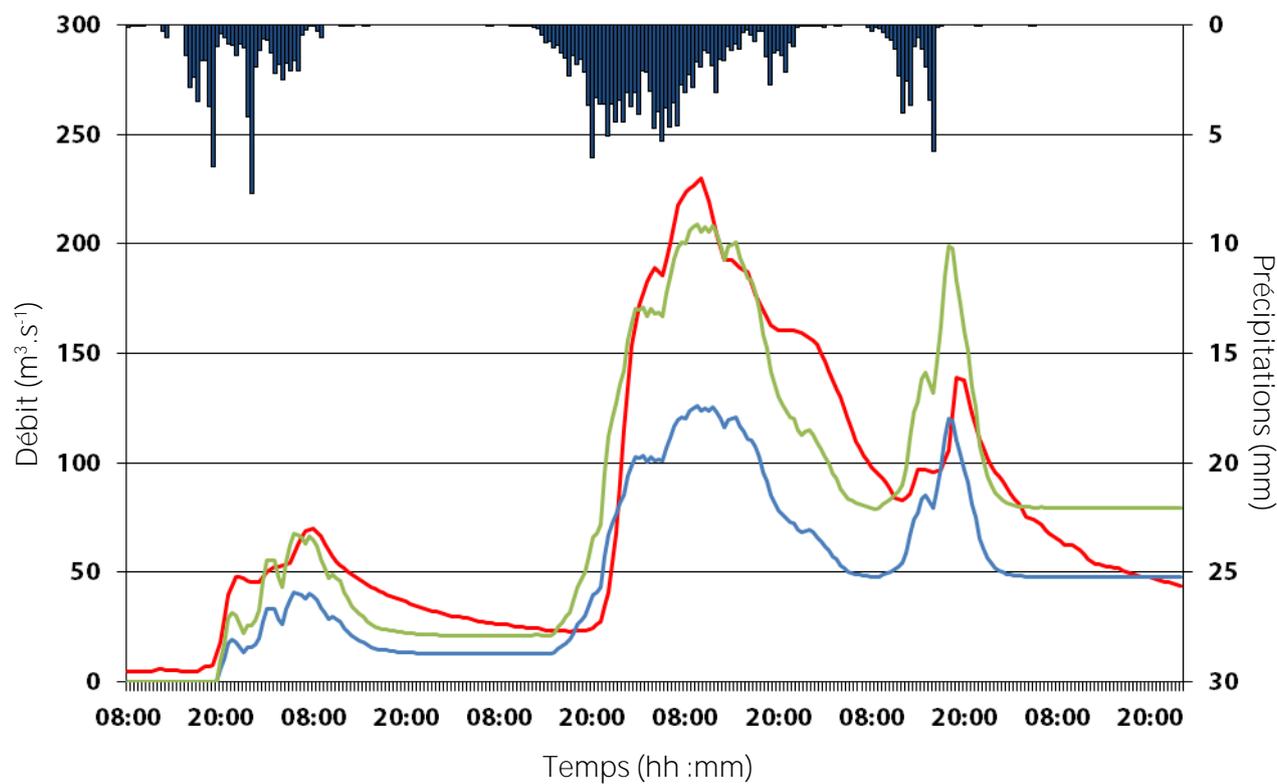


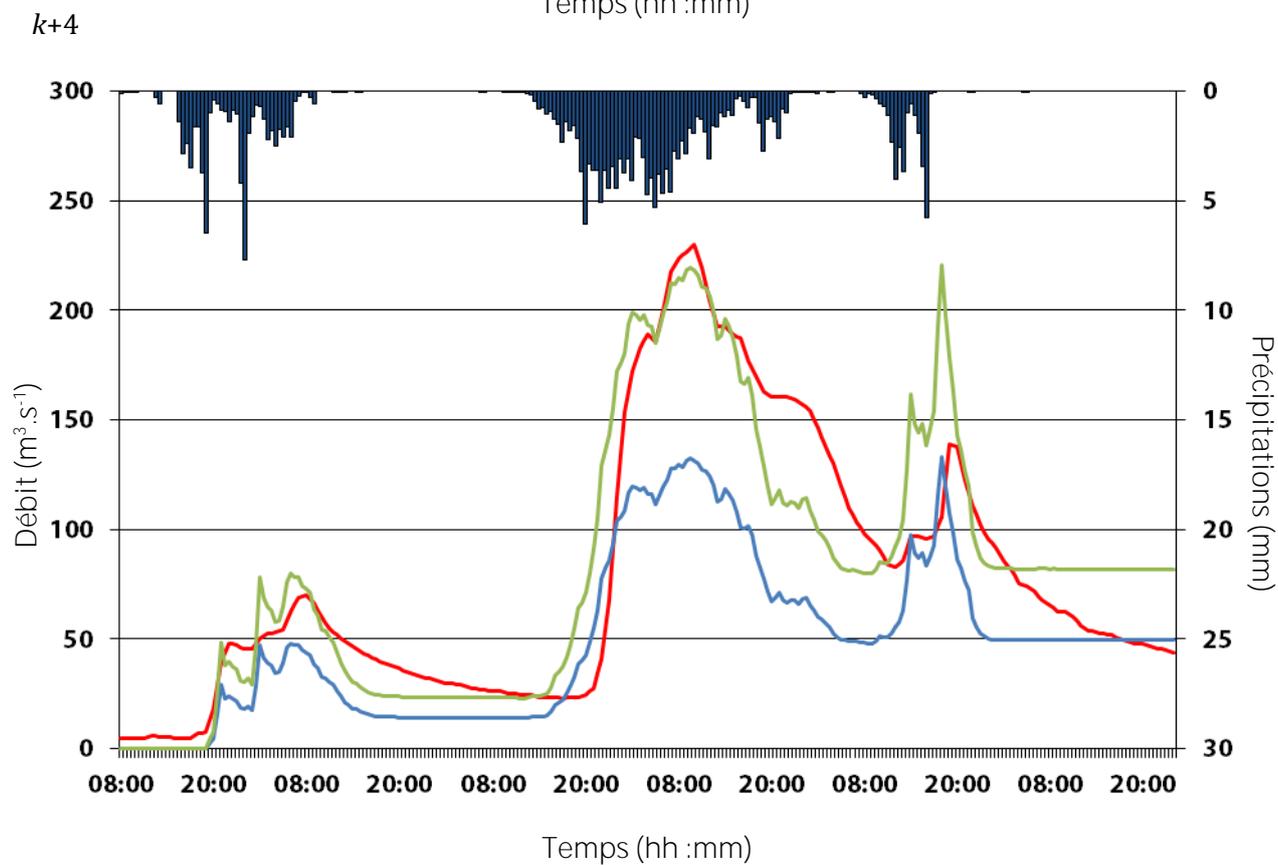
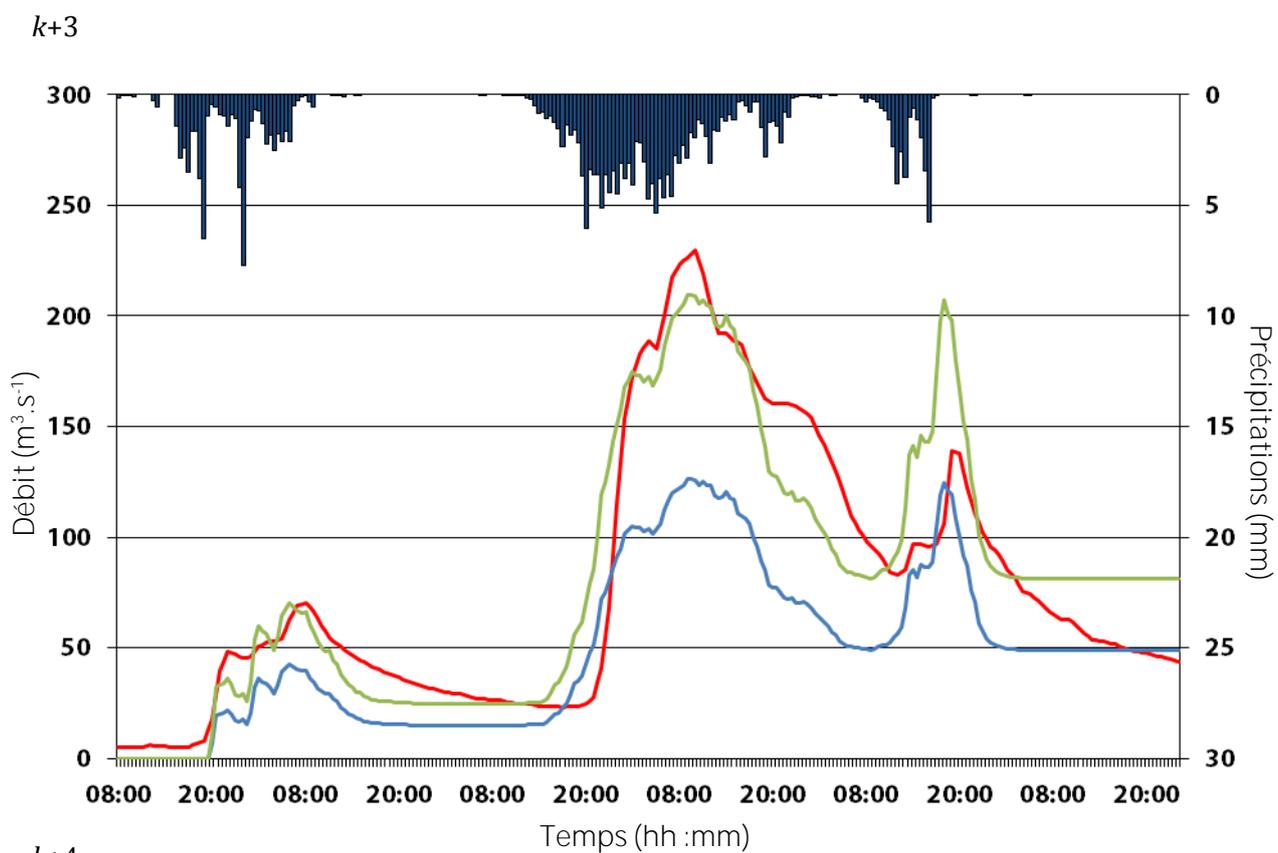
Bassin de la Baume, évènement n°906 :

k+1



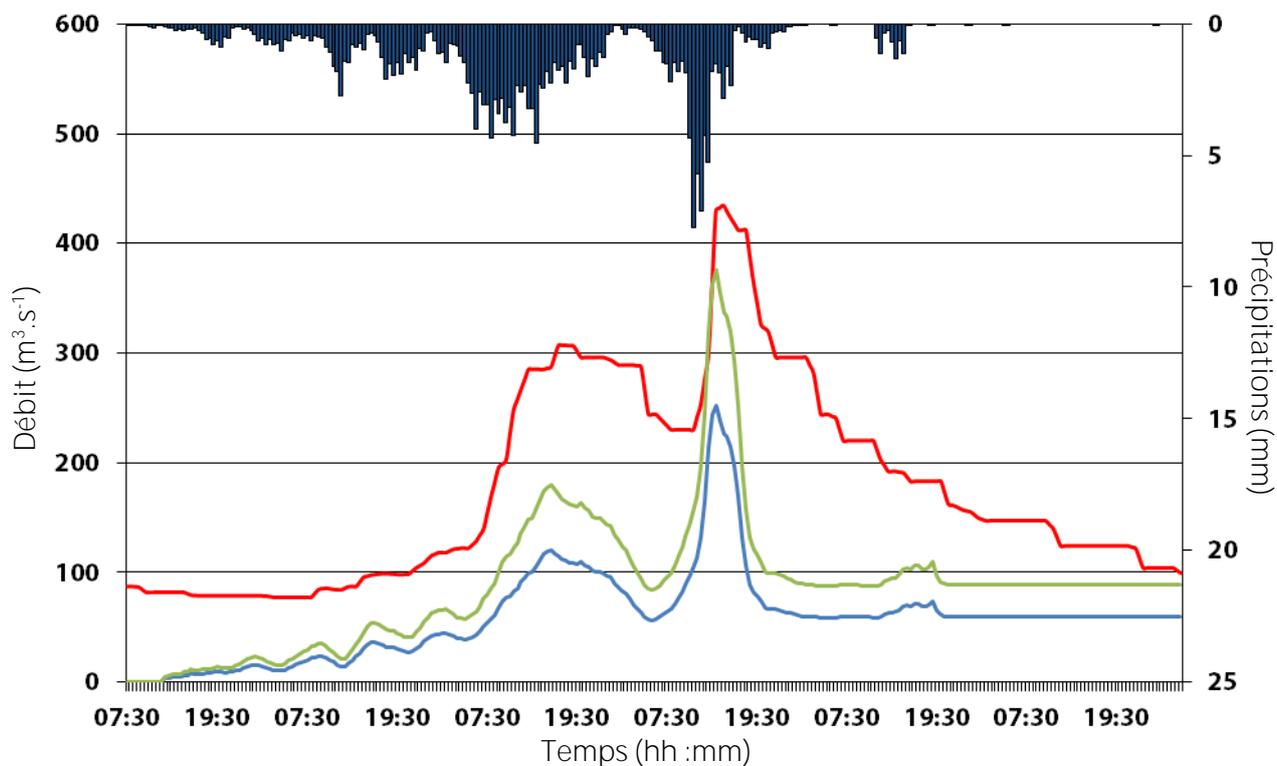
k+2



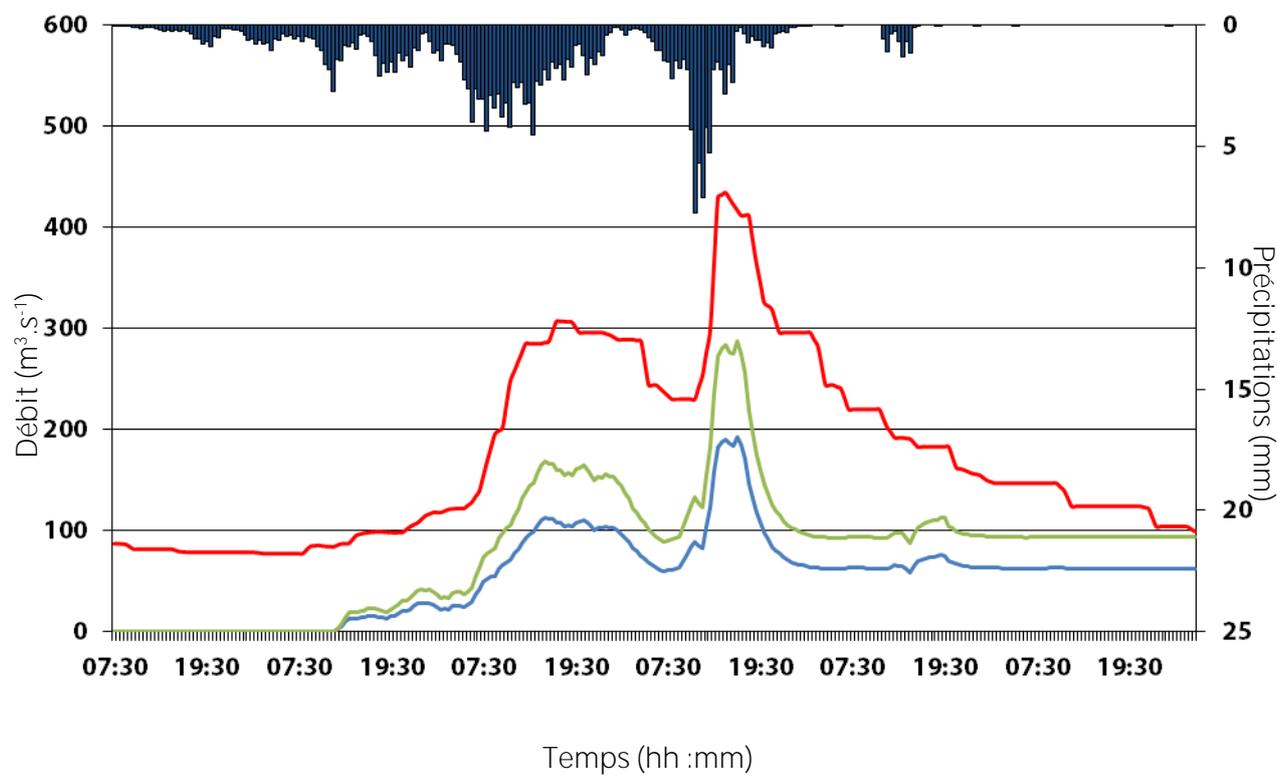


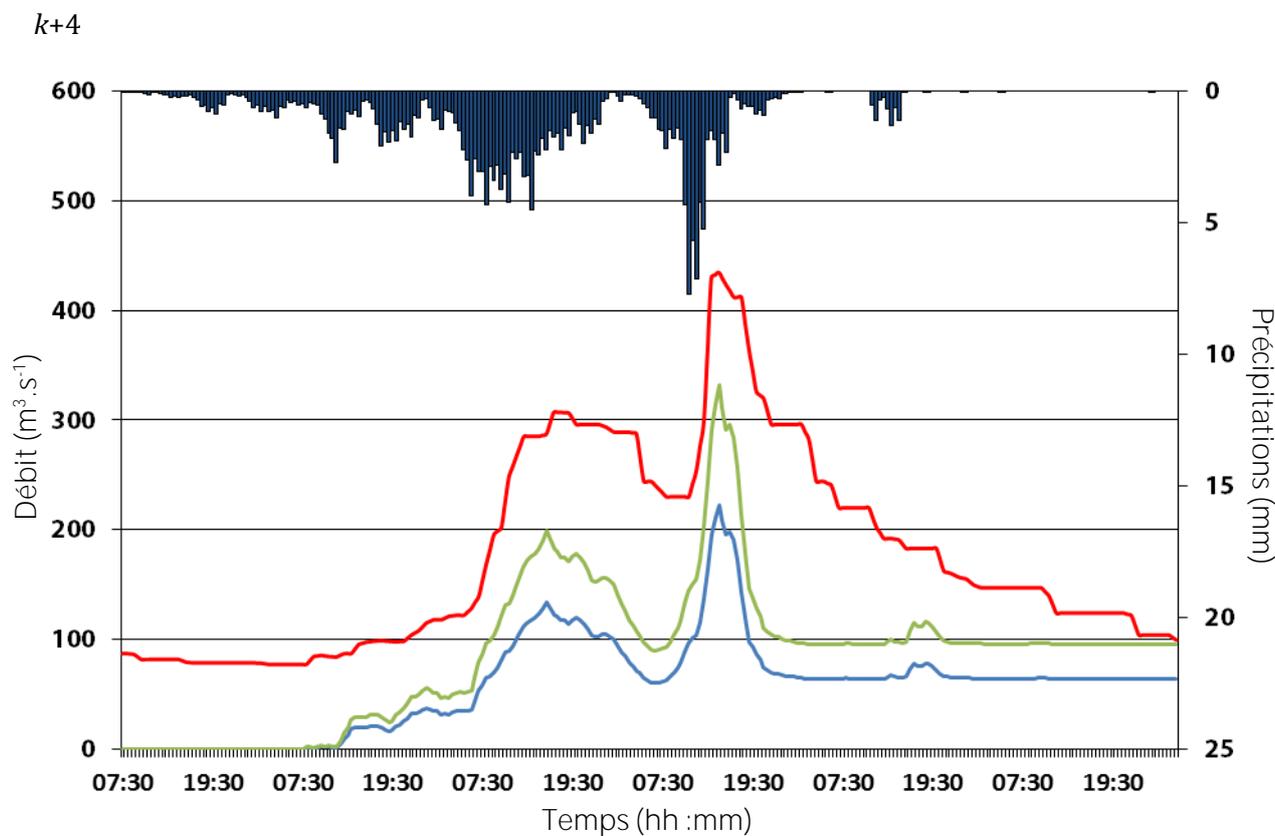
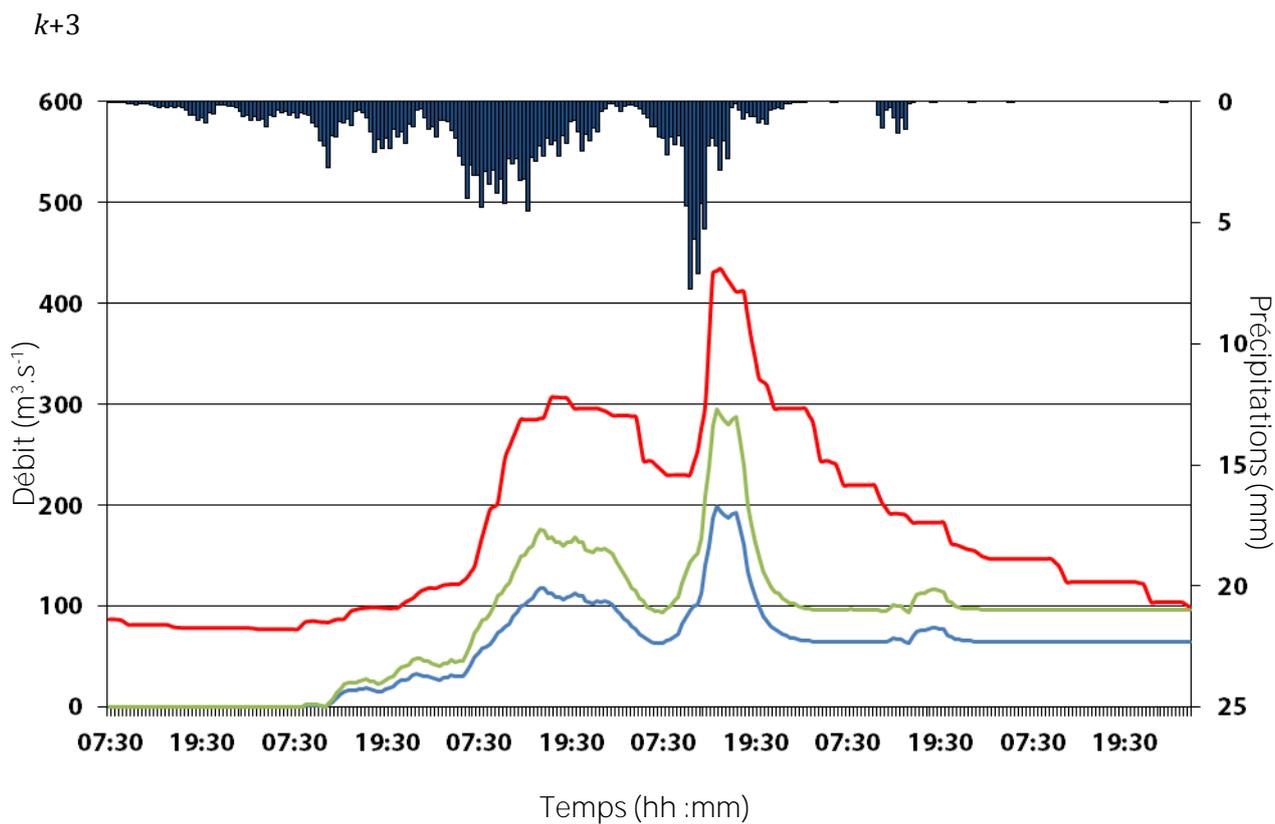
Bassin de l'Ardèche, évènement n°901 :

$k+1$



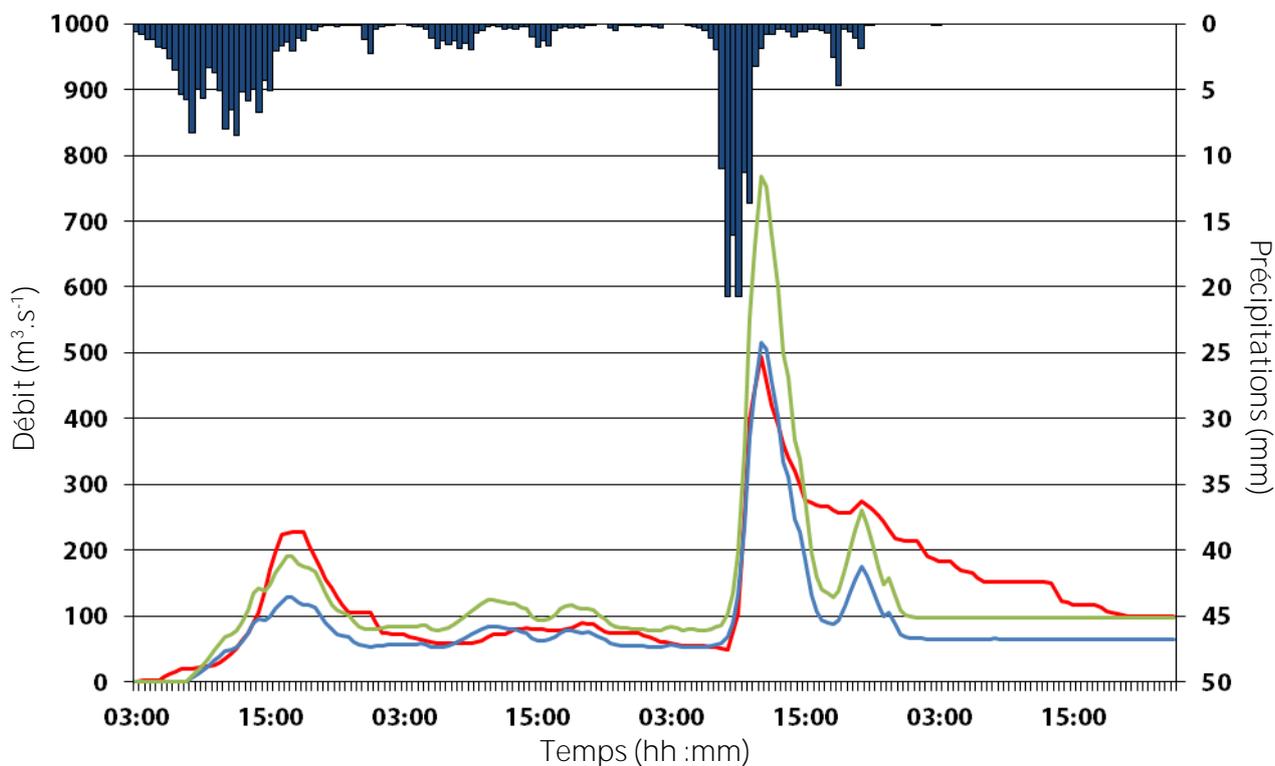
$k+2$



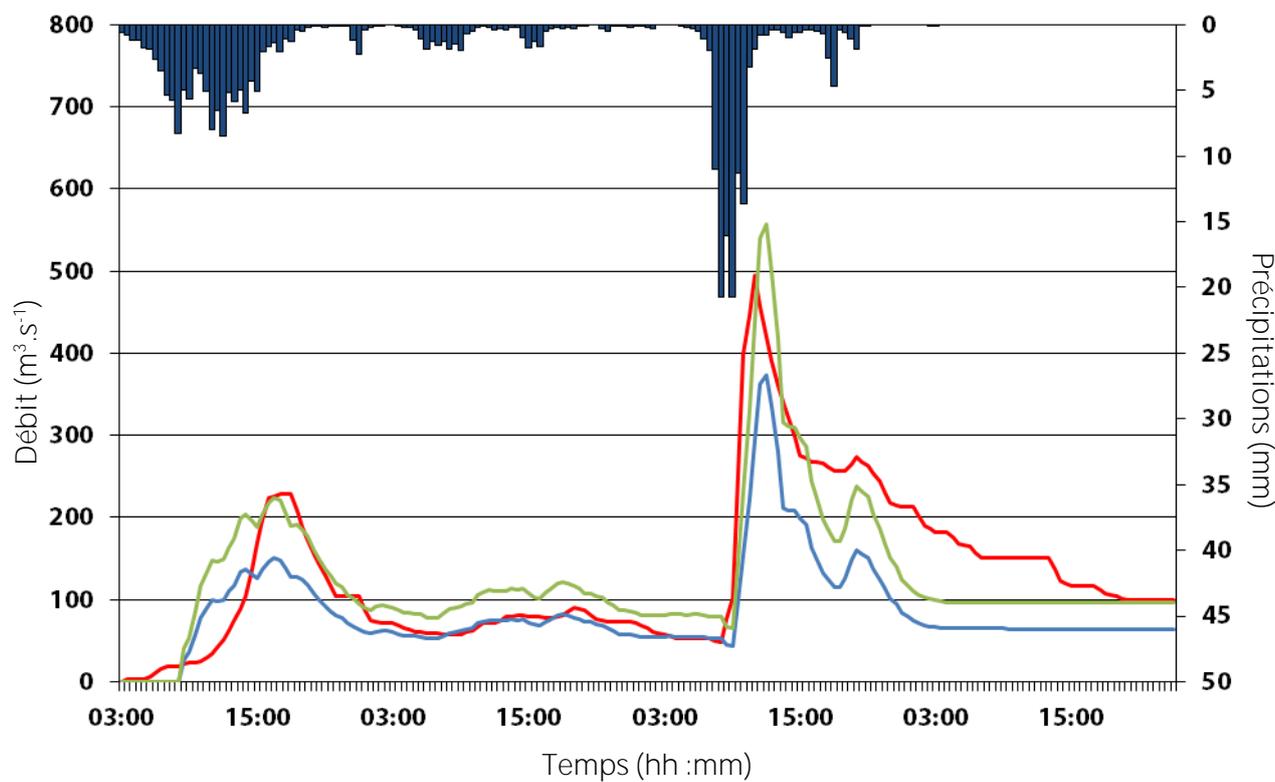


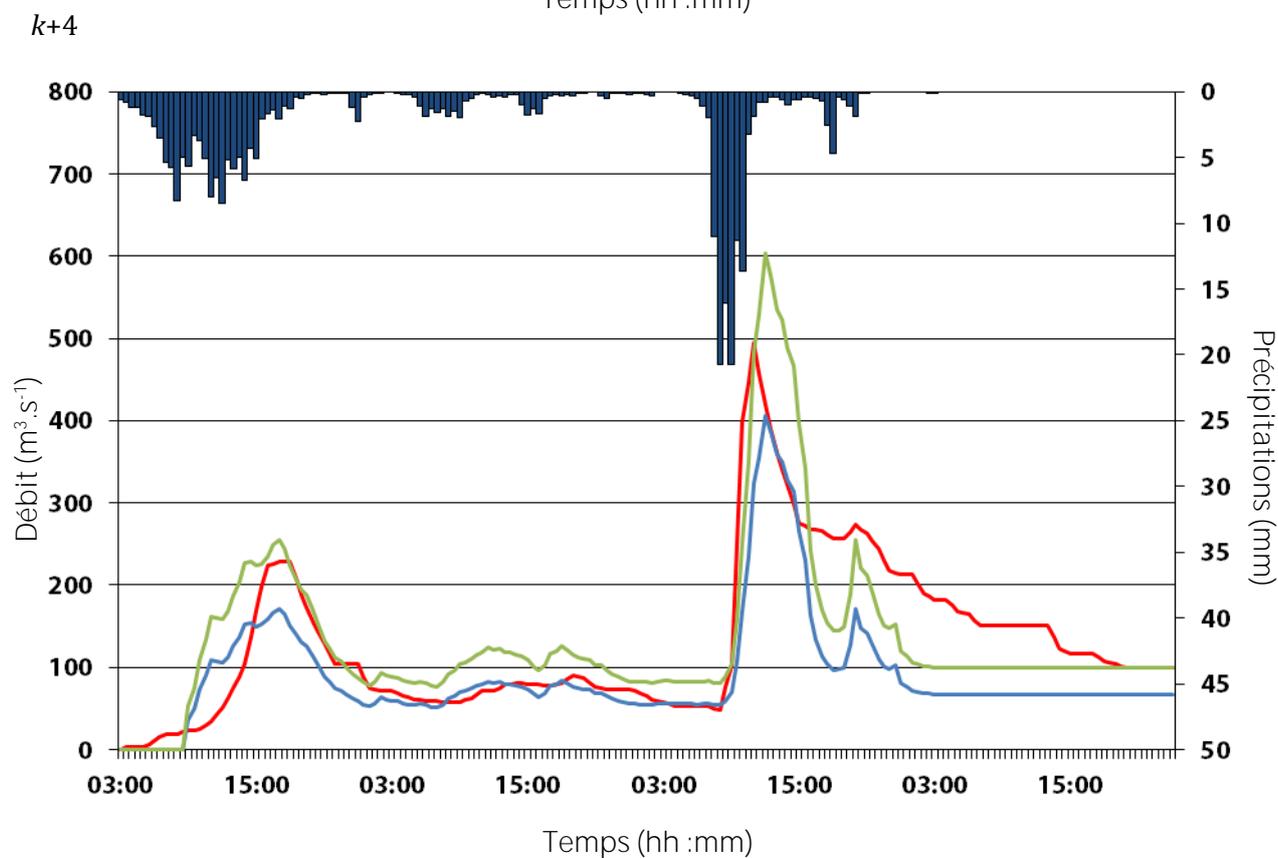
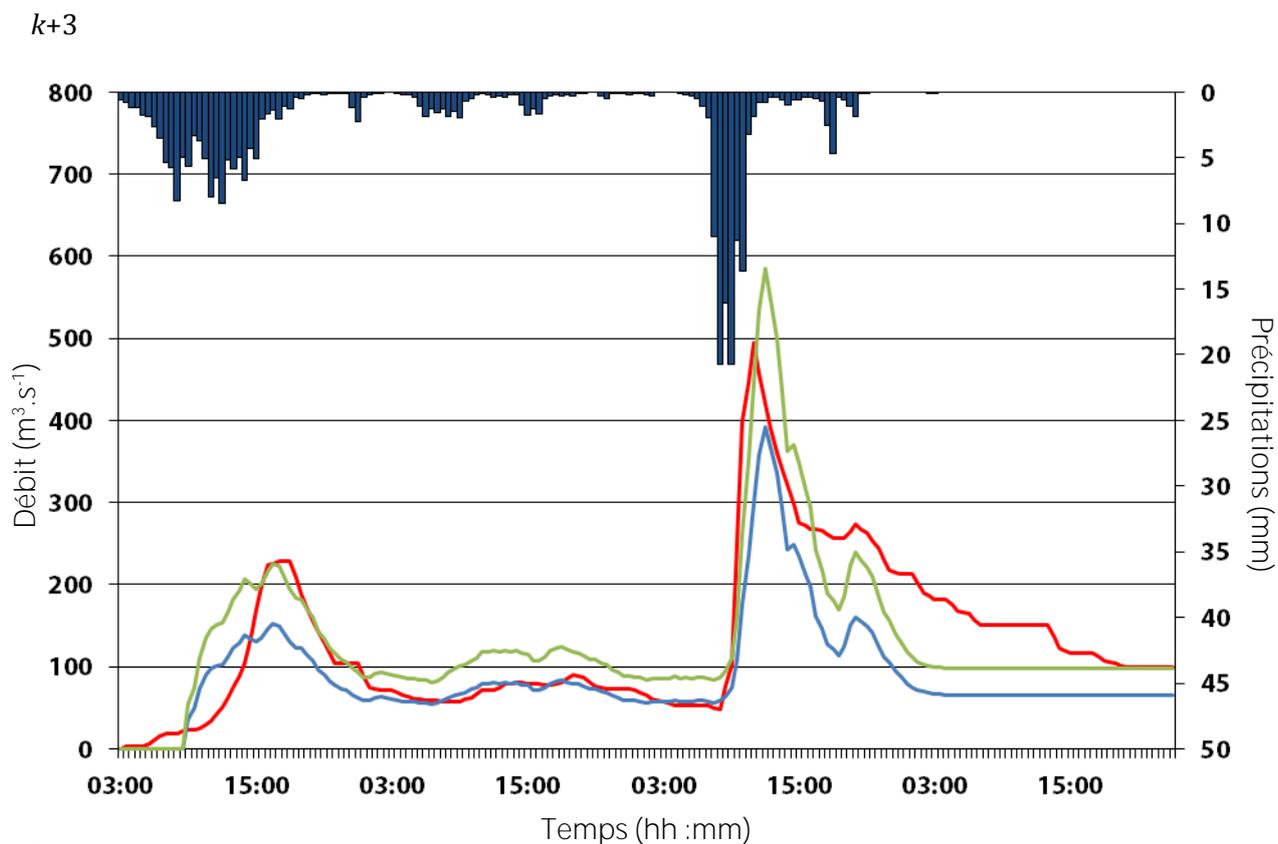
Bassin de l'Ardèche, évènement n°902 :

k+1



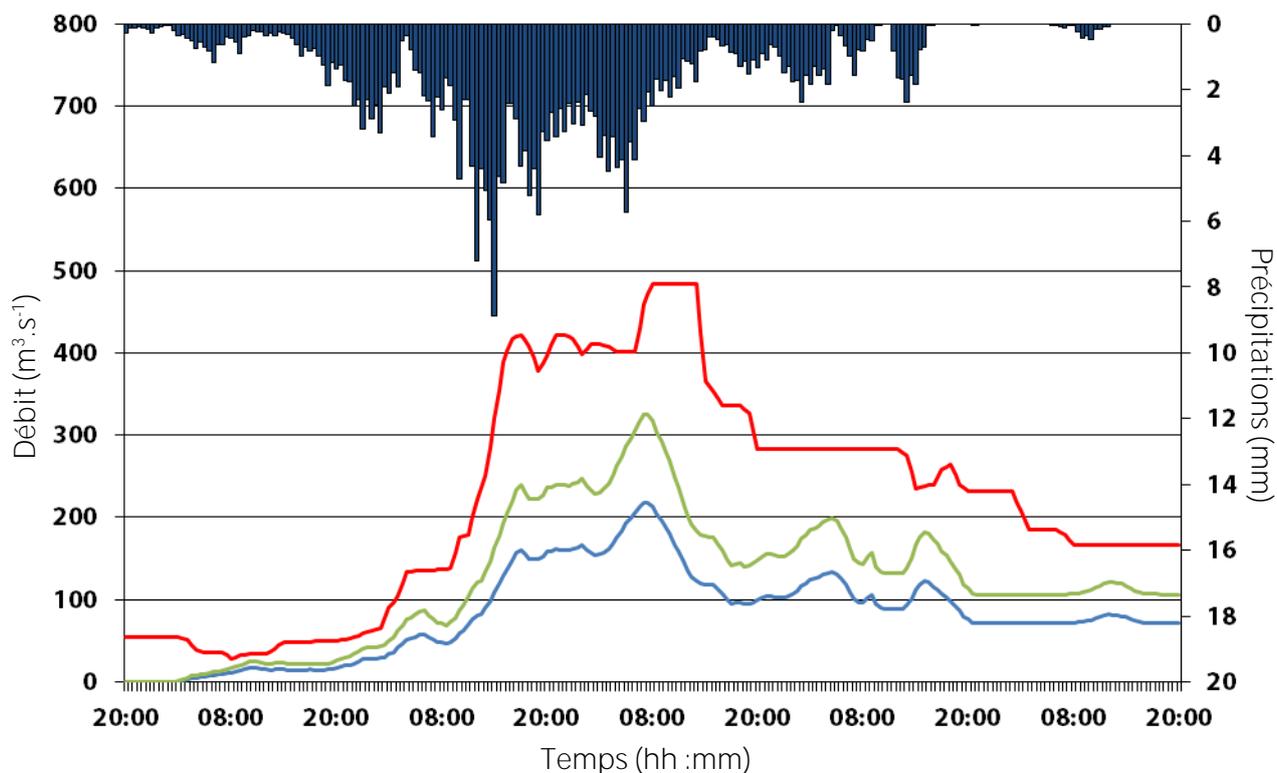
k+2



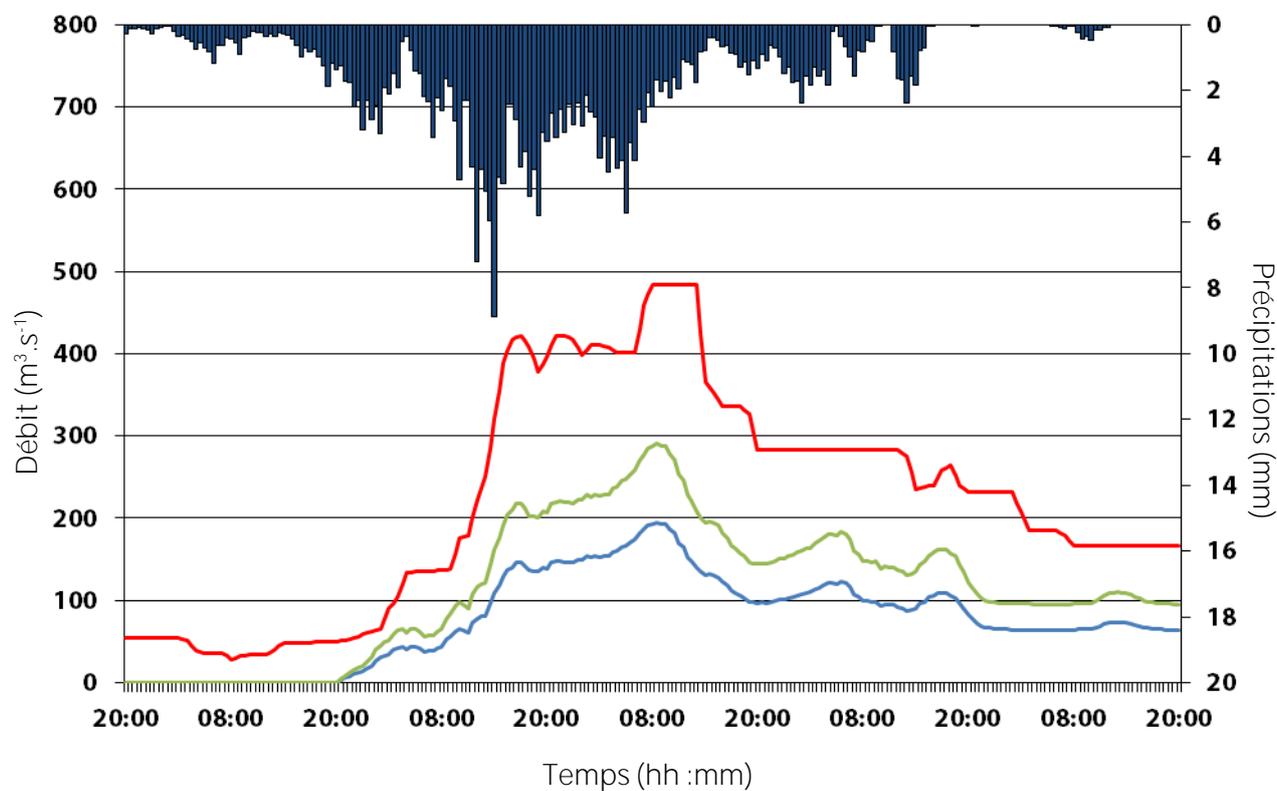


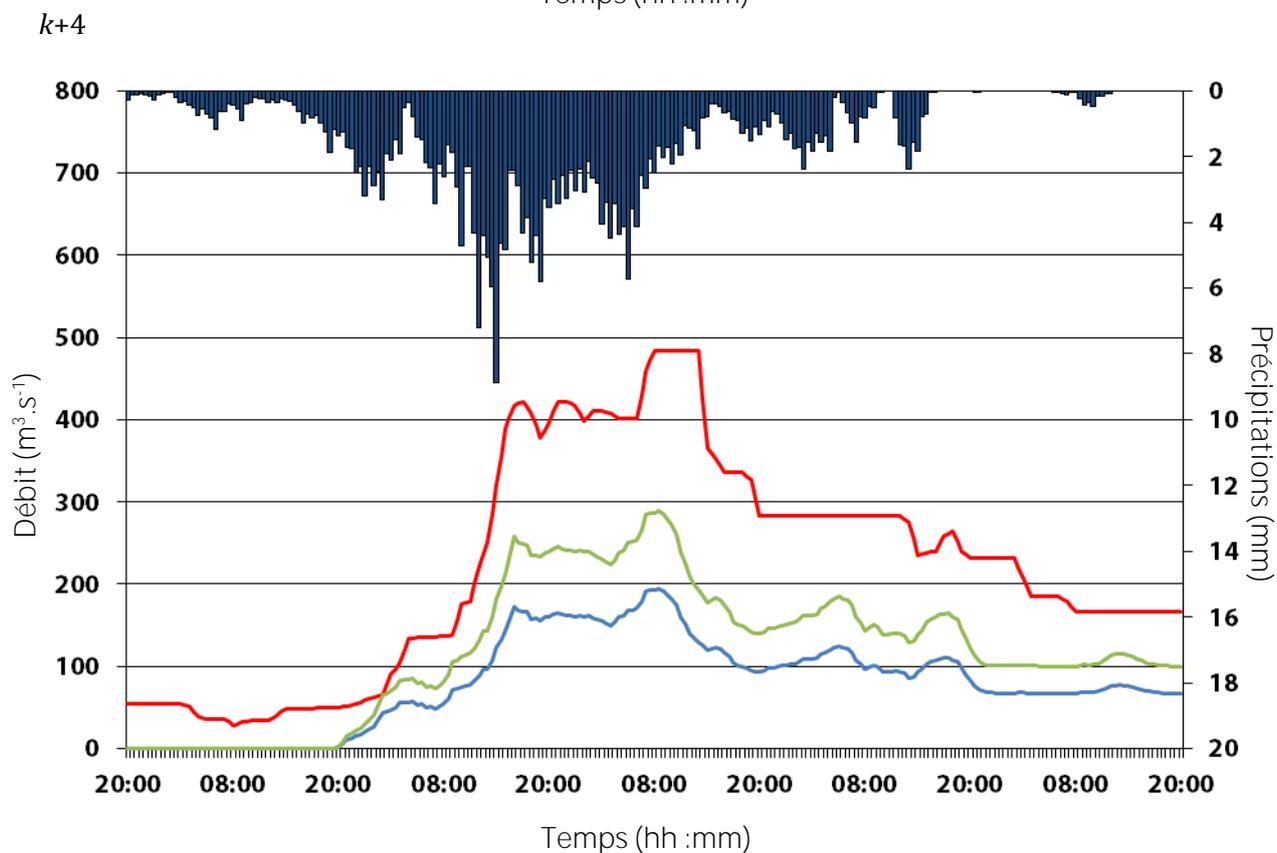
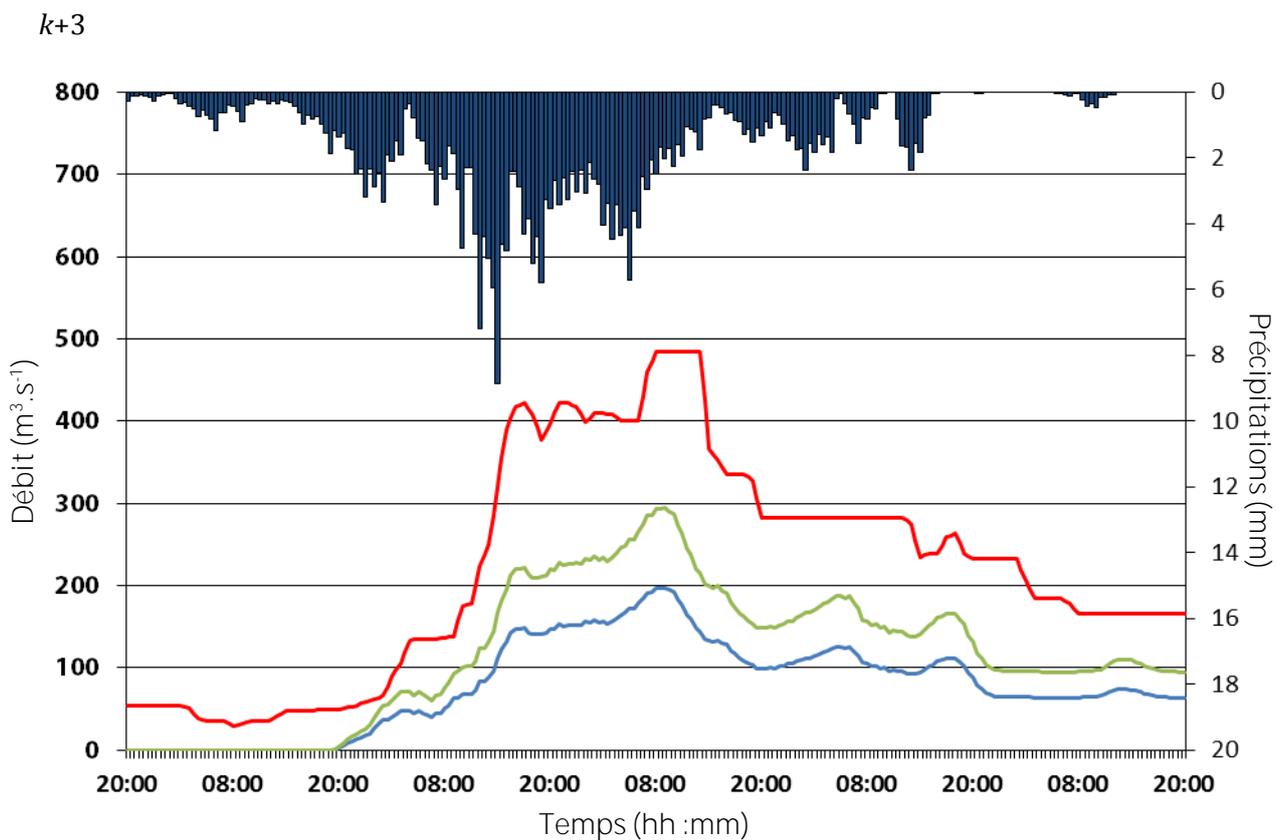
Bassin de l'Ardèche, évènement n°903 :

k+1



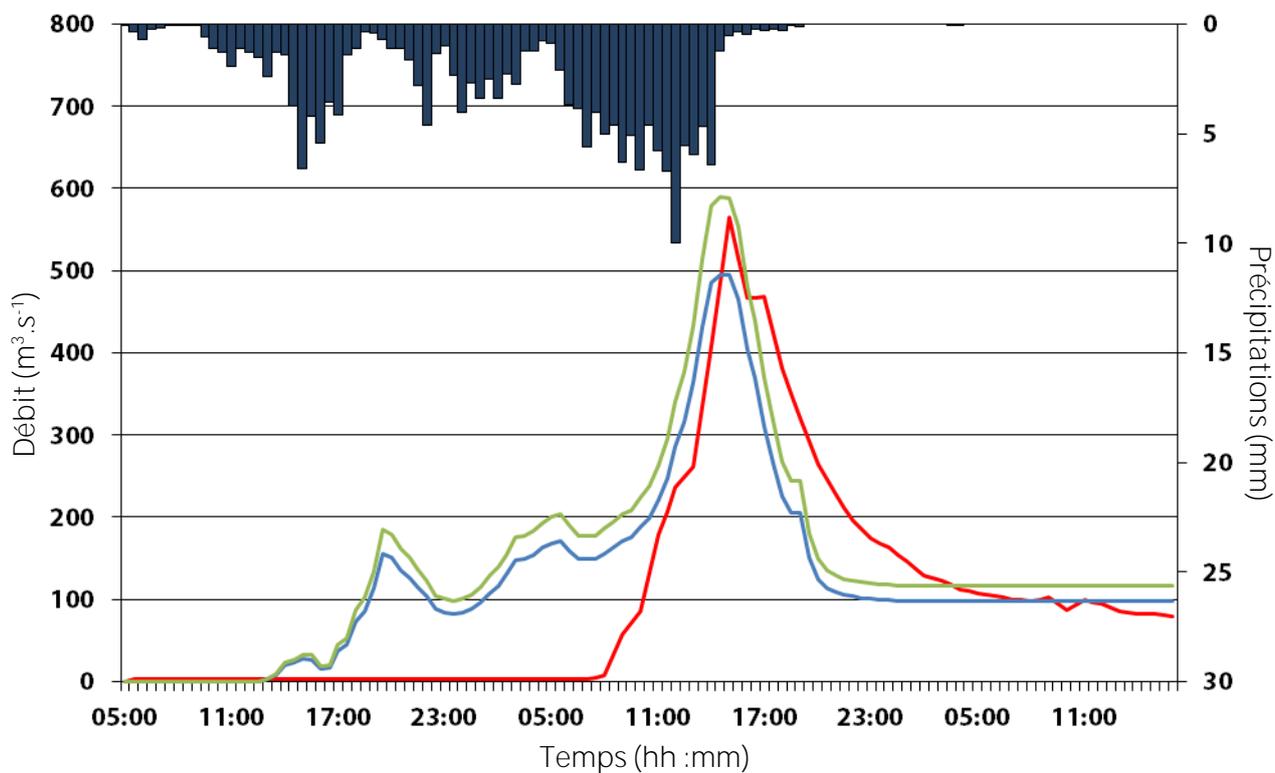
k+2



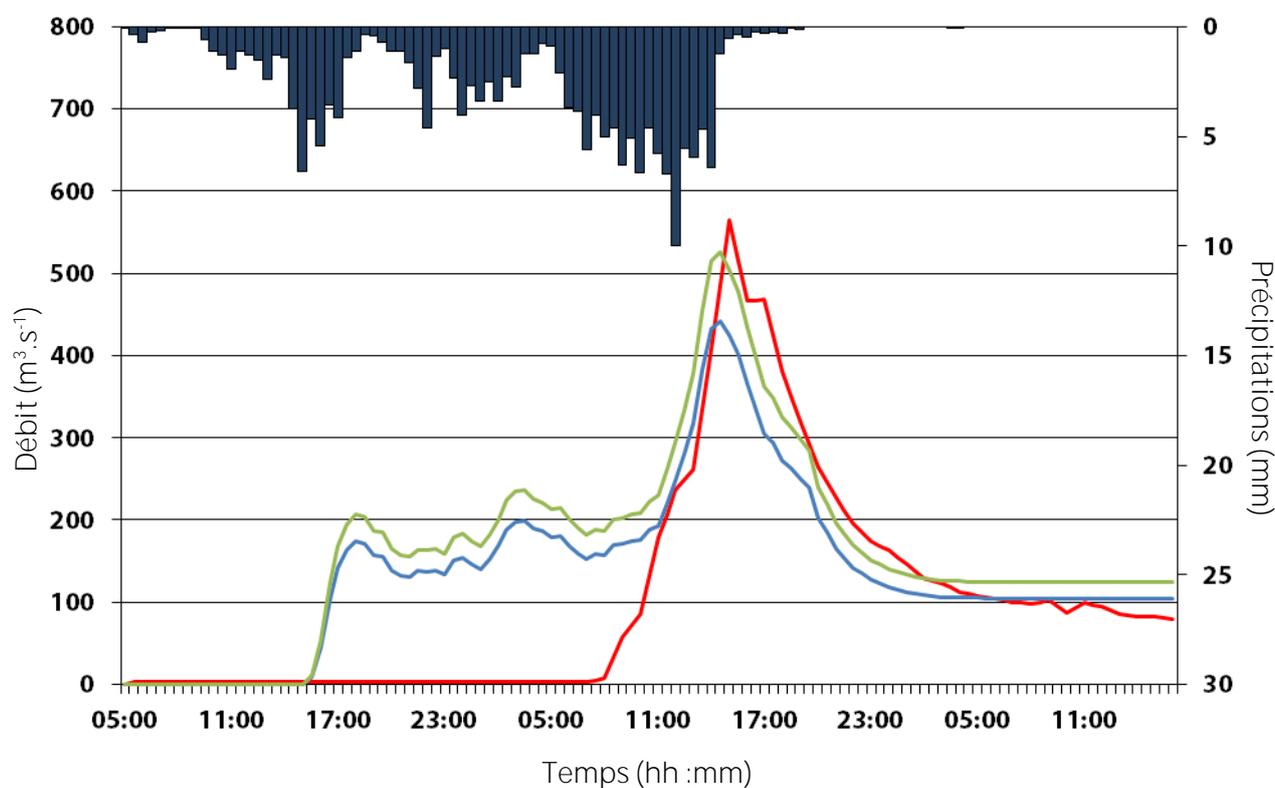


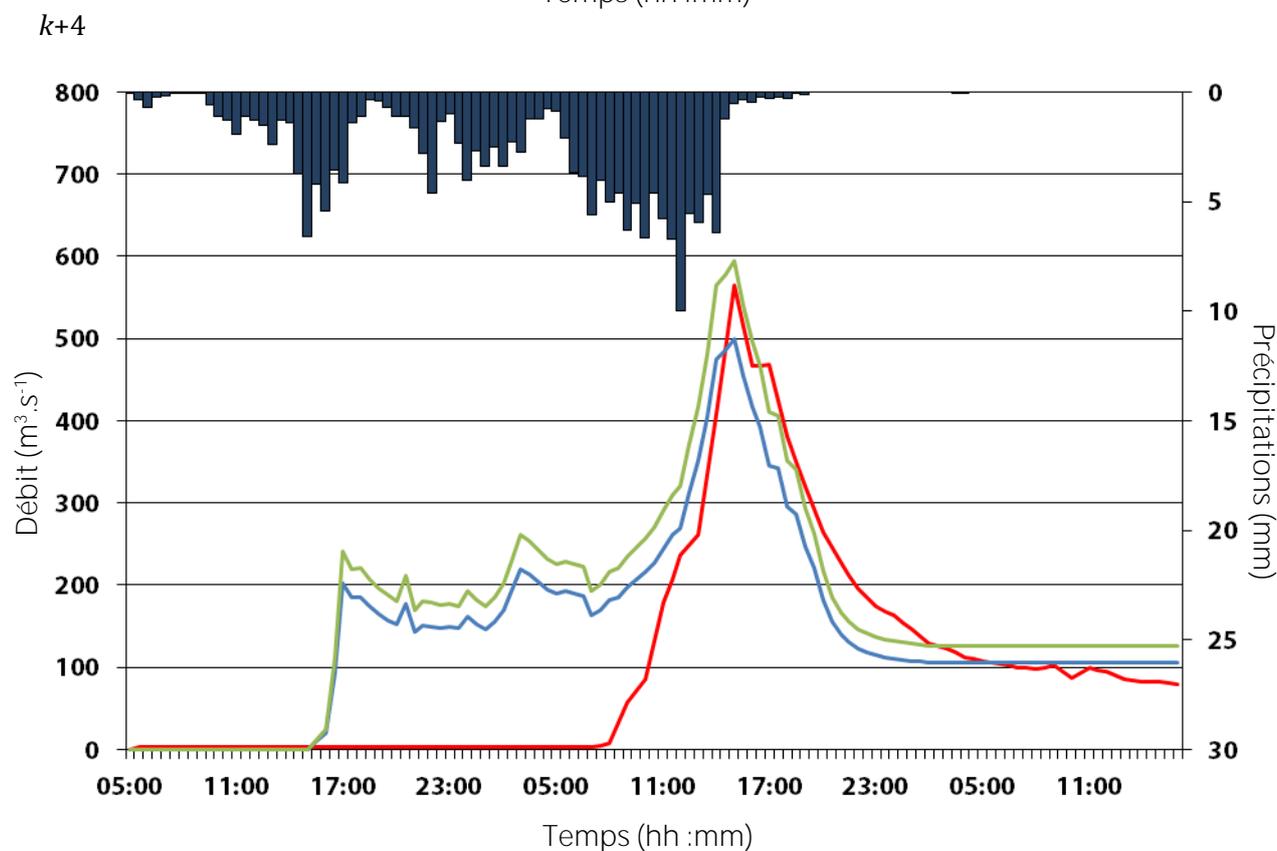
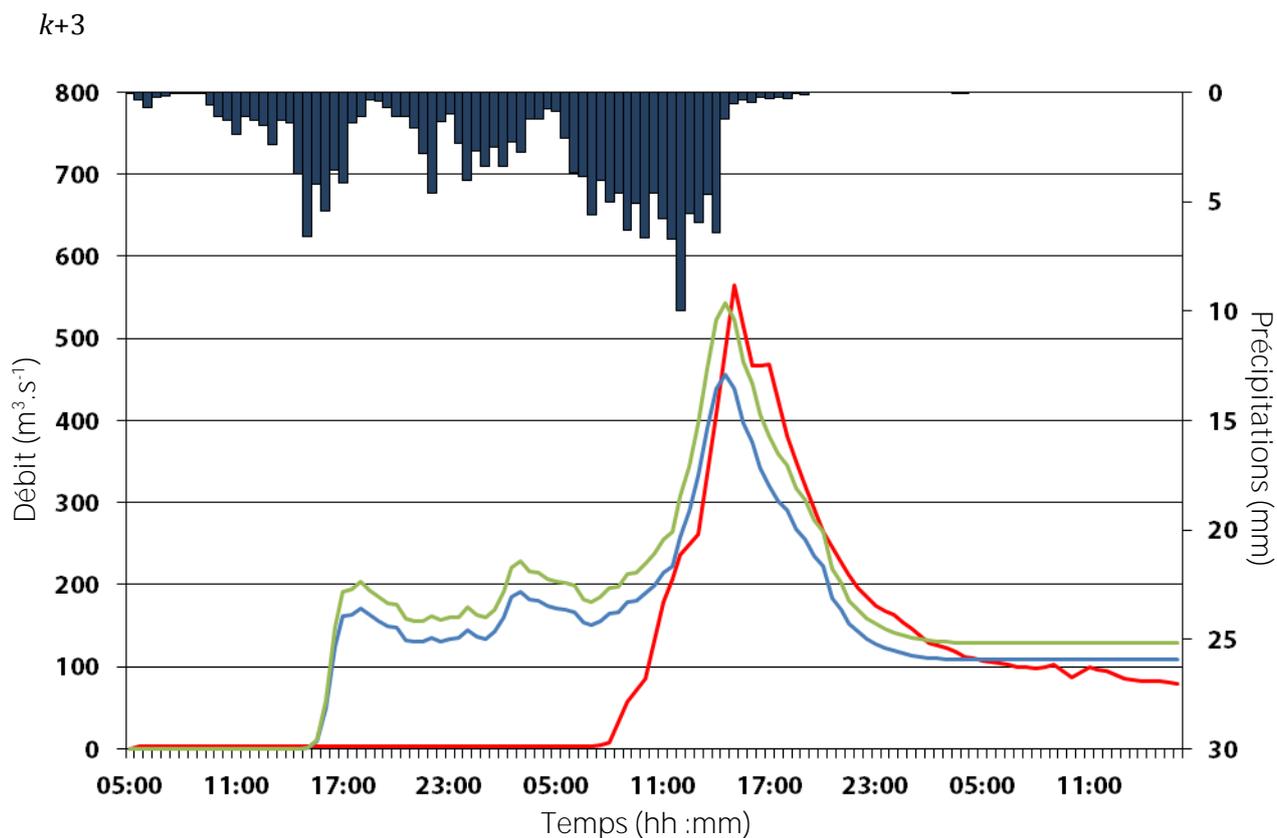
Bassin du Chassezac, évènement n°907 :

k+1



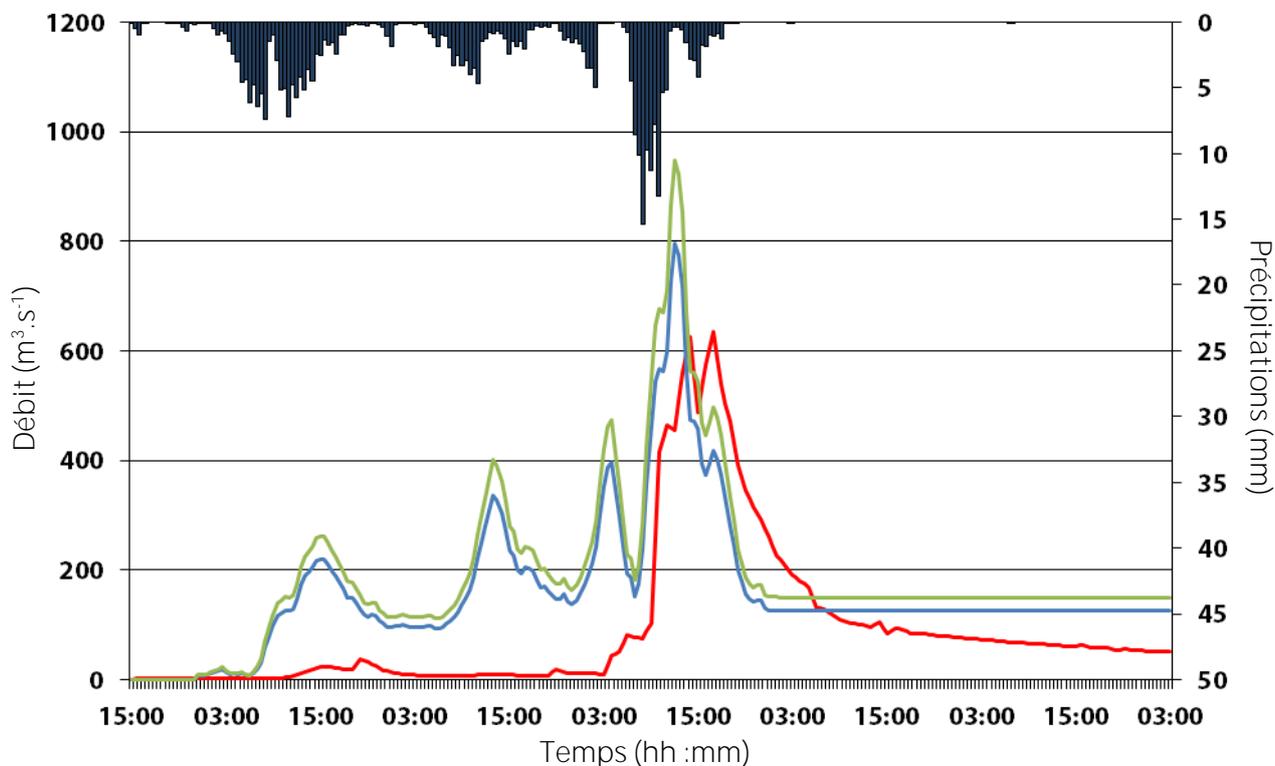
k+2



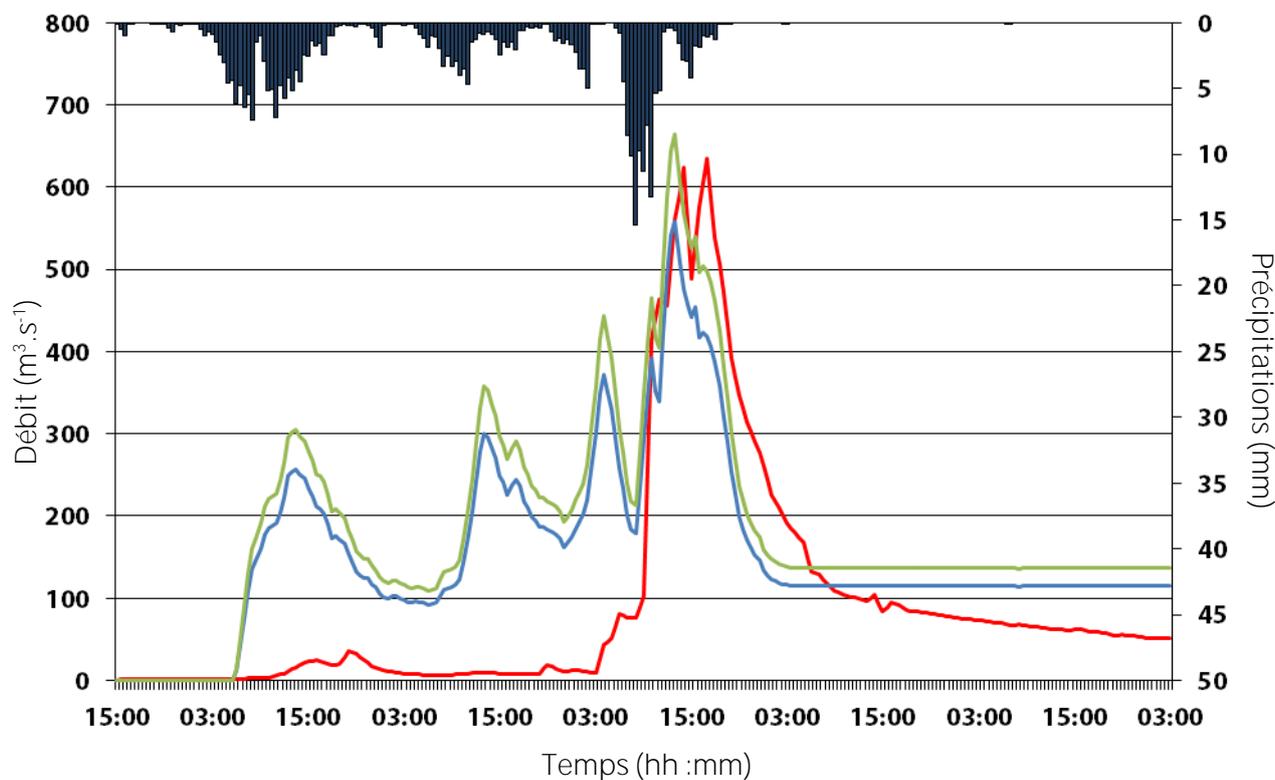


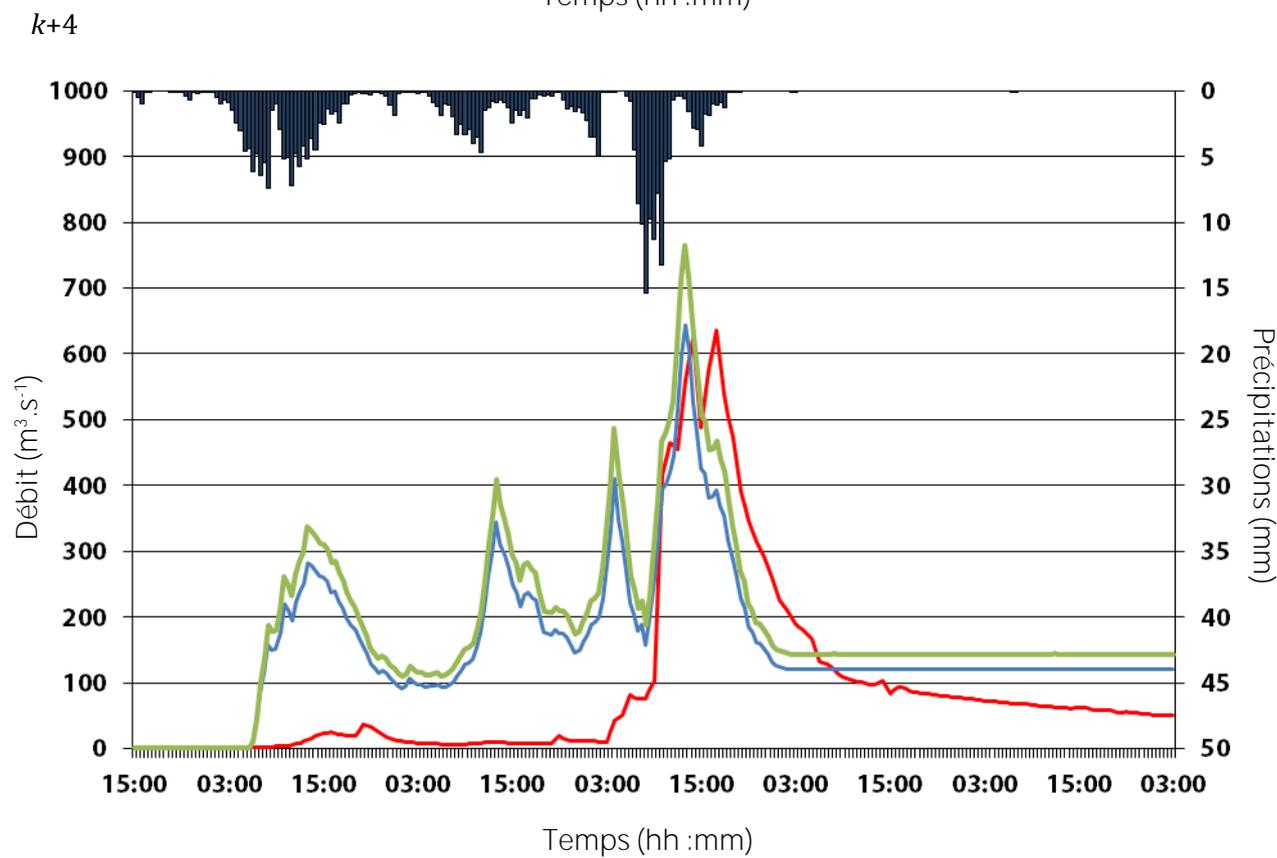
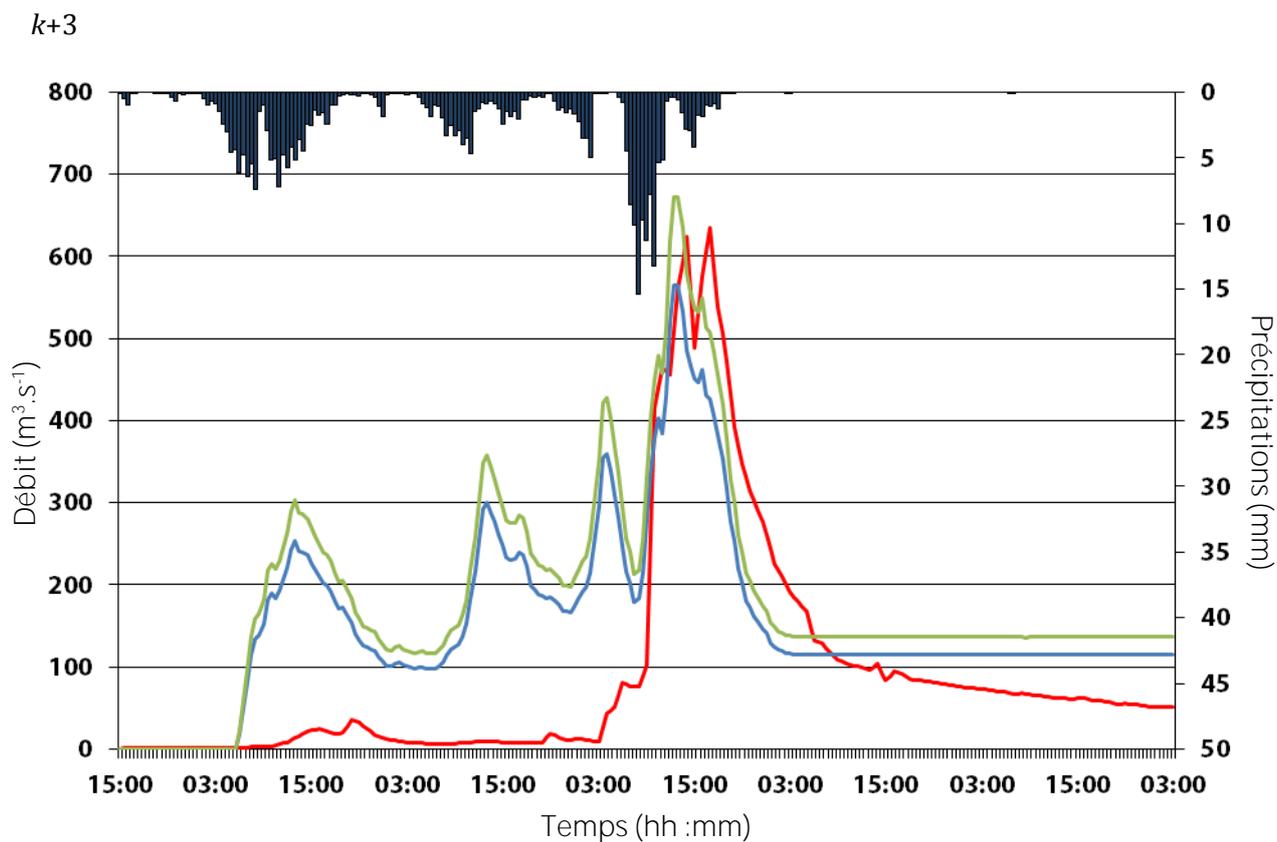
Bassin du Chassezac, évènement n°908 :

k+1



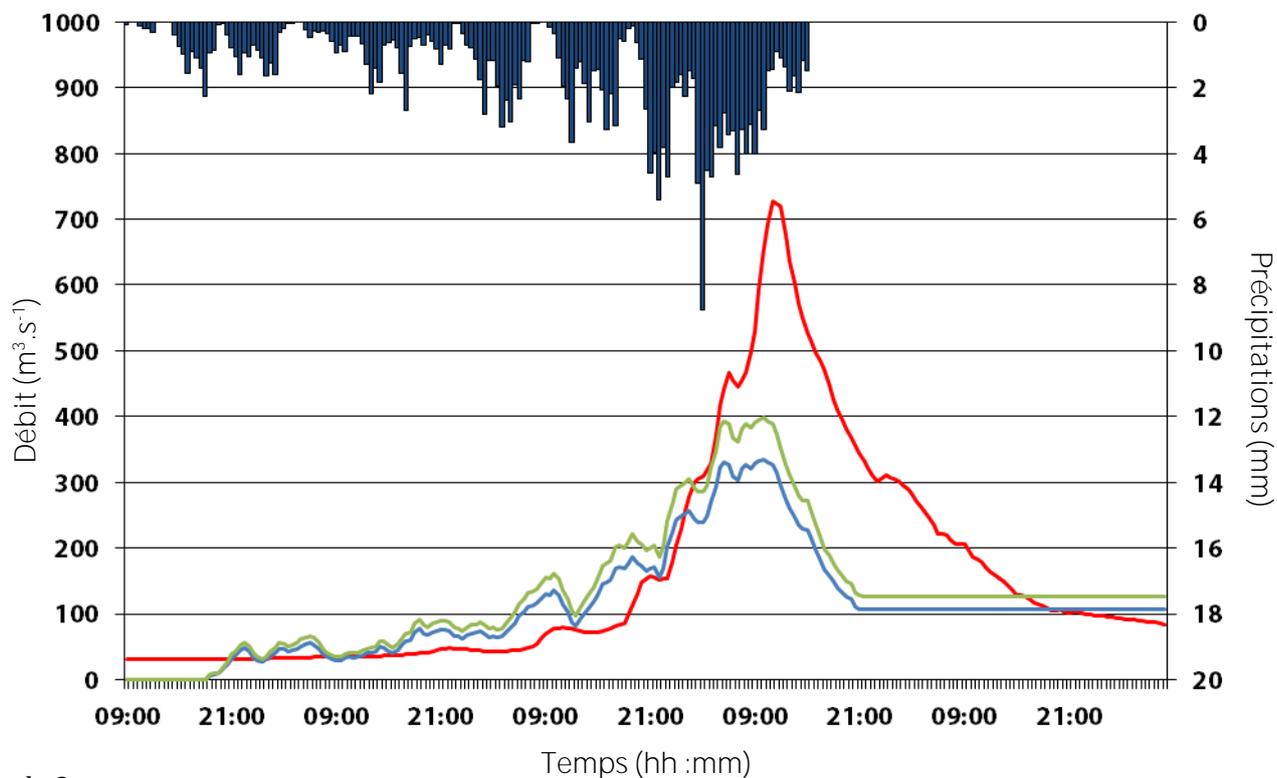
k+2



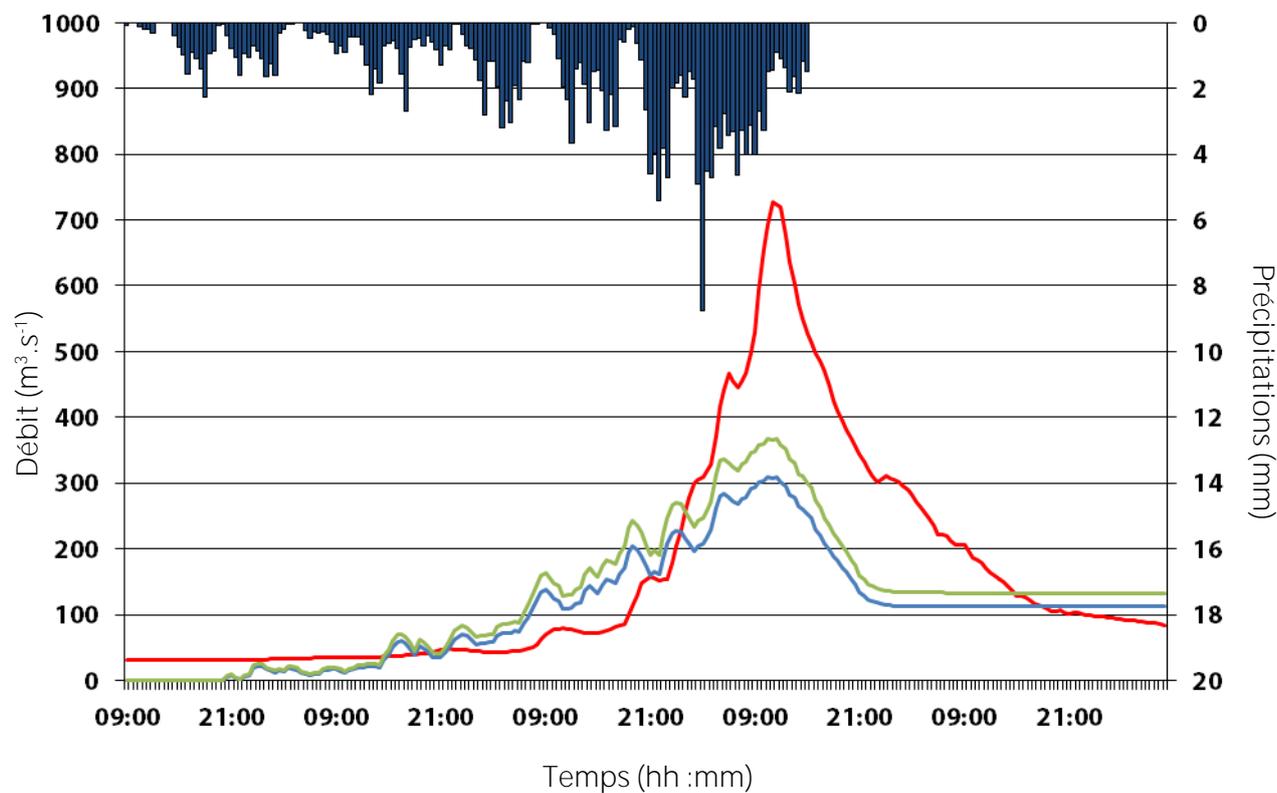


Bassin du Chassezac, évènement n°909 :

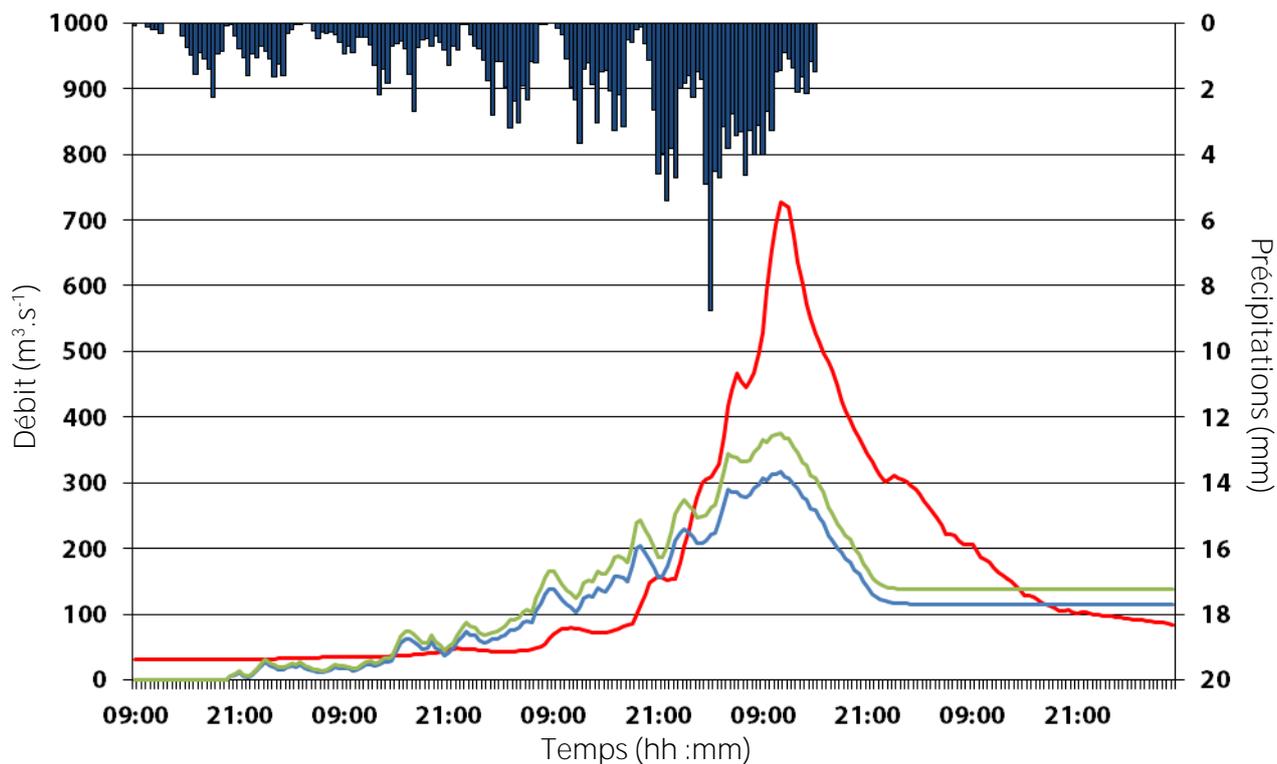
k+1



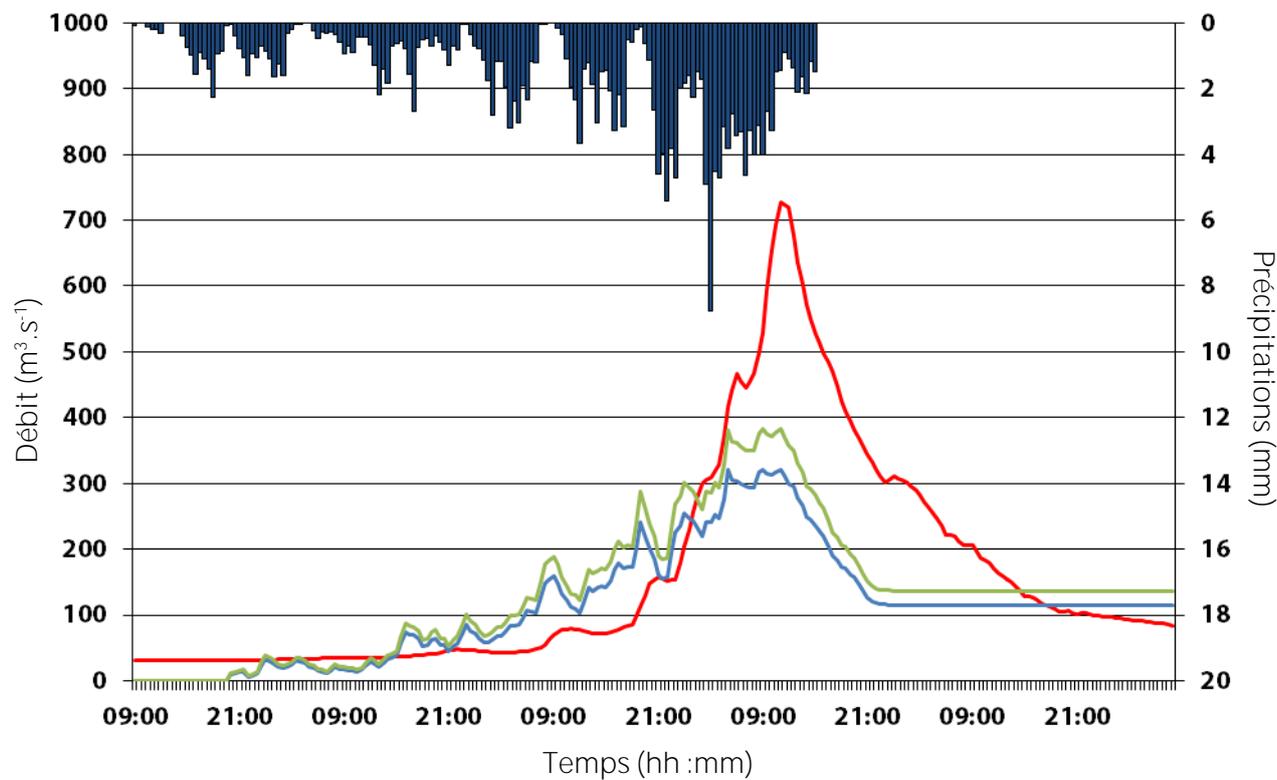
k+2



$k+3$

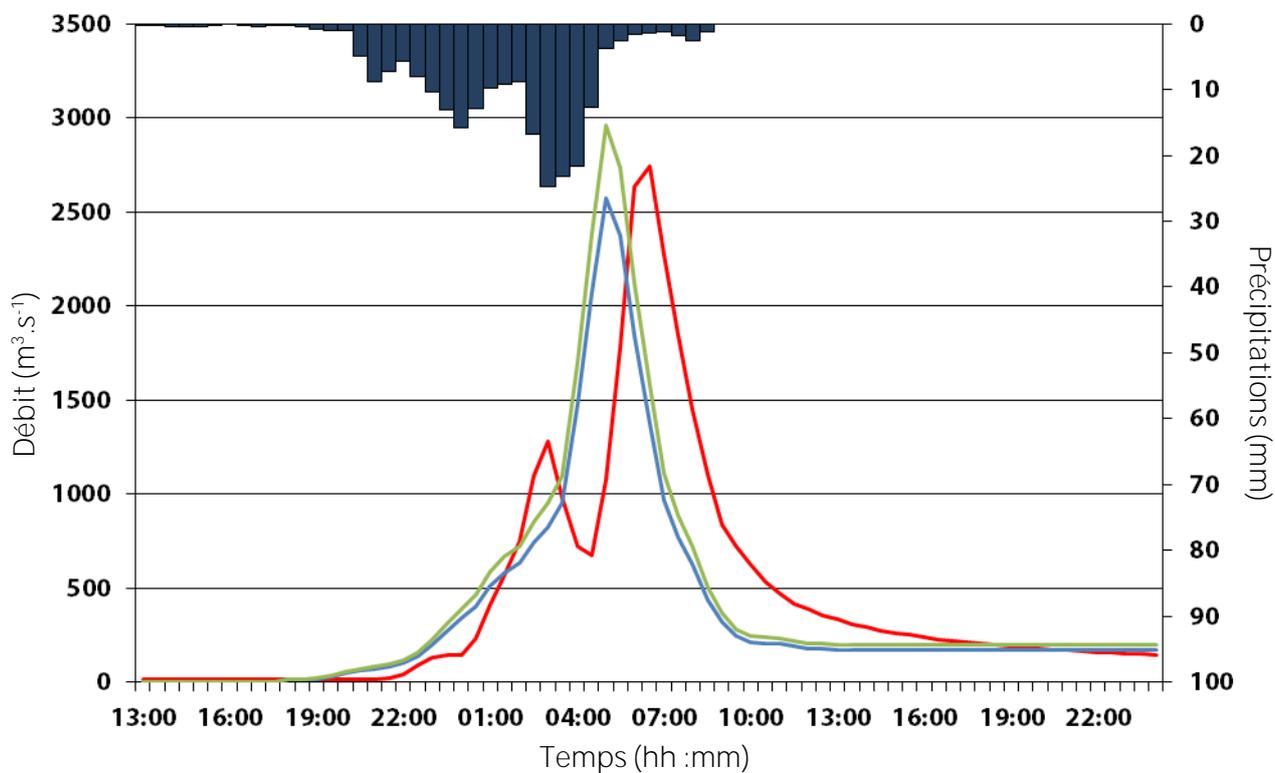


$k+4$

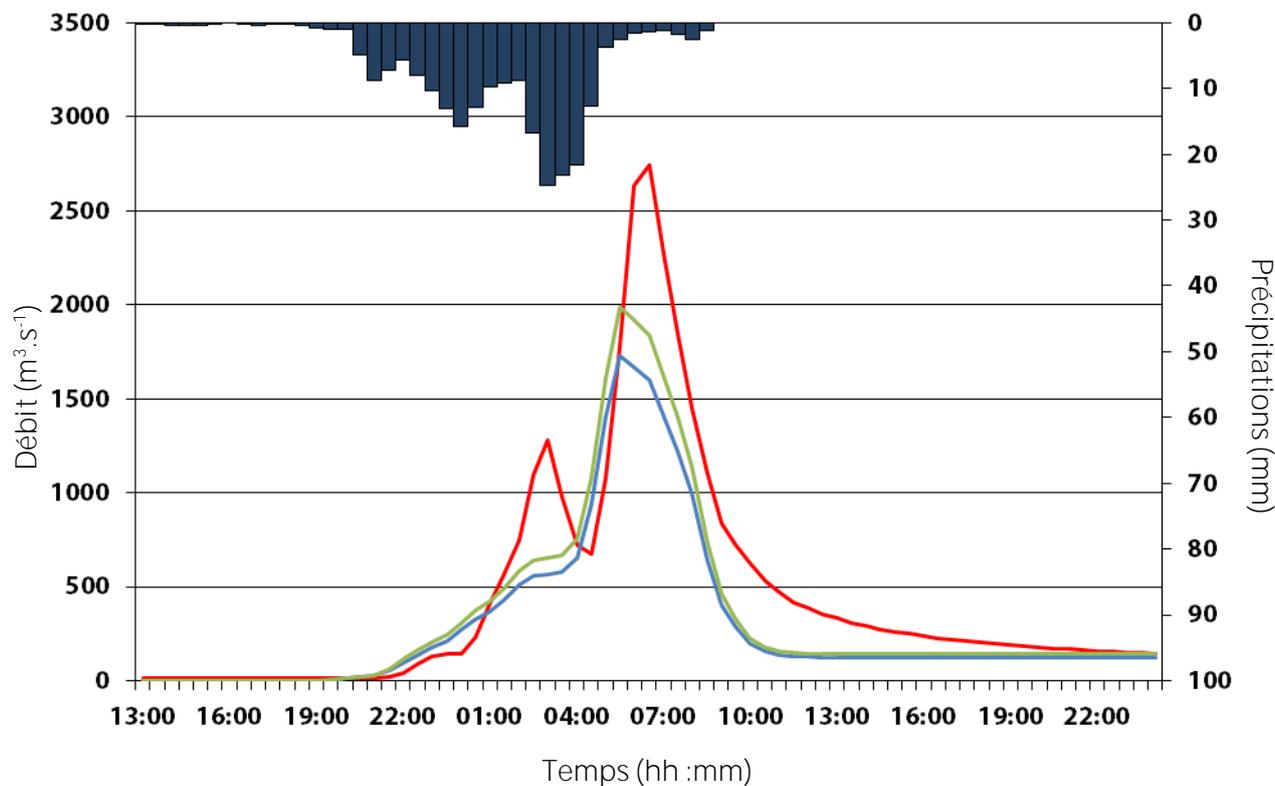


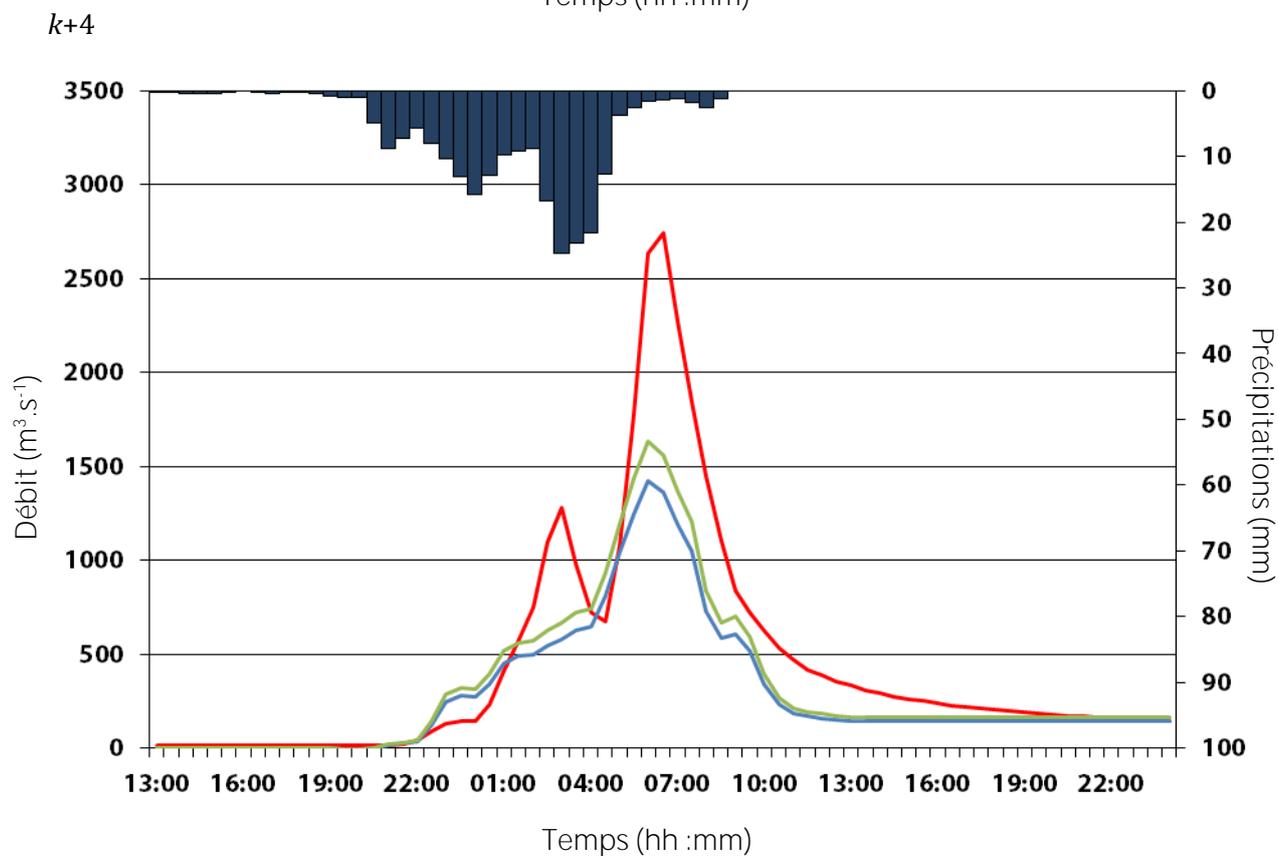
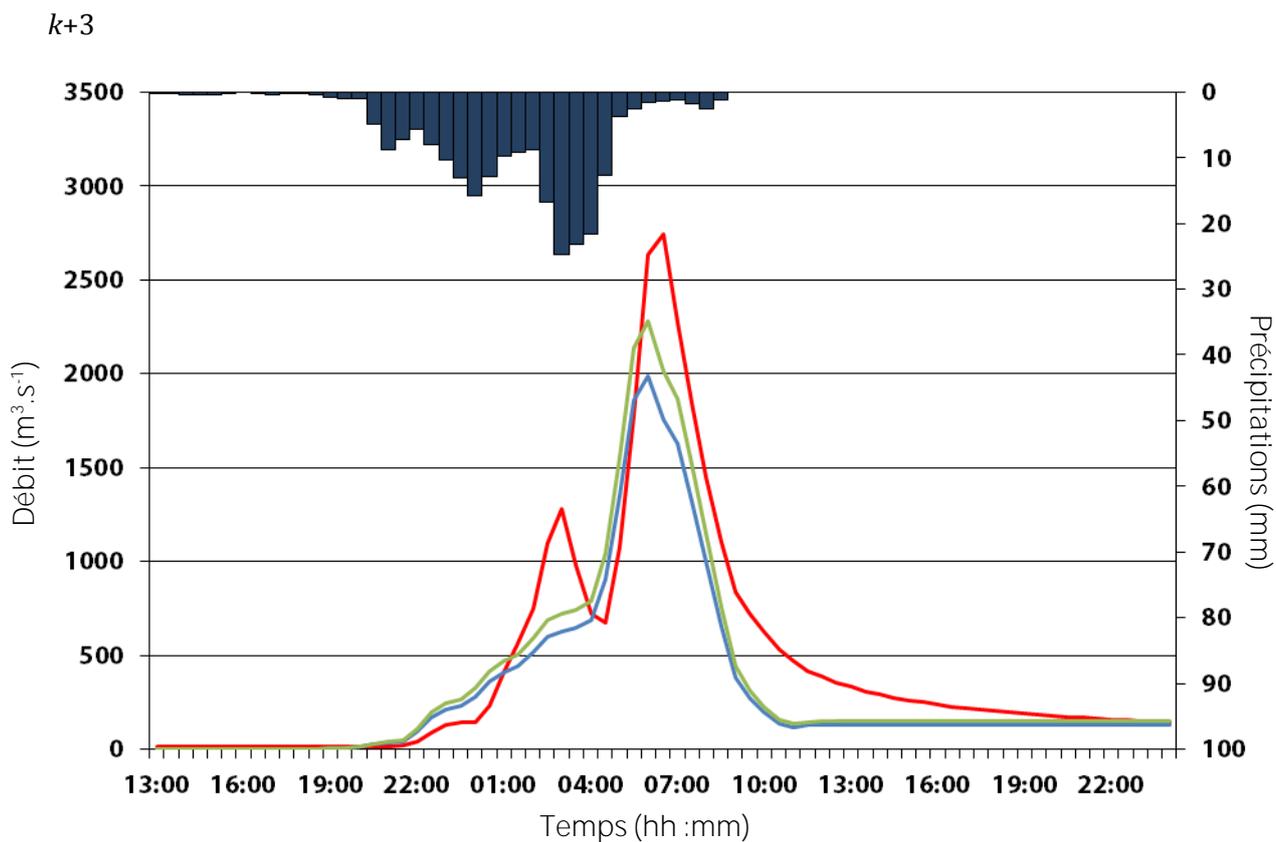
Bassin du Gardon à Anduze, évènement n°934 :

k+1



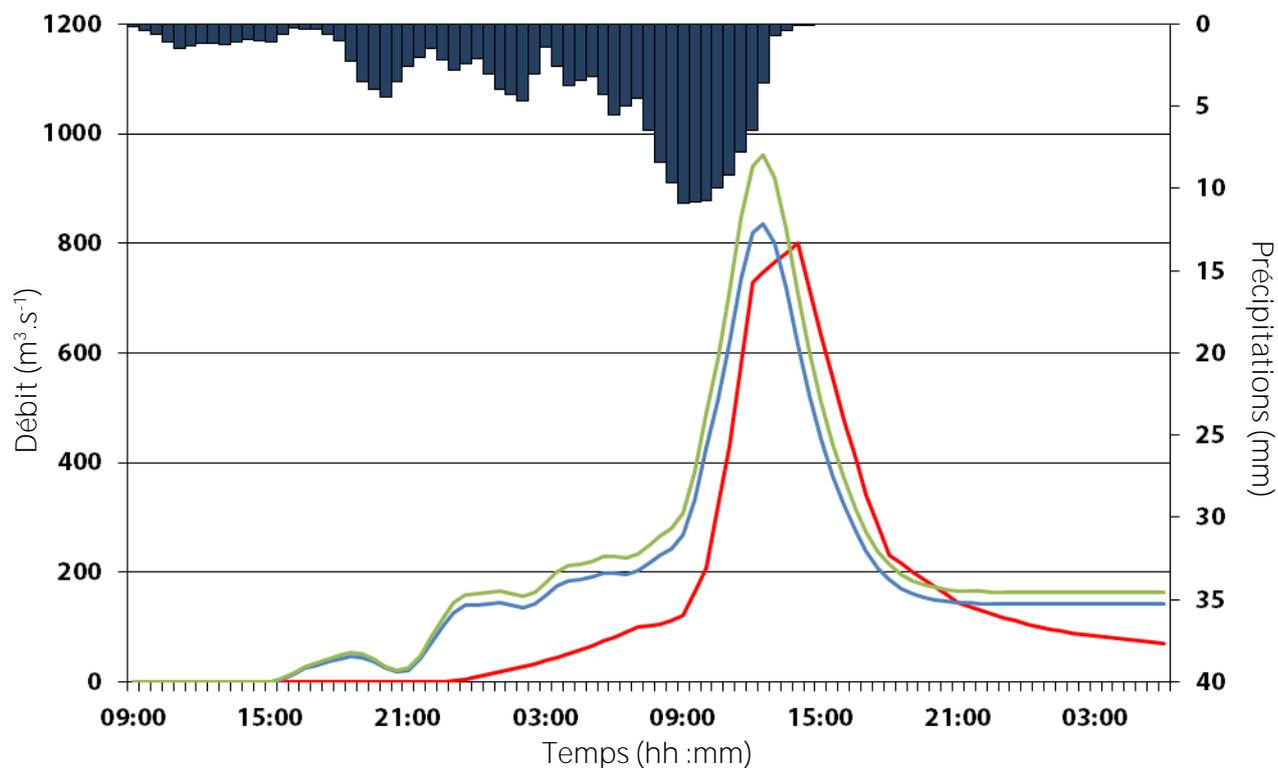
k+2



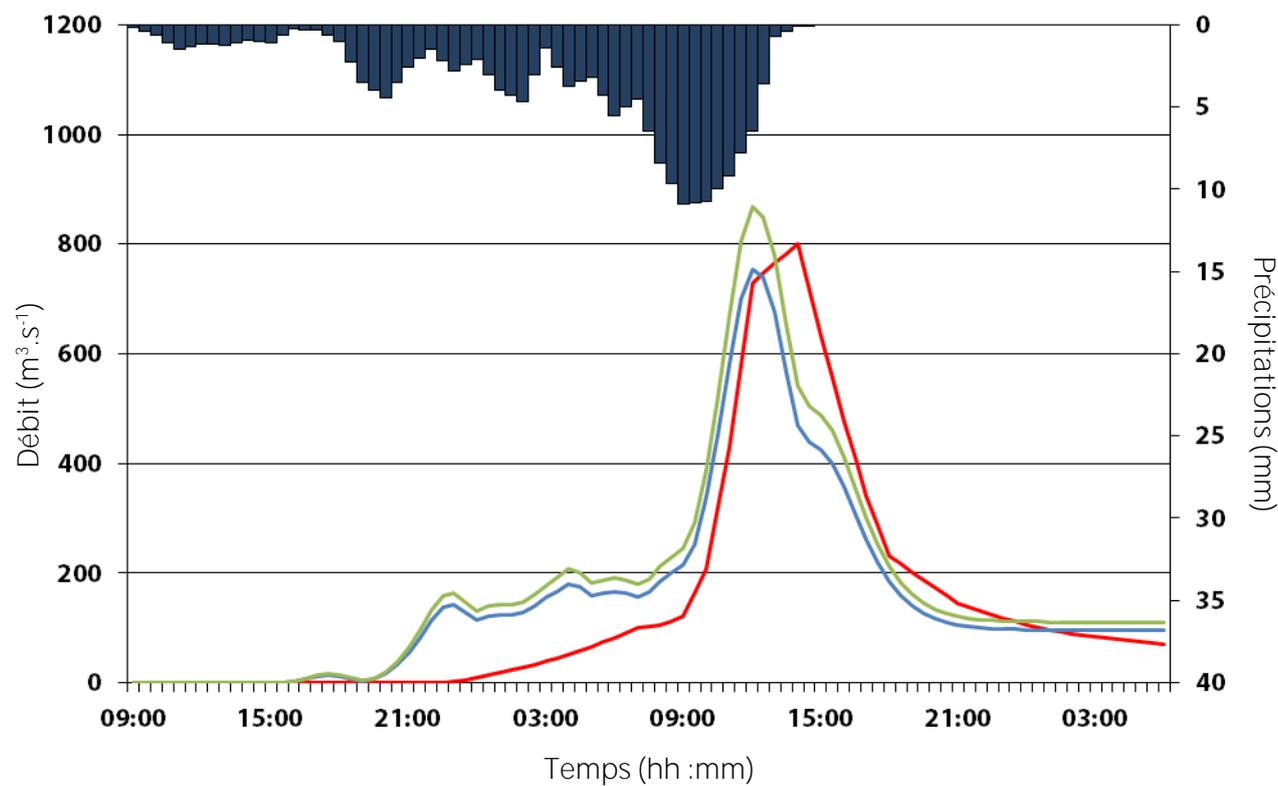


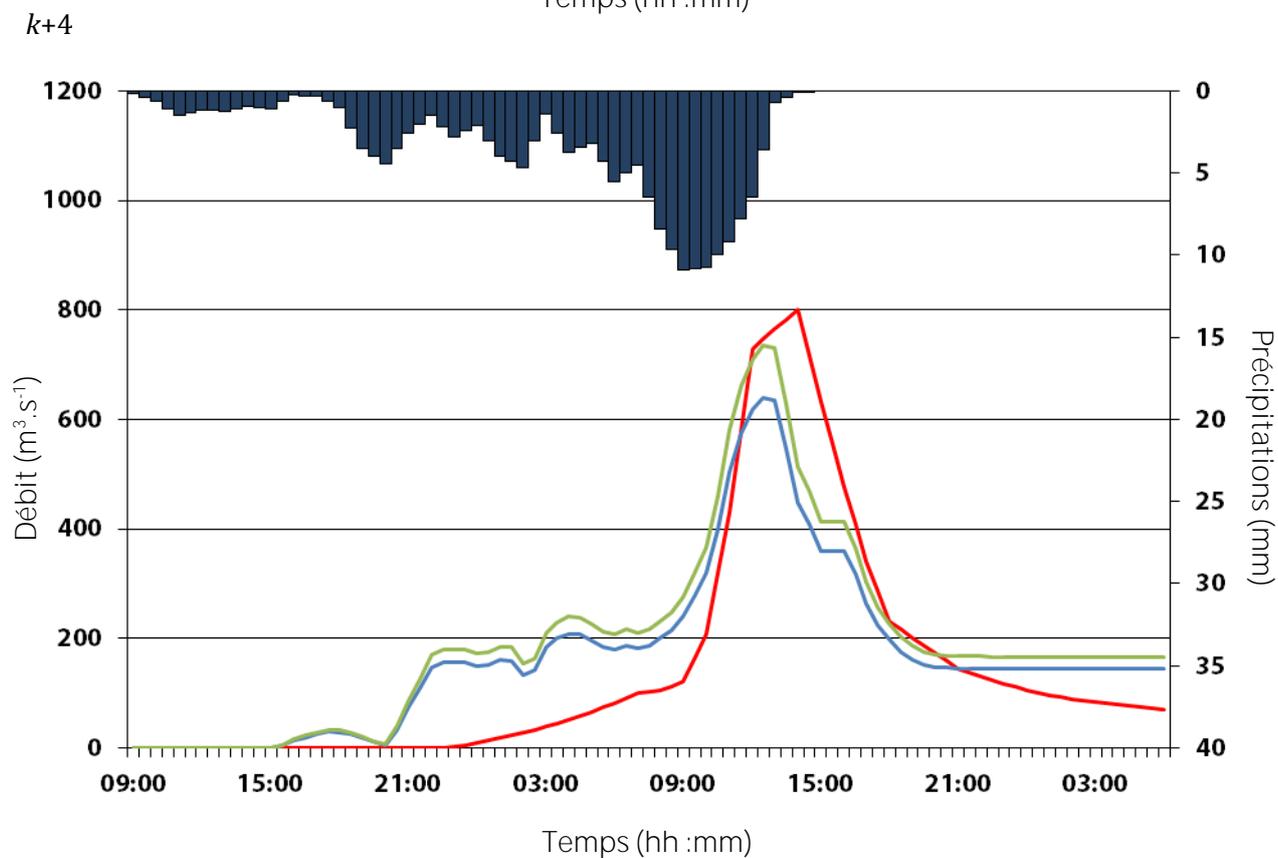
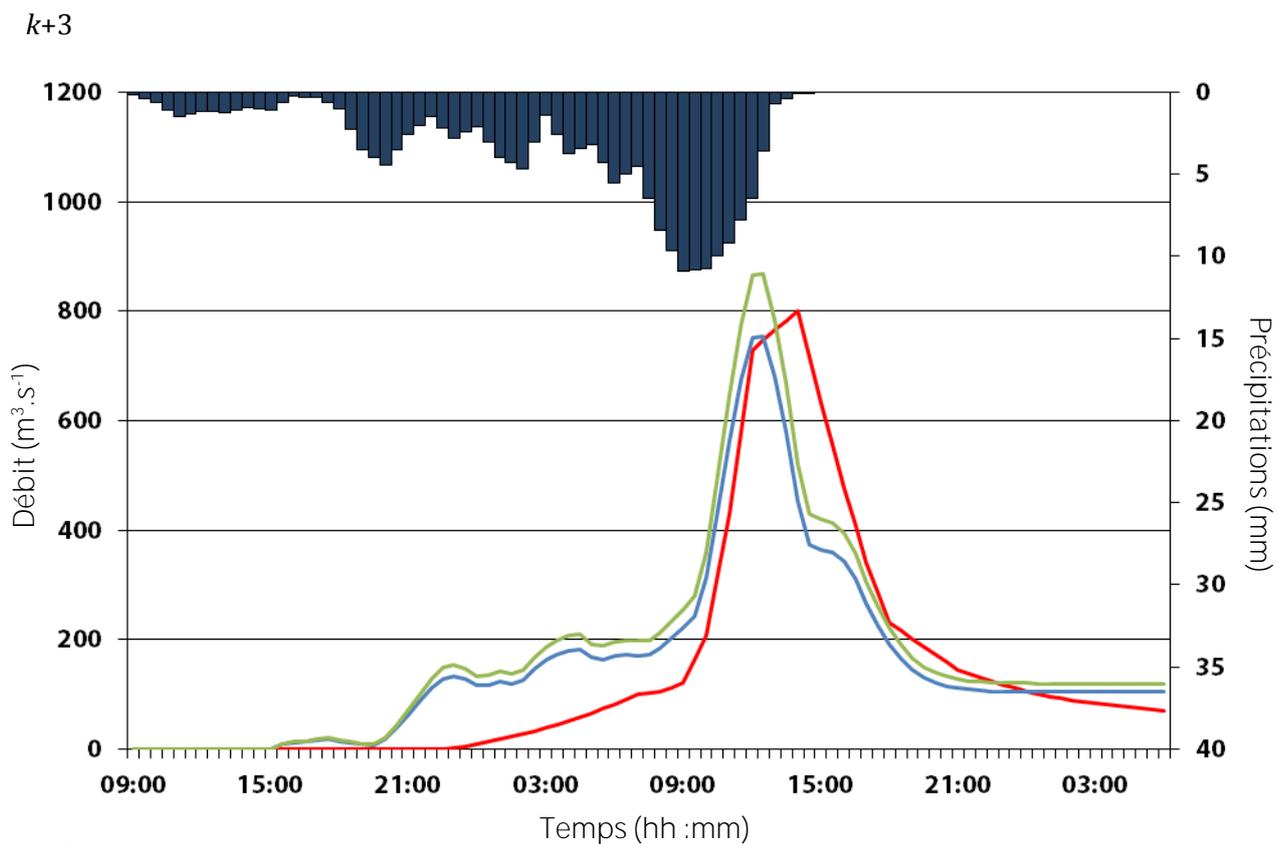
Bassin du Gardon à Anduze, évènement n°935 :

k+1



k+2





**ANNEXE 7** : Tableau des résultats détaillés en prévision sur les évènements du bassin de la Ganière.

Evènement	Horizon	Nash	Persistance	SPPD	PPD	Décalage (h)
910	k+1	0,70	-17,32	112%	112%	0,00
910	k+2	0,87	-1,21	84%	84%	0,00
910	k+3	0,84	-0,37	86%	88%	0,50
910	k+4	0,84	0,18	95%	97%	1,00
911	k+1	0,81	-14,33	73%	80%	2,00
911	k+2	0,87	-2,07	86%	86%	-3,50
911	k+3	0,86	0,12	84%	89%	-3,50
911	k+4	0,86	0,39	83%	87%	0,50
912	k+1	0,68	-20,11	75%	79%	1,00
912	k+2	0,75	-5,23	78%	80%	0,50
912	k+3	0,73	-3,20	80%	82%	0,50
912	k+4	0,78	-1,54	86%	89%	1,00
913	k+1	-2,62	-39,14	44%	65%	1,00
913	k+2	-0,34	-2,54	64%	78%	1,00
913	k+3	-0,90	-1,18	63%	80%	1,00
913	k+4	-1,26	-0,97	66%	91%	1,00
914	k+1	0,25	-40,19	21%	25%	2,50
914	k+2	0,50	-11,50	33%	33%	0,50
914	k+3	0,54	-8,87	35%	35%	0,50
914	k+4	0,60	-2,41	39%	39%	0,00
915	k+1	-0,01	-27,55	131%	137%	1,00
915	k+2	-0,80	-14,56	159%	169%	-1,50
915	k+3	-1,01	-3,98	171%	173%	-1,00
915	k+4	-1,52	-4,07	184%	191%	-1,00
916	k+1	0,81	-12,89	73%	79%	1,50
916	k+2	0,85	-0,10	74%	77%	1,50
916	k+3	0,81	0,24	78%	79%	2,00
916	k+4	0,81	0,26	88%	88%	1,50
917	k+1	0,71	-7,81	86%	87%	1,00
917	k+2	0,80	-0,23	69%	72%	1,00
917	k+3	0,76	0,09	72%	72%	1,00
917	k+4	0,78	0,52	81%	82%	1,00
918	k+1	0,69	-25,45	87%	87%	0,00
918	k+2	0,74	-5,66	83%	83%	0,00
918	k+3	0,67	-0,90	82%	82%	0,00
918	k+4	0,67	0,02	88%	88%	0,00
919	k+1	0,63	-3,13	127%	127%	0,00
919	k+2	0,76	-0,12	121%	122%	0,50
919	k+3	0,64	0,65	116%	116%	0,00
919	k+4	0,61	0,71	115%	117%	0,50
920	k+1	0,64	-9,82	43%	55%	2,50
920	k+2	0,65	-1,34	52%	57%	1,00
920	k+3	0,73	-0,35	58%	64%	1,00
920	k+4	0,78	-0,20	66%	72%	1,50

## Liste des figures

Figure 1: localisation de la zone d'étude en France (en rouge). Le cadre rouge représente l'emprise approximative des cartes suivantes. Altitude d'après IGN BD ALTI 250, départements d'après IGN BD Geofla.....	16
Figure 2: topographie de la zone d'étude. Altitudes d'après ASTER GDEM, un produit of METII and NASA, départements d'après IGN BD Geofla, cours d'eau d'après IGN BD Carthage.....	17
Figure 3: pentes de la zone d'étude. Altitudes utilisées pour le calcul des pentes : IGN BD ALTI 250, limites départementales d'après IGN BD Geofla. ....	18
Figure 4: hydrographie de la zone d'étude. Cours d'eau, limites de bassin et ligne de partage des eaux d'après IGN BD Carthage. Les cours d'eau principaux sont de classification 1 et 2, les cours d'eau secondaires 3, 4 et 5 et les cours d'eau très secondaires de classification supérieure (cf. Tableau 1). ....	20
Figure 5: carte géologique simplifiée de la zone d'étude. Géologie d'après cartes géologiques de la France 1/50 000, 1/250 000 et 1/1 000 000, BRGM, feuilles diverses ; cours d'eau d'après IGN BD Carthage. ....	21
Figure 6: carte simplifiée de l'occupation des sols de la zone d'étude. Occupation des sols d'après Corine Land Cover 2006 ; limites départementales d'après IGN BD Geofla.....	23
Figure 7: carte de densité de la population sur la zone d'étude. Lissage de type « spline » à partir des données communales. Densité d'après IGN RGC 2010 ; limites départementales d'après IGN BD Geofla ; cours d'eau d'après IGN BD Carthage. ....	24
Figure 8: évolution de la population entre 1999 et 2008. Lissage de type « spline » à partir des données communales. Données démographiques d'après IGN RGC 2010 ; limites départementales d'après IGN BD Geofla.....	25
Figure 9: climatologie simplifiée des températures et des précipitations à la station de Saint-Christol-lez-Alès (30) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990). ....	26
Figure 10: climatologie simplifiée des températures et des précipitations à la station du Mont-Aigoual (Valleraugue, 30) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990). ....	27
Figure 11: climatologie simplifiée des températures et des précipitations à la station d'Issanlas (07) sur les normales climatiques 1951-1980 (Kessler et Chambraud, 1990).....	28
Figure 12: représentation schématique d'un épisode cévenol en coupe entre Golfe du Lion et Cévennes. .	29
Figure 13: nombre d'épisodes (par département) de plus de 200mm en une journée climatologique de 1958 à 2010 inclus (Météo France, 2012).....	30
Figure 14: répartition mensuelle des épisodes de plus de 200mm en une journée climatologique sur la France et durant la période 1958-2010 (Météo France, 2012). ....	31
Figure 15: nombre d'épisodes de plus de 200mm en une journée climatologique par an sur la France et par département, durant la période 1958-2010 (Météo France, 2012). ....	32
Figure 16: écoulements mensuels (naturels) calculés sur 43 ans pour le Gardon d'Anduze à Anduze (surface du bassin : 545km <sup>2</sup> ). Données de débit d'après Banque Hydro. ....	32
Figure 17: exemple de carte de vigilance crues le 3 novembre 2011 sur le territoire du SPC Grand Delta (en blanc). Les tronçons règlementaires sont colorés en fonction du niveau de risque qui pourrait être atteint dans les prochaines 24 heures.....	37
Figure 18: (a) : exemple de carte de vigilance crue sur le nord du Gard et le sud de l'Ardèche, les tronçons en vert sont dits tronçons règlementaires et sont dotés d'un niveau de vigilance, ici minimal. (b) : ensemble des cours d'eau de catégorie trois et moins du nord du Gard et du sud de l'Ardèche (conception	

Artigue 2010, sources : (a) : vigicrues.gouv.fr ; (b) : altitude d'après ASTER GDEM, un produit de METII and NASA, cours d'eau d'après IGN BD CARTHAGE, limites départementales d'après IGN BD GEOFLA).....	40
Figure 19: aspect de l'hydrogramme observé en sortie de deux bassins de formes différentes soumis à des précipitations semblables (d'après Musy et al., 2009). .....	53
Figure 20: représentation schématique d'un modèle hydrologique mathématique pluie-débit.....	55
Figure 21: représentation schématique d'un neurone formel. $\mathbf{x}$ représente le vecteur des variables d'entrée, $\mathbf{w}$ la matrice des paramètres, ou coefficients synaptiques, $f(.)$ est la fonction d'activation et $s(v)$ la sortie du neurone. ....	67
Figure 22: représentation schématique du perceptron multicouche : ici, réseau non bouclé à une couche de neurones cachés. ....	69
Figure 23: représentation schématique d'un réseau de neurones bouclé d'ordre 1.....	69
Figure 24: représentation schématique de l'approximateur universel : un réseau de neurones à $n+1$ entrées, $p$ neurones et un neurone de sortie linéaire. La variable appelée « entrée constante » est nécessaire pour assurer que la sortie puisse être non nulle même dans le cas où toutes les variables sont nulles.....	70
Figure 25: représentation schématique du processus de validation croisée sur $D$ sous-ensembles.....	74
Figure 26: score de validation croisée en fonction du nombre de neurones cachés en trait plein : sans régularisation ; en tirets rouges : avec la modération des poids ; en pointillés bleus : avec l'arrêt précoce. D'après Kong-A-Siou et al., 2012.....	75
Figure 27: situation du bassin versant du Gardon de Mialet dans la zone d'étude. Cours d'eau d'après IGN BD Carthage ; départements d'après IGN BD Geofla.....	82
Figure 28: 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.....	83
Figure 29: carte géologique simplifiée du bassin versant du Gardon de Mialet à Mialet. Géologie d'après carte géologique de la France au 1/50 000, BRGM ; Cours d'eau d'après IGN BD Carthage ; effet de relief d'après ASTER GDEM, un produit de METII et NASA. ....	84
Figure 30: courbes de tarage de la station limnimétrique de Mialet pour les faibles hauteurs (bleu) et pour les hauteurs élevées (rouge). Le marqueur vert représente l'estimation du débit de la crue de septembre 2002 effectuée par SIEE en 2003.....	88
Figure 31 : répartition des évènements conservés entre 1992 et 2008. ....	89
Figure 32: répartition mensuelle des évènements de la base de données.....	89
Figure 33: durées des épisodes par classes. ....	90
Figure 34: doubles cumuls appliqués à chaque couple de pluviomètres du bassin. La corrélation linéaire entre la droite de régression est les points est à chaque fois supérieure à 0.99. ....	91
Figure 35: histogramme de la répartition des cumuls par classes de 100mm pour les trois pluviomètres du bassin. ....	92
Figure 36: représentation en nuage de points de l'ensemble des relevés semi-horaires des précipitations sur les trois couples de pluviomètres avec équation et coefficient de corrélation linéaire de la droite de régression linéaire.....	93
Figure 37: représentation en nuage de points de l'ensemble des cumuls de précipitations par épisode sur les trois couples de pluviomètres avec équation et coefficient de corrélation linéaire de la droite de régression linéaire.....	94

Figure 38: représentation des intensités moyennes sur 1h et 1/2h maximales, moyennes et médianes sur l'ensemble de la base pour chaque poste pluviométrique. ....	95
Figure 39: histogramme de la répartition des débits spécifiques de pointe à la station de Mialet.....	96
Figure 40: répartition en trois classes des rapports de forme des évènements de la base. ....	97
Figure 41: hydrogrammes des deux évènements présentant les rapports de forme extrêmes : à gauche, l'évènement n°230 et à droite l'évènement n°3.....	97
Figure 42: coefficient d'écoulement moyen par mois (la population prise en compte pour la moyenne est indiquée dans les barres). ....	98
Figure 43: hydrogramme et hyétogramme de l'évènement n°212. ....	99
Figure 44: modèle utilisé dans les premières simulations. En vert : les entrées ; en bleu : la couche cachée ; en orange : le débit estimé (entrée récurrente) ; en rouge la sortie. ....	101
Figure 45: modèle statique utilisé dans la prévision des crues du Gardon de Mialet. ....	103
Figure 46: modèle récurrent utilisé dans la prévision des crues du Gardon de Mialet.....	104
Figure 47: histogramme de fréquence d'apparition de différents temps de réponse entre les précipitations enregistrées à Barre-des-Cévennes et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h). ....	106
Figure 48: histogramme de fréquence d'apparition de différents temps de réponse entre précipitations enregistrées à Saint-Roman-de-Tousque et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h). ....	106
Figure 49: histogramme de fréquence d'apparition de différents temps de réponse entre précipitations enregistrées à Mialet et le débit mesuré à Mialet par classes de trois pas de temps (un pas de temps vaut 0.5h). ....	107
Figure 50: évènement n°208 présentant une pluie clairement reliée à une réponse hydrologique, comme une réponse à une impulsion. ....	109
Figure 51: évènement n°222 de janvier 1998 utilisé comme ensemble d'arrêt. ....	113
Figure 52: hydrogrammes d'observation et de prévision à $k+3$ des évènements 19 (à gauche) et 13 (à droite) avec le modèle statique dirigé.....	115
Figure 53 : représentation incréments de débit prévus en fonction des incréments de débit observés pour les quatre évènements aux quatre horizons de prévision.....	117
Figure 54: hydrogrammes observés et prévus par le modèle récurrent pour les évènements 8 à $k+1$ , 13 à $k+2$ , 302 à $k+3$ et 19 à $k+4$ . ....	119
Figure 55: comparaison des résultats moyens et extrêmes du modèle ensembliste avec les résultats du modèle déterministe pour les quatre évènements à quatre horizons de prévision. ....	120
Figure 56: comparaison des résultats du modèle récurrent déterministe non dirigé sans ou avec SIM. ...	122
Figure 57: schématisation du raisonnement conduisant à la généralisation aux bassins non jaugés.....	127
Figure 58: localisation des bassins cibles ainsi que du bassin du Gardon de Mialet à Mialet, le bassin pilote (numéros par ordre de surface croissante). ....	129
Figure 59: nombre d'évènements par classes de débits spécifiques sur un total de 35 évènements.....	132
Figure 60: nombre d'épisodes par classes de cumul moyen. ....	132
Figure 61: décalage temporel moyen en fonction de la surface des bassins considérés. Le marqueur bleu symbolise le décalage théorique du bassin pilote. ....	134

Figure 62: décalage temporel moyen en fonction de la pente du drain des bassins considérés. Le marqueur bleu symbolise le décalage théorique du bassin pilote.....	134
Figure 63: décalage temporel moyen en fonction de la longueur du drain du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote.....	135
Figure 64: décalage temporel en fonction de la longueur du drain du bassin considéré.....	135
Figure 65: décalage temporel en fonction de la pente des versants du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote.....	136
Figure 66: décalage temporel en fonction de la dénivelée du bassin considéré. Le marqueur bleu symbolise le décalage théorique du bassin pilote. ....	136
Figure 67: PPD moyen en fonction de la surface du bassin considéré.....	137
Figure 68: 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). ....	140
Figure 69: représentation des rapports des débits spécifiques de pointe et des surfaces des bassins aux valeurs respectives du bassin de Mialet. ....	141
Figure 70: carte du bassin versant du ruisseau du Lyonnais.....	143
Figure 71: prévision de l'évènement des 8 et 9 septembre 2002 sur le ruisseau du Lyonnais à $k+3$ .....	144
Figure 72: carte du bassin versant de l'Alzon à Saint-Christol-lez-Alès.....	146
Figure 73: prévision de l'évènement des 8 et 9 septembre 2002 sur l'Alzon à $k+2$ . ....	147
Figure 74: carte du bassin versant de la Salandre. ....	148
Figure 75: prévision de l'évènement des 8 et 9 septembre 2002 sur la Salandre à $k+2$ .....	149
Figure 76: carte du bassin versant de l'Amous.....	151
Figure 77: prévision de l'évènement de septembre 2002 sur l'Amous à $k+3$ .....	152
Figure 78: carte du bassin versant de l'Auzonnet. ....	153
Figure 79: prévision de l'évènement de septembre 2002 sur l'Auzonnet à $k+2$ . ....	154
Figure 80: carte du bassin du Galeizon amont.....	156
Figure 81: prévision de l'évènement de septembre 2002 sur le Galeizon amont à $k+2$ . ....	157
Figure 82: carte du bassin versant de la Ganière.....	158
Figure 83: prévision de l'évènement d'octobre 2004 sur la Ganière à $k+4$ . ....	159
Figure 84: prévision de l'évènement de septembre 2002 sur la Ganière à $k+2$ .....	160
Figure 85: carte du bassin versant de l'Avène. ....	161
Figure 86: prévision de l'évènement de septembre 2002 sur l'Avène à $k+3$ . ....	162
Figure 87: carte du bassin versant du Galeizon. ....	164
Figure 88: carte du bassin du Gardon à Saumane. ....	165
Figure 89: prévision de l'évènement de septembre 2002 sur le Gardon à Saumane à $k+4$ .....	166
Figure 90: carte du bassin versant de la Baume. ....	168
Figure 91: prévision de l'évènement d'octobre 2004 sur la Baume à Rosières à $k+4$ . ....	169
Figure 92: carte du bassin versant de l'Ardèche à Pont-de-Labeaume. ....	170
Figure 93: prévision de l'évènement d'octobre 2001 à un horizon de $k+3$ .....	171

Figure 94 : carte du bassin versant du Chassezac.....	173
Figure 95: prévision de l'évènement de septembre 2000 à $k+4$ .....	174
Figure 96: carte du bassin versant du Gardon d'Anduze. ....	176
Figure 97: prévision de l'évènement de septembre 2002 à $k+3$ .....	177



## Liste des tableaux

Tableau 1: classification des cours d'eau de la BD Carthage (IGN, 2006). .....	40
Tableau 2: quelques statistiques sur les cumuls enregistrés sur les différents pluviomètres, pour les évènements sélectionnés .....	92
Tableau 3: coefficients de corrélation linéaire entre les débits de pointe et les pluies en cumuls et en intensités horaires et semi-horaires. ....	98
Tableau 4 : indication de l'architecture du premier type de modèle utilisé. ....	102
Tableau 5: meilleurs résultats obtenus avec un modèle non linéaire simple sur l'évènement n°13. ....	102
Tableau 6: temps de réponse et intervalles d'investigation pour les optimisations, par pluviomètre. ....	107
Tableau 7 : quelques caractéristiques des évènements 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. ....	110
Tableau 8: architectures optimales pour les deux types de modèles utilisés. Les valeurs sont indiquées en nombre de pas de temps (soit en 0,5h). ....	111
Tableau 9 : Seuils de débit approximatifs correspondant aux niveaux de vigilance. ....	113
Tableau 10: Récapitulatif des résultats obtenus avec le modèle statique dirigé (Evt. correspond au numéro de l'évènement, $l$ à l'horizon de prévision, Pers. au critère de persistance, $V_p$ à la vigilance prévue par le modèle, $V_o$ à la vigilance correspondant au débit observé à $k+l$ et $D$ au décalage temporel). ....	114
Tableau 11 : pourcentages d'anticipation des seuils de vigilance crue pour chaque évènement et chaque horizon de prévision. L'évènement n°8 étant à deux pics, on distingue l'anticipation pour chacun d'eux. ....	116
Tableau 12: récapitulatif des résultats obtenus avec le modèle récurrent non dirigé (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). ....	118
Tableau 13: principales caractéristiques topographiques des bassins sélectionnés (les statistiques en bas de tableau sont calculées, Gardon de Mialet exclu, seulement présenté ici comme référence ; les cellules en rouge représentent les valeurs minimales et les vertes les valeurs maximales). ....	130
Tableau 14: nombre d'épisodes par bassin, type de données et source de ces données (Rex 2002 : retour d'expérience de l'épisode de septembre 2002 ; SIEE / Syndicat du Galeizon : données issues des simulations de l'épisode de septembre 2002 par SIEE commandées par le syndicat du Galeizon ; OHMCV : données issues de la base de données Sevnol de l'Observatoire Hydrométéorologique Cévennes Vivarais ; BVNE : données fournies dans le cadre du projet BVNE). ....	131
Tableau 15: coefficient correctif correspondant à chaque bassin visé par l'étude. ....	142
Tableau 16: résultats détaillés en prévision de l'évènement de septembre 2002 sur le ruisseau du Lyonnais. ....	144
Tableau 17: résultats détaillés en prévision de l'évènement de septembre 2002 sur l'Alzon. ....	147
Tableau 18: résultats détaillés en prévision de l'évènement de septembre 2002 sur la Salandre. ....	149
Tableau 19: résultats détaillés en prévision de l'évènement de septembre 2002 sur l'Amous. ....	152
Tableau 20: description synthétique des trois évènements disponibles pour l'Auzonnet. $RR$ représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	154
Tableau 21: résultats détaillés en prévision des évènements 921, 922 et 923 sur l'Auzonnet. ....	155

Tableau 22: résultats détaillés en prévision de l'évènement de septembre 2002 sur le Galeizon amont. .	157
Tableau 23: description synthétique des 11 évènements disponibles pour la Ganière. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	159
Tableau 24: résultats détaillés en prévision de l'évènement de septembre 2002 sur l'Avène. ....	162
Tableau 25: résultats détaillés en prévision de l'évènement de septembre 2002 sur le Galeizon. ....	164
Tableau 26: description synthétique des trois évènements disponibles pour le Gardon à Saumane. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	166
Tableau 27: résultats détaillés en prévision des évènements 931, 932 et 933 sur le Gardon à Saumane. .	167
Tableau 28: description synthétique des trois évènements disponibles pour la Baume. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	168
Tableau 29: résultats détaillés en prévision des évènements 904, 905 et 906 sur la Baume. ....	169
Tableau 30: description synthétique des trois évènements disponibles pour l'Ardèche. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	171
Tableau 31 : résultats détaillés en prévision des évènements 901, 902 et 903 sur l'Ardèche. ....	172
Tableau 32 : description synthétique des trois évènements disponibles pour le Chassezac. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	174
Tableau 33: résultats détaillés en prévision des évènements 907, 908 et 908 sur le Chassezac. ....	175
Tableau 34: description synthétique des deux évènements disponibles pour le Gardon à Anduze. RR représente le cumul moyen, $I_{xl}$ l'intensité maximale moyenne en 1h, $I_{xc}$ l'intensité maximale moyenne en 0.5h, $Q_x$ le débit de pointe, $q_x$ le débit spécifique de pointe, Pluie le type de données de pluie disponible et Dispo. le nombre d'appareils disponibles pour l'évènement. ....	177
Tableau 35: résultats détaillés en prévision des évènements 934 et 935 sur le Gardon à Anduze. ....	177
Tableau 36 : <i>SPPD</i> moyens et médians et taux d'atteinte de <i>SPPD</i> compris entre 70 et 130% pour différents types d'évènements et de bassins, décrits dans le tableau. ....	179
Tableau 37 : augmentation en valeur des critères de Nash, du <i>PPD</i> et du <i>SPPD</i> pour l'ensemble des évènements testés avec le traitement appliqué. ....	180

## Liste des photos

Photo 1: à gauche : barrage écrêteur de crues et de soutien d'étiage de Sénéchas, sur la Cèze, à sa confluence avec l'Homol ;.....	19
Photo 2: à droite : barrage écrêteur de crues de Sainte-Cécile-d'Andorge, sur le Gardon d'Alès. Photos : Artigue, 2007. ....	19
Photo 3: à gauche : terrains karstiques dans les gorges de la Cèze, sur la commune de Méjannes-le-Clap (30), photo Artigue, 2007.....	22
Photo 4: à droite : affleurement de micaschistes au cœur des Cévennes, sur la commune de Saint-André-de-Valborgne (48). Photos : Artigue, 2007.....	22
Photo 5 : vignoble en terrain argileux avec sillons dans le sens de la pente, Saint-Hilaire-de-Brethmas (30) ; importants ruissellements suite à un orage qui a laissé 62mm dont 42mm en 25 minutes le 24/09/2012. Photo : Artigue 2012.....	48
Photo 6: <i>faïsses</i> près de Gravières (07) en aval du Chassezac ; Photo : Artigue 2012.....	49
Photo 7: pluviomètre de Saint-Roman-de-Tousque ; photo : Artigue, 2011.....	83





## Résumé

Dans les régions méditerranéennes françaises, des épisodes pluvieux diluviens se produisent régulièrement et provoquent des crues très rapides et volumineuses que l'on appelle crues éclair. Elles font fréquemment de nombreuses victimes et peuvent, sur un seul évènement, coûter plus d'un milliard d'euros. Face à cette problématique, les pouvoirs publics mettent en place des parades parmi lesquelles la prévision hydrologique tient une place essentielle.

C'est dans ce contexte que le projet BVNE (Bassin Versant Numérique Expérimental) a été initié par le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) dans le but d'améliorer la prévision des crues rapides. Ces travaux s'inscrivent dans le cadre de ce projet et ont trois objectifs principaux : réaliser des prévisions sur des bassins capables de ces réactions qu'ils soient correctement jaugés, mal jaugés ou non jaugés.

La zone d'étude choisie, le massif des Cévennes, concentre la majorité de ces épisodes hydrométéorologiques intenses en France. Ce mémoire la présente en détails, mettant en avant ses caractéristiques les plus influentes sur l'hydrologie de surface.

Au regard de la complexité de la relation entre pluie et débit dans les bassins concernés et de la difficulté éprouvée par les modèles à base physique à fournir des informations précises en mode prédictif sans prévision de pluie, l'utilisation de l'apprentissage statistique par réseaux de neurones s'est imposée dans la recherche d'une solution opérationnelle.

C'est ainsi que des modèles à réseaux de neurones ont été synthétisés et appliqués à un bassin de la zone cévenole, dans des contextes bien et mal jaugés. Les bons résultats obtenus ont été le point de départ de la généralisation à 15 bassins de la zone d'étude.

A cette fin, une méthode de généralisation est développée à partir du modèle élaboré sur le bassin jaugé et de corrections estimées en fonction des caractéristiques physiques des bassins.

Les résultats de l'application de cette méthode sont de bonne qualité et ouvrent la porte à de nombreux axes de recherche pour l'avenir, tout en démontrant encore que l'utilisation de l'apprentissage statistique pour l'hydrologie peut constituer une solution pertinente.

## Summary

In the French Mediterranean regions, heavy rainfall episodes regularly occur and induce very rapid and voluminous floods called flash floods. They frequently cause fatalities and can cost more than one billion euros during only one event. In order to cope with this issue, the public authorities' implemented countermeasures in which hydrological forecasting plays an essential role.

In this context, the French Flood Forecasting Service (called SCHAPI for Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) initiated the BVNE (Digital Experimental Basin, for Bassin Versant Numérique Expérimental) project in order to enhance flash flood forecasts. The present work is a part of this project and aim at three main purposes: providing flash flood forecasts on well-gauged basins, poorly gauged basins and ungauged basins.

The study area chosen, the *Cévennes* range, concentrates the major part of these intense hydrometeorological events in France. This dissertation presents it precisely, highlighting its most hydrological-influent characteristics.

With regard to the complexity of the rainfall-discharge relation in the focused basins and the difficulty experienced by the physically based models to provide precise information in forecast mode without rainfall forecasts, the use of neural networks statistical learning imposed itself in the research of operational solutions.

Thus, the neural networks models were designed and applied to a basin of the *Cévennes* range, in the well-gauged and poorly gauged contexts. The good results obtained have been the start point of a generalization to 15 basins of the study area.

For this purpose, a generalization method was developed from the model created on the gauged basin and from corrections estimated as a function of basin characteristics.

The results of this method application are of good quality and open the door to numerous paths of inquiry for the future, while demonstrating again that the use of statistical learning for hydrology can be a relevant solution.