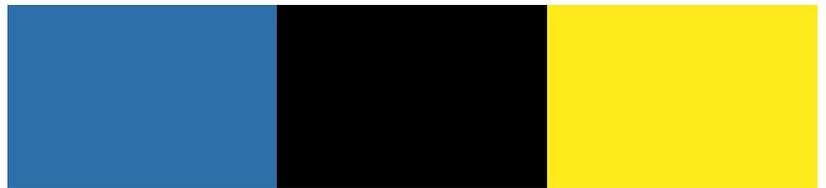




31 décembre 2013

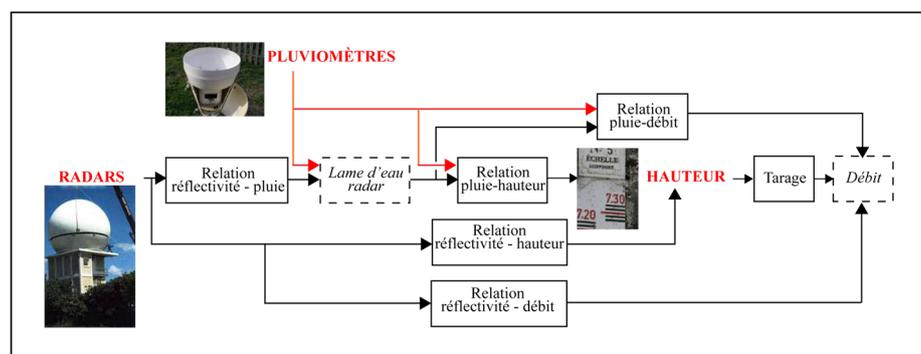
Contributeurs :

Anne Johannet, Guillaume Artigue, Pierre-Alain Ayrat, Audrey Bornancin-Plantier, Gérard Dreyfus, Benjamin Silva, Pierre Roussel, Georges-Marie Saulnier, Isabelle Tachot, Yann Visserot.



Tâche 02

Modélisation de la chaîne hydrométéorologique



Sommaire

1. INTRODUCTION	3
2. PRE-TRAITEMENT DES DONNEES	4
3. IDENTIFICATION DES DIFFERENTES RELATIONS	7
3.1 CRITERES DE PERFORMANCE	7
3.1.1 RMSE	7
3.1.2 Critère de Nash	8
3.1.3 Critère pondéré de pic.....	9
3.1.4 Critères d'estimation de la qualité du pic de crue.....	9
3.2 METHODOLOGIE DE CONCEPTION DE MODELES A RESEAUX DE NEURONES	10
3.2.1 Choix du modèle	10
3.2.2 Méthodologie de conception	11
3.3 RELATION PLUIE PLUVIOMETRE - HAUTEUR	11
3.3.1 Mise en évidence de la sensibilité à l'initialisation des paramètres.....	11
3.3.2 Recherche de l'amélioration des prévisions.....	13
3.4 GENERALISATION AU BASSIN DU GARDON A REMOULINS	18
3.4.1 Site d'étude	18
3.4.2 Types de modèles utilisés	19
3.4.3 Conclusion.....	20
3.5 RELATION REFLECTIVITE - HAUTEUR	21
3.5.1 Données et stratégie de modélisation	21
3.5.2 Résultats et discussion.....	22
3.5.3 Conclusion et perspectives	23
4. RELATION PLUIE-DEBIT SUR LES PETITS BASSINS VERSANTS	24
5. FIABILITE DE LA CHAINE DE PREVISION	24
5.1 INTEGRATION DE LA PLUIE RADAR	24
5.1.1 Données et stratégie de modélisation	24
5.1.2 Résultats et discussion.....	26
5.1.3 Conclusion et perspectives	28
5.2 STRATEGIE D'INTEGRATION DE L'IMAGE RADAR	29
5.2.1 Données et stratégie de modélisation	29
5.2.2 Résultats et discussion.....	30
5.2.3 Conclusion.....	31

6. CONCLUSION	31
7. REFERENCES.....	33
LISTE DES FIGURES	34
LISTE DES TABLEAUX.....	35
ANNEXE 1 : PLUVIOGRAPHES VS PLUIE RADAR	38

1. INTRODUCTION

La tâche 2 du projet ANR FLASH (**F**lood forecasting with machine **L**earning, data **A**ssimilation and **S**emi-**p**hysical modeling) a pour objectif de réaliser la modélisation complète de la chaîne hydrométéorologique, représentée avec des chemins multiples telle que sur la figure 1, avec les réseaux de neurones. Le principe est le suivant : dans une première étape, chacune des boîtes décrivant une relation particulière, par exemple réflectivité-hauteur, peut être modélisée, et ses performances évaluées sur des événements qui n'ont pas été utilisés pour l'apprentissage (exemples de test). Dans un second temps, on cherchera à optimiser la fiabilité de la chaîne, en évaluant le meilleur estimateur des grandeurs qui peuvent être évaluées au moyen de plusieurs chemins.

Dans le projet FLASH initial la tâche 2 comporte 4 sous-tâches :

- T2.1 Pré-traitement des données
- T2.2 Identifications des relations de la chaîne hydrométéorologique
- T2.3 Changement d'échelle pour les très petits bassins
- T2.4 Fiabilité de la chaîne hydrométéorologique

Les trois premières sous-tâches ont pu être traitées dans le projet, néanmoins du fait de l'extrême sensibilité des modèles neuronaux à l'initialisation des paramètres du réseau avant l'apprentissage, sensibilité mise en évidence par la thèse d'Audrey Bornancin-Plantier, la recherche de fiabilisation de la chaîne n'a pu qu'être abordée en tant que telle. L'accent a donc été mis sur la sous-tâche 2.2 afin d'améliorer la robustesse des prévisions fournies pour la relation pluie-hauteur.

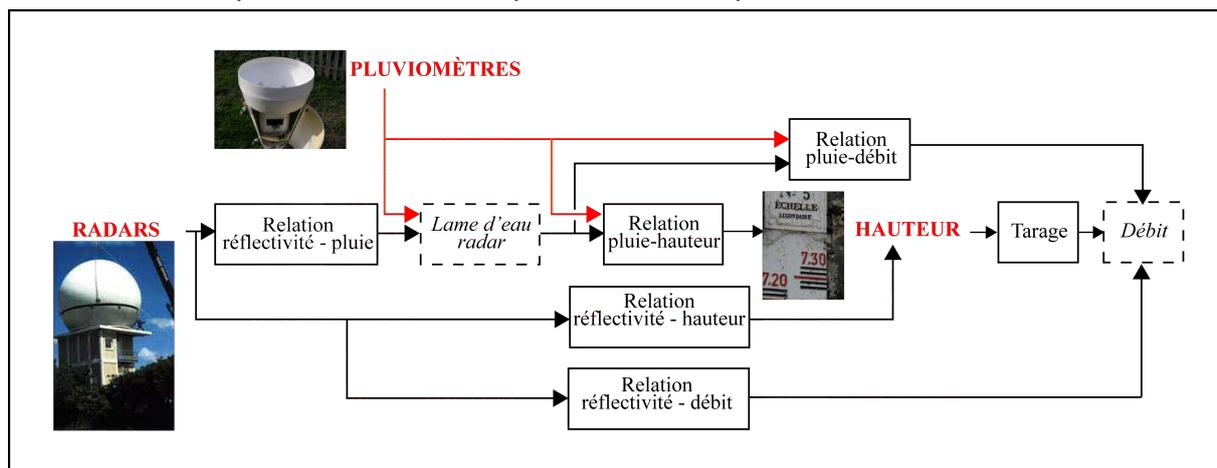


Figure 1: La chaîne hydrométéorologique

Le parti pris choisi est de ne pas présenter l'état de l'art, classique, déjà décrit dans la proposition de projet afin de pouvoir se concentrer sur les apports, scientifiques, méthodologiques ou applicatifs auxquels a abouti cette tâche. Les études et développements effectués pour chacune des sous-tâches sont donc présentés ci-dessous en conservant la proposition initiale du projet. Néanmoins des renvois sont effectués à d'autres rapports si la présentation détaillée de l'étude peut être envisagée avec plus de cohérence par ce moyen.

2. PRE-TRAITEMENT DES DONNEES

Le prétraitement est une phase imposée par l'utilisation des réseaux de neurones. Il s'agit de présenter au réseau des variables conditionnées pour que l'apprentissage se déroule de manière optimale, c'est à dire avec la prise en compte de toutes les variables utiles et sans saturation des sigmoïdes.

Quatre prétraitements ont été identifiés dans la proposition du projet :

1. Réduction de la dimensionnalité des images RADAR

La réduction de la dimensionnalité des images RADAR est décrite dans le présent rapport (section 5.2).

2. Extraction des événements de crue majeurs

L'extraction des événements de crue majeurs a été réalisée par Mohamed Toukourou dans le cadre de sa thèse (Toukourou, 2009) puis complétée par Audrey Bornancin-Plantier avec les événements de 2006 (Tableau 1).

N°	Date	Durée (h)	Pluie moyenne cumulée (mm)	Hauteur maximum (m)	Niveau de vigilance [SPCGD 2010]
1	22-24 Septembre 1994	35	189	3,71	vert
3	3-6 Octobre 1995	53	319	5,34	jaune
4	13-15 Octobre 1995	38	194	5	jaune
8	10-13 Novembre 1996	81	157	2,71	vert
107	4-7 Novembre 1997	89	273	4,20	jaune
108	24-28 Novembre 1997	81	107	2,58	vert
109	16-20 Décembre 1997	103	376	5,37	jaune
117	20-21 Octobre 1999	33	117	3,64	vert
13	28-30 Septembre 2000	45	184	4,80	jaune
120	12-15 Novembre 2000	70	125	2,77	vert
19	8-9 Septembre 2002	35	406	9,71	rouge
22	24-25 Septembre 2006	37	61	2,24	vert
23	18-20 Octobre 2006	54	198	6,61	orange
24	16-18 Novembre 2006	33	59	2,75	vert
25	20-23 Novembre 2007	70	124	2,69	vert
26	21-23 Octobre 2008	51	173	5,57	jaune
27	31 Oct. – 3 Nov. 2008	81	272	5,53	jaune

Tableau 1 : Les événements de la base de données dont les pluies sont fournies par les pluviomètres (Bornancin-Plantier, 2013)

La base comprend donc 17 événements sélectionnés parmi les événements les plus intenses. Il est à noter que le travail d'Audrey Bornancin-Plantier, débuté en 2009, s'est effectué avec une base d'événements sélectionnés sur les hauts débits. Afin de ne pas ajouter de variabilité sur ses travaux du fait de la base de données, celle-ci a été conservée à l'identique durant tous ses travaux, même si une seconde extraction d'événements a été réalisée par Thomas Darras sur un critère lié à la pluviométrie (FLASH_Rapport T1). La liste des événements utilisés par Audrey Bornancin-Plantier pour le travail relatif à la tâche 2 est présentée dans le tableau 1.

Une seconde extraction des événements de crue majeurs a été réalisée par Yann Visserot dans le cadre de son stage de Master 2. La différence avec la précédente est qu'il s'est intéressé aux informations issues des RADARs météorologiques (pluie, réflectivité). Du fait de contraintes technologiques, les deux listes d'événements ne sont pas identiques mais comportent une intersection malheureusement réduite.

N°	Date début (TU)	Date fin (TU)	Durée (h)	Limni max (m)	Vigilance
12	19/09/2000 05:00	20/09/2000 20:00	39,00	2,72	vert
13	27/09/2000 06:00	30/09/2000 09:30	51,50	4,80	jaune
120	11/11/2000 05:00	15/11/2000 09:00	76,00	2,77	vert
15	14/03/2002 03:00	16/03/2002 10:00	55,00	3,67	jaune
19	08/09/2002 09:30	10/09/2002 03:00	41,50	10,43	rouge
200	15/11/2003 18:00	18/11/2003 03:30	57,50	3,81	jaune
201	21/11/2003 14:30	25/11/2003 08:30	90,00	5,06	jaune
21	05/09/2005 04:00	10/09/2005 07:30	123,50	2,06	vert
22	24/09/2006 05:30	25/09/2006 15:00	33,50	2,27	vert
23	18/10/2006 13:30	21/10/2006 07:30	66,00	6,54	orange
24	16/11/2006 00:00	18/11/2006 07:00	55,00	2,72	vert

Tableau 2 : Les événements de la base de données dont les pluies sont fournies par les RADARs (Visserot, 2012)

Enfin une troisième extraction a eu lieu concernant les événements de crue à Remoulins, à l'aval du bassin versant du Gardon, ils sont présentés dans le Tableau 3.

<i>N°</i>	<i>Date</i>	<i>Durée (h)</i>	<i>Cumul de pluie à Remoulins (mm)</i>	<i>Hauteur maximum (m)</i>	<i>Niveau de vigilance [SPCGD 2010]</i>
3	3-6 Octobre 1995	66	67	3,80	jaune
4	13-15 Octobre 1995	54	36	3,80	jaune
8	10-14 Novembre 1996	105	1	3,00	jaune
107	4-8 Novembre 1997	101	36	2,70	vert
108	24-27 Novembre 1997	72	4	1,87	vert
109	15-21 Décembre 1997	142	58	4,74	jaune
117	20-22 Octobre 1999	59	107	1,26	vert
13	28-30 Septembre 2000	60	52	2,89	vert
120	12-15 Novembre 2000	80	10	2,10	vert
22	23-26 Septembre 2006	70	63	1,55	vert
23	18-21 Octobre 2006	64	39	4,10	jaune
24	16-18 Novembre 2006	47	49	2,79	vert
25	20-24 Novembre 2007	84	85	2,37	vert
26	21-23 Octobre 2008	53	10	3,93	jaune
27	31 Octobre – 4 Novembre 2008	95	36	4,14	jaune

Tableau 3 : Base de données comportant le numéro de l'événement, la date, la durée, le cumul de pluie à Remoulins, la hauteur maximale à Remoulins et le niveau de vigilance (Bornancin-Plantier, 2013)

3. Normalisation des variables.

La normalisation des variables est effectuée de la manière suivante : chacune des variables (pluies) est réduite dans l'intervalle 0-0,9. Compte-tenu de la distribution statistique très spécifique des variables hydrométéorologiques, cette normalisation s'est avérée plus efficace que la normalisation classique dans l'intervalle [-1,1].

4. Elaboration d'une typologie.

Deux typologies différentes ont été établies durant le projet : la première est effectuée sur des critères hydrologiques et est présentée dans le Tableau 4, la seconde est présentée dans le rapport de la tâche 4, elle utilise une classification hiérarchique avec des machines à vecteurs supports (Boukharouba, 2013).

Plus d'information sur le principe de la typologie peut être trouvé dans (Audrey Bornancin-Plantier, 2013).

Intense	Long	Monopic	Classe
Oui	Oui	Oui	A
Oui	Oui	Non	B
Oui	Non	Oui	C
Oui	Non	Non	D
Non	Oui	Oui	E
Non	Oui	Non	F
Non	Non	Oui	G
Non	Non	Non	H

Tableau 4 : Caractérisation des événements (Bornancin-Plantier, 2013)

3. IDENTIFICATION DES DIFFERENTES RELATIONS

Sur l'ensemble des relations proposées dans le schéma-bloc de la Figure 1, la relation la plus étudiée a été la relation Pluie-hauteur, celle-ci distinguée en 3 relations possibles :

Réfectivité-hauteur, dont l'étude est présentée en section 3.5 de ce rapport.

Pluie RADAR-hauteur, dont l'étude est présentée en section 5.1 de ce rapport.

Pluie pluviomètre-hauteur, dont l'étude a été menée exhaustivement par la thèse d'Audrey-Bornancin Plantier. Cette étude est présentée synthétiquement en section 3.3 ci-dessous et plus exhaustivement dans son manuscrit de thèse (Audrey Bornancin-Plantier, 2013).

Trois autres relations devaient être étudiées ; les relations **pluie-débit**, sans passer par la variable hauteur-débit (pour les trois entrées de pluie identifiées ci-dessus) : l'idée était de concevoir des modèles pluie-débit directement ; et de comparer les prévisions obtenues par les deux chemins : pluie-hauteur-débit et pluie-débit. Il est apparu, en première analyse, que les comportements des relations pluie-hauteur et pluie-débit étaient sensiblement identiques ; ainsi la conception de la relation **hauteur-débit** perdait de son intérêt ; elle n'a pas été réalisée exhaustivement.

La relation pluie-hauteur a également été étudiée sur le bassin versant du Gardon à Remoulins.

3.1 CRITERES DE PERFORMANCE

Afin d'estimer la qualité des résultats d'un modèle et de pouvoir comparer ses performances avec celles d'autres modèles, on utilise des critères de performance. Parmi tous les critères existants nous en avons sélectionné quatre qui paraissent appropriés pour évaluer les crues et les performances en prévision.

3.1.1 RMSE

Le RMSE ou Root Mean Squared Error, est la racine carrée de l'erreur quadratique moyenne. Ce critère a la même dimension que la grandeur étudiée, ce qui donne un ordre de grandeur de l'erreur commise.

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_k^p - y_k)^2} \quad \text{Eq. 1}$$

avec :

- y_k^p sortie mesurée à l'instant k
- y_k sortie calculée à l'instant

Cependant, ce critère ne permet pas de comparer la performance du modèle entre plusieurs événements d'amplitudes très différentes. En effet un même RSME sur un événement avec une hauteur de pic importante correspond à un meilleur résultat relatif que pour un événement de faible hauteur de pic. C'est pourquoi il est intéressant de calculer également des critères normés comme le critère de Nash ou la Persistance.

3.1.2 Critère de Nash

Le R^2 , ou critère de Nash [Nash 1970] rapporte l'erreur quadratique moyenne à la variance des mesures.

$$C_N = 1 - \frac{\sum_{k=1}^n (y_k^p - y_k)^2}{\sum_{k=1}^n (y_k^p - \overline{y^p})^2} \quad \text{Eq. 2}$$

avec :

- y_k^p sortie mesurée à l'instant k
- y_k sortie calculée à l'instant k
- $\overline{y^p}$ moyenne de la sortie mesurée sur l'ensemble considéré

Les valeurs du critère de Nash sont comprises entre $-\infty$ et 1. Le critère de Nash vaut 1 quand la sortie calculée est égale à la sortie mesurée, donc quand le modèle est parfait. En hydrologie, le critère de Nash est considéré comme bon lorsqu'il a une valeur supérieure à 0,8.

Ce critère est intéressant car il permet de comparer des événements pluvieux de caractéristiques différentes. Néanmoins il souffre de deux inconvénients : il prend en compte toutes les valeurs de l'hydrogramme alors que celles du pic de crue sont plus importantes, d'autre part il ne tient pas spécifiquement compte de la synchronisation du pic de crue. On peut donc obtenir de très bons critères de Nash avec des prévisions parfaitement inutiles en termes opérationnels. C'est pourquoi on utilise également d'autres critères.

Le critère de persistance [Kitadinis 1980] permet de comparer la prévision du modèle à la prévision naïve, c'est-à-dire la prévision qui donne à la grandeur concernée une valeur future égale à la valeur actuelle.

$$C_p = 1 - \frac{\sum_{k=1}^n (y_{k+h_p}^p - y_{k+h_p})^2}{\sum_{k=1}^n (y_{k+h_p}^p - y_k^p)^2} \quad \text{Eq. 3}$$

avec :

- $y_{k+h_p}^p$ sortie mesurée à l'instant $k+h_p$ où h_p est l'horizon de prévision
- y_{k+h_p} sortie calculée à l'instant $k+h_p$
- y_k^p sortie mesurée à l'instant k

Le critère de persistance peut prendre des valeurs comprises entre $-\infty$ et 1. Lorsque le critère de persistance est égal à 1, cela signifie que les valeurs calculées par le modèle sont égales aux valeurs mesurées ; le modèle est le meilleur possible. Si le critère de persistance est nul, alors les valeurs calculées par le modèle sont de qualité équivalente aux valeurs de la prévision naïve qui consiste à prédire que la grandeur à l'instant $k+1$ aura la même valeur que la grandeur mesurée à l'instant k . Si le critère de persistance est négatif, le modèle donne de moins bons résultats que la prévision naïve : il présente donc peu d'intérêt.

3.1.3 Critère pondéré de pic

Le critère Peak-Weighted Root Mean Squared Error ou PWRMSE [Amengual 2007], pondère l'erreur quadratique par un coefficient qui augmente avec la valeur de la grandeur concernée. Ce coefficient est donc plus grand pour les valeurs du pic de crue, ce qui donne plus d'importance à l'erreur quadratique au niveau du pic, et par conséquent reflète mieux la performance du modèle au niveau du pic, qui est la zone la plus intéressante pour la prévision des crues.

$$PWRMSE = \left[\frac{1}{2n} \sum_{k=1}^n \left[(y_k^p - y_k)^2 \left(1 + \frac{y_k^p}{y^p} \right) \right] \right]^{\frac{1}{2}} \quad \text{Eq. 4}$$

avec :

- y_k^p sortie mesurée à l'instant k
- y_k sortie calculée à l'instant k
- y^p moyenne de la sortie mesurée

3.1.4 Critères d'estimation de la qualité du pic de crue

Pour estimer la qualité des prévisions au niveau du pic de crue, nous avons défini trois critères qui évaluent l'amplitude du pic prévu et sa synchronisation par rapport au pic observé :

- le Percentage Peak Discharge, *PPD*, qui correspond au pourcentage de la hauteur maximale calculée du pic par rapport à la hauteur maximale du pic observée,
- le Synchronous Percentage Peak Discharge, *SPPD*, qui est le pourcentage de la hauteur calculée à l'instant de la hauteur maximale du pic observée par rapport à la hauteur maximale du pic observée,
- le décalage temporel entre la hauteur maximale du pic prévu et la hauteur maximale du pic observé, que l'on appellera Retard H_{pic} .

Ces critères ont été utilisés pour effectuer (i) l'apprentissage, (ii) la sélection du modèle et (iii) pour comparer les différentes prévisions effectuées.

3.2 METHODOLOGIE DE CONCEPTION DE MODELES A RESEAUX DE NEURONES.

3.2.1 Choix du modèle

Le réseau de neurones utilisé est un perceptron multicouche comportant une couche de neurones cachés avec des tangentes hyperboliques comme fonction d'activation et un neurone de sortie linéaire.

Le prédicteur retenu (Bornancin-Plantier, 2013) est celui qui est associé à l'hypothèse de bruit d'état. Ce modèle utilise les valeurs précédentes observées de la sortie qu'il doit estimer : l'apprentissage du réseau est dirigé (Figure 2). Cela évite les divergences que pourrait subir un modèle associé à l'hypothèse de bruit de sortie qui calcule sa sortie à partir des valeurs précédentes de celle-ci. Ces valeurs calculées étant soumises à des erreurs, les suivantes risquent de s'éloigner encore plus des sorties observées.

Le modèle réalise la fonction suivante :

$$y(k+h_p) = \varphi_{RN} (y^p(k), \dots, y^p(k-r+1), \mathbf{u}(k), \dots, \mathbf{u}(k-f_T+1), \mathbf{w}) \quad \text{Eq. 5}$$

avec :

- $y(k+h_p)$ la sortie calculée par le modèle : la hauteur prévue à Anduze pour l'instant $k+h_p$
- h_p l'horizon de prévision
- $y^p(k)$ la sortie observée (la hauteur d'eau à Anduze) à l'instant k
- f_T la fenêtre temporelle des pluies
- r la fenêtre temporelle des hauteurs
- \mathbf{u} le vecteur les variables exogènes : les pluies mesurées aux 6 pluviomètres
- \mathbf{w} le vecteur des paramètres

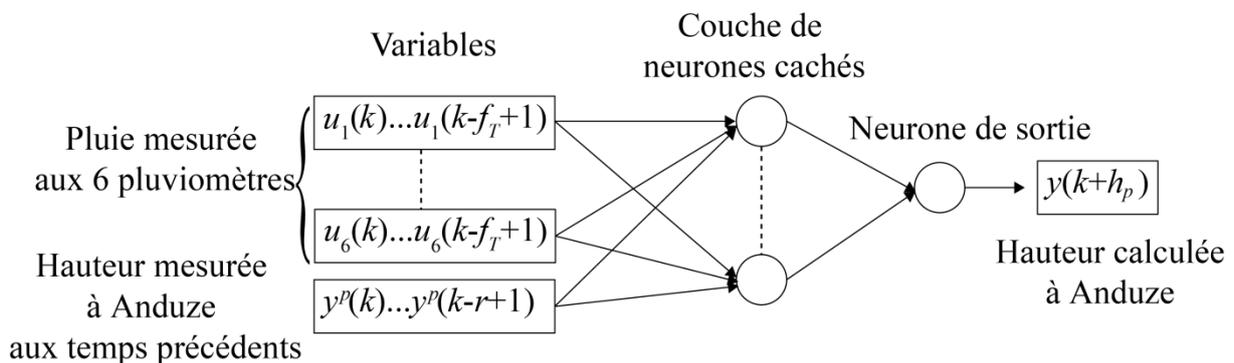


Figure 2 : Schéma du perceptron multicouche utilisé (Bornancin-Plantier, 2013)

Afin de prendre en compte les différents horizons de prévision, et en l'absence de prévisions de pluies qui sont indisponibles de manière fiable pour de si courts laps de temps et une localisation si précise, le principe est d'effectuer la prévision des hauteurs d'eau futures jusqu'à 5h, en prenant en compte les valeurs passées et actuelles de la hauteur d'eau et des précipitations. Les pluies prévues ne sont pas utilisées. Ceci nous a donc conduit à réaliser 6 modèles de prévision pour les valeurs d'anticipation suivantes : 1/2h, 1h, 2h, 3h, 4h, 5h.

3.2.2 Méthodologie de conception

La méthodologie de conception de chaque modèle est la suivante. Comme la complexité du modèle joue un rôle important dans le dilemme biais-variance : un modèle de faible complexité donne de mauvais résultats en apprentissage et par conséquent en généralisation, alors qu'un modèle trop complexe est trop ajusté à l'ensemble d'apprentissage et ne peut pas généraliser de manière satisfaisante. Il convient de trouver un compromis : nous avons donc cherché à optimiser la complexité du modèle, qui, pour les réseaux de neurones, est liée essentiellement au nombre de neurones cachés et au nombre de variables. Pour cela deux méthodes de régularisation ont été utilisées : l'arrêt précoce et la validation croisée.

La validation croisée permet d'obtenir le modèle de plus faible variance. Elle a été utilisée, dans sa forme de validation croisée partielle, variante de la validation croisée n'utilisant que les événements les plus intenses pour calculer le score de validation croisée (définie dans Toukourou, 2009), pour optimiser le nombre de neurones cachés et le nombre des variables. Puis elle a été utilisée pour étudier l'influence de deux hyperparamètres de l'algorithme d'apprentissage sur la capacité de généralisation : l'initialisation du pas d'apprentissage de l'algorithme de Levenberg-Marquardt et le facteur décroissance de ce pas au fil des itérations. Enfin elle a permis de définir la meilleure initialisation des paramètres du réseau.

3.3 RELATION PLUIE PLUVIOMETRE - HAUTEUR

3.3.1 Mise en évidence de la sensibilité à l'initialisation des paramètres.

La relation pluie-hauteur lorsque la pluie tombée est estimée à partir de mesures issues de pluviomètres a occupé l'essentiel des efforts dévolus à la tâche 2. En effet, les études précédentes ont mal estimé l'importance essentielle que joue l'initialisation préliminaire des paramètres du réseau de neurones dans la qualité des modèles réalisés. Ce point a été soulevé par la thèse d'Audrey Bornancin qui a pu mettre ce phénomène en évidence de manière spectaculaire.

Les pluviomètres utilisés sont ceux du Service de Prévision des Crues Grand Delta (SPCGD) basé à Nîmes et présentés dans le rapport de la tâche 1 (Marchandise *et al*, 2013) : Anduze, Mialet, Saumane, Barre-des-Cévennes, Saint-Roman-de-Tousque et Soudorgues.

Les événements de la base de données présentée dans le tableau 1 ont été répartis comme suit dans les différents ensembles :

- L'ensemble d'apprentissage comprend tous les événements de la base excepté les événements 13, 19, 26 et 27.
- L'événement 13 est choisi pour l'arrêt précoce car c'est l'événement qui donne les meilleurs résultats lorsqu'il est utilisé en test. Il est donc bien représentatif de l'ensemble de la base d'apprentissage et il y a une cohérence à le choisir pour arrêter l'apprentissage sans ajouter une spécialisation liée à cet événement.
- Les événements 19, 26 et 27 sont réservés pour le test car ce sont des événements intenses. D'ailleurs, les événements 26 et 27 ont été choisis comme événements tests par le SCHAPI dans le projet BVNE (Bassins Versants Numériques Expérimentaux). Nous avons également retenu l'événement 19 car il s'agit de l'événement le plus intense de la base de données. Cependant les mesures ont été en grande partie reconstruites a posteriori : les erreurs sur les valeurs de hauteur sont donc probablement très importantes.

- Les événements servant d'ensemble de « validation » pour la validation croisée partielle sont les événements 3, 4, 109 et 23, événements parmi les plus intenses de la base d'apprentissage puisque que l'on cherche à sélectionner les modèles qui prévoient le mieux les événements les plus intenses.

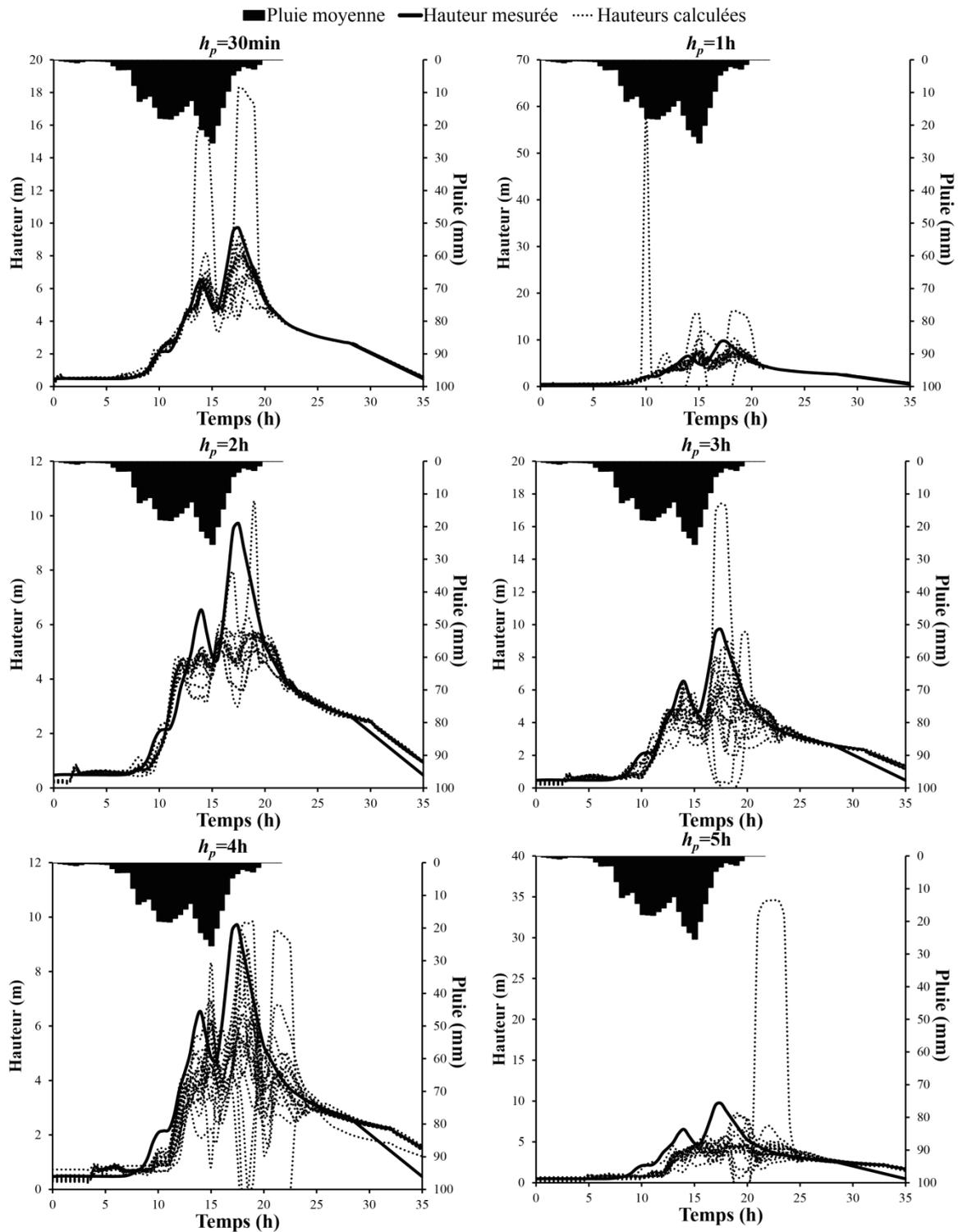


Figure 3 : Événement 19, les 20 hauteurs prévues par les 20 modèles sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013)

Une fois l'architecture du modèle sélectionnée grâce à la méthodologie précédemment présentée, on regarde quelle est la sensibilité de la prévision en fonction de l'initialisation du modèle.

On trouve que, par exemple, pour l'événement 19 (septembre 2002), le plus intense de la base, les prévisions sont très dispersées en fonction de l'initialisation des paramètres et peuvent atteindre des valeurs aberrantes (prévision d'une hauteur de 18m pour l'horizon de prévision de 1/2h, alors que la hauteur maximale de la base est de 9,70m comme reporté en tableau 1). Il apparaît donc que la validation croisée, si elle conduit à des modèles ayant un sens physique compréhensible lorsqu'elle permet de sélectionner par exemple les historiques de pluie, en revanche échoue à sélectionner la meilleure initialisation des paramètres. Les meilleurs scores de validation croisée, ne conduisent pas en général à des modèles efficaces sur les ensembles de test alors que ces modèles existent dans l'ensemble des modèles avant la sélection.

Face à ce constat il a été nécessaire de trouver des méthodes de sélection plus efficaces ou des stratégies de construction de modèles différentes.

3.3.2 Recherche de l'amélioration des prévisions

Impact de l'utilisation de différents critères d'apprentissage ou de sélection

Une première voie investiguée afin de construire ou de sélectionner de meilleurs modèles a été de définir d'autres critères que celui de la seule erreur quadratique moyenne. Parmi les critères présentés en 3.1, on peut distinguer plusieurs groupes. Le premier regroupe les critères de l'erreur quadratique moyenne, de Nash, de la persistance et de PWRMSE qui sont fondés sur une erreur quadratique avec quelques variantes et qui peuvent être utilisés pour réaliser l'apprentissage ou la sélection de modèles. En ce qui concerne l'apprentissage, on n'a pas noté de différences notables à utiliser l'un ou l'autre des critères ; il est donc proposé de rester sur l'erreur quadratique moyenne qui a l'avantage d'être d'un usage très répandu. En ce qui concerne la sélection de modèles, on peut noter sur l'exemple fourni en figure 4, que tous les critères de ce premier groupe fournissent des scores qui évoluent dans le même sens ; aucun n'a un avantage distinctif des autres. Considérant le second groupe des critères définis : celui des critères relatifs à la qualité du pic, ces derniers ne sont pas facilement automatisables pour des événements complexes, mais les premières investigations « manuelles » n'ont pas non plus permis de dégager d'avantage clair dans leur utilisation pour la sélection de modèle.

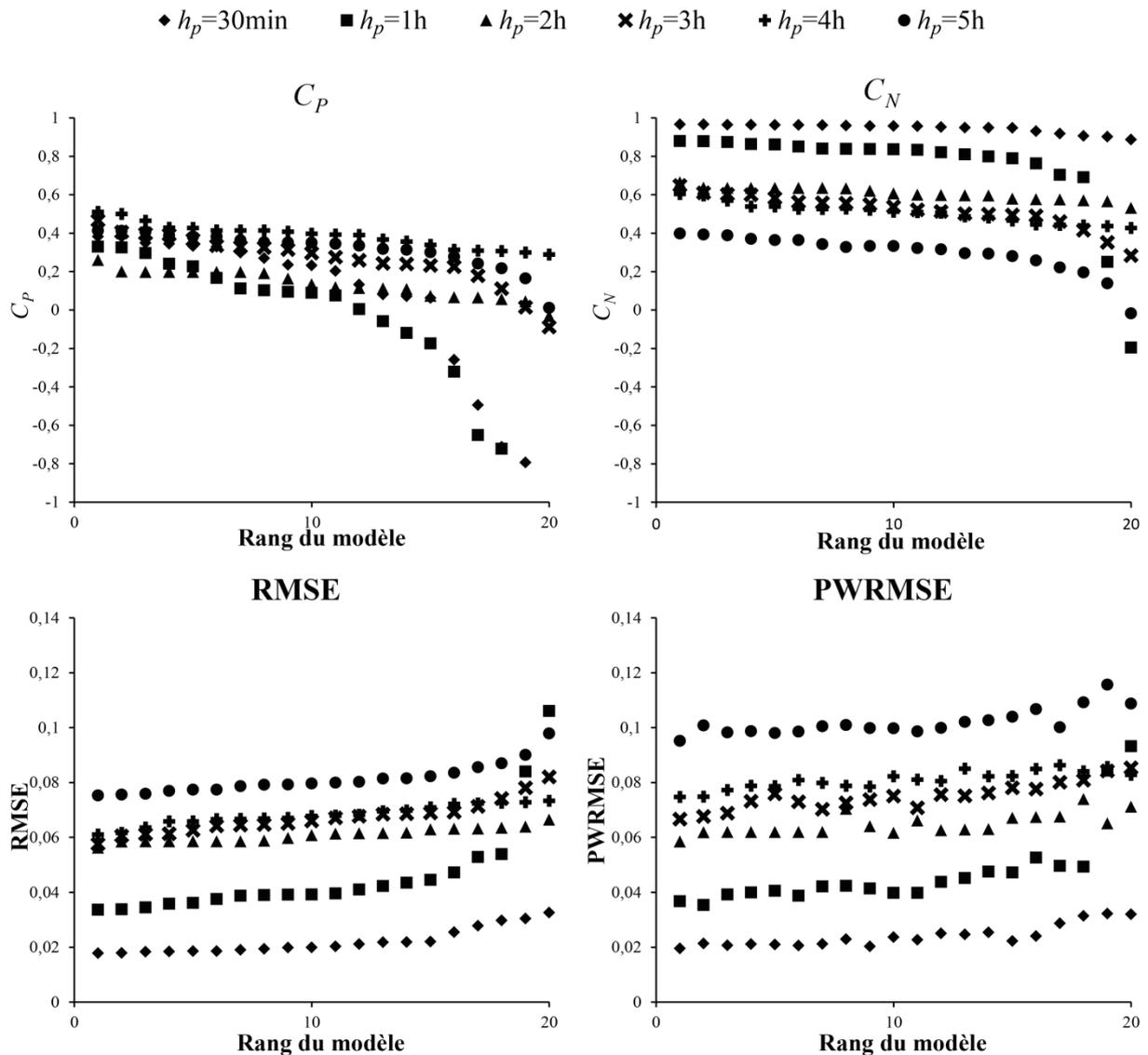


Figure 4 : Valeurs des critères C_p , C_N , RMSE et PWRMSE pour le test de l'événement 26 sur les 20 meilleurs modèles, au sens de la validation croisée partielle, pour chaque horizon de prévision. L'axe des abscisses correspond au rang du modèle par C_p décroissants (Bornancin-Plantier, 2013)

Calcul de prévision d'ensemble

Un moyen d'obtenir une prévision robuste, bien que de qualité moyenne, consiste à calculer la médiane des prévisions délivrées par un ensemble de modèles. Cette solution a été proposée et évaluée dans le projet FLASH. Il apparaît qu'elle apporte effectivement une stabilisation de la prévision obtenue, ce qui rend ainsi l'approche utilisable en opérationnel. A titre d'illustration des résultats ainsi obtenus, on reporte en figure 5 les prévisions délivrées par le modèle d'ensemble pour l'événement le plus intense de la base. La Figure 5 est à rapprocher de la Figure 3 pour évaluer l'apport de l'approche ensembliste.

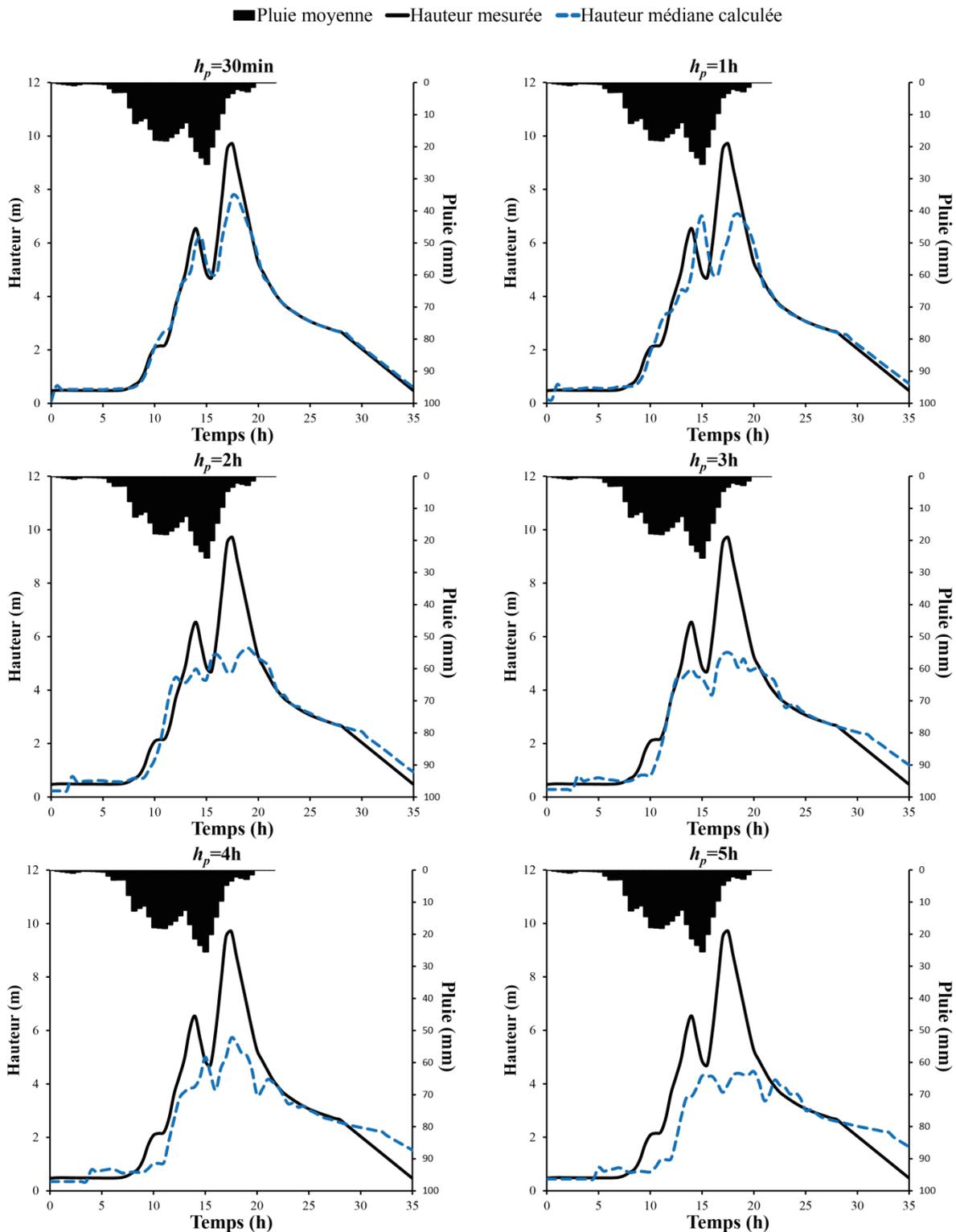


Figure 5 : Événement 19, hauteur médiane calculée sur les hauteurs prévues par les 20 modèles sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013)

Modèles comportant des liens directs

L'étude de la sensibilité de la prévision obtenue en fonction de la valeur initiale des paramètres du modèle a conduit à l'hypothèse que la surface de minimisation non convexe comportait trop de minima locaux empêchant ainsi la convergence vers un minimum global. Afin d'aider cette minimisation, deux pistes ont été envisagées :

rechercher une meilleure stratégie d'initialisation ou rendre la surface « plus convexe ». La deuxième voie a été choisie en ajoutant des termes linéaires, ou liens directs, au modèle initial tel que représenté en figure 6.

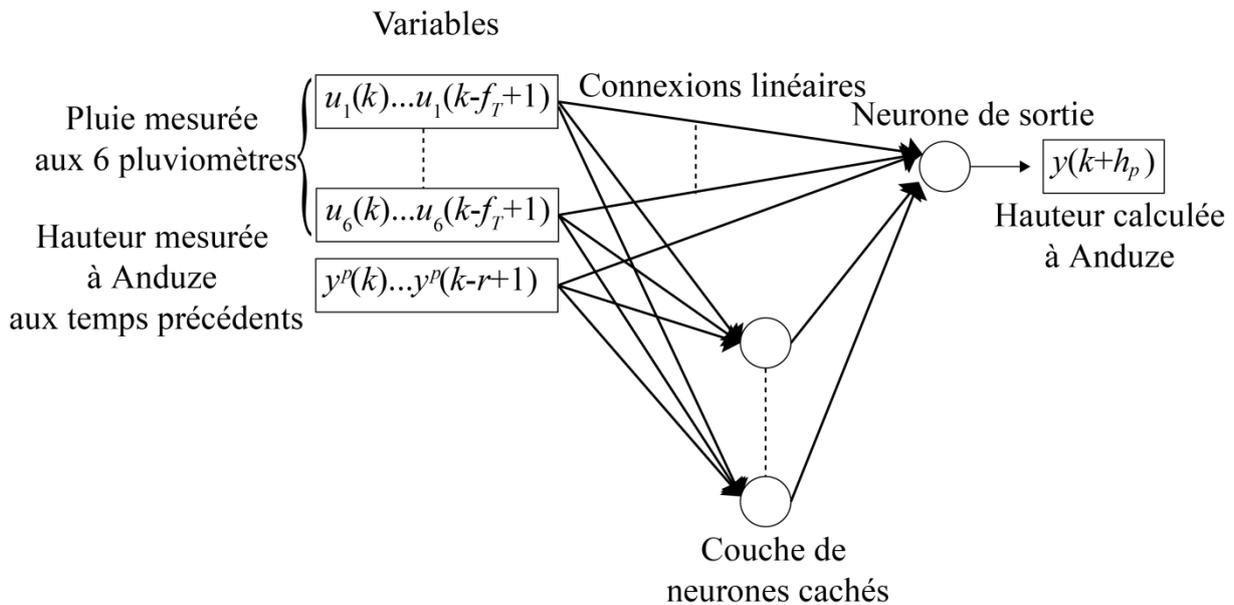


Figure 6: Modèle modifié avec introduction d'une partie linéaire (Bornancin-Plantier, 2013)

En cumulant les deux améliorations : lien directs et modèle d'ensemble, une illustration des prévisions obtenues est représentée en Figure 7. On peut y noter une nette amélioration quantifiée dans le Tableau 5.

Médianes		$h_p=0,5h$	$h_p=1h$	$h_p=2h$	$h_p=3h$	$h_p=4h$	$h_p=5h$
C_P	PM+PL	0,54	0,48	0,70	0,81	0,68	0,65
	PM	0,00	0,08	0,46	0,64	0,69	0,62
C_N	PM+PL	0,98	0,92	0,88	0,88	0,73	0,59
	PM	0,96	0,86	0,78	0,78	0,73	0,56
PPD (%)	PM+PL	99	97	83	88	68	89
	PM	80	73	57	56	59	46
SPPD (%)	PM+PL	99	83	69	88	59	49
	PM	80	63	48	56	59	41
Retard H_{pic} (h)	PM+PL	0	0,5	2	0	1,5	1,5
	PM	0	1	1,5	0	0	2,5

Tableau 5 : Valeurs du C_P , C_N , PPD, SPPD et Retard H_{pic} suivant l'horizon de prévision pour la hauteur prévue médiane des sorties sur l'événement 19. Comparaison entre les modèles à perceptron multicouche uniquement (PM) et ceux avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)

On peut ainsi noter que, sur l'événement le plus intense de la base, les prévisions ainsi obtenues en l'absence de prévision de précipitation permettent d'envisager une utilisation de cette méthode par le SCHAPI dans sa chaîne de prévision temps réel.

A cette fin, un prototype a été réalisé durant le projet et livré au SCHAPI. Il est actuellement implémenté et en phase de test dans cette chaîne de prévision.

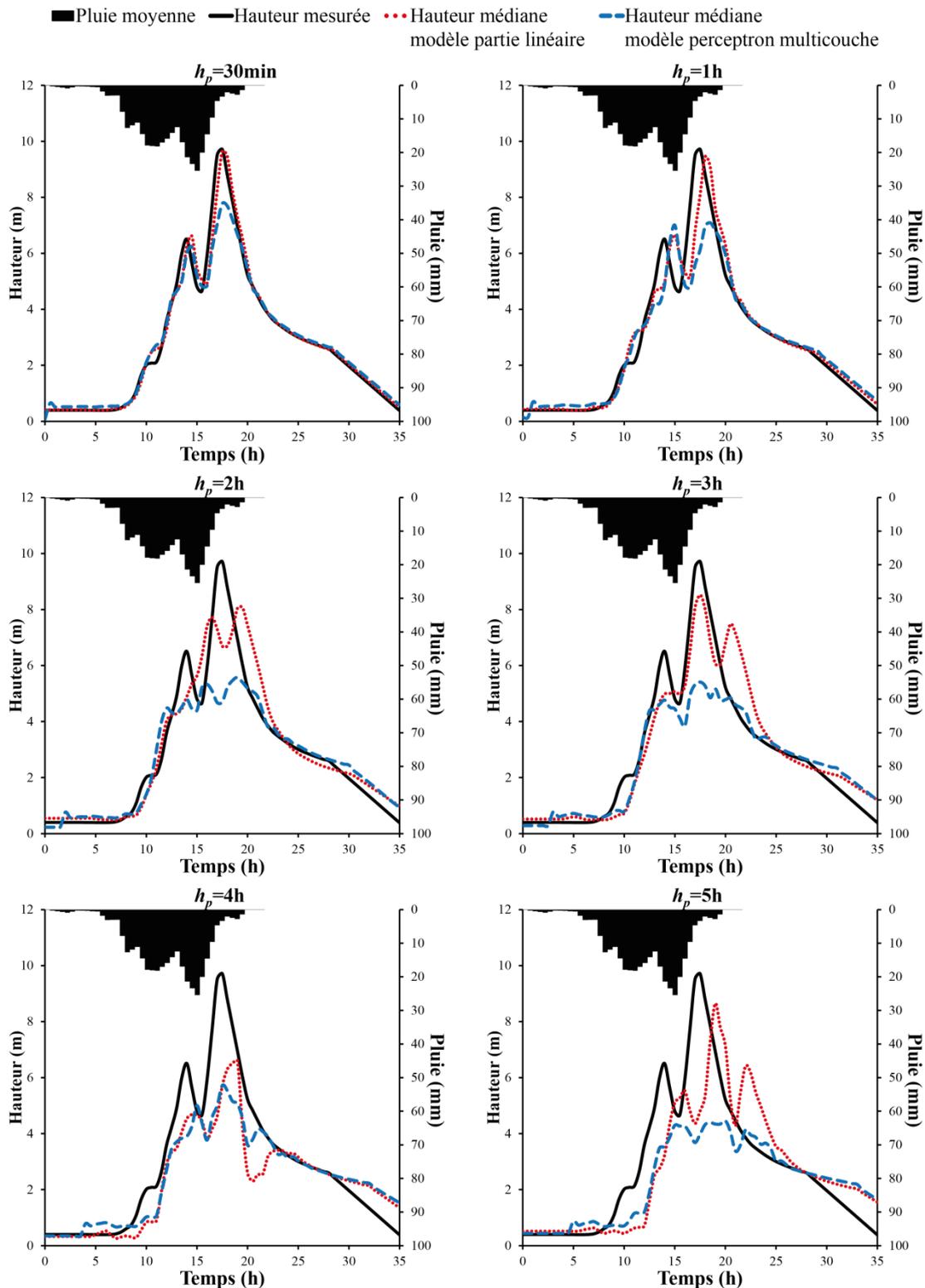


Figure 7 : Événement 19, hauteurs médianes sur les hauteurs prévues par les 20 modèles avec partie linéaire et les 20 modèles à perceptron multicouche sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013).

Ce travail a été soumis dans une publication à la revue neurocomputing (Bornancin-Plantier, 2013b).

3.4 GENERALISATION AU BASSIN DU GARDON A REMOULINS

Au-delà de la prévision des crues rapides sur les amonts du Gardon, comme à Anduze, le projet FLASH prévoyait d'étudier la généralisation des méthodes développées à la modélisation de la relation pluie-hauteur à l'ensemble du bassin versant du Gardon c'est-à-dire au bassin versant du Gardon à Remoulins. En suivant la méthodologie définie sur le bassin versant du Gardon d'Anduze (modèle d'ensemble avec connexions linéaires), on créera des modèles à différents horizons de prévision qui prévoiront la hauteur du Gardon à Remoulins.

3.4.1 Site d'étude

Le bassin versant du Gardon à Remoulins, représenté Figure 8, a une superficie de 1848 km². Le Gardon d'Anduze et le Gardon d'Alès se rejoignent un peu avant Ners pour former le Gardon. La topographie du bassin versant du Gardon peut être décrite en quatre parties. De l'amont vers l'aval :

- les Cévennes avec de fortes pentes et des vallées encaissées,
- le piémont cévenol avec des reliefs moins prononcés et des plaines assez étroites,
- la Gardonnenque, qui commence peu après Ners, est une large plaine constituée d'alluvions du tertiaire et du quaternaire,
- les gorges calcaires qui s'étendent sur environ 25 km et qui se terminent quelques kilomètres avant Remoulins ; le lit du Gardon y est très encaissé.
- Ces variations importantes de topographie ont pour conséquences des régimes hydrauliques très différents qui doivent être pris en compte dans les modèles hydrologiques.



Figure 8 : Bassin versant du Gardon (conception G. Artigue)

Les stations de mesures que nous avons utilisées sont indiquées sur la Figure 8. La prévision de la hauteur à Remoulins est réalisée grâce aux hauteurs d'eau amont mesurées à Ners, Alès et Anduze, ainsi qu'aux précipitations reçues par l'aval du bassin à Ners et à Remoulins. De cette manière la conversion des pluies intense reçues par les amonts en hauteur à Anduze et Alès n'est plus à réaliser par le

modèle, et, comme nous le verrons, le temps de propagation de la crue depuis les stations d'Alès et d'Anduze confère une capacité d'anticipation suffisante au modèle.

3.4.2 Types de modèles utilisés

Les deux types de modèles utilisés précédemment pour la prévision des crues du Gardon d'Anduze ont été étudiés, seuls les résultats du modèle couplant les liens linéaires et non linéaires sont reportés dans ce rapport ; le travail complet pourra être trouvé dans le mémoire de thèse (Bornancin-Plantier, 2013).

Le modèle utilisé est représenté en (Figure 9).

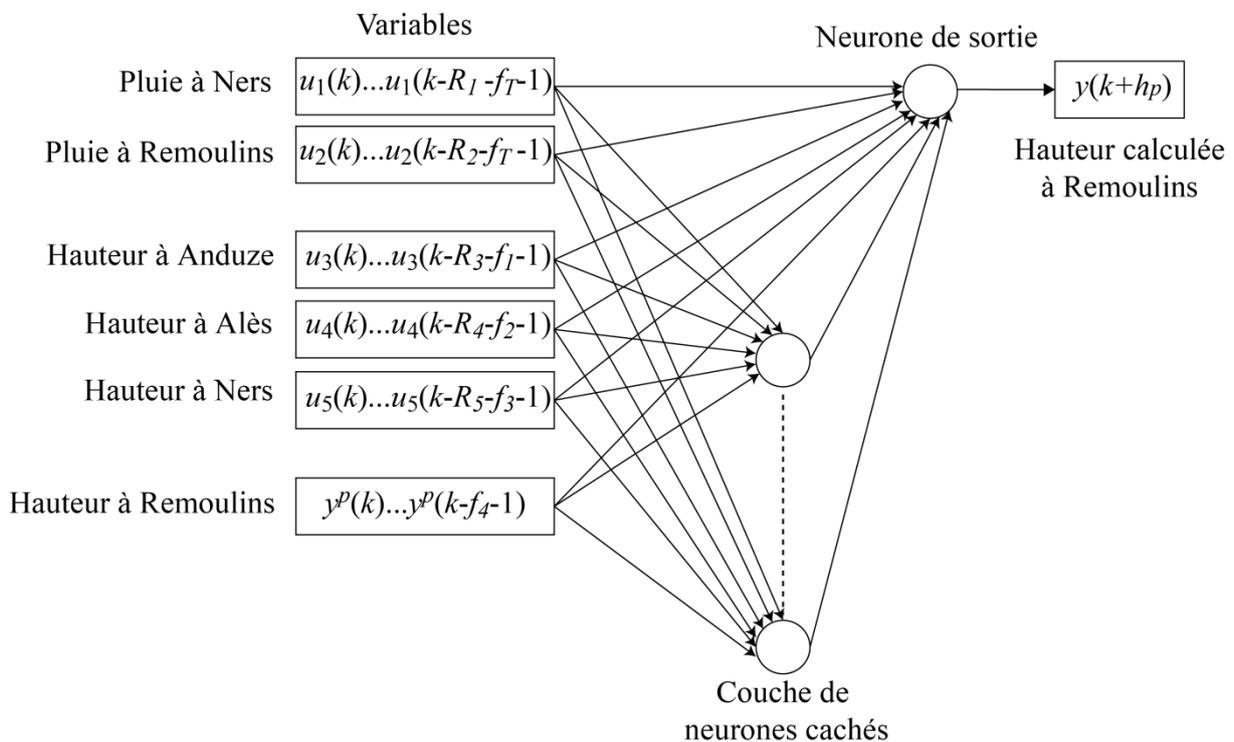


Figure 9 : Modèle avec ajout d'une partie linéaire au perceptron multicouche avec h_p l'horizon de prévision, f_T la fenêtre temporelle des pluies, f_1 à f_4 les fenêtres temporelles pour les hauteurs correspondantes et R_1 à R_5 les retards appliqués aux pluies et hauteurs correspondantes (Bornancin-Plantier, 2013)

On réalise des modèles pour chaque horizon de prévision : 1h, 3h, 5h, 7h, 9h et 11h. Dans chaque cas, on choisit les fenêtres temporelles (f_T), retard initial (R_i) et hyperparamètres comme présenté précédemment avec la validation croisée partielle.

On compare les performances sur deux événements de test : les événements 26 et 27 communs avec la base d'Anduze, l'événement 19 n'étant pas disponible sur Remoulins. Les résultats sont synthétisés dans les (Tableau 6, Tableau 7).

<i>Événement 26</i>		$h_p=1h$	$h_p=3h$	$h_p=5h$	$h_p=7h$	$h_p=9h$	$h_p=11h$
C_P	PM	0,81	0,85	0,87	0,89	0,91	0,87
	PM+PL	0,77	0,80	0,89	0,92	0,93	0,87
C_N	PM	0,99	0,93	0,87	0,82	0,80	0,66
	PM+PL	0,99	0,91	0,89	0,86	0,84	0,66
<i>PPD (%)</i>	PM	101	94	87	83	105	105
	PM+PL	106	87	96	80	95	118
<i>SPPD (%)</i>	PM	100	86	65	82	88	42
	PM+PL	96	85	80	79	79	50
Retard H_{pic} (h)	PM	1	1	1	0,5	0,5	2,5
	PM+PL	0,5	0,5	0,5	0,5	0,5	2,5

Tableau 6 : Pour chaque horizon de prévision, critères de performance calculés sur les prévisions médianes de l'événement 26 obtenues par les modèles à perceptron multicouche (PM) et avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)

<i>Événement 27</i>		$h_p=1h$	$h_p=3h$	$h_p=5h$	$h_p=7h$	$h_p=9h$	$h_p=11h$
C_P	PM	0,74	0,86	0,92	0,94	0,96	0,93
	PM+PL	0,67	0,83	0,89	0,96	0,96	0,90
C_N	PM	0,99	0,98	0,97	0,97	0,97	0,93
	PM+PL	0,99	0,97	0,96	0,98	0,97	0,90
<i>PPD (%)</i>	PM	100	99	96	95	96	103
	PM+PL	102	98	99	91	95	108
<i>SPPD (%)</i>	PM	99	94	88	81	89	102
	PM+PL	98	92	84	89	89	102
Retard H_{pic} (h)	PM	1	1,5	2	1,5	-2	1
	PM+PL	1	1,5	1,5	2,5	-0,5	1

Tableau 7 : Pour chaque horizon de prévision, critères de performance calculés sur les prévisions médianes de l'événement 27 obtenues par les modèles à perceptron multicouche (PM) et avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)

Les limnigrammes de prévision des événements 26 et 27 pourront être trouvés dans le mémoire de thèse (Bornancin-Plantier, 2013).

3.4.3 Conclusion

En conclusion de cette première tentative de généralisation des modèles de prévision de la relation pluie-hauteur depuis le bassin d'Anduze vers un bassin 3,5 fois plus grand, il apparaît que si les modèles ne sont pas directement transposables, en revanche, la méthodologie de conception est tout à fait adaptée et permet d'envisager la réalisation de systèmes d'alerte performants jusqu'à 7 heures à l'avance pour la station de Remoulins. Le remplacement des pluies mesurées à l'amont par des hauteurs pourrait être appliqué de manière inverse sur des bassins plus petits : par exemple sur le bassin d'Anduze avec les stations amont de Mialet, de Saumane et de St Jean du Gard. Cette dernière, installée en 2003, n'a pour l'instant pas permis d'enregistrer suffisamment d'événements pour envisager un apprentissage statistique.

3.5 RELATION REFLECTIVITE - HAUTEUR

Une première approche de la relation Réflectivité / Hauteur a été réalisée dans le cadre de ce projet sur le bassin versant du Gardon d'Anduze. Cette section fait état des premiers résultats obtenus.

3.5.1 Données et stratégie de modélisation

La base de données utilisée comprend 11 événements détaillés dans le Tableau 2. Le bassin versant du Gardon d'Anduze est couvert par 545 pixels (de 1 km²) de l'image RADAR. Il convenait donc de proposer un découpage de l'image RADAR permettant de limiter le nombre de variables pour le modèle. De manière à pouvoir comparer les résultats obtenus à partir des données de pluie des 6 pluviographes du bassin versant (section 3.3) avec ceux obtenus à partir de l'image RADAR, un découpage en 6 zones hydrologiquement pertinentes (en respectant une logique de sous-bassin versant amont et aval) a été réalisé (Figure 10). La définition d'une discrétisation optimale de l'image RADAR a été recherchée dans le cadre du projet et sera présentée dans la section 5.2 de ce présent rapport.

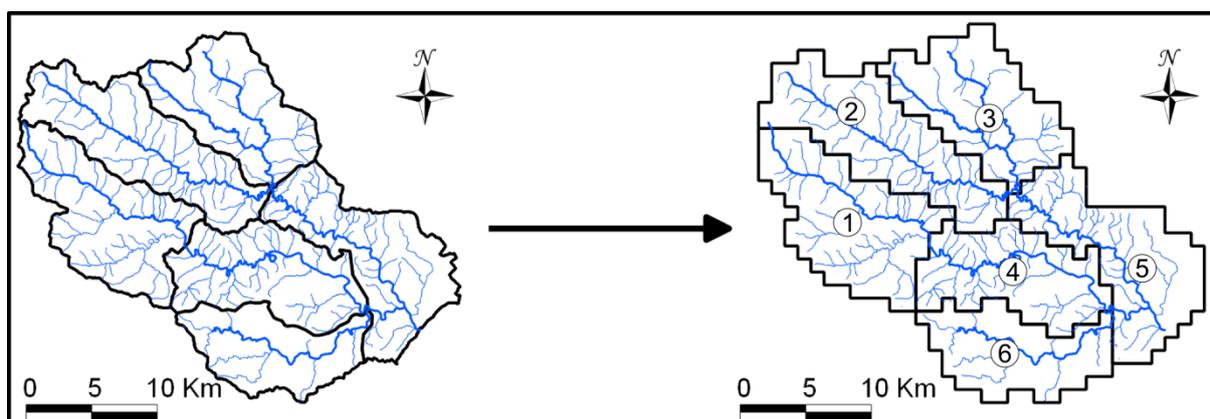


Figure 10 : Découpage du bassin versant du Gardon d'Anduze (Visserot, 2012)

Contrairement à la donnée de pluie, la mesure de réflectivité ne se prête pas bien au ré-échantillonnage spatial autant que temporel du fait de la très forte non-linéarité qui lie la réflectivité à la pluie. Néanmoins, dans l'objectif de pouvoir comparer avec l'approche développée précédemment (section 3.3. relation pluie-hauteur), le choix a été fait dans ce premier travail de ne pas modifier les périodes de travail du modèle et donc de ré-échantillonner la réflectivité sur 30 minutes. Le schéma suivant synthétise les traitements et la modélisation effectués dans cette opération.



Figure 11 : Relation Réflectivité (Z) - Hauteur (H), RN est le réseau de neurones

Le modèle utilisé est un perceptron multicouches sans liens linéaires ni calcul de la médiane, les conclusions d'Audrey Bornancin-Plantier étant intervenues postérieurement aux travaux présentés ici. La meilleure configuration du réseau est

recherchée en termes de fenêtres temporelles (de $k-2$ à $k-4$) et de nombre de neurones cachés (de 2 à 4) ; 50 initialisations différentes sont effectuées pour chaque expérience de validation croisée partielle. La configuration dont le score de validation croisée partielle fondé sur le critère de persistance le plus proche de 1 est sélectionnée et évaluée en test.

Les ensembles d'apprentissage, d'arrêt précoce et de test sont définis comme suit :

- Ensemble d'apprentissage : 12, 120, 15, 19, 20a, 20b, 21, 22, 24
- Ensemble d'arrêt précoce : 13
- Ensemble de test : 23

La prévision à 2h et à 3h est ici recherchée.

3.5.2 Résultats et discussion

Le tableau suivant présente les résultats obtenus en test sur l'événement 23 qui est le second événement le plus intense de la base (niveau de vigilance orange).

Horizon de prévision	Nb. neurones cachés	Nash	Persistance	Retard	PPD	Commentaires
2 h	2	0,89	0,55	45 min	74 %	Léger retard mais sous estimation du pic
3 h	2	0,79	0,51	1h30	77%	Retard + sous estimation

Tableau 8 : Résultats relation Réflectivité – Hauteur (Visserot, 2012)

Réalisées à partir de la réflectivité, les prévisions à 2h et 3h de la hauteur d'eau à Anduze présentent des critères de Nash satisfaisants et des critères de persistance qui sont supérieurs à 0,5. Il faut constater pour les deux horizons de prévision un léger retard du pic de crue qui s'accroît avec l'augmentation de l'horizon de prévision. La figure suivante propose une comparaison des prévisions à 2h de la hauteur d'eau à partir des données de réflectivité (Figure 12-a) et des données des 6 pluviographes du bassin versant du Gardon d'Anduze (Figure 12-b). Le modèle utilisé est identique dans les deux cas.

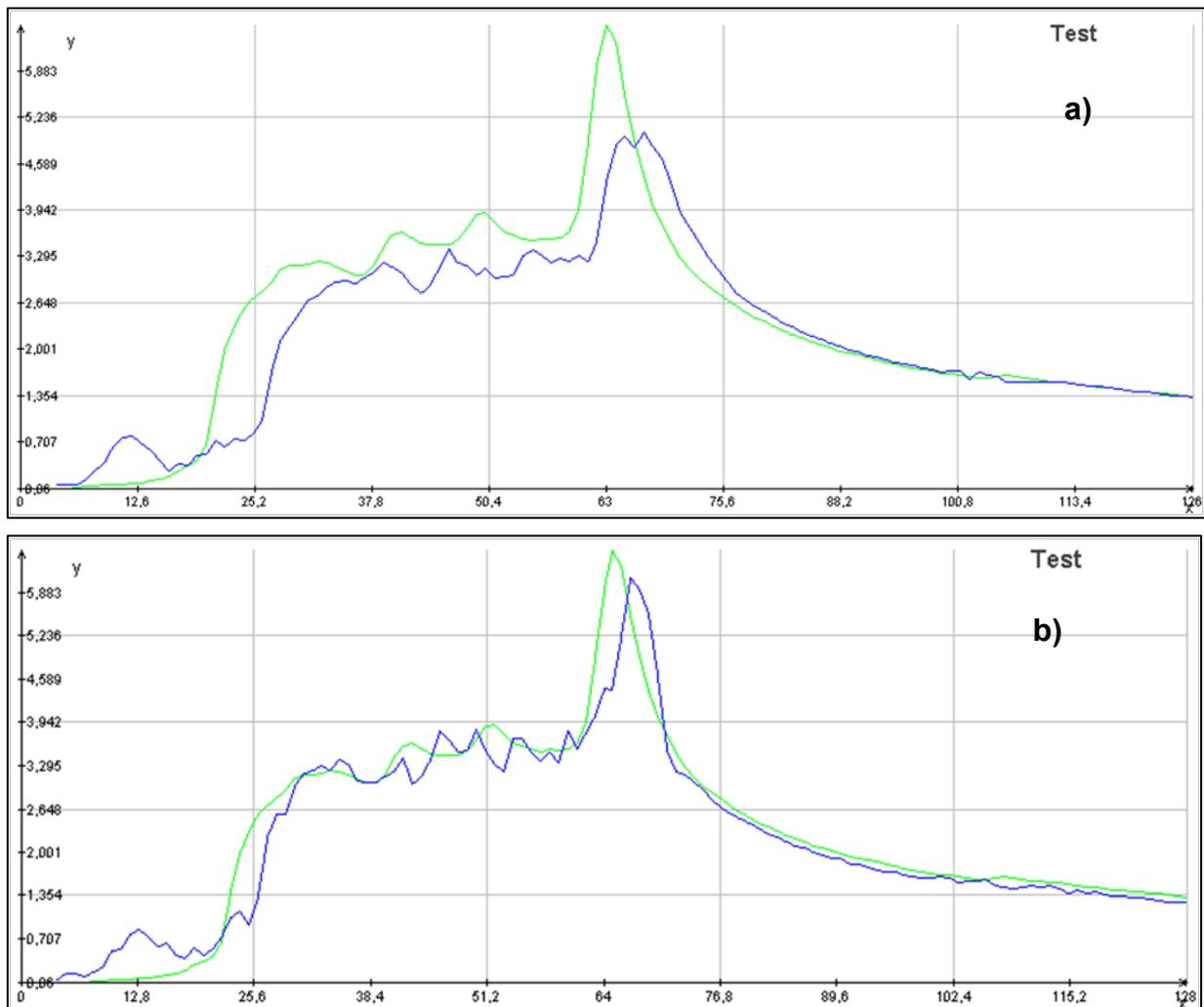


Figure 12 : Résultat de la prévision à 2h00 : a) Réflectivité – Hauteur et b) Pluie – Hauteur. Avec les hauteurs d'eau en m sur l'axe y, le temps sur l'axe des x, la hauteur observée en vert et la hauteur prévue en bleu

La prévision effectuée à 2h à partir des données des pluviographes est nettement meilleure que celle réalisée à partir des données de réflectivité RADAR. Cependant, il convient de noter que la cinétique du limnigramme prévue avec les données de réflectivité est globalement bien respectée et que seule une sous estimation de la hauteur d'eau est avérée. Le pic de crue est prévu dans les deux cas avec un léger retard. Les mêmes remarques peuvent être faites sur la prévision à 3 heures.

3.5.3 Conclusion et perspectives

Cette première approche de la relation réflectivité RADAR - hauteur d'eau est tout à fait prometteuse. En effet, dans ce cas le modèle a été placé dans des conditions défavorables où il devait :

- Compenser la non linéarité du passage de la réflectivité 5 minutes à la réflectivité 30 minutes
- Compenser la non linéarité du ré-échantillonnage spatial en 6 zones.
- Evaluer la relation réflectivité - pluie
- Evaluer la relation pluie - réflectivité

Des recherches vont être poursuivies sur ce thème par le LGEI (ARMINES) dans un futur proche en s'attachant tout d'abord à modifier le modèle pour permettre d'intégrer au réseau directement la réflectivité 5 minutes ce qui devrait permettre une meilleure prévision en éliminant l'étape du passage de la réflectivité 5 minutes (mesurée par le RADAR) à une réflectivité de 30 minutes. Cette étape permettra également de s'approcher des conditions opérationnelles où la donnée de réflectivité RADAR est disponible au pas de temps 5 minutes pour les Services de Prévision des Crues (SPC).

4. RELATION PLUIE-DEBIT SUR LES PETITS BASSINS VERSANTS

En adoptant la méthodologie de travail définie en 3.3 de ce rapport il est nécessaire de disposer d'une base de données importante pour réaliser l'apprentissage. Or cette base de données n'a pas pu être constituée durant la durée du projet du fait de difficultés d'instrumentations et du manque d'épisodes pluvieux très intenses sur le bassin ciblé à l'origine du projet (petits bassins de Valescure). Néanmoins le point concernant la généralisation des méthodes fondées sur l'apprentissage statistique à des petits bassins versants a pu être abordé dans le travail de thèse de Guillaume Artigue et est présenté dans le rapport relatif à la tâche 4 du projet (Artigue, 2013). En synthèse, le modèle développé par Guillaume Artigue est un modèle récurrent qui réalise l'apprentissage sur un bassin versant représentatif des types de bassins visés et grâce à une transformation qui ne dépend que des caractéristiques physiques des deux bassins considérés peut être transposé à un bassin versant cible dépourvu d'instrumentation et appelé en hydrologie « bassin versant non jaugé ».

5. FIABILITE DE LA CHAINE DE PREVISION

5.1 INTEGRATION DE LA PLUIE RADAR

L'objectif est ici de réaliser une comparaison de prévisions de hauteur d'eau à Anduze obtenues à partir des données des pluviographes et des données de pluie issues des images RADARs.

5.1.1 Données et stratégie de modélisation

La base de données comprend les 11 événements présentés dans le tableau 2. Les données de pluie RADAR sont ré-échantillonnées de 5 minutes en 30 minutes et discrétisées spatialement en 6 zones suivant le schéma présenté sur la Figure 10 (section 3.5).

Le type de modèle utilisé est identique à ceux utilisés dans la section 3.4. Il s'agit d'un perceptron multicouche à une couche de neurones cachés sans liens linéaires ni calcul d'ensemble. Ce modèle non bouclé utilise les hauteurs mesurées précédentes comme variables d'entrées. Deux variables d'entrée seront donc testées : la pluie mesurée par les 6 pluviographes (modèle « pluvio ») et la lame d'eau RADAR ré-échantillonnée et spatialisée (modèle « RADAR »).

Les sous-ensembles de la base de données ont été définis de la façon suivante :

- L'évènement 19, qui est le plus intense, sera l'évènement de test. Comme cet évènement est unique et exceptionnel et que son hydrogramme a été reconstitué

artificiellement après la crue du fait de l'endommagement de l'instrumentation, le modèle sera également testé sur l'évènement 13 qui est plus commun. Quand l'évènement 13 est utilisé pour le test, l'évènement 19 passe dans le sous ensemble d'apprentissage et vice-versa ;

- Le sous-ensemble d'arrêt est constitué de l'évènement 200. Ce choix s'est fait par validation croisée : c'est cet évènement qui présentait l'erreur quadratique moyenne la plus faible. Cela veut dire que c'est lui qui représente le mieux l'ensemble d'apprentissage, il y a donc une cohérence à utiliser cet ensemble pour arrêter l'apprentissage.
- Les 9 autres évènements font partie du sous-ensemble d'apprentissage.

La sélection de l'architecture s'est faite au moyen de la validation croisée partielle au moyen du score fondé sur le critère de persistance puisqu'il renseigne la qualité de la synchronisation des valeurs prédites et des valeurs mesurées. La meilleure architecture est choisie de la façon suivante :

Essais de différentes architectures sur la base de 50 initialisations des paramètres. La fin de chaque phase d'apprentissage est déterminée par un nombre maximal de présentations soit 50 itérations.

L'architecture retenue est celle dont la moyenne des dix meilleurs scores de validation croisée est la plus grande.

Pour mémoire, les hyperparamètres déterminés pour chaque horizon de prévision (ici 30 min, 1h, 2h, 3h, 4h et 5h) sont : la profondeur des fenêtres temporelles de pluie et de débit, et le nombre de neurones cachés (N_c).

Les tableaux suivants synthétisent les résultats de la validation croisée partielle qui permettent de définir l'architecture des deux modèles (Tableau 9 et Tableau 10).

Horizon de prévision (h_p)	0,5 h	1 h	2 h	3 h	4 h	5 h
Fenêtre temporelle pluies f_t (heure)	0,5	5	4,5	4,5	5	4,5
Fenêtre temporelle hauteurs précédents (heure)	1,5	4,5	4,5	1	0,5	1
Complexité N_c	2	2	2	2	1	1

Tableau 9 : Architecture du modèle "pluviomètres" (Visserot, 2012)

Horizon de prévision (h_p)	0,5 h	1 h	2 h	3 h	4 h	5 h
Fenêtre temporelle pluies f_t (heure)	0,5	0,5	0,5	0,5	0,5	4,5
Fenêtre temporelle hauteurs précédents (heure)	1	1	1	2	2,5	1
Complexité N_c	1	3	4	3	4	1

Tableau 10 : Architecture du modèle "RADAR" (Visserot, 2012)

Après avoir défini l'architecture des deux modèles, on constate qu'il y a peu de points communs entre ces deux réseaux de neurones en fonction de la source des données. Compte-tenu des conclusions du travail présenté en partie 3.3, ceci peut être expliqué par la grande dépendance des modèles à la sélection de l'initialisation des modèles. Cette étude devrait donc être ré-examinée avec d'une part des liens linéaires supplémentaires, et d'autre part au moyen du calcul d'un modèle d'ensemble.

Néanmoins, il apparaît que le modèle alimenté par les pluies RADAR ne prend quasiment pas en compte les précipitations antérieures, comme si ces dernières étaient peut significatives. Si ce résultat était confirmé par une étude comportant le calcul d'ensemble cela signifierait qu'un travail de fond doit être entrepris pour estimer les précipitations au moyen des RADARs.

Ces deux modèles ont été testés sur les événements 13 et 19, ce qui a permis de comparer leurs performances en prévision. La meilleure initialisation a été sélectionnée par la validation croisée partielle sur la base de 100 initialisations. Les limnigrammes ainsi obtenus et les critères de performances associés sont présentés dans la partie suivante.

5.1.2 Résultats et discussion

Les résultats présentés ici reprennent les prévisions à 2 heures et à 3 heures pour les événements 13 (septembre 2000, 1 000 m³/s ; Figure 13) et 19 (septembre 2002, 3 500 m³/s ; Figure 14). Les résultats obtenus pour l'ensemble des horizons de prévision sont présentés en annexe.

Afin de juger de la qualité des limnigrammes en dehors de l'analyse visuelle, quatre critères ont été utilisés (présentés en 3.1, leurs valeurs étant reportées sur les différentes figures) :

- Le critère de Nash.
- Le critère de persistance.
- Le PPD (Percentage Peak Discharge).
- Le SPPD (Synchronous Percentage Peak Discharge).

Sur l'événement 13, les résultats obtenus à partir des pluies « RADAR » sont moins bons que les résultats obtenus avec les pluies « pluvios ». En effet, même si la forme du limnigramme est tout à fait satisfaisante, le modèle spatialisé (RADAR) accuse un retard systématique durant la phase de montée de la crue. A cause de ce retard le modèle est peu exploitable pour un prévisionniste de crue. Ce mauvais *timing* provient du fait que le modèle spatialisé n'utilise pas d'historiques des précipitations. Il ne peut donc pas anticiper les hauteurs d'eau de façon pleinement satisfaisante. Cet état de fait est particulièrement vrai pour les premiers horizons de prévision (30 minutes et 1h, reportés en annexe 1).

Enfin, il est intéressant de voir que, quelle que soit la source d'informations pluviométriques, les limnigrammes prévus perdent de leur amplitude au niveau du pic pour les l'horizon de prévision de deux et trois heures. Pour les deux derniers horizons de prévisions seul le modèle « pluvio » parvient à proposer un limnigramme qui se rapproche de celui observé.

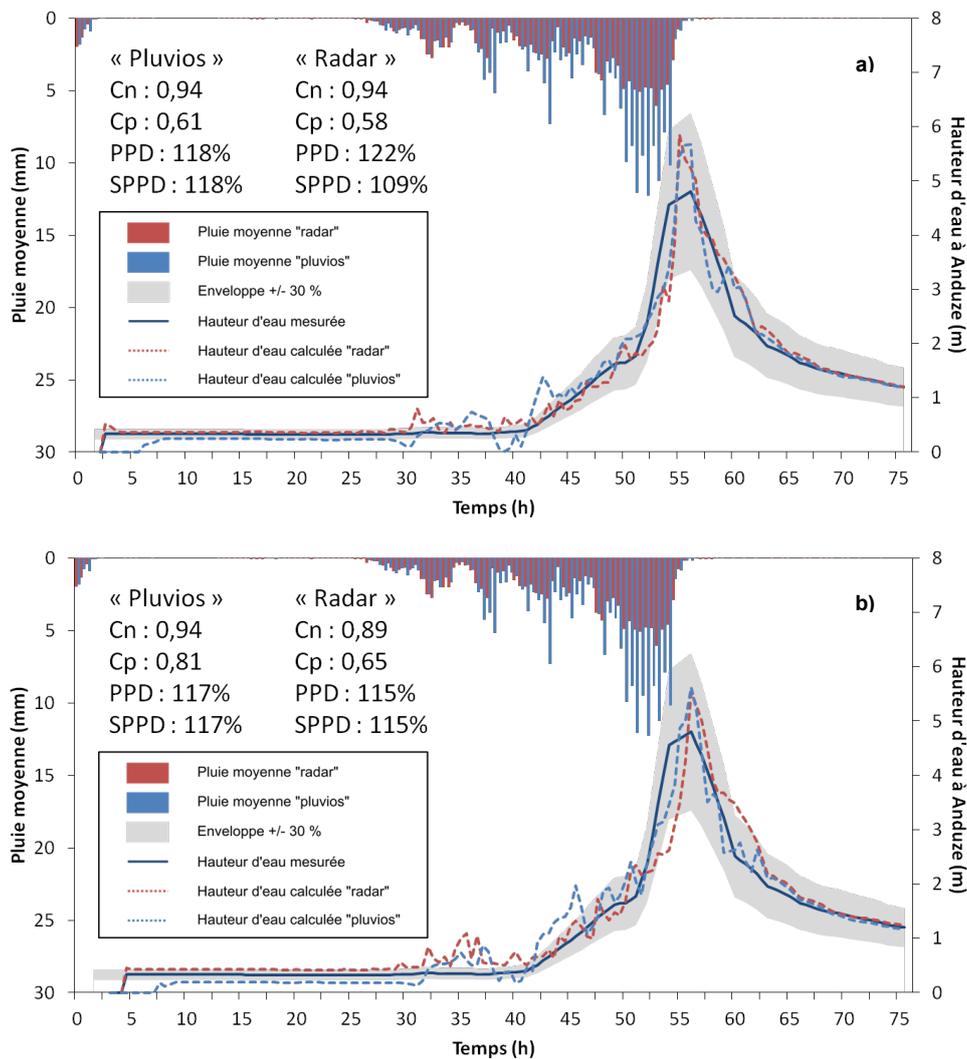


Figure 13 : Résultat de la prévision à 2h a) et 3h b) pour l'événement 13 (Visserot, 2012)

Si l'on considère maintenant l'événement 19, il est intéressant de voir sur la figure 11 que sur cet événement intense, le modèle utilisant les pluies « RADAR » délivre de meilleures prévisions que le modèle « pluvious » jusqu'à l'horizon de prévision de trois heures. Contrairement au résultat obtenu sur l'événement 13. Pour l'événement 19 les prévisions issues du modèle spatialisé accusent un retard moins important que le modèle « pluvious ». Cette remarque est particulièrement vraie pour la phase de montée qui se situe après le premier « petit » pic : le modèle réussit à « rattraper » son retard, notamment pour un horizon de prévision de trois heures. On observe également que le modèle « pluvious » n'arrive pas à prévoir le pic à partir de l'horizon de prévision une heure alors que le modèle « RADAR » arrive à retranscrire la forme du limnigramme jusqu'à l'horizon de prévision trois heures. Ceci évoque les résultats présentés en section 3.3 développés postérieurement à l'étude présente.

Les prévisions sont ici difficilement utilisables à partir de l'horizon de prévision 2 heures, car le modèle spatialisé accuse un retard important dans la phase de montée de crue située avant le petit pic, puis dans la seconde phase de montée le modèle anticipe mieux les hauteurs futures. Malgré ces retards, il est intéressant de voir que la prévision du modèle « RADAR » se situe globalement dans une tranche de

précision de +/- 30 % autour des valeurs observées, intervalle généralement admis comme incertitude sur les mesures de débit des crues rapides.

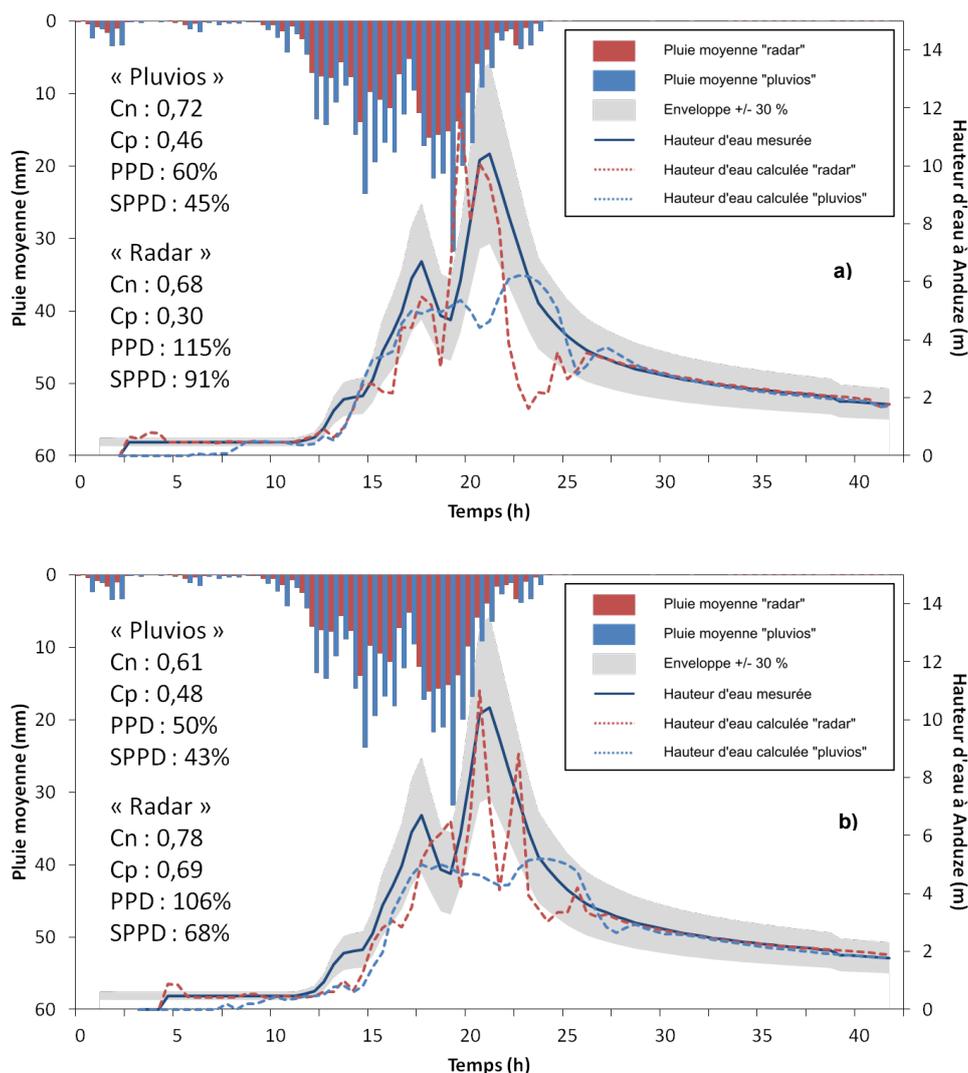


Figure 14 : Prédiction à 2h a) et 3h b) pour l'événement 19 (Visserot, 2012)

5.1.3 Conclusion et perspectives

Les données spatialisées de pluie n'ont pas permis d'améliorer significativement les prévisions par rapport aux modèles utilisant les informations pluviométriques. Cet état de fait peut trouver plusieurs éléments d'explication.

En premier lieu, une critique de la base de données utilisée et de sa qualité est nécessaire. En effet, le nombre peu important d'évènements (11 au total) est un obstacle dans le cadre de l'apprentissage statistique.

Il est toutefois possible de tirer quelques enseignements de la comparaison entre le modèle utilisant des données ponctuelles (pluviomètre) et celui utilisant des données spatialisées (RADAR).

Bien que la base de données « pluviomètres » aient été tronquée pour permettre la comparaison sur un ensemble d'évènements identiques, elle fournit de meilleurs résultats que le modèle spatialisé sur un évènement important, mais simple. Effectivement, le fait d'utiliser un historique des pluies conséquent permet au modèle

issu des entrées pluviométriques de limiter le retard observé dans la prévision du limnigramme (Tableau 9). Le modèle spatialisé s'appuyant principalement sur les hauteurs observées, il lui est impossible d'anticiper correctement la montée de la crue (Tableau 10).

Pour terminer cette conclusion sur les résultats obtenus, on mettra en avant la qualité des prévisions du modèle spatialisé sur l'évènement intense de septembre 2002, le plus important de la base de données. Bien qu'imparfaites, les prévisions obtenues peuvent être jugées comme correctes puisqu'elles se situent dans une enveloppe de +/- 30 % par rapport aux valeurs observées jusqu'à un horizon de prévision de trois heures. Ces résultats sont encourageants, particulièrement si l'on rappelle qu'ils sont obtenus en l'absence de prévision de pluies.

Pour conclure, les résultats de ces travaux permettent de dégager de nouvelles pistes pour l'utilisation des données spatialisées dans le cadre la prévision des crues éclair par réseaux de neurones. Ces travaux pourraient éventuellement s'axer sur la recherche d'autres méthodes de spatialisation : un nouveau découpage pourrait par exemple s'effectuer sur plus de 6 zones ou bien en en fonction d'isochrones, comme le font Pereira Filho & Dos Santos (2006) dans leurs travaux.

5.2 STRATEGIE D'INTEGRATION DE L'IMAGE RADAR

L'objectif ici est de tester différents découpages de l'image RADAR pour mieux prendre en compte les différences spatiales de précipitations qui peuvent être très importantes pour les événements très intenses.

5.2.1 Données et stratégie de modélisation

Le type de modèle utilisé est identique à ceux utilisés précédemment. Pour mémoire, il s'agit d'un perceptron multicouche à une couche de neurones cachés. Ce modèle non bouclé utilise les hauteurs mesurées précédentes comme variables d'entrée avec les données de pluie.

La base de données utilisée comprend les 11 événements présentés dans le tableau 2. La stratégie de modélisation est identique à celle développée dans la section 3.4 (Relation réflectivité - hauteur), avec notamment les ensembles d'apprentissage, d'arrêt précoce et de test qui sont définis comme suit :

- Ensemble d'apprentissage : 12, 120, 15, 19, 20a, 20b, 21, 22, 24
- Ensemble d'arrêt précoce : 13
- Ensemble de test : 23

Dans ces travaux, un découpage suivant une logique hydrologique a été recherché et ce découpage a permis de dégager et de tester 4, 6, 8, 10 et 12 zones. Sur ces zones la pluie moyenne et la réflectivité moyenne ont été calculées et ont permis d'alimenter le réseau. Les prévisions à 2 heures et à 3 heures ont été recherchées. La Figure 15 présente les différents découpages testés.

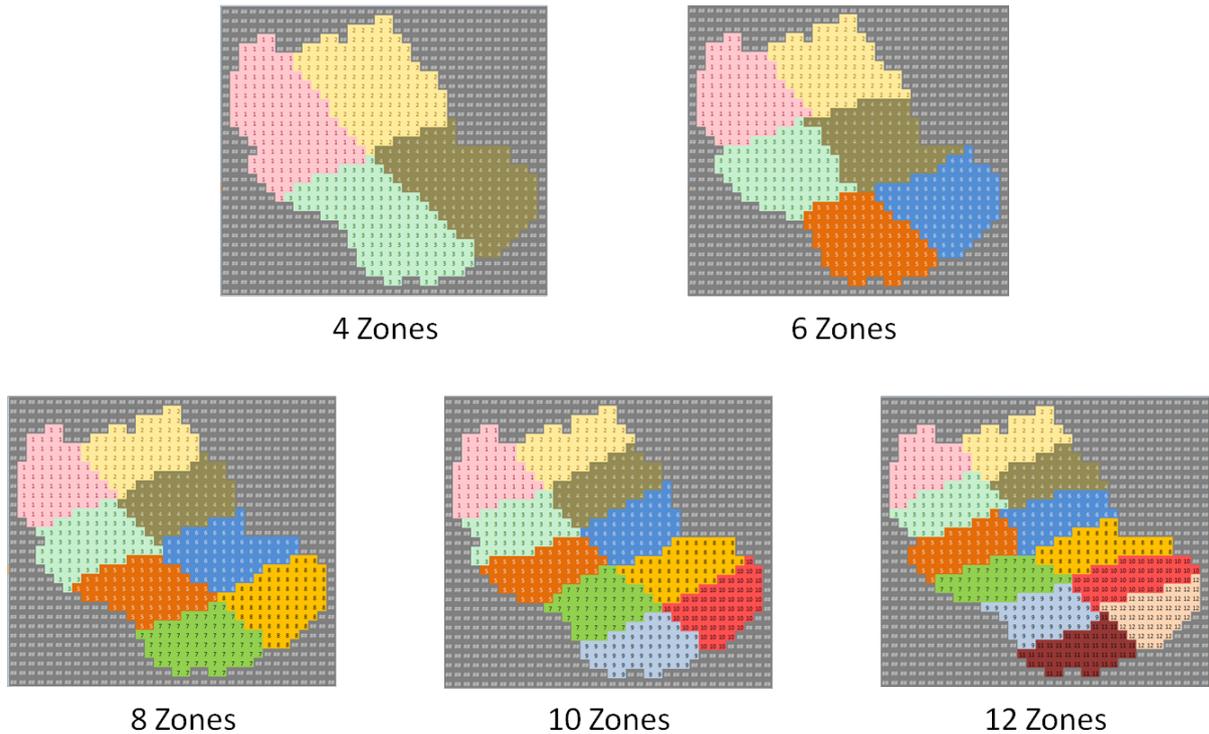


Figure 15 : Proposition de découpage de l'image RADAR, de 4 à 12 zones (Visserot, 2012)

5.2.2 Résultats et discussion

Le tableau ci-dessous présente les résultats obtenus pour différents découpages de l'image RADAR. La donnée d'entrée est ici la pluie RADAR.

4 zones						
H_p	N_C	Nash	Persistance	Retard	PPD	
2h	2	0,85	0,41	45min	90%	
3h	3	0,77	0,45	1h30	98%	
6 zones						
H_p	N_C	Nash	Persistance	Retard	PPD	
2h	2	0,85	0,42	1h30	89%	
3h	2	0,74	0,38	2h30	80%	
10 zones						
H_p	N_C	Nash	Persistance	Retard	PPD	
2h	2	0,87	0,48	1h	111%	
3h	2	0,71	0,31	2h15	87%	
12 zones						
H_p	N_C	Nash	Persistance	Retard	PPD	
2h	2	0,8	0,21	45 min	112%	
3h	2	0,7	0,4	1h30	107%	

Tableau 11 : Résultats de la relation pluie RADAR - hauteur d'eau à Anduze pour différentes zones (Visserot, 2012)

Pour les différentes configurations les résultats restent globalement assez proches en termes de critères de Nash et de persistance. La prévision à 2h semble être

optimale pour 10 zones alors que celle à 3 heures est optimale pour 4 zones et 12 zones.

Ces constatations vont dans le sens des conclusions de la section précédente qui montre que le réseau qui prend comme entrée des données issues de l'imagerie RADAR est finalement très conditionné par la hauteur d'eau aux pas de temps précédents et donne donc moins de poids à l'entrée « pluie ».

5.2.3 Conclusion

Ces résultats doivent être approfondis mais prouve que la prochaine étape sur ce thème de l'intégration opérationnelle de l'imagerie RADAR devra passer par une modification de la structure du réseau (pas de temps et/ou structure) pour mieux prendre en compte l'historique des données de pluie, peu utilisé par le réseau dans la configuration actuelle ou *a minima* par un nouveau type de discrétisation soit matriciel (diminution de la résolution de l'image) ou de type isochrone (Pereira Filho & Dos Santos, 2006).

6. CONCLUSION

Ce rapport a présenté les travaux effectués dans le cadre du projet FLASH financé par l'ANR. Pour faciliter la lecture en relation avec la proposition initiale, ce rapport a repris le découpage en tâches et sous-tâches qui y étaient proposé. Ainsi ont pu être présentées les différentes études relatives à la modélisation systémique de la chaîne hydrométéorologique : (i) le pré-traitement des données, (ii) l'identification des relations de la chaîne hydrométéorologique, (iii) la fiabilité de la chaîne hydrométéorologique, (iv) le changement d'échelle pour les très petits bassins étant quand à lui présenté dans le rapport de la tâche T4.

Ce travail a été réalisé principalement grâce à la thèse de Audrey Bornancin-Plantier, financée par le projet FLASH, et le stage Master 2 de Yann Visserot. Nous incitons le lecteur intéressé à se reporter à leurs mémoires respectifs.

La chaîne hydrométéorologique peut être décomposée en plusieurs relations dont la principale est la relation pluie-hauteur. L'étude de la modélisation de cette relation par un réseau de neurones pour la prévision des crues rapides et dévastatrices du bassin versant d'Anduze a été envisagée au moyen de plusieurs sources d'estimation des précipitations : (i) les pluies mesurées par les six pluviomètres du bassin, (ii) les précipitations spatialisées estimées par les RADARs météorologiques, (iii) la réflectivité brute mesurée par le RADAR.

Le modèle neuronal choisi est un perceptron multicouche non-récurrent, connu comme le meilleur prédicteur en hydrologie du fait des incertitudes et bruits considérables relatifs aux variables hydrométéorologiques. Sur ce modèle, alimenté par les mesures de précipitation et de hauteur d'eau observées avant le temps courant, il a pu être montré que le principal facteur déterminant la qualité de la prévision était le choix de l'initialisation préliminaire des paramètres du réseau. Ce facteur est plus déterminant que la sélection des variables (le dimensionnement de la fenêtre temporelle d'historique des précipitations ou celui des hauteurs d'eau précédentes) ou que le choix de la complexité optimale (nombre de neurones cachés). Ce comportement atypique, absent de la bibliographie sur les crues éclair, a induit un manque de robustesse dans la prévision, qui a rendu les comparaisons entre les différents types d'estimation des précipitations difficiles à réaliser.

Néanmoins une solution a été proposée pour pallier ce défaut majeur : une prévision de type ensembliste est réalisée. Ainsi il est proposé d'utiliser 20 modèles différents en parallèle, la prévision obtenue étant celle calculée par la médiane des différents modèles.

Une autre avancée originale pour les applications en hydrologie a été de proposer de coupler deux modèles : l'un linéaire et l'autre non-linéaire afin d'effectuer la prévision des crues les plus intenses. Le sous-réseau linéaire prend en charge la réponse aux très grandes impulsions de pluie, tandis que le sous-réseau non linéaire prend en compte les réponses au début et à la fin de l'événement, réponses essentiellement non-linéaires.

Pour finaliser la comparaison des performances associées aux trois méthodes d'estimation les précipitations tombant sur le bassin, ces deux avancées majeures devront être appliquées également aux modèles recevant en entrées les informations issues des RADARs. Ce travail va être réalisé dans les mois à venir.

Enfin, compte-tenu de la qualité des prévisions délivrées par le modèle ensembliste intégrant les liens linéaires et non linéaire, ce modèle a été développé et livré à l'utilisateur final de ces prévisions et partenaire du projet : le SCHAPI. Ce modèle EMEAU est actuellement en test dans la chaîne de prévision temps réel du SCHAPI et est utilisable sous forme d'un démonstrateur accessible sur le site internet du projet.

7. REFERENCES

- (Artigue, 2012) G. Artigue, *Prévision des crues éclair par réseaux de neurones : généralisation aux bassins non jaugés*, thèse de doctorat, Université Montpellier II, 2012.
- (Bornancin-Plantier, 2011) A. Bornancin-Plantier, A. Johannet, P. Roussel-Ragot, G. Dreyfus, *Flash Flood Forecasting using Neural Networks without Rainfall Forecasts: Model Selection and Generalization capability*, EGU 2011-1794, présentation orale, 2011.
- (Bornancin-Plantier, 2013) thèse de doctorat, *Conception de modèles de prévision des crues éclair par apprentissage artificiel*, Université Pierre et Marie Curie, Ecole Doctorale EDITE, 2013.
- (Bornancin-Plantier, 2013b) Bornancin-Plantier A., Johannet A., Roussel-Ragot P., Dreyfus G., *Improving Flash Flood Predictions by Neural Networks using Mixed Linear/Non Linear architecture*, soumise à Neurocomputing.
- (Boukharouba, 2013) FLASH-Rapport T₄.
- (Johannet, 2013) FLASH-Rapport T₆.
- (Marchandise, 2013) FLASH-Rapport T₁.
- (Nash, 1970) J.E. Nash, J.V. Sutcliffe, *River Flow Forecasting through Conceptual Models. Part I – A Discussion of Principles*, Journal of Hydrology, vol 10, p. 282-290, 1970.
- (Nerrand, 1993) O. Nerrand, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, S. Marcos, *Neural networks and non –linear adaptative filtering: unifying concepts and new algorithms*, Neural computation, vol5, p. 165-197, 1993.
- (Pereira Filho, 2006) Pereira Filho A.J., Dos Santos C.C. (2006) – *Modeling a densely urbanized watershed with an artificial neural network, weather RADAR and telemetric data* – Journal of Hydrology, n°317, pp. 31-48.
- (Stone, 1974) M. Stone, *Cross-validators choice and assessment of statistical predictions*, Journal of the Royal Statistical Society, B36, p. 111-147, 1974.
- (Toukourou, 2009.a) M.S. Toukourou, *“Application de l’apprentissage artificiel à la prévision des crues éclair”*, thèse de l’Ecole Nationale Supérieure des Mines de Paris, Spécialité Informatique Temps Réel, Robotique et Automatique, 175p, 2009.
- (Vigicrues) <http://www.vigicrues.gouv.fr/>
- (Visserot, 2012) Y. Visserot, *“Intégration de l’information hydrométéorologique spatialisée pour la modélisation et la prévision des crues éclair à l’aide des réseaux de neurones”*, Mémoire de Master Pro Gestion des Catastrophes et des Risques Naturels, Département de Géographie UFR III de l’université Paul Valérie, Montpellier III, 2012.

LISTE DES FIGURES

<i>Figure 1: La chaîne hydrométéorologique</i>	3
<i>Figure 2 : Schéma du perceptron multicouche utilisé (Bornancin-Plantier, 2013)</i>	10
<i>Figure 3 : Événement 19, les 20 hauteurs prévues par les 20 modèles sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013)</i>	12
<i>Figure 4 : Valeurs des critères C_p, C_N, RMSE et PWRMSE pour le test de l'événement 26 sur les 20 meilleurs modèles, au sens de la validation croisée partielle, pour chaque horizon de prévision. L'axe des abscisses correspond au rang du modèle par C_p décroissants (Bornancin-Plantier, 2013)</i>	14
<i>Figure 5 : Événement 19, hauteur médiane calculée sur les hauteurs prévues par les 20 modèles sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013)</i>	15
<i>Figure 6: Modèle modifié avec introduction d'une partie linéaire (Bornancin-Plantier, 2013)</i>	16
<i>Figure 7 : Événement 19, hauteurs médianes sur les hauteurs prévues par les 20 modèles avec partie linéaire et les 20 modèles à perceptron multicouche sélectionnés pour chaque horizon de prévision (Bornancin-Plantier, 2013)</i>	17
<i>Figure 8 : Bassin versant du Gardon (conception G. Artigue)</i>	18
<i>Figure 9 : Modèle avec ajout d'une partie linéaire au perceptron multicouche avec h_p l'horizon de prévision, f_T la fenêtre temporelle des pluies, f_1 à f_4 les fenêtres temporelles pour les hauteurs correspondantes et R_1 à R_5 les retards appliqués aux pluies et hauteurs correspondantes (Bornancin-Plantier, 2013)</i>	19
<i>Figure 10 : Découpage du bassin versant du Gardon d'Anduze (Visserot, 2012)</i>	21
<i>Figure 11 : Relation Réflectivité (Z) - Hauteur (H), RN est le réseau de neurones</i>	21
<i>Figure 12 : Résultat de la prévision à 2h00 : a) Réflectivité – Hauteur et b) Pluie – Hauteur. Avec les hauteurs d'eau en m sur l'axe y, le temps sur l'axe des x, la hauteur observée en vert et la hauteur prévue en bleu</i>	23
<i>Figure 13 : Résultat de la prévision à 2h a) et 3h b) pour l'événement 13 (Visserot, 2012)</i>	27
<i>Figure 14 : Prévision à 2h a) et 3h b) pour l'événement 19 (Visserot, 2012)</i>	28
<i>Figure 15 : Proposition de découpage de l'image RADAR, de 4 à 12 zones (Visserot, 2012)</i>	30

LISTE DES TABLEAUX

Tableau 1 : Les événements de la base de données dont les pluies sont fournies par les pluviomètres (Bornancin-Plantier, 2013)	4
Tableau 2 : Les événements de la base de données dont les pluies sont fournies par les RADARs (Visserot, 2012)	5
Tableau 3 : Base de données comportant le numéro de l'événement, la date, la durée, le cumul de pluie à Remoulins, la hauteur maximale à Remoulins et le niveau de vigilance (Bornancin-Plantier, 2013)	6
Tableau 4 : Caractérisation des événements (Bornancin-Plantier, 2013)	7
Tableau 5 : Valeurs du C_p , C_N , PPD, SPPD et Retard H_{pic} suivant l'horizon de prévision pour la hauteur prévue médiane des sorties sur l'événement 19. Comparaison entre les modèles à perceptron multicouche uniquement (PM) et ceux avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)	16
Tableau 6 : Pour chaque horizon de prévision, critères de performance calculés sur les prévisions médianes de l'événement 26 obtenues par les modèles à perceptron multicouche (PM) et avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)	20
Tableau 7 : Pour chaque horizon de prévision, critères de performance calculés sur les prévisions médianes de l'événement 27 obtenues par les modèles à perceptron multicouche (PM) et avec une partie linéaire ajoutée (PM+PL) (Bornancin-Plantier, 2013)	20
Tableau 8 : Résultats relation Réflectivité – Hauteur (Visserot, 2012)	22
Tableau 9 : Architecture du modèle "pluviomètres" (Visserot, 2012)	25
Tableau 10 : Architecture du modèle "RADAR" (Visserot, 2012)	25
Tableau 11 : Résultats de la relation pluie RADAR - hauteur d'eau à Anduze pour différentes zones (Visserot, 2012)	30

Annexe

ANNEXE 1 : PLUVIOGRAPHES VS PLUIE RADAR

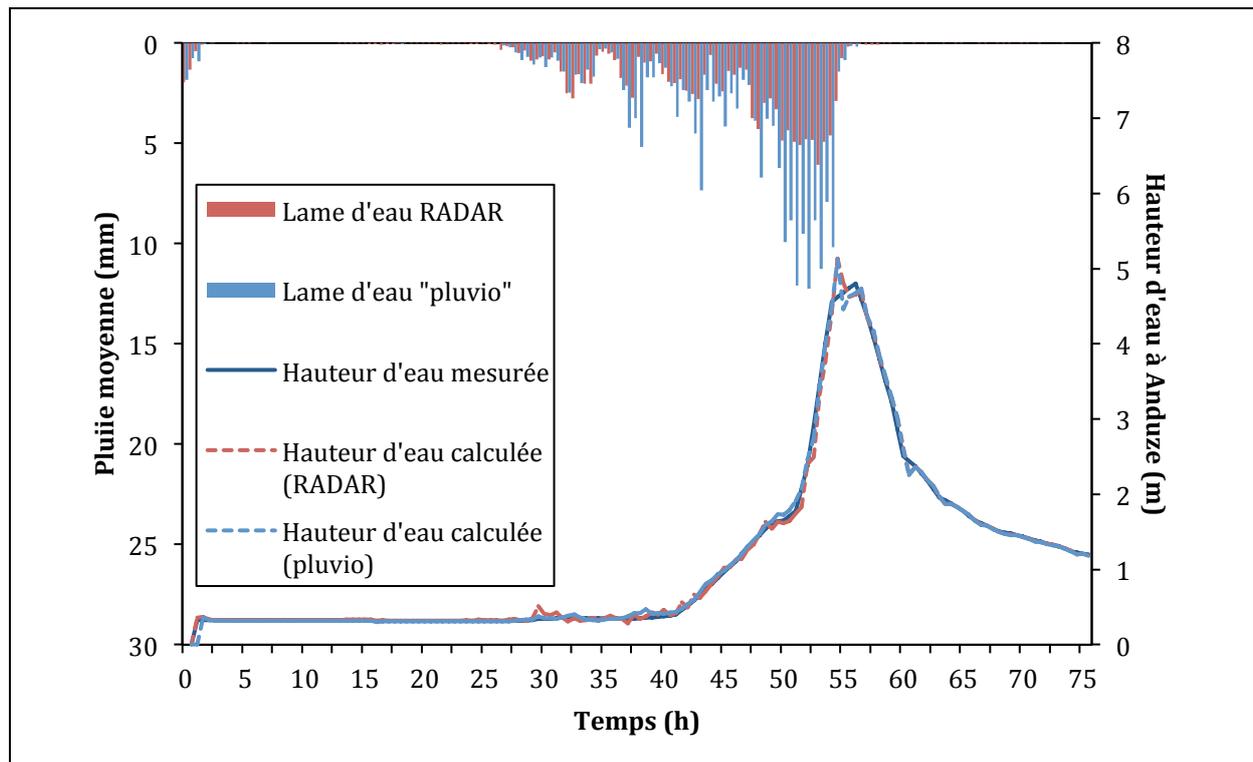


Figure 1 – Prévisions pour $h_p = 0,5$ h (ev. 13)

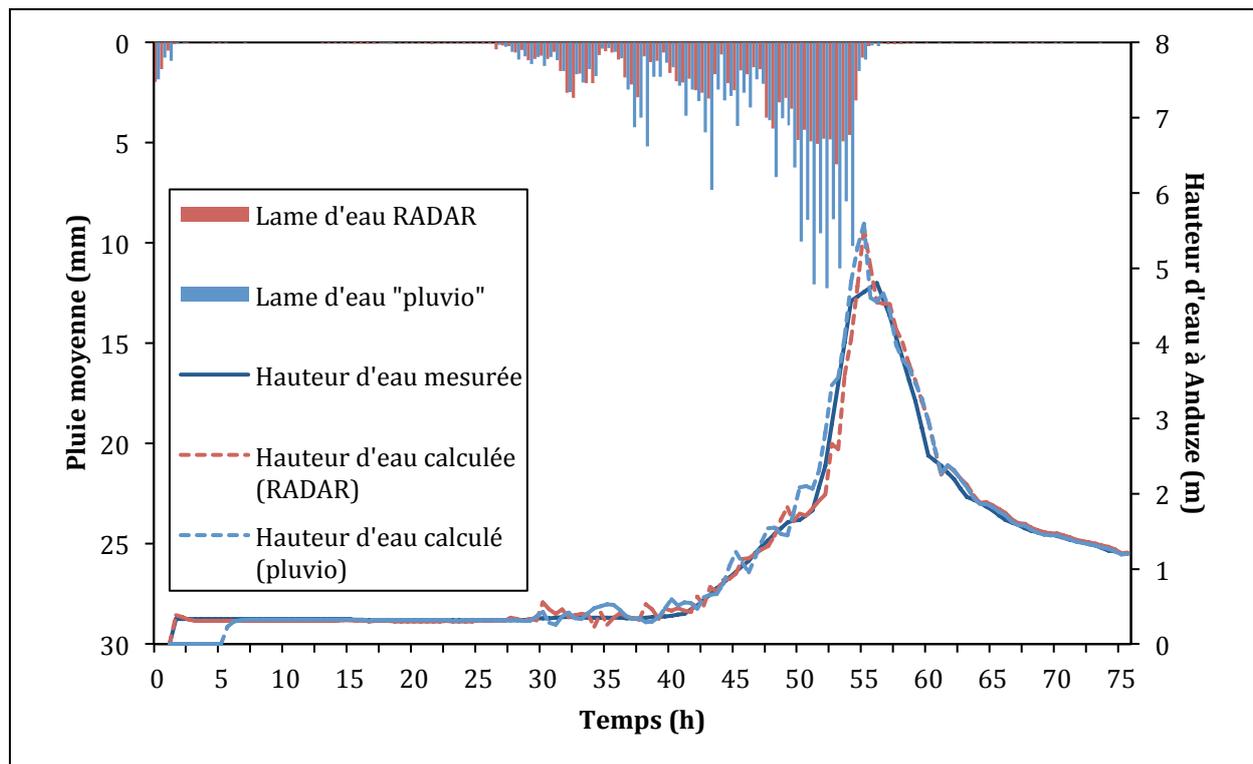


Figure 2 – Prévisions pour $h_p = 1$ h (ev. 13)

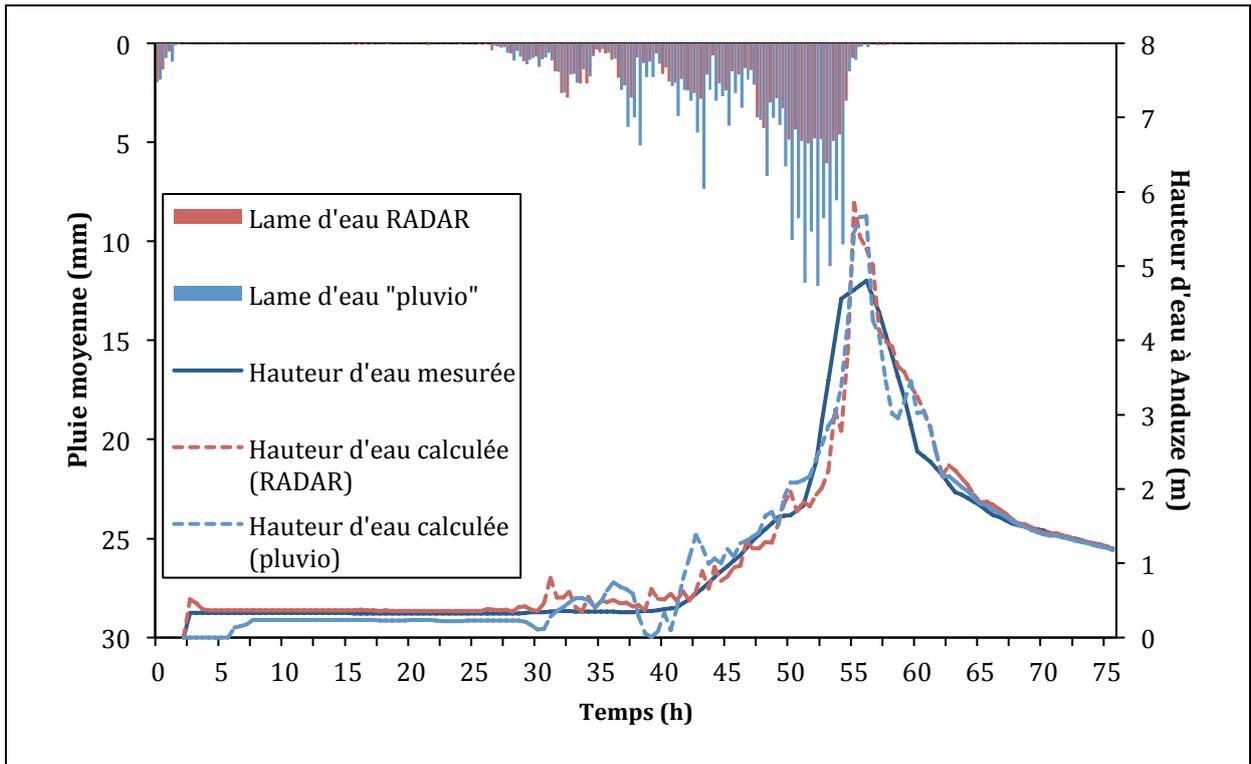


Figure 3 – Prévisions pour $h_p = 2$ h (ev. 13)

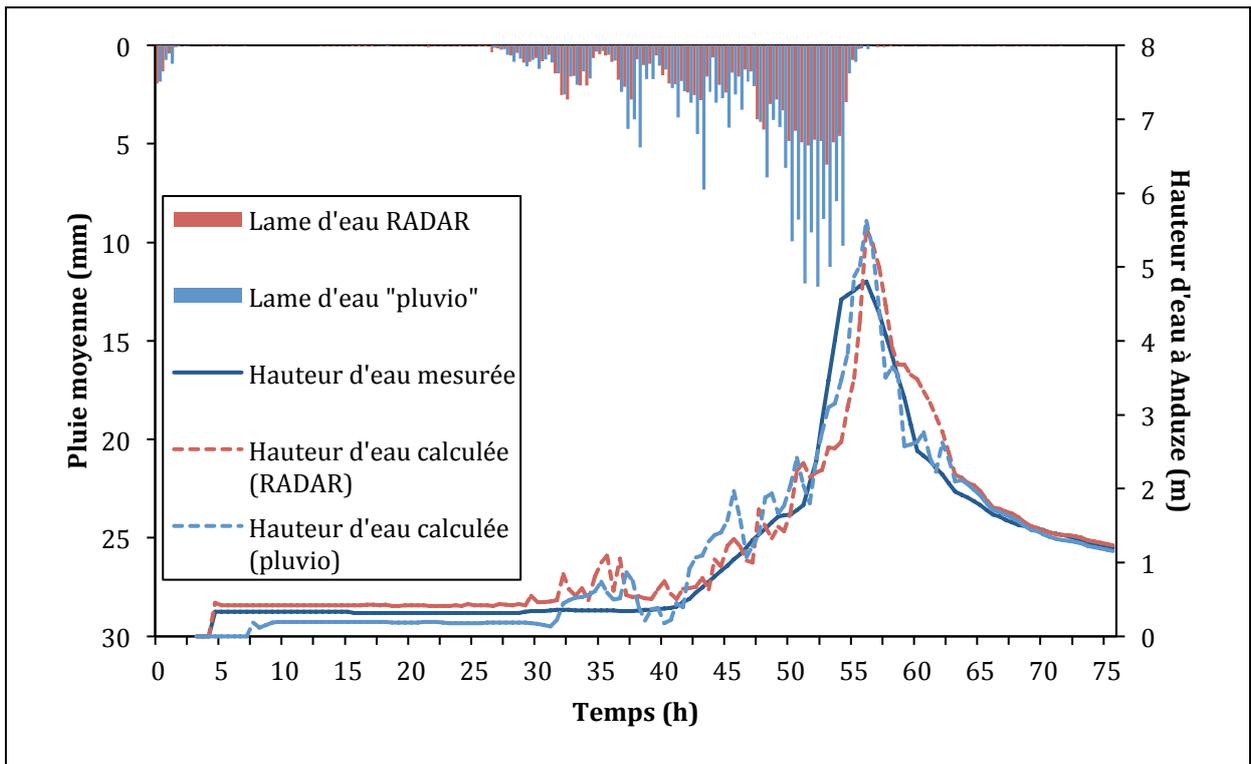


Figure 4 – Prévisions pour $h_p = 3$ h (ev. 13)

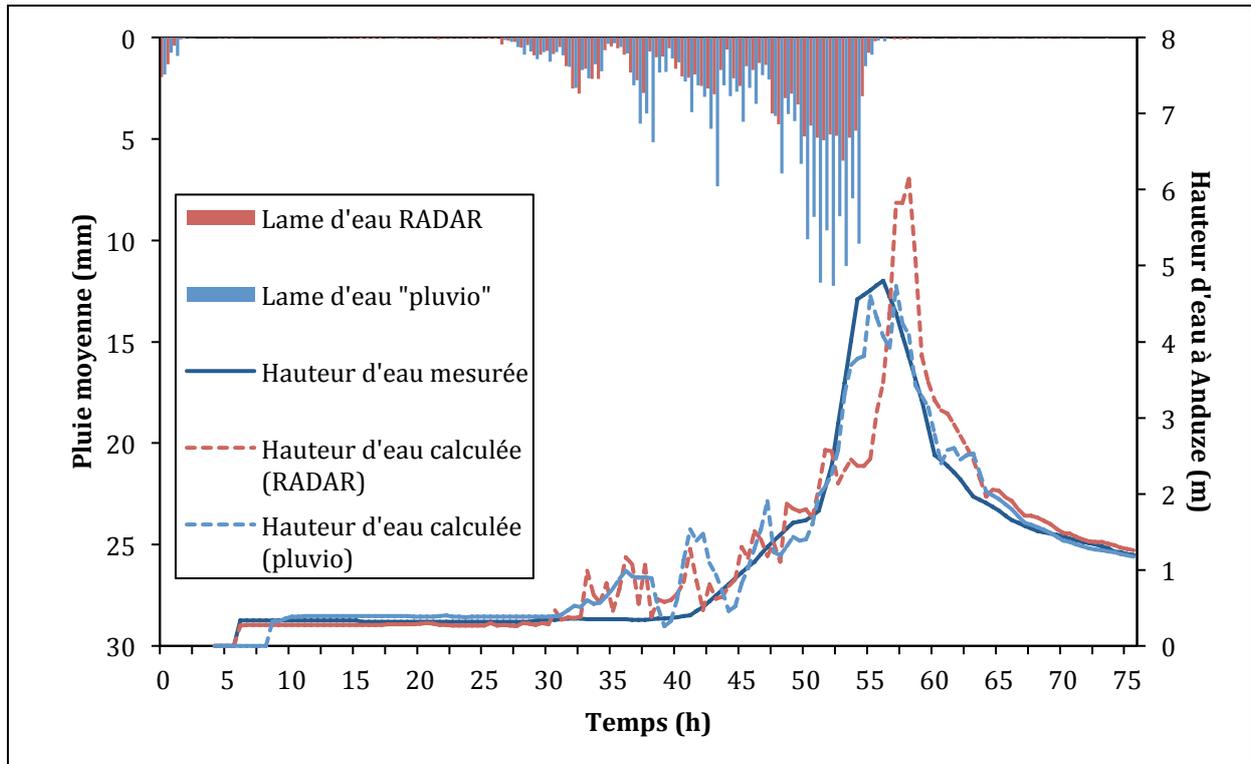


Figure 5 – Prévisions pour $h_p = 4$ h (ev. 13)

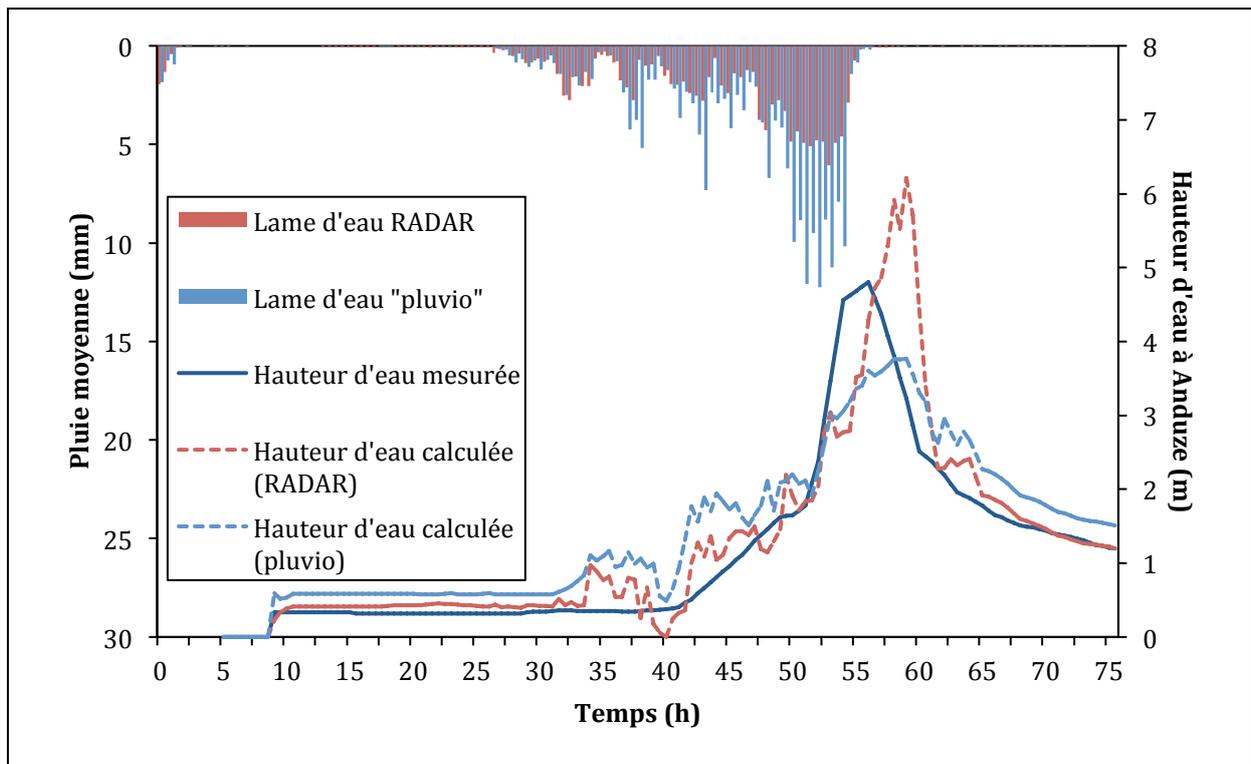


Figure 6 – Prévisions pour $h_p = 5$ h (ev. 13)

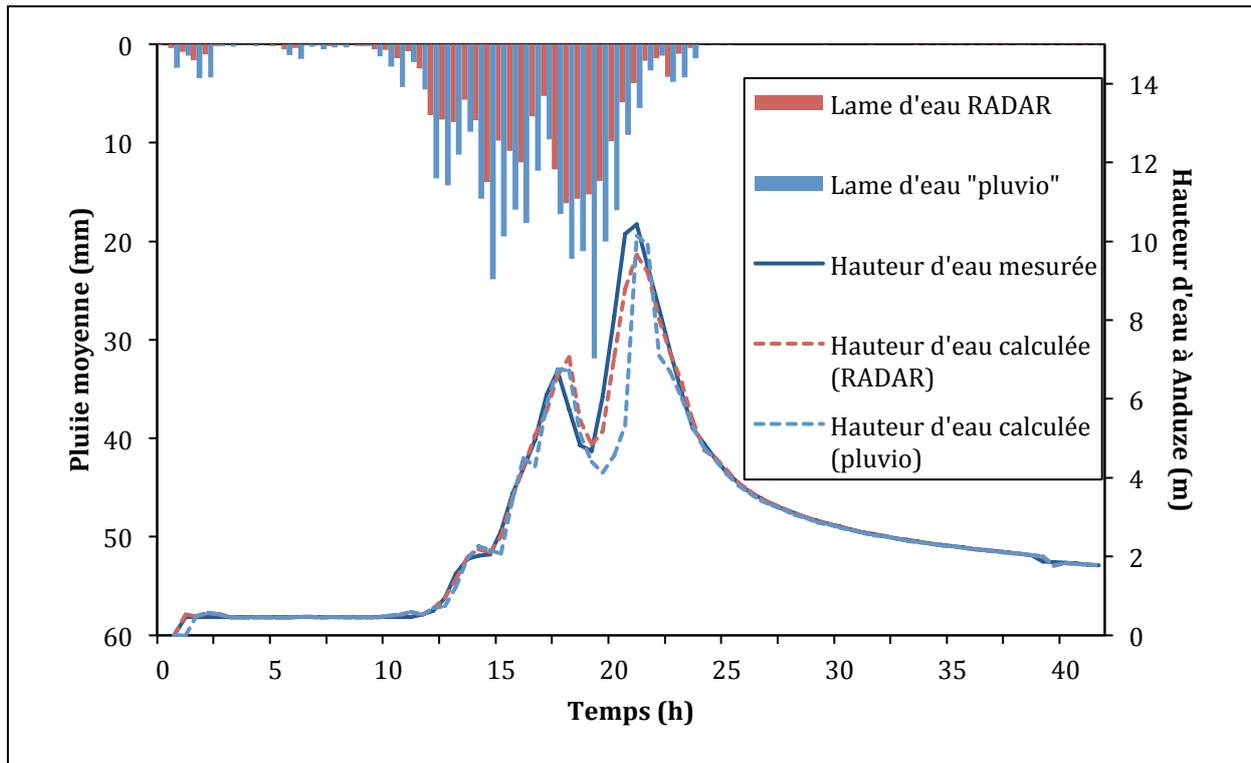


Figure 7 – Prévisions pour $h_p = 0,5$ h (ev. 19)

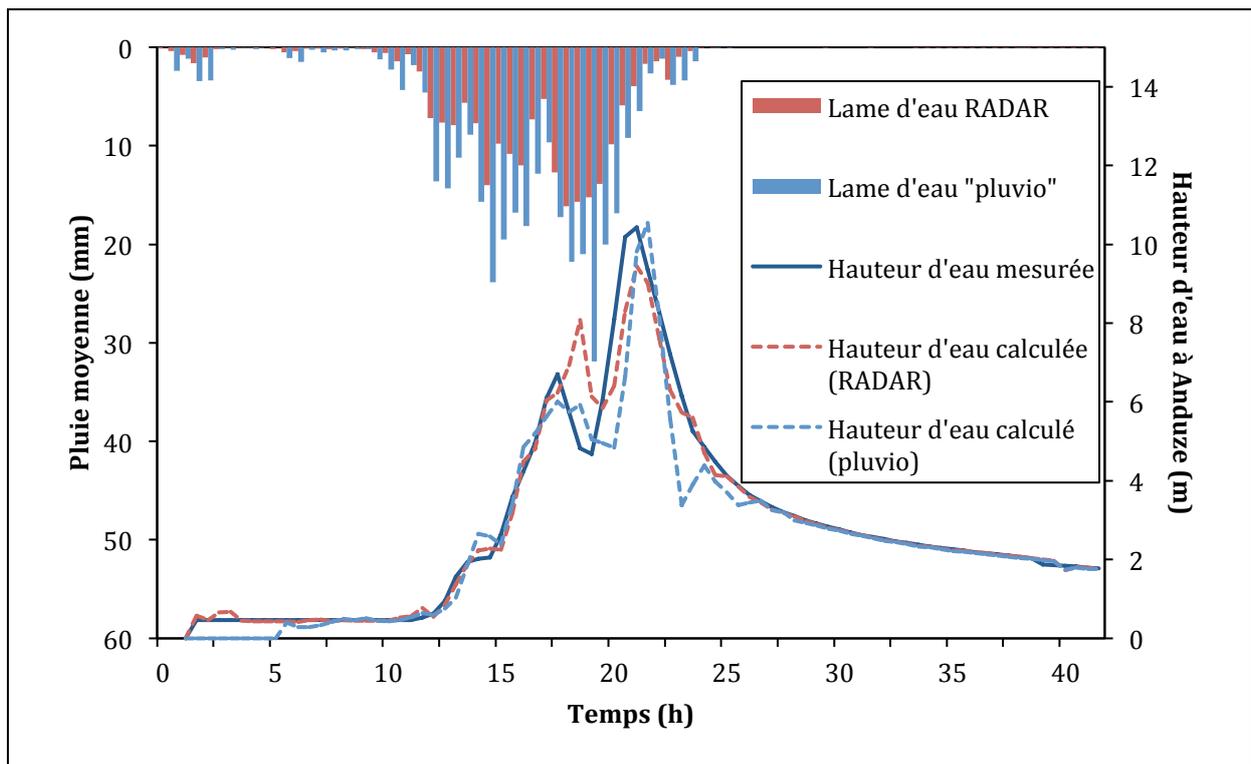


Figure 8 – Prévisions pour $h_p = 1$ h (ev. 19)

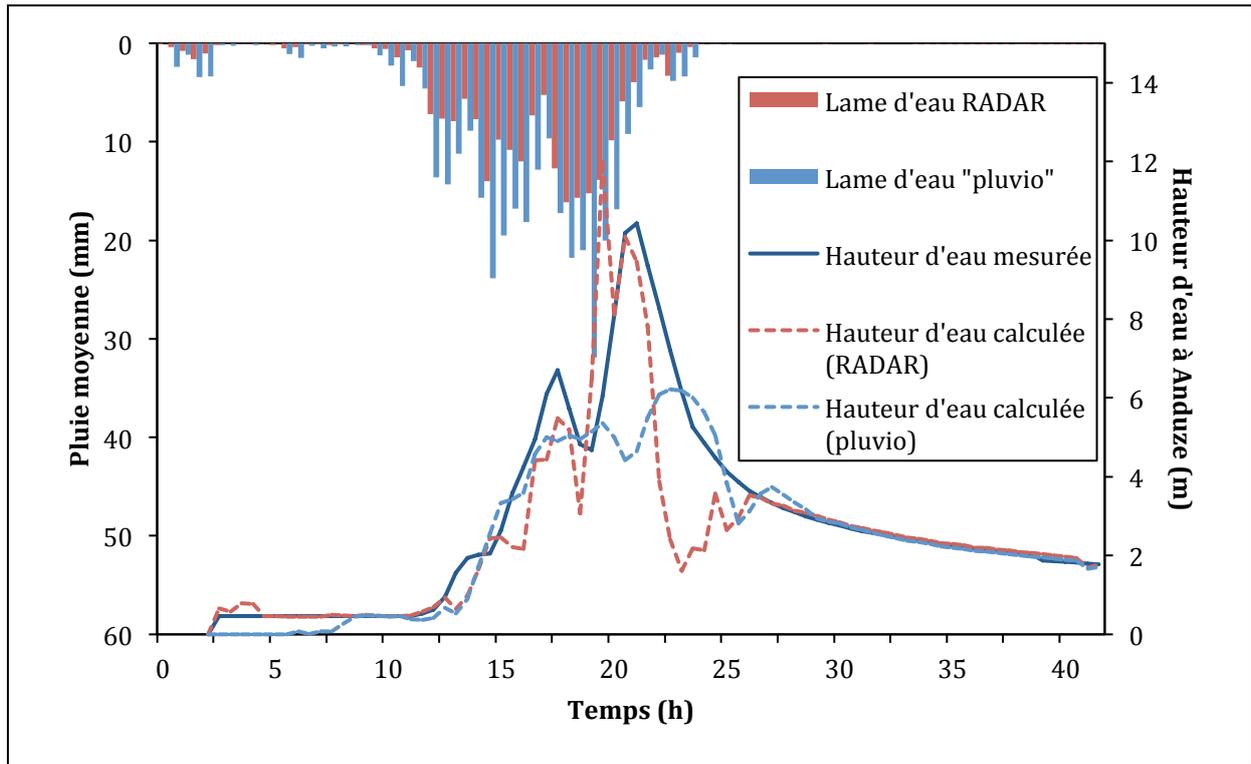


Figure 9 – Prévisions pour $h_p = 2$ h (ev. 19)

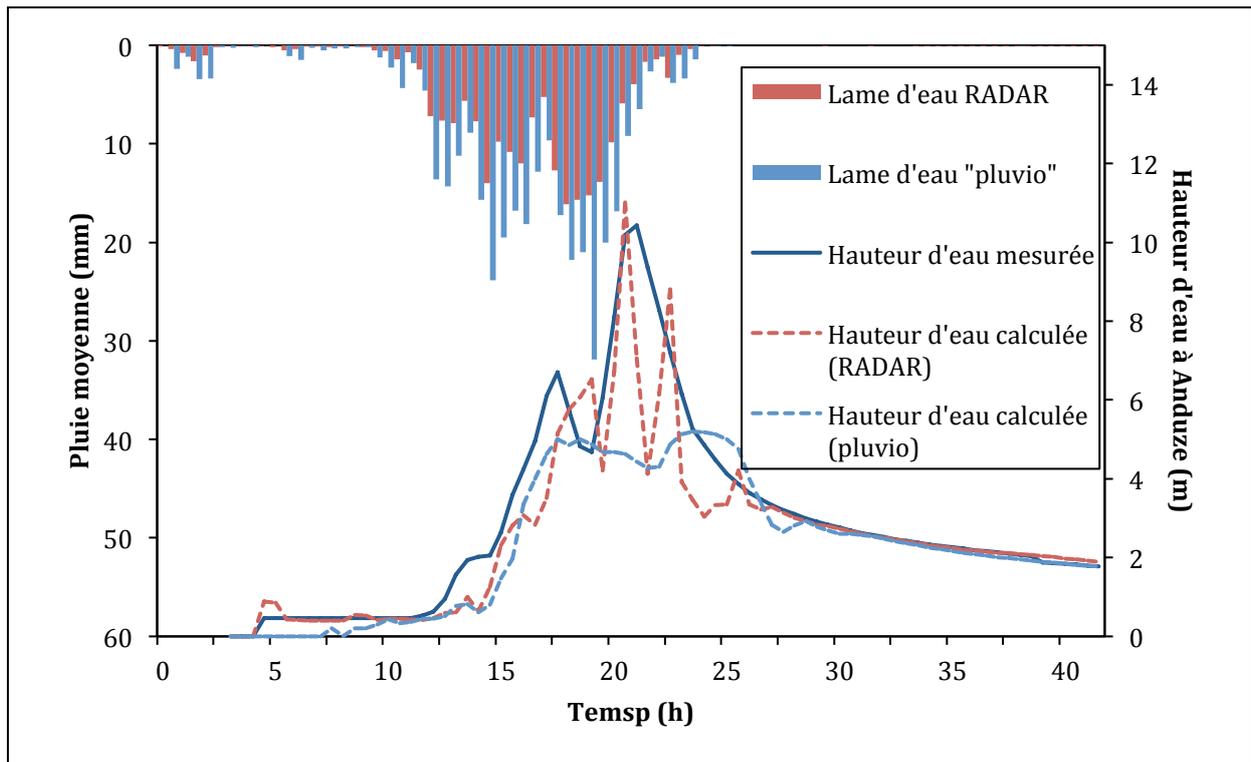


Figure 10 – Prévisions pour $h_p = 3$ h (ev. 19)

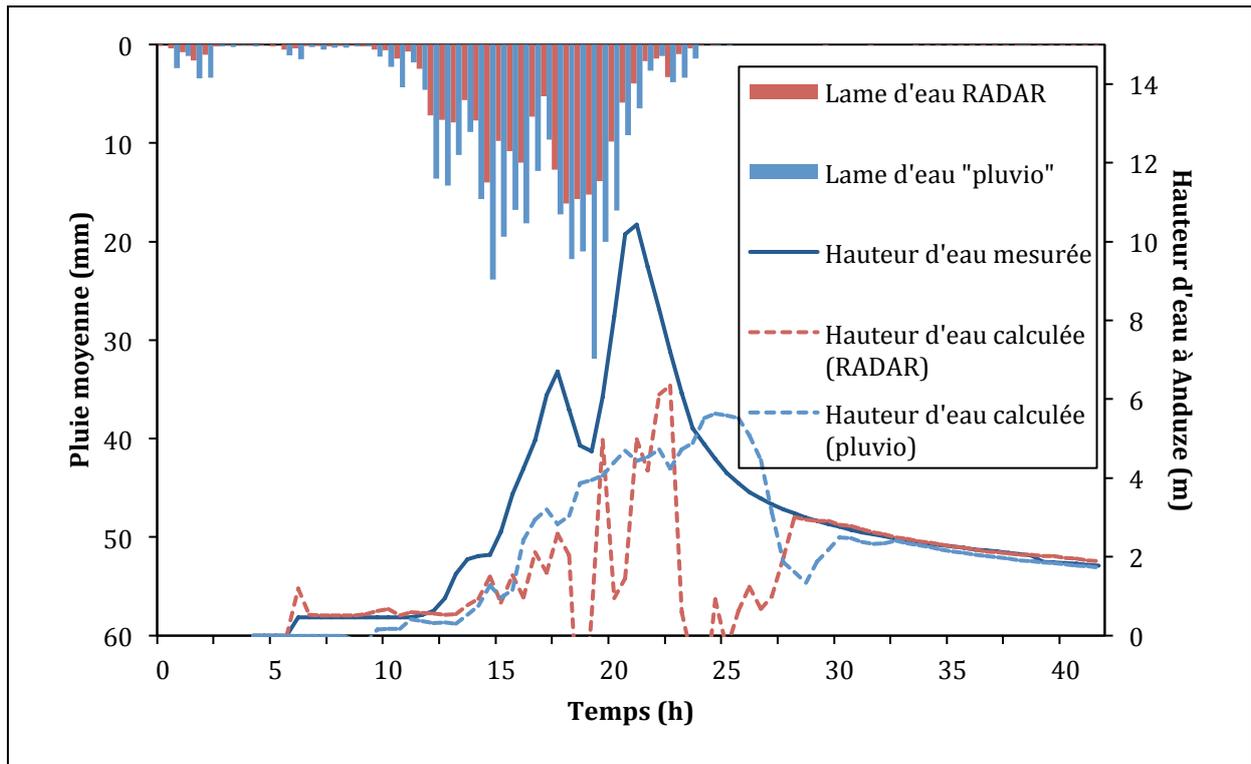


Figure 11 – Prévisions pour $h_p = 4$ h (ev. 19)

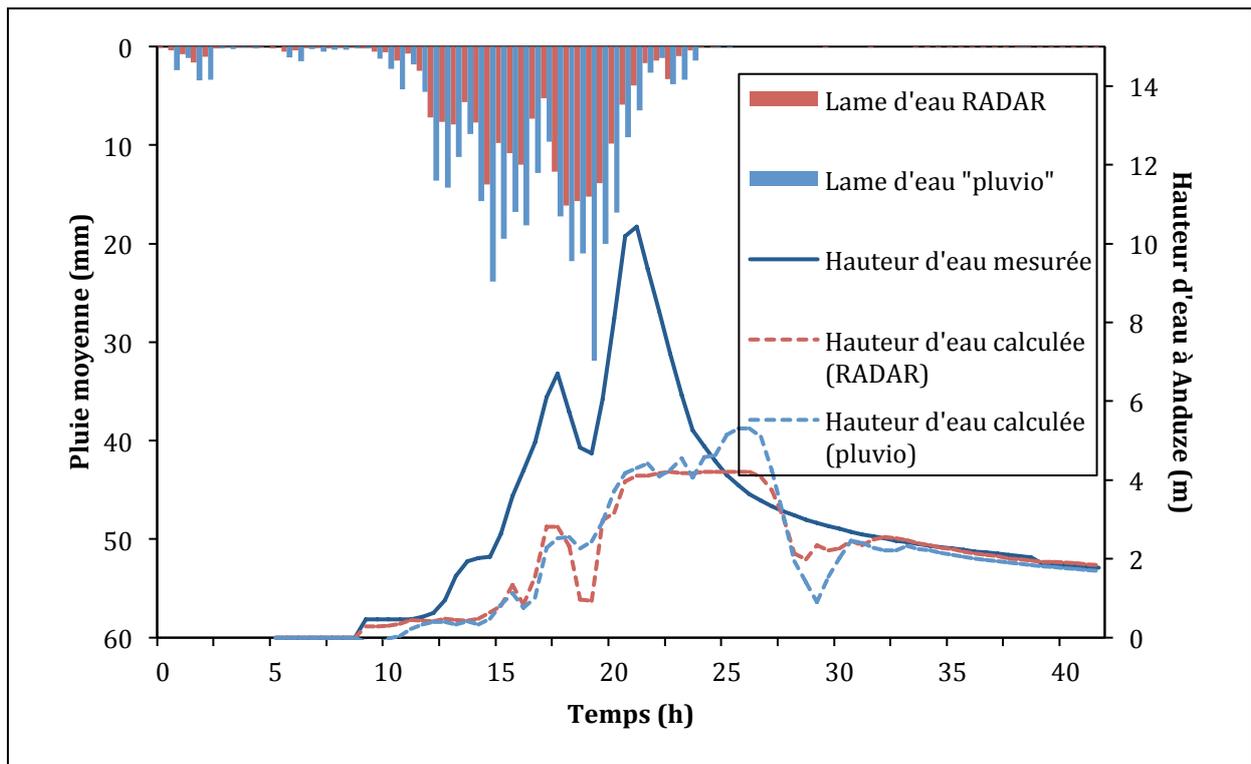


Figure 12 – Prévisions pour $h_p = 5$ h (ev. 19)

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

