DOCTORAT ECOLOGIE BRETAGNE GEOSCIENCES LOIRE AGRONOMIE ALIMENTATION

Nantes Université

THESE DE DOCTORAT DE

NANTES UNIVERSITE

ECOLE DOCTORALE N° 600
Ecole doctorale Ecologie, Géosciences, Agronomie et Alimentation
Spécialité : *Sciences de la Terre*

Par Maryse CHARPENTIER- -NOYER

Évaluation de prévisions hydrométéorologiques d'ensemble des crues soudaines à partir de données sur les impacts

Thèse présentée et soutenue à Nantes, le 15 décembre 2022
Unité de recherche : Université Gustave Eiffel, GERS-LEE

Rapporteurs avant soutenance :

Sandrine ANQUETIN	Directrice de recherche, CNRS-IGE
Freddy VINET	Professeur des Universités, Université Paul Valéry Montpellier 3

Composition du Jury :

Président :	Agnès BALTZER	Professeur des Universités, Université de Nantes
Examinateurs :	Sandrine ANQUETIN	Directrice de recherche, CNRS-IGE
	Freddy VINET	Professeur des Universités, Université Paul Valéry Montpellier 3
	Isabella ZIN	Maître de conférences, Université Grenoble-Alpes, IGE
Dir. de thèse :	Éric GAUME	Ingénieur général des ponts, eaux et forêts, UGE, Nantes
Co-dir. de thèse :	Olivier PAYRASTRE	Ingénieur divisionnaire des travaux publics de l'Etat, UGE, Nantes

Invités

François BOUTTIER	Ingénieur général des ponts, eaux et forêts, Météo-France, CNRM
Bruno JANET	Ingénieur en chef des ponts, eaux et forêts, SCHAPI

Table des matières

Table des matières **i**

Liste des figures **vi**

Liste des tableaux **xv**

1 Introduction Générale **9**
- 1.1 Le risque inondation, un risque majeur . . . 10
- 1.2 Les crues soudaines, un phénomène fréquent en France . . . 11
  - 1.2.1 Typologie des crues et définition des crues soudaines . . . 11
  - 1.2.2 Les crues soudaines observées dans l'arc méditerranéen . . . 13
- 1.3 Les systèmes actuels de prévision et d'alerte aux inondations . . 15
  - 1.3.1 Un dispositif d'avertissement précoce sur les cours d'eau principaux bien établi 15
  - 1.3.2 L'apparition de systèmes d'anticipation et d'avertissement pour les petits cours d'eau . . . 17
- 1.4 Questions scientifiques soulevées et objectifs de la thèse . . . 18
- 1.5 Organisation de la thèse . . . 20

2 La prévision immédiate des crues-éclair : de la prévision des pluies, à celle des débits, des zones inondées et des impacts associés **23**
- 2.1 Définition d'une prévision immédiate . . . 24
- 2.2 Principe de fonctionnement d'une chaîne de prévision des crues . 24
- 2.3 Les différentes observations de précipitations mobilisables . . . 25
- 2.4 Les différentes méthodes de prévision déterministe des précipitations . . . 27
  - 2.4.1 La prévision météorologique numérique . . . 27
  - 2.4.2 L'assimilation de données pour améliorer la prévision immédiate . . . 30
  - 2.4.3 L'extrapolation des lames d'eau radar observées pour la prévision immédiate 31
- 2.5 Vers une approche probabiliste des prévisions . . . 32
- 2.6 La production de prévisions hydrologiques . . . 34
  - 2.6.1 Utilisation des prévisions de pluie en entrée d'un modèle hydrologique . . . 34
  - 2.6.2 Développement des prévisions d'ensemble des débits . . . 36
- 2.7 Le développement de prévisions des inondations et des impacts . . 37
  - 2.7.1 Une identification des secteurs inondés difficile . . . 37
  - 2.7.2 Développement de méthodes de prévisions d'impact . . . 38

2.7.3 Outils de cartographie automatisée 39
2.8 Les méthodes d'évaluation d'une prévision d'ensemble 40
2.8.1 Scores statistiques classiques de vérification des prévisions d'ensemble 40
2.8.2 Enjeux de l'évaluation événementielle des prévisions d'ensemble 42
2.9 Cadre méthodologique retenu pour l'évaluation événementielle des prévisions hydrométéorologiques d'ensemble 43
2.9.1 Étape 1 : Fiabilité des ensembles de pluie 45
2.9.2 Étape 2 : Capacité d'anticipation de seuils de débit pendant la montée de crue 45
2.9.3 Étape 3 : Prévisions détaillées à certains exutoires 45
2.9.4 Étape 4 : Capacité d'anticipation de seuil d'impacts 46
2.9.5 Étape 5 : Prévisions d'impacts détaillées sur certains tronçons 46

3 Évènements étudiés, données et modèles utilisés **49**
3.1 Trois crues soudaines majeures de l'année 2018 retenues comme cas d'étude 50
3.1.1 Les crues de l'Aude en octobre 2018 50
3.1.2 Les crues de la partie aval de l'Ardèche et de la Cèze, en août 2018 51
3.1.3 Les crues de l'Argens, en octobre-novembre 2018 55
3.2 Données disponibles pour l'étude des trois événements 56
3.2.1 Observations pluviométriques disponibles 56
3.2.2 Produits de prévision d'ensemble des précipitations basés sur le modèle AROME-FR 56
3.2.3 Données hydrologiques 59
3.2.4 Données d'occupation des sols 60
3.3 Les modèles utilisés 61
3.3.1 Le modèle hydrologique Cinecar 61
3.3.2 Catalogue d'emprises inondées utilisé pour la modélisation des impacts . . . 65

4 Évaluation des produits de prévision d'ensembles des pluies sur la base des pluies observées **69**
4.1 Méthodologie retenue pour l'évaluation événementielle des prévisions d'ensemble de pluie 71
4.1.1 Principe général : une évaluation en 3 étapes 71
4.1.2 Le Diagramme de rang (ou diagramme de Talagrand) 72
4.1.3 Le CRPS (Continuous Rank Probability Score) 73
4.2 Comparaison des pluies prévues et observées du point de vue temporel et spatial 74
4.2.1 Évolution temporelle des pluies 74
4.2.2 Localisation des fortes intensités de pluies 79
4.3 Évaluation de la fiabilité et la finesse des prévisions d'ensemble de pluie horaires 83
4.3.1 Résultats obtenus pour les diagrammes de rang 84
4.3.2 Résultats obtenus pour les CRPS 86
4.3.3 Influence du nombre de membres sur les scores de CRPS obtenus 89

5 Évaluation des prévisions d'ensemble des débits de crue **95**
5.1 Principes de l'évaluation des prévisions de débits 97
5.1.1 Préparation des prévisions d'ensemble des débits et définition des simulations de référence 97
5.1.2 Méthodologie d'évaluation des prévisions de débits 100
5.2 Résultats de l'évaluation sur l'événement de l'Aude 104
5.2.1 Détection des dépassements de seuils 104
5.2.2 Délai d'anticipation pour les trois produits de prévision d'ensemble 106
5.2.3 Analyse détaillée des prévisions pour quelques exutoires 108
5.3 Résultats de l'évaluation sur l'événement de l'Ardèche 112
5.3.1 Détection des dépassements de seuils de débits 112
5.3.2 Délais d'anticipation pour les trois produits de prévision d'ensemble 116
5.3.3 Analyse détaillée pour quelques exutoires 117
5.4 Résultats de l'évaluation sur l'événement du Var 121
5.4.1 Détection des dépassements de seuils de débits 121
5.4.2 Délai d'anticipation pour les trois produits de prévision d'ensemble 124
5.4.3 Analyse détaillée pour quelques exutoires 125
5.5 Conclusions sur l'analyse de l'anticipation des débits par les prévisions d'ensemble 128
5.5.1 Comment la méthodologie mise en place répond-elle aux besoins de l'évaluation de prévisions d'ensemble adaptées aux crues soudaines ? 128
5.5.2 Qu'en est-il de la performance des trois ensembles de prévision testés ? . . . 130

6 Élaboration et évaluation de prévisions d'impacts des inondations **135**
6.1 Passage de la prévision des débits à la prévision des impacts . . . 137
6.1.1 Choix méthodologiques pour la construction des courbes d'impacts 137
6.1.2 Obtention de courbes débit-impacts continues pour chaque tronçon de cours d'eau 139
6.2 Évaluation de la chaîne de simulation pluie-débit-impacts obtenue 141
6.2.1 Emprise inondée de référence pour la crue d'octobre 2018 142
6.2.2 Comparaison de la simulation pluie-débit-nombre de bâtiments avec le nombre de bâtiments situés dans l'emprise de référence 142
6.3 Évaluation des prévisions d'ensembles des impacts sur les bâtiments 144
6.3.1 Courbes ROC obtenues pour l'anticipation de seuils de bâtiments inondés . . 145
6.3.2 Analyse spatiale de l'anticipation d'un seuil de bâtiments inondés pour le percentile 75% 146
6.3.3 Délais d'anticipation des seuils de bâtiments inondés à partir des prévisions d'ensembles 149
6.4 Comparaison des résultats de l'anticipation de seuils de débits et de seuils de bâtiments inondés 149
6.4.1 Première comparaison à partir des aires sous les courbes ROC (AUC) 149
6.4.2 Origines possibles des différences de résultats 150

6.4.3 Deuxième comparaison avec association des seuils de débits et des seuils d'impacts . . . 151

7 Intérêt des prévisions d'ensemble des impacts pour les prises de décision d'un service de secours **157**
7.1 Étude de cas retenue : la gestion de la crue d'octobre 2018 par le SDIS de l'Aude . . . 158
7.1.1 Organisation générale du SDIS de l'Aude . . . 158
7.1.2 Chronologie des interventions réalisées par le SDIS lors de l'événement . . . 160
7.2 Description et mise en œuvre de la modélisation des interventions du SDIS . . . 162
7.2.1 Principe général et modalités de déclenchement des interventions . . . 162
7.2.2 Représentation des temps de trajet vers les sites d'intervention . . . 167
7.2.3 Choix des autres paramètres de modélisation . . . 168
7.2.4 Mesure de l'efficacité des interventions déclenchées (Fonction de coût) . . . 169
7.2.5 Choix des paramètres de modélisation . . . 171
7.2.6 Quelques améliorations possibles du modèle . . . 171
7.3 Première simulation avec un nombre illimité d'équipes de secours 172
7.3.1 Évolution spatio-temporelle des équipes d'intervention pendant l'événement . 173
7.3.2 Évolution du nombre de bâtiments secourus selon les différents scénarios . . 175
7.3.3 Calcul du coût en bâtiments non secourus, pour toutes les modélisations . . 177
7.3.4 Évolution temporelle du nombre d'équipes mobilisées . . . 179
7.3.5 Bilan de la première simulation avec des moyens d'intervention illimités . . . 180
7.4 Simulations avec des moyens d'intervention limités . . . 180
7.4.1 Résultats pour une situation proche de celle d'octobre 2018 . . . 181
7.4.2 Influence de la valeur du seuil d'intervention . . . 185
7.4.3 Influence de la capacité d'intervention choisie . . . 191
7.4.4 Influence du nombre d'équipes mobilisables . . . 198
7.5 Bilan sur l'utilisation des impacts pour les prises de décision d'un service de secours . . . 201
7.5.1 Adaptabilité de la méthode à d'autres cas d'études et d'autres utilisateurs . 201
7.5.2 Questionnement sur le choix de la fonction de coût . . . 201

8 Conclusions et perspectives **207**
8.1 Bilan du travail de thèse . . . 208
8.2 Conclusions sur les résultats obtenus . . . 209
8.2.1 Performances des trois produits de prévision d'ensemble testés . . . 209
8.2.2 Intérêts et limites de la méthodologie d'évaluation des prévisions de pluie et débits . . . 211
8.2.3 Plus-value de l'évaluation basée sur les impacts . . . 212
8.3 Des compléments encore nécessaires pour aboutir à une véritable chaîne opérationnelle de prévision immédiate des impacts de crues soudaines . . . 213
8.3.1 Aspects relatifs à l'évaluation des pluies . . . 213

8.3.2 Aspects relatifs à l'évaluation des débits 213
8.3.3 Aspects relatifs à la prévision des impacts 214

Bibliographie **217**

A Un cadre méthodologique pour l'évaluation événementielle des prévisions hydrométéorologiques des crues éclair à courte échéance **233**

Liste des figures

1.1 Sinistralité liée aux catastrophes naturelles (hors sinistralité automobile) de 1882 à 2017 en France (CCR 2018b) 10
1.2 Illustration des différentes causes de crues : a) Crue à Trans-en-Provence (Var) en juin 2010 après d'intenses précipitations (source : P Cantet, INRAE-Aix), b) Crue nivale sur la Seine après plusieurs jours de précipitations de pluie et de neige en février 2018 (source : AFP/Stéphane De Sakutin) et c) Débâcles de glace sur le Danube près de Belgrade après plusieurs semaines de grand froid (source : AFP/ Alexa Stankovic) 12
1.3 Dommages suite à des crues soudaines : a) Crue des cadereaux à Nîmes en octobre 1988 (source : Pluies extrêmes), b) Gard, septembre 2002 (source : France 3 Région Occitanie) et c) Dégâts après le passage de la tempête Alex, sur le Boréon à Saint-Martin-Vésubie (source : France Info) 14
1.4 Intensification des fortes pluies sur les régions méditerranéennes (Soubeyroux 2021) 14
1.5 Chaîne de prévision hydrologique d'EFAS (source : efas.eu) 15
1.6 Chaîne de modélisation des impacts intégrée à EFAS (Dottori *et al.* 2017) . . 15
1.7 Couverture mondiale de GloFAS et échéances de prévision (en jours) (source : Alfieri *et al.* (2013)) 16
1.8 Principe de fonctionnement de Vigicrues Flash (inspiré de INRAE) 18
1.9 Organisation du projet PICS (source : PICS) 19

2.1 Schéma général de fonctionnement d'une chaîne de prévision des crues, depuis les prévisions météorologiques jusqu'aux impacts potentiels des précipitations et des crues 24
2.2 Localisation des différentes stations pluviométriques représentées par les points rouges (série pluviométrique complète de 1958 à 2022) et bleus (série pluviométrique incomplète) (source : Pluies Extrêmes) 26
2.3 Résolution horizontale en kilomètres du modèle ARPEGE : minimum 5 km, moyenne 11 km et maximum 24 km (CNRM) 28
2.4 Précipitations cumulées sur 3 h (en mm) entre 18 h et 21 h (UTC) sur le sud-est de la France pour la situation du 3 octobre 2015 : a) Prévision AROME-France issue de la prévision faite à 3 h UTC (échéance 24 h), b) Analyse Antilope de précipitations cumulées sur 3 h. (CNRM) 29
2.5 Relations entre les prévisions AROME-FR et les prévisions AROME-PI (Météo-France 2020) 31

2.6 Principe théorique de la prévision d'ensemble (D'après DWD) 32
2.7 Illustration d'une prévision d'ensemble composée de 12 membres (AROME-EPS), pour des cumuls pluviométriques sur 3 heures 33
2.8 Illustration de la mise en œuvre du modèle d'impact sur un tronçon de rivière : a) catalogue des zones inondées et b)relation débit – impact interpolé (Le Bihan *et al.* 2017) 38
2.9 Cartes des débits de pointe (périodes de retour) et des impacts associés (nombre de polices d'assurance inondées) simulées par le modèle, et des nombres de sinistres extraits de la base de données CCR pour Alès 2002 (Le Bihan *et al.* 2017) 38
2.10 Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble 44

3.1 Cumul de pluie (lame d'eau ANTILOPE J+1) du 14 octobre au 16 octobre 2018 51
3.2 Cumul de pluie (lame d'eau ANTILOPE J+1) du 9 août 2018 53
3.3 Cumul de pluie (lame d'eau ANTILOPE J+1) du 29 octobre au 2 novembre 2018 55
3.4 Illustration de la combinaison des prévisions pour former les différents ensembles étudiés 57
3.5 Évaluation des prévisions d'ensemble pour des pluies excédant 5 mm/h a) Diagramme de fiabilité et b) Courbes ROC 58
3.6 Représentation schématique d'un bassin versant dans Cinecar. Chaque sous-bassin est constitué d'un thalweg de section rectangulaire et d'un ou deux versants rectangulaires (Gaume 2002) 61
3.7 Découpage en sous-bassins réalisé avec TauDEM où chaque couleur correspond à un sous-bassin : a) pour l'Aude, b) pour l'Ardèche et c) pour le Var 63
3.8 Valeurs de CN retenues pour reproduire les débits pour l'événement de l'Aude 2018 64
3.9 Écarts entre les débits spécifiques observés ou relevés avec les débits spécifiques simulés a) pour l'événement de l'Aude, b) pour l'événement de l'Ardèche et c) pour l'événement du Var 65

4.1 Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, est représentée l'étape d'évaluation menée dans le Chapitre 4. 70
4.2 Illustration des différentes formes caractéristiques des diagrammes de rang . . 72
4.3 Illustration du calcul du CRPS avec les fonctions de répartition observée (en vert) et prévue (en rose) d'une variable 73
4.4 Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement de l'Aude 75

4.5 Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement d'août 2018 sur l'Ardèche 76
4.6 Cumuls de pluie pour Antilope J+1 a) Du 29 octobre 00 h au 30 octobre 12 h et b) du 30 octobre 14 h au 1er novembre 23 h 77
4.7 Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement de fin octobre 2018 dans le Var (secteur de l'Argens) 78
4.8 Comparaison des cumuls de précipitations observées et de la somme des pluies horaires prévues (pour une échéance de prévision d'une heure) sur 15 heures (du 14 octobre 20 h au 15 octobre 10 h) pour l'événement de l'Aude : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. 80
4.9 Comparaison des cumuls de précipitations observées et de la somme des pluies horaires prévues (pour une échéance de prévision de six heures) sur 15 heures (du 14 octobre 20 h au 15 octobre 10 h) pour l'événement de l'Aude : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. . . . 80
4.10 Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision d'une heure) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. 81
4.11 Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision de six heures) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. 81
4.12 Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision d'une heure) pour l'événement du Var : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. 82

4.13 Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision de six heures) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%. 82
4.14 Diagrammes de rang sur l'événement de l'Aude, pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures 84
4.15 Diagrammes de rang sur l'événement de l'Ardèche, pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures 85
4.16 Diagrammes de rang sur l'événement du Var pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures 85
4.17 CRPS calculés sur l'événement de l'Aude pour AROME-EPS, pepi et pertDpepi 87
4.18 CRPS calculés sur l'événement de l'Ardèche pour AROME-EPS, pepi et pertDpepi 87
4.19 CRPS calculés sur l'événement du Var pour AROME-EPS, pepi et pertDpepi 88
4.20 Comparaison des CRPS obtenus sur les ensembles pepi (en orange) et pertDpepi (en violet) originaux (en abscisse) et sur les ensembles réduits (en ordonnée), toutes échéances confondues, pour l'événement de l'Aude 90
4.21 Comparaison des différences de CRPS entre AROME-EPS et les ensembles pepi, pertDpepi, pepi-réduit et pertDpepi-réduit, toutes échéances confondues, pour l'événement de l'Aude 90

5.1 Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, sont représentées les deux étapes d'évaluation menées dans le Chapitre 5. 96
5.2 Combinaison entre les pluies Antilope J+1 et les prévisions d'ensemble (exemple d'AROME-EPS) pour former les données d'entrée de Cinecar 98
5.3 Configurations possibles en cas de dépassement de seuil par RS 101
5.4 Configurations possibles en cas de non-dépassement de seuil par RS 101
5.5 Hydrogrammes obtenus pour l'exutoire 1168, à partir des prévisions d'ensemble pertDpepi, de 21 h à 2 h le 15 octobre 2018 dans l'Aude 103
5.6 Courbes ROC obtenues pour les trois ensembles et RF0 pour la période de retour 10 ans. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0 105
5.7 Cartes d'anticipation (0-6 h) du seuil de période de retour 10 ans : a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation minimale de 15 minutes 106
5.8 Comparaison des délais d'anticipation du dépassement de seuil de période de retour 10 ans pour les percentiles a) 20%, b) 50%, c) 75% et d) 95% 107

5.9 Localisation des six exutoires retenus pour une analyse détaillée des prévisions de débits pour l'événement de l'Aude 108
5.10 Hydrogrammes de prévisions obtenues avec le modèle Cinecar pour l'exutoire n°1 (216 km²) a) AROME-EPS, b) pepi et c) pertDpepi 109
5.11 Hydrogrammes obtenus pour l'exutoire n°2 (197 km²) a) AROME-EPS, b) pepi et c) pertDpepi 109
5.12 Hydrogrammes obtenus pour l'exutoire n°3 (85 km²) a) AROME-EPS, b) pepi et c) pertDpepi 110
5.13 Hydrogrammes obtenus pour l'exutoire n°4 (173 km²) a) AROME-EPS, b) pepi et c) pertDpepi 111
5.14 Hydrogrammes obtenus pour l'exutoire n°5 (263 km²) a) AROME-EPS, b) pepi et c) pertDpepi 111
5.15 Hydrogrammes obtenus pour l'exutoire n°6 (257 km²) a) AROME-EPS, b) pepi et c) pertDpepi 112
5.16 Résultats de l'anticipation pour RF0 pour différents seuils de débit a) Débit de période de retour 2 ans Q_2, b) 0.75 * Q_2 et c) 0.5 * Q_2 113
5.17 Courbes ROC obtenues pour les trois ensembles de prévision des débits, et RF0 pour la période de retour 0.75*Q_2. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0 avec a) Vue globale des valeurs des courbes ROC et b) Zoom sur les valeurs obtenues. 114
5.18 Cartes d'anticipation (0-6 h) du seuil de période de retour 0.75*Q_2 : a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation supérieure à 15 minutes 115
5.19 Comparaison des délais d'anticipation du dépassement de seuil de période de retour 0,75*Q_2 pour les percentiles a) 20%, b) 50%, c) 75% et d) 95% pour l'événement d'août 2018. 116
5.20 Localisation des quatre exutoires retenus pour une analyse détaillée des prévisions des débits, pour l'événement de l'Ardèche 117
5.21 Hydrogrammes obtenus pour l'exutoire n°1 (107 km²) a) AROME-EPS, b) pepi et c) pertDpepi 118
5.22 Hydrogrammes obtenus pour l'exutoire n°2 (15 km²) a) AROME-EPS, b) pepi et c) pertDpepi 119
5.23 Hydrogrammes obtenus pour l'exutoire n°3 (42 km²) a) AROME-EPS, b) pepi et c) pertDpepi 119
5.24 Hydrogrammes obtenus pour l'exutoire n°4 (31 km²) a) AROME-EPS, b) pepi et c) pertDpepi 120
5.25 Résultats de détection des dépassements de seuils de débits, pour RF0 et des seuils de a) Période de retour 2 ans et b) Période de retour 5 ans 121
5.26 Courbes ROC obtenues pour les trois ensembles de prévision et RF0 pour la période de retour 5 ans. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0. 122

5.27 Cartes d'anticipation (0-6 h) du seuil de période de retour 5 ans : a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation supérieure à 15 minutes 123
5.28 Comparaison des délais d'anticipation du dépassement de seuil de période de retour 5 ans pour les percentiles a) 20%, b) 50%, c) 75% et d) 95% 125
5.29 Localisation des quatre exutoires retenus pour une analyse détaillée des prévisions des débits pour l'événement du Var 125
5.30 Hydrogrammes obtenus pour l'exutoire n°1 (2.3 km²) a) AROME-EPS, b) pepi et c) pertDpepi 126
5.31 Hydrogrammes obtenus pour l'exutoire n°2 (57 km²) a) AROME-EPS, b) pepi et c) pertDpepi 126
5.32 Hydrogrammes obtenus pour l'exutoire n°3 (33 km²) a) AROME-EPS, b) pepi et c) pertDpepi 127
5.33 Hydrogrammes obtenus pour l'exutoire n°4 (93 km²) a) AROME-EPS, b) pepi et c) pertDpepi 128

6.1 Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, sont représentées les deux étapes d'évaluation menées dans le Chapitre 6. 136
6.2 Méthodologie mise en place pour dénombrer les bâtiments situés dans une emprise d'inondation à partir de la couche bâtiments de la BD TOPO 139
6.3 Comparaison des emprises inondées au niveau de Carcassonne pour les périodes de retour 2 ans (en rouge) et 5 ans (en bleu) 140
6.4 Exemple de quatre courbes d'impacts obtenues pour différents biefs de cours d'eau a) Carcassonne, b) Amont de l'Orbiel, c) Aude aval et d) Amont du Fresquel 141
6.5 Emprise inondée relevée à la suite des crues d'octobre 2018 dans l'Aude . . . 142
6.6 Comparaison du nombre de bâtiments inondés avec a) l'emprise de référence observée de la crue d'octobre 2018 dans l'Aude et b) l'emprise estimée à partir de la chaine de simulation pluie-débit-impacts 143
6.7 Comparaison du nombre de bâtiments au sein des emprises d'inondation entre l'emprise de référence et l'emprise obtenue à partir de l'interpolation débit/bâtiments 143
6.8 Courbes ROC obtenues pour les trois produits de prévision d'ensemble pour les quatre seuils d'impacts considérés : a) 10, b) 20, c) 50 et d) 100 bâtiments inondés. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur obtenue pour la prévision RF0 146
6.9 Cartes d'anticipation (0-6 h) du seuil de 10 bâtiments inondés : a) pour le scénario RF0, b) pour le percentile 75% d'AROME-EPS, c) pour le percentile 75% de pepi et d) pour le percentile 75% de pertDpepi. Les succès sont comptabilisés pour une anticipation supérieure à 15 minutes 147

6.10 Cartes d'anticipation (0-6 h) du seuil de 20 bâtiments inondés : a) pour le scénario RF0, b) pour le percentile 75% d'AROME-EPS, c) pour le percentile 75% de pepi et d) pour le percentile 75% de pertDpepi. Les succès sont comptabilisés pour une anticipation supérieure à 15 minutes 148
6.11 Comparaison des délais d'anticipation du dépassement de seuil de 10 bâtiments pour les percentiles 75% et 95% 149
6.12 Périodes de retour des débits associées à différents seuils de bâtiments inondés 151
6.13 Histogramme de comparaison des deux évaluations sur les débits et les impacts (pour le seuil de 10 bâtiments) 152

7.1 Moyens d'intervention (en nombre de véhicules de type CCF) disponibles sur le département de l'Aude par centre de secours d'après le schéma départemental d'analyse et de couverture des risques du SDIS 11 159
7.2 Localisation des opérations de secours menées par le SDIS 11 et les renforts extra-départementaux lors de la crue d'octobre 2018. Les interventions sont représentées par les cercles noirs. 160
7.3 Principe de fonctionnement de la modélisation des interventions du SDIS . . 163
7.4 Localisation des interventions simulées pour la crue de 2018 dans l'Aude sur l'ensemble de la durée de l'événement pour l'ensemble pepi avec les paramètres de simulation suivants : 10 bâtiments pour le seuil d'intervention et 100 bâtiments/équipe pour la capacité d'intervention 165
7.5 Extrait d'un fichier log pour un rejeu avec l'ensemble pepi avec les paramètres suivants : 10 bâtiments pour le seuil d'intervention et 100 bâtiments pour la capacité d'intervention 166
7.6 Illustration du nombre de bâtiments effectivement secourus sur le tronçon 4038 pour l'ensemble pepi avec les paramètres suivants :10 bâtiments pour le seuil d'intervention et 25 bâtiments pour la capacité d'intervention 167
7.7 Illustration de la méthode de calcul de la fonction de coût sur un cas théorique 170
7.8 Évolution spatio-temporelle des équipes (nombre illimité) en intervention pour la simulation de référence RS (sans anticipation des impacts), pour la prévision de pluie future nulle, RF0 et pepi 174
7.9 Bâtiments effectivement secourus avec un nombre d'équipes illimité pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la simulation d'impacts à partir de la pluie observée (RS) 175
7.10 Bâtiments inondés, capacité d'intervention et bâtiments effectivement secourus pour le tronçon 716 pour a) la simulation de référence d'impacts basée sur les pluies Antilope J+1, RS, b) la prévision de pluie future nulle, RF0 et c) la prévision d'ensemble pepi 176
7.11 Fonctions de coût calculées sur l'ensemble des tronçons pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la simulation d'impacts à partir de la pluie observée (RS) 177

7.12 Fonctions de coût calculées pour le tronçon 697 pour a) AROME-EPS et b) pertDpepi 178
7.13 Nombre d'équipes mobilisées au cours de l'événement pour la modélisation en moyens illimités, pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la modélisation de référence à partir du nombre de bâtiments simulé RS (sans prévision) 179
7.14 Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 17 h à 22 h 182
7.15 Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 23 h à 04 h 183
7.16 Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 05 h à 10 h 184
7.17 Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 11 h à 13 h 185
7.18 Évolution du nombre de bâtiments réellement secourus avec un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments 186
7.19 Évolution de la mobilisation des équipes de secours pour un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments 188
7.20 Résultats de modélisation à partir d'AROME-EPS sur le tronçon 2067 avec a) un seuil d'intervention à 10 bâtiments et b) un seuil d'intervention à 50 bâtiments 189
7.21 Évolution du coût en bâtiments non secourus pour un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments 190
7.22 Évolution du nombre de bâtiments effectivement secourus pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments 192
7.23 Résultats de modélisation à partir d'AROME-EPS sur le tronçon 2967 avec a) une capacité d'intervention de 10 bâtiments et b) une capacité d'intervention de 150 bâtiments 193
7.24 Évolution de la mobilisation des équipes de secours pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments 195
7.25 Évolution du coût en bâtiments non secourus pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments 197
7.26 Évolution du nombre de bâtiments effectivement secourus pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe 198
7.27 Évolution du nombre d'équipes mobilisées pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe 199
7.28 Coût des bâtiments non secourus à temps pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe 200

NOTE :

Les sources des figures sont précisées dans chaque légende, sinon il s'agit d'une production personnelle.

Liste des tableaux

5.1 Cas possibles pour un tableau de contingence 102

6.1 Nombre de tronçons considérés pour chaque valeur seuil de bâtiments 145
6.2 Aires sous les courbes ROC (AUC) calculées à partir des résultats de prévision d'ensemble des débits et des impacts 150
6.3 Nombre de tronçons sans évolution avec l'utilisation des débits ou des impacts 153

7.1 Liste des différentes valeurs de variables testées 169
7.2 Liste des différentes valeurs de variables testées 171
7.3 Valeurs retenues pour les différents paramètres de simulation 181

Résumé

Les crues soudaines (ou crues-éclair) sont responsables chaque année de décès et d'importants dégâts matériels. En France, ce type d'événement concerne principalement l'arc méditerranéen et ses petits cours d'eau. Depuis plusieurs années, d'importantes évolutions sur les différents modèles et méthodes pour la prévision immédiate des crues-éclair ont été réalisées.

Dans ce travail, trois produits météorologiques de prévisions d'ensemble ont été utilisés comme données d'entrée pour une modélisation hydrologique à partir du modèle Cinecar et évalués à partir d'un cadre méthodologique pensé pour l'évaluation événementielle. Cette méthodologie a été appliquée sur trois cas de crues soudaines survenues en 2018 dans l'Aude, l'Ardèche et le Var et repose sur une adaptation de la méthodologie de construction des courbes ROC. Cette approche permet en particulier de représenter spatialement les anticipations de dépassement de seuil de débits.

À partir des simulations de débits et d'un catalogue de scénario d'inondation, des prévisions d'ensemble d'impacts ont été réalisées et à leur tour évaluées avec le même cadre méthodologique basé sur les courbes ROC. Les prévisions d'impacts ont également été évaluées du point de vue de leur capacité à déclencher des prises de décisions pertinentes et efficaces via un rejeu de l'événement de l'Aude 2018 du point de vue du SDIS.

L'ensemble de ce travail a permis de mettre en évidence les performances de ces trois ensembles pour la prévision des crues-éclair, mais aussi de proposer un cadre méthodologique intéressant à exploiter dans le cadre d'évaluations événementielles, ce cadre restant encore perfectible. Enfin, le passage à des simulations d'impacts a montré l'intérêt de l'utilisation de ce genre de données pour l'anticipation et la gestion de crise de ce type d'événement.

Mots clés : *Crue rapide, prévision d'ensemble, impacts*

Abstract

Flash floods are responsible each year for deaths and important material damages. In France, these events mainly concern the Mediterranean arc and its small rivers. Since several years, important evolutions on the models and methods for the immediate forecasting of flash floods have been realized.

In this work, three meteorological products of ensemble forecasts have been used as input data for a hydrological simulation based on the Cinecar model and evaluated from a methodological framework designed for the event evaluation. This methodology was applied to three flash floods that occurred in 2018 in the Aude, Ardèche and Var. The methodology is based on an adaptation of the ROC curve construction. Specifically, this approach is designed for spatial representation of the thresholds exceedances anticipations.

From the discharge simulations and a catalog of flooding scenarios, ensemble impact forecasts were made and evaluated with the same methodological framework based on the ROC curves. The impact forecasts were also evaluated from an emergency service point of view, with a specific attention to the ability of the impact forecasts to be relevant and decision-making.

This work allowed to highlight the performance of these three ensembles for flash flood forecasting, but also to propose an interesting methodological framework which can be useful for event-based evaluations, this framework still remaining perfectible. Finally, the transition to impact simulations has shown the interest of using this kind of data for the anticipation and crisis management of this type of event.

Keywords : *flash floods, ensemble forecasts, impacts*

À ma famille pour leurs encouragements, à mes amis musiciens de la Compagnie des Arts qui ont éclairé mes mercredis soirs.

À Brieuc pour son soutien sans faille.

Remerciements

Sans la collaboration de nombreux acteurs en amont de la thèse, les travaux présentés ici n'auraient pu aboutir.

Merci à l'ensemble des collaborateurs du projet ANR-PICS ayant produit données et modèles nécessaires à ce travail de thèse :

- Météo-France et CNRM, en particulier François Bouttier et Axelle Fleury, pour la mise à disposition des observations de pluie et la création des produits de prévision d'ensemble de pluie à courte échéance.
- Nabil Hocini et Pierre Nicolle (UGE) pour la mise en place d'un catalogue de scénarios d'inondation à partir de FLOODOS.

Mes remerciements vont également à Daniela Peredo (INRAE) en particulier, et plus généralement à tous les co-auteurs qui ont contribué aux résultats de l'article "A methodological framework for the evaluation of short-range flash-flood hydrometeorological forecasts at the event scale" ainsi qu'à sa rédaction.

Merci aussi au comité de suivi de la thèse, composé de Maria-Helena Ramos (INRAE), Pierre Javelle (INRAE) et François Bouttier (CNRM) pour leur regard pertinent et bienveillant sur l'avancement de mes travaux.

Je remercie les rapporteurs, Sandrine Anquetin (CNRS) et Freddy Vinet (Université Paul Valéry, Montpellier 3) pour leur lecture détaillée, ainsi que les autres membres du jury, Agnès Baltzer (Université de Nantes), François Bouttier (CNRM), Bruno Janet (SCHAPI) et Isabella Zin (Université Grenoble Alpes), pour leurs intérêt et questions lors de la soutenance.

Enfin, un grand merci à Oliver Payrastre et Eric Gaume pour avoir encadré et dirigé les travaux pendant ces trois dernières années.

Chapitre 1

Introduction Générale

1.1 Le risque inondation, un risque majeur

L'année 2022 est riche en événements climatiques intenses, avec notamment des épisodes de précipitations et d'inondations spectaculaires. Les inondations les plus marquantes sont certainement celles du Pakistan au mois de juin, qui ont provoqué près de 1 800 décès, presque 13 000 blessés (ifrc.org) et près de 40 milliards de dollars de dégâts (reuters.org). Le cumul des pluies de la mousson annuelle, combiné à la fonte des glaciers suite à une forte vague de chaleur, ont été à l'origine de ces inondations. Il est également possible de citer les inondations dans la région de Sydney (Australie) survenues au mois de juillet, avec de moindres conséquences humaines (un décès), mais de nombreux dégâts matériels et des conséquences économiques remarquables. L'Europe n'est pas en reste pour les épisodes de précipitations intenses avec près de 20 événements recensés dans 14 pays. Il faut noter que ces événements européens très récents restent mineurs et peu meurtriers en comparaison avec les événements cités précédemment. Néanmoins, des inondations intenses et meurtrières se produisent également en Europe, comme l'ont montré les inondations de juillet 2021 en Belgique et en Allemagne (Météo-France). Cet épisode est responsable de 196 décès dont 189 en Allemagne (Thieken *et al.* 2022).

Les inondations sont un phénomène naturel touchant la totalité du globe et constituent le type de catastrophe naturelle le plus fréquent. Depuis 2011, les inondations représentent plus d'un tiers des catastrophes naturelles (Bevere et Remondi 2022). En plus d'être le phénomène le plus fréquent, c'est également le plus meurtrier au niveau mondial, avec en moyenne 50 000 décès par an (Hu *et al.* 2018).

La France est aussi régulièrement touchée par ces phénomènes, les inondations étant le phénomène naturel le plus fréquent sur le territoire comme le montre la Figure 1.1. En moyenne, les inondations des quarante dernières années ont coûté près de 540 millions d'euros par an en dommages assurés (CCR 2018b).

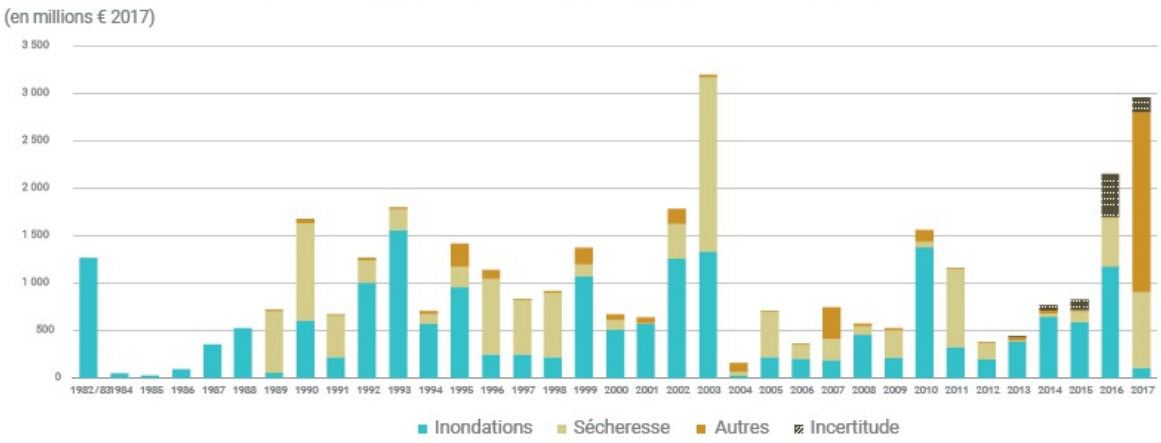

FIGURE 1.1 – *Sinistralité liée aux catastrophes naturelles (hors sinistralité automobile) de 1882 à 2017 en France (CCR 2018b)*

Les inondations peuvent être causées par différents événements comme :

- Le débordement temporaire d'un cours d'eau. La crue de la Seine survenue en mai 2016 (Perrin *et al.* 2017), considérée comme une crue lente, et les crues soudaines de la Tinée, de la Vésubie et de la Roya en octobre 2020 (Melun *et al.* 2022) en sont des exemples. Il s'agit de la cause la plus fréquente d'inondations.
- Le ruissellement d'origine pluviale.
- La rupture d'ouvrages comme des digues ou barrages. À titre d'exemple, il est possible de citer la rupture du barrage de Malpasset en 1959 (Rollet *et al.* 2012).
- Un tsunami ou une submersion marine touchant des régions côtières comme lors de la tempête Xynthia en 2010 (Bersani *et al.* 2010).
- La remontée du niveau d'une nappe comme les crues de la Somme en 2001 (Sénat 2002).

1.2 Les crues soudaines, un phénomène fréquent en France

1.2.1 Typologie des crues et définition des crues soudaines

Différentes causes peuvent être à l'origine de crues majeures. Il peut s'agir de crues survenues après des précipitations exceptionnelles par leur intensité et/ou leur cumul (Figure 1.2 a)). Il peut également s'agir de crues de fonte nivale, éventuellement accompagnée de précipitations (Figure 1.2 b)). Enfin, plus rarement, il est possible d'observer des crues d'embâcles et de débâcles, liées à une obstruction temporaire de l'écoulement dans le lit mineur du cours d'eau (embâcle) souvent en amont d'un pont et à la rupture (débâcle), provoquant une augmentation très rapide du niveau d'eau. Les embâcles peuvent être le fait d'accumulation de blocs de glace, de troncs d'arbres et débris végétaux, de voitures et autres objets emportés par les flots (Figure 1.2 c)).

Les crues se différencient également par leur dynamique temporelle. Selon le forçage pluviométrique et les caractéristiques du bassin versant touché, il est possible de distinguer les crues dites « lentes » et les crues rapides (aussi appelées crues-éclair). Les crues lentes affectent de grands cours d'eau drainant des bassins versants de plusieurs milliers de km^2 comme la Loire ou la Seine. Ces crues surviennent généralement après plusieurs jours de pluie continue ou encore du fait de la fonte d'un manteau neigeux (Figure 1.2 b)). Les temps de réponse élevés des bassins versants et de propagation des écoulements dans le réseau hydrographique facilitent l'anticipation. Les observations disponibles à l'amont du bassin versant permettent de renseigner en avance sur l'intensité possible de la crue qui surviendra à l'aval. La durée de plusieurs jours, voire de plusieurs semaines pour des crues de nappes, peut s'avérer source de dégâts matériels et d'impacts sur les enjeux exposés du fait de la stagnation prolongée des eaux.

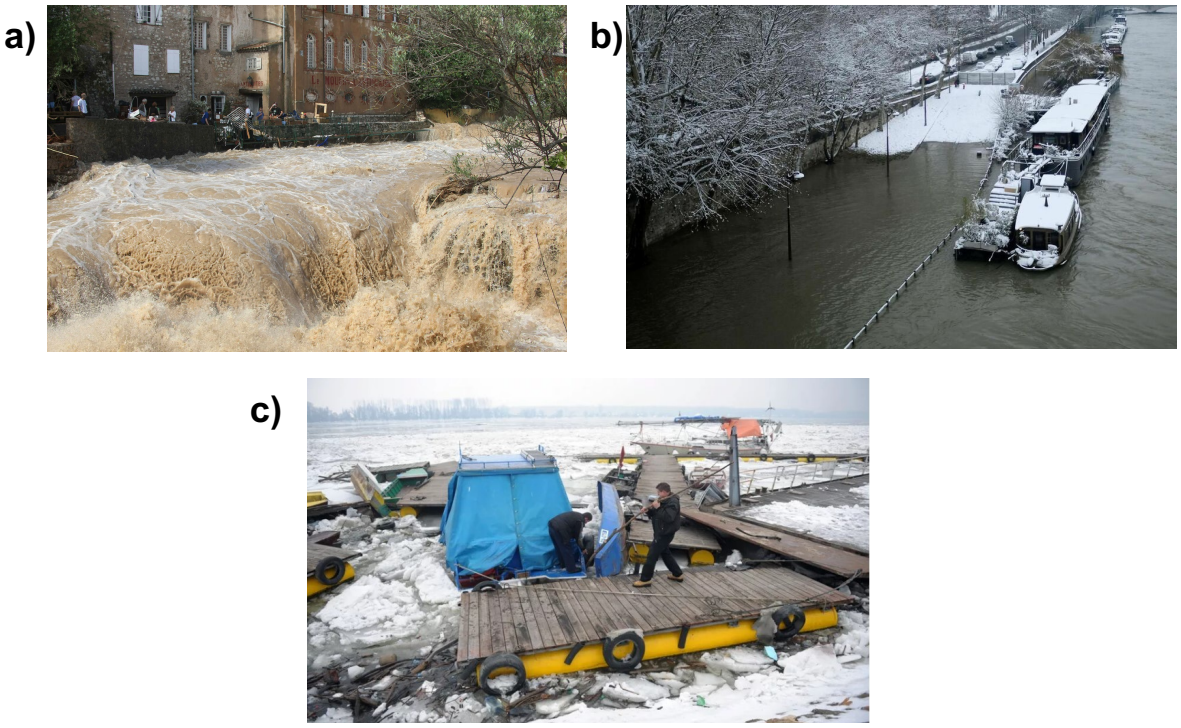

FIGURE 1.2 – *Illustration des différentes causes de crues : a) Crue à Trans-en-Provence (Var) en juin 2010 après d'intenses précipitations (source : P Cantet, INRAE-Aix), b) Crue nivale sur la Seine après plusieurs jours de précipitations de pluie et de neige en février 2018 (source : AFP/Stéphane De Sakutin) et c) Débâcles de glace sur le Danube près de Belgrade après plusieurs semaines de grand froid (source : AFP/ Alexa Stankovic)*

Les crues rapides sont quant à elles souvent provoquées par des pluies intenses et convectives, d'extension spatiale souvent limitée et entraînant parfois du ruissellement, en particulier sur les bassins urbanisés. Plus rarement, ces crues peuvent survenir après des ruptures d'ouvrages comme des digues ou des barrages. Il s'agit de « crues dont l'apparition est soudaine, souvent difficilement prévisible, de temps de montée rapide et de débit spécifique relativement important. Ces crues sont donc généralement liées à des épisodes pluvieux intenses et se manifestent souvent sur des bassins de taille modérée » d'après la définition donnée par l'Association Internationale des Sciences Hydrologiques (IAHS 1974). Les bassins versants concernés par ces crues sont donc de plus petite taille (de quelques km^2 à quelques centaines, voire milliers de km^2) et ont un temps de réponse limité, tandis que les débits unitaires ($m^3/s/km^2$) s'avèrent plus importants que pour les crues lentes. Lorsque le temps de réponse (délai entre le maximum d'intensité des pluies et le pic de crue) est inférieur à 6 h, la crue est généralement appelée crue-éclair.

1.2.2 Les crues soudaines observées dans l'arc méditerranéen

L'arc méditerranéen français est régulièrement touché durant l'automne par des événements de fortes précipitations, pouvant conduire à des crues soudaines. On parle alors d'épisodes méditerranéens ou encore d'événements cévenols lorsqu'ils concernent les reliefs des Cévennes.

Lors de ces événements, l'équivalent de plusieurs mois de précipitations peut tomber en l'espace de quelques heures ou quelques jours. Ces épisodes sont liés à des remontées d'air chaud, humide et instable en provenance de la mer Méditerranée, qui peuvent générer de violents systèmes orageux (parfois stationnaires) apportant des précipitations intenses. Ces orages présentent une saisonnalité, avec une fréquence d'apparition plus forte à l'automne, quand la Mer Méditerranée est plus chaude et génère donc une plus forte évaporation.

Deux phénomènes météorologiques sont générateurs d'épisodes de pluies intenses sur les régions méditerranéennes :

- **Le forçage orographique :** Ce premier cas correspond à une modification du régime de précipitations liée au relief. Le relief fait office de barrière naturelle aux masses d'air chaud et humide, en provenance de la Méditerranée. Cet air, poussé par les vents de basse altitude, se soulève alors le long des reliefs. En prenant de l'altitude, la masse d'air se refroidit, la vapeur d'eau qu'elle contient se condense alors et l'eau retombe sous la forme de précipitations. Le rôle de relief est joué par les Cévennes, mais également les Pyrénées, la Montagne Noire, les Alpes ou encore le relief corse.

- **Le forçage convectif :** Dans ce cas, des cellules orageuses (ou convectives) se forment du fait de l'instabilité des masses d'air, et génèrent de fortes précipitations. Elles couvrent de faibles étendues, mais engendrent des intensités de précipitations pouvant dépasser les 100 mm/h. Ces systèmes convectifs peuvent également être stationnaires et s'auto-alimenter, en créant une goutte d'air froid proche du sol, qui agit comme le relief. Cette goutte froide peut entraîner le soulèvement de la masse d'air plus chaud et humide comme dans le cas d'un forçage orographique.

Des événements de pluie intenses de nature similaire surviennent à l'automne sur tout le pourtour méditerranéen, en Italie, dans le sud de la France, l'est de l'Espagne ainsi qu'au Maghreb.

Dans la liste des événements méditerranéens les plus violents des trente dernières années en France, il est possible de citer la crue des cadereaux de Nîmes (Fabre 1990) en octobre 1988, avec près de 420 mm de cumuls de pluie en moins de 12 heures, soit près de 6 mois de précipitations, les crues particulièrement dévastatrices dans l'Aude de novembre 1999 (Lefrou *et al.* 2000), les crues des Gardons (Huet *et al.* 2003) en septembre 2002 (ponctuellement 687 mm de pluie en moins de 36 heures) ou encore la tempête Alex dans les Alpes-Maritimes en octobre 2020 (Melun *et al.* 2022) avec 500 mm de pluie enregistrés en 24 heures à Saint-Martin-Vésubie. La Figure 1.3 présente des exemples de dommages survenus après ces différents événements.

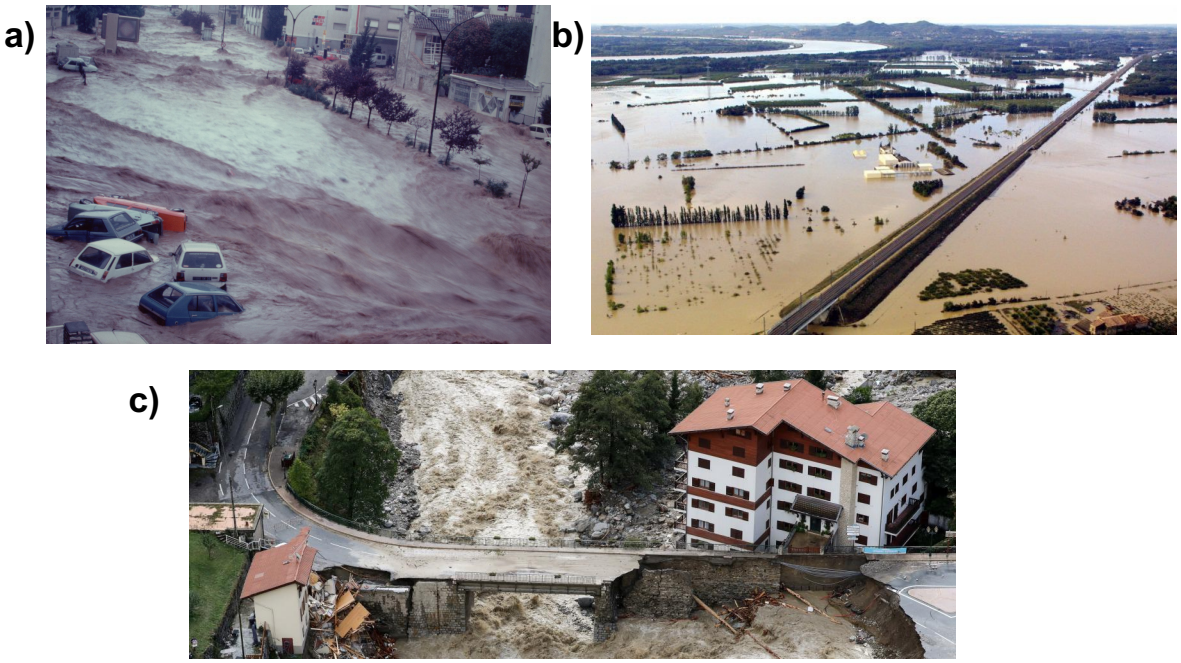

FIGURE 1.3 – *Dommages suite à des crues soudaines : a) Crue des cadereaux à Nîmes en octobre 1988 (source : Pluies extrêmes), b) Gard, septembre 2002 (source : France 3 Région Occitanie) et c) Dégâts après le passage de la tempête Alex, sur le Boréon à Saint-Martin-Vésubie (source : France Info)*

En France, la moitié des dommages aux biens assurés liés aux inondations concernent des crues soudaines (CCR 2018b), et avec le changement climatique, ces dommages pourraient s'intensifier, même si les pluies moyennes annuelles dans le secteur méditerranéen sont plutôt à la baisse (Raymond *et al.* 2016). Une augmentation de la fréquence des pluies extrêmes est déjà observable en région méditerranéenne comme l'indique le graphique de la Figure 1.4.

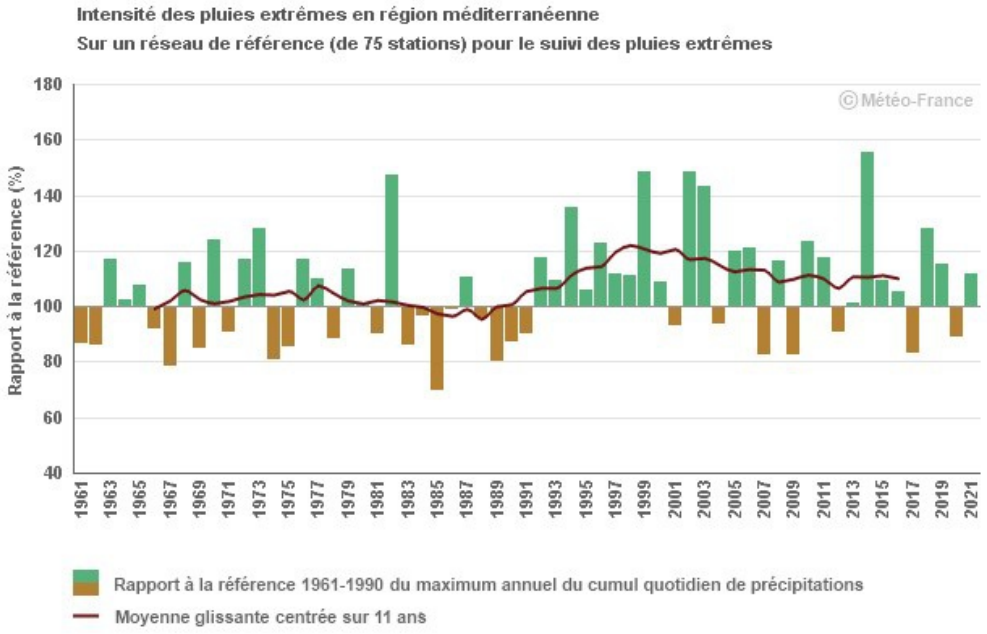

FIGURE 1.4 – *Intensification des fortes pluies sur les régions méditerranéennes (Soubeyroux 2021)*

Du fait de la rapidité du phénomène, les crues soudaines laissent peu de temps pour agir (vigilances et alertes auprès de la population...), il est donc particulièrement important d'améliorer les capacités d'anticipation de ces crues et de leurs conséquences.

1.3 Les systèmes actuels de prévision et d'alerte aux inondations

1.3.1 Un dispositif d'avertissement précoce sur les cours d'eau principaux bien établi

Différents projets nationaux et européens ont abouti à la mise en place de systèmes d'avertissements précoce destinés aux grands cours d'eau et fleuves. Ainsi, le système européen EFAS - European Flood Awareness System (Thielen *et al.* 2009), débuté en 2003, est conçu pour fournir des prévisions d'inondations à l'échelle européenne avec une échéance de 10 jours. EFAS constitue également le premier réseau de prévision hydrologique opérationnel en Europe. La Figure 1.5 présente le fonctionnement de ce système de prévision.

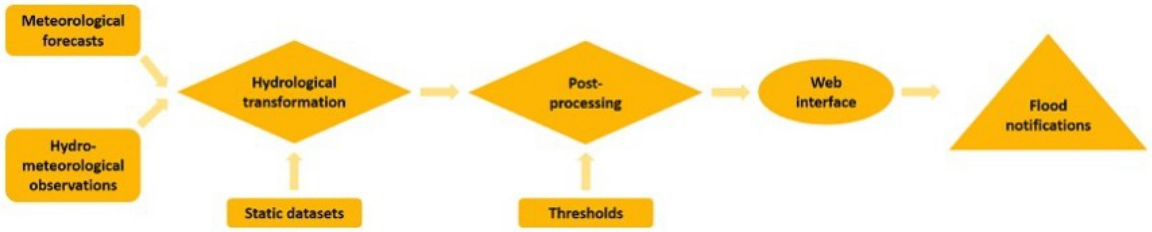

FIGURE 1.5 – *Chaîne de prévision hydrologique d'EFAS (source : efas.eu)*

EFAS comporte aussi un volet de prévision des impacts (Dottori *et al.* 2017) présentée sur la Figure 1.6.

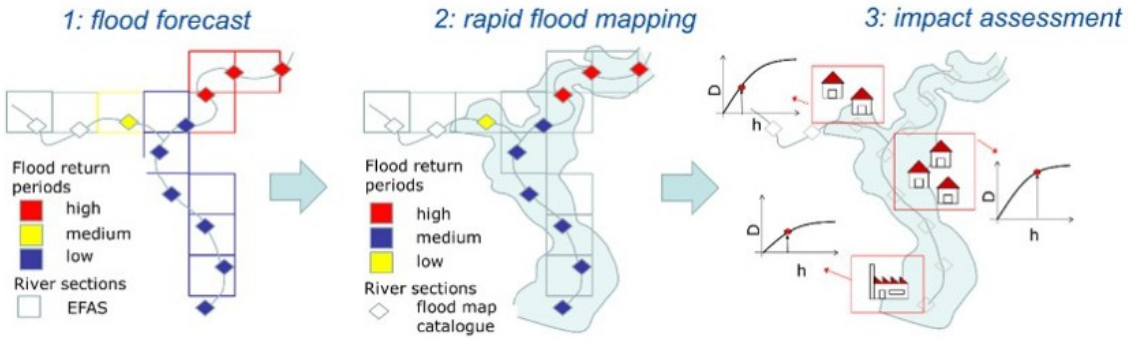

FIGURE 1.6 – *Chaîne de modélisation des impacts intégrée à EFAS (Dottori* et al. *2017)*

GloFAS (Global Flood Awareness System - Alfieri *et al.* (2013)) est un équivalent mondial d'EFAS, fournissant également des avertissements plusieurs jours à l'avance. La Figure 1.7 présente les échéances de prévision de GloFAS.

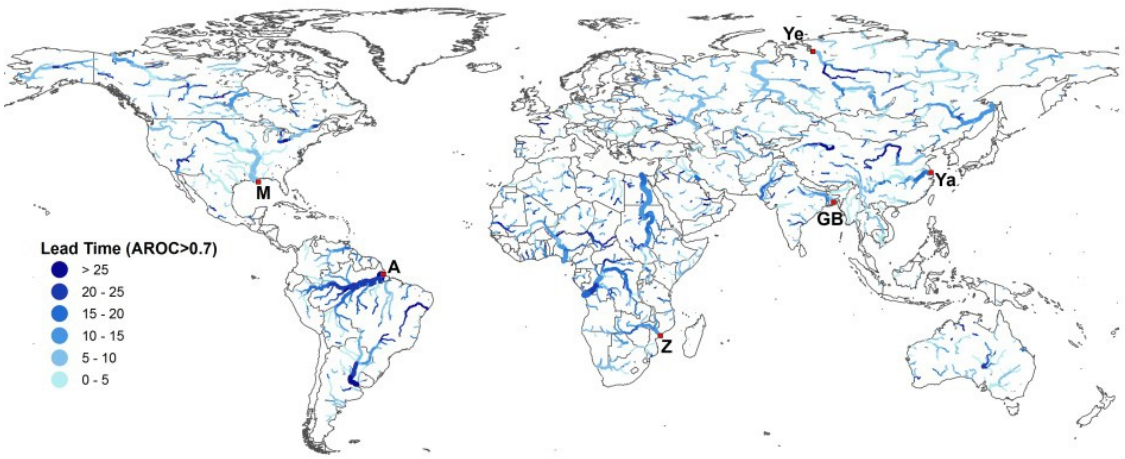

FIGURE 1.7 – *Couverture mondiale de GloFAS et échéances de prévision (en jours) (source : Alfieri* et al. *(2013))*

En France, après les crues catastrophiques de la Loire de 1846, 1856 et 1866, le Conseil Général des Ponts et Chaussées a installé les premières échelles de crue sur les grands fleuves. Le premier service d'hydrométrie de transmission de prévisions aux ingénieurs riverains est créé en 1854 sur la Seine. Pour la crue de 1876, ce service a permis de produire une prévision assez précise du niveau maximum de la crue à Paris à échéance 5 jours. Ce service d'alerte s'est progressivement développé sur les autres rivières principales françaises sous la forme de SAC, Service d'Annonce des Crues à partir de 1876 (Houdré 2001).

En France, aujourd'hui la plateforme Vigicrues, mise en service en 2006, publie deux fois par jour une carte de vigilance concernant les risques de crues sur actuellement 21 700 km de cours d'eau principaux, le linéaire de cours d'eau surveillés étant en constante évolution. La vigilance crue correspond à un avertissement précoce sur les risques de crues dommageables pour les 24 h à venir. Il existe 4 niveaux de vigilance, allant du vert (pas de risque particulier), au jaune, orange, puis rouge (risque de crue majeure). La vigilance est élaborée par les Services d'hydrométrie et de Prévision des Crues (SPC), à partir des observations météorologiques et limnimétriques sur les cours d'eau surveillés et parfois de résultats de modèles permettant d'effectuer des prévisions de débits/hauteurs à échéance de quelques heures à quelques jours suivant les stations et les cours d'eau. Vigicrues couvre des cours d'eau dont le temps de réaction est supérieur à 6 h, et il ne concerne donc pas les nombreux petits cours d'eau sur lesquels sont observées les crues soudaines (on estime à 120 000 km le linéaire de cours d'eau de plus de 1 m de large). Ce système a largement évolué au cours des dernières années avec, notamment, la production et la mise à disposition de scénarios d'inondations, permettant aux utilisateurs d'interpréter les hauteurs et débits prévus aux stations limnimétriques (zones d'inondation potentielle ou ZIP). Une évolution vers une vigilance "surfacique", couvrant les petits cours d'eau, est également à l'étude.

1.3.2 L'apparition de systèmes d'anticipation et d'avertissement pour les petits cours d'eau

La couverture des plus petits cours d'eau par des systèmes de prévision des crues est apparue plus récemment. Une majorité des systèmes existants est basée sur le Flash Flood Guidances System (FFGS) développé aux États-Unis, ayant pour but de fournir aux prévisionnistes des informations en temps réel sur le risque de crues-éclair à petite échelle. Depuis 2017, FFGS est déployé en Amérique Centrale, à Haïti, au Moyen-Orient ainsi qu'en Afrique Australe. Les prévisions de précipitations (ou les observations de précipitations) sont comparées à un volume de précipitation de référence nécessaire pour provoquer un début d'inondation (Carpenter et Georgakakos 2004, Georgakakos *et al.* 2021).

D'autres modèles dédiés aux crues soudaines ont été développés, toujours aux États-Unis, comme FLASH et FFMP (Gourley *et al.* 2017), proposant une surveillance continue des précipitations et une alerte automatique quand une situation dangereuse est en cours sur un cours d'eau ou un bassin versant donné.

L'Europe a également mis en place des systèmes dédiés aux crues soudaines : développement d'ERIC (European Runoff Index based on Climatology) et d'ERICHA (European Rainfall-InduCed Hazard Assessment). ERIC (Alfieri *et al.* 2013) est basé sur un indicateur, appelé indice de ruissellement extrême, qui est calculé à partir des données d'entrée (pluie observée) par une analyse statistique. Il est conçu pour être utilisé dans des domaines vastes ou mal jaugés, dans la mesure où aucune connaissance locale ou observation in situ n'est nécessaire pour sa mise en place. ERICHA est quant à lui implémenté dans EFAS pour la surveillance en temps réel des phénomènes de crues soudaines.

En France, pour compléter la couverture du système Vigicrues, les avertissements Vigicrues Flash ont été mis en place en mars 2017. Ce système cible des petits cours d'eau non instrumentés avec un temps de réponse supérieur à 1 h 30 (SCHAPI 2018). Il couvre environ 10 000 petits bassins versants et 30 000 km de cours d'eau non couverts par la vigilance crues. La Figure 1.8 présente le principe général de fonctionnement de Vigicrues Flash. Ce système déclenche un avertissement envoyé de manière automatique aux communes abonnées, si les seuils de débit simulés par un modèle pluie-débit sont dépassés. Néanmoins, Vigicrues Flash se base uniquement sur une observation des pluies par radar, ce qui limite l'anticipation, et il ne précise pas l'heure prévue de dépassement du seuil. De plus, la couverture spatiale de Vigicrues Flash reste limitée, et il n'y a pas de prise en compte dans les avertissements émis des secteurs potentiellement inondés (emprises des inondations) et des impacts associés, puisque l'avertissement se base uniquement sur l'intensité des crues (débit prévisible par rapport aux précipitations). Vigicrues Flash constitue donc une première réponse apportée pour l'anticipation des crues rapides, mais ce système peut encore être largement amélioré, notamment sur l'anticipation des phénomènes et la prise en compte des enjeux vulnérables.

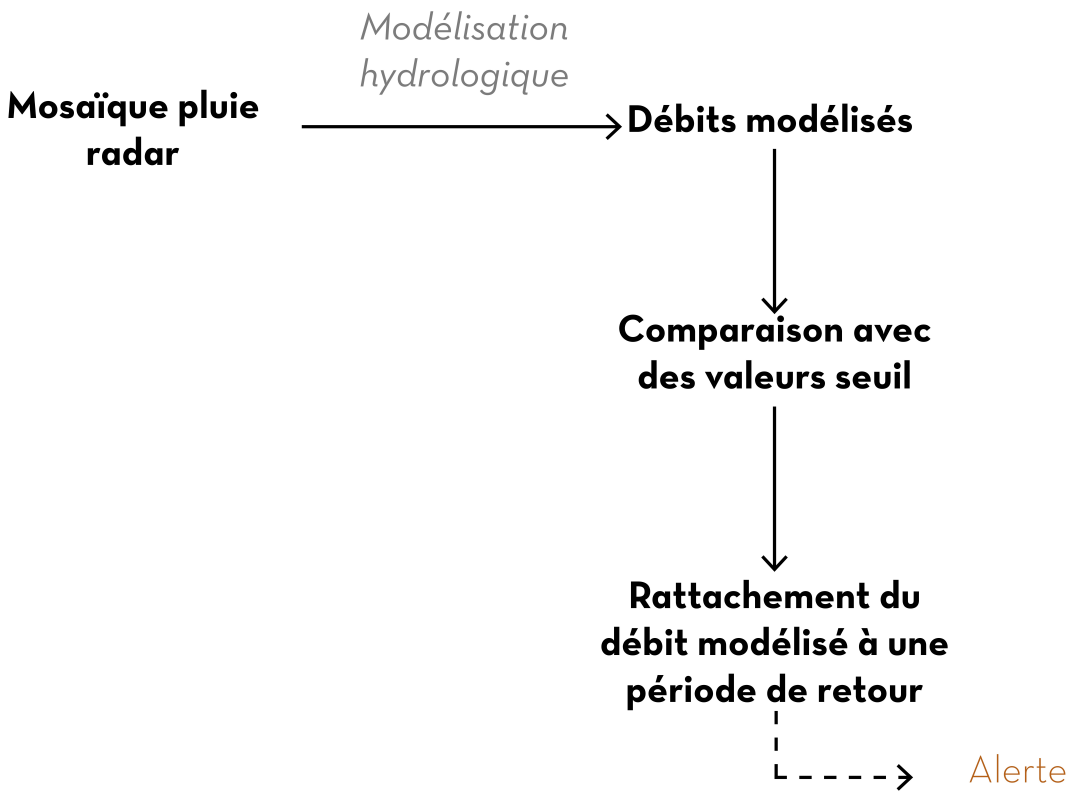

FIGURE 1.8 – *Principe de fonctionnement de Vigicrues Flash (inspiré de INRAE)*

Il est également possible de citer le système d'avertissement APIC (Avertissement Pluies Intenses à l'échelle des Communes), uniquement basé sur l'observation des pluies. Ce système, opérationnel dans 93% des communes du territoire métropolitain depuis fin 2011, a pour but d'anticiper les précipitations exceptionnellement intenses et pouvant présenter un danger potentiel lié à du ruissellement ou des crues de petits cours d'eau. Cette méthode repose sur le réseau de radars météorologiques ARAMIS et les pluviomètres en transmission temps réel avec une réactualisation toutes les 15 minutes. Lors de précipitations intenses dépassant une certaine période de retour (détermination avec la méthode AIGA-pluie), les communes abonnées reçoivent un SMS, message vocal et/ou mail d'avertissement "précipitations intenses" (précipitations supérieures à une période de retour 10 ans) ou "précipitations très intenses"(précipitations supérieures à une période de retour 50 ans).

1.4 Questions scientifiques soulevées et objectifs de la thèse

Le développement des systèmes de prévision particulièrement adaptés aux phénomènes de crues-éclair est un enjeu majeur du fait du grand nombre de bassins versants concernés, de leur petite taille, mais aussi de leur temps de réponse réduit (de l'ordre de quelques heures au plus), ainsi que de la faible connaissance actuelle des enjeux exposés à ces phénomènes. La probable augmentation de l'occurrence des crues soudaines avec le changement climatique laisse également penser que le développement des systèmes d'anticipation est nécessaire pour faire face aux événements futurs.

La thèse s'intègre au projet de recherche ANR-PICS (Prévision Immédiate Intégrée des Impacts des Crues Soudaines) qui a débuté en 2018. Ce projet visait à réaliser et évaluer des chaînes de prévision des crues soudaines permettant jusqu'à 6 h d'anticipation et une modélisation directe des impacts de l'événement. Les résultats devraient permettre d'améliorer le service Vigicrues Flash du SCHAPI, en apportant de nouveaux progrès liés à l'utilisation des prévisions immédiates d'ensemble des pluies, à l'amélioration de la prévision des débits et à l'évaluation des champs d'inondation et des impacts de crues. La Figure 1.9 présente les différentes étapes de ce projet.

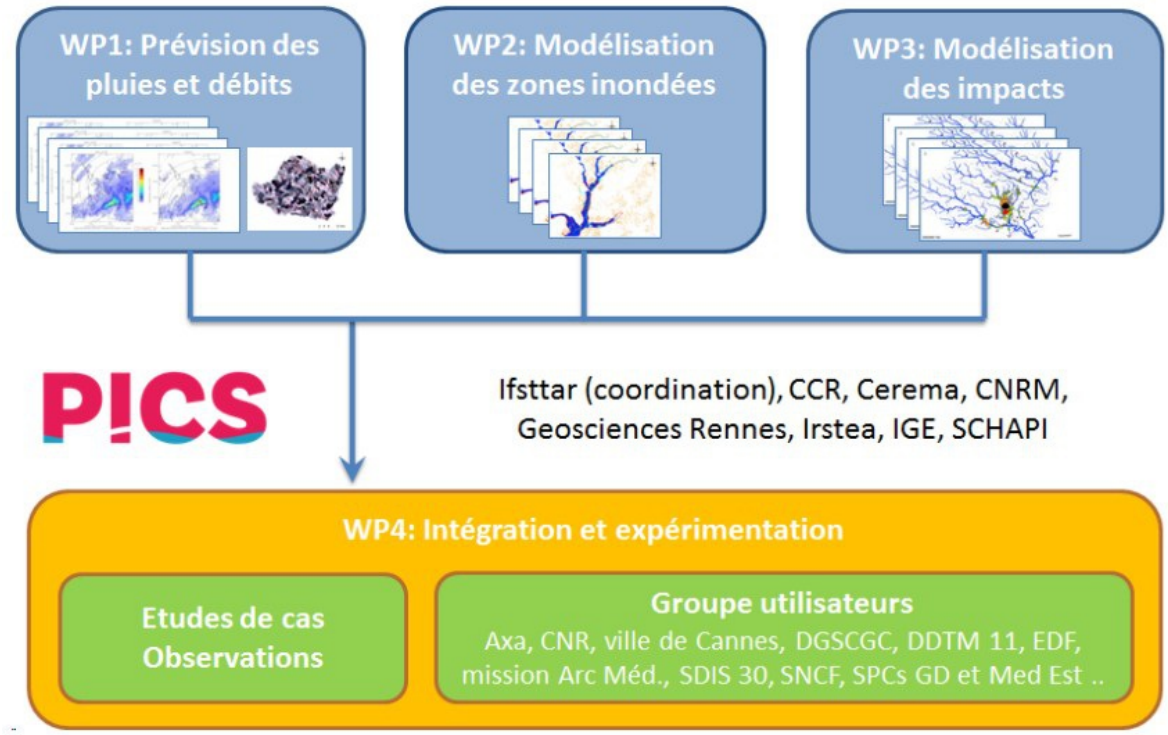

FIGURE 1.9 – *Organisation du projet PICS (source : PICS)*

Un des objectifs du projet PICS était de mieux évaluer les capacités d'anticipation des crues soudaines à partir des prévisions de pluie. Sur cette question, plusieurs travaux ont déjà illustré les limites importantes liées à la localisation imparfaite des pluies intenses par les modèles numériques de prévision météorologique. Cette incertitude de localisation des pluies doit impérativement être prise en compte et propagée, généralement sous forme de prévisions d'ensemble, pour obtenir des prévisions exploitables. De nouveaux produits sont en cours de développement à Météo-France et le projet PICS a été l'occasion de les utiliser comme supports à des prévisions immédiates d'ensemble des crues soudaines, permettant de propager les incertitudes sur les pluies prévues. Le deuxième objectif du projet PICS était de développer des outils de prévision des champs d'inondation et des impacts associés. Une information directe sur les champs d'inondation et les impacts s'avère en effet essentielle pour les utilisateurs des prévisions (gestionnaires de crise, de réseaux, sociétés d'assurances...). La mise en œuvre des chaînes de prévision des impacts doit également permettre de confronter les résultats de modélisation à des données d'impacts observés à l'échelle des territoires sinistrés. Les observations d'impacts (coupures des réseaux de transports, interventions des

SDIS, activité sur les réseaux sociaux, informations diffusées par les médias et dommages assurés) étant susceptibles d'être mobilisées sur chaque tronçon de cours d'eau, elles offrent des possibilités de validation beaucoup plus étendues que les enregistrements de débits aux quelques stations hydrométriques disponibles.

Cette thèse a pour but de contribuer aux travaux entrepris dans le projet PICS, en combinant des outils de prévision immédiate d'ensemble des pluies avec des modèles pluie-débit et des modèles de prévision des impacts. Cette chaîne de prévision sera mise en œuvre et évaluée sur les crues-éclair qui se sont produites en 2018 dans l'Aude (octobre), l'Ardèche (août) et le Var (octobre). Pour ces événements, de nouvelles approches de prévision immédiate d'ensemble des pluies ont été proposées par Météo France au cours du projet PICS. Ces prévisions seront utilisées comme données d'entrée de la chaîne de prévision évaluée.

Le principal objectif de la thèse sera d'illustrer la plus-value des prévisions immédiates des pluies et de l'évaluation des impacts pour la prévision des inondations. Pour atteindre cet objectif, l'enjeu principal consistera à mettre en place une méthodologie appropriée d'évaluation des prévisions d'ensemble des crues et de leurs impacts pour les trois événements majeurs de l'année 2018, en se focalisant dans un premier temps sur la capacité à anticiper les débits des crues, puis sur la capacité à anticiper les impacts, de façon à mieux renseigner les utilisateurs des prévisions. À plus long terme, la thèse s'inscrit dans une volonté de faire progresser l'anticipation des crues soudaines, anticipation qui est cruciale afin de limiter les conséquences de ces crues, qui sont souvent très dévastatrices.

1.5 Organisation de la thèse

Les prochains Chapitres de ce document présenteront les différents travaux menés au long de ces trois ans. Le Chapitre 2 retrace l'état de l'art concernant les évolutions des différents modèles et méthodes pour la prévision immédiate des crues-éclair. Dans le Chapitre 3, les données mobilisées ainsi que les études de cas retenues sont introduites. Le contenu des Chapitres 4 et 5 se concentre sur les résultats d'évaluation des prévisions d'ensemble de pluie (pour le Chapitre 4) et de débits obtenus par simulation hydrologique (pour le Chapitre 5). Le Chapitre 6 présente également une évaluation des prévisions d'ensemble, non plus sur des prévisions hydrologiques, mais sur des prévisions d'impacts obtenus via l'utilisation d'un catalogue de scénario d'inondation. Enfin, le Chapitre 7 propose une approche de rejeu d'un événement passé, les crues de l'Aude d'octobre 2018, du point de vue d'un service de secours pour lequel les interventions seraient basées sur les prévisions d'impacts simulées dans le Chapitre 6. Pour terminer, le Chapitre 8 reprend les principales conclusions et propose des perspectives à l'issue de ce travail de recherche.

Chapitre 2

La prévision immédiate des crues-éclair : de la prévision des pluies, à celle des débits, des zones inondées et des impacts associés

Objectifs du chapitre :

Ce Chapitre présente les différents modèles et méthodes mobilisés pour la prévision immédiate des crues-éclair. Dans un premier temps, les définitions de prévision immédiate et de chaîne de prévision sont présentées. Les différents types d'observations de pluie sont ensuite décrites. Les parties suivantes sont consacrées à la description des différentes méthodes de prévision des pluies et aux différentes modélisations pluie-débits. Une troisième section se concentre sur l'utilisation des prévisions d'impacts et leurs intérêts et complémentarités par rapport à des prévisions de pluie et de débit. Enfin, une dernière partie traite des problématiques de l'évaluation des prévisions d'ensemble.

2.1 Définition d'une prévision immédiate

En hydrométéorologie, une prévision est définie comme l'estimation de l'évolution de la valeur d'une variable sur une période délimitée et à venir. L'intervalle de temps entre l'émission d'une prévision et l'instant futur considéré est appelé l'échéance de prévision. Pour l'Organisation Météorologique Mondiale, une prévision est dite immédiate pour une échéance de quelques minutes à deux ou trois heures (WMO). Une prévision allant jusqu'à une échéance de six heures est généralement appelée prévision à très courte échéance. Néanmoins, le terme prévision immédiate est très régulièrement utilisé pour des prévisions allant jusqu'à six heures d'échéance, cette dénomination englobe donc le terme de prévision à très courte échéance.

Une prévision immédiate repose généralement sur les dernières observations météorologiques instantanées, et nécessite donc d'être rafraîchie fréquemment pour tenir compte des évolutions de ces observations. Les temps d'acquisition et de mise à disposition de ces observations se doivent d'être les plus réduits possibles afin de garantir que les prévisions puissent être émises avant l'échéance de prévision.

2.2 Principe de fonctionnement d'une chaîne de prévision des crues

Un système de prévision des crues est généralement composé d'un ensemble d'observations, de décisions et d'opérations successives formant une chaîne, dont les sorties permettent aux utilisateurs de prévoir des variables caractéristiques de la crue (débits, hauteurs), et de prendre une décision sur l'alerte et les secours en fonction des risques identifiés sur un territoire donné. La Figure 2.1 est un exemple de chaîne théorique de prévision des crues.

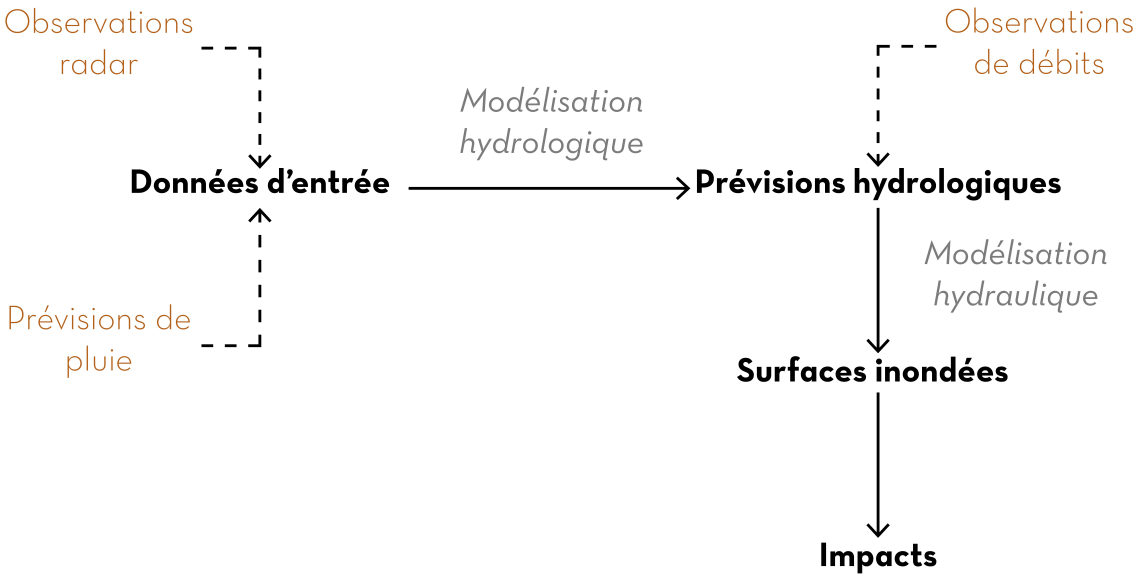

FIGURE 2.1 – *Schéma général de fonctionnement d'une chaîne de prévision des crues, depuis les prévisions météorologiques jusqu'aux impacts potentiels des précipitations et des crues*

En entrée, se trouvent généralement des informations sur les précipitations passées et/ou à venir. Celles-ci peuvent être constituées d'observations des pluies par un réseau de pluviomètres, des radars ou des satellites, ou bien de prévisions obtenues par modélisation numérique. Les méthodes et modèles disponibles pour obtenir ces données d'entrée seront détaillés dans la Section 2.4.

Les données de pluie servent le plus souvent d'entrée à une modélisation hydrologique. L'utilisation d'un modèle hydrologique peut se faire à l'échelle locale d'un bassin versant, ou plus largement à l'échelle d'un pays ou d'un continent, en intégrant les différents bassins versants présents sur le territoire modélisé. La Section 2.6 traitera des différents modèles hydrologiques de recherche et opérationnels, en particulier ceux utilisés pour la prévision immédiate des crues.

La plupart des chaînes de prévision s'arrêtent à l'étape de modélisation hydrologique et se basent sur le dépassement de seuil d'un débit de référence pour déclencher une alerte ou une recommandation de vigilance. Néanmoins, l'identification des zones inondées peut également s'avérer utile pour estimer les impacts potentiels des crues prévues. La Section 2.7 s'intéressera plus précisément aux approches de modélisation hydrauliques adaptées aux crues soudaines pour permettre l'identification des zones susceptibles d'être inondées, ainsi qu'au développement des modèles d'impacts de ces crues.

2.3 Les différentes observations de précipitations mobilisables

Historiquement, les observations de précipitations se sont développées à partir de réseaux pluviographiques, installés au sol, et permettant de mesurer les précipitations ponctuellement à pas de temps fin. Ces mesures comportent trois sources d'incertitudes principales, liées d'une part à l'appareil lui-même, à la localisation et à l'environnement dans lequel se trouve le pluviomètre et à la nécessité de recourir à l'interpolation spatiale des mesures ponctuelles pour reconstituer des lames d'eaux spatiales. Depuis une quarantaine d'années, les mesures pluviographiques sont automatisées et permettent une mesure au pas de temps minimum de 6 minutes, voire 1 minute pour les équipements du projet RADOME (Météo-France). La Figure 2.2 présente la localisation des différentes stations pluviographiques (au nombre de 1738 en 2022) présentes sur le territoire métropolitain français.

Les réseaux pluviographiques présentent néanmoins l'inconvénient de capturer de façon imparfaite les phénomènes de pluies intenses et très localisées, qui sont à l'origine des crues-éclair. Les méthodes d'interpolation spatiale, longtemps utilisées pour reconstituer des lames d'eau spatiales, ne compensaient pas ce manque d'observations. La prévision des crues-éclair ne s'est véritablement développée que depuis que des champs de pluie spatialement continus peuvent être véritablement observés. Ces champs de pluie sont le plus souvent obtenus à partir des enregistrements de radar météorologiques : on parle généralement de lames d'eau radar.

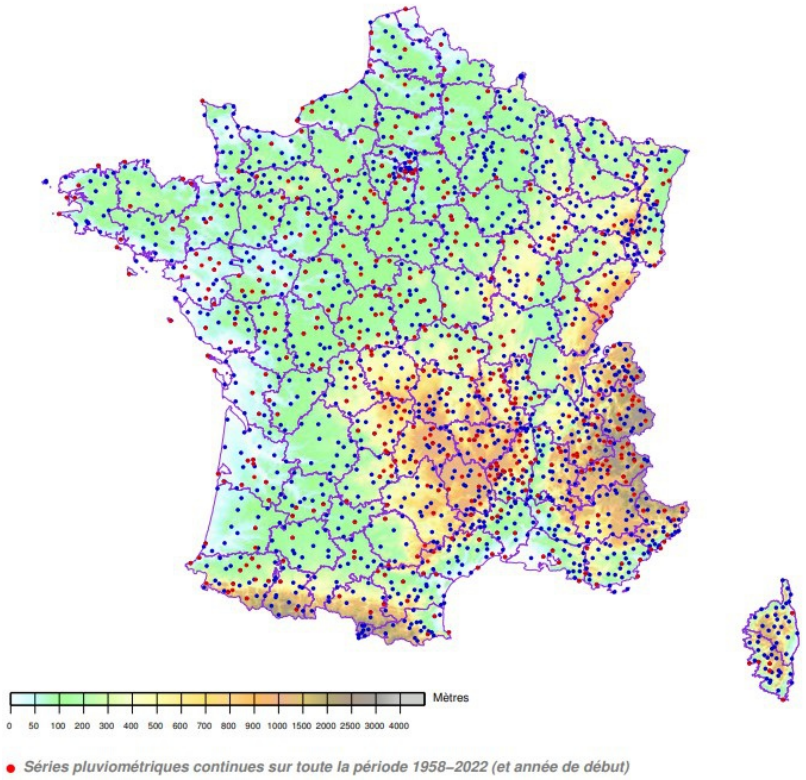

FIGURE 2.2 – *Localisation des différentes stations pluviométriques représentées par les points rouges (série pluviométrique complète de 1958 à 2022) et bleus (série pluviométrique incomplète) (source : Pluies Extrêmes)*

L'utilisation des observations provenant des radars météorologiques est désormais très commune pour les services opérationnels dans les pays développés. Une lame d'eau radar correspond à une estimation spatialisée des cumuls de pluie basée sur les observations de réflectivité d'un radar météorologique. Les ondes émises par le radar, lorsqu'elles rencontrent une cellule pluvieuse, sont rétro diffusées, le signal renvoyé est appelé réflectivité est corrélée à l'intensité des précipitations. Il est ainsi possible d'estimer les cumuls ou intensités pluviométriques observés pour chaque pixel de la lame d'eau radar, dont la résolution est généralement de 1 x 1 km.

Météo-France fournit actuellement plusieurs produits "lames d'eau radar", à partir des observations du réseau de radars ARAMIS couvrant le territoire métropolitain français (Champeaux *et al.* 2009) :

- La lame d'eau PANTHERE (Projet ARAMIS Nouvelles Technologies en Hydrométéorologie Extension et Renouvellement) est une lame d'eau calculée au pas de temps 5 minutes. PANTHERE amène une meilleure localisation des forts cumuls (Tabary 2007, Tabary *et al.* 2007; 2013). Les biais éventuels de la lame d'eau PANTHERE sont corrigés depuis 2007 par calage régulier des estimations radars à partir des enregistrements pluviographiques disponibles pour les pas de temps précédents. La lame d'eau PANTHERE est disponible en temps quasi réel.

- La lame d'eau ANTILOPE (Champeaux *et al.* 2009), opérationnelle depuis 2009, est une lame d'eau calculée aux pas de temps 15 minutes et 1 heure. Cette lame est construite par fusion de la lame d'eau PANTHERE et d'une lame d'eau issue du krigeage - interpolation spatiale - des stations pluviographiques disponibles lors du calcul de la lame d'eau. Cette approche permet notamment une estimation plus fiable des pluies intenses convectives. Le délai de mise à disposition de la lame d'eau ANTILOPE est de 10 minutes. Le nombre de d'enregistrements pluviographiques disponibles en temps réel étant assez limité, il existe une deuxième lame d'eau appelée ANTILOPE 'J+1', produite à partir de l'ensemble des données pluviométriques disponibles le lendemain et qui permet d'inclure 40% de pluviomètres supplémentaires.

Au niveau européen, il est possible de citer la lame d'eau radar OPERA, obtenue à partir de l'aggrégation des lames d'eau des réseaux de radars météorologiques nationaux (Saltikoff *et al.* 2019). Les données OPERA sont des observations à haute résolution (grille carrée de 2 km) avec un rafraîchissement fréquent toutes les 15 minutes.

2.4 Les différentes méthodes de prévision déterministe des précipitations

Il existe plusieurs méthodes permettant d'effectuer des prévisions de précipitations :

- Les modèles de prévision numérique du temps (Numerical Weather Prediction, NWP), qui sont généralement utilisées pour effectuer des prévisions des échéances courtes à moyennes (de quelques heures à quelques jours). L'assimilation des données observées dans les modèles de prévisions numériques permet d'augmenter leurs performances.
- L'extrapolation temporelle des observations, en particulier de données radar, qui permet d'effectuer des prévisions pour des échéances de quelques dizaines de minutes.

Les deux approches de prévision numérique et d'extrapolation des observations peuvent être combinées. Ces méthodes permettent d'obtenir des prévisions continues depuis des échéances très courtes (produites par extrapolation) jusqu'à des échéances moyennes (issues des modèles de prévision numérique). C'est l'objet du produit PIAF (Prévision Immédiate Agrégée Fusionnée - Météo-France). Cette possibilité ne sera pas traitée plus en détail dans les Sections suivantes.

2.4.1 La prévision météorologique numérique

La prévision numérique (Numerical Weather Prediction, NWP) s'est imposée comme un moyen de prévision fiable du temps à venir et deviendra certainement une approche privilégiée (Price *et al.* 2018). Une multitude de modèles existe avec différentes résolutions spatiales et temporelles, selon les variables à prévoir, le type de phénomène ciblé, le domaine spatial couvert et les échéances visées.

Des modèles mondiaux, basés sur l'hypothèse hydrostatique

Avec l'hypothèse hydrostatique, l'accélération verticale n'est pas résolue explicitement. La vitesse verticale est alors calculée avec l'équation de conservation de la masse. Cette approximation reste pertinente tant que les vitesses ascensionnelles des masses d'air restent limitées ou pour représenter les phénomènes atmosphériques à large échelle : échelle synoptique.

Il existe ainsi des modèles de prévision mondiaux, avec une grille de calcul recouvrant la surface de la Terre, comme le modèle GFS (Global Forecast System - NCEI NOAA) du National Weather Service (National Centers for Environmental Prediction (U.S.) 2004). Les prévisions du modèle GFS s'étendent sur une période de 16 jours et le modèle est initialisé 4 fois par jour (5 h 30, 11 H 30, 17 h 30 et 23 h 30 UTC). Le calcul s'effectue en deux phases avec une grille de résolution 30 km jusqu'à l'échéance 192 heures puis avec une résolution plus faible de 70 km pour les échéances plus lointaines. ARPEGE (Action de Recherche Petite Échelle Grande Échelle - Pailleux *et al.*) est aussi un modèle de prévision numérique planétaire. Il est utilisé en opérationnel par Météo-France (en collaboration avec l'ECMWF - Centre européen pour les prévisions météorologiques à moyen terme) depuis plus de 30 ans (Courtier *et al.* 1991, Déqué *et al.* 1994).

Comme pour le modèle GFS, ARPEGE est initialisé 4 fois par jour (00 h, 06 h, 12 h et 18 h UTC) et a une échéance maximale de prévision de 102 heures. La résolution horizontale du modèle ARPEGE est variable. Elle est de 5 km environ sur la France et atteint 24 km aux antipodes, comme le montre la Figure 2.3.

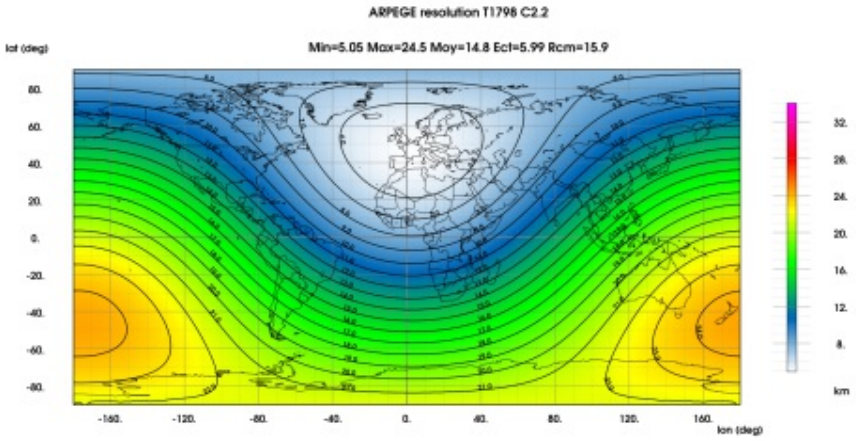

FIGURE 2.3 – *Résolution horizontale en kilomètres du modèle ARPEGE : minimum 5 km, moyenne 11 km et maximum 24 km (CNRM)*

Le modèle ALADIN (Pailleux *et al.* 2011, Coppola *et al.* 2021) est une version d'ARPEGE, appliquée au seul territoire européen. Le principal intérêt de cette version régionale est de limiter le coût de calcul des différentes simulations.

Ces modèles de prévision à l'échelle mondiale ou européenne offrent la possibilité de prévoir les conditions météorologiques sur des zones très étendues et sur une période de temps assez longue. Néanmoins, la précision diminue assez fortement avec l'échéance.

Développement des modèles non-hydrostatiques à haute résolution (convection-permitting models)

Pour des représentations plus fines de l'atmosphère et des épisodes pluvieux convectifs, il devient important de décrire plus explicitement les processus non-hydrostatiques. On parle alors de modèles non-hydrostatiques ou encore à haute résolution (convection-permitting models). Ces modèles ont des mailles de calcul plus petites et sont conçus pour des territoires plus restreints. Différents modèles ont été développés, comme le modèle AROME-FR ou le modèle COSMO.

COSMO (Baldauf 2013) est un modèle opérationnel qui possède différentes configurations :

- COSMO-EU couvrant la majorité de l'Europe avec une résolution horizontale de 7 km (proche de celle proposée par ARPEGE).
- COSMO-DE est imbriqué dans COSMO-EU avec une maille d'environ 2,8 km. D'autres versions de COSMO avec des résolutions horizontales similaires sont également utilisées dans d'autres pays européens (Grèce, Italie, Pologne, Russie, Roumanie, Suisse).

AROME-FR (Application of Research to Operations at Mesoscale) est un modèle numérique de prévision, opérationnel depuis 2008 (Seity *et al.* 2011, Auger *et al.* 2015, Brousseau *et al.* 2016). Il a été conçu pour des horizons de prévision inférieurs à deux jours. Les prévisions sont mises à jour toutes les trois heures à 00 h, 03 h, 06 h, 09 h, 12 h, 15 h, 18 h et 21 h UTC, avec un délai de mise à disposition des prévisions situé entre 2 h 45 et 5 h 05 (pour la version actuelle d'AROME-FR). La résolution du modèle est passée de 2,5 km à 1,3 km en 2015. Les prévisions ARPEGE servent de conditions limites au modèle AROME-FR. La Figure 2.4 ci-dessous, illustre une comparaison entre les AROME-FR et les lames d'eau radar observées.

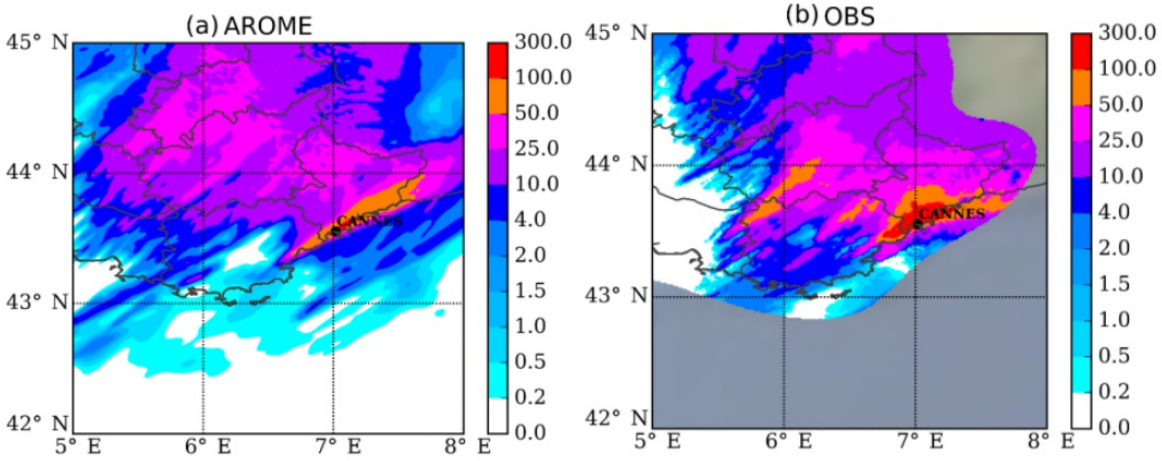

FIGURE 2.4 – *Précipitations cumulées sur 3 h (en mm) entre 18 h et 21 h (UTC) sur le sud-est de la France pour la situation du 3 octobre 2015 : a) Prévision AROME-France issue de la prévision faite à 3 h UTC (échéance 24 h), b) Analyse Antilope de précipitations cumulées sur 3 h. (CNRM)*

Une quinzaine de pays européens utilisent également AROME comme modèle opérationnel à maille fine.

Malgré ces avancées en termes de résolutions et de mises à jour plus fréquentes, des incertitudes importantes subsistent sur les prévisions de pluie déterministes réalisées par ces modèles. Si les forts cumuls pluviométriques sur des pas de temps courts sont généralement assez bien anticipés, leur localisation géographique est souvent imprécise.

2.4.2 L'assimilation de données pour améliorer la prévision immédiate

Pour réduire les incertitudes des modèles numériques de prévision à haute résolution (convection-permitting models) pour les échéances très courtes, des méthodes d'assimilation (Davolio *et al.* 2017, Lagasio *et al.* 2019, Poletti *et al.* 2019) des données ont été développées ayant pour but de rapprocher de l'état initial des modèles de la réalité des champs de pluie observés.

Ainsi, des schémas d'extrapolation et d'assimilation des précipitations, basés sur les profils d'humidité, ont été appliqués à un modèle météorologique convectif à haute résolution et évalués. D'après les travaux de Davolio *et al.* (2017), l'impact de l'assimilation des données pluviométriques est significatif pour les prévisions concernant la période de 3 heures suivant la période d'assimilation, mais l'effet de l'assimilation disparaît pour les échéances de prévision supérieures à 6 heures. De plus, l'impact de l'assimilation dépend des caractéristiques de l'environnement. AROME-FR est enrichi avec les observations via un système d'assimilation variationnel tridimensionnel (3D-Var) pour estimer l'état de l'atmosphère en prenant en compte les observations disponibles et la dernière prévision. Il s'agit d'une assimilation horaire. Grâce à son mode de fonctionnement et ses caractéristiques d'assimilation, AROME-FR est particulièrement adapté pour décrire la convection à l'échelle du kilomètre. AROME-FR représente notamment de manière très réaliste les événements avec les précipitations intenses caractéristiques des régions méditerranéennes (Fourrié *et al.* 2015).

AROME-PI (Météo-France 2020), PI pour prévision immédiate, est basé sur le modèle AROME-FR et en a les principales caractéristiques, la résolution, le domaine de calcul, le code numérique et la méthode d'assimilation (3D-Var). Afin de répondre aux besoins liés à la prévision immédiate (six heures d'échéance de prévision et disponibilité en moins de trente minutes), certaines adaptations ont été faites :

- AROME-PI produit des prévisions à une échéance maximale de 6 heures avec un pas de temps de 15 minutes.
- La prévision de l'heure H est au maximum disponible à H + 35 minutes.
- L'état initial de la prévision (faite à l'heure H) est calculé à partir d'une assimilation des données (3D-Var) basée sur la dernière prévision AROME-FR disponible et les observations récupérées entre H - 10 minutes et H + 10 minutes, le temps de calcul permettant (35 minutes de mise à disposition) d'intégrer les observations les plus récentes pour les échéances lointaines (6 heures au maximum).

La Figure 2.5 ci-dessous illustre le mode de génération des réseaux AROME-PI selon la disponibilité d'AROME-FR.

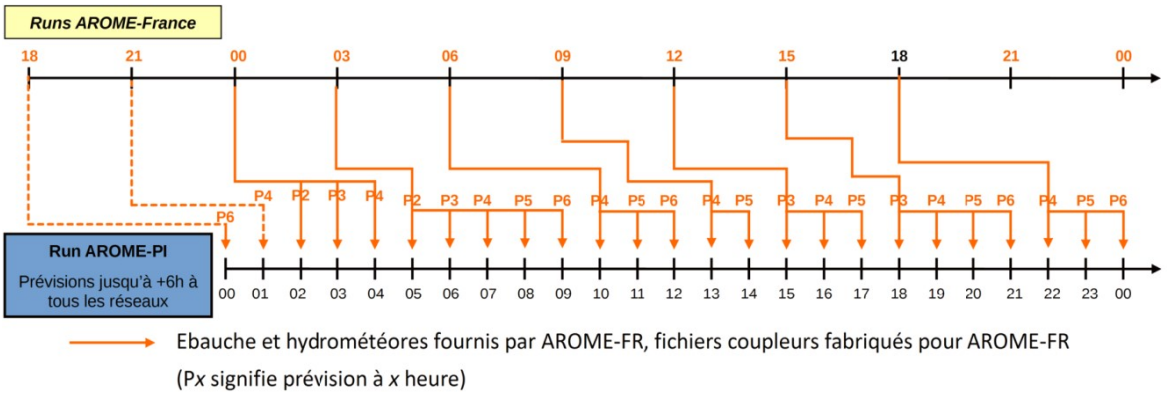

FIGURE 2.5 – *Relations entre les prévisions AROME-FR et les prévisions AROME-PI (Météo-France 2020)*

Les horizons de prévision plus réduits d'AROME-PI s'avèrent nettement plus adaptés à la prévision des crues soudaines puisqu'ils permettent de prendre en compte la rapidité de ces phénomènes, en produisant des assimilations et prévisions plus fréquentes.

Il est également possible de citer le modèle de prévision HRRR (High-resolution Rapid Refresh - (James *et al.* 2022)). Il s'agit d'un modèle convectif couvrant les États-Unis avec des prévisions générées toutes les heures pour une utilisation en temps réel.

2.4.3 L'extrapolation des lames d'eau radar observées pour la prévision immédiate

Dans les années 1950, les extrapolations radars ont commencé à être utilisées (Ligda 1953) pour la prévision immédiate des pluies à très courte échéance. Depuis, cette méthode basée sur l'extrapolation a évolué avec l'utilisation d'approches automatiques pour mieux anticiper les précipitations. Ainsi, la méthode 2PIR (Prévision Immédiate de Précipitations par Imagerie Radar, Jauffret (2016)), développée par Météo-France, est basée sur une comparaison entre les images observées successives. Les déplacements des champs de précipitations sont identifiés et un champ de déplacement peut alors être défini.

L'extrapolation de ces champs de déplacements pour les pas de temps futurs permet de produire des champs de précipitations estimés pour ces pas de temps. Cette extrapolation repose toutefois sur des hypothèses de stationnarité de l'intensité et de l'extension des zones pluvieuses, qui ne s'avèrent pertinentes que pour des horizons de quelques dizaines de minutes. Par ailleurs, la formation de nouvelles cellules pluvieuses intenses n'est pas représentée.

D'autres méthodes reposent sur la reconnaissance des cellules convectives pour suivre individuellement chaque cellule pluvieuse, qui se révèlent assez efficaces pour les précipitations intenses. Ainsi, il est possible de citer de nouveau les lames d'eau radar OPERA, qui sont utilisées pour identifier les crues-éclair localisées, induites par les précipitations à l'échelle européenne (Park *et al.* 2019). De nouvelles méthodes d'extrapolation radar basées sur des réseaux de neurones sont également en cours de développement (Shen *et al.* 2022).

2.5 Vers une approche probabiliste des prévisions

Ces modèles conservent néanmoins des limites, liées notamment à la localisation imprécise des cellules pluvieuses intenses. Ces erreurs de localisation (Vincendon *et al.* 2011) sont particulièrement pénalisantes pour les petits bassins versants. Quatre sources d'incertitudes affectent de prévision des modèles numériques météorologiques (Lovat *et al.* 2020, Vié 2013) :

- Les incertitudes sur les conditions aux limites des modèles,
- Les incertitudes sur les conditions atmosphériques initiales,
- Les limitations liées à la paramétrisation physique du modèle.
- Les incertitudes sur les conditions de surface (température de surface de la mer en particulier et des terres).

Plutôt que de considérer un scénario unique (approche déterministe), il est possible de tenir compte de ces incertitudes, en utilisant la prévision d'ensemble, qui fournit plusieurs scénarios d'évolution correspondant à différentes hypothèses plausibles concernant les conditions aux limites, initiales en particulier. La Figure 2.6 ci-dessous présente le principe théorique de la prévision d'ensemble.

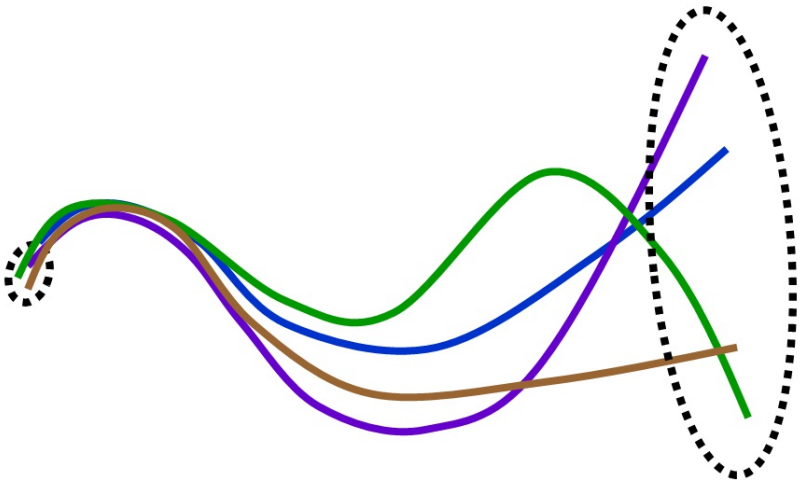

FIGURE 2.6 – *Principe théorique de la prévision d'ensemble (D'après DWD)*

La prévision d'ensemble consiste à élaborer différents scénarios de prévision afin d'obtenir une distribution statistique des scénarios d'évolution possibles, appelés membres. La Figure 2.7 illustre une prévision d'ensemble composée de 12 membres.

Il existe plusieurs façons de générer des ensembles de prévisions de pluies :

- Les ensembles peuvent être générés de manière stochastique en déplaçant le champ de pluie établi par un modèle numérique de prévision météo et en modifiant l'intensité (méthode PERT décrite par Vincendon *et al.*).

- L'état initial de l'atmosphère peut être modifié avant de lancer le modèle météo pour générer le champ de pluie. Chaque champ de pluie ainsi généré est différent du précédent.
- L'utilisation de plusieurs échéances successives d'une prévision déterministe permet également créer un ensemble (méthode de time-lag).

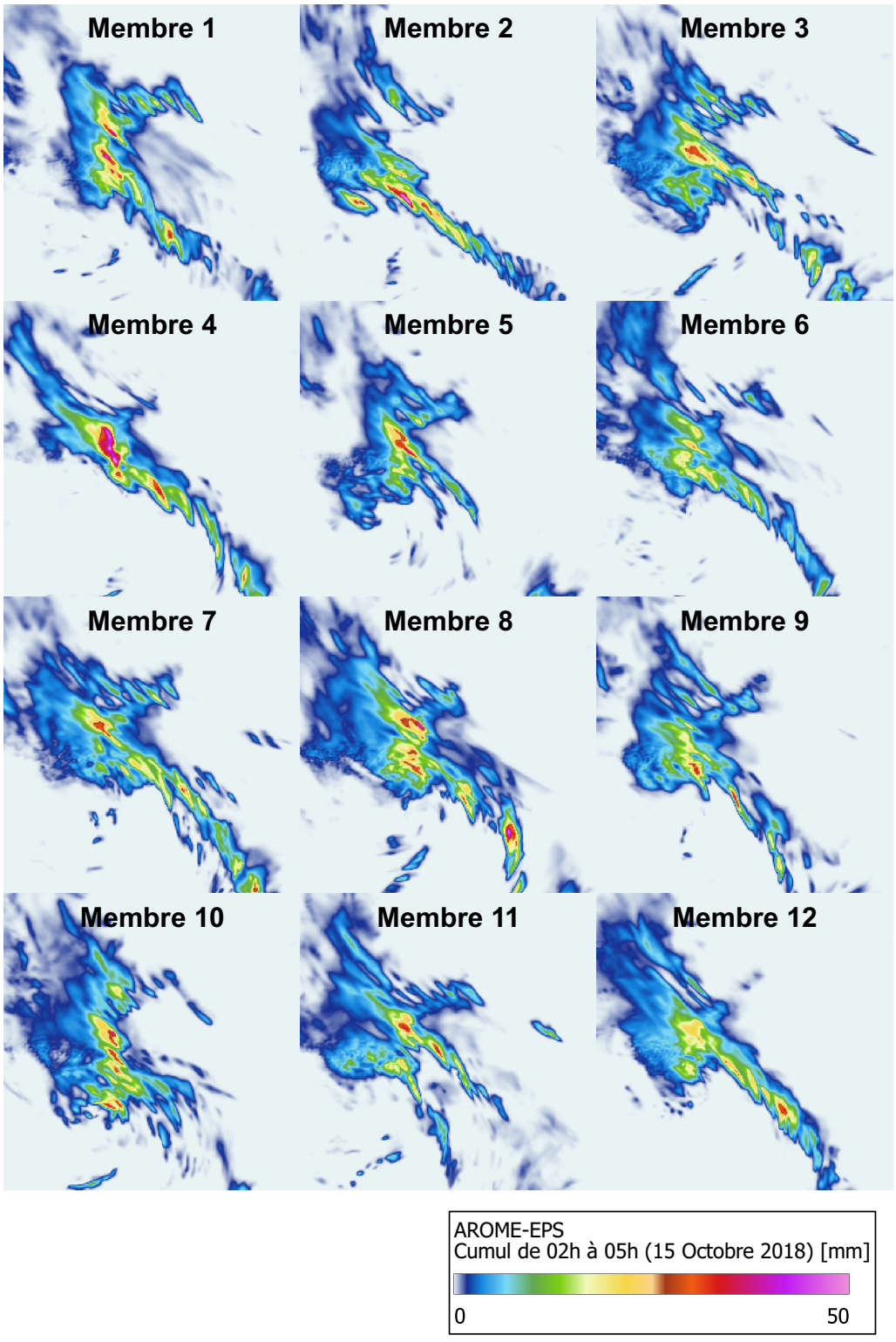

FIGURE 2.7 – *Illustration d'une prévision d'ensemble composée de 12 membres (AROME-EPS), pour des cumuls pluviométriques sur 3 heures*

L'intégration des incertitudes dans les prévisions météorologiques n'est pas une nouvelle approche. En effet, Lorenz, dès 1963, avait mis en avant la difficulté à obtenir des prévisions déterministes fiables. Tous les modèles numériques cités dans la Section 2.4.1 possèdent leur configuration de prévision d'ensemble.

PEARP (Prévision d'Ensemble ARPEGE), repose sur le modèle ARPEGE. Constituée de 35 membres, cette configuration d'ARPEGE a une résolution horizontale de 7,5 km sur la France allant jusqu'à 36 km aux antipodes.

COSMO-DE possède aussi sa propre configuration en prévision d'ensemble, COSMO-D2-EPS opérationnel pour l'Allemagne depuis 2018, avec une échéance maximale de 27 heures (et 45 heures pour les prévisions calculées à 03 h UTC). Une version antérieure, COSMO-DE-EPS, était disponible en opérationnel depuis mai 2012. COSMO-D2-EPS a une résolution horizontale de 2,2 km et est composé de 20 membres.

Enfin, il existe un produit de prévision d'ensemble opérationnel développé par Météo-France, AROME-EPS. AROME-EPS est issu du couplage entre le modèle AROME-FR et le modèle global ARPEGE qui fournit les conditions aux limites. Différentes études ont été menées sur ce modèle sur la représentation probabiliste des prévisions de pluie (Bouttier *et al.* 2012) et également sur la sensibilité initiale du modèle aux conditions de l'atmosphère (Bouttier *et al.* 2016) afin de mieux représenter les incertitudes atmosphériques des prévisions. Cet ensemble possède 12 membres (depuis juillet 2019, AROME-EPS a 16 membres) avec une échéance de prévision de 45 h à 51 h selon l'heure de prévision. La résolution horizontale d'AROME-EPS est de 2,5 km.

2.6 La production de prévisions hydrologiques

2.6.1 Utilisation des prévisions de pluie en entrée d'un modèle hydrologique

Principe général

La prévision hydrologique peut se définir comme la prévision du débit à venir d'un cours d'eau. Ces prévisions sont classiquement obtenues en alimentant un modèle hydrologique pluie-débit à partir de pluies observées ou de prévisions de pluie afin d'augmenter les échéances de prévision.

Il existe une grande diversité de modèles hydrologiques dont les différences reposent principalement sur la résolution spatiale (globale ou distribuée) et le type de modélisation des processus (physique ou conceptuelle). La représentation des processus de transformation pluie-débit et d'écoulement peut se faire en prenant appui sur les lois à base physique, ou bien par conceptualisation via un système de réservoirs inter-connectés. Dans le cas des modèles à base physique, le calage s'avère un peu moins déterminant, même s'il reste le plus souvent indispensable, au contraire des paramètres des modèles conceptuels qui doivent impérativement être calés en minimisant les écarts entre débits observés et simulés. Différents critères de mesure des écarts entre débits observés et simulés sont proposés dans la littérature.

On parle de modèle hydrologique global lorsque le bassin versant d'étude est décrit comme une seule unité homogène et avec un nombre de paramètres limité. Au contraire, un modèle hydrologique distribué tient compte de l'hétérogénéité spatiale des pluies et tente de rendre compte de la variabilité spatiale possible des réponses des bassins versants. Il existe également des modèles dits semi-distribués, pour lesquels les processus d'écoulement sont représentés par une combinaison de modèles globaux avec des paramètres propres pour chaque unité hydrologique élémentaire.

Le choix du modèle se fait en partie selon l'usage recherché. En effet, selon le contexte spatio-temporel et le niveau de détail souhaité de la représentation des processus, différents modèles pourront être privilégiés. Les méthodes pluie-débit sont nombreuses et applicables à différents contextes, avec des résolutions spatio-temporelles adaptées. Pour la prévision des crues soudaines, un maillage relativement fin donné par un modèle distribué ou semi-distribué peut s'avérer préférable, car il permet de mieux tenir compte des effets et processus associés à la forte variabilité spatio-temporelle des pluies.

Des modèles adaptés aux crues-éclair

L'amélioration de la compréhension des conditions atmosphériques, le développement des systèmes de surveillance des paramètres météorologiques et hydrologiques, ainsi que l'utilisation de modèles météorologiques et hydrologiques à résolution fine, sont des développements contribuant à l'amélioration de la capacité à prévoir les crues-éclair (Zanchetta et Coulibaly 2020). Les modèles hydrologiques à base physiques, ou à tout le moins les modèles distribués, sont souvent préférés aux modèles conceptuels, statistiques ou encore aux réseaux de neurones, car ils donnent des résultats plus réalistes (Hapuarachchi *et al.* 2011), même si certaines chaînes de prévision utilisent des modèles conceptuels comme Vigicrues Flash (SCHAPI 2018). Ainsi, le modèle GRD (Javelle *et al.* 2010, Demargne *et al.* 2019), intégré à la plateforme de prévision opérationnelle Vigicrues Flash, est un modèle conceptuel de type GR, distribué avec des mailles de 1 km^2 et un transfert immédiat de l'amont vers l'aval du bassin versant modélisé. Les modèles hydrologiques physiques utilisés dans les travaux sur la prévision des crues soudaines sont souvent distribués, comme ISBA-TOPMODEL dans les travaux de Edouard (2016) ou encore G2G (Bell *et al.* 2007), un modèle physique utilisé au Royaume-Uni.

Dans le contexte de prévision des crues-éclair dans lequel se place ce travail, les fortes intensités pluviométriques et leur grande variabilité spatio-temporelle nécessitent généralement l'utilisation de prévisions de pluie émises par des modèles à haute résolution représentant la convection (convection permitting models). L'utilisation de prévisions immédiates de pluie peut également s'avérer particulièrement intéressante, car permettant un rafraîchissement très régulier des prévisions de pluies et débits en fonction de l'évolution des observations. Plusieurs auteurs, comme Berenguer *et al.* et Heuvelink *et al.*, ont ainsi démontré que l'utilisation de prévisions immédiates des pluies permet de gagner en anticipation des débits.

Cependant, il existe encore assez peu de modèles capables de produire des prévisions fiables de crues-éclair, car la plupart des bassins concernés sont non jaugés, ce qui rend nettement plus délicate la mise en œuvre de modélisations hydrologiques fiables. L'assimilation des données observées est, comme l'utilisation des prévisions immédiates, une piste prometteuse pour limiter les incertitudes des modèles hydrologiques (Collier 2007, Jay-Allemand 2021).

2.6.2 Développement des prévisions d'ensemble des débits

Compte tenu des incertitudes associées à l'intensité et la localisation des fortes pluies prévues, l'utilisation de prévisions d'ensemble s'avère importante.

Des projets collaboratifs comme HEPEX (Hydrologic Ensemble Prediction EXperiment), ou EFAS (European Flood Alert System), ou CHROME (Couplage Hydrométéorologique RégiOnal Multi-Ensemble) ont été entrepris pour travailler sur la prise en compte des incertitudes associées aux prévisions de pluie dans les prévisions hydrologiques. HEPEX (Schaake *et al.* 2007) est un projet international mis en place en 2004, visant à faire progresser les techniques de prévision hydrologique d'ensemble, regroupant chercheurs, prévisionnistes, gestionnaires et utilisateurs. EFAS (Thielen *et al.* 2009), débuté en 2003, est un projet européen pour l'amélioration de l'alerte précoce aux inondations sur les grands bassins fluviaux européens.

Néanmoins, ces projets, en particulier EFAS, se positionnent à des échelles spatiales très larges (échelle européenne) ainsi qu'à une échéance de prévision assez grande, entre 3 et 10 jours (Thielen *et al.* 2009). Ce système EFAS est pensé et conçu pour la prévision des crues lentes des grands bassins fluviaux bien plus que pour celle des petits bassins sujets aux crues-éclair. Il existe également un dispositif d'alerte précoce dédiés aux phénomènes intenses et localisés, développé dans le cadre d'EFAS - EPIC - Precipitation Index Based on Simulated Climatology (Alfieri *et al.* 2019). Mais ce dispositif repose aujourd'hui sur les prévisions de pluie et non sur les débits prévus aujourd'hui.

D'autres travaux (Vincendon *et al.* 2011, Edouard 2016, Nardo-Caseri 2017, Lovat 2019) ont déjà été réalisés pour combiner des prévisions d'ensemble de pluies à des modèles pluie-débit afin d'obtenir des prévisions d'ensemble des débits, adaptées au cas des crues-éclair. Ainsi, des prévisions d'ensemble de pluie ont été spécifiquement mises au point par Vincendon *et al.* et Edouard (2016) et utilisées en entrée au modèle hydrologique ISBA-TOPMODEL (Bouilloud *et al.* 2010). Dans le cas des travaux de Nardo-Caseri, les modélisations hydrologiques ont été réalisées à partir d'un modèle GR (GR4H - Perrin *et al.*). Ces travaux ont mis en avant l'intérêt de l'utilisation des prévisions d'ensemble pour expliciter les incertitudes liées à la prévision immédiate des précipitations dans le cas de la prévision des crues-éclair. Il est également possible de citer le projet CHROME (Sorbet *et al.* 2014), débuté en 2011 par le SCHAPI et Météo-France. Ce projet se concentre sur la construction d'une chaîne de prévision des débits sur les bassins versants méditerranéens français, spécifiquement orientée vers les crues-éclair.

2.7 Le développement de prévisions des inondations et des impacts

L'utilisation de données autres que les débits et les hauteurs d'eau (données indirectes, incluant notamment les données d'impacts) sur l'intensité des crues dans les bassins non jaugés, pour le calage et la validation des modèles hydrologiques de prévision des crues, a commencé à être abordée dans plusieurs travaux récents (Nardo-Caseri 2017, Naulin *et al.* 2013, Randrianasolo 2009, Versini 2007). Cette utilisation présente néanmoins des difficultés, non seulement pour acquérir les données d'impacts, nécessaires au calage et à l'évaluation des modélisations, mais également pour développer des modèles d'impacts adaptés.

2.7.1 Une identification des secteurs inondés difficile

L'identification des secteurs inondés lors d'une crue est nécessaire pour l'estimation des impacts potentiels. En effet, le débit est une information disponible dans le lit mineur des cours d'eau, mais en cas de forts débits, l'eau peut s'étendre au-delà de son lit mineur. La modélisation hydraulique est nécessaire pour prévoir l'étendue de l'inondation. Le passage a une modélisation hydraulique permet :

- D'estimer la hauteur d'eau atteinte en tout point du domaine de modélisation.
- D'estimer les vitesses d'écoulement.

L'étendue des zones inondées dépend inévitablement des débits prévus par la modélisation hydrologique. Il existe donc déjà une incertitude sur les sorties de débits obtenues, associée aux limites de la modélisation hydrologique.

Le calage des modèles hydrauliques est également une difficulté, car cela nécessite d'avoir une bonne connaissance des caractéristiques du domaine de modélisation comme la rugosité et la topographie. La modélisation hydraulique est très souvent combinée à une approche historique. L'approche historique permet d'établir des cartes après un événement à partir de laisses de crues, de repères de crues, de témoignages ou encore de photos de l'événement. Ces éléments ne sont pas toujours aisés à collecter et comportent également leur part d'incertitudes (déplacement post-événement des laisses de crues, subjectivité des témoignages...). De plus, les résolutions des MNT (Modèles Numériques de Terrain) utilisés dans le cadre des modélisations hydrauliques étaient, jusqu'à présent, relativement souvent incompatibles avec la représentation des petits cours d'eau.

Depuis quelques années, plusieurs travaux ont été réalisés sur des modélisations hydrauliques adaptées à la prévision des crues soudaines (Le Bihan 2016, Hocini 2022), en particulier pour résoudre les difficultés associées à ce type de modélisation lorsqu'elles sont déployées à large échelle (échelle régionale), avec des possibilités réduites de calage préalable des modèles hydrauliques.

2.7.2 Développement de méthodes de prévisions d'impact

Dans ce contexte, Le Bihan a proposé une méthode d'évaluation directe des impacts des crues soudaines à partir des sorties d'un modèle pluie-débit distribué et d'un calcul automatisé des surfaces inondées (Le Bihan 2016, Le Bihan *et al.* 2017). Pour chaque bief de cours d'eau, un catalogue d'emprises d'inondation est élaboré, pour différentes valeurs de débits (différentes périodes de retour, à partir d'une modélisation hydraulique 1D automatisée et d'un MNT. À partir de la cartographie et d'une base de donnée des bâtiments (BD Topo de l'IBN), il est possible de d'évaluer un nombre de bâtiments dans la zone de crue selon la période de retour de cette dernière. La Figure 2.8 illustre les zones inondées obtenues (Figure 2.8 a)) ainsi que le nombre de bâtiments inondés associés par période de retour (Figure 2.8 b)).

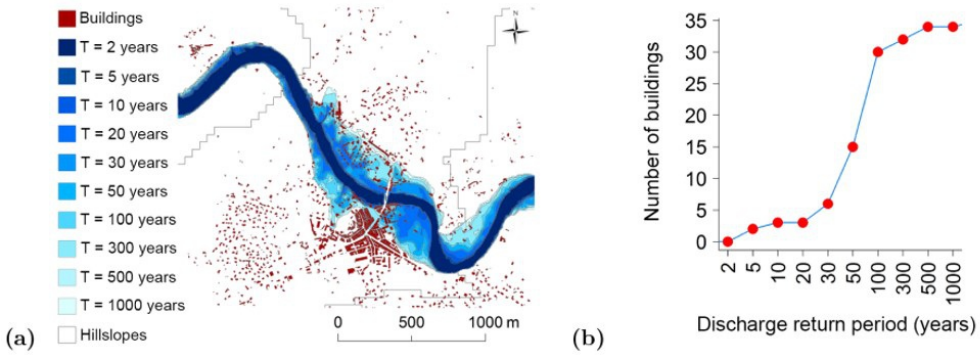

FIGURE 2.8 – *Illustration de la mise en œuvre du modèle d'impact sur un tronçon de rivière : a) catalogue des zones inondées et b)relation débit – impact interpolé (Le Bihan* et al. *2017)*

La méthode mise en place par Le Bihan a déjà été évaluée en comparant les impacts simulés aux données de sinistralité d'assurance disponibles. Comme le montre la Figure 2.9, les impacts estimés pour le TRI d'Alès (Figure 2.9 b)) sont assez proches du décompte de sinistres déclarés (Figure 2.9 c)).

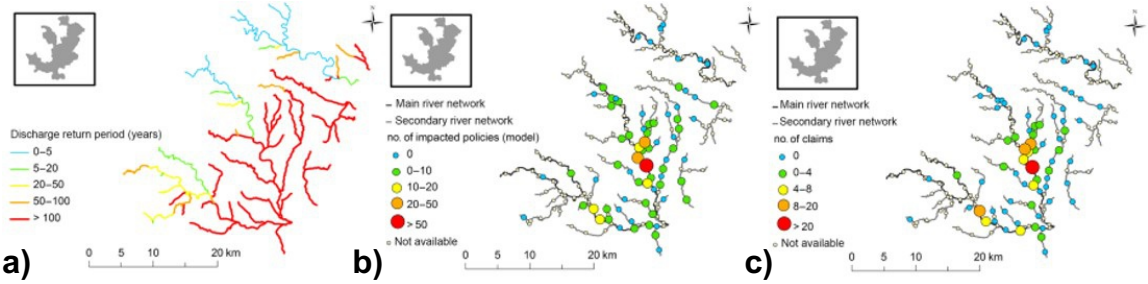

FIGURE 2.9 – *Cartes des débits de pointe (périodes de retour) et des impacts associés (nombre de polices d'assurance inondées) simulées par le modèle, et des nombres de sinistres extraits de la base de données CCR pour Alès 2002 (Le Bihan* et al. *2017)*

Ces premiers résultats ont été jugés encourageants, malgré les difficultés liées à la précision des données topographiques (MNT), aux simplifications introduites dans l'approche hydraulique, et à la qualité variable des données de sinistralité (taux de couverture faible et déclaration des sinistres non systématique).

Concernant les données disponibles pour construire des modèles d'impacts, il est assez aisé de collecter des données sur les impacts potentiels (ou enjeux) qui peuvent être des personnes, des infrastructures ponctuelles (bâtiments, établissements de santé) et linéaires (coupures de route, réseau électrique). Ces informations se trouvent exhaustivement dans la BD TOPO (disponible sur geoservices.ign.fr) qui est une description vectorielle de la composition du territoire national. Des travaux ont déjà été menés sur le poids à donner à ces différentes données sur les enjeux pour obtenir une vision pertinente du niveau d'impacts générés par une inondation (Saint-Martin 2018). Différents types de données d'impacts observés ont déjà été utilisés comme :

- Des données de coupure de routes, enregistrées par le Conseil Général, ont été utilisées par Naulin *et al.* sur près de 2000 points de croisement dans le département du Gard.
- Des données de sinistralité issues de la base de données de la CCR ont été exploitées par Le Bihan, en particulier pour les crues soudaines survenues dans le Gard en 2002.
- Des données issues d'enquêtes de terrain ont été collectées par Saint-Martin pour la base de données DamaGIS (dommages survenus depuis 2011 sur le sud de la France).
- Des données sur la mortalité ont également servi dans les travaux de Petrucci *et al.*, Papagiannaki *et al.*, Vinet *et al.*.

Néanmoins, ces données, observées spécifiquement sur des événements et nécessaires à l'évaluation des résultats de modélisation hydraulique, restent relativement difficiles à rassembler.

2.7.3 Outils de cartographie automatisée

Ces dernières années encore, les outils de cartographie et de prévisions des inondations étaient peu disponibles, ce qui est en train de changer. L'augmentation de la capacité de calcul et une plus grande précision des données de topographie ont permis d'appliquer des modèles d'inondation à grande échelle et à haute résolution (Wood *et al.* 2011). Des plaines inondables à partir d'un catalogue de débits (obtenu par des données de forçage météo sur une période de 30 ans) ont été évaluées à l'échelle européenne (Pappenberger *et al.* 2011) mais également à l'échelle mondiale en modélisant des cartes d'inondation avec les impacts associés (Winsemius *et al.* 2013). D'autres études ont été faites à une échelle plus locale avec des topographies fines avec une volonté d'application en temps réel (de Almeida *et al.* 2018, Neal *et al.* 2011, Pedrozo-Acuña *et al.* 2015).

Des travaux plus récents ont été menés pour évaluer différentes méthodes de cartographie des inondations adaptées aux petits cours d'eau (Hocini 2022). L'évolution de la qualité des MNT et des outils de simulation hydraulique (approche 2D) a permis d'atteindre des niveaux d'incertitudes jugés acceptables, incertitudes qui restent principalement dues à certains défauts subsistant dans les données topographiques.

2.8 Les méthodes d'évaluation d'une prévision d'ensemble

La vérification des prévisions hydrométéorologiques permet aux prévisionnistes de connaître les qualités et défauts de leurs prévisions. Cette vérification est très importante pour gagner en confiance dans les résultats de prévisions futures. Cela se répercute directement sur la facilité des utilisateurs à pouvoir prendre des décisions à partir des prévisions.

Alors qu'une prévision déterministe se compose uniquement d'un seul scénario, une prévision d'ensemble se compose de plusieurs scénarios considérés comme équiprobables, comme cela a déjà été évoqué dans la Section 2.5. Un scénario unique peut être comparé assez facilement aux observations en choisissant un critère mesurant l'écart entre les valeurs simulées et observées. Dans le cas des prévisions d'ensemble, il faut évaluer une distribution prédictive fournie par l'ensemble par rapport à la valeur de l'observation. Ainsi, l'évaluation de la qualité d'un ensemble se fait à l'aide de scores statistiques développés d'abord en météorologie, mais qui peuvent être transposés aux prévisions hydrologiques (Jolliffe et Stephenson 2012).

2.8.1 Scores statistiques classiques de vérification des prévisions d'ensemble

Ce paragraphe présente ces différentes méthodes dédiées à la vérification des prévisions d'ensemble.

D'une façon générale, la vérification permet d'évaluer la précision des prévisions, d'en trouver les biais et points de faiblesse et comparer la qualité de différents ensembles de prévision (Marshall 2007). Il existe deux types de vérifications faites sur les prévisions d'ensemble pour évaluer leur qualité :

- Évaluer des propriétés propres à l'ensemble comme la dispersion de ses membres : variance et distance inter-quantiles.
- Vérifier la qualité des prévisions d'ensemble au regard des observations (Jolliffe et Stephenson 2012).

Ces deux vérifications sont complémentaires, l'une se concentrant uniquement sur les qualités intrinsèques de l'ensemble, l'autre évaluant la vraisemblance de l'ensemble par rapport aux observations.

L'évaluation est réalisée en fonction de l'objectif recherché. Même si un prévisionniste ou un utilisateur des prévisions ne poursuivent pas des objectifs d'évaluation identiques, les mêmes attributs classiques seront néanmoins souvent regardés.

Il est possible de distinguer 6 attributs principaux recherchés (Golding 2022) pour une prévision d'ensemble :

- La fiabilité : la fiabilité décrit la cohérence statistique entre les observations et les prévisions d'ensemble, soit la capacité du système de prévision d'ensemble à rendre compte des probabilités d'occurrence des observations (Wilks 1995; 2011). Des défauts de fiabilité peuvent être liés soit à un biais des prévisions - c'est-à-dire tendance générale à surestimer ou sous-estimer le phénomène prévu - soit à une sous ou sur dispersion des ensembles.
- La finesse : la finesse est une mesure de la dispersion des ensembles.
- La précision (exactitude) : la précision mesure l'écart moyen entre la prévision et l'observation (Wilks 1995; 2011). C'est un attribut synthétique de la qualité globale d'une prévision d'ensemble.
- La discrimination : la discrimination évalue la capacité d'une prévision à distinguer un événement (c'est-à-dire un dépassement de seuil par exemple), d'un non-événement.
- La résolution : la résolution, d'après Golding (2022), désigne la capacité de la prévision à diviser les événements en sous-ensembles dont les résultats diffèrent de manière caractéristique.
- La compétence : la compétence mesure l'écart des performances d'un modèle de prévision par rapport à un modèle de référence, souvent le modèle "naïf" consistant à reconduire la dernière valeur observée.

Pour Murphy, Candille et Talagrand, les deux attributs essentiels d'une prévision d'ensemble sont la fiabilité et la finesse. Ainsi, un ensemble très dispersé, a une finesse faible, peut être statistiquement juste, cependant en pratique, il sera peu informatif et difficile à valoriser, car il traduit une grande incertitude. Il est donc intéressant de produire des prévisions qui soient les plus fines possibles. Cependant, avoir des prévisions fines, mais peu fiables, ne présente que peu d'intérêt, car la probabilité émise ne correspond pas à la probabilité réelle de l'événement. Gneiting *et al.* précisent qu'il est donc préférable de chercher d'abord à s'assurer de la fiabilité avant de s'intéresser à la finesse des prévisions.

Il existe un grand nombre de scores adaptés à la vérification des prévisions d'ensemble :

- Scores numériques basés sur la moyenne de l'ensemble : biais, MAE (mean average error) ou encore RMSE (root mean square error).
- Scores numériques sur la totalité de l'ensemble : score de Brier, CRPS (continuous ranked probability score) ou encore le FSS (fractions skill score).
- Outils graphiques : diagrammes d'attribut, de fiabilité, d'acuité et de rang.

Tous ces scores sont calculés à partir d'une comparaison des prévisions avec une référence. Cette référence peut être des observations (lame d'eau, débits aux stations hydrométriques...) ou bien des proxys (sorties d'un modèle météorologique, hydrologique...).

Une description détaillée des principaux scores (liste non exhaustive) de vérification des prévisions d'ensemble est proposée par Anctil et Ramos dans le Chapitre "Verification Metrics for Hydrological Ensemble Forecasts" extrait de l'ouvrage *Handbook of Hydrometeorological Ensemble Forecasting*.

Même s'il existe une multitude de scores d'évaluation, il n'y a pas de nécessité à tous les mettre en œuvre pour réaliser l'évaluation d'une prévision d'ensemble. Il s'agit de sélectionner quelques scores ou représentations graphiques, permettant d'évaluer prioritairement la fiabilité, et dans un second temps la finesse des prévisions d'ensemble. Pour ce qui est de la fiabilité, les diagrammes d'attribut, de fiabilité et de rang font partie des outils évaluant uniquement ce critère. Pour ce qui est de la finesse, le CRPS est le critère le plus fréquemment utilisé, ce critère dépendant aussi de la fiabilité.

2.8.2 Enjeux de l'évaluation événementielle des prévisions d'ensemble

Les scores statistiques, présentés au paragraphe précédent, nécessitent de disposer d'une grande quantité de données sur laquelle effectuer l'évaluation. Ces scores demandent donc de disposer d'une large période de rejeu des prévisions à évaluer. Dans le cas de la prévision des crues, il est néanmoins souvent utile d'évaluer les prévisions d'ensemble à l'échelle d'un événement de crue particulier. Cette évaluation événementielle est difficile à mener à partir des scores classiques de la vérification, le nombre de couples prévision - observation étant réduit dans le temps et dans l'espace. Les évaluations événementielles reposent donc le plus souvent sur de l'inspection visuelle des prévisions à partir des hydrogrammes observés ou bien sur l'anticipation de dépassements de seuil de débits prédéfinis au niveau de stations hydrométriques ou de points d'estimation des débits (Vincendon *et al.* 2011, Vié *et al.* 2012, Davolio *et al.* 2013, Hally *et al.* 2015, Nuissier *et al.* 2016, Amengual *et al.* 2017, Lagasio *et al.* 2019, Sayama *et al.* 2020).

Lorsque différentes séries de prévisions sont disponibles pour le même événement et pendant toute sa durée, comme des prévisions à court terme générées par des modèles de prévision numérique à partir de différents cycles d'initialisation des prévisions, et/ou lorsque différents bassins sont affectés par le même événement, des scores statistiques tels que le RMSE, le CRPS, des tableaux de contingence ou des courbes ROC, peuvent également être utilisés pour fournir une évaluation synthétique de la performance des prévisions pour l'événement évalué (Davolio *et al.* 2017, Poletti *et al.* 2019, Sayama *et al.* 2020).

Bien que largement utilisés ces dernières années, en particulier dans les études post-événements, ces cadres d'évaluation soulèvent plusieurs questions méthodologiques :

- L'accent est mis sur un événement en particulier ou quelques épisodes typiques et graves. Cela peut générer une évaluation spécifique et non reproductible pour d'autres événements. Rien ne garantit donc des performances équivalentes pour des événements futurs.

- L'évaluation se concentrant le plus souvent sur quelques exutoires jaugés, il n'y a qu'une vue partielle des performances de la prévision, notamment lorsque des impacts sont également observés sur des sites non jaugés et/ou lorsque des décalages spatiaux importants existent entre les précipitations observées et prévues.

- Les scores (trop) synthétiques, qui agrègent les performances des différentes prévisions, peuvent dissimuler une disparité des performances de la prévision du point de vue spatial et temporel.

Par conséquent, l'évaluation des prévisions à court terme des crues soudaines à l'échelle de l'événement nécessite non seulement de prendre en compte la qualité des prévisions évaluées sur les exutoires jaugés avec des observations de débits, mais également de réaliser une évaluation à une échelle plus large, au niveau régional, en incluant des secteurs non jaugés pour obtenir une évaluation plus robuste des performances de la prévision (Anderson *et al.* 2019, Davolio *et al.* 2015, Silvestro et Rebora 2012).

2.9 Cadre méthodologique retenu pour l'évaluation événementielle des prévisions hydrométéorologiques d'ensemble

Un cadre d'évaluation a été proposé et appliqué lors de cette thèse, pour déterminer si les prévisions d'ensemble anticipent correctement l'intensité et la localisation des crues et des inondations. Il n'est pas attendu de cette évaluation qu'elle se focalise sur les points d'observation des débits, mais plutôt qu'elle donne une vision globale des capacités d'anticipation sur l'ensemble du réseau hydrographique. Comme cela a déjà été évoqué précédemment (Section 2.8.2), l'évaluation événementielle des prévisions d'ensemble est relativement complexe du fait d'un nombre de données limité à la fois spatialement, mais aussi temporellement sur un événement unique. Il faut donc sélectionner un nombre limité de critères agrégés pour tirer des conclusions solides sur la performance des prévisions.

Pour cela, une approche en cinq étapes a été proposée. Dans un premier temps, une évaluation des prévisions pluviométriques se focalise sur les fenêtres temporelles et spatiales des forts cumuls de pluie observés et prévus. Ensuite, une analyse géographique des capacités d'anticipation de seuils de débits est réalisée sur un grand nombre d'exutoires non jaugés, en se concentrant sur la phase la plus critique des crues (montée de crue). L'un des objectifs principaux de la thèse étant l'évaluation de la plus-value des prévisions immédiates de pluie pour la prévision des crues soudaines, les débits simulés à partir des pluies observées seront pris pour référence. Ce choix permettra d'évaluer les prévisions dans les secteurs non jaugés. Sur la base de cette deuxième étape, une évaluation détaillée de la performance des prévisions de débits est effectuée au niveau des certains points jugés représentatifs du bassin versant (étape 3). À cette étape, les débits observés pourront être considérés et les incertitudes de prévision liées à la modélisation pluie-débit discutées.

Puis dans une quatrième étape, les prévisions d'impacts des crues, obtenues à partir des prévisions d'ensemble des débits, sont à leur tour évaluées sur leur capacité à anticiper des seuils d'impacts. Enfin, ces prévisions d'impacts sont également plus spécifiquement étudiées pour certains tronçons représentatifs (étape 5).

La Figure 2.10 présente la méthodologie mise en place pour l'évaluation hydrologique événementielle de prévisions d'ensemble.

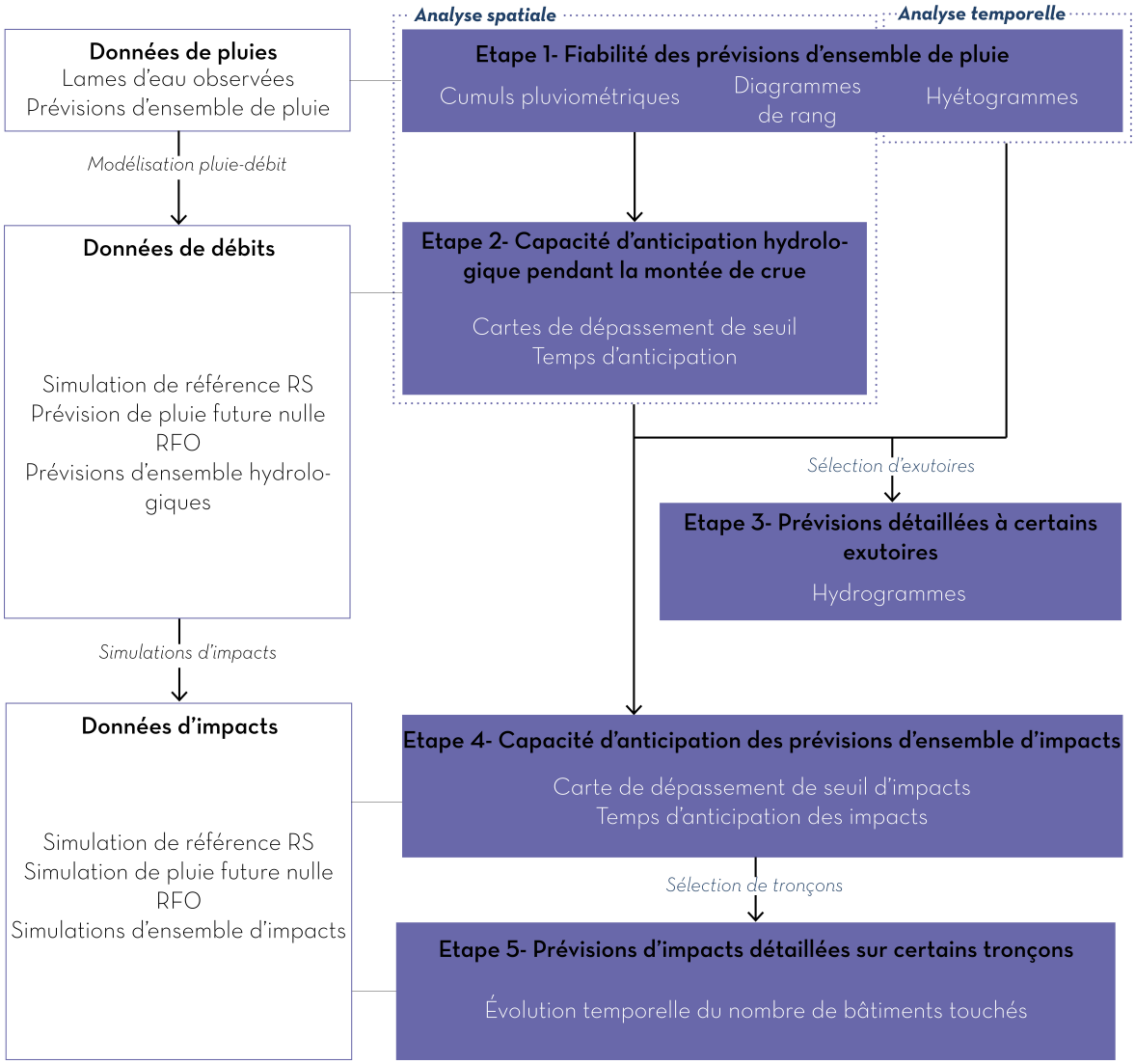

FIGURE 2.10 – *Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble*

Les différentes étapes de la méthodologie sont décrites plus en détail dans les Sections 2.9.1 à 2.9.5.

2.9.1 Étape 1 : Fiabilité des ensembles de pluie

Le double objectif de cette phase initiale est d'analyser les qualités des prévisions pluviométriques et de définir les échelles spatiales et temporelles pertinentes pour les analyses suivantes. Le choix a donc été fait de réaliser une évaluation en deux étapes avec :

- Une comparaison des données de pluie des prévisions d'ensemble avec la lame d'eau radar observée via des hyétogrammes horaires moyens ainsi que des cartes de cumuls des précipitations observées.
- Une évaluation basée sur des diagrammes de rang et CRPS, deux scores synthétiques adaptés à l'évaluation de la prévision d'ensemble, en sachant que les conclusions de cette évaluation seront partielles et à considérer avec précaution au regard des réserves émises sur le volume limité de données évaluées.

Cette évaluation des prévisions d'ensemble de pluie sera présentée dans le Chapitre 4.

2.9.2 Étape 2 : Capacité d'anticipation de seuils de débit pendant la montée de crue

L'objectif de cette deuxième étape est de caractériser la capacité d'anticipation de dépassements de seuils de débits pour la fenêtre spatio-temporelle prédéfinie lors de l'étape 1, et plus particulièrement pendant la phase de montée de crue, qui est la période critique pour un utilisateur des prévisions d'ensemble.

L'évaluation est essentiellement basée sur une approche classique de tableau de contingence (Wilks 2011), avec quelques adaptations importantes visant à concentrer l'analyse sur la fenêtre temporelle la plus critique du point de vue de l'utilisateur, et à agréger les prévisions émises durant cette fenêtre temporelle, indépendamment des échéances.

Il est également possible à partir de cette évaluation de calculer des délais d'anticipation de ces dépassements de seuils.

Le Chapitre 5 est consacré à la présentation détaillée de cette évaluation basée sur une adaptation des tableaux de contingence et des courbes ROC.

2.9.3 Étape 3 : Prévisions détaillées à certains exutoires

Cette étape 3 permet de faire le lien entre les résultats de l'anticipation des dépassements de seuils de débit et les caractéristiques détaillées des prévisions d'ensemble. Certains exutoires d'intérêt sont plus particulièrement étudiés à l'aide d'hydrogrammes, comme cela est le plus souvent fait lors d'évaluations hydrologiques de prévisions d'ensemble. Les débits observés ou débits maximums de crues estimés seront aussi considérés.

Cette étape, très liée à l'anticipation des seuils de débits (étape 2), est également présentée dans le Chapitre 5.

2.9.4 Étape 4 : Capacité d'anticipation de seuil d'impacts

L'avant-dernière étape s'intéresse à l'identification des secteurs touchés par l'inondation, et plus particulièrement des impacts qui en découlent. En effet, au-delà des débits de crue, il est important d'évaluer si les prévisions permettent d'anticiper la localisation et l'ampleur des impacts observés.

La méthode d'évaluation appliquée est la même que celle utilisée pour les débits, inspirée des courbes ROC, et est présentée dans le Chapitre 6.

2.9.5 Étape 5 : Prévisions d'impacts détaillées sur certains tronçons

La dernière étape, très semblable à l'étape 3 pour l'évaluation des débits, consiste à regarder sur quelques tronçons d'intérêt le comportement détaillé des prévisions d'impacts au cours de l'événement, afin de confirmer ou d'infirmer les qualités des prévisions déjà repérées pour l'anticipation des dépassements de seuils de débits.

Comme pour les débits, cette étape directement en lien avec l'évaluation de l'anticipation de seuils d'impacts, et est présentée dans le Chapitre 6.

Synthèse

Ce Chapitre a permis de présenter les différents enjeux liés à la prévision des crues soudaines sur le pourtour méditerranéen. Ces prévisions doivent en effet répondre aux caractéristiques spatio-temporelles spécifiques de ce type d'événement : phénomènes convectifs rapides et très localisés. Les critères spécifiques utilisés pour l'évaluation de la prévision d'ensemble ont également été présentés.

Le développement des modèles de prévision météorologiques à haute résolution, et de l'assimilation des données d'observations, permettent d'obtenir des prévisions immédiates de pluies (prévision déterministe) comme AROME-PI. L'utilisation de l'approche probabiliste avec les prévisions d'ensemble est également prometteuse pour l'obtention de prévisions immédiates tenant compte des incertitudes dans un contexte de crues soudaines.

Le couplage pluie-débit, très largement utilisé dans les chaînes de prévision, peut également être adapté au contexte de crues-éclair, en particulier avec l'utilisation de modèles distribués donnant une haute résolution spatiale, nécessaire pour ce type d'événements hydrologiques. Comme pour les prévisions d'ensemble de pluies, les prévisions d'ensemble de débits ont jusqu'à présent plutôt porté sur les grands bassins fluviaux et les crues lentes, mais des travaux récents illustrent l'intérêt des prévisions d'ensemble de débits pour rendre compte des incertitudes liées à la prévision immédiate des crues soudaines.

Même si le développement des prévisions d'impacts reste pour le moment plus limité, l'identification des secteurs inondés et l'évaluation rapide des impacts potentiels sont cruciaux pour l'organisation des opérations de secours. Compte tenu de la rapidité des crues soudaines, cette étape doit idéalement pouvoir être aussi automatisée. Plusieurs travaux récents, mobilisant plusieurs sources de données d'impacts (mortalité, sinistralité, coupures de routes), montrent l'importance, pour les utilisateurs des prévisions, de ce type de modélisation.

Cet aperçu de l'état de l'art, de la production d'une prévision de pluie à son utilisation en tant que donnée d'entrée de modélisations des débits ou des impacts, a permis de mettre en évidence l'intérêt du développement de chaînes de prévision immédiate d'ensemble, adaptées aux phénomènes de crues soudaines, mais également d'illustrer l'importance et la difficulté de l'évaluation de ce type de produit et l'intérêt d'une évaluation événementielle détaillée. En effet, il semble essentiel de se demander à quels besoins doivent répondre ces prévisions pour l'utilisateur final du produit, et d'évaluer les prévisions en fonction de ces besoins. Cette question sera abordée dans le Chapitre 7 de cette thèse.

Chapitre 3

Évènements étudiés, données et modèles utilisés

Objectifs du chapitre :

Ce chapitre présente les trois principales crues soudaines de l'année 2018 retenues pour l'étude et l'évaluation des prévisions hydrométéorologiques d'ensemble. Un inventaire des données pluviométriques et hydrologiques disponible est également établi dans un second temps. Une troisième section présente les produits de prévision d'ensemble qui seront par la suite évalués. Enfin, le modèle pluie-débit utilisé pour les simulations hydrologiques, Cinecar, est décrit, ainsi que le catalogue d'emprises obtenu à partir d'une modélisation hydraulique.

3.1 Trois crues soudaines majeures de l'année 2018 retenues comme cas d'étude

L'année 2018 a été marquée d'un point de vue hydrologique par un début d'année particulièrement pluvieux avec des inondations de plusieurs semaines sur le bassin de la Marne et de la Seine (CCR 2018b). Par la suite, d'autres événements se sont succédé, avec une crue majeure de la Loue (Doubs), de nombreux orages générateurs de crues (Aquitaine, Bretagne, Aveyron, Haute-Marne, Haute-Corse), ainsi que trois événements de pluie intense dans le Gard et l'Ardèche, l'Aude, puis le Var (Météo-France, base de données des pluies extrêmes). Dans le cadre du projet ANR-PICS (pics.ifsttar.fr), trois produits de prévision immédiate d'ensemble ont été développés et appliqués sur quelques événements de pluie intense de l'année 2018. Pour cette raison, trois événements de crues soudaines observés lors de cette fin d'été et automne 2018 ont été retenus ici comme cas d'étude. Ces trois événements sont présentés en détail, par intensité décroissante, dans les paragraphes qui suivent.

3.1.1 Les crues de l'Aude en octobre 2018

L'événement pluvieux le plus marquant de l'année 2018 est survenu mi-octobre, sur le bassin de l'Aude. Des pluies intenses se sont abattues entre les 14 et 16 octobre 2018, avec des cumuls pluviométriques particulièrement importants dans la nuit du 14 au 15, atteignant jusqu'à 300 mm en 12 h aux alentours de Carcassonne. Ces pluies ont généré des crues exceptionnelles occasionnant des dégâts matériels estimés à 200 millions d'euros (CCR 2018a). La pluviométrie horaire a été très importante localement (60 mm/h), mais ce sont surtout les cumuls pluviométriques sur une durée de six heures qui rendent l'événement assez exceptionnel, puisque la période de retour centennale a été dépassée sur ce pas de temps dans la bande pluvieuse la plus intense avec, par exemple, 213 mm en six heures à Villegailhenc, alors que le cumul centennal est de 101 mm (Ayphassorho *et al.* 2019). La Figure 3.1 représente le cumul des précipitations observées lors de cet événement.

D'un point de vue hydrologique, l'événement s'est déroulé en deux temps avec :

- La crue des petits cours d'eau affluents dans la partie intermédiaire du bassin de l'Aude (Trapel, Rieu Sec, Orbiel, Fresquel) ainsi que de l'Orbieu et la Cesse dans la partie aval. Pour ces cours d'eau, les pluies intenses ont directement contribué aux augmentations des débits, qui ont été très rapides.
- La propagation des ondes de crue en provenance de ces différents affluents, ainsi que de la partie amont du bassin versant de l'Aude, vers l'exutoire du bassin versant.

Pour les affluents de l'Aude, il s'agit de crues exceptionnelles d'une période de retour qui semblerait être proche de la centennale (Ayphassorho *et al.* 2019). Pour l'Aude, le rôle prépondérant des affluents a contribué à une crue exceptionnelle dans la partie intermédiaire du bassin, à l'aval de Trèbes, alors que les crues dans les basses plaines de l'Aude peuvent seulement être qualifiées d'importantes.

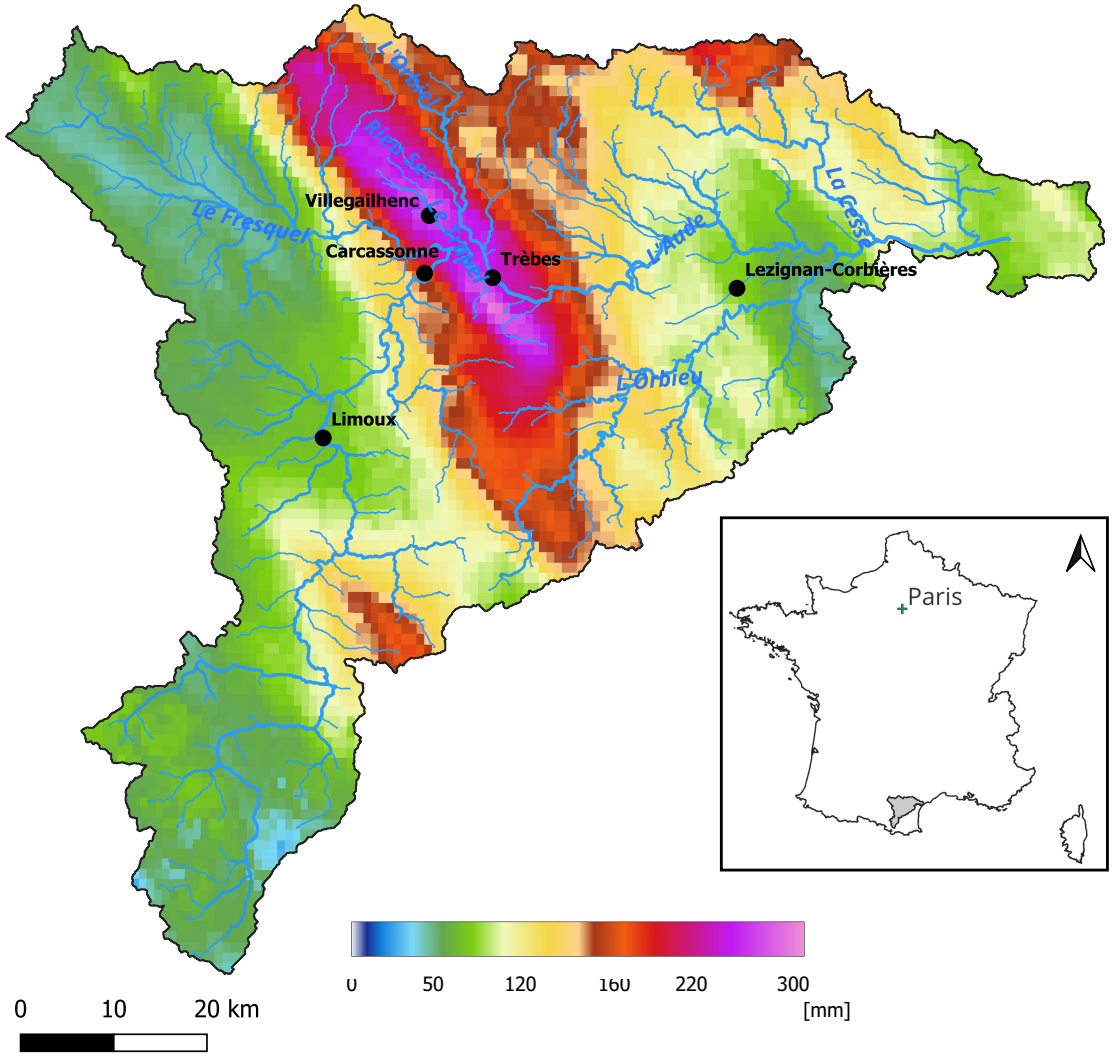

FIGURE 3.1 – *Cumul de pluie (lame d'eau ANTILOPE J+1) du 14 octobre au 16 octobre 2018*

Ce secteur est propice à ce type d'événement. Historiquement, les crues dévastatrices de novembre 1999 (Gaume *et al.* 2004) ressemblent à celles de 2018 par leur intensité, même si les affluents touchés n'étaient pas exactement les mêmes. Dans le secteur de Trèbes néanmoins, seule la crue de 1891 se rapproche, voire dépasse en intensité celle observée en 2018.

3.1.2 Les crues de la partie aval de l'Ardèche et de la Cèze, en août 2018

De fortes intempéries ont eu lieu dans le nord du Gard et le sud de l'Ardèche du 8 au 10 août 2018, avec des orages et des pluies particulièrement intenses au cours de la journée du 9 août. Avec le passage d'une dépression d'ouest en est, l'air froid d'altitude (lié à la dépression) et la remontée d'un air très chaud et très humide provenant de Méditerranée ont provoqué une forte instabilité. Le flux s'est maintenu au sud pendant plusieurs heures, et a généré des orages quasi stationnaires qui ont apporté des pluies intenses, particulièrement dans les vallées aval de la Cèze et de l'Ardèche. Le 10 août, la perturbation s'est décalée vers

le sud-est, en perdant en intensité. Les événements pluvieux après cette date ne sont plus considérés comme des épisodes de pluie intense.

La carte de la Figure 3.2 montre les cumuls pluviométriques estimés pour la journée du 9 août, en combinant les enregistrements radars et pluviométriques (lame d'eau Antilope J+1), sur les principaux bassins touchés par l'événement : l'Ardèche, la Cèze et l'Eyrieux en rive droite du Rhône, ainsi que l'Aigue, le Jabron et la Drôme (rive gauche du Rhône). La rive gauche du Rhône a été globalement moins impactée, avec des cumuls pluviométriques moins importants, au maximum autour de 150 mm en 24 h vers Bollène. Les secteurs Ardèche et Cèze, en rive droite, ont été pour leur part touchés par des cumuls particulièrement importants supérieurs à 240 mm en 24 h, avec un débordement de la cellule pluvieuse intense au Nord-Ouest (Beaumont) et à l'est.

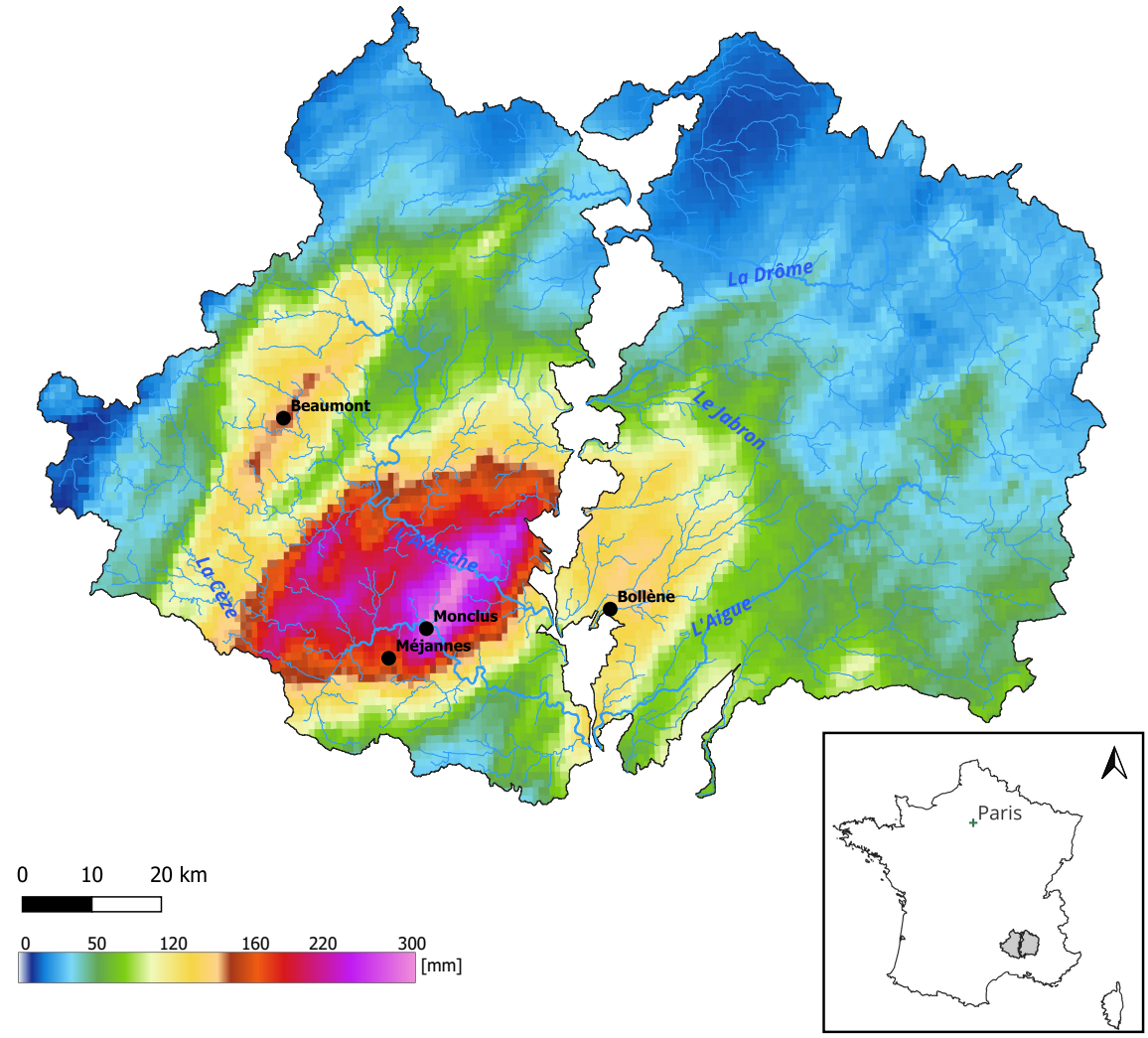

FIGURE 3.2 – *Cumul de pluie (lame d'eau ANTILOPE J+1) du 9 août 2018*

Pour ce qui est de la pluviométrie horaire, elle a été particulièrement importante en rive droite du Rhône, avec exemple 113 mm en 1 h à Monclus (bassin de la Cèze). Les cumuls pluviométriques sur 12 h rendent également l'événement assez conséquent avec :

- 292 mm à Monclus.
- 173 mm à Méjannes.

Les conséquences de ces pluies intenses ont été des crues importantes de certains affluents de

la Cèze et de l'Ardèche, qui ont occasionné des dommages matériels significatifs, ainsi qu'un décès.

Cet événement reste malgré tout relativement modéré en termes de précipitations et de dégâts observés, par rapport à ce qui peut être observé régulièrement dans la région. On peut ainsi citer les précipitations de septembre 1992 (620 mm sur le nord de l'Ardèche en 24 h) sur l'Ardèche et le Vaucluse entrainant des crues dévastatrices et quatre décès (Mottet *et al.* 1993).

3.1.3 Les crues de l'Argens, en octobre-novembre 2018

Le troisième événement conséquent de l'année 2018 s'est produit dans le département du Var entre le 29 octobre et le 2 novembre. Comme les deux événements précédents, il se caractérise par une pluviométrie relativement importante, observée cette fois sur le bassin versant de l'Argens (Figure 3.3), avec près 255 mm de précipitations cumulées en 4 jours, dont des cumuls atteignant 130 mm dans le secteur de Draguignan entre le 29 et le 30 octobre, ainsi que 162 mm sur la journée du 31 octobre à Entrecasteaux (pluiesextremes.meteo.fr).

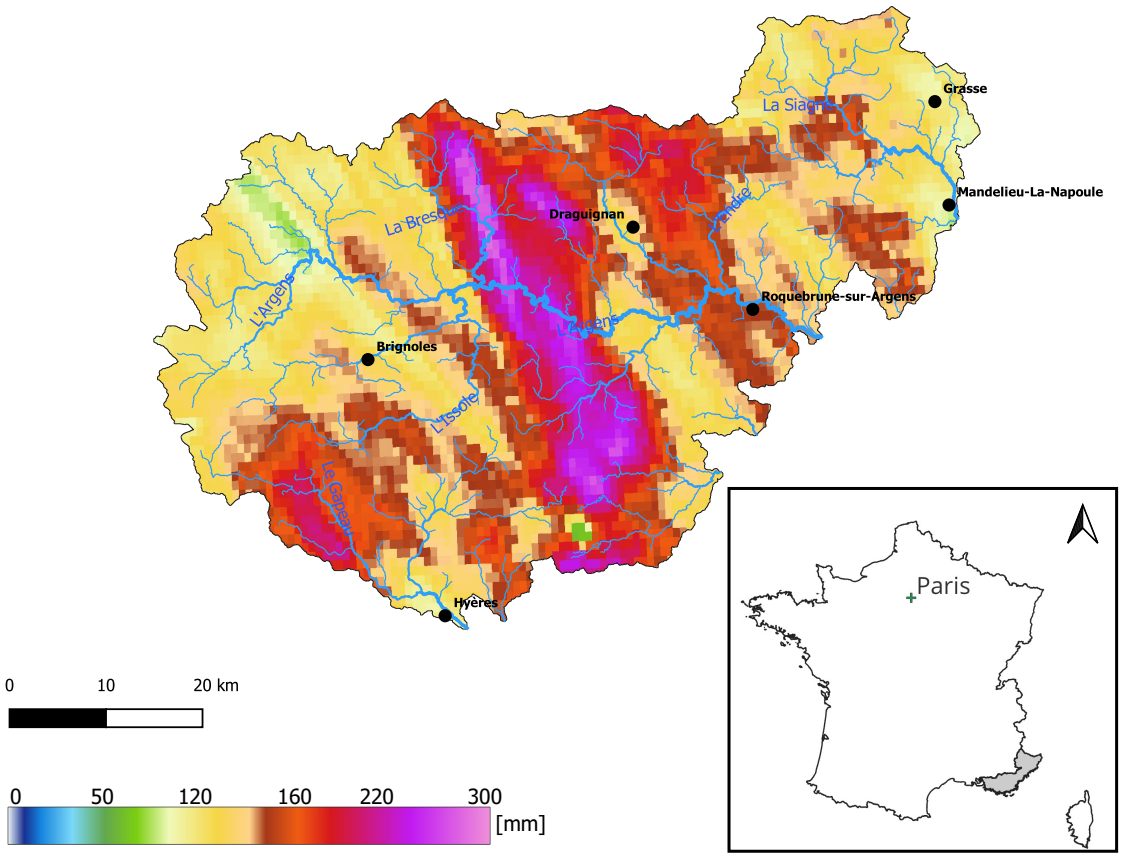

FIGURE 3.3 – *Cumul de pluie (lame d'eau ANTILOPE J+1) du 29 octobre au 2 novembre 2018*

Cet épisode méditerranéen reste d'une ampleur moindre par rapport à celui survenu dans l'Aude quelques semaines auparavant, avec un mort (Dantec et Pipien 2019) et relativement peu de dégâts matériels à déplorer.

3.2 Données disponibles pour l'étude des trois événements

De nombreuses données de différents types (pluviométrique, hydrologique, hydraulique, d'impacts...) sont disponibles pour les événements décrits ci-dessus. Ces données s'avèrent d'une richesse et d'un niveau de détail variables selon l'intensité de l'événement et de ses conséquences en termes humains et matériels. Elles sont décrites dans les paragraphes qui suivent.

3.2.1 Observations pluviométriques disponibles

Le premier type de données utilisables concerne la pluviométrie. Comme cela a été décrit dans la Section 2.4, il existe deux lames d'eau, PANTHERE et ANTILOPE (ainsi qu'ANTILOPE J+1). ANTILOPE J+1 intégrant plus de données des pluviomètres, en particulier les informations des pluviomètres qui ne sont pas collectées en temps réel, soit environ 40% de postes pluviométriques supplémentaires (Laurantin 2008), c'est cette dernière estimation qui est à priori la plus fiable. C'est celle-ci qui sera utilisée, par la suite, pour l'évaluation des produits de prévision d'ensemble expérimentaux.

3.2.2 Produits de prévision d'ensemble des précipitations basés sur le modèle AROME-FR

Les prévisions de précipitations qui vont être évaluées par la suite sont toutes basées sur le modèle AROME-FR développé par Météo-France (Seity *et al.* 2011, Auger *et al.* 2015, Brousseau *et al.* 2016). AROME-FR est un modèle opérationnel à aire limitée qui fournit des prévisions météorologiques déterministes jusqu'à deux jours à l'avance. Sa grande résolution horizontale de 1.3 km (Météo-France 2015) permet de résoudre explicitement la convection profonde, ce qui est bien adapté à la prévision de fortes précipitations. Plus de détails sur ce produit opérationnel ont été donnés dans la Section 2.4.1.

Trois différents produits de prévision d'ensemble à court terme, basés sur AROME-FR, ont été testés dans le cadre de ce travail et sont décrits dans les paragraphes qui suivent. La Figure 3.4 illustre la construction de ces trois produits.

3.2.2.1 AROME-EPS

AROME-EPS (Raynaud et Bouttier 2016, Bouttier *et al.* 2012) correspond à la version ensembliste d'AROME-France. C'est un ensemble de 12 simulations AROME-FR, construites en perturbant chaque prévision AROME-FR de façon à représenter les incertitudes de prévision provenant de l'observation, de l'analyse et de la modélisation. En 2018, AROME-EPS est une prévision d'ensemble mise à jour toutes les 6 heures (soit quatre fois par jour : à 3 h, 9 h, 15 h et 21 h UTC) ; il fournit des prévisions sur un horizon de 48 h. Il s'agit du seul produit de prévision d'ensemble opérationnel à Météo-France dont la résolution permet de prendre en compte la convection et est donc compatible avec la prévision des crues soudaines.

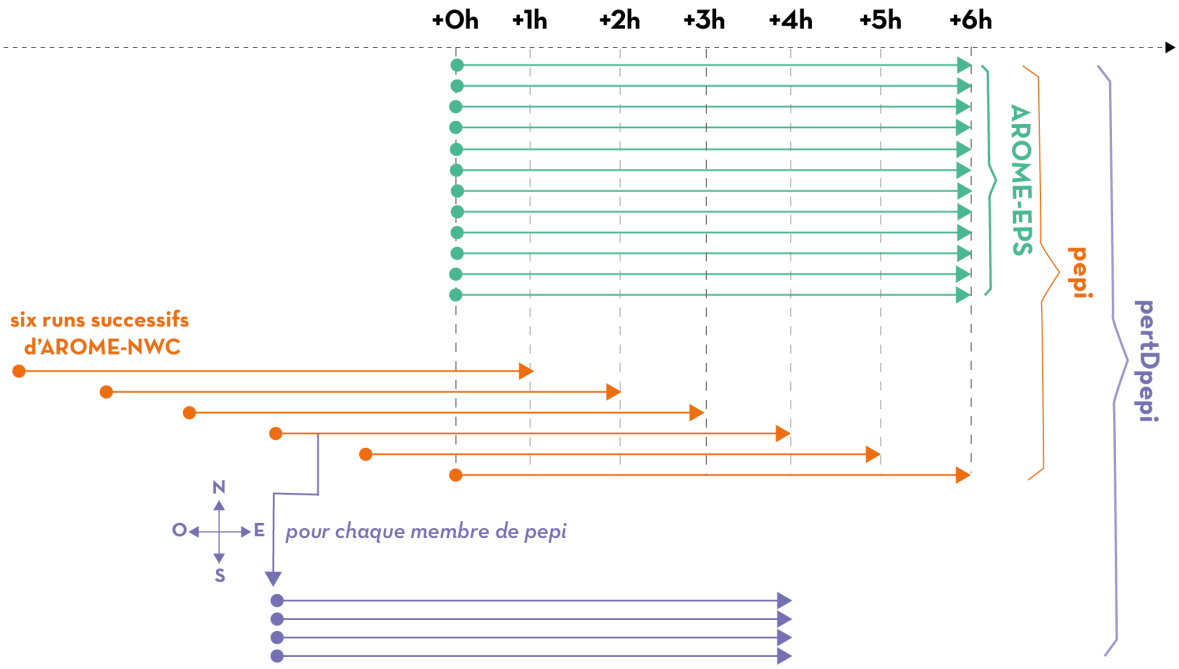

FIGURE 3.4 – *Illustration de la combinaison des prévisions pour former les différents ensembles étudiés*

Il existe un autre produit de prévisions d'ensemble, ARPEGE, qui est plus ancien, mais peu adapté à l'objectif de prévision immédiate sur de petits bassins versants, ARPEGE n'étant pas conçu pour représenter la convection.

3.2.2.2 Pepi

Pepi est une combinaison d'AROME-EPS et d'AROME-PI (Auger *et al.* 2015). AROME-PI est une configuration d'AROME-France conçue pour les prévisions immédiates ; elle est mise à jour toutes les heures et fournit des prévisions fraîches jusqu'à un horizon de 6 heures. Afin de prendre en compte les changements météorologiques soudains, AROME-PI est intégré à l'ensemble pepi avec un décalage temporel (Lu *et al.* 2007, Osinski et Bouttier 2018), qui consiste à utiliser les six dernières simulations successives d'AROME-PI au lieu de n'utiliser que la simulation la plus récente. Cette combinaison est illustrée sur la Figure 3.4. Pepi est donc composé de 13 à 18 membres selon l'échéance considérée du fait de l'ajout des simulations successives d'AROME-PI. Ainsi pour l'échéance +1 heure, pepi contient 18 membres (les 12 membres d'AROME-EPS ainsi que les 6 simulations les plus récents d'AROME-PI) et pour l'échéance +6 heures, pepi contient 13 membres (les 12 membres d'AROME-EPS ainsi que la simulation la plus récente d'AROME-PI).

3.2.2.3 PertDpepi

Le dernier ensemble utilisé, appelé pertDpepi, correspond à l'ensemble pepi auquel des perturbations spatiales ont été ajoutées. Des membres supplémentaires sont donc créés en déplaçant les membres originaux de 20 km au nord, au sud, à l'est et à l'ouest, afin de prendre en compte les incertitudes de localisation des précipitations. Le nombre de membres obtenu varie de 65 à 90 en fonction de l'échéance considérée. Ainsi, pour l'échéance +1 heure, pertDpepi contient 90 membres (les 18 membres de pepi ainsi que les 4*18 membres créés par les perturbations spatiales dans les quatre directions cardinales), et pour l'échéance +6 heures, pertDpepi contient 65 membres (les 13 membres de pepi ainsi que les 4*13 membres créés par les perturbations spatiales dans les quatre directions cardinales). Ce produit est basé sur la méthodologie de perturbation proposée par Vincendon *et al.* (2011), qui a néanmoins été appliquée de façon très simplifiée, puisqu'il s'agit d'un produit-test spécifiquement conçu pour cette étude. Deux points particuliers ont été simplifiés :

- Seules les perturbations spatiales ont été considérées, car elles ont souvent le plus fort impact sur la performance des prévisions d'ensemble (Vincendon *et al.* 2011).
- Seuls des décalages fixes de 20 km (sans composante aléatoire) ont été considérés, car ils représentent déjà une incertitude spatiale importante tout en restant réalistes. Un décalage plus important mélangerait l'influence des zones de plaine et de montagne, ce qui pénaliserait fortement l'ensemble en lui faisant perdre du réalisme.

3.2.2.4 Évaluation initiale des prévisions d'ensemble de pluie par Météo-France

Les 3 produits utilisés fournissent des prévisions jusqu'à un horizon de 6 h, et à une résolution de 2,5 km². Ils sont mis à jour toutes les heures afin de faire bénéficier pepi et pertDpepi des mises à jour horaires d'AROME-PI. Les simulations d'AROME-PI et d'AROME-EPS utilisés sont les plus récents disponibles d'après les horaires de mises à jour de chaque produit, sans tenir compte du temps d'attente nécessaire à leur mise à disposition.

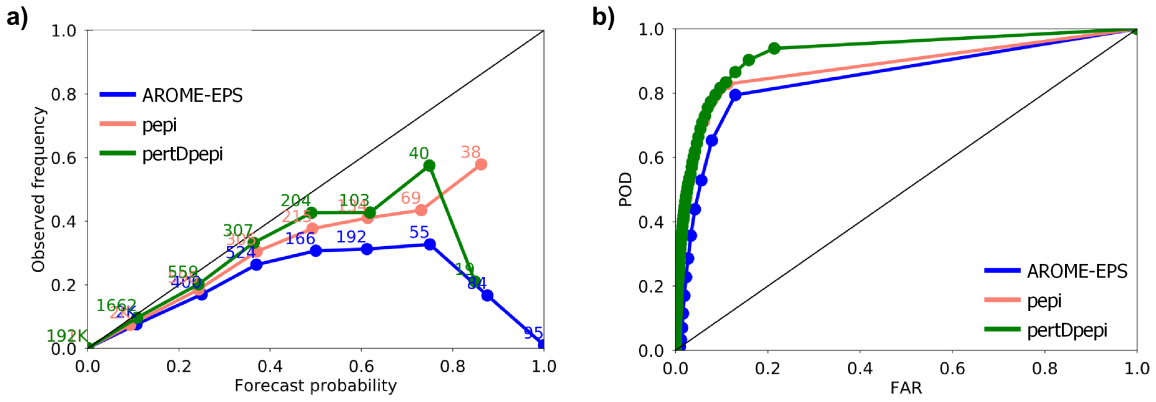

FIGURE 3.5 – *Évaluation des prévisions d'ensemble pour des pluies excédant 5 mm/h a) Diagramme de fiabilité et b) Courbes ROC*

Ces ensembles ont été évalués par Météo-France sur l'ensemble du territoire métropolitain pour l'année 2018, à l'aide de scores classiques et adaptés à l'évaluation de prévision d'ensemble (diagramme de fiabilité et courbe ROC). La Figure 3.5 montre les résultats de la comparaison des trois produits pour des intensités de pluie supérieures à 5 mm/h. Cette intensité a spécialement été choisie pour l'évaluation puisque ce sont des précipitations relativement intenses, pouvant possiblement conduire à des crues soudaines. Sur cette Figure, les scores sont moyennés sur les échéances de 2 h à 6 h. Le diagramme de fiabilité (Figure 3.5 a)) permet de comparer la probabilité d'un événement (en abscisse) à la fréquence observée (en ordonnées). Plus un système de prévision est fiable, et plus le diagramme de fiabilité s'approche de la droite $y = x$. Ainsi, pour la courbe verte, il y a 192 000 prévisions de risque nul, et ainsi de suite. Pour des probabilités de prévision inférieures à 0.5, les trois produits de prévision sont proches de la bissectrice. Pour des probabilités plus grandes, pepi et pertDpepi restent autour de la bissectrice alors qu'AROME-EPS s'en éloigne. La Figure 3.5 b) présente les courbes ROC obtenues pour les trois produits de prévision. PertDpepi est l'ensemble le plus proche de la courbe idéale attendue. Ces résultats préliminaires suggèrent une performance légèrement supérieure du produit pertDpepi, en termes de diagramme de fiabilité ainsi que de courbes ROC.

3.2.3 Données hydrologiques

Deux principaux types de données hydrologiques ont été utilisés : i) les observations de débits issues des enregistrements de débits aux stations hydrométriques de l'Hydro Portail (hydro.eaufrance.fr) ainsi que des relevés post-crue réalisés dans la cadre du programme HyMeX (Hydrological cycle in the Mediterranean eXperiment, hymex.org) ; ii) des valeurs statistiques de débits de crue pour plusieurs périodes de retour, obtenues via la méthode SHYREG (Lavabre *et al.* 2003, Organde *et al.* 2013, Arnaud *et al.* 2013).

Les observations de débits issues des stations hydrométriques sont toutes issues de l'Hydro Portail, qui assure l'archivage des données hydrométriques d'environ 3 000 stations hydrométriques en service sur les cours d'eau métropolitains (Leleu *et al.* 2014). Ces données ont servi au calage du modèle hydrologique ainsi qu'à la validation des données simulées par ce modèle. Les relevés HyMEX ne sont disponibles que pour certains événements particulièrement intenses, et qui ont de ce fait été suivis d'une campagne de relevés post-crue. Cela a notamment été le cas pour le bassin versant de l'Aude après les crues d'octobre 2018 (Lebouc *et al.* 2019a), et pour l'aval des bassins de la Cèze et de l'Ardèche suite aux crues d'aout 2018 (Lebouc *et al.* 2019b). HyMEX est un projet de recherche international qui a notamment permis de collecter des données complémentaires à celles fournies par les stations hydrométriques (Drobinski *et al.* 2014). Les données fournies dans le cadre d'HyMeX sont surtout une manière de compléter les informations parfois incertaines et incomplètes des débits de pointe mesurés par les stations hydrométriques, qui ne couvrent pas les petits cours d'eau, et sont parfois endommagées par les crues ou ont une courbe de tarage non fiable pour les grands débits. Ces données HyMeX sont des estimations des débits de pointe (pas d'hydrogrammes complets), obtenues à partir de mesures ponctuelles de hauteurs d'eau maximales et de sections d'écoulement correspondantes, à partir des laisses de crues pour un

événement spécifique. La hauteur mesurée est ensuite transformée en débit, le plus souvent, à dire d'expert, via l'application de la formule de Manning-Strickler en régime uniforme, qui est une méthode souvent qualifiée de Slope-Area method (Lumbroso et Gaume 2012).

La méthode SHYREG quant à elle a été développée pour constituer une base de données des quantiles de débits de période de retour 2 à 1 000 ans, couvrant près de 140 000 exutoires sur la France métropolitaine. La superficie des bassins versants varie entre 2 et 10 000 km², avec une majorité entre 5 et 5 000 km². Par conséquent, lorsque SHYREG est appliquée à des bassins de superficie inférieure à 5 km² ou supérieure à 5 000 km², les résultats sont de moindre qualité. Même si la méthode se veut régionale, il n'y a pas de prise en compte des spécificités géologiques comme les karsts, ni des aménagements hydrauliques ou encore de l'urbanisation à l'échelle du bassin. Dans ce travail, les débits de la méthode SHYREG seront utilisés comme valeurs de seuils de débits pour l'évaluation des prévisions d'ensemble. Ainsi, cela permettra d'avoir des valeurs de seuils de débits relativement homogènes du point de vue de l'intensité des crues, pour des unités de modélisation ayant des surfaces drainées très différentes.

3.2.4 Données d'occupation des sols

Afin de localiser les potentielles zones d'impacts, il est nécessaire d'avoir une description précise des zones urbanisées ou aménagées, dans les secteurs d'étude. La BD TOPO produite par l'institut national de l'information géographique et forestière (geoservices.ign.fr) couvre l'ensemble du territoire métropolitain avec une description des éléments bâtis et paysagers sous une forme vectorielle de précision métrique. Une mise à jour trimestrielle est réalisée sur les différents items. C'est plus particulièrement la couche "bâti", réalisée à partir des plans cadastraux, qui sera utilisée par la suite. Pour l'étude des événements survenus durant l'année 2018, et plus particulièrement sur celui d'octobre, le choix s'est logiquement porté sur la version de l'année 2018 de la BD TOPO (mise à jour en septembre 2018).

3.3 Les modèles utilisés

3.3.1 Le modèle hydrologique Cinecar

Le modèle hydrologique pluie-débit Cinecar (Le Bihan 2016, Naulin *et al.* 2013, Versini 2007) a été utilisé dans ce qui suit pour réaliser l'ensemble des simulations pluie-débit, ainsi que les prévisions d'ensemble des débits à partir des différents produits de prévision de pluie.

3.3.1.1 Principe de fonctionnement

Le modèle hydrologique Cinecar est un modèle développé pour simuler la dynamique des crues soudaines à l'échelle régionale, et spécifiquement adapté à la simulation des crues-éclair sur les bassins non instrumentés. Ce modèle a déjà été utilisé dans des travaux de thèse antérieurs sur les crues de septembre 2002, de septembre 2005, et d'octobre-novembre 2008 dans le Gard (Le Bihan 2016, Naulin *et al.* 2013, Versini 2007), et sur la crue de juin 2010 de l'Argens dans le Var (Le Bihan 2016).

Chaque bassin versant est représenté par un réseau de biefs de profils en travers rectangulaires pour le lit mineur et le lit majeur. À chaque bief sont raccordés deux versants en rives gauches et droites, dont la géométrie est simplifiée : plans inclinés de forme rectangulaire.

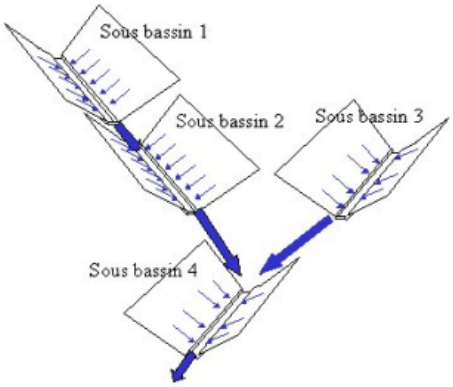

FIGURE 3.6 – *Représentation schématique d'un bassin versant dans Cinecar. Chaque sous-bassin est constitué d'un thalweg de section rectangulaire et d'un ou deux versants rectangulaires (Gaume 2002)*

La transformation pluie-débit s'effectue en deux étapes avec d'abord la transformation de la pluie brute en pluie efficace puis la propagation des pluies efficaces vers l'aval. Le passage de la pluie brute à la pluie efficace se fait via une fonction de production qui est ici le modèle conceptuel du Soil Conservation Service – Curve Number (SCS-CN). Ce modèle décrit la saturation progressive des sols et avec l'augmentation du ratio pluie nette - pluie brute. La fonction est basée sur les deux équations suivantes :

$$\begin{cases} \mathrm{P} = I_a + F + Q \\ \frac{Q}{P - I_a} = \frac{F}{S} \end{cases} . \tag{3.1}$$

Où F est le déficit d'écoulement cumulé, Q le ruissellement direct cumulé, S le potentiel maximal de rétention, P la pluie brute cumulée et I_a les pertes initiales. En combinant les deux équations, il est possible d'estimer l'écoulement direct (volume cumulé de pluie nette) Q :

$$Q = \frac{(P - I_a)^2}{P - I_a + S} \tag{3.2}$$

Et en considérant que les pertes initiales sont une fraction du potentiel de rétention : $I_a = \lambda S$ et en considérant que $\lambda \approx 0.2$:

$$Q = \frac{(P - 0.2S)^2}{P + 0.8S} \tag{3.3}$$

La pluie efficace peut donc ainsi être déterminée sur chaque versant à partir de la valeur du potentiel maximal de rétention S. Cette pluie est ensuite propagée dans l'ensemble des bassins versants par la fonction de transfert qui s'applique sur les versants et dans les biefs. En pratique, la valeur de S est déterminée à partir d'un paramètre, le Curve Number, CN, qui est défini comme :

$$S = \frac{1000}{CN} - 10 \tag{3.4}$$

Ce paramètre varie selon la nature et l'occupation du sol : type de substratum, imperméabilisation, occupation ainsi que l'antériorité pluviométrique (paramètre proche de l'humidité). Les valeurs du CN sont autour de 60 à 90 pour des sols agricoles et peu végétalisés, et plutôt au-delà de 50 pour des surfaces ayant des sous-sols avec une grande capacité de stockage (comme les karsts). C'est principalement la nature du substrat qui influence la valeur du CN, ainsi une couverture végétale dense ne modifie pas significativement ce paramètre (Gaume *et al.* 2004). Le CN est surtout considéré comme un paramètre de calage pour le modèle Cinecar.

Le principal avantage du modèle SCS-CN est le fait qu'il repose sur un unique paramètre (le Curve Number). Néanmoins, cette méthode présente des inconvénients, dont le fait qu'elle est plus adaptée pour simuler des cours d'eau avec un débit de base négligeable, ou en encore pour simuler des crues d'intensité significative (Michel *et al.* 2005, Ponce et Hawkins 1996). Ce dernier inconvénient n'en est pas vraiment un dans le cas des trois cas d'étude simulés ici puisqu'il s'agit de reproduire des événements relativement exceptionnels.

Le transfert de l'eau sur les versants et la propagation dans le réseau hydrographique se font par la résolution en une dimension de l'équation d'onde cinématique (Le Bihan 2016). Cette simplification des équations de Saint-Venant néglige les termes inertiels, en particulier l'accélération et la décélération. En raison de la géométrie simplifiée des versants et des biefs, ce modèle est relativement simple puisqu'il ne demande comme paramètres que les coefficients de rugosité des versants et des biefs (coefficients de Manning-Strickler) et la largeur des biefs (obtenue par relevé topographique sur le terrain ou estimée via l'ordre de Strahler des différents biefs).

3.3.1.2 Découpage des trois territoires d'étude

Avant de pouvoir exécuter la modélisation pluie-débit, il faut extraire, en partant d'un Modèle Numérique de Terrain (MNT), les caractéristiques géométriques du bassin versant, de ses sous-bassins et de son réseau hydrographique. Ces étapes sont réalisées à l'aide d'un script de préparation du modèle, qui s'appuie sur les librairies TauDEM (Yıldırım *et al.* 2015) pour réaliser le découpage hydrologique à partir du MNT.

Il est ainsi possible d'extraire les directions de drainage, de définir des exutoires à partir de la Base Nationale des Bassins Versants, BNBV, ou encore de créer un réseau hydrographique et de définir une surface drainée des sous-bassins (selon une valeur minimale qu'il est possible de fixer). Pour chaque sous-bassin découpé, un exutoire est créé. Le positionnement des limites des sous-bassins versants est défini par :

- Une surface minimale de drainage à atteindre.
- Chaque confluence, embouchure et station hydrométrique référencée.

Ce travail de préparation du modèle a été réalisé antérieurement à la thèse dans le cas de l'Aude (Hocini *et al.* 2021), et a été reproduit pour les besoins de la thèse pour le Var et l'Ardèche. La Figure 3.7 illustre les découpages obtenus en 1 174 sous-bassins pour l'Aude (5 160 km²), 2 358 pour l'Ardèche (3 080 km²) et 1 093 pour le Var (1 960 km²). Pour l'événement de l'Ardèche, le choix a été fait de retirer le couloir rhodanien du secteur d'étude qui se retrouve donc découpé en deux parties (Cèze/Ardèche et Drôme). Ce choix est motivé par le fait que le Rhône sort du cadre des petits cours d'eau réactifs aux pluies intenses.

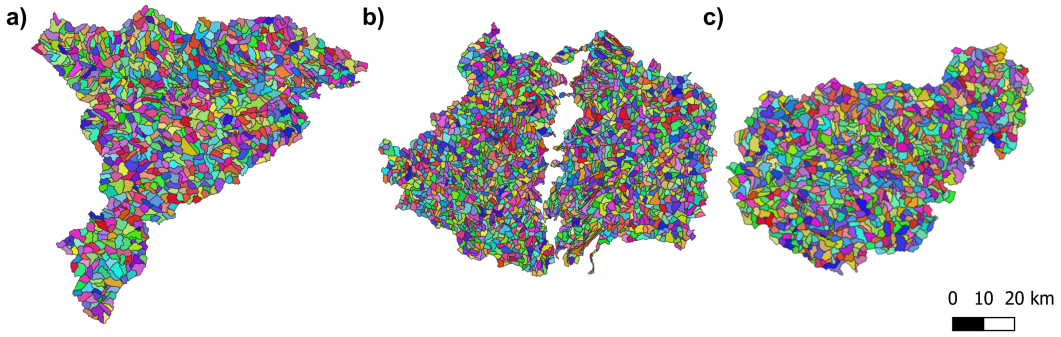

FIGURE 3.7 – *Découpage en sous-bassins réalisé avec TauDEM où chaque couleur correspond à un sous-bassin : a) pour l'Aude, b) pour l'Ardèche et c) pour le Var*

3.3.1.3 Calage sur les trois événements étudiés

Les valeurs par défaut des Strickler, qui sont de 15 sur les versants et 20 dans les cours d'eau, ont été conservées ici et n'ont pas été calées. En effet, des études précédentes (Naulin *et al.* 2013) et des études post-événements (Lebouc *et al.* 2019a;b) justifient l'utilisation de ces valeurs.

Comme cela a été évoqué dans la Section 3.3.1.1, le Curve Number CN est utilisé comme paramètre de rugosité. Tout d'abord, plusieurs simulations spatialement homogènes, avec des CN dans une gamme de valeurs entre 35 (pour des sols avec un fort stockage) et 100 (pour un écoulement intégral), ont été testés. Puis une carte spatialisée ajustée est produite en prenant appui sur des cartes pédologiques et/ou géologiques. La carte correspondante pour l'événement de l'Aude 2018 est présentée sur la Figure 3.8.

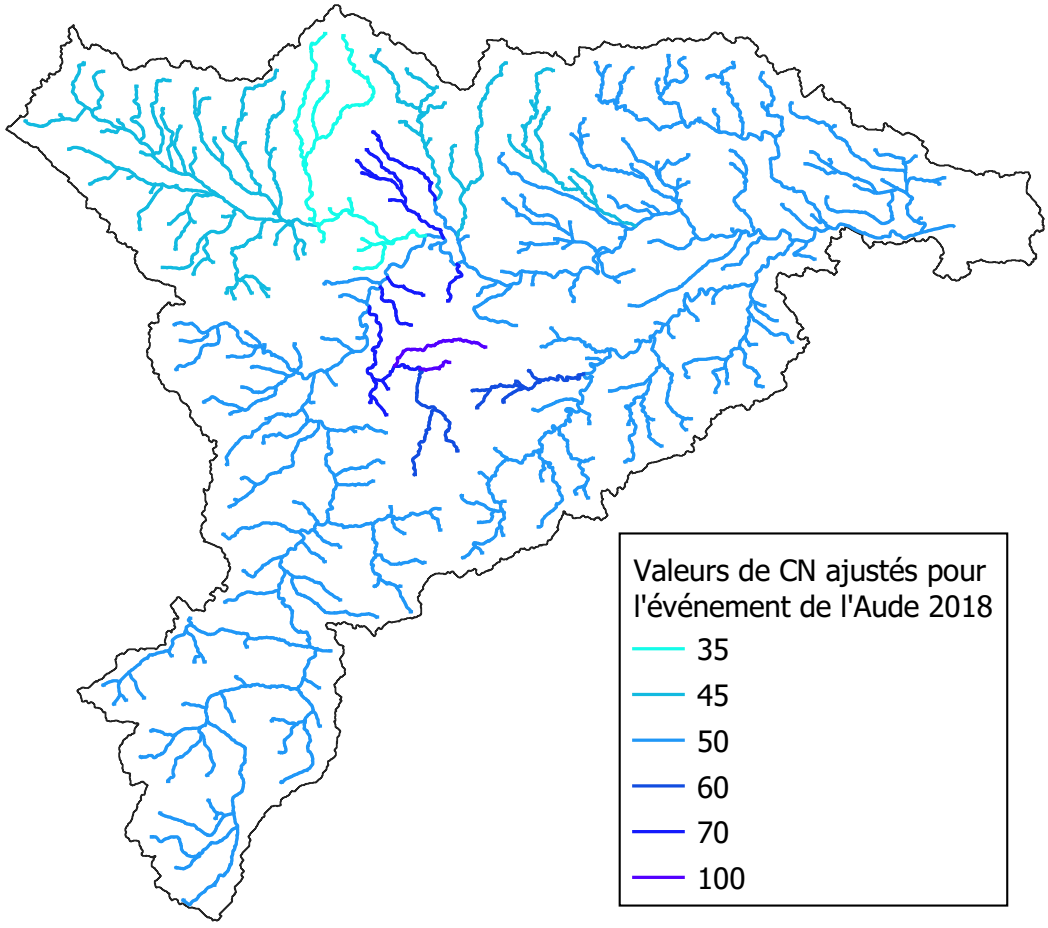

FIGURE 3.8 – *Valeurs de CN retenues pour reproduire les débits pour l'événement de l'Aude 2018*

La Figure 3.9 illustre les écarts entre débits spécifiques observés, aux stations de mesure (en rouge) et aux relevés HyMEX (en noir), et simulés avec le modèle Cinecar ajusté pour les 3 crues considérées.

Comme cela est visible sur la Figure 3.9, la majorité des écarts se trouve dans la fourchette +/- 20% d'écart en débit spécifique, ce qui n'est pas surprenant puisqu'il y a un ajustement des CN pour chacun des événements.

Il faut noter que dans le cas de l'événement du Var, il n'y a pas de données HyMEx disponibles, car aucune campagne de terrain n'a été réalisée à la suite de ces crues. Le calage se fait donc uniquement sur les stations hydrométriques disponibles.

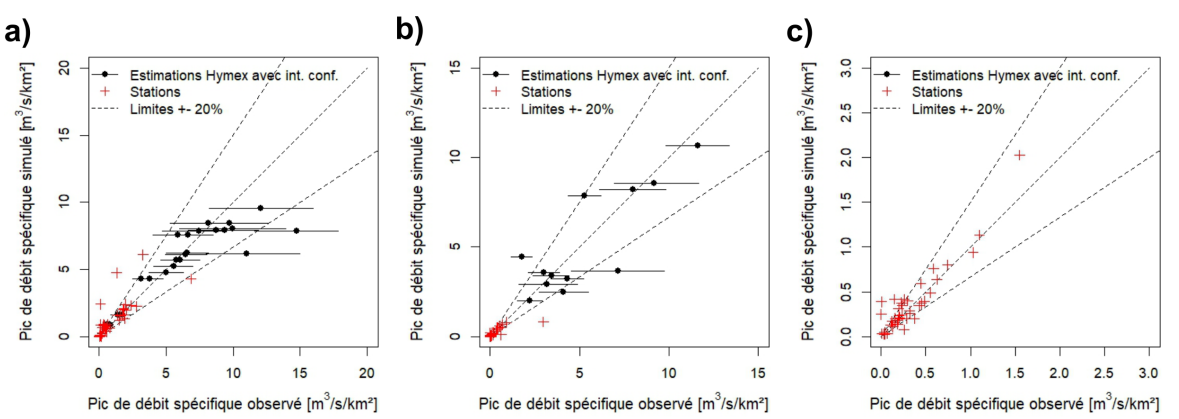

FIGURE 3.9 – *Écarts entre les débits spécifiques observés ou relevés avec les débits spécifiques simulés a) pour l'événement de l'Aude, b) pour l'événement de l'Ardèche et c) pour l'événement du Var*

3.3.2 Catalogue d'emprises inondées utilisé pour la modélisation des impacts

Dans le cadre de cette thèse, les simulations d'impacts ont été réalisées à partir de courbes débit/impact estimées sur chaque tronçon de cours d'eau. Ces courbes ont été obtenues à partir d'un catalogue de scénarios d'inondation réalisé par Nicolle (2021). Ce catalogue représente les zones inondées estimées pour 8 débits de pointe différents de période de retour 2 à 1000 ans correspondant aux quantiles de débits estimés par la méthode SHYREG (Aubert *et al.* 2014), et a été produit en appliquant le modèle hydraulique 2D Floodos (Davy *et al.* 2017) sur le MNT du RGE Alti à 5 m de résolution. En se plaçant dans un objectif de prévision des impacts, l'utilisation d'un catalogue des zones inondées pour la création de courbe débit/impact a pour avantage de pouvoir s'affranchir de la modélisation hydraulique et de pouvoir directement, à partir d'un débit simulé, obtenir les impacts attendus pour chaque tronçon de cours d'eau.

Synthèse

Dans ce Chapitre, les événements pluviométriques marquants de l'année 2018 qui seront par la suite étudiés dans la thèse, ainsi que les données disponibles et les différents modèles utilisés pour l'étude de ces trois événements sont présentés.

Concernant les événements sélectionnés, celui de l'Aude est particulièrement marquant du fait de son intensité et de ses conséquences dramatiques. Les deux autres études de cas dans l'Ardèche et le Var correspondent à des événements également intenses, mais qui s'avèrent néanmoins moins exceptionnels, et également de moindre ampleur en termes de dégâts occasionnés.

Les trois produits de prévision d'ensemble qui seront étudiés par la suite ont été créés à partir des différentes prévisions issues du modèle AROME-FR. À partir de l'ensemble d'AROME-EPS comportant 12 membres, 1 à 6 réseaux issus d'AROME-PI sont ajoutés pour obtenir l'ensemble pepi. Des perturbations spatiales dans les quatre directions cardinales sont ensuite appliquées à tous les membres de pepi pour obtenir l'ensemble pertDpepi.

Pour obtenir des prévisions hydrologiques d'ensemble, le choix s'est porté sur le modèle pluie-débit Cinecar, déjà éprouvé sur des événements méditerranéens intenses, et qui a aussi l'avantage de demander peu de paramètres et de nécessiter des temps de calcul très courts.

Enfin, la prévision des impacts s'appuiera sur un catalogue de 8 scenarios d'inondations, produits avec le modèle hydraulique Floodos et couvrant des périodes de retour des débits de 2 à 1000 ans.

Finalement, un grand nombre de données observées sont disponibles pour l'évaluation des prévisions d'ensemble sur les événements sélectionnés, comme la lame d'eau radar Antilope ré-analysée à J+1 avec les données des pluviomètres disponibles, les données hydrologiques détaillées issues des enquêtes HyMeX, et enfin les données de sinistralité et d'occupation des sols permettant d'évaluer l'importance des impacts.

Chapitre 4

Évaluation des produits de prévision d'ensembles des pluies sur la base des pluies observées

Objectifs du Chapitre :

Dans ce Chapitre 4, la première étape de la méthodologie d'évaluation, concernant la fiabilité des ensembles de pluie, est mise en œuvre.

Cette étape (Figure 4.1, dans le cadre corail) décrite dans le Chapitre 2 inclut l'étude de la répartition temporelle des précipitations observées et prévues avec des hyétogrammes, puis l'étude de leur répartition spatiale à partir de cartes de précipitations observées et prévues, et enfin une évaluation avec des scores classiques évaluant finesse et fiabilité des ensembles : le diagramme de rang et le CRPS (Continuous Rank Probability Score).

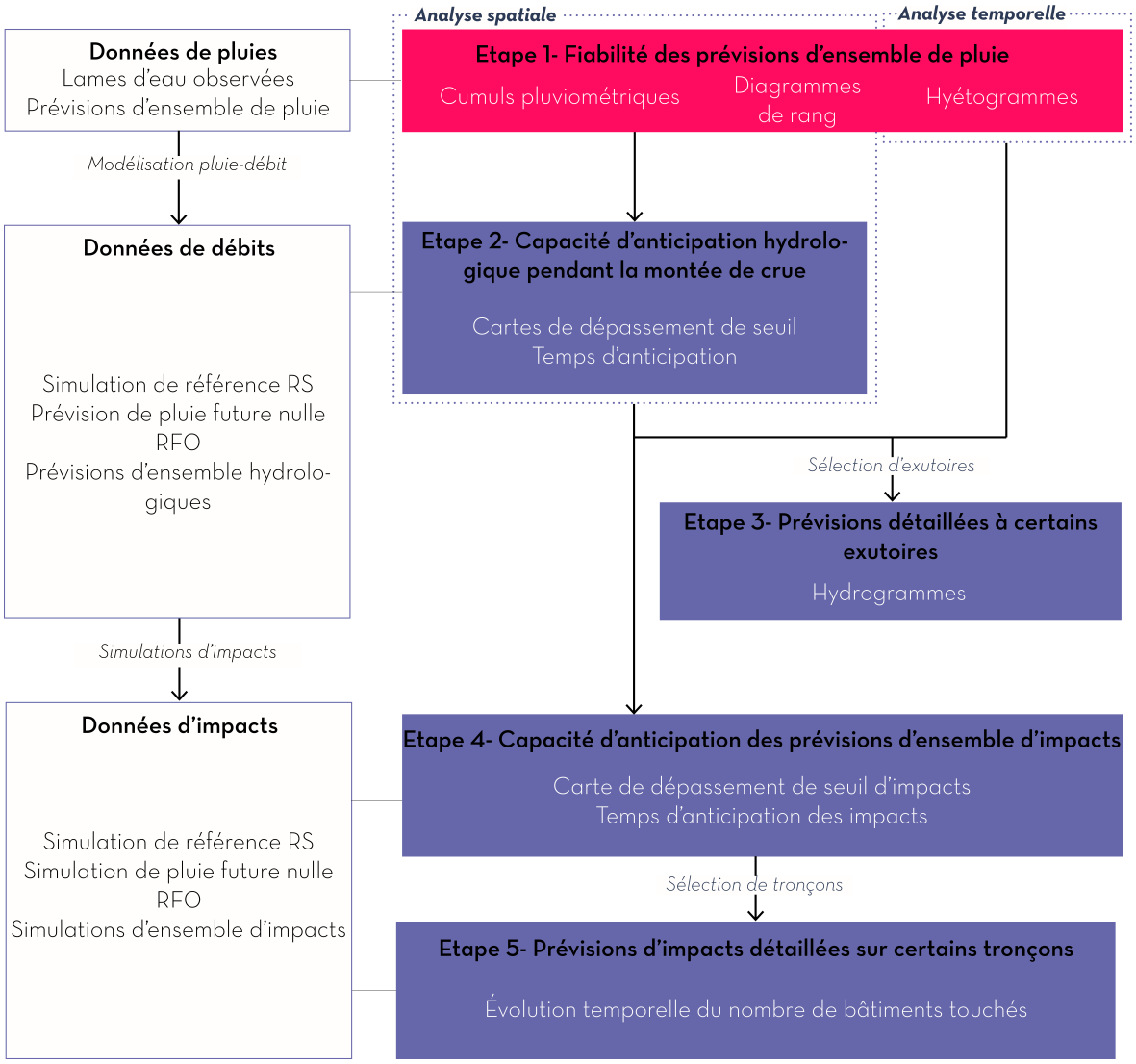

FIGURE 4.1 – *Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, est représentée l'étape d'évaluation menée dans le Chapitre 4.*

4.1 Méthodologie retenue pour l'évaluation événementielle des prévisions d'ensemble de pluie

4.1.1 Principe général : une évaluation en 3 étapes

Pour rappel, les ensembles de prévision de pluie ont déjà été évalués par Météo-France (Section 3.2.2.4) à l'échelle de l'année 2018 sur tout le territoire métropolitain via des diagrammes de fiabilité et des courbes ROC. Dans ce Chapitre, le but n'est pas de refaire une évaluation telle que celle qui a déjà été menée par Météo-France, mais d'évaluer les ensembles de prévisions de pluie disponibles pour chacun des événements sélectionnés (Aude, Ardèche et Var pour l'année 2018).

Trois aspects différents ont par conséquent été considérés pour une comparaison des pluviométries observées et prévues :

- Des hyétogrammes d'intensité moyenne des précipitations sont d'abord tracés pour la prévision d'ensemble évaluée et pour toutes les échéances de prévision. Ces hyétogrammes sont utilisés pour évaluer si en moyenne la temporalité et l'intensité des précipitations sont bien capturées par le produit de prévision. Cette étape permet également de sélectionner une fenêtre temporelle réduite où des intensités significatives sont prévues ou mesurées sur laquelle la suite de l'évaluation sera focalisée.

- Des cartes de précipitations moyennes prévues et observées sont réalisées pour évaluer si les zones où de fortes accumulations de pluie ont été prévues et se sont effectivement produites coïncident. Une carte est faite par produit de prévision, par délai de prévision et par percentile de l'ensemble. Ces cartes comparent la distribution spatiale des précipitations obtenue à partir de l'agrégation (cumul) de toutes les prévisions fournies pendant l'événement. Ce sont donc les totaux moyens de précipitations qui sont d'abord évalués. Ces cartes permettent également de délimiter les zones affectées par des totaux élevés de précipitations mesurées ou prévues.

- Une évaluation sur des scores classiques de la prévision d'ensemble à partir de la fenêtre spatio-temporelle définie lors des deux points précédents. Les scores retenus pour cette évaluation des prévisions d'ensemble sont le diagramme de rang (ou diagramme de Talagrand) pour la fiabilité, et le CRPS pour la finesse et la fiabilité. Ces deux scores sont présentés dans les sections qui suivent. L'objectif est de détecter ou de confirmer rapidement les éventuels biais systématiques ou le manque de variabilité des produits pluviométriques de l'ensemble des prévisions.

4.1.2 Le Diagramme de rang (ou diagramme de Talagrand)

Le diagramme de rang est un critère qualitatif qui évalue la fiabilité des prévisions d'ensemble. Il permet de rendre compte de la dispersion globale de la prévision en comparant l'ensemble des valeurs prises par les membres, avec celle prise par l'observation (Talagrand et Vautard 1997). Le diagramme de rang indique la fréquence des positions prises par la valeur de l'observation dans le classement de toutes les valeurs des membres. La Figure 4.2 présente les différents types de diagrammes de rang qu'il est possible d'obtenir.

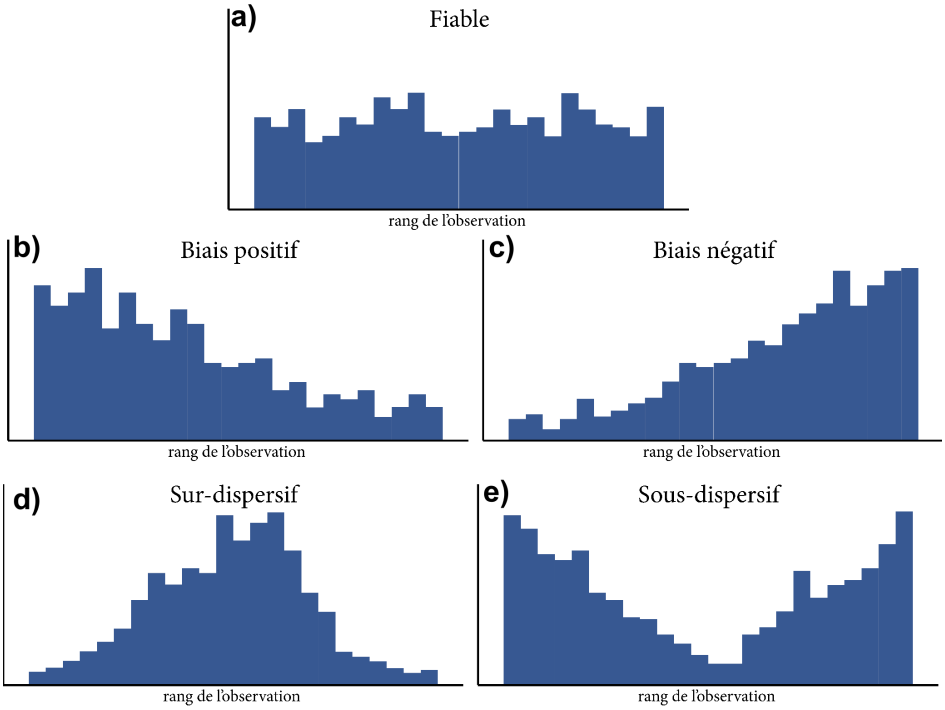

FIGURE 4.2 – *Illustration des différentes formes caractéristiques des diagrammes de rang*

Ici, dans le cas de l'évaluation de prévisions de pluies, le diagramme considéré comme le plus fiable possible Hamill (2001) est un histogramme uniforme correspondant à la Figure 4.2 a). Si le diagramme n'est pas uniforme, la forme qu'il prend renseigne sur les défauts de l'ensemble évalué. Ainsi :

- Si le diagramme est asymétrique, l'ensemble considéré contient un biais. Le biais de l'ensemble peut être positif si de nombreuses observations se situent parmi les plus petites valeurs de l'ensemble (Figure 4.2 b)), cela traduit une tendance à la surestimation. Le biais peut aussi être négatif, illustrant une tendance à la sous-estimation si la majorité des observations se trouvent dans les valeurs supérieures de l'ensemble (Figure 4.2 c)).
- Si le diagramme est en forme de dôme, il est alors dit sur-dispersif : l'ensemble est trop dispersé, la plupart des observations sont près du centre de l'ensemble (Figure 4.2 d)).
- À l'inverse, si le diagramme est en forme de U, il est dit sous-dispersif : l'ensemble est trop peu dispersé et de nombreuses observations correspondent aux valeurs extrêmes des ensembles qui ne sont pas assez représentées (Figure 4.2 e)).

Le diagramme de rang est un critère très agrégatif qui peut masquer des disparités spatiales et/ou temporelles importantes. Un diagramme de rang peut être visuellement uniforme, mais finalement, avoir des valeurs maximales sous-estimées et des valeurs minimales également sous-estimées.

4.1.3 Le CRPS (Continuous Rank Probability Score)

Le CRPS est un score de vérification évaluant la performance globale de prévisions d'ensemble en mesurant la distance quadratique entre la fonction de répartition de la prévision et celle de l'observation (Hersbach 2000). Il existe une solution analytique uniquement pour les distributions normales (ce qui n'est pas forcément le cas pour les prévisions de pluies). Dans le cas des prévisions de pluies, la distribution (empirique) est obtenue en classant les membres par ordre croissant (décomposition de Hersbach). Pour chaque membre i d'un ensemble de valeurs, le $CRPS$ associé, $CRPS_i$, est calculé par l'équation 4.1 :

$$CRPS_i(F_i, y_i) = \int_{-\infty}^{\infty} (F_i(z) - H(z - y_i))^2 dz \tag{4.1}$$

Avec F_i, la fonction de répartition prévue, H, la fonction d'Heaviside telle que $H = 1$ sur $\mathbb{R}+$ et $H = 0$ sinon, et $H(z - y_i)$ qui est assimilée à la fonction de répartition de la valeur observée y_i.

Le CRPS peut être vu comme une moyenne pondérée des écarts quadratiques entre les membres de l'ensemble et la valeur observée, son unité est donc la même que celle des prévisions et de l'observation (équation 4.2).

$$CRPS = \frac{1}{p} \sum_{i=1}^{p} CRPS_i \tag{4.2}$$

La valeur du CRPS reflète l'aire située entre les deux fonctions de répartition comme le montre la Figure 4.3. Pour une prévision parfaite, le CRPS est nul.

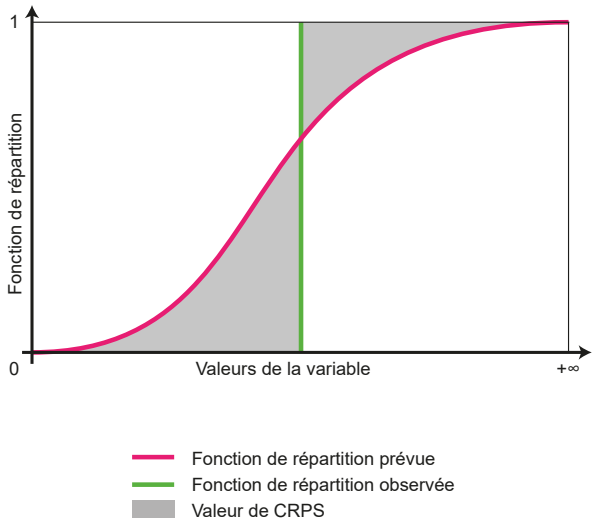

FIGURE 4.3 – *Illustration du calcul du CRPS avec les fonctions de répartition observée (en vert) et prévue (en rose) d'une variable*

Le CRPS permet d'évaluer à la fois la fiabilité des ensembles, mais aussi la finesse, c'est-à-dire la dispersion de l'ensemble par rapport à l'observation (Bourgin 2014). C'est un critère quantitatif et plus synthétique, complémentaire du diagramme de rang, qui est plus un outil de diagnostic qualitatif.

Cependant, le CRPS est un critère qui est sensible aux biais liés à la taille de l'ensemble évalué. Des méthodes de dé-biaisement ont été proposées dans la littérature, en particulier par Ferro *et al.*. Ce point sera étudié plus en détail dans la Section 4.3.3.

Dans le cas des pluies utilisées ici, le CRPS est calculé sur chaque pixel. Il est le plus souvent ensuite moyenné afin d'avoir un critère synthétique.

4.2 Comparaison des pluies prévues et observées du point de vue temporel et spatial

Avant de se lancer dans le calcul de scores synthétiques, le choix a été fait de comparer les prévisions aux observations de façon visuelle, en termes de temporalité, d'intensité, de localisation, pour les trois événements étudiés à l'aide de hyétogrammes et de cartes de pluie cumulée.

Cette première étape a également pour objectif de définir des cadres temporels et spatiaux pertinents pour l'évaluation hydrologique. Le cadre temporel doit comprendre à la fois les périodes de fortes intensités observées (soit la lame d'eau choisie qui est ici Antilope J+1) mais également les périodes de pluies intenses prévues par les trois ensembles de prévision. De la même façon, d'un point de vue géographique, il est nécessaire d'inclure les zones de précipitations intenses observées ainsi que les zones de précipitations prévues, sans pour autant ajouter de zones avec des faibles intensités qui i) présentent peu d'intérêt pour l'évaluation de l'événement ; ii) peuvent masquer les principales caractéristiques des ensembles.

4.2.1 Évolution temporelle des pluies

Le premier outil utilisé de comparaison des pluies observées et prévues est un hyétogramme horaire des pluies moyennées sur l'ensemble du bassin versant étudié. Ce hyétogramme permet de vérifier à la fois la précision de la temporalité des prévisions d'ensemble, c'est-à-dire repérer un éventuel retard ou au contraire une avance de la prévision, mais aussi comparer l'intensité prévue par l'ensemble par rapport à celle observée. Enfin, ce type de figure permet également de mettre en avant les caractéristiques propres ou communes aux trois ensembles de prévision évalués, comme la dispersion des membres, et ainsi révéler l'influence de l'ajout de membres via AROME-PI et via la perturbation spatiale.

La Figure 4.4 illustre l'évolution temporelle des précipitations moyennes horaires, mesurées et prévues (pour des échéances de 1 h, 3 h et 6 h) sur l'ensemble du bassin de l'Aude (5 160 km²).

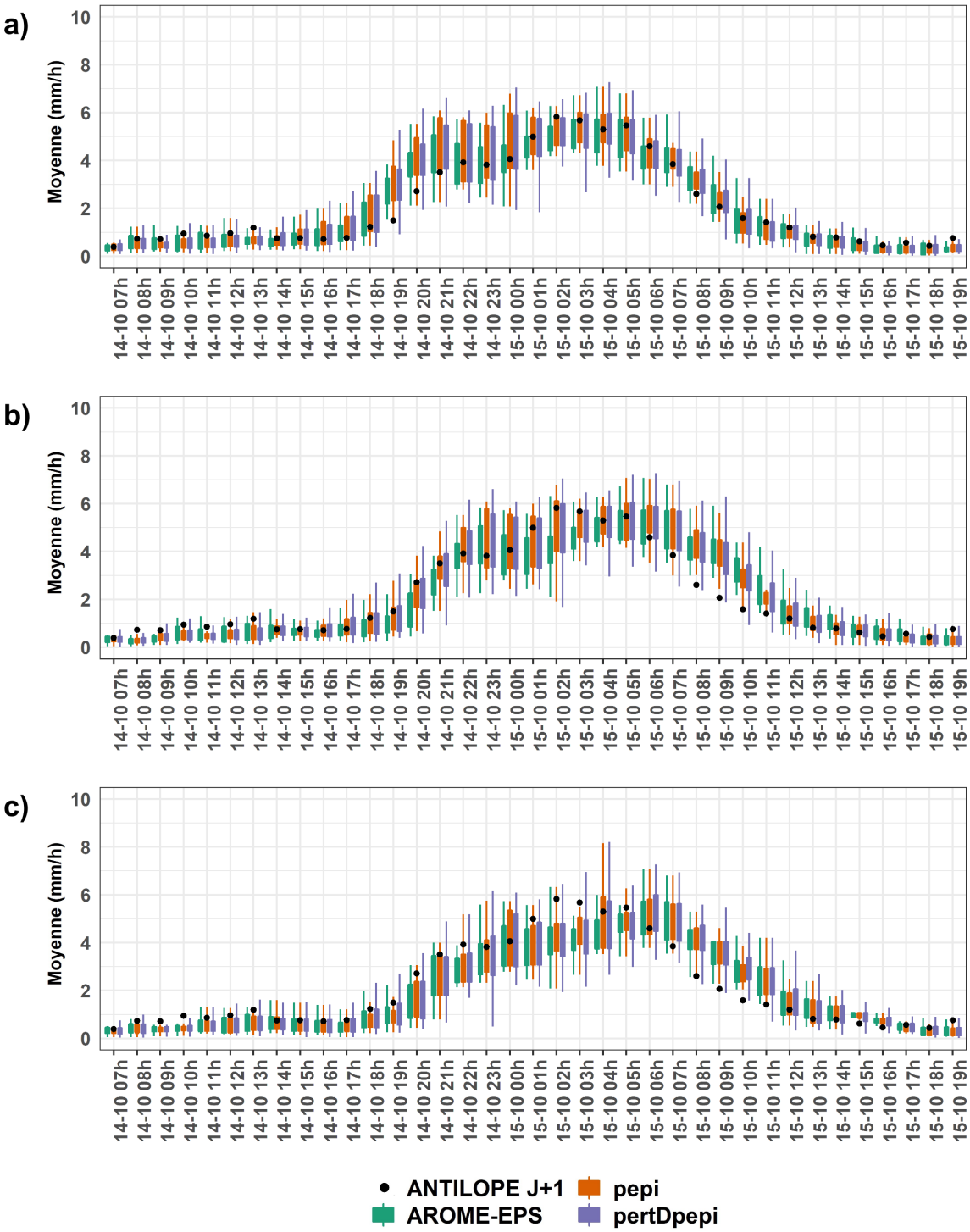

FIGURE 4.4 – *Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement de l'Aude*

La Figure 4.5 illustre l'évolution temporelle des précipitations moyennes horaires, mesurées et prévues (pour des échéances de 1 h, 3 h et 6 h) sur l'ensemble des bassins de l'Ardèche et de la Cèze (3 080 km²).

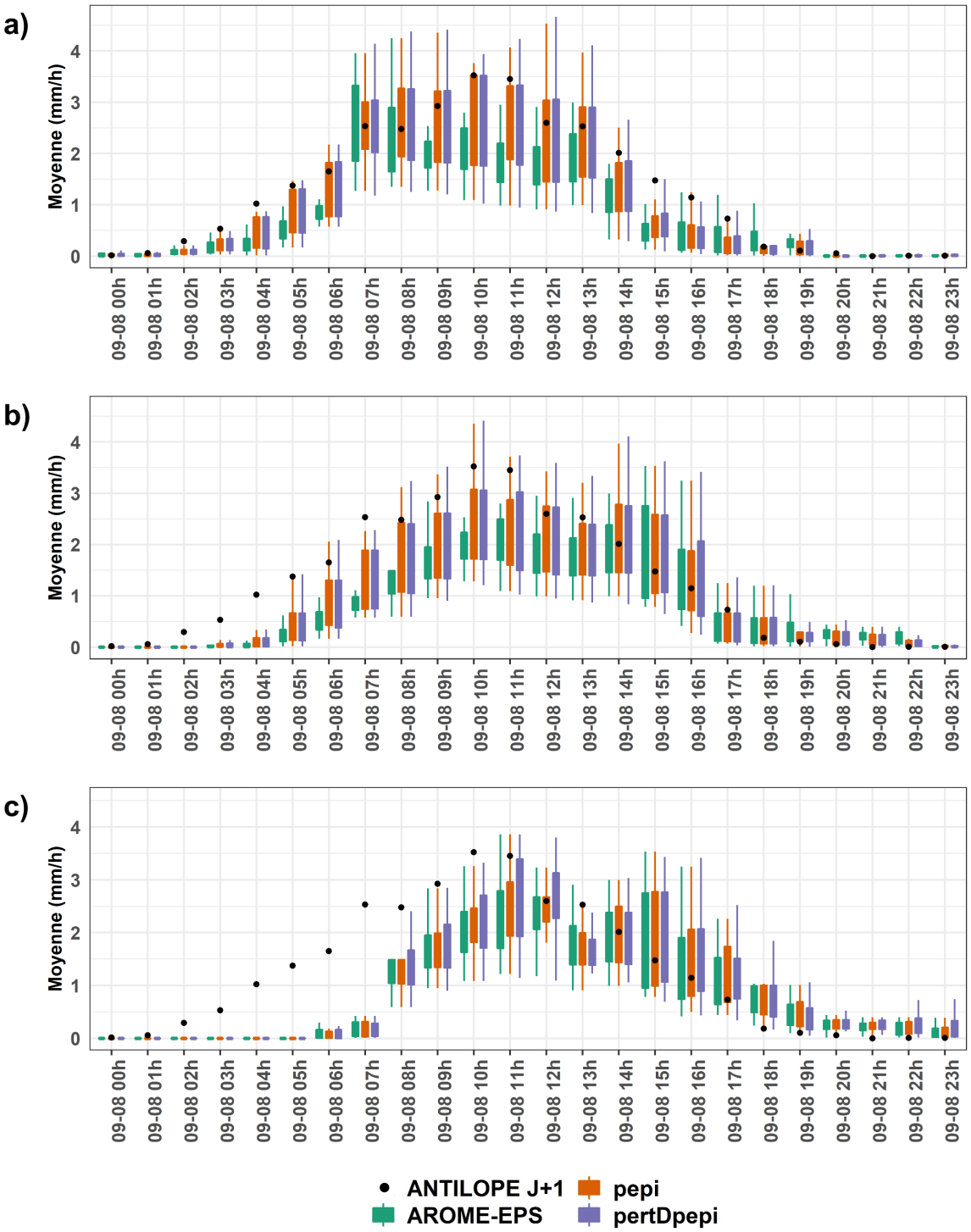

FIGURE 4.5 – *Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement d'août 2018 sur l'Ardèche*

Le cas de l'événement du Var est un légèrement différent, car il se divise en deux pics de précipitations séparés par presque 24 heures sans pluie. Il est donc possible de considérer chaque pic de pluie comme un événement distinct. Les cumuls observés pour les deux pics sont présentés sur la Figure 4.6.

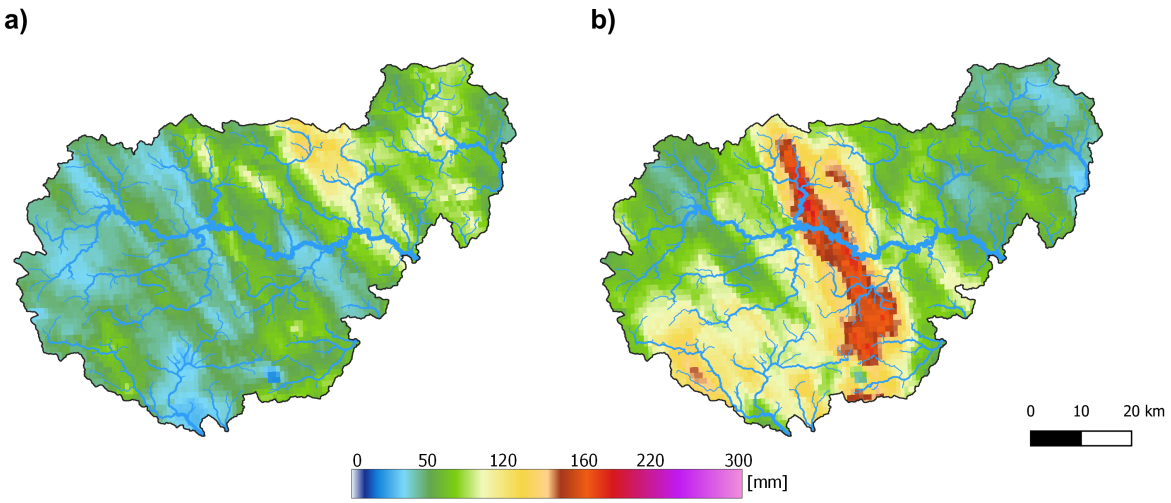

FIGURE 4.6 – *Cumuls de pluie pour Antilope J+1 a) Du 29 octobre 00 h au 30 octobre 12 h et b) du 30 octobre 14 h au 1er novembre 23 h*

Le premier constat est que le premier pic de précipitation est très réduit en termes de durée, et est également assez peu intense en cumul observé. La deuxième partie de l'événement atteint un cumul pluviométrique plus important dépassant localement les 180 mm sur la période du 30 octobre, 14 h, au 1er novembre, 23 h. Partant de ce dernier constat et en prenant en compte le fait que la majorité des dégâts observés ont eu lieu pendant ce deuxième pic de précipitations (d'après les différentes coupures de journaux encore disponibles - Le Parisien), le choix a été fait de ne considérer que cette période du 30 octobre 14 h au 1er novembre 23 h pour l'évaluation qui suit.

La Figure 4.7 illustre l'évolution temporelle des précipitations moyennes horaires mesurées et prévues (pour des échéances de 1 h, 3 h et 6 h) pour l'événement de fin octobre, début novembre sur le bassin de l'Argens et les autres bassins limitrophes concernés par l'événement.

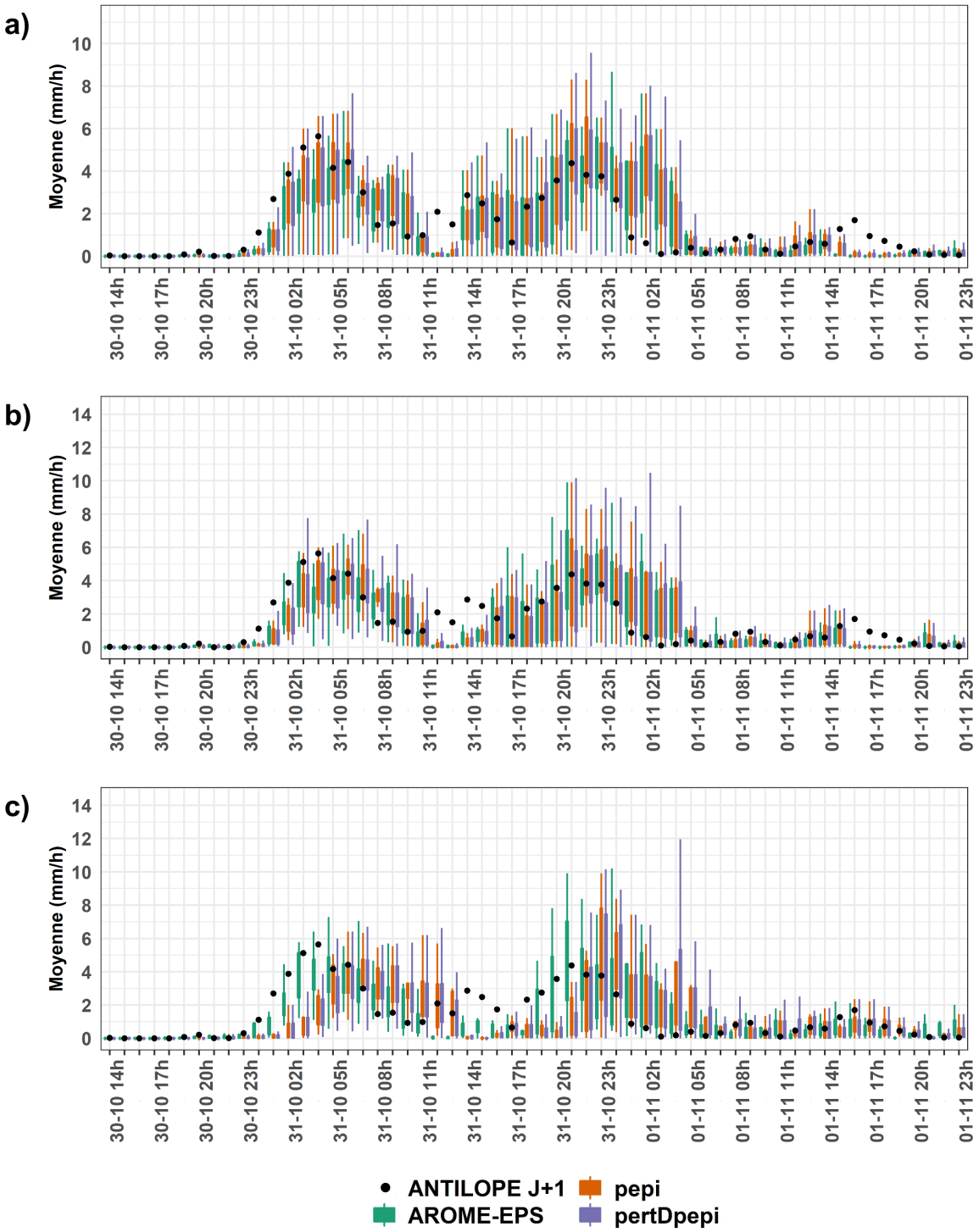

FIGURE 4.7 – *Évolution temporelle des pluies moyennes horaires observées (Antilope J+1) et prévues par AROME-EPS (vert), pepi (orange) et pertDpepi (violet) pour les échéances 1 h (a)), 3 h (b)) et 6 h (c)) pour l'événement de fin octobre 2018 dans le Var (secteur de l'Argens)*

Les pluies moyennes horaires observées sont globalement très bien anticipées par les trois produits de prévision. Les prévisions d'ensemble de pluie apparaissent généralement très proches les unes des autres, sauf dans les cas des intensités de pluie les plus grandes. Pour les périodes de pluies intenses, la valeur ajoutée du modèle AROME-PI apparaît plus nettement, en augmentant la dispersion des prévisions d'ensemble, en particulier pour les cas de l'Aude et

l'Ardèche. En général, les pluies moyennes horaires observées sont contenues dans l'intervalle 25%-75% des prévisions d'ensemble. Néanmoins, deux exceptions peuvent être faites à cette observation générale : i) Dans les cas de l'Aude et du Var, il y a une surestimation de la pluie prévue pour l'échéance + 1 h à la fin de l'événement pluvieux (Figures 4.4 a) et 4.7 a)). Cette situation n'aura probablement qu'un impact limité sur l'anticipation des inondations puisqu'elle ne concerne que la fin de l'événement pluvieux ; ii) Dans le cas de l'Ardèche, les prévisions d'ensemble ont plutôt tendance à la sous-estimation sur la totalité de l'événement, avec une importance moindre puisque les cumuls observés sont globalement moins grands que pour les deux autres cas d'étude. Pour les trois cas évalués, la forme du hyétogramme des précipitations moyennes est également bien anticipée par les prévisions même à une échéance lointaine (Figures 4.4 c), 4.5 c) et 4.7 c)), mais avec un décalage temporel de 3-4 h pour l'Aude et plutôt 5-6 h pour l'Ardèche et le Var (même si dans ce dernier cas, le décalage est plus difficile à lire du fait de la succession de pics de précipitations). Sans être totalement parfaites, ces prévisions à échéance assez lointaine (pour de la prévision d'événements méditerranéens) restent très intéressantes, car elles permettent d'espérer une anticipation des pluies (le retard restant inférieur à l'échéance de prévision).

À partir de ces hyétogrammes moyens, il est possible de définir une fenêtre temporelle comprenant les pluies d'intensité forte pour chaque cas d'étude. Par la suite, les prévisions seront uniquement considérées sur la période du 14 octobre 20 h au 15 octobre 10 h pour l'événement de l'Aude, de 00 h à 23 h le 9 août pour l'Ardèche et enfin du 30 octobre 14 h au 1er novembre 23 h pour l'événement du Var.

4.2.2 Localisation des fortes intensités de pluies

Le deuxième outil utilisé pour comparer les pluies observées et prévues est un ensemble de cartes représentant la localisation des pluies : cumuls de pluie observés sur l'événement, et somme temporelle des prévisions moyennes et des percentiles 75% et 95% des trois ensembles de prévision, pour une échéance de prévision fixée. Ces cartes permettent de comparer les intensités prévues et observées, mais aussi d'évaluer les différences de localisation des pluies observées et prévues. De plus, elles permettent de mettre en évidence certaines caractéristiques des ensembles de pluies prévues.

Ce travail sur la localisation des cumuls de précipitations permet également de définir les secteurs de précipitations intenses sur lesquels l'évaluation des prévisions d'ensemble des débits pourra être mené pour les trois événements. Les secteurs retenus sont représentés avec un contour noir sur les cartes qui suivent. Ils correspondent à la fois aux secteurs de cumuls observés importants, mais également aux zones de pluies importantes prévues.

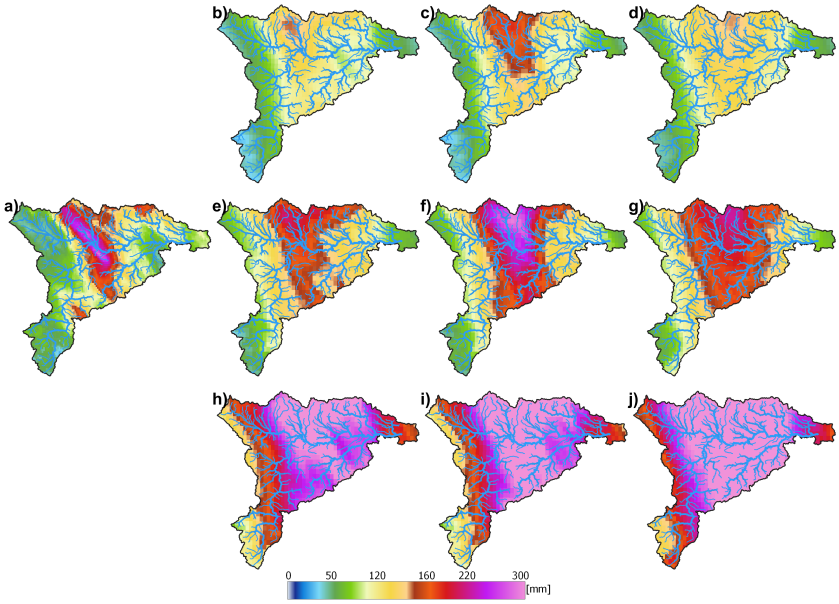

FIGURE 4.8 – *Comparaison des cumuls de précipitations observées et de la somme des pluies horaires prévues (pour une échéance de prévision d'une heure) sur 15 heures (du 14 octobre 20 h au 15 octobre 10 h) pour l'événement de l'Aude : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

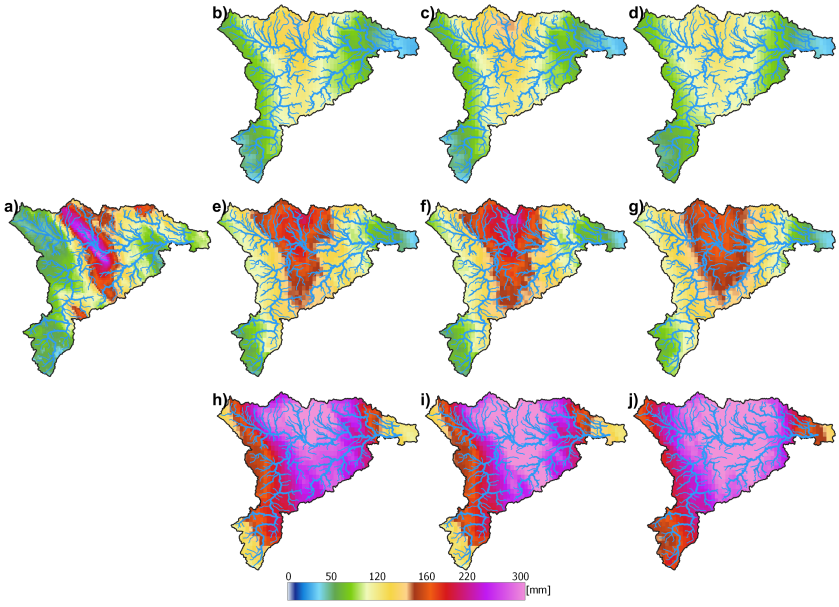

FIGURE 4.9 – *Comparaison des cumuls de précipitations observées et de la somme des pluies horaires prévues (pour une échéance de prévision de six heures) sur 15 heures (du 14 octobre 20 h au 15 octobre 10 h) pour l'événement de l'Aude : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

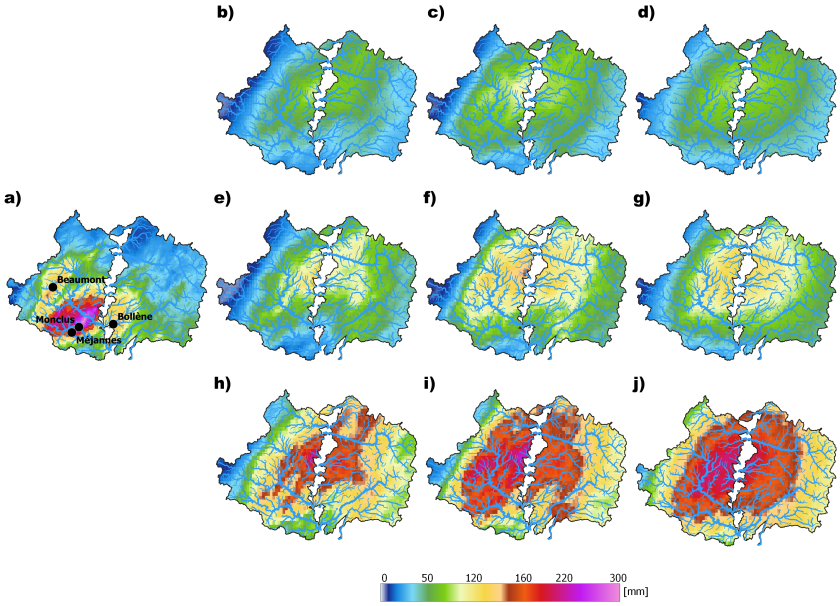

FIGURE 4.10 – *Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision d'une heure) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

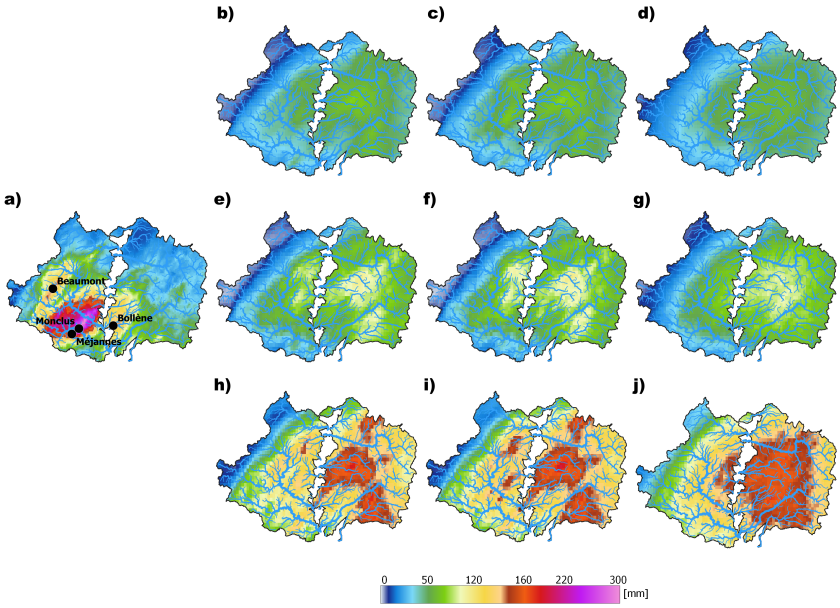

FIGURE 4.11 – *Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision de six heures) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

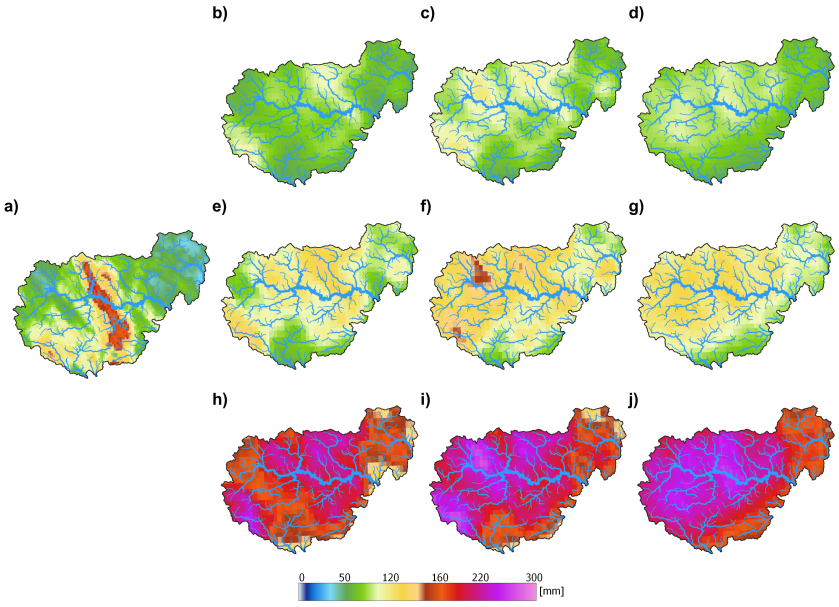

FIGURE 4.12 – *Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision d'une heure) pour l'événement du Var : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

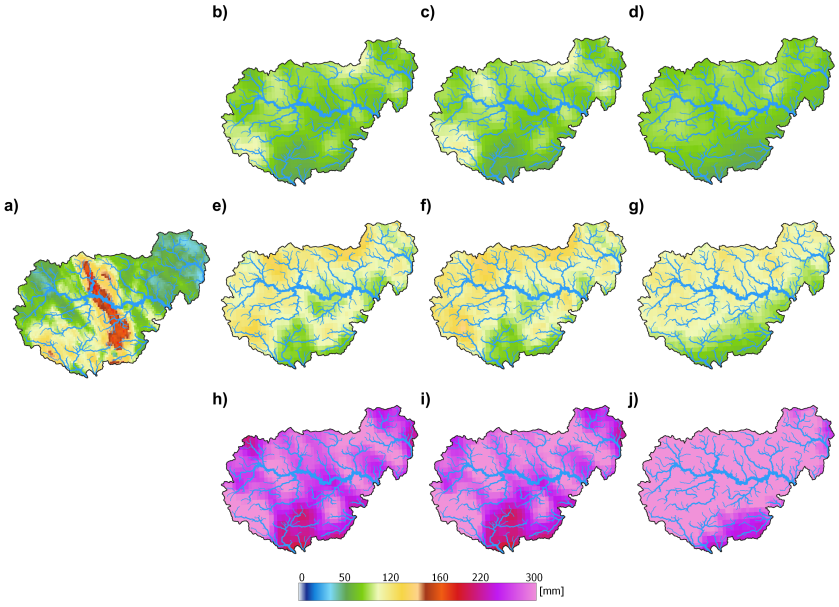

FIGURE 4.13 – *Comparaison des précipitations cumulées observées et des sommes des pluies horaires prévues (pour une échéance de prévision de six heures) pour l'événement de l'Ardèche : a) précipitations observées, b) e) h) moyenne de l'ensemble AROME-EPS, percentiles 75% et 95%, c) f) i) moyenne de l'ensemble pepi, percentiles 75% et 95%, d) g) j) moyenne de l'ensemble pertDpepi, percentiles 75% et 95%.*

Le bilan de cette première visualisation est que les prévisions d'ensemble sont globalement de très bonne qualité et retracent relativement bien les trois événements étudiés. Plusieurs constats peuvent être tirés des cartes de localisation des pluies observées et prévues.

La moyenne des ensembles, quel que soit le produit considéré, aboutit à des quantités de précipitations prévues plutôt faibles par rapport à la situation observée et ne permet pas de délimiter des zones de fortes précipitations prévues (faible contraste spatial). Avec des percentiles plus élevés, la distinction entre les zones de pluies peu intenses et fortes est plus nette. C'est même avec le percentile 75% que les quantités de pluie prévues s'approchent le plus du cumul observé (quel que soit l'ensemble de prévision). Plus le percentile augmente et plus la pluie forte « s'étale » en surface, cela est particulièrement visible sur le percentile 95% présenté sur la dernière ligne (h), i) et j)) des Figures 4.8, 4.9, 4.10, 4.11, 4.12 et 4.13. Se focaliser sur des percentiles très grands (respectivement très petits) peut conduire à sélectionner un membre extrême de la prévision, et conduire à une forte sur-estimation (respectivement sous-estimation) des intensités pluviométriques. C'est ici une difficulté de l'évaluation d'événements rares et extrêmes avec des prévisions d'ensemble. Pour la suite de l'évaluation de ces prévisions d'ensemble, le percentile 75% sera le plus souvent utilisé comme valeur indicative pour juger des performances de l'ensemble.

Le deuxième point concerne la construction des ensembles incluant AROME-PI (pepi et pertDpepi). Pour les événements étudiés, l'ajout de membres à partir d'AROME-PI permet d'obtenir des prévisions de pluie plus intenses pour un même percentile. Cette influence diminue avec les percentiles les plus élevés pour lesquels la zone de pluie intense est tellement étendue qu'il n'y a plus de distinction entre les trois produits de prévision. Enfin, le dernier point concerne l'utilisation des perturbations spatiales pour construire l'ensemble pertDpepi, qui entraîne un lissage spatial des cumuls pour ce produit.

4.3 Évaluation de la fiabilité et la finesse des prévisions d'ensemble de pluie horaires

Cette section s'intéresse à évaluation des prévisions d'ensemble à partir du CRPS et du diagramme de rang. Ces deux outils permettent d'évaluer les deux attributs principaux des prévisions d'ensemble, qui sont la fiabilité et la finesse, comme le proposent Gneiting *et al.* (2007). S'agissant ici d'évaluations événementielles, les deux scores seront calculés sur des chroniques temporelles de longueur réduite, ainsi que des secteurs géographiques restreints. Ces résultats seront donc à considérer avec précaution.

Les précipitations ont été stratifiées en trois sous-catégories correspondant à des intensités de pluie différentes. Dans un contexte de vérification, la stratification consiste à découper l'échantillon prévisions - observation en plusieurs sous-échantillons homogènes (Murphy 1995).

Ces seuils ont été définis à partir des percentiles 20% et 80% de l'ensemble pertDpepi. Cette division de l'ensemble des données de prévisions en différentes catégories de pluie prévues permet d'évaluer le comportement des pluies selon leur intensité. La stratification est faite sur les prévisions, car des études antérieures (Bellier 2018) ont montré que la stratification sur les prévisions était judicieuse et permettait de se rapprocher de la définition de la fiabilité (attribut recherché de la prévision), la stratification sur l'observation étant à proscrire puisque occasionnant des perturbations sur le diagramme de rang.

Les trois sous-catégories d'intensités retenues sont :

- Les prévisions de pluies inférieures à 2 mm/h.
- Les prévisions de pluies comprises entre 2 et 5 mm/h.
- Enfin, les prévisions de pluies supérieures à 5 mm/h.

4.3.1 Résultats obtenus pour les diagrammes de rang

Les Figures 4.14, 4.15 et 4.16 présentent les diagrammes de rang obtenus pour les trois événements étudiés et pour les échéances de prévision 1 h, 3 h et 6 h.

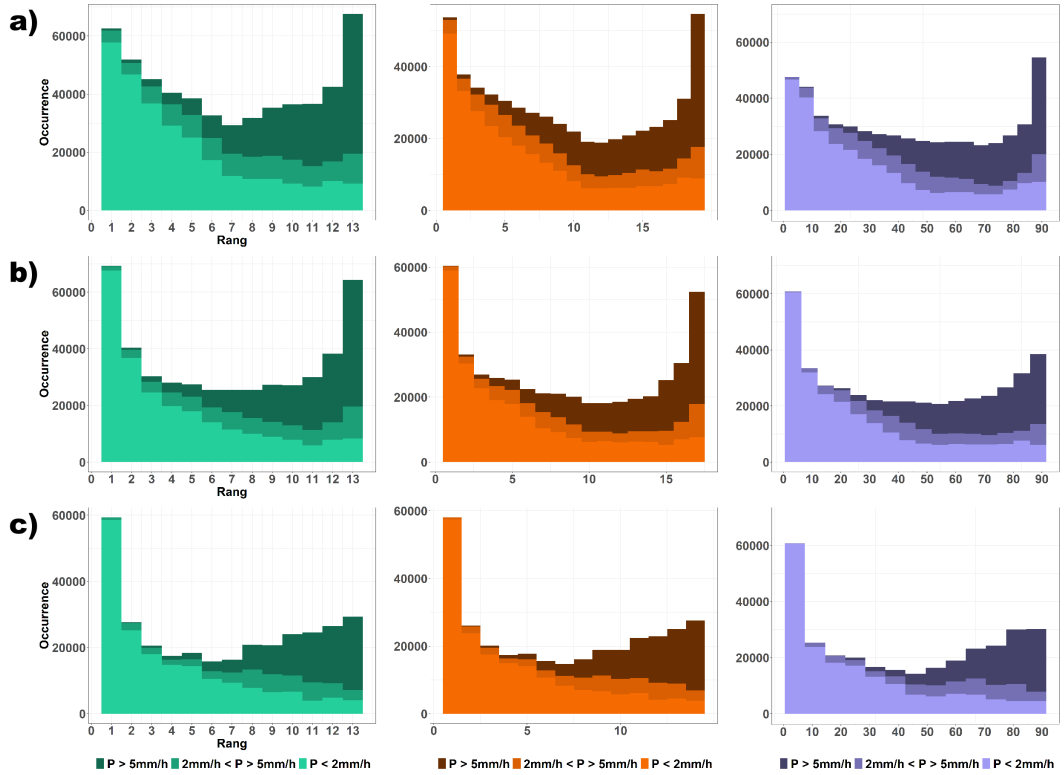

FIGURE 4.14 – *Diagrammes de rang sur l'événement de l'Aude, pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures*

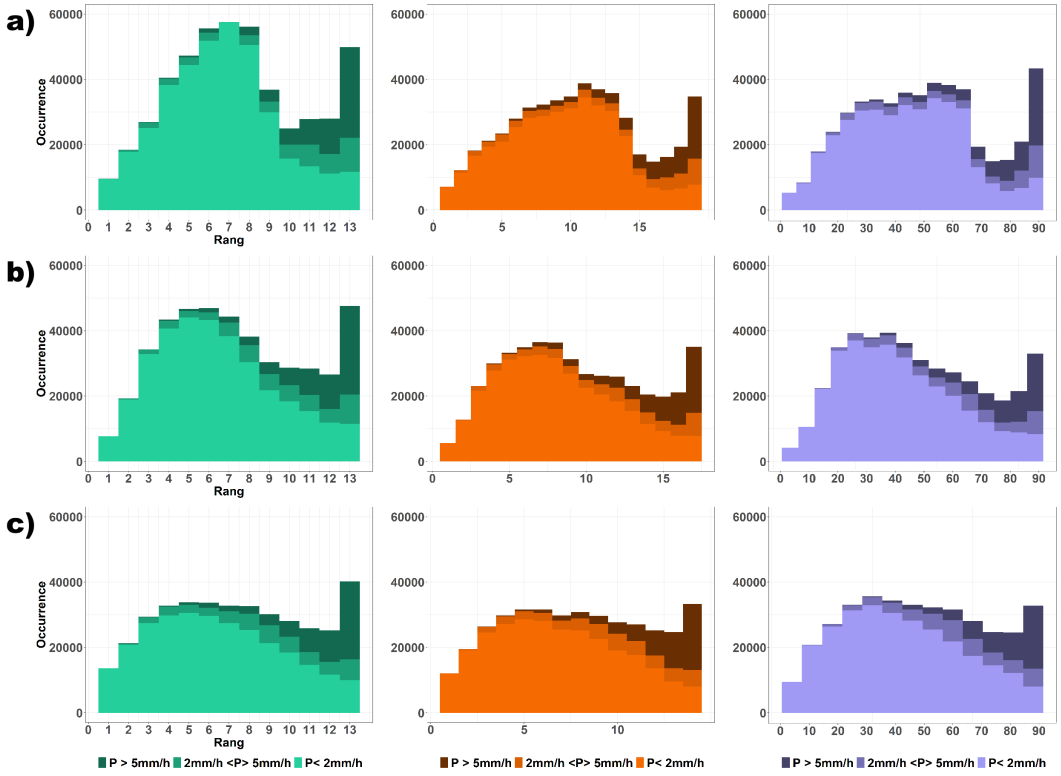

FIGURE 4.15 – *Diagrammes de rang sur l'événement de l'Ardèche, pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures*

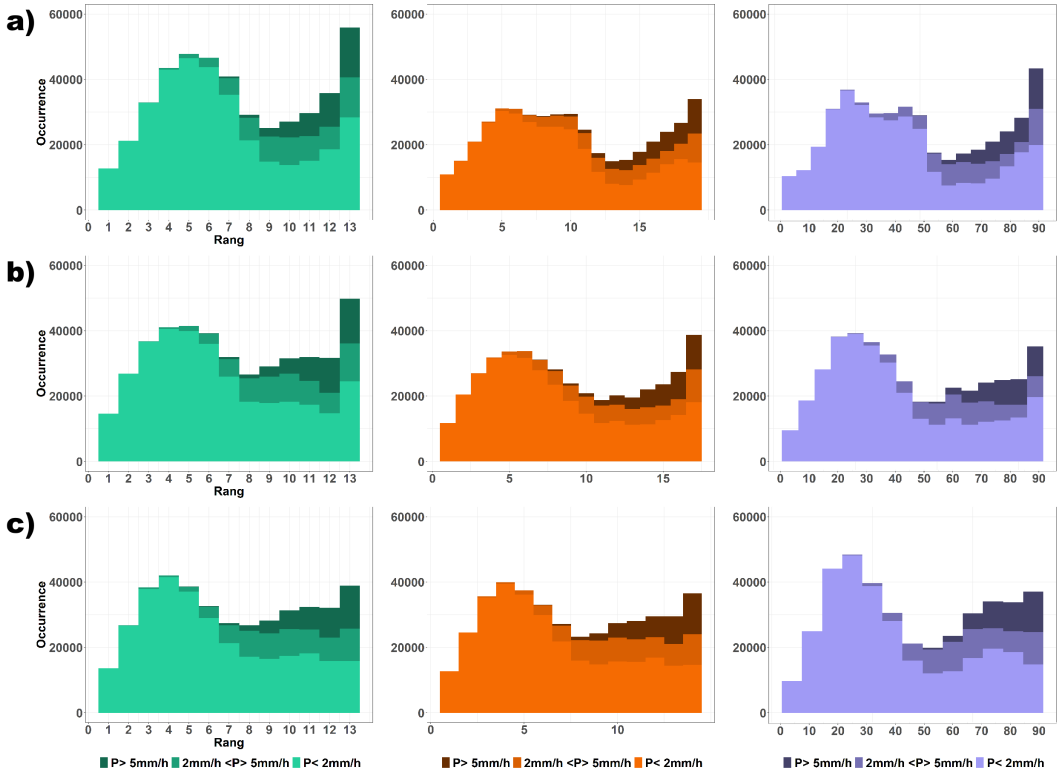

FIGURE 4.16 – *Diagrammes de rang sur l'événement du Var pour AROME-EPS, pepi et pertDpepi, et pour des échéances de prévision de a) une heure, b) trois heures, c) six heures*

Sur la forme générale des diagrammes de rang (sans tenir compte de la stratification sur les précipitations prévues), deux types de diagrammes se distinguent :

- Des classes extrêmes (percentiles inférieurs à 20 % et supérieurs à 80 %), plus complètes que les classes intermédiaires, donnent un caractère sous-dispersif aux prévisions dans le cas de l'Aude (Figure 4.14). Ce trait s'atténue avec l'augmentation de l'échéance, faisant plutôt apparaitre un biais positif avec des classes supérieures très peu remplies (Figure 4.14 c)).

- Au contraire, pour les cas de l'Ardèche et du Var, les classes intermédiaires (percentiles 50 %) des diagrammes de rang sont plus denses (Figures 4.15 et 4.16), montrant un caractère sur-dispersif des prévisions d'ensemble et ce quelque soit l'échéance de prévision regardée. Ainsi, une majorité des observations se trouve dans les percentiles 40 - 60 %, et il y a une faible représentation des valeurs extrêmes. Pour les diagrammes de rang du Var, une petite différence peut être notée sur la forme générale du diagramme, avec une majorité des observations se situant parmi les percentiles 60 - 75 %, plutôt qu'autour des percentiles 50 %.

Concernant le détail des trois sous-échantillons issus de la stratification, d'autres tendances apparaissent selon l'intensité de la pluie. Pour les pluies inférieures à 2 mm/h, les prévisions présentent un biais positif (sur-représentation des percentiles peu élevés) pour l'Aude et sont sur-dispersées dans les cas de l'Ardèche et du Var. Cela est valable pour toutes les échéances de prévision. Pour les pluies comprises entre 2 et 5 mm/h, la répartition dans les différentes classes est uniforme. Enfin, concernant le sous-échantillon des précipitations supérieures à 5 mm/h, les prévisions présentent pour tous les cas d'étude un biais négatif.

Dans l'ensemble, si on ne tient pas compte de la stratification, les diagrammes de rang sont néanmoins assez fiables avec des biais peu marqués. Au niveau de la dispersion, les résultats sont variables avec une sous-dispersion pour l'Aude et au contraire, une sur-dispersion pour l'Ardèche et le Var. Concernant les biais, il y a pour les trois cas un biais négatif, en particulier pour les pluies supérieures à 5 mm/h. Ce biais constaté met en évidence l'importance de définir très précisément les zones d'évaluation.

4.3.2 Résultats obtenus pour les CRPS

Les Figures 4.17, 4.18 et 4.19 présentent les CRPS pour les trois événements étudiés et pour toutes les échéances de prévision, de 1 h à 6 h.

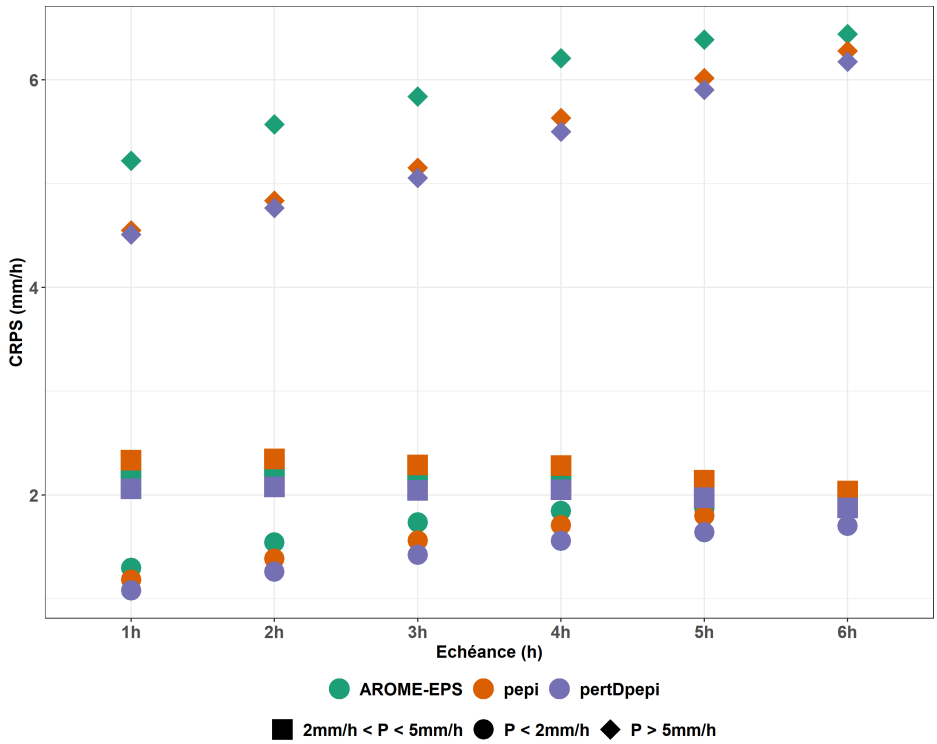

FIGURE 4.17 – *CRPS calculés sur l'événement de l'Aude pour AROME-EPS, pepi et pertDpepi*

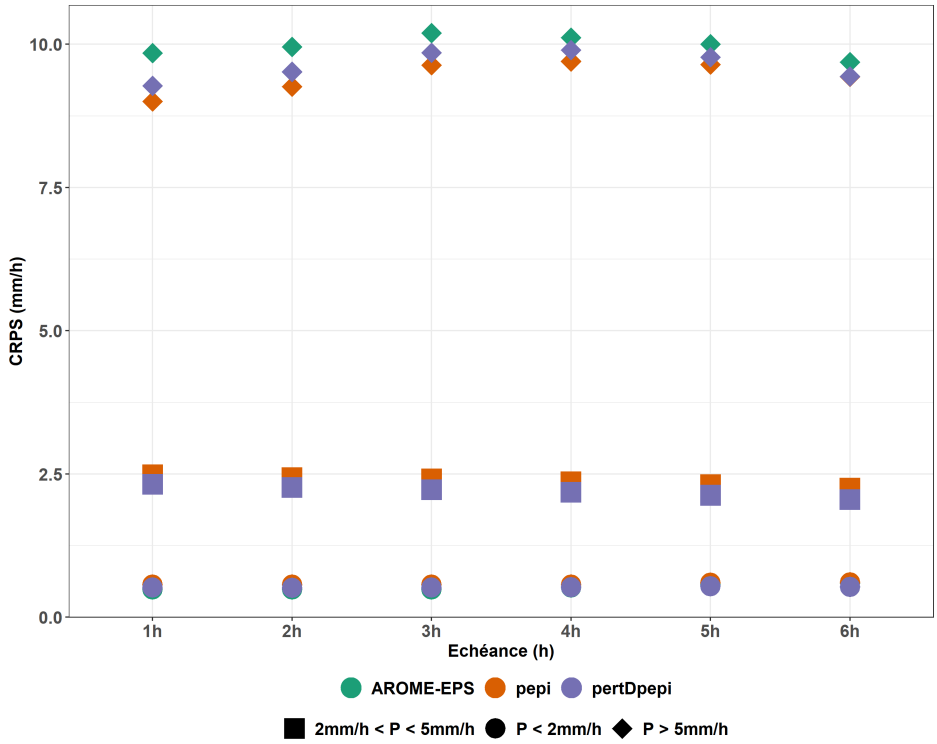

FIGURE 4.18 – *CRPS calculés sur l'événement de l'Ardèche pour AROME-EPS, pepi et pertDpepi*

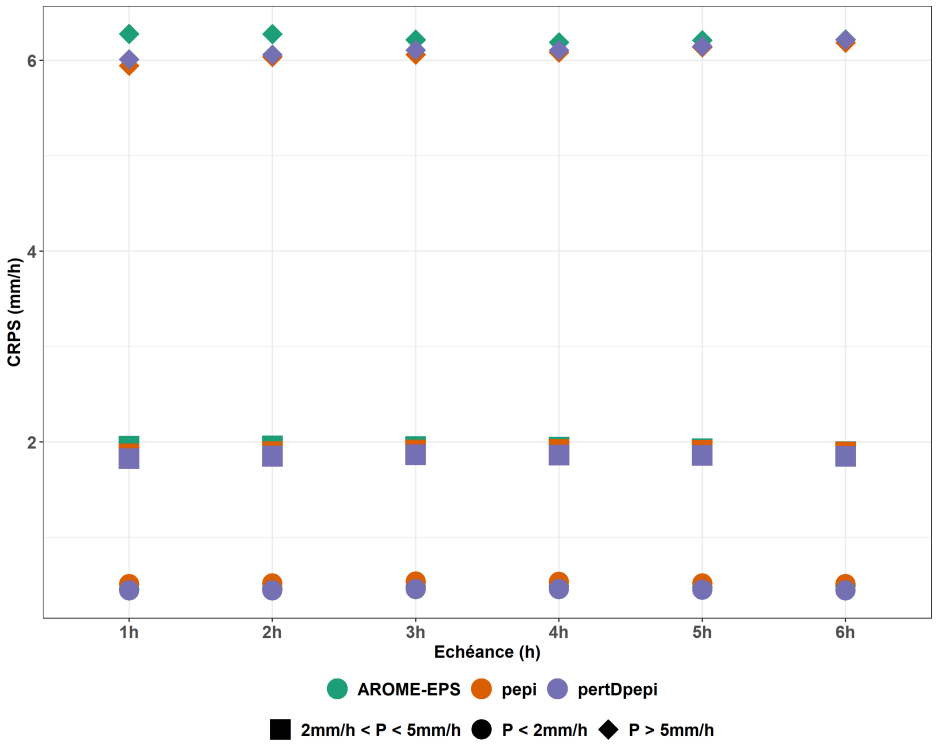

FIGURE 4.19 – *CRPS calculés sur l'événement du Var pour AROME-EPS, pepi et pertDpepi*

Les conclusions sur l'analyse des CRPS sont les mêmes pour les trois cas d'étude. Le premier constat qui peut être fait est que l'on retrouve un classement des CRPS selon l'intervalle de précipitations regardé (stratification), avec des CRPS proches de 0 (la valeur optimale du score) pour les précipitations inférieures à 2 mm/h, puis se dégradent pour l'intervalle 2-5 mm/h, et encore plus pour les précipitations supérieures à 5 mm/h. Cela s'explique par le fait que l'intervalle de pluie inférieure à 2 mm/h contient plus de pluie nulle, ce qui fait tendre le CRPS vers sa valeur idéale. En effet, les pluies nulles sont souvent très bien prévues. Cet argument explique aussi la dégradation observée, pour les autres intervalles d'intensité pluvieuse : l'absence de pluies nulles dans les autres classes fait largement augmenter la valeur du CRPS.

Un deuxième constat possible est à propos de l'évolution du CRPS selon l'échéance regardée. En ce qui concerne l'échéance, la conclusion est assez variable selon le sous-échantillon d'intensités pluviométriques considéré. Ainsi, pour les pluies supérieures à 5 mm/h, la valeur du CRPS augmente sensiblement pour les échéances les plus lointaines. C'est également le cas du CRPS pour les pluies inférieures à 2 mm/h. Néanmoins, les CRPS pour les pluies comprises entre 2 et 5 mm/h sont en légère amélioration avec les échéances lointaines et passent à un CRPS autour de 2 mm/h. Cette amélioration peut s'expliquer par une plus grande dispersion des prévisions d'ensembles pour les échéances les plus lointaines, réduisant ainsi le nombre de prévisions situées dans cette classe. Cette amélioration avait déjà été observée sur les diagrammes de rang des échéances plus lointaines (Figures 4.14 c), 4.15 c) et 4.16 c)).

Deux distinctions peuvent néanmoins être faites sur les différents cas d'études. Pour l'Ardèche, l'augmentation du CRPS avec l'échéance est moins importante que pour l'Aude.

De plus, les valeurs de CRPS atteintes pour l'Ardèche sont plus élevées allant jusqu'à 10 mm/h (contrairement aux 6,5 mm/h pour le Var et l'Aude).

Enfin, le troisième constat concerne la distinction des produits, celle-ci ne se fait pas très nettement sur les pluies inférieures à 2 mm/h, mais plus nettement quand les valeurs du CPRS augmentent. Au regard du CRPS, pertDpepi est l'ensemble le plus performant avant pepi et AROME-EPS. Les CRPS se dégradent d'autant plus que l'on regarde les précipitations plus intenses (supérieures à 5 mm/h), cela indique que les ensembles sont moins fiables et moins fins pour les pluies les plus intenses. Les différences entre les trois produits ne sont pas très importantes, mais elles existent. Cette différence est particulièrement visible sur les courtes échéances pour le CRPS et sur les précipitations les plus intenses (supérieures à 5 mm/h). L'ajout d'AROME-PI (pour pepi) et des perturbations spatiales (pour pertDpepi) influencent positivement la finesse et la fiabilité des prévisions. Il faut souligner un léger avantage de pertDpepi sur le score du CRPS qui apparait plus fiable et plus fin que les deux autres produits.

4.3.3 Influence du nombre de membres sur les scores de CRPS obtenus

Une des conclusions de l'étude des CRPS pour les trois produits de prévision d'ensemble est que l'ensemble pertDpepi semble plus fin et plus fiable que ne l'est pepi et lui-même que AROME-EPS. Il est intéressant de noter que le classement des produits se fait aussi selon le nombre de membres avec pertDpepi contenant entre 65 et 90 membres, pepi, contenant entre 13 et 18 membres et AROME-EPS, 12 membres. Il a été montré que l'augmentation de la taille d'un ensemble influençait positivement l'évolution du CRPS (Buizza et Palmer 1998). Les ensembles contenant plus de membres peuvent donc être privilégiés du fait même de leur nombre de membres, indépendamment de leur finesse et leur fiabilité. Afin de vérifier si seule la taille de l'ensemble est responsable d'un CRPS plus proche de 0, des ensembles réduits ont été créés à partir de pepi et pertDpepi de manière à les ramener à 12 membres (correspondant au nombre de membres de l'ensemble le plus petit, AROME-EPS). Pour réduire la taille des ensembles, les membres ont été classés par ordre croissant pour chaque pas de temps et chaque échéance, puis, 12 membres ont été tirés à intervalles réguliers dans l'ensemble des membres classés. Cette méthode permet de garder la dispersion et la distribution statistique de l'ensemble original, et de ne pas perdre les données caractérisant l'ensemble. Afin de voir quelles différences il y a entre les CRPS sur les ensembles entiers et sur les ensembles réduits, les CRPS obtenus pour les ensembles orignaux pepi et pertDpepi sont tracés avec les nouveaux CRPS obtenus pour les mêmes ensembles réduits. Les résultats obtenus sont présentés sur la Figure 4.20 pour l'événement de l'Aude. Par commodité, seront appelés :

- pepi-réduit : l'ensemble pepi ramené à 12 membres.
- pertDpepi-réduit : l'ensemble pertDpepi ramené à 12 membres.

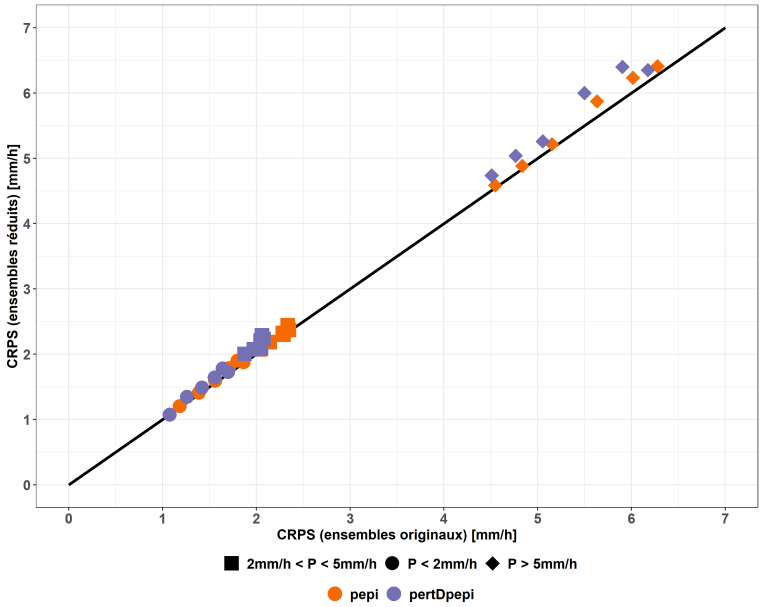

FIGURE 4.20 – *Comparaison des CRPS obtenus sur les ensembles pepi (en orange) et pertDpepi (en violet) originaux (en abscisse) et sur les ensembles réduits (en ordonnée), toutes échéances confondues, pour l'événement de l'Aude*

La première remarque générale est que les points sont tous très proches de la droite x=y ce qui signifie que les valeurs de CRPS de pertDpepi et pepi sont semblables à celles calculées pour pepi-réduit et pertDpepi-réduit. Il faut noter que les valeurs de CRPS, pour les ensembles réduits, sont légèrement augmentées dans le cas de pepi-réduit et pertDpepi-réduit, en particulier pour les précipitations supérieures à 5 mm/h. En regardant maintenant par rapport aux performances du CRPS pour AROME-EPS, il est possible de tracer la Figure 4.21.

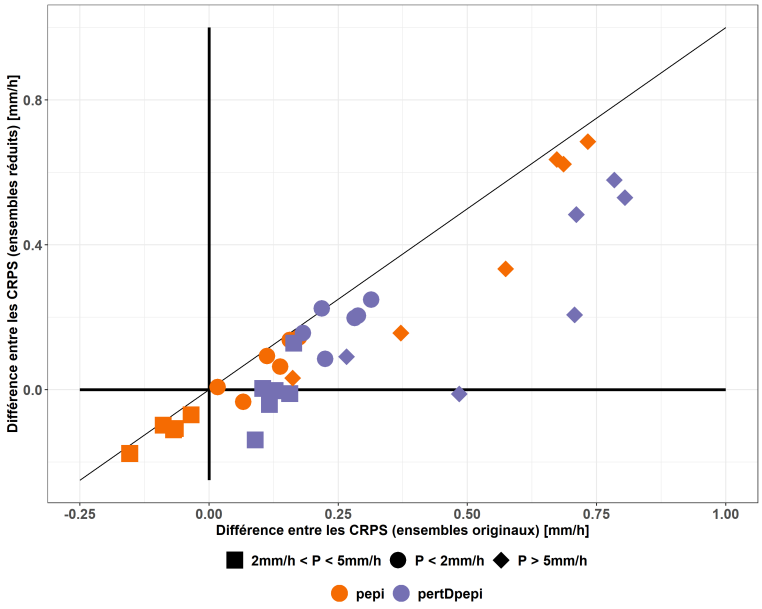

FIGURE 4.21 – *Comparaison des différences de CRPS entre AROME-EPS et les ensembles pepi, pertDpepi, pepi-réduit et pertDpepi-réduit, toutes échéances confondues, pour l'événement de l'Aude*

La majorité des points se trouvent dans la partie positive du graphique, ce qui signifie que même si les CRPS de pepi-réduit et pertDpepi réduit sont plus grands que les CRPS originaux, la différence est toujours en faveur de pepi et pertDpepi par rapport au CRPS d'AROME-EPS. Il y a cependant quelques cas, en particulier les précipitations d'intensité moyenne (en carré sur la Figure 4.21), où la valeur du CRPS est suffisamment dégradée pour être plus grande que le CRPS d'AROME-EPS.

À partir de ces résultats, il est possible de constater que même si le nombre de membres d'un ensemble impacte le calcul du CRPS en l'améliorant (Figure 4.20), ce n'est pas ici l'unique élément explicatif des meilleurs résultats de CRPS observés pour les ensembles pepi et pertDpepi. En effet, même en diminuant le nombre de membres, les résultats sont presque toujours meilleurs (Figure 4.21) pour ces deux produits par rapport à l'ensemble AROME-EPS.

Synthèse sur l'évaluation des prévisions d'ensemble de pluie

Dans ce Chapitre, une évaluation de prévisions d'ensemble des pluies, complémentaire à celle initialement effectuée par Météo-France, a été menée. Cette évaluation est focalisée sur les événements de crue qui seront étudiés par la suite, et elle combine une visualisation des données et des critères plus classiques d'évaluation des prévisions d'ensemble.

La visualisation temporelle des prévisions d'ensemble a été réalisée à partir de hyétogrammes d'intensités moyennes horaires permettant de visualiser la dispersion des ensembles et de les comparer à la pluie observée Antilope J+1. En ce qui concerne l'aspect spatial, la moyenne ainsi que les percentiles 75% et 95% de chaque ensemble ont été sommés sur la durée de l'événement pour une échéance donnée, et comparés au cumul de pluie obtenu à partir d'Antilope J+1. Cette première évaluation visuelle, réalisée sur les trois événements étudiés, a permis d'identifier les caractéristiques générales des prévisions d'ensemble et de définir un cadre spatio-temporel pour la suite de l'étude. Dans l'ensemble, les pluies sont très bien anticipées par les trois produits de prévision, avec une difficulté pour capturer les fortes intensités pluviométriques, qui se retrouve sur tous les scores d'évaluation utilisés. De plus, la visualisation spatiale permet de mettre en avant le fait qu'une prévision présentant un léger décalage spatial d'un point de vue pluviométrique, comme c'est le cas ici (en particulier pour le Var), pourrait s'avérer très problématique pour réaliser des prévisions hydrologiques. En effet, le décalage des précipitations pourrait affecter un bassin-versant voisin et fausser complètement la prévision hydrologique. Le Chapitre 5 s'intéressera plus particulièrement à ce point avec l'évaluation des prévisions d'ensemble hydrologiques.

Des scores de fiabilité et de finesse (diagrammes de rang et CRPS) ont été calculés sur les périodes et secteurs sélectionnés à partir des visualisations. Autant le diagramme de rang s'avère assez facile à interpréter et informatif sur le comportement des prévisions d'ensemble, autant le CRPS s'avère être un score difficile à utiliser. Sans référence comparable, le CRPS ne permet pas d'émettre un jugement de valeur sur les performances de prévisions d'ensemble dans l'absolu. Néanmoins, il rend possible la comparaison des trois produits de prévision entre eux. La distinction entre ces trois ensembles se fait surtout sur les pluies intenses. Cependant, cette plus grande dispersion n'est pas autant visible sur les diagrammes de rang, si ce n'est pour le Var où une moins grande sur-dispersivité est observée pour pepi. À l'aune de ces critères, les produits pepi et pertDpepi semblent apporter des améliorations, par rapport à AROME-EPS.

Chapitre 5

Évaluation des prévisions d'ensemble des débits de crue

Objectifs du chapitre :

Après l'évaluation événementielle des prévisions de pluie, réalisée dans le Chapitre 4, une évaluation complémentaire portant sur les prévisions hydrologiques est proposée dans ce Chapitre. Un nouveau cadre méthodologique d'évaluation est proposé pour répondre aux questions des échelles d'analyse spatiales et temporelles spécifiques aux crues soudaines. Ces deux étapes sont mises en évidence sur la Figure 5.1.

Cette évaluation hydrologique des trois produits de prévision d'ensemble est conduite pour les événements de l'Aude, de l'Ardèche et du Var de 2018. Elle est menée en se concentrant dans un premier temps sur l'anticipation d'un dépassement de seuil de débit, en utilisant des courbes ROC (Relative Operating Characteristics diagram), des cartes de détection anticipée des dépassements. Par la suite, plusieurs exutoires sont sélectionnés sur lesquels une évaluation détaillée des prévisions hydrologiques est réalisée, en comparant des hydrogrammes de crues simulés sur la base des pluies observées et des hydrogrammes prévus simulés sur la base des prévisions de pluies.

L'objectif principal est d'évaluer l'apport des trois produits de prévision d'ensemble pour l'anticipation des débits, ainsi que les effets de la construction de ces trois ensembles sur les résultats obtenus.

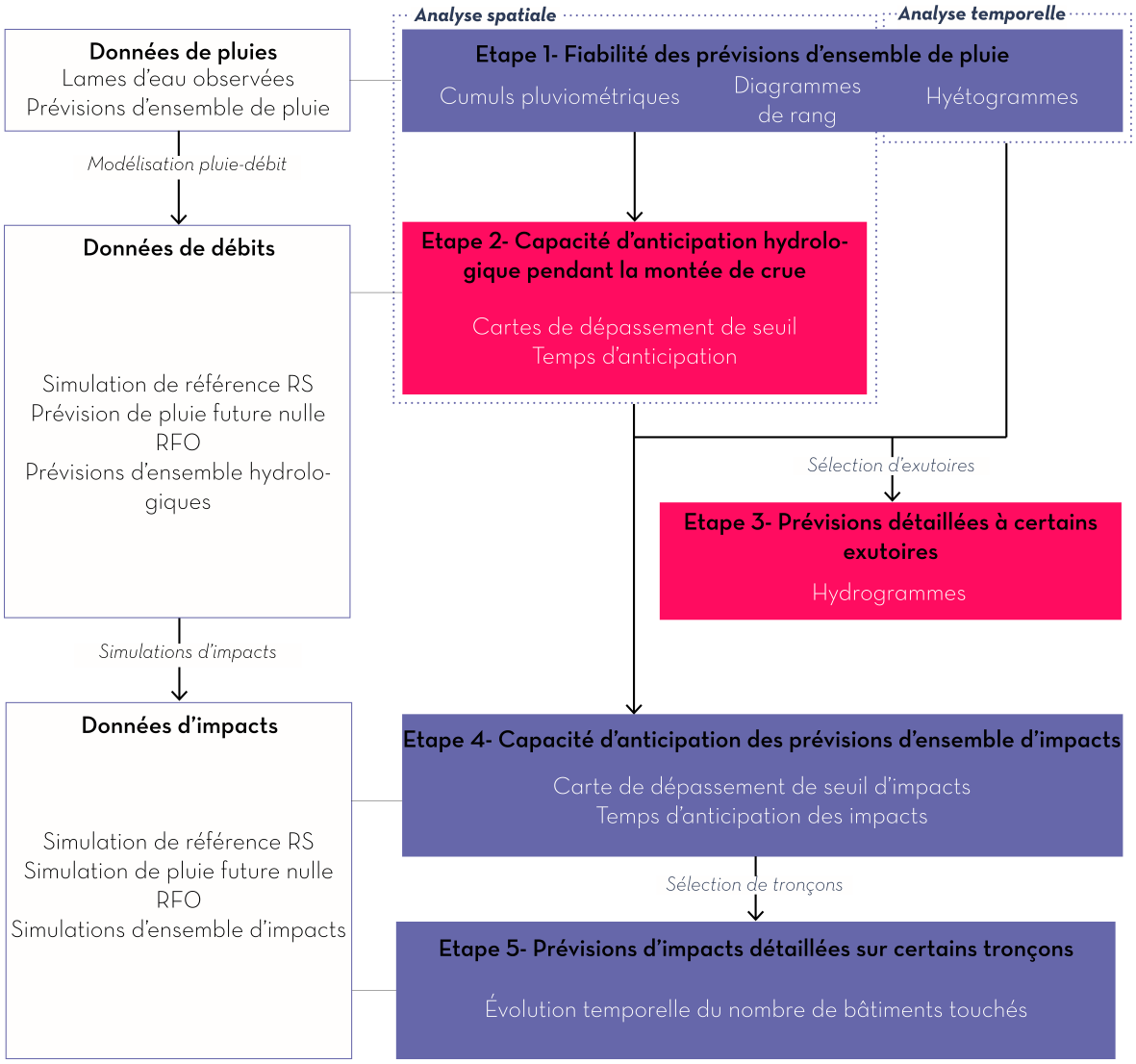

FIGURE 5.1 – *Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, sont représentées les deux étapes d'évaluation menées dans le Chapitre 5.*

5.1 Principes de l'évaluation des prévisions de débits

Le cadre d'évaluation proposé a pour but de déterminer si l'intensité des crues peut être correctement anticipée, en se basant sur des prévisions d'ensemble de pluie qui ne capturent pas parfaitement toute la complexité spatiale et temporelle de la pluie observée, comme cela a été montré dans le Chapitre 4. Plus précisément, il s'agit de traiter la question de l'anticipation de seuils de débits, non seulement sur les tronçons de cours d'eau jaugés, mais également sur l'ensemble du domaine d'étude qui comporte un grand nombre de secteurs non jaugés. Le but est de rapprocher l'évaluation de l'utilisation réelle des données de prévision d'ensemble dans un contexte de gestion de crues et des inondations associées. Il s'agit donc d'évaluer ces prévisions du point de vue de l'anticipation des dépassements de débit ou d'une hauteur d'eau. Ces dépassements conduisent souvent au déclenchement d'actions de gestion de crise : information des populations, fermetures d'axes routiers, évacuations ou confinement, intervention des services de secours.

5.1.1 Préparation des prévisions d'ensemble des débits et définition des simulations de référence

5.1.1.1 Préparation des pluies

Pour réaliser les prévisions hydrologiques, le choix s'est porté sur le modèle pluie-débit distribué Cinecar (Section 3.3.1.1), pour sa simplicité et sa capacité à simuler des crues soudaines à l'échelle régionale (Naulin *et al.* 2013). Un traitement préalable sur les prévisions d'ensemble des pluies a été effectué afin de disposer de prévisions horaires. En effet, les prévisions de pluie étaient initialement fournies de façon cumulée, soit, pour une prévision réalisée à l'instant t, 6 fichiers disponibles (pour chaque membre) :

- Cumul de pluie de t+00 h à t+01 h.
- Cumul de pluie de t+00 h à t+02 h.
- Cumul de pluie de t+00 h à t+03 h.
- Cumul de pluie de t+00 h à t+04 h.
- Cumul de pluie de t+00 h à t+05 h.
- Cumul de pluie de t+00 h à t+06 h.

Cinecar ayant besoin de fichiers de pluviométrie au quart d'heure, la première étape a été d'obtenir les cumuls prévus au pas de temps horaires en décumulant les fichiers de prévision. Les intensités pluviométriques infra-horaires ont ensuite été considérées comme uniformes pour générer des simulations avec le modèle Cinecar à une résolution temporelle de 15 minutes.

Pour générer une prévision des débits à l'instant t, des séries pluviométriques continues sont ensuite utilisées, combinant la lame d'eau observée du début de l'événement pluvieux à l'instant t et les prévisions d'ensemble de pluie de t à $t+6h$. La Figure 5.2 reprend le principe de combinaison des prévisions d'ensemble de pluie avec la lame d'eau observée pour alimenter les prévisions de débits.

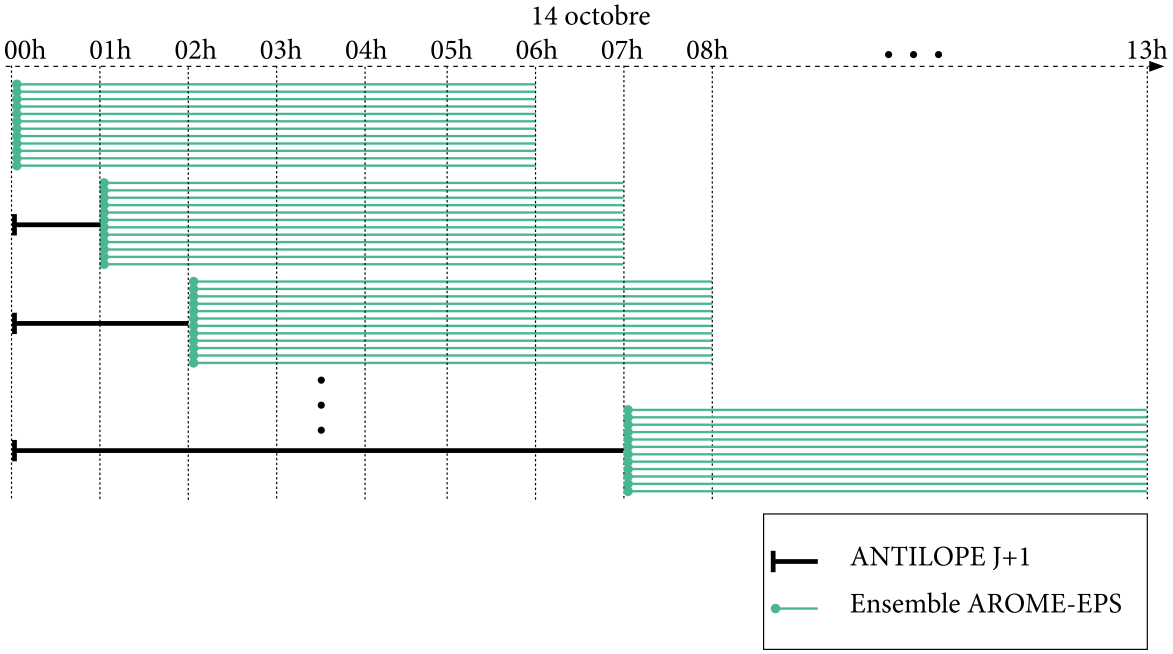

FIGURE 5.2 – *Combinaison entre les pluies Antilope J+1 et les prévisions d'ensemble (exemple d'AROME-EPS) pour former les données d'entrée de Cinecar*

Finalement, le nombre de prévisions de débits à lancer correspond au nombre de simulations de prévision de pluie (1 simulation par heure) multiplié par le nombre de membres dans l'ensemble.

À partir du découpage en sous-bassins versants du modèle Cinecar pour les trois études de cas (Section 3.3.1.1), il est possible de générer un hydrogramme pour chacun des exutoires du modèle (1 174 pour l'Aude, 2 358 pour l'Ardèche et 1 093 pour le Var). Pour chaque prévision effectuée, 1 174 hydrogrammes (respectivement 2 358 et 1 093 pour l'Ardèche et le Var) sont donc obtenus pour chaque membre de chaque ensemble.

5.1.1.2 Définition des simulations de référence pour l'évaluation

Afin d'évaluer les prévisions d'ensemble hydrométéorologiques, deux simulations de référence ont été considérées.

Simulation de référence RS

Afin de pouvoir évaluer plus facilement les prévisions hydrologiques réalisées à partir de Cinecar, le choix a été fait d'utiliser les résultats de la simulation réalisée à partir de la lame d'eau observée Antilope J+1 (appelée par la suite RS pour "Reference Simulation"), comme hydrogramme de référence pour évaluer les débits obtenus avec les prévisions d'ensemble. Ce choix présente plusieurs avantages :

- Les éventuelles erreurs et approximations du modèle pluie-débit ne viendront pas perturber l'évaluation des prévisions puisqu'elles se retrouveront à la fois sur les prévisions à partir des ensembles de pluie prévue, mais également sur la simulation à partir de la pluie observée.
- Le modèle Cinecar étant un modèle distribué, l'évaluation peut se faire sur un grand nombre de couples simulation-prévision. Ce nombre de couples aurait été plus restreint avec seulement l'utilisation d'hydrogrammes observés comme référence.
- Cela permet également de s'affranchir des incertitudes liées aux débits observés en période de crue (courbes de tarage, embâcles, appareils de mesure endommagés...) et à l'estimation des pluies.

Simulation avec une pluie future nulle, RF0

La prévision de pluie future nulle (RF0, Reference Forecast 0) est une prévision reposant uniquement sur les débits déjà constitués sur le bassin versant et en cours de propagation dans le réseau hydrographique, en amont de l'exutoire considéré. Cette prévision est générée à partir d'une nouvelle prévision de pluie, constituée de la manière suivante : jusqu'à l'heure de prévision t, la pluie Antilope J+1 est utilisée, puis à partir de t, la pluie est considérée comme égale à 0. Cela permet d'avoir un scénario de référence correspondant à une absence de prévision de pluie, utile pour évaluer les apports de la prévision d'ensemble des pluies.

5.1.2 Méthodologie d'évaluation des prévisions de débits

Principe général

La méthodologie utilisée ici pour évaluer les prévisions de débits se focalise sur la détection anticipée des dépassements de seuils de débit. Cette méthode s'inspire directement du principe de la méthode des courbes ROC développée par Mason (1982), Levine et Wilks (2000). La courbe ROC (Relative Operating Characteristics diagram) est un critère évaluant la capacité d'un système de prévision à détecter l'occurrence (réciproquement la non-occurrence) d'un événement. En hydrologie, l'événement correspond généralement au dépassement d'un seuil de débit, et la courbe ROC est alors construite pour résumer de longues séries temporelles de prévision à l'exutoire d'un bassin versant. Pour chaque échéance et chaque probabilité de l'ensemble, il est possible de comparer la position respective de la prévision avec celle de l'observation (ou ici de la simulation de référence RS). Cela permet de remplir un tableau de contingence 5.1 et de calculer la probabilité de détection (POD, Probability Of Detection) et le taux de fausses alarmes (FAR, False Alarm Rate) (Levine et Wilks 2000, Jolliffe et Stephenson 2012). La courbe ROC est ensuite tracée en utilisant les couples de points (FAR ; POD) obtenus en choisissant successivement les différents percentiles des ensembles de prévision des débits.

Dans ce travail d'évaluation, le calcul conventionnel des courbes ROC a été adapté en se focalisant sur l'analyse du premier dépassement d'un seuil de débit, aux échelles spatiales et temporelle de chaque événement étudié. Cela a conduit à introduire les évolutions suivantes :

- Tous les exutoires situés dans la fenêtre spatiale d'étude (1 174 pour l'Aude, 2 358 pour l'Ardèche et 1 093 pour le Var) sont considérés pour construire la table de contingence. Les hydrogrammes simulés (RS) avec les précipitations observées sont utilisés comme hydrogrammes de référence, et les seuils de débit sont adaptés à chaque exutoire en fonction d'une période de retour donnée.

- Un unique pas de temps est considéré pour chaque exutoire, c'est-à-dire le premier moment où l'hydrogramme de référence dépasse le seuil considéré (ou le moment du pic de l'hydrogramme de RS, en cas de non-dépassement du seuil). La réduction drastique du nombre de pas de temps considérés est compensée par la prise en compte d'un grand nombre d'exutoires.

- Toutes les échéances de prévision sont considérées simultanément pour évaluer l'anticipation du dépassement de seuil. En pratique, une détection est considérée comme prévu si le dépassement du seuil est prévu au moins une fois à l'une des six échéances de prévision considérées.

Plus concrètement, pour chaque exutoire, un seuil de débit est d'abord fixé en fonction d'une période de retour donnée. Les valeurs de débit correspondantes ont été directement extraites dans la base de données statistiques nationale SHYREG (Aubert *et al.* 2014).

Quatre cas de figure sont alors possibles et peuvent être découpés en deux catégories :

- **Catégorie n° 1 : Dépassement du seuil.** Le seuil est dépassé par la simulation de référence RS, la date (jour et heure) du premier dépassement observé est identifiée. Ce cas correspond à la Figure 5.3 ci-dessous. Les six prévisions de débit émises avant cette date (correspondant à l'échéance maximale de 6 h de chaque prévision) sont sélectionnées. Pour un percentile donné, un succès est comptabilisé dans la table de contingence si au moins une des six simulations dépasse le seuil de débit à n'importe quel moment (Figure 5.3 a)), et un échec est compté si aucun des six hydrogrammes prévus ne dépasse le seuil (Figure 5.3 b)). Ce dépassement prévu peut se produire, d'après le modèle de prévision, après l'instant effectif du dépassement observé, l'intérêt étant ici que le dépassement du seuil ait été prévu avant son occurrence. Cela s'apparente aux méthodes de voisinage discutées par Schwartz et Sobash (2017), la différence étant qu'ici, il s'agit d'un voisinage temporel.

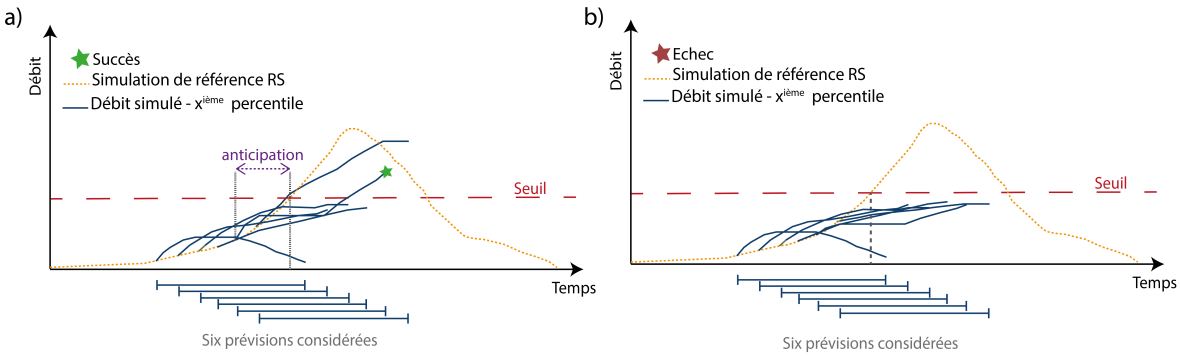

FIGURE 5.3 – *Configurations possibles en cas de dépassement de seuil par RS*

- **Catégorie n° 2 : Non-dépassement du seuil.** Le seuil de débit n'est pas dépassé par la simulation de référence RS. C'est alors la date du pic de débit de RS qui est utilisée. Cette configuration correspond à la Figure 5.4. Les six prévisions de débit émises avant cette date sont sélectionnées. Une fausse alarme est comptabilisée si au moins une des prévisions sélectionnées dépasse le seuil de débit à n'importe quelle échéance (Figure 5.4 a)), et un rejet correct si aucune des séries sélectionnées ne dépasse le seuil de débit à n'importe quelle échéance (Figure 5.4 b)).

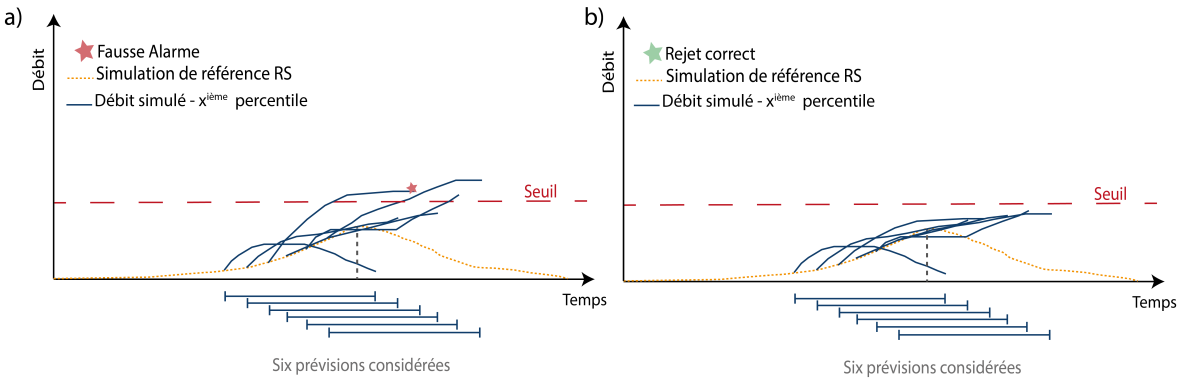

FIGURE 5.4 – *Configurations possibles en cas de non-dépassement de seuil par RS*

Un tableau de contingence, comme celui présenté ci-dessous, (Tableau 5.1) est finalement obtenu pour chaque percentile de la prévision d'ensemble des débits, chaque tableau contenant autant de valeurs (succès, échecs, fausses alarmes et rejets corrects) que le nombre d'exutoires dans la zone étudiée.

TABLEAU 5.1 – *Cas possibles pour un tableau de contingence*

	$\mathbf{Q_{RS} > Seuil}$	$\mathbf{Q_{RS} < Seuil}$
$\mathbf{Q_{percentile\ choisi} > Seuil}$	Succès	Fausse Alarme
$\mathbf{Q_{percentile\ choisi} < Seuil}$	Échec	Rejet Correct

La courbe ROC peut alors être tracée en dérivant les scores POD et FAR des tables de contingence, en suivant la méthode traditionnelle décrite ci-dessus. Le calcul des POD et FAR est explicité dans les équations 5.1 et 5.2.

$$POD = \frac{Succes}{Succes + echec} \tag{5.1}$$

$$FAR = \frac{FausseAlarme}{FausseAlarme + RejetCorrect} \tag{5.2}$$

Un avantage de cette façon de construire les tables de contingence est que les résultats de la détection à chaque exutoire peuvent être cartographiés. Cela permet d'analyser la répartition géographique des succès, échecs, fausses alarmes ou rejets corrects. De plus, pour chaque succès, le délai d'anticipation peut également être évalué. Il correspond à la différence entre l'instant du dépassement du seuil selon la simulation de référence (RS), et l'instant de la première prévision indiquant un dépassement de seuil à venir. On peut ainsi dessiner la distribution des anticipations obtenues sur les exutoires considérés, pour les différents percentiles de la prévision d'ensemble.

Exemple d'application de la méthode sur un exutoire dans le cas de l'événement de l'Aude

La Figure 5.5 présente, à titre d'exemple, les séries d'hydrogrammes obtenus pour l'exutoire n°1168 sur l'Aude. Si un seuil de période de retour 10 ans est considéré, la valeur seuil de débit à détecter est de 21,3 m ^{3}/s d'après la base de données SHYREG. Dans le cas de ce sous-bassin, ce seuil de débit est dépassé par la simulation de référence RS, à 2 h 15 le 15 octobre.

Le dépassement du seuil place l'exutoire dans le **Catégorie n° 1** décrit au-dessus. Les six prévisions de l'ensemble (ici pertDpepi) précédant le dépassement sont sélectionnés, soit les prévisions de 21 h, 22 h, 23 h, 00 h, 01 h et 2 h. Si l'on choisit de s'intéresser au percentile 50% (soit la médiane de l'ensemble - en tirets violet foncé sur la Figure 5.5), le seuil de débit (ligne horizontale pointillée) n'est jamais dépassé. Un échec est alors compté.

En considérant le percentile 75%, le seuil est dépassé pour la simulation de 1h du matin (Figure 5.5 e)) et donc un succès est compté. Il est alors possible de calculer l'anticipation liée à cette prévision pour cet exutoire, qui est ici de 1 h 15.

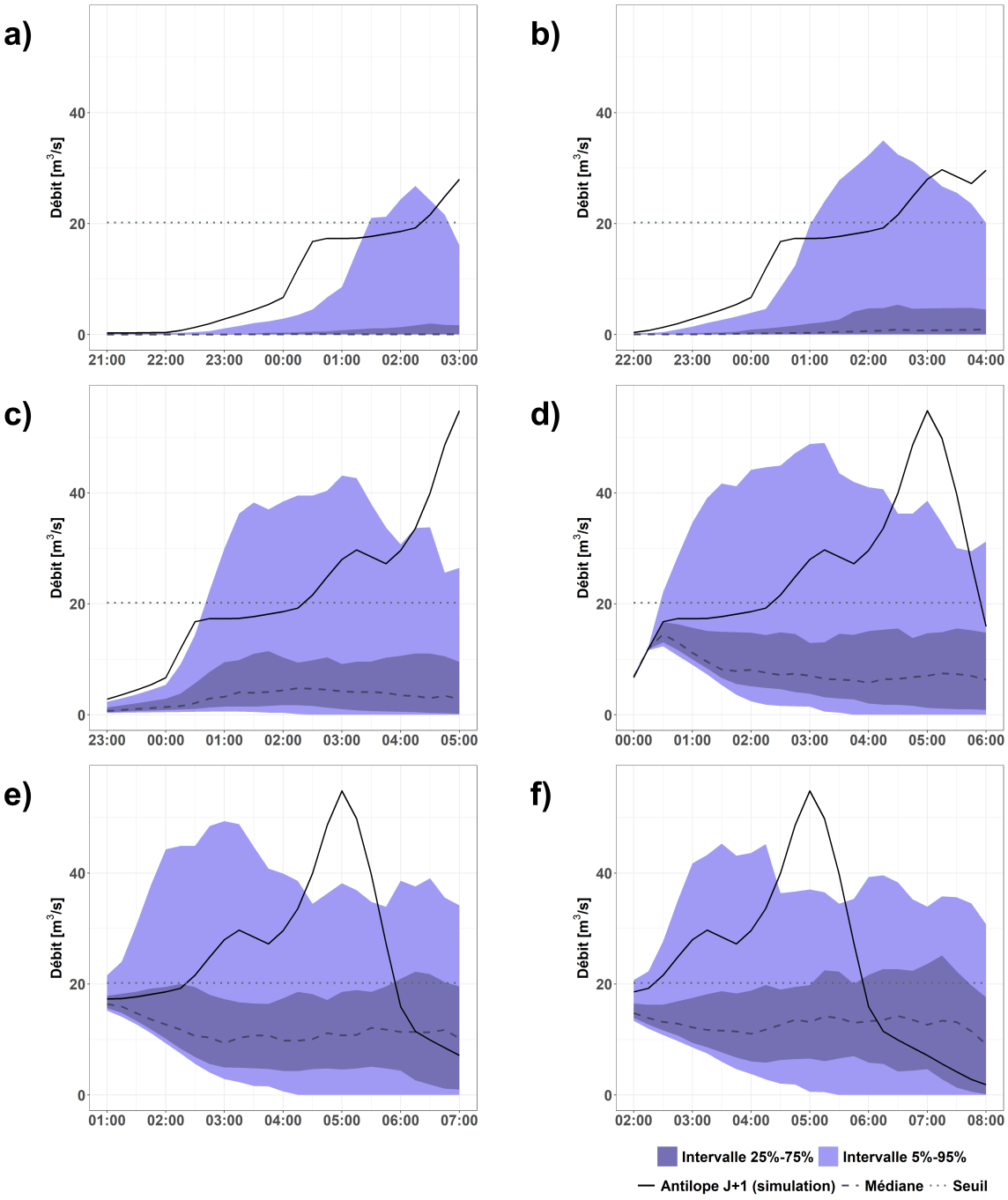

FIGURE 5.5 – *Hydrogrammes obtenus pour l'exutoire 1168, à partir des prévisions d'ensemble pertDpepi, de 21 h à 2 h le 15 octobre 2018 dans l'Aude*

5.2 Résultats de l'évaluation sur l'événement de l'Aude

5.2.1 Détection des dépassements de seuils

5.2.1.1 Courbes ROC obtenues pour les trois produits de prévision d'ensemble des pluies

Dans le cas de l'événement de l'Aude, les seuils de débits considérés correspondent à une période de retour de 10 ans, et ont été estimés grâce à la base de données SHYREG. Le choix de cette période de retour s'explique par deux raisons principales :

- Il s'agit d'une période de retour pour lesquels des débordements de cours d'eau et des dommages importants sont susceptibles d'être observés sur de nombreux cours d'eau.
- Près de la moitié du bassin de l'Aude a été touchée par des crues dont le débit de pointe dépasse la période de retour 10 ans, d'après la simulation RS. Ainsi, les échantillons de cours d'eau affectés ou non par des crues dépassant la valeur seuil auront des tailles similaires sur la zone étudiée.

La Figure 5.6 illustre les courbes ROC obtenues pour les trois ensembles de prévisions hydrométéorologiques testés. La valeur ajoutée de l'ensemble pepi par rapport à l'ensemble AROME-EPS, qui a déjà été observée sur l'analyse des précipitations prévues (Section 4.2.1), est à nouveau clairement visible ici : pour un percentile d'ensemble sélectionné, les deux ensembles fournissent des ratios de fausses alarmes similaires, mais l'ensemble pepi a une probabilité de détection plus élevée. Ceci est illustré sur la Figure 5.6 sur la base des valeurs affichées pour le percentile de 75%.

Pour une probabilité de détection équivalente (POD), la probabilité de fausse alarme (FAR) est également réduite pour pertDpepi par rapport à pepi. Le losange gris correspondant à RF0 permet de visualiser la plus-value des trois produits de prévision évalués puisque même pour le percentile 5%, le POD est meilleur pour les ensembles que pour RF0, avec un taux de fausses alarmes nul.

5.2.1.2 Analyse cartographique de l'anticipation du dépassement de seuil sur le percentile 75%

La distribution spatiale des échecs et des fausses alarmes, illustrée sur la Figure 5.7 pour le cas du quantile 75%, offre un point de vue plus détaillé sur les résultats d'anticipation des dépassements de seuil.

La comparaison avec la simulation RF0 (Figure 5.7 a)) permet d'évaluer la valeur ajoutée des prévisions d'ensemble par rapport au cas d'une absence de prévisions de pluie. Ainsi, sur la zone centrale des pluies intenses observées, on retrouve un certain nombre de succès, même dans le cas de RF0. Ce résultat provient du temps de réponse des bassins et s'observe principalement sur les bassins où l'on retrouve le réseau hydrographique principal (avec des surfaces drainées les plus importantes).

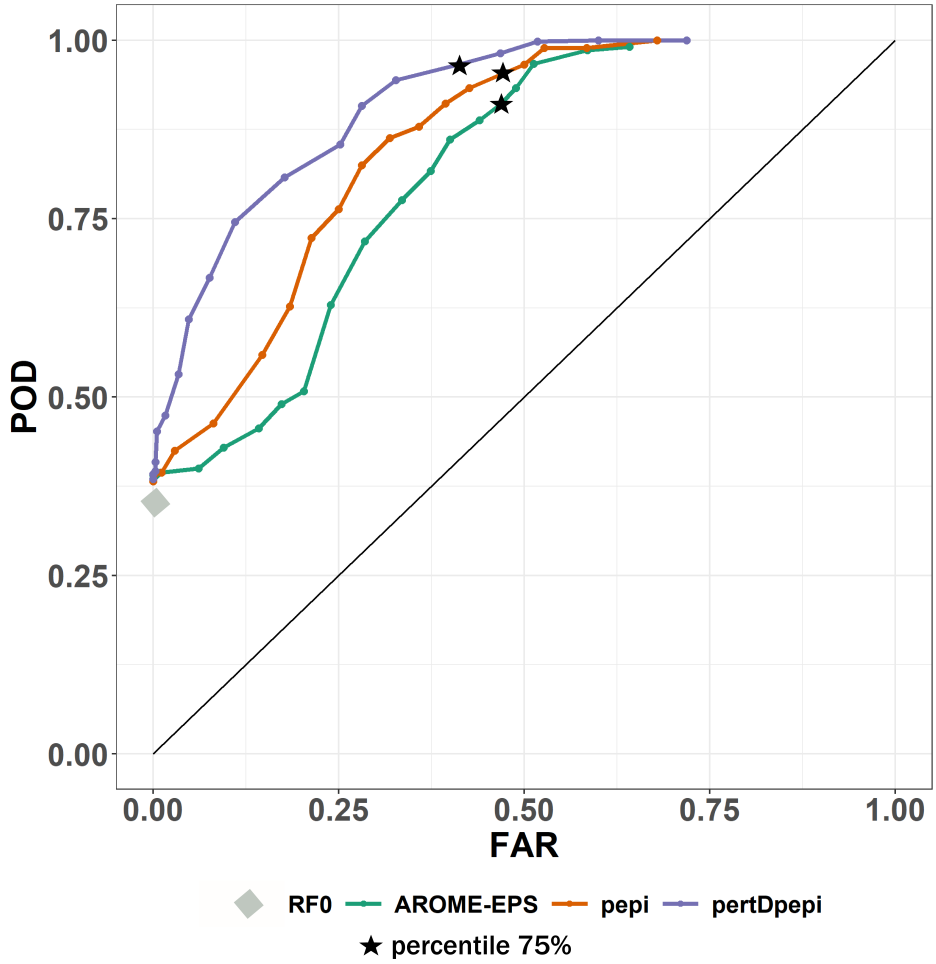

FIGURE 5.6 – *Courbes ROC obtenues pour les trois ensembles et RF0 pour la période de retour 10 ans. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0*

L'apport des trois produits de prévision d'ensemble par rapport à RF0, déjà noté sur la Figure 5.6, apparaît d'autant plus nettement sur les Figures 5.7 b), c) et d), avec une disparition quasi complète des exutoires en échec (en rouge foncé) dans la zone de pluies intenses. Néanmoins, ce gain en matière de détection s'accompagne d'une augmentation du nombre de fausses alarmes (en rouge clair) au sud-est du territoire d'étude. Concernant la comparaison entre les produits, les exutoires en échec sont moins nombreux avec pepi qu'AROME-EPS dans la zone des fortes pluies (Figures 5.7 b) et c)). L'ensemble pertDpepi conduit étonnamment à une réduction significative du nombre de fausses alarmes (255) (Figure 5.7 d)) en comparaison avec les deux autres ensembles (290 pour AROME-EPS et 292 pour pepi) pour un taux de détection proche de celui de pepi, ce qui n'était pas évident à prévoir à partir de l'analyse des pluies prévues (Figures 4.8 d), g) et j)). Ces conclusions ne sont néanmoins valables que pour le percentile 75%. On peut par exemple remarquer sur la Figure 5.6 que le taux de fausses alarmes est toujours plus bas pour pertDpepi, sauf pour le percentile 95%.

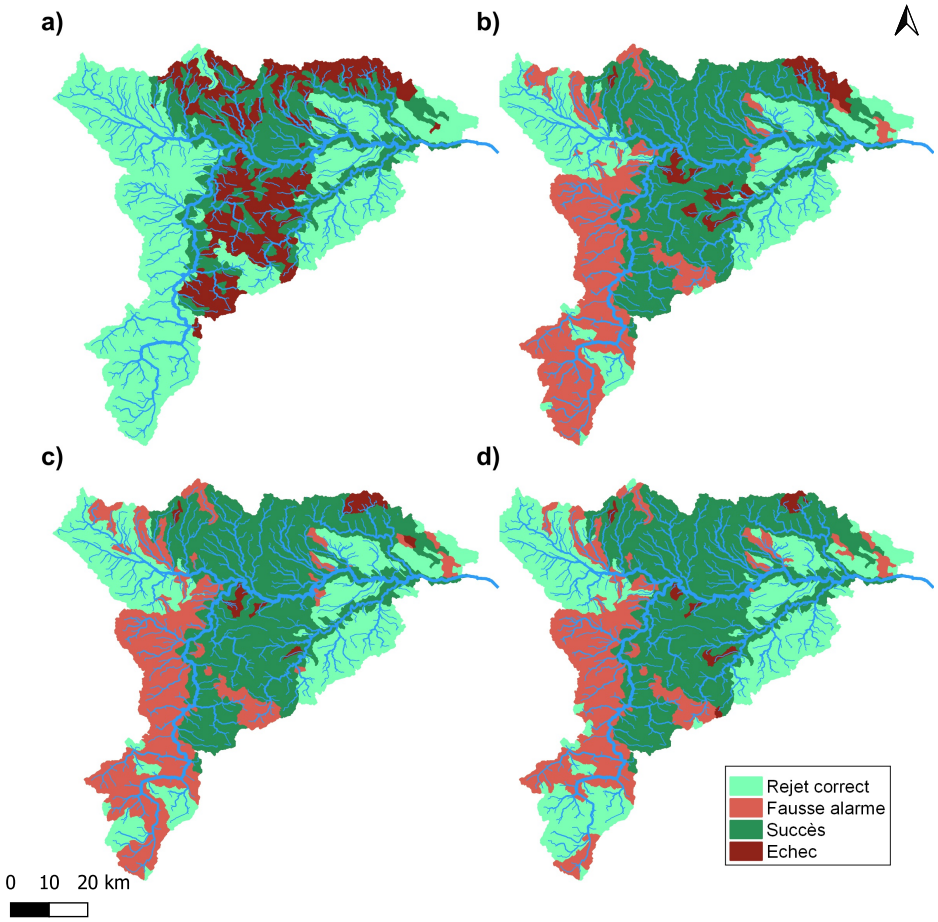

FIGURE 5.7 – *Cartes d'anticipation (0-6 h) du seuil de période de retour 10 ans : a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation minimale de 15 minutes*

5.2.2 Délai d'anticipation pour les trois produits de prévision d'ensemble

Après avoir vu que les trois produits de prévision d'ensemble, AROME-EPS, pepi et pertDpepi, permettaient d'améliorer la détection des dépassements de seuil par rapport à RF0 (Section 5.2.1), il est intéressant de voir dans quelles mesures ils ont amélioré les temps d'anticipation. La Figure 5.8 illustre l'équilibre à trouver entre l'amélioration des délais d'anticipation des prévisions, et les performances de détection (succès, fausses alarmes et échecs).

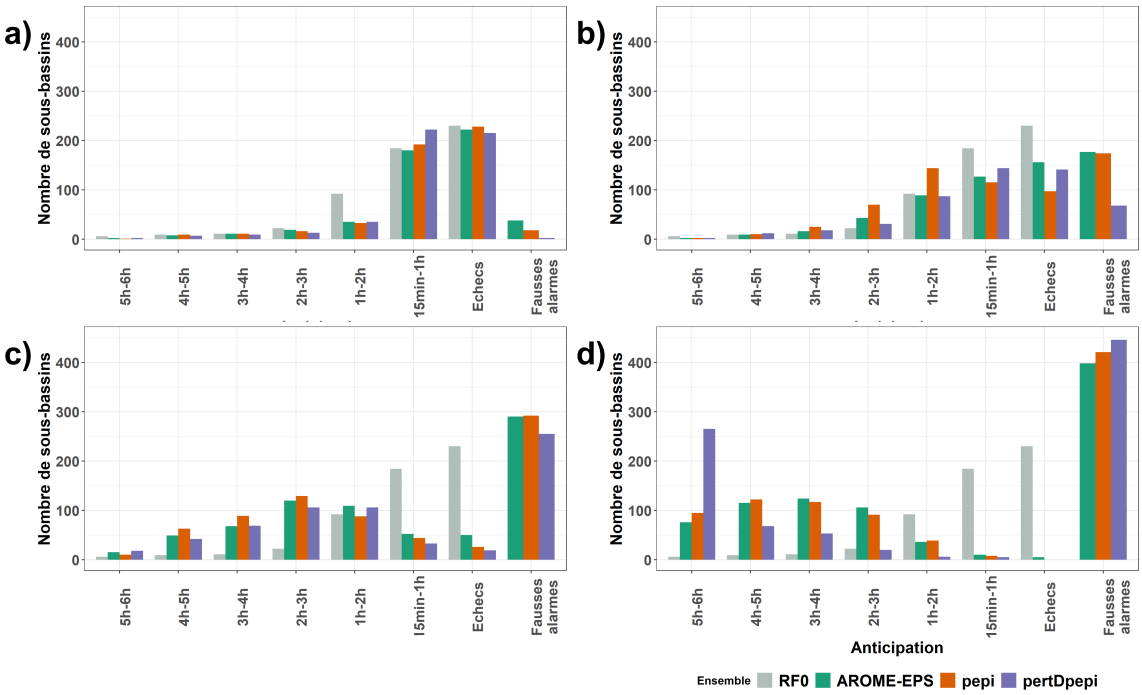

FIGURE 5.8 – *Comparaison des délais d'anticipation du dépassement de seuil de période de retour 10 ans pour les percentiles a) 20%, b) 50%, c) 75% et d) 95%*

Dès le percentile 20% (Figure 5.8 a)), les temps d'anticipation obtenus avec les ensembles sont globalement plus élevés et augmentent d'autant plus que le percentile considéré est élevé (Figures 5.8 b), c) et d)). Néanmoins, comme déjà observé précédemment, si regarder un percentile assez grand permet d'augmenter la possibilité d'anticipation et réduire le nombre de cas où le dépassement n'est pas anticipé (nombre d'échecs), cela fait également considérablement augmenter le nombre de fausses alarmes, quel que soit l'ensemble considéré. Pour un faible percentile (20% pour la Figure 5.8 a)), le nombre de fausses alarmes est proche de 0, mais le nombre d'échecs se situe autour de 220 (1174 bassins étudiés au total), et l'anticipation possible dépasse rarement les 2-3 h. Au contraire, si l'on regarde le percentile 75%, le nombre d'échecs est quasiment réduit à 0, et le nombre de bassins ayant une anticipation supérieure à 3 h augmente considérablement, passant pour AROME-EPS (respectivement pepi et pertDpepi) de 21 sous-bassins (respectivement 21 et 22 sous-bassins) à 132 sous-bassins (respectivement 162 et 129 sous-bassins). Mais le nombre de fausses alarmes se situe dans ces cas aux alentours de 400. Enfin, la Figure 5.8 illustre également le fait qu'il est assez difficile de distinguer les trois produits de prévision de pluie utilisés du point de vue des délais d'anticipation offerts.

5.2.3 Analyse détaillée des prévisions pour quelques exutoires

Pour compléter les résultats de l'évaluation basée sur l'anticipation des dépassements de seuil, une analyse détaillée des prévisions de débits obtenues avec le modèle Cinecar est réalisée ici, au niveau de six exutoires sélectionnés et couvrant l'ensemble des cas possibles (succès, échec, fausse alarme, rejet correct et succès avec RF0). La Figure 5.9 présente les six exutoires retenus.

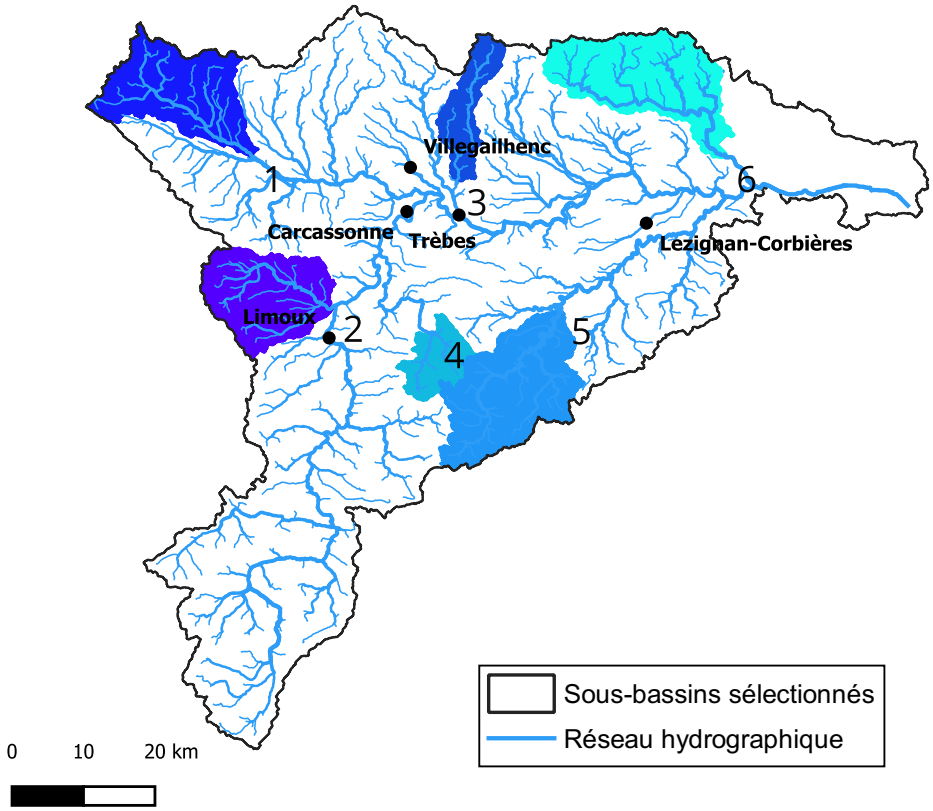

FIGURE 5.9 – *Localisation des six exutoires retenus pour une analyse détaillée des prévisions de débits pour l'événement de l'Aude*

Exutoires avec une faible réaction hydrologique - Exutoires n°1 et n°2

Les exutoires n°1 et n°2 correspondent à des secteurs à faibles réactions hydrologiques, c'est-à-dire que le débit de pointe du scénario de référence RS reste largement inférieur au seuil de débit décennal dans les deux cas.

La Figure 5.10 représente les hydrogrammes obtenus pour la prévision de 4 h du matin le 15 octobre, pour les trois produits de prévision d'ensemble et pour l'exutoire n°1.

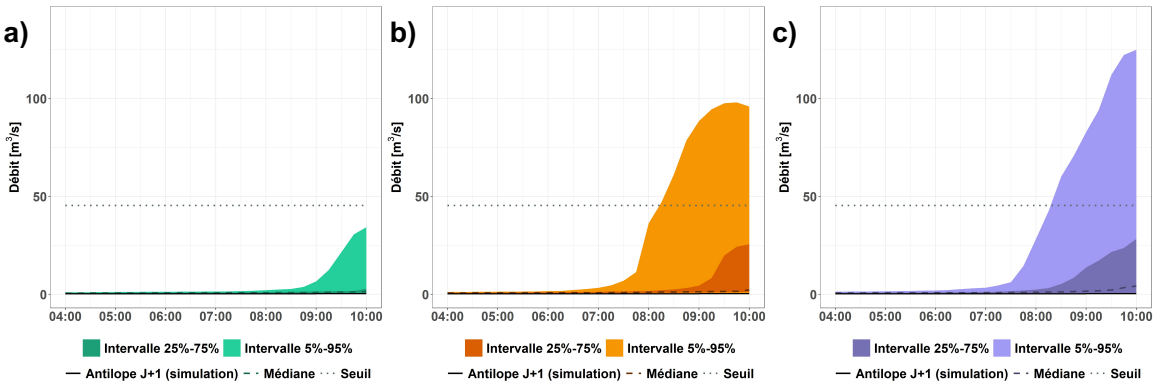

FIGURE 5.10 – *Hydrogrammes de prévisions obtenues avec le modèle Cinecar pour l'exutoire n°1 (216 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

L'exutoire n°1 se situe dans une zone de rejet correct pour les trois produits (Figures 5.7 b), c) et d)) pour le percentile 75%. Cela se retrouve sur les hydrogrammes présentés. L'exutoire se trouve à proximité de zones touchées par les pluies intenses et l'intervalle de pertDpepi est logiquement plus large (effet des perturbations spatiales). Un autre constat observable est la plus grande dispersion des produits pepi et pertDpepi par rapport à AROME-EPS qui amène à dépasser le seuil de débit pour des percentiles élevés (supérieurs à 75%).

La Figure 5.11 présente les hydrogrammes obtenus pour la prévision de 4 h du matin le 15 octobre, pour les trois produits de prévision d'ensemble et pour l'exutoire n°2.

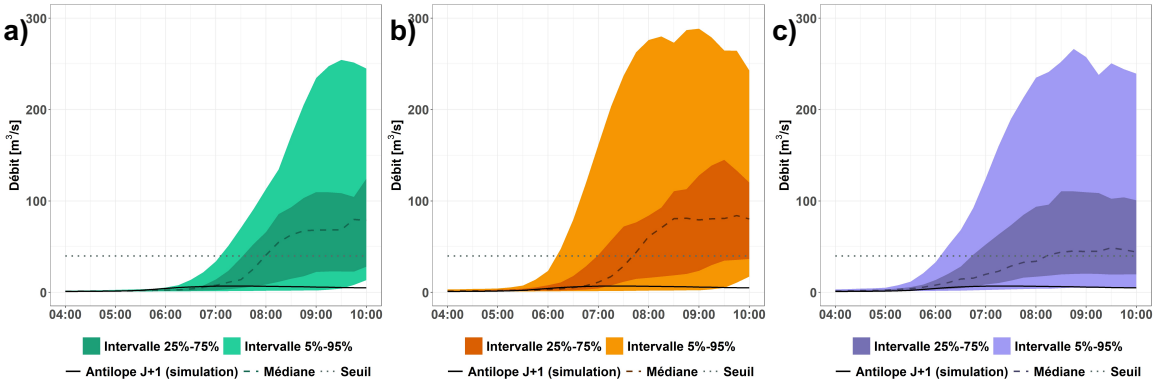

FIGURE 5.11 – *Hydrogrammes obtenus pour l'exutoire n°2 (197 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Contrairement à l'exutoire n°1, l'exutoire n°2 se trouve dans la zone de fausses alarmes (Figures 5.7 b), c) et d)) pour le percentile 75%. Ces fausses alarmes sont observées pour les trois produits de prévision d'ensemble. Une forte surestimation est observée, particulièrement pour les percentiles supérieurs. Cette surestimation présente sur les hydrogrammes explique les fausses alarmes observées sur la Figure 5.7. Ce résultat est directement lié au lissage des prévisions d'ensemble observé sur le cumul pluviométrique présenté dans la Section 4.2.1. La dispersion des ensembles pepi et pertDpepi est confirmée par les hydrogrammes ainsi que sur la courbe ROC avec une forte augmentation du taux de fausses alarmes 5.6, mais reste

particulièrement visible pour les percentiles extrêmes (5% et 95%) tandis que les percentiles allant de 20% à 75% montrent une différence limitée avec le produit de l'ensemble AROME-EPS.

Exutoires de bassins amonts dans la zone de précipitations intenses - Exutoires n°3 et n°4

Les exutoires 3 et 4 correspondent à des sous-bassins situés dans la zone de précipitations intenses observées, et dans lesquelles les dépassements du seuil décennal de débit peuvent difficilement être détectés à l'avance en l'absence de prévisions pluviométriques (temps d'anticipation limités pour la prévision RF0 du fait de la surface limitée des bassins versants amonts).

L'exutoire n°3 se situe dans la zone centrale, dans laquelle les prévisions d'ensemble se traduisent par des succès. La Figure 5.12 illustre les hydrogrammes obtenus pour les trois produits de prévision d'ensemble, pour la prévision de 02 h le 15 octobre.

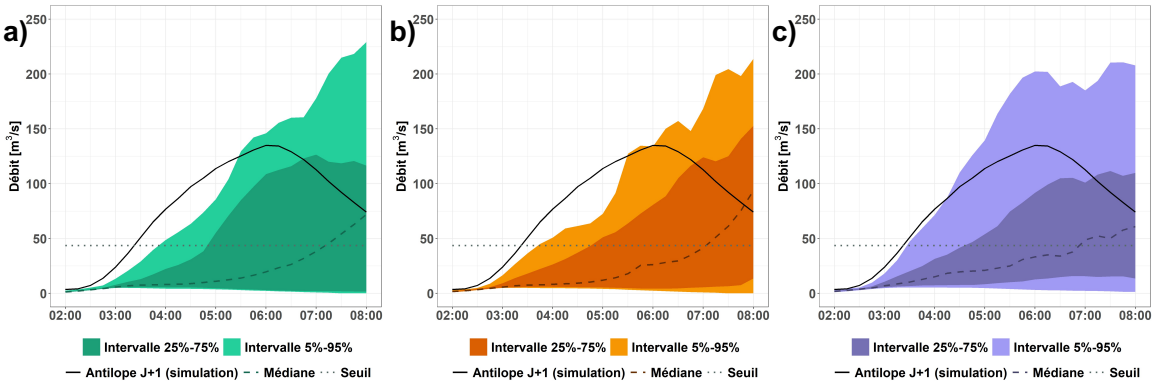

FIGURE 5.12 – *Hydrogrammes obtenus pour l'exutoire n°3 (85 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Le succès est très facilement observable pour les trois produits de prévision, le dépassement de seuil ayant lieu dès le début de la montée de crue. Le premier élément qui ressort est le retard du dépassement de seuil de débit pour les prévisions, par rapport à l'observation. En effet, le retard est de 1 h 30 pour les ensembles AROME-EPS et pepi, et de 1 h 15 pour pertDpepi pour le percentile 75%. Le deuxième élément remarquable est la plus forte dispersion des percentiles élevés (supérieurs à 75%) pour l'ensemble pertDpepi, dispersion qui, par ailleurs, permet de diminuer le retard du dépassement de seuil de 15 minutes par rapport à la simulation de référence RS.

L'exutoire n°4 se situe dans la même zone que l'exutoire n°3 (Figure 5.9), mais plus au sud. La Figure 5.13 illustre les hydrogrammes obtenus pour les trois produits de prévision d'ensemble pour la prévision de 01 h le 15 octobre.

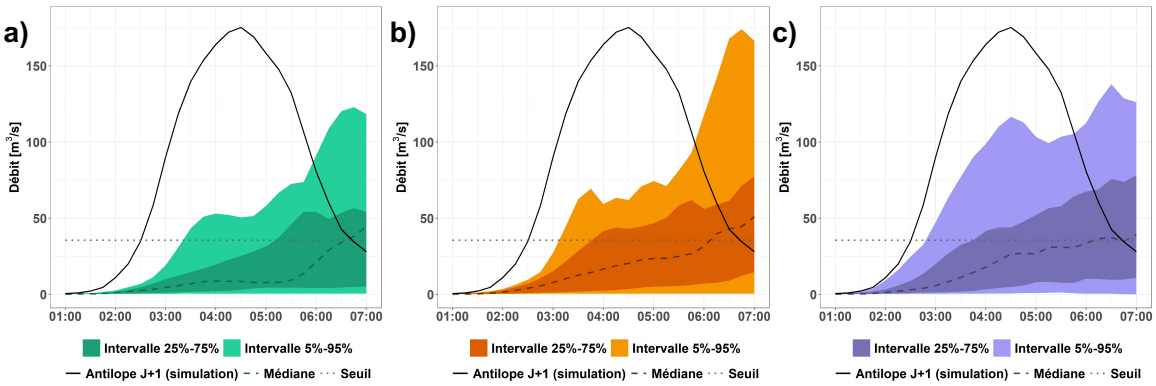

FIGURE 5.13 – *Hydrogrammes obtenus pour l'exutoire n°4 (173 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Le retard est beaucoup plus marqué dans le cas de cet exutoire et est associé dans ce cas à une sous-estimation significative de l'ampleur de l'événement. Ceci semble être lié à la tendance générale des ensembles de pluie à sous-estimer les intensités et les cumuls de pluie les plus élevés dans le cas de l'événement de l'Aude (Section 4.2.1). La sous-estimation est légèrement corrigée pour pertDpepi.

Exutoires de bassins avals dans la zone de précipitations intenses - Exutoires n°5 et n°6

Les exutoires n°5 et n°6 sont situés en aval du domaine d'étude, dans les secteurs où même sans prévisions de précipitations (RF0), le dépassement de seuil est correctement anticipé. Ce sont des cas où les prévisions d'ensemble peuvent simplement aider à augmenter les délais d'anticipation.

La Figure 5.14 montre les trois hydrogrammes de prévisions obtenus pour l'exutoire n°5 (Figure 5.9), pour la prévision de 01 h le 15 octobre.

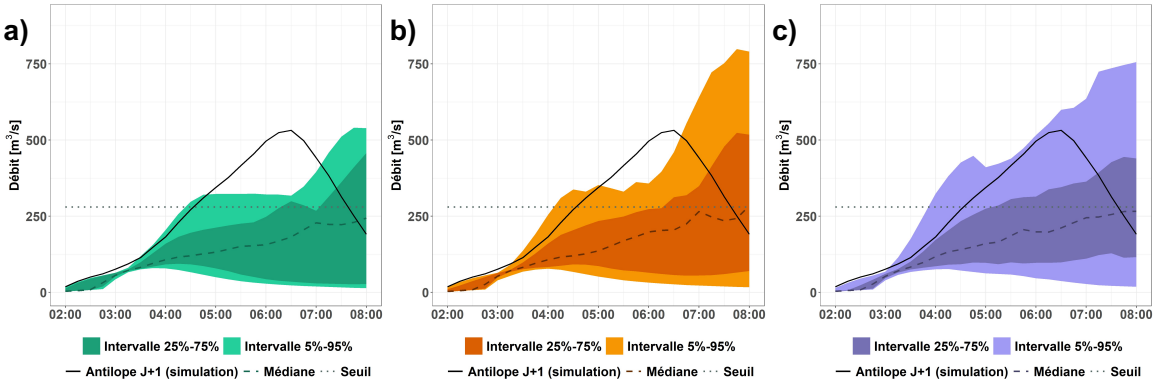

FIGURE 5.14 – *Hydrogrammes obtenus pour l'exutoire n°5 (263 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Un retard important entre prévisions et la simulation de référence est observé avec un décalage de 1 h 45 pour AROME-EPS et pepi, et 30 min pour pertDpepi. C'est également un cas où la plus-value de la grande dispersion de pertDpepi apparait en termes d'anticipation.

De plus, en s'intéressant au débit de pointe, pertDpepi est l'ensemble qui s'en approche le plus rapidement et qui l'atteint pour les percentiles les plus élevés (supérieurs à 75%).

Les trois hydrogrammes obtenus pour l'exutoire n°6 (Figure 5.9) pour la prévision de 06 h le 15 octobre sont présentés sur la Figure 5.15.

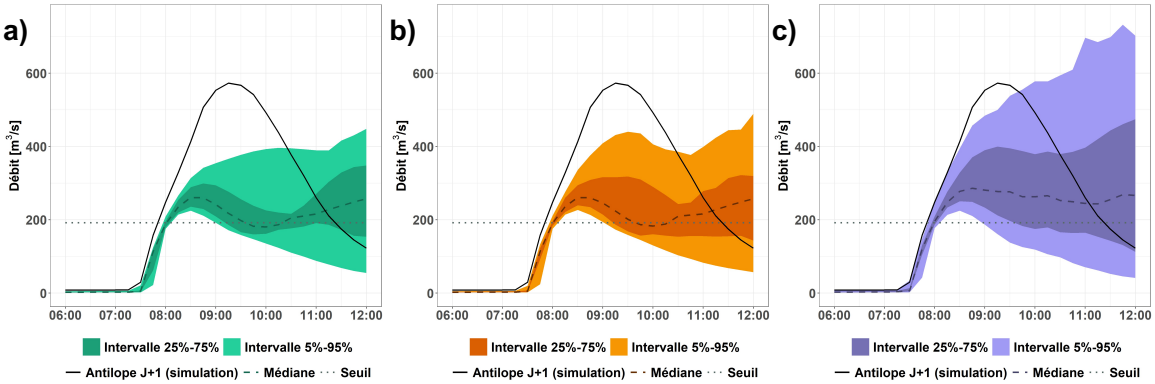

FIGURE 5.15 – *Hydrogrammes obtenus pour l'exutoire n°6 (257 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Le succès est très facilement observable pour les trois produits de prévision, le dépassement de seuil ayant lieu dès le début de la montée de crue. Contrairement aux autres cas de succès, il n'y a pas dans ce cas de retard entre la simulation de référence RS et les trois produits de prévision. Cela s'explique principalement par le fait que le seuil est assez facilement dépassé même pour RF0 (Figure 5.7). Cependant, là encore, il semble difficile de prévoir correctement l'intensité du débit de pointe, en particulier pour les ensembles AROME-EPS et pepi. Cette Figure met 5.15 également en évidence le fait que les modèles de prévision de pluie ont du mal à anticiper les pluies les plus intenses à très courte échéance.

5.3 Résultats de l'évaluation sur l'événement de l'Ardèche

5.3.1 Détection des dépassements de seuils de débits

5.3.1.1 Choix de la période de retour des seuils de débits

Les critères et les cartes présentés pour l'évaluation des prévisions d'ensemble dans le cas de l'événement d'août 2018 incluent l'ensemble des 2358 sous-bassins modélisés avec Cinecar.

Cet événement, contrairement à celui de l'Aude, est d'une moindre intensité, et les seuils de débits de période de retour 10 ans ne sont pas dépassés. La Figure 5.16 ci-dessous présente les résultats cartographiques de détection pour RF0, et pour des seuils correspondant à des débits de périodes de retour 2 ans, puis à ces mêmes seuils multipliés par 0.75 et 0.5.

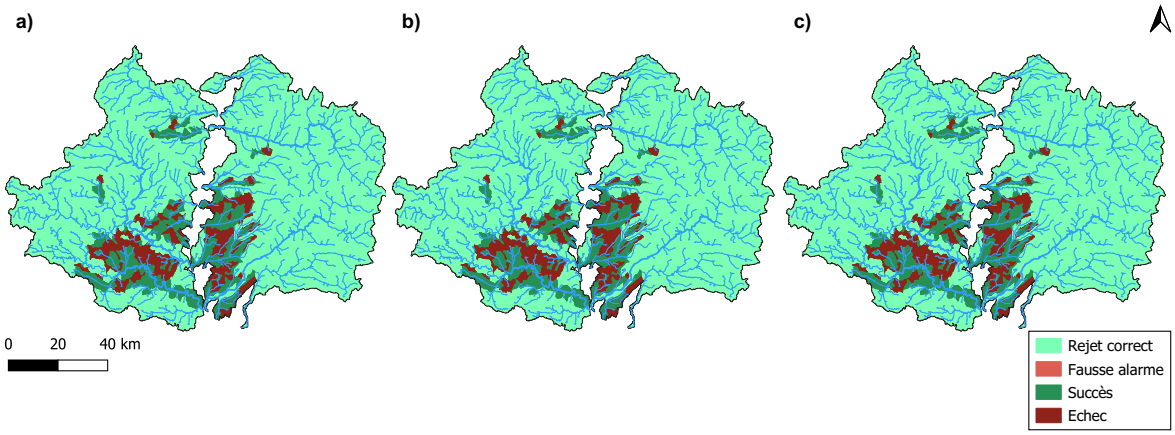

FIGURE 5.16 – *Résultats de l'anticipation pour RF0 pour différents seuils de débit a) Débit de période de retour 2 ans Q_2, b) 0.75 * Q_2 et c) 0.5 * Q_2*

Le constat est que peu de bassins versants contenus dans la zone d'étude réagissent, seuls ceux vraiment au cœur de la zone de précipitations intenses sont concernés, même pour un seuil de débit très faible. Ainsi, même si le secteur géographique étudié est défini comme la zone de pluie intense observée et prévue (Section 4.2.1), les précipitations ne sont ici pas suffisantes pour entraîner un dépassement de seuil de débit pour un grand nombre de sous-bassins. Pour maximiser le nombre de sous-bassins avec dépassement, il faudrait abaisser le seuil de dépassement de débit. Néanmoins, à trop baisser la valeur du seuil, celui-ci perd de son sens pour deux raisons :

- Un seuil de débit trop faible ferait sortir du contexte d'évaluation sur des crues significatives, qui est visé ici.
- Choisir une période de retour trop faible fait également diminuer la capacité de détection dans la zone de pluie intense (fausses alarmes probables).

La période de retour retenue pour l'évaluation est donc de $0.75*Q_2$, qui semble être un bon compromis pour le nombre de sous-bassins dépassant le seuil.

5.3.1.2 Courbes ROC obtenues pour les trois produits de prévision d'ensemble

Pour rappel, le territoire d'étude de l'événement de l'Ardèche a été découpé en 2358 sous-bassins.

La Figure 5.17 représente les courbes ROC obtenues pour les trois ensembles de prévision de débits testés.

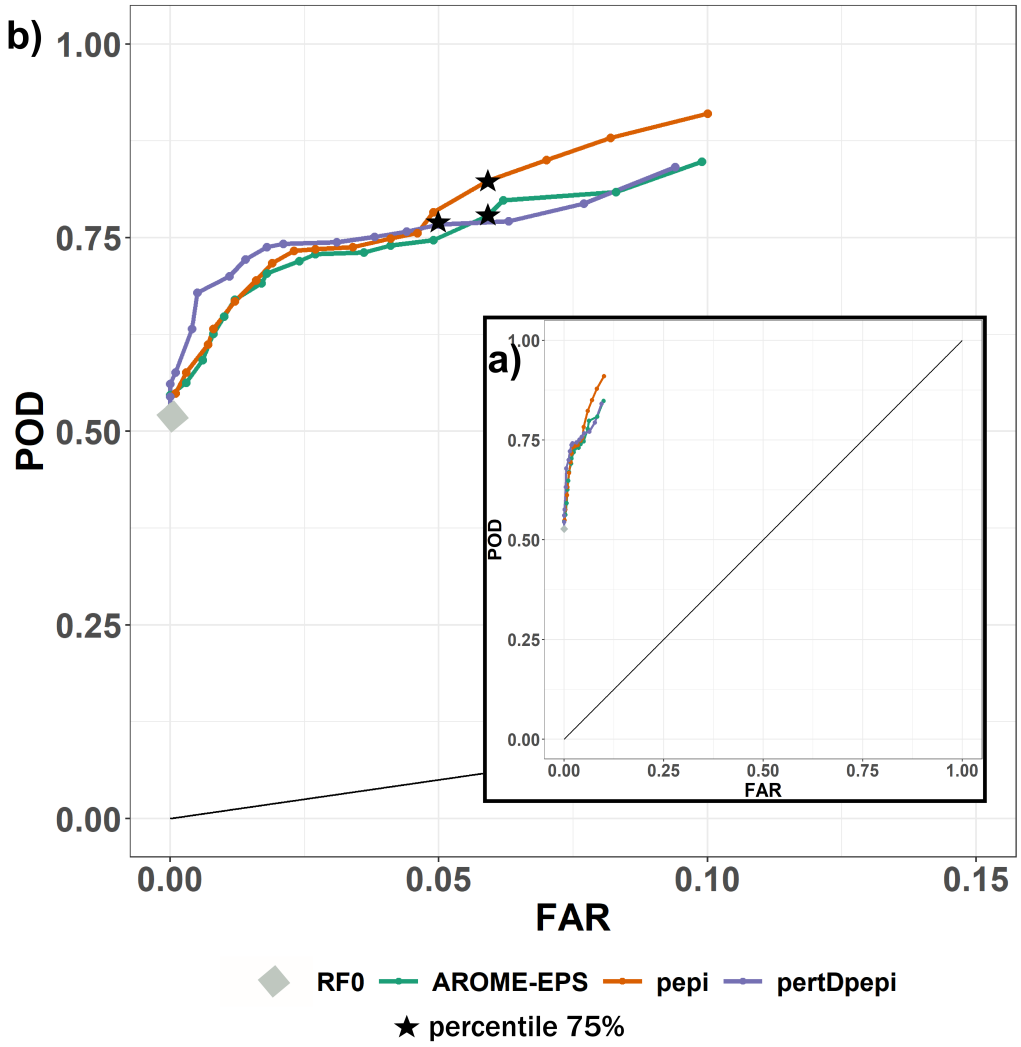

FIGURE 5.17 – *Courbes ROC obtenues pour les trois ensembles de prévision des débits, et RF0 pour la période de retour 0.75*Q_2. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0 avec a) Vue globale des valeurs des courbes ROC et b) Zoom sur les valeurs obtenues.*

La différence entre les trois produits de prévision apparaît moins nettement sur cet événement, par rapport à ce qui a été observé pour l'Aude (Section 5.2.1). Ainsi, si l'on considère les premiers percentiles (jusqu'à la médiane), les courbes d'AROME-EPS et pepi sont très proches et pertDpepi se distingue un peu avec un taux de fausses alarmes plus faible. Au contraire, pour les percentiles les plus élevés (au-delà de la médiane), les courbes ROC pour AROME-EPS et pertDpepi sont semblables et c'est pepi qui s'améliore avec une probabilité de détection plus grande pour un taux de fausses alarmes équivalent. Cela ne semble pas très étonnant d'après la Figure 4.10 sur laquelle, entre les trois produits, peu de différences étaient observées en termes de quantités de précipitations prévues. Comme pour l'événement de l'Aude, les ensembles de prévision montrent une plus-value par rapport au scénario RF0. Toujours en comparaison avec l'Aude, le taux de fausses alarmes est, pour l'Ardèche, beaucoup plus faible.

5.3.1.3 Analyse cartographique de l'anticipation du dépassement de seuil pour le percentile 75%

La distribution spatiale de l'anticipation du dépassement de seuil de débit de période de retour $0.75*Q_2$ pour le percentile 75% est présentée sur la Figure 5.18.

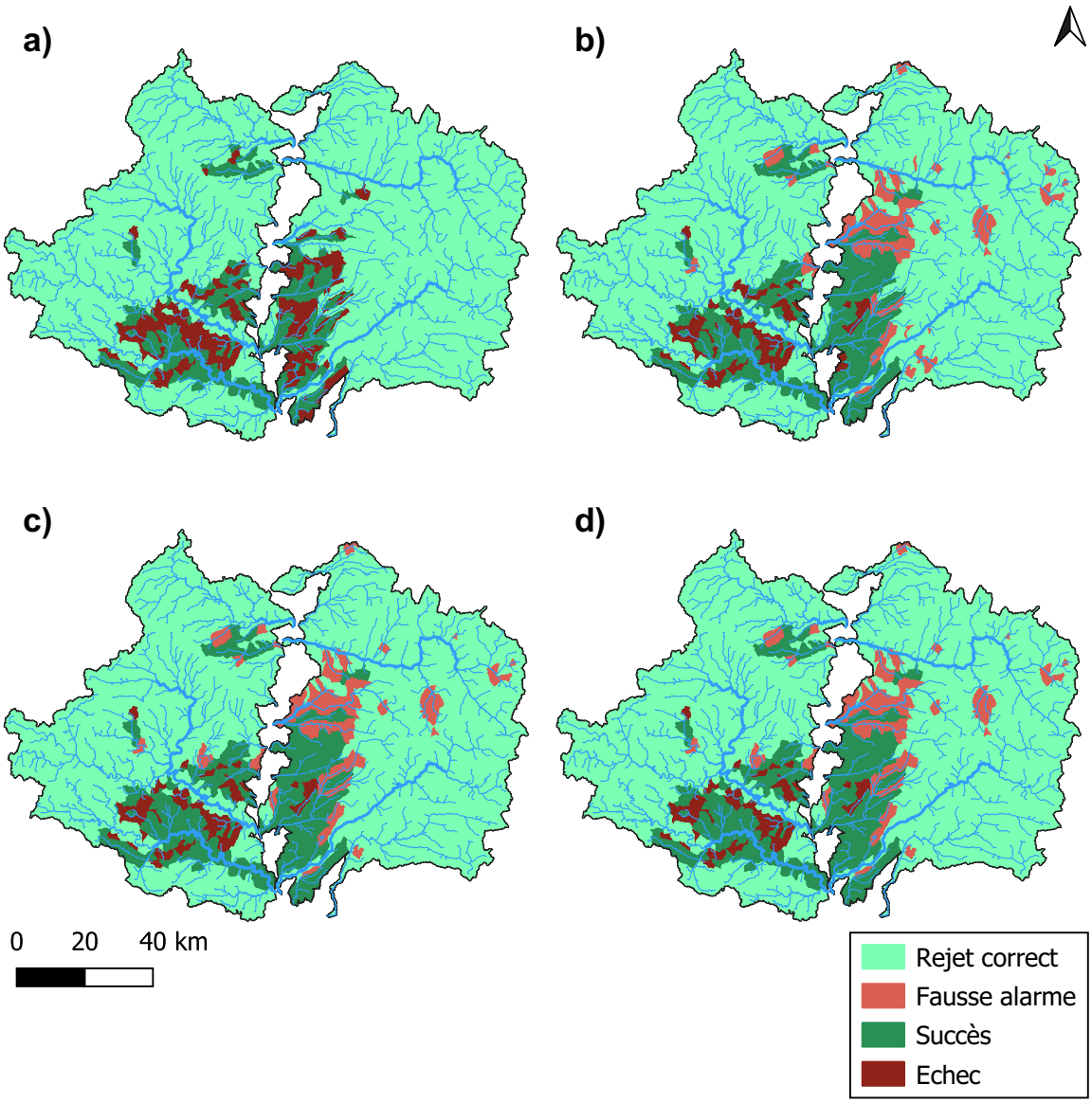

FIGURE 5.18 – *Cartes d'anticipation (0-6 h) du seuil de période de retour $0.75*Q_2$: a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation supérieure à 15 minutes*

Comme cela a déjà été noté dans la Section 5.3.1.1, peu de bassins réagissent au seuil de débit considéré, ceux réagissant sont principalement situés sur l'Ardèche et la Cèze. Une majorité de sous-bassins se trouve donc dans le cas d'un rejet correct (77% des sous-bassins pour AROME-EPS). Par rapport à l'évaluation menée avec le même percentile (75%) pour l'événement de l'Aude, le nombre de fausses alarmes est très réduit ($\approx$ 5%) pour cet épisode.

En comparant le résultat de l'évaluation pour RF0 (Figure 5.18 a)) avec celle des trois ensembles (Figures 5.18 b), c) et d)), le constat est qu'un grand nombre de succès est lié uni-

quement aux phénomènes de propagation amont-aval et non pas à la qualité de la prévision des pluies. Enfin, une zone de fausses alertes à la limite est de la zone de précipitations intenses est observable pour les trois produits de prévision d'ensemble. Les erreurs de localisation de la pluie prévue (4.10) se retrouvent, impactant les possibilités d'anticipation de dépassement d'un seuil de débit. Cependant, ce problème de localisation semble moins impactant en s'intéressant aux débits, avec en particulier un faible taux de fausses alarmes.

Pour ce qui est des performances des ensembles et en prenant appui sur les courbes ROC, il est assez difficile de conclure sur une éventuelle différence entre les trois ensembles. Néanmoins, pour le percentile 75%, pepi semble avoir le meilleur rapport détection/fausse alarme. Les trois ensembles permettent cependant une meilleure détection du dépassement de seuil, en particulier sur le secteur est de la zone d'étude, par rapport au scénario RF0.

5.3.2 Délais d'anticipation pour les trois produits de prévision d'ensemble

La Figure 5.19 illustre les délais d'anticipation obtenus pour les trois prévisions d'ensemble et pour la prévision RF0.

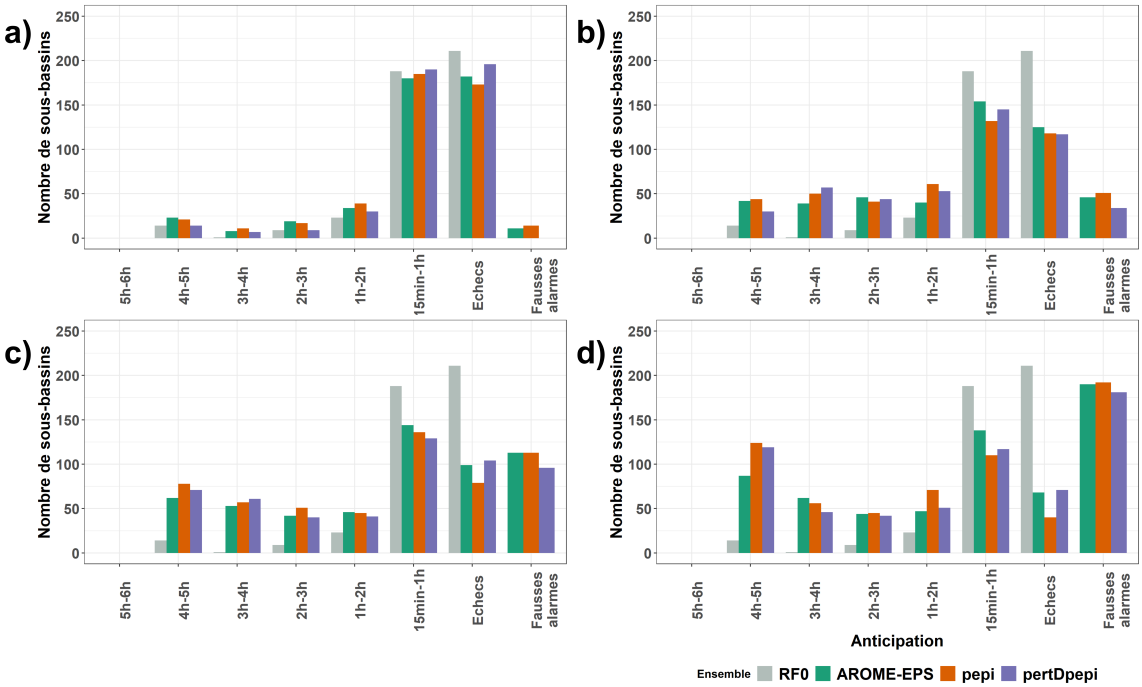

FIGURE 5.19 – *Comparaison des délais d'anticipation du dépassement de seuil de période de retour 0,75*Q_2 pour les percentiles a) 20%, b) 50%, c) 75% et d) 95% pour l'événement d'août 2018.*

Le premier constat pour le scénario RF0 est le délai d'anticipation excède rarement l'heure, avec près de 80% des exutoires pour lesquels un succès est observé ayant un délai d'anticipation inférieur à une heure. Même si les taux de détection ne sont pas significatifs pour les prévisions d'ensemble de cette étude de cas, les prévisions se révèlent intéressantes sur les délais d'anticipation.

De manière assez prévisible, une augmentation des délais d'anticipation avec l'augmentation du percentile regardé est nettement visible, avec des délais d'anticipation allant jusqu'à 5 h pour les percentiles élevés. Pour AROME-EPS (respectivement pepi et pertDpepi), le nombre d'exutoires avec un délai d'anticipation entre 4 et 5 h passe de 212 sous-bassins (respectivement 189 et 141) pour le percentile 20% (Figure 5.19 a)) à 542 sous-bassins (respectivement 731 et 755) pour le percentile 95% (Figure 5.19 b), c) et d)). Cependant, ce gain d'anticipation va de pair avec l'augmentation du nombre de fausses alarmes, passant de 12 sous-bassins pour le percentile 20% et pour AROME-EPS (respectivement 12 pour pepi et 0 pour pertDpepi) à 377 sous-bassins pour le percentile 95% AROME-EPS (respectivement 189 pour pepi et pertDpepi). Le nombre d'échecs reste relativement réduit pour cet événement. Ainsi, pour un faible percentile (20% - Figure 5.19 a)), le nombre d'échecs pour AROME-EPS (respectivement pepi et pertDpepi) est autour de 189 sous-bassins (respectivement 165 et 189).

En ce qui concerne la distinction entre les trois produits de prévision, il est difficile de tirer des conclusions définitives concernant les délais d'anticipation offerts, la différence entre les produits étant d'une dizaine d'exutoires.

5.3.3 Analyse détaillée pour quelques exutoires

Quelques exutoires ont été sélectionnés pour illustrer via des hydrogrammes la capacité d'anticipation des dépassements de seuil de débits. Quatre exutoires ont été retenus et sont présentés sur la Figure 5.20.

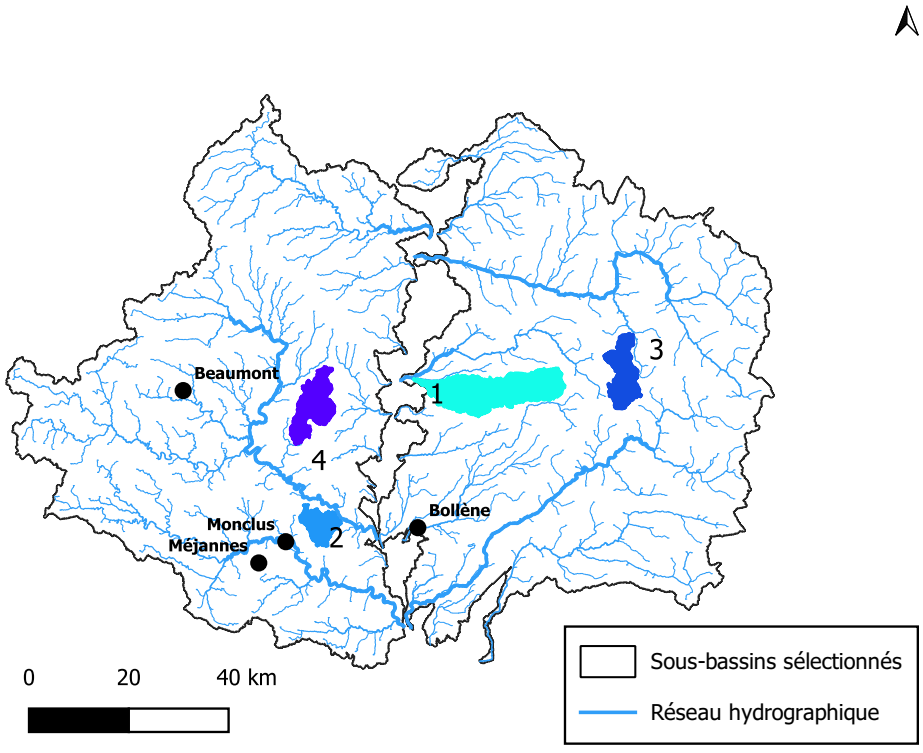

FIGURE 5.20 – *Localisation des quatre exutoires retenus pour une analyse détaillée des prévisions des débits, pour l'événement de l'Ardèche*

Exutoires avec une forte réaction hydrologique - Exutoires n°1 et n°2

Les exutoires n°1 et n°2 correspondent à des secteurs avec une réaction hydrologique assez forte puisque le seuil retenu est largement dépassé par la simulation de référence à partir de la pluie observée (RS). La Figure 5.21 représente les hydrogrammes obtenus pour la prévision de 8 h du matin le 9 août, pour les trois produits de prévision d'ensemble, et pour l'exutoire n°1.

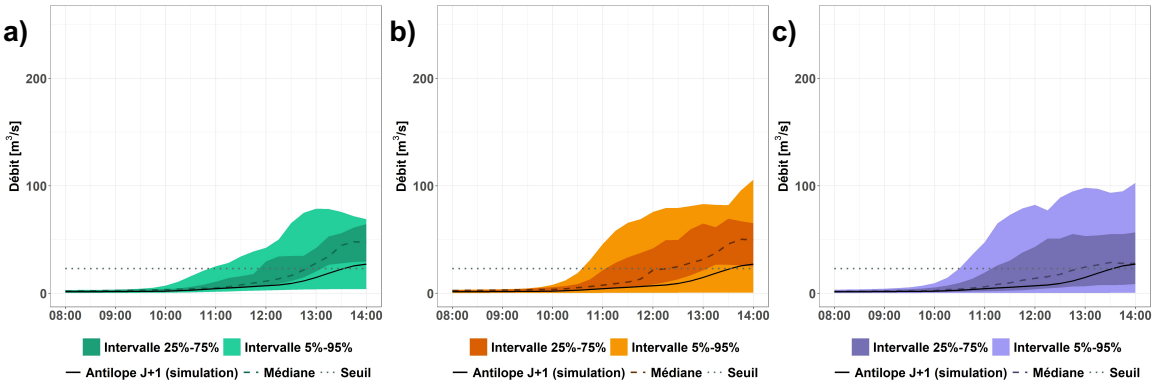

FIGURE 5.21 – *Hydrogrammes obtenus pour l'exutoire n°1 (107 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

L'exutoire n°1 se trouve à la limite de la cellule de pluie intense. Un succès est observé pour les trois ensembles de prévisions des débits.

La première remarque est que ce dépassement du seuil est très bien anticipé, puisque dès la simulation de 8 h, le dépassement de 13 h 30 pour RS est déjà visible. Le dépassement est même prévu prématurément pour le percentile 75%, dès 11 h 45 pour AROME-EPS, (respectivement 11 h pour pepi et pertDpepi). Concernant la différence entre les trois ensembles, plusieurs points peuvent être notés :

- Le débit maximal atteint par le quantile 95% est d'environ 76 m^3/s pour AROME-EPS et de 105 m^3/s pour pepi et pertDpepi. Cela illustre la plus grande dispersion atteinte pour pepi et pertDpepi par l'ajout de nouveaux membres issus d'AROME-PI.
- Même si la dispersion est plus grande avec pepi et pertDpepi, le percentile 50% est nettement plus proche du débit simulé avec Antilope J+1 pour pertDpepi que pour les deux autres ensembles.

La Figure 5.22 présente les résultats de simulation pour la prévision du 9 août à 7 h, pour les trois ensembles et pour l'exutoire n°2.

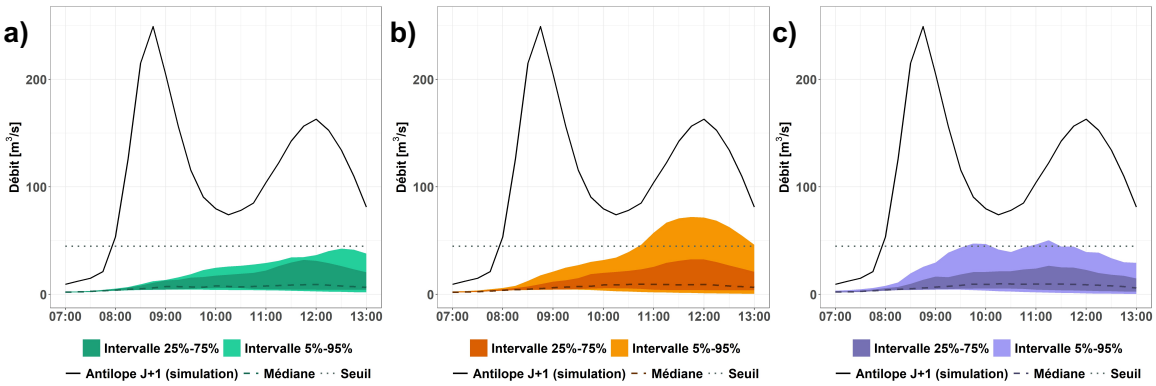

FIGURE 5.22 – *Hydrogrammes obtenus pour l'exutoire n°2 (15 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

L'exutoire n°2 se trouve au cœur de la zone de pluie intense. Aucun dépassement n'est observé pour les trois ensembles, il s'agit d'un échec de la prévision du dépassement de seuil de débit dans le cas de cet exutoire, pour le percentile 75%. Comme pour l'exutoire précédent, une plus forte dispersion des percentiles élevés (95%) est observée pour pepi et pertDpepi, ce qui permet à ces percentiles d'atteindre le seuil de débit. Néanmoins, ce secteur reste un endroit où la prévision de débit est loin de pouvoir anticiper le dépassement de seuil. Cela s'explique par la localisation défaillante des pluies intenses, mais aussi par une sous-estimation des pluies les plus intenses, qui était un peu masquée par l'effet de moyenne sur les hyétogrammes de cumuls spatiaux.

Exutoires avec une faible réaction hydrologique - Exutoires n°3 et n°4

Les exutoires n°3 et n°4 réagissent peu et se situent dans des secteurs limitrophes aux pluies intenses prévues (exutoire n°3) et observées (exutoire n°4).

Les résultats obtenus pour l'exutoire n°3 sont présentés sur la Figure 5.23 pour la simulation de midi le 9 août 2018.

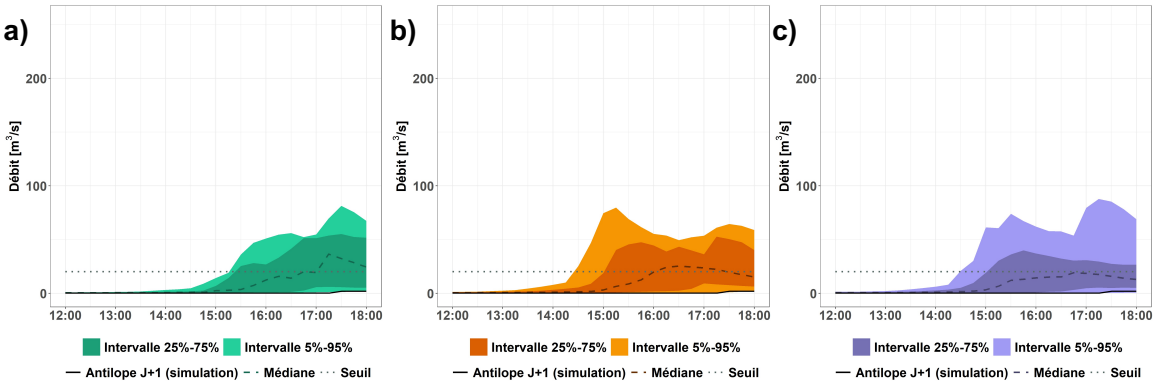

FIGURE 5.23 – *Hydrogrammes obtenus pour l'exutoire n°3 (42 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Cet exutoire se trouve dans un secteur où les pluies prévues sont nettement surestimées par rapport à la pluie de la lame d'eau Antilope J+1. La principale conséquence est une très grande surestimation du débit prévu comme le montre la Figure 5.23. Cela est valable pour les trois produits de prévision pour lesquels une fausse alarme est constatée. Contrairement aux deux exutoires précédents, où la dispersion des plus grands percentiles étaient plus élevée, elle est ici plus faible pour pepi et pour pertDpepi. Le plus grand nombre de membres permet d'atténuer la surestimation des membres intermédiaires pour pertDpepi et le percentile 50% se retrouve sous le seuil de débit, ce qui n'est pas le cas pour les deux autres ensembles.

La Figure 5.24 présente les résultats de simulation pour la prévision du 9 août à 10 h, pour les trois ensembles, pour l'exutoire n°4.

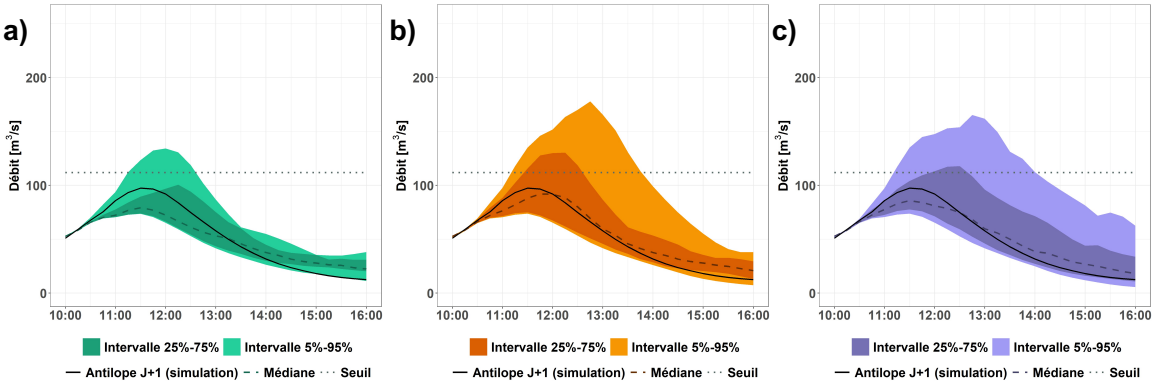

FIGURE 5.24 – *Hydrogrammes obtenus pour l'exutoire n°4 (31 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Pour cet exutoire, les résultats de l'anticipation du dépassement de seuil sont différents selon l'ensemble considéré. Ainsi, pour AROME-EPS est en rejet correct pour le percentile 75%, alors que pepi et pertDpepi sont en fausse alarme. Cet exutoire est limitrophe à la zone de pluie intense observée. En se basant de nouveau sur les cumuls pluviométriques présentés sur la Figure 4.10 dans le Chapitre précédent, ces fausses alarmes retrouvent une explication avec une pluviométrie prévue plus importante pour pepi et pertDpepi que pour AROME-EPS. Le débit maximum est proche du seuil et les pluies prévues sont très satisfaisantes d'après les hydrogrammes générés. Ce cas est une bonne illustration de l'impact des incertitudes qui, même modérées, deviennent déterminantes lorsque les débits observés s'approchent du seuil retenu sans le dépasser.

Le bilan de l'évaluation de l'anticipation pour l'Ardèche reste proche des observations faites pour le cas de l'Aude. La perturbation spatiale introduite par l'ensemble pertDpepi avait initialement pour but de remplacer un pic de pluie par 5 pics décalés les uns les autres sur la surface d'étude. Cela augmente donc la surface couverte, mais réduit globalement la probabilité d'observation du pic. Et c'est exactement ce que montrent les résultats avec une majorité des exutoires où les ensembles pepi et pertDpepi semblent avantager par rapport à AROME-EPS par leur plus grande dispersion induite par les membres ajoutés à partir d'AROME-PI et la perturbation spatiale. Néanmoins, il y a certains exutoires où cette plus grande dispersion est source de fausses alarmes.

5.4 Résultats de l'évaluation sur l'événement du Var

5.4.1 Détection des dépassements de seuils de débits

Cette section présente les résultats de l'évaluation des prévisions des débits menée sur le dernier cas d'étude de l'année 2018, l'événement du Var. Pour cet épisode de crues soudaines, 1 093 sous-bassins ont été modélisés sur le territoire touché.

Cet événement est d'une ampleur modérée par rapport à ce qui avait été observé sur l'Aude ce même mois. La Figure 5.25 illustre les résultats de détection des dépassements de seuils, pour les périodes de retour 2 ans et 5 ans et pour la prévision à pluie nulle RF0.

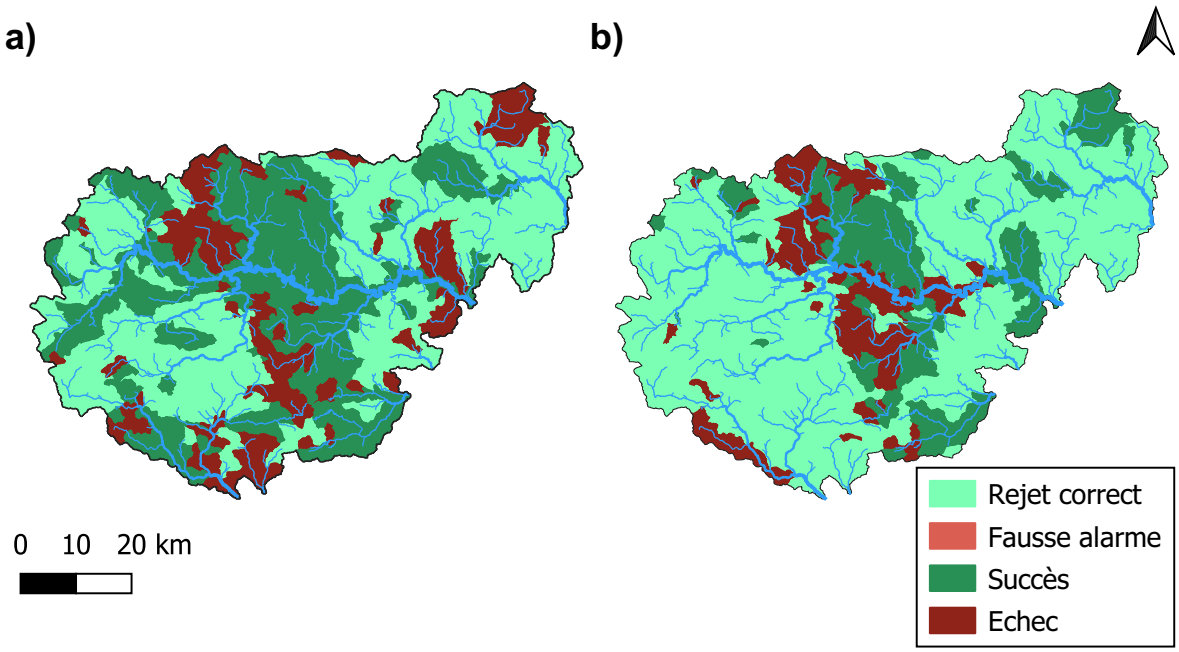

FIGURE 5.25 – *Résultats de détection des dépassements de seuils de débits, pour RF0 et des seuils de a) Période de retour 2 ans et b) Période de retour 5 ans*

Contrairement au cas de l'Ardèche, un grand nombre de succès est observé pour le scénario de pluie nulle RF0, avec 444 sous-bassins pour la période de retour 2 ans et 214 pour la période de retour 5 ans. Néanmoins, il y a plus d'échec avec la période de retour 5 ans que 2 ans. Cela indique que pour un bon nombre d'exutoires ont une période de retour entre 2 ans et 5 ans. Le choix du seuil de débit s'est porté sur la période de retour 5 ans qui permet d'avoir un nombre important de bassins qui dépassent le seuil, sans que cette situation soit majoritaire. Cela semble être un bon compromis pour le nombre de sous-bassins qui doivent être détectés.

5.4.1.1 Courbes ROC obtenues pour les trois produits de prévision d'ensemble

Les courbes ROC présentées sur la Figure 5.26 ont été calculées pour des seuils de débits de période de retour 5 ans sur les 1093 exutoires.

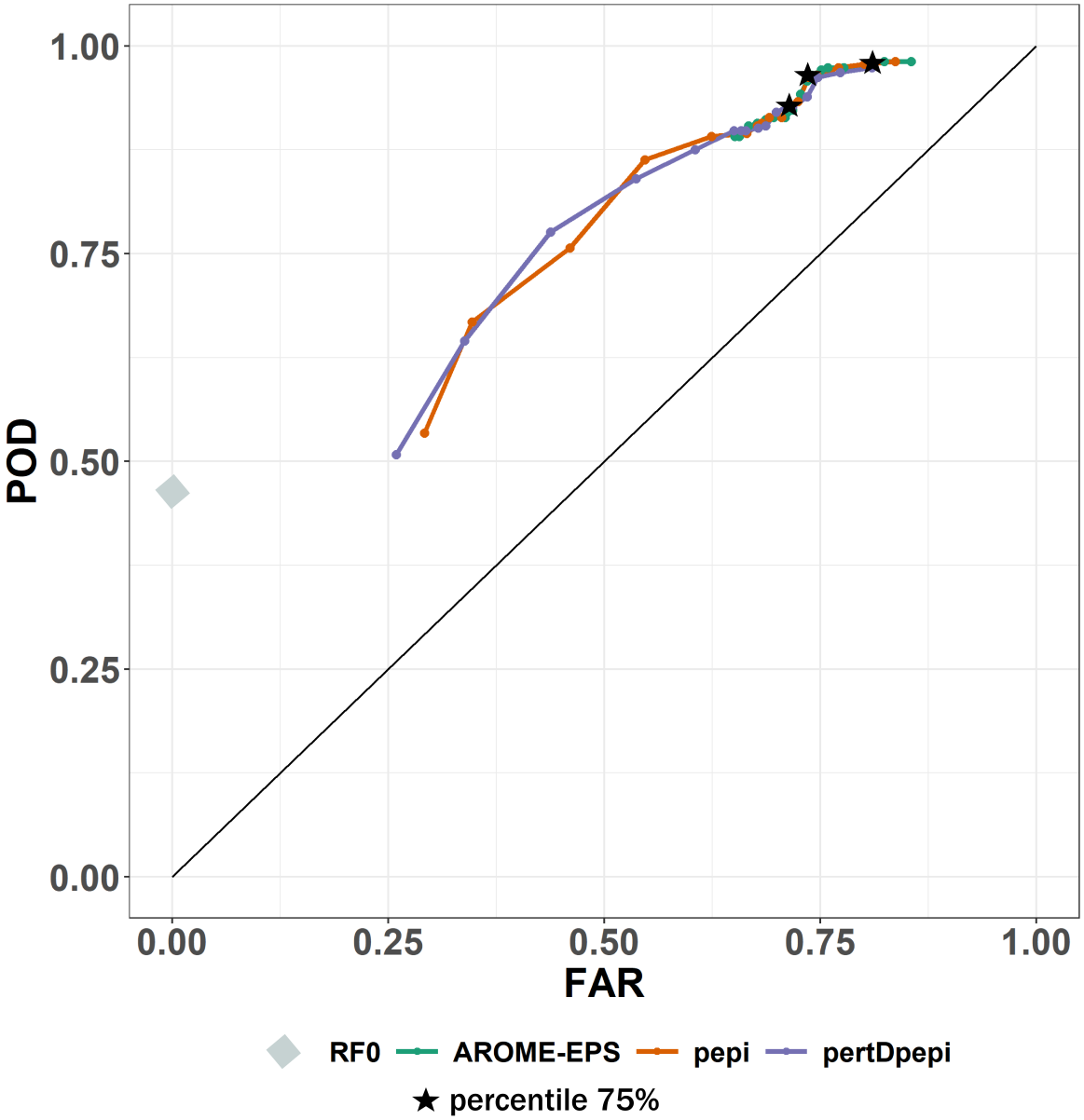

FIGURE 5.26 – *Courbes ROC obtenues pour les trois ensembles de prévision et RF0 pour la période de retour 5 ans. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur pour RF0.*

Cette Figure diffère de celles précédemment obtenues (Figures 5.6 et 5.17) pour les percentiles les plus petits de l'ensemble. En effet, alors que pour l'Aude et l'Ardèche le taux de fausses alarmes restait très faible, il est ici dans le cas du Var relativement important (0.65, 0.29 et 0.26 pour le percentile 10% pour AROME-EPS, pepi et pertDpepi). Cela indique, pour cet événement, une surestimation des débits prévus par rapport à ceux obtenus avec RS, et donc des dépassements de seuils plus fréquents et souvent non pertinents. Cette surestimation avait déjà été observée sur les cartes de cumuls prévus (Figure 4.12) et indique qu'une surestimation modérée pénalise fortement la détection d'un dépassement de seuil, en particulier pour l'ensemble AROME-EPS. En regardant des percentiles plus élevés, le constat est le même pour les trois ensembles, avec des probabilités de détection qui augmentent, mais qui se compensent par un nombre équivalent de fausses alarmes (courbes presque parallèles

à $y = x$). Enfin, pour les percentiles les plus élevés, supérieurs à 75%, le taux de détection s'approche de 1 (plus que dans le cas des deux autres événements), soit presque une détection parfaite des dépassements de seuil.

Concernant le scénario de pluie nulle, RF0, la probabilité de détection est relativement proche de celle obtenue pour l'Ardèche, soit autour 0.5.

5.4.1.2 Analyse cartographique de l'anticipation du dépassement de seuil sur le percentile 75%

La Figure 5.27 illustre la répartition spatiale des succès, échecs, fausses alarmes et rejets corrects pour le percentile 75% avec comme seuil de débit retenu la période de retour 5 ans.

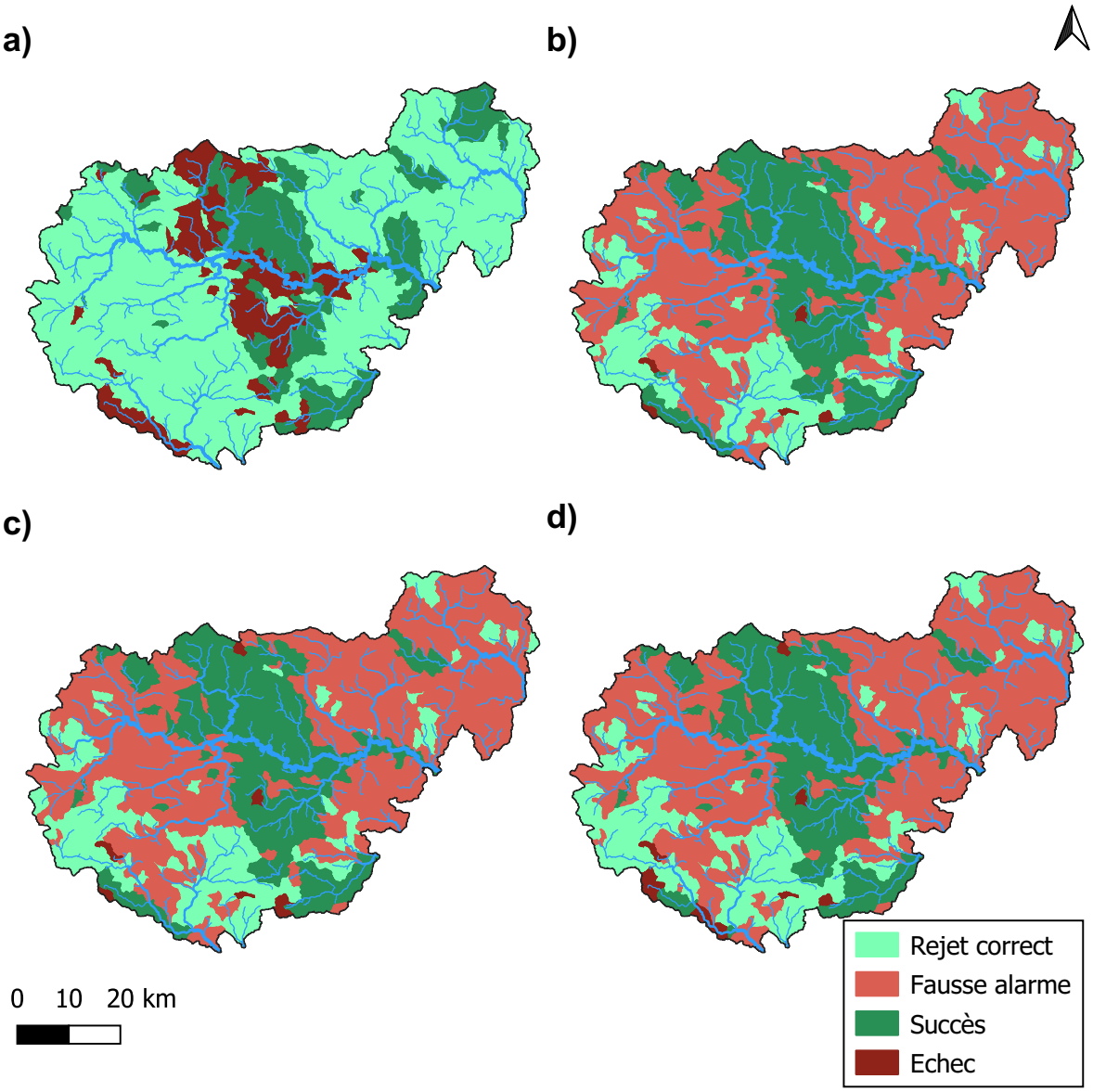

FIGURE 5.27 – *Cartes d'anticipation (0-6 h) du seuil de période de retour 5 ans : a) pour le scénario RF0, b) pour AROME-EPS, c) pour pepi et d) pour pertDpepi. Les succès correspondent à une anticipation supérieure à 15 minutes*

Comme le laissait présager les courbes ROC obtenues (Figure 5.26) le nombre de fausses alarmes est très important, 590 sous-bassins pour AROME-EPS (respectivement 579 et 568 pour pepi et pertDpepi), pour un taux d'échecs quasiment nul, 8 pour AROME-EPS (respectivement 11 pour pepi et pertDpepi). Le nombre de succès est équivalent d'un ensemble à un autre (294, 302 et 305 pour AROME-EPS, pepi et pertDpepi). Ces succès sont principalement localisés dans la zone de pluie intense (Figure 4.12). C'est l'événement où le plus grand nombre de fausses alarmes est observé pour ce percentile.

En comparant avec les résultats de l'évaluation pour RF0 (Figure 5.27 a)), le principal constat est que les trois produits de prévision permettent d'améliorer la détection du dépassement de seuil sur les têtes de bassin, mais globalement, le nombre de succès n'augmente pas autant que pour les deux autres événements. Ainsi, dans la majorité des cas, il y a des échecs pour RF0 et amont et, par propagation amont-aval, un succès pour les exutoires plus en aval. Les prévisions d'ensemble permettent de détecter les dépassements de seuil. Cependant, cet avantage, sans doute lié à la surestimation des précipitations prévues, est également un inconvénient puisque sur les bassins situés aux alentours de la zone de pluie intense, des fausses alarmes sont observées.

Concernant les performances relatives des trois ensembles, il est encore une fois difficile d'établir une différence significative entre les trois produits de prévision. En effet, les courbes ROC sont semblables et spatialement, sur le percentile 75%, peu de différences apparaissent. De plus, contrairement à l'événement de l'Ardèche où la propagation amont-aval n'était pas toujours suffisante pour avoir un dépassement de seuil avec RF0, ici c'est plus souvent le cas.

5.4.2 Délai d'anticipation pour les trois produits de prévision d'ensemble

Même si l'augmentation du nombre de succès obtenus avec les prévisions d'ensemble en comparaison avec ceux obtenus avec RF0 est correcte pour cet événement. La Figure 5.28 présente les histogrammes d'anticipation pour quatre percentiles différents, 20%, 50%, 75% et 95%.

Pour rappel, l'anticipation est toujours calculée comme la différence entre l'heure de dépassement pour la simulation de référence (Antilope J+1), et l'heure de la première prévision pour laquelle un dépassement de seuil a été prévu. Dès le percentile 20% (Figure 5.28 a)), les délais d'anticipation obtenus avec les ensembles sont assez grands, et majoritairement dans la classe 4 h - 5 h, et cette classe se remplit d'autant plus que le percentile considéré est élevé (Figures 5.28 b), c) et d)). Il faut également noter qu'un certain nombre d'exutoires a un délai d'anticipation assez grand (supérieur à 4 h) avec la prévision de pluie nulle. Comme cela a déjà été remarqué dans la section précédente (Section 5.4.1), le nombre de fausses alarmes est important, même pour les faibles percentiles. Le nombre de cas où le dépassement n'est pas anticipé (nombre d'échecs) n'étant pas très élevé, même pour les faibles percentiles, l'apport des percentiles plus élevés n'est pas aussi évident que dans les cas précédents. Enfin, la Figure 5.28 confirme le fait qu'il est très difficile d'établir une hiérarchie entre les trois produits de prévision de pluie utilisés du point de vue des délais d'anticipation offerts.

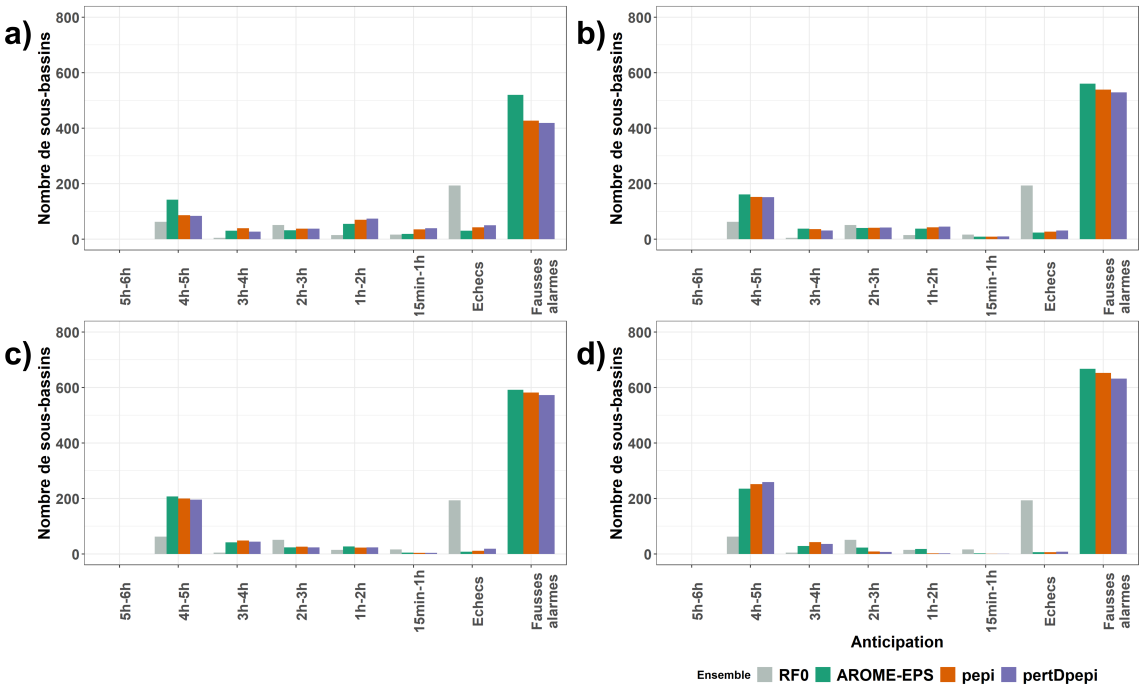

FIGURE 5.28 – *Comparaison des délais d'anticipation du dépassement de seuil de période de retour 5 ans pour les percentiles a) 20%, b) 50%, c) 75% et d) 95%*

5.4.3 Analyse détaillée pour quelques exutoires

Comme pour les deux autres cas d'études, certains exutoires ont été retenus pour étudier plus précisément le comportement des prévisions d'ensemble pour l'anticipation des dépassements de seuils de débits. Les exutoires sélectionnés sont présentés sur la Figure 5.29 ci-dessous.

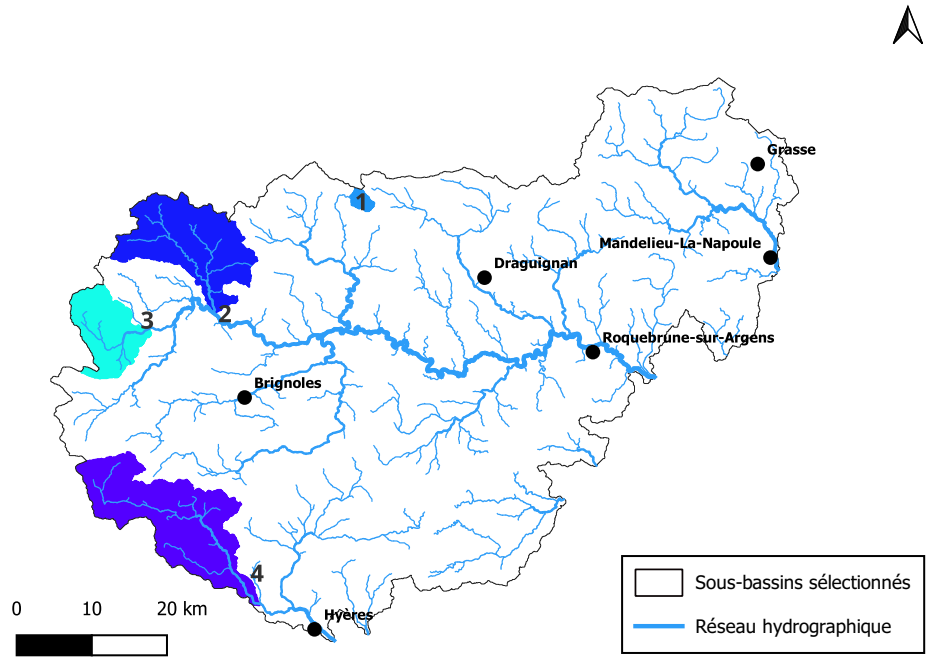

FIGURE 5.29 – *Localisation des quatre exutoires retenus pour une analyse détaillée des prévisions des débits pour l'événement du Var*

Exutoires dans la zone de forts cumuls pluviométriques - Exutoires n°1 et n°2

Les exutoires n°1 et n°2 se trouvent dans la zone de cumul pluviométrique la plus intense, pour laquelle il est possible de s'attendre à de fortes réactions hydrologiques. La Figure 5.30 représente les hydrogrammes obtenus pour la prévision de 9 h du matin le 31 octobre pour les trois produits de prévision d'ensemble pour l'exutoire n°1.

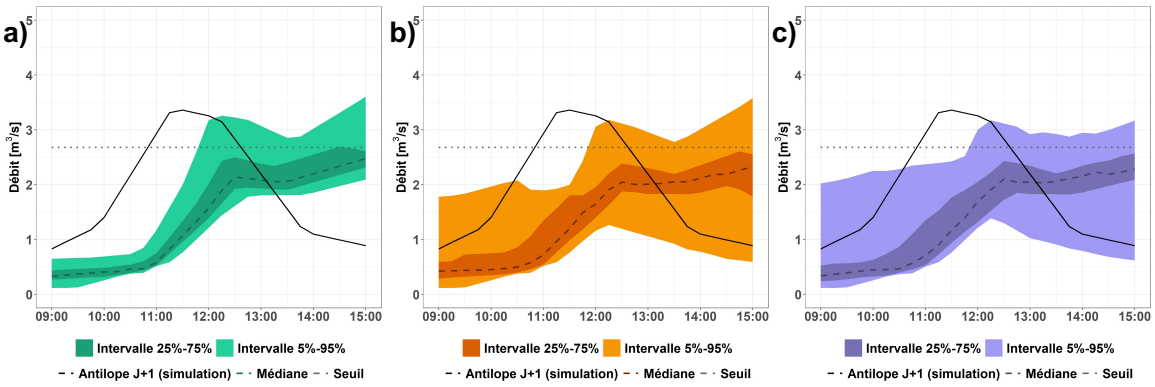

FIGURE 5.30 – *Hydrogrammes obtenus pour l'exutoire n°1 (2.3 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

L'exutoire n°1 se trouve très en amont sur le domaine d'étude, mais dans le secteur de pluie intense. Un succès est observé pour AROME-EPS, mais ce sont des échecs d'anticipation pour pepi et pertDpepi sur cet exutoire. Le premier constat est que le débit seuil est très peu élevé (inférieurs à 5 m^3/s) : c'est un exutoire où le seuil de débit n'est pas forcément représentatif d'une situation potentiellement dangereuse. Le deuxième constat est que même si le seuil de débit n'est pas dépassé pour pepi et pertDpepi pour le percentile 75%, les valeurs de débits obtenues ne sont pas si éloignées de celles d'AROME-EPS.

La Figure 5.31 représente les hydrogrammes obtenus pour la prévision de 7 h du matin le 31 octobre pour les trois produits de prévision d'ensemble pour l'exutoire n°2.

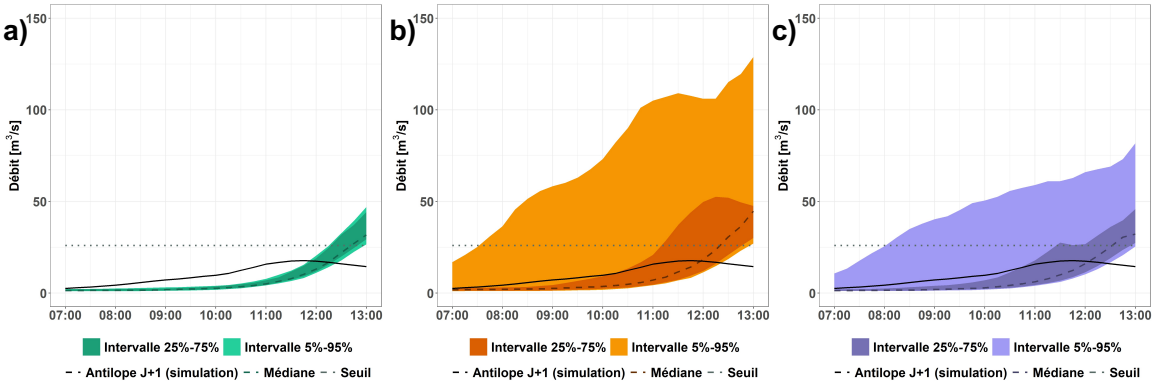

FIGURE 5.31 – *Hydrogrammes obtenus pour l'exutoire n°2 (57 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Cet exutoire se trouve à la limite de la zone pluvieuse intense observée, et en particulier dans la zone de pluies les plus intenses prévues pour le percentile 75% pour l'ensemble pepi

(Figure 4.12 f)). Pour les trois produits de prévision, une fausse alarme est observée pour le percentile 75%. Cela n'est pas une surprise pour l'ensemble pepi qui présentait une grosse erreur de localisation des pluies intenses. Cette erreur de localisation à l'échelle de l'évaluation des pluies n'apparaissait pas aussi nettement, néanmoins, en passant à l'évaluation hydrologique par exutoire, un décalage de quelques kilomètres de la cellule de pluie intense peut avoir des conséquences sur la prévision de débit, comme ici où il y a une forte surestimation des débits. L'ajout de membres à partir d'AROME-PI modifie considérablement la distribution des débits pour l'ensemble pepi par rapport à AROME-EPS. A contrario, pertDpepi est moins dispersé. Cela permet de mettre en avant une des qualités de pertDpepi qui est l'atténuation des maximums de pluie mal localisés, par perturbation spatiale.

Exutoires hors de la zone de pluie intense observée - Exutoires n°3 et n°4

Les exutoires n°3 et n°4 se trouvent dans la zone de pluie moins intense. La Figure 5.32 présente les hydrogrammes obtenus pour la prévision de 11 h le 31 octobre 2018 pour l'exutoire n°3.

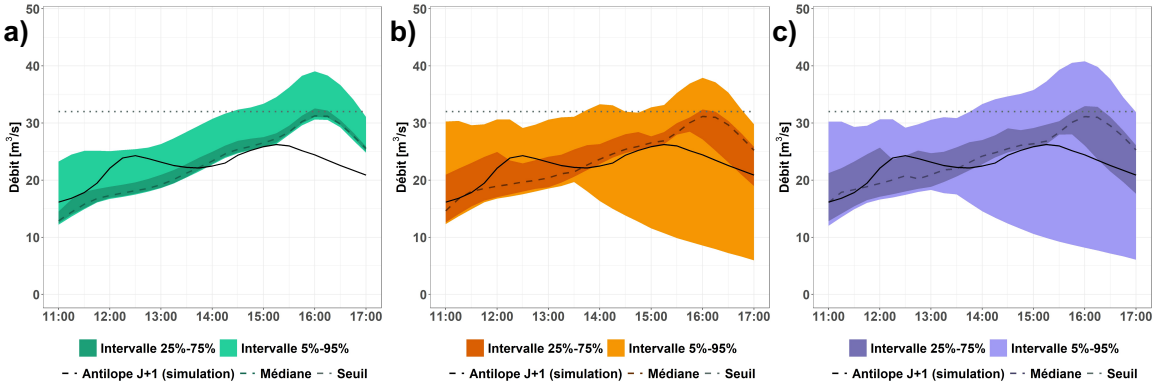

FIGURE 5.32 – *Hydrogrammes obtenus pour l'exutoire n°3 (33 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Ce sous-bassin est dans une zone de précipitations prévues et observées de moindre intensité. Ainsi un rejet correct est observé pour AROME-EPS, mais ce sont des fausses alarmes pour pepi et pertDpepi. C'est encore un cas où la plus grande dispersion des valeurs de débits pour les ensembles pepi et pertDpepi n'est pas un avantage pour détecter le dépassement de seuil. Ce constat peut être nuancé dans le cas d'AROME-EPS puisque même s'il n'y a pas de dépassement de seuil observé en prenant les 6 prévisions précédant le maximum de RS, le seuil est en réalité dépassé pour la prévision de 13 h pour le percentile 75%. Ce cas met en évidence un défaut de la méthode d'évaluation utilisée ici, qui fait les prévisions successives ne sont pas parcourues entièrement.

Enfin, la Figure 5.33 présente les hydrogrammes obtenus pour la prévision de 8 h le 31 octobre 2018.

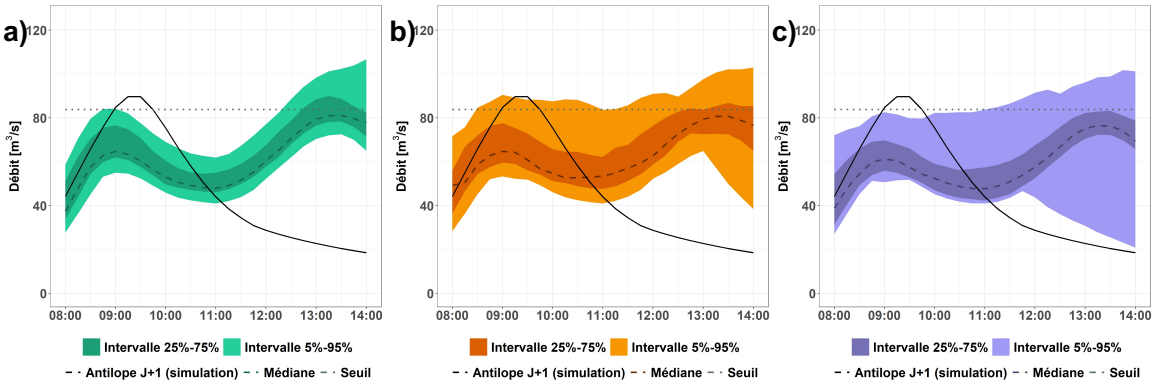

FIGURE 5.33 – *Hydrogrammes obtenus pour l'exutoire n°4 (93 km²) a) AROME-EPS, b) pepi et c) pertDpepi*

Pour cet exutoire n°4, un succès est observé pour AROME-EPS et pepi et un échec pour pertDpepi. La forme globale des trois hydrogrammes est proche, mais un décalage apparait entre les trois prévisions et l'hydrogramme RS. Comme pour l'exutoire n°1, même si ce n'est pas un succès pour le percentile 75% pour l'ensemble pertDpepi, les valeurs de débits obtenues sont proches de celles d'AROME-EPS et pepi. L'utilisation d'un autre percentile aurait probablement conduit à des résultats identiques pour les trois ensembles. En restant dans la logique d'un utilisateur potentiel de ce type de donnees, c'est encore une fois l'utilisation qui en sera faite qui est déterminante et qui dictera la décision de l'utilisateur.

5.5 Conclusions sur l'analyse de l'anticipation des débits par les prévisions d'ensemble

5.5.1 Comment la méthodologie mise en place répond-elle aux besoins de l'évaluation de prévisions d'ensemble adaptées aux crues soudaines ?

Comme cela a été expliqué dans la Section 5.1.2, l'évaluation est basée sur la méthode des courbes ROC, et donc sur des tableaux de contingence. L'utilisation de ces tableaux est recommandée par plusieurs auteurs (Silvestro *et al.* 2011, Anderson *et al.* 2019, Sayama *et al.* 2020). Quelques différences sont à noter entre la méthode traditionnelle des courbes ROC et celle employée ici :

- L'évaluation se focalise uniquement sur la montée de crue, soit la période critique dans un contexte opérationnel (Anderson *et al.* 2019).
- L'analyse réalisée ici se fait en deux temps : d'abord l'analyse spatiale en utilisant les cartes associées de détection des dépassements de seuil, puis une analyse temporelle avec les histogrammes des délais d'anticipation. Cela permet d'avoir une information synthétique, à la fois géographique et temporelle, sur les gains possibles en matière

d'anticipation, par rapport à la simulation de pluie future nulle RF0, mais aussi entre les trois ensembles eux-mêmes.

Cette méthode reste adaptable à chaque événement et à ses caractéristiques en ajustant la période de retour des seuils de débits. Le seuil peut également être ajusté en fonction de l'utilisation souhaitée des ensembles, entre la prévision d'un événement relativement courant (période de retour de 2 ou 5 ans) ou plus rare.

Enfin, l'étape finale de l'évaluation, basée sur les hydrogrammes, s'avère essentielle pour apporter une vue détaillée des ensembles évalués. Cette étape permet notamment de mettre en avant les risques liés à la plus grande dispersion de certains ensembles (en particulier pertDpepi), et à l'importance des décalages temporels, qui sont difficiles à identifier sans passer par cette phase d'analyse des hydrogrammes.

Un autre avantage de la procédure d'évaluation proposée ici, est que la première étape, via les cartes de détection des dépassements de seuils, apporte des informations utiles pour la sélection des exutoires à analyser à partir des hydrogrammes. Ces cartes permettent ainsi d'éviter de sélectionner des exutoires pour lesquels une anticipation importante est possible sans prévision de pluie (correspondant à un succès avec RF0), et elles permettent de sélectionner un panel varié de situations de prévision (succès, échec, fausse alarme, et rejet correct). L'étude plus précise de quelques exutoires judicieusement choisis permet finalement d'apporter des précisions sur les caractéristiques des prévisions d'ensemble et d'expliquer les tendances observées sur les cartes de détection des dépassements de seuil.

Néanmoins, la méthode proposée comporte également quelques limites. Premièrement, la manière dont le délai d'anticipation a été calculé peut être considérée comme optimiste. Comme l'explique Richardson *et al.* 2020, avoir une séquence de plusieurs prévisions cohérentes est une qualité souhaitée. Ici, en parcourant chronologiquement les différentes prévisions disponibles, le délai d'anticipation est calculé en considérant uniquement la première prévision qui dépasse le seuil de débit, sans prendre en considération la cohérence ou non des prévisions successives. D'autres pistes d'évaluation pourraient être explorées en conservant la même méthode d'évaluation tout en ajoutant la prise en considération du nombre de simulations consécutives, anticipant correctement les dépassements (respectivement les non-dépassements) du seuil de débit. Il serait également envisageable d'utiliser des critères spécifiques pour la cohérence des prévisions (Ehret et Zehe 2011, Pappenberger *et al.* 2011).

5.5.2 Qu'en est-il de la performance des trois ensembles de prévision testés ?

5.5.2.1 Observations générales

La contribution de la prévision d'ensemble de pluie à l'anticipation du dépassement d'un seuil de débit a été illustrée clairement pour les événements de l'Aude et de l'Ardèche, par comparaison à une situation sans prévision de pluie (scénario de pluie future nulle RF0). L'apport réside principalement dans la forte réduction du nombre d'échecs sur les cartes de détection des dépassements de seuil (Figures 5.7 et 5.18), et un gain significatif des délais d'anticipation pour presque tous les exutoires (Figures 5.8 et 5.19). Il faut cependant noter que le gain en anticipation est quelque peu contrebalancé par l'apparition d'un nombre important de fausses alarmes. Pour ce qui est de l'événement du Var, l'apport de ces trois produits de prévision d'ensemble est plus nuancé en raison du nombre de fausses alarmes obtenues, nettement plus important que dans les deux autres cas (Figure 5.27). Néanmoins, le gain en temps d'anticipation (Figure 5.28) est bel et bien existant, comme pour les deux autres événements.

Pour pouvoir juger de la plus-value pour les utilisateurs des outils de prévisions d'ensemble des pluies, il faut pouvoir juger des gains et des coûts que représentent pour chaque utilisateur des prévisions, l'augmentation de l'anticipation d'une part, mais aussi de l'augmentation du nombre de fausses alarmes. Cette problématique sera reprise dans les Chapitres 6 et surtout 7, dans lesquels la prévision des impacts plutôt que des débits sera étudiée, et permettra de mesurer ce rapport cout/bénéfice des prévisions.

5.5.2.2 Existe-t-il une différence significative entre les trois produits de prévision d'ensemble au regard de l'évaluation menée ?

Apport d'AROME-PI au sein de l'ensemble pepi

L'ajout de un à six membres supplémentaires (en fonction de l'échéance considérée) basés sur de nouveaux scénarios extraits d'AROME-PI, pour construire l'ensemble pepi à partir d'AROME-EPS, a une influence intéressante sur la qualité de l'ensemble, avec une tendance à la diminution du nombre d'échecs pour un nombre équivalent de fausses alarmes dans les cas de l'Aude et de l'Ardèche. Pour le Var, il est difficile de faire une observation semblable du fait du très faible nombre d'échecs observés, même pour un percentile faible (Figure 5.28). Cette qualité de pepi peut être directement liée à une meilleure capture des périodes de forts cumuls de pluie et de leur extension spatiale (Figures 4.8 et 4.10), qualité qui manque à AROME-EPS sur les cas de l'Aude et de l'Ardèche. Finalement, dans le cas de ces deux événements, pepi apporte une valeur ajoutée pour la caractérisation de la cellule pluvieuse intense, sans dégradation significative de la performance de l'ensemble dans les zones environnantes, où les accumulations de pluie étaient beaucoup plus limitées. Pour l'événement du Var, les apports de pepi n'apparaissent pas sur l'évaluation hydrologique, principalement, car il n'y a pas de grande différence en termes de cumuls et de localisation de la cellule pluvieuse principale. Un point peut néanmoins être souligné : sur la Figure 4.12, une incertitude importante de

localisation des pluies intenses (Figure 4.12 f)), sans conséquence sur l'évaluation des pluies, mais pouvant avoir des conséquences sur les prévisions de débits en cas de mauvais sous-bassin touché, avait été remarquée pour le percentile 75%. Toutefois, les effets sur les prévisions de débits n'ont pas été identifiés, tout le secteur étant en fausse alarme quel que soit le produit de prévision.

L'amélioration globale de la détection des seuils ne se traduit pas pour autant par une évolution significative des délais d'anticipation obtenus avec pepi par rapport à AROME-EPS (Figures 5.8, 5.19 et 5.28).

Influence des perturbations spatiales de pertDpepi

L'ajout de perturbations spatiales dans les quatre directions cardinales aux membres de l'ensemble pepi conduit à augmenter largement le nombre de membres de l'ensemble pertDpepi résultant, avec de 65 à 90 membres en fonction de l'échéance. Globalement, il en résulte un effet de lissage plus important de l'ensemble, particulièrement visible sur les percentiles intermédiaires. Cet effet de lissage des ensembles est plus ou moins un avantage selon les cas regardés. Ainsi, pour l'événement de l'Aude, la période de forte accumulation de pluie est toujours bien capturée (Figure 4.8), mais la zone de pluie maximale n'est pas aussi bien localisée que pour l'ensemble pepi. Pour l'événement de l'Ardèche, le lissage permet au contraire de réduire la zone de pluie intense prévue, trop éloignée de pluie observée (Figure 4.10) et donc de diminuer le nombre de fausses alarmes. En contrepartie, sur cet événement, l'atténuation spatiale des forts cumuls fait augmenter le nombre d'échecs pour l'ensemble pertDpepi (Figure 5.19). Pour le Var, le constat est le même, le nombre de fausses alarmes se réduit légèrement aux dépens du nombre grandissant d'échecs, ce qui pour ce cas est en faveur de pertDpepi.

Les temps d'anticipation semblent être légèrement réduits avec pertDpepi si on le compare aux deux autres ensembles. Là encore, l'étape d'analyse détaillée des hydrogrammes confirme que pertDpepi ne compense pas efficacement la représentation retardée des hydrogrammes. D'un point de vue plus technique, l'ajout de perturbations spatiales conduit à une augmentation significative du nombre de membres, ce qui augmente sensiblement les coûts de calcul des prévisions hydrologiques d'ensemble. Ainsi, le bilan coût/bénéfice de l'ajout d'un très grand nombre de membres supplémentaires dans pertDpepi (jusqu'à 72 membres supplémentaires par rapport à pepi) ne semble pas être forcément pertinent pour les trois événements étudiés.

Synthèse sur l'évaluation des prévisions d'ensemble des débits et sur la méthodologie mise en œuvre

Le cadre d'évaluation mis en œuvre dans ce chapitre a été défini pour évaluer si l'intensité des crues est correctement anticipée avec des simulations d'ensemble des débits, en sachant que les prévisions d'ensemble de pluie ne capturent pas parfaitement l'intensité ni la spatialisation de l'événement.

Après avoir mené une évaluation préliminaire sur les ensembles de pluie, l'objectif était d'évaluer les apports de ces ensembles pour les besoins de la prévision des crues. En ce qui concerne la méthode, le choix d'un cadre spatio-temporel a permis de sélectionner avec précision la zone d'intérêt et de focaliser l'évaluation sur les secteurs concernés. De plus, les cartes de détection des dépassements permettent d'avoir une vue cartographique des performances des différents produits de prévision d'ensemble évalués.

Pour ce qui est de la performance des trois ensembles testés, il y a une nette amélioration des délais d'anticipation par rapport au scénario de pluie nulle mais qui est contrebalancé par l'apparition d'un certain nombre de fausses alarmes (d'autant plus grand que le percentile regardé est élevé). Le nombre de fausses alarmes est d'ailleurs directement lié à l'extension de la zone de pluie intense, qui s'accentue avec les percentiles les plus élevés. En ce qui concerne la distinction entre les trois produits de prévision, pepi est assez intéressant, avec de meilleurs résultats qu'AROME-EPS, tout en conservant un nombre de nouveaux membres assez limité. Pour ce qui est de pertDpepi, le constat est plus nuancé du fait de l'ajout d'un très grand nombre de membres pour une plus-value pas toujours très importante.

Globalement, les prévisions tendent à sous-estimer les pluies intenses et donc à sous-estimer les débits. Un taux de détection élevé de dépassement, et des délais d'anticipation significatifs nécessitent par conséquent de choisir des percentiles élevés, ce qui en contrepartie implique l'augmentation du nombre de fausses alarmes. Il faut également souligner que cette évaluation est très partielle puisqu'elle se concentre sur des événements pluviométriques intenses.

Chapitre 6

Élaboration et évaluation de prévisions d'impacts des inondations

Objectifs du Chapitre :

Le Chapitre 6 s'intéresse à l'élaboration de prévisions d'ensembles d'impacts à partir des prévisions de débits obtenues dans le Chapitre 5 ainsi qu'à leur évaluation. Ces deux étapes sont mises en évidence sur la Figure 6.1.

Dans ce Chapitre, un unique indicateur basé sur le nombre de bâtiments dans l'emprise inondée est considéré comme représentant les impacts. Tout d'abord, une relation univoque débit - nombre de bâtiments dans l'emprise inondée est établie pour chaque tronçon de cours d'eau à partir d'un catalogue d'emprises de périodes de retour 2 ans à 1000 ans. Ces courbes sont une façon d'obtenir, pour une large gamme de débit, les impacts associés et donc de convertir directement les prévisions d'ensemble de débits en prévisions d'ensemble d'impacts.

La méthodologie d'évaluation appliquée dans le Chapitre 5, est reprise dans ce Chapitre pour évaluer les prévisions d'ensemble sous le prisme de la représentation des impacts potentiels. Le principal objectif est d'évaluer si les ensembles sont performants pour localiser de façon anticipée les zones d'impacts, soit les zones avec un certain nombre de bâtiments dans l'emprise inondée, et d'identifier si cette évaluation basée sur les impacts conduit à des conclusions différentes de l'évaluation des prévisions d'ensembles sur les débits effectuée dans le Chapitre 5. Dans ce Chapitre et le suivant (Chapitre 7), les calculs ont été menés uniquement sur l'événement de l'Aude qui est le plus marquant en termes d'impacts observés, par rapport aux deux autres événements de l'année 2018 observés sur l'Ardèche et le Var.

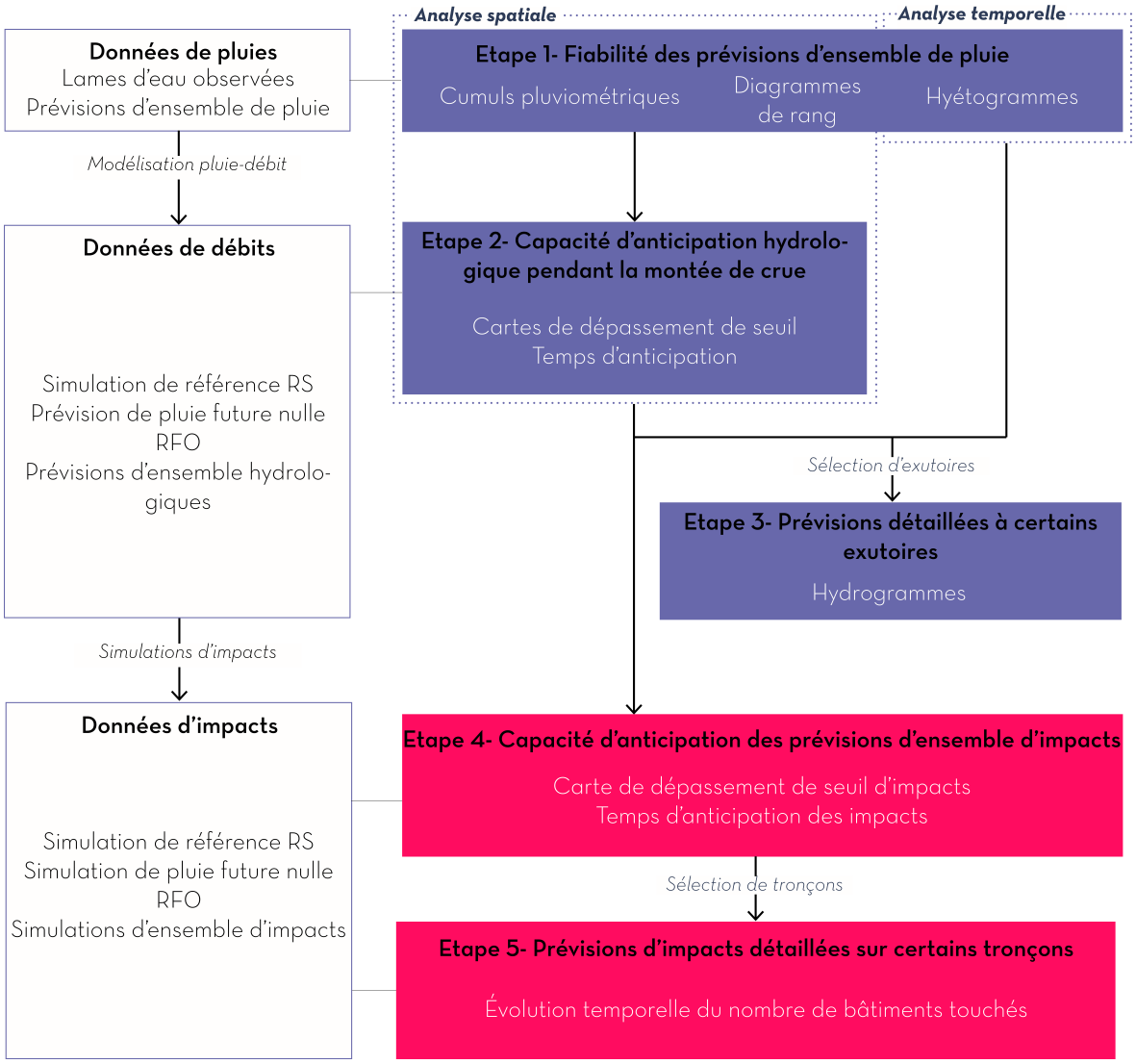

FIGURE 6.1 – *Schéma du principe de la méthodologie mise en place pour l'évaluation hydrologique événementielle des prévisions d'ensemble. En corail, sont représentées les deux étapes d'évaluation menées dans le Chapitre 6.*

6.1 Passage de la prévision des débits à la prévision des impacts

6.1.1 Choix méthodologiques pour la construction des courbes d'impacts

Pourquoi utiliser des courbes d'impacts ?

Passer d'un hydrogramme à une donnée sur les impacts associés au débit prévu demande la représentation de l'emprise inondée associée de la crue à venir. L'approche la plus logique et la plus commune est de réaliser une modélisation hydraulique à partir de l'hydrogramme simulé prévu. Dans le cas d'une prévision d'ensemble couvrant un réseau hydrographique détaillé, un calcul hydraulique doit être fait par tronçon de cours d'eau et par hydrogramme (donc pour tous les membres de la prévision d'ensemble).

L'application de cette approche s'avèrerait extrêmement lourde en temps de calcul. Elle semble donc très peu compatible avec une application dans une chaîne de prévision en temps réel.

L'approche qui a été adoptée dans la thèse est donc beaucoup plus simple et probablement plus pragmatique. Elle s'avère analogue à l'approche proposée par Le Bihan (2016). Elle consiste à construire, sur chaque tronçon de cours d'eau, la relation entre le débit et l'emprise de l'inondation à partir de scénarios d'inondation correspondant à une large gamme de débits. L'utilisation de ces scénarios permet d'évaluer sur chaque tronçon les impacts potentiels pour différentes valeurs de débits (c'est-à-dire le nombre de bâtiments dans l'emprise de l'inondation), et finalement de construire une relation continue entre le débit et les impacts générés, en interpolant entre les différents scénarios d'inondation disponibles (Figure 2.8). Une fois cette relation débit-impacts établie, elle permet de convertir très rapidement les hydrogrammes prévus en évolution temporelle des impacts.

Cette approche est basée sur deux hypothèses simplificatrices essentielles, en comparaison avec l'approche hydraulique demandant de modéliser individuellement l'inondation associée à chaque hydrogramme de la prévision d'ensemble :

- En premier lieu, les scénarios d'inondation pré-calculés sont établis en régime permanent, ce qui peut amener à surestimer les débordements et impose des simplifications dans la représentation des conditions aux limites au niveau des confluences.
- En outre, le nombre de scénarios d'inondation étant limité, l'évolution des impacts entre deux scénarios d'inondation est interpolée linéairement en fonction du débit. Il n'y a donc pas de représentation fine de l'évolution de l'inondation entre les scénarios pré-établis.

Scénarios d'inondations utilisés

Sur le territoire touché par la crue d'octobre 2018 dans l'Aude, un catalogue de scénarios d'emprises d'inondations a déjà été obtenu par Nicolle (2021). Ce catalogue a été présenté dans le Chapitre 3 (Section 3.3.2). Il a été produit à partir de percentiles de débits SHYREG et couvre une large gamme de périodes de retour allant de 2 ans à 1000 ans. Il couvre par

ailleurs l'ensemble des petits cours d'eau du bassin versant de l'Aude, et les zones inondées sont restituées à une résolution de 5 m ce qui permet d'identifier assez facilement le risque d'inondation d'enjeux localisés.

Ce catalogue répond donc très bien aux critères requis pour la construction de courbes débit-impacts sur les cours d'eau touché par la crue d'octobre 2018 dans l'Aude.

Sélection d'une variable représentative des impacts

Les courbes d'impacts calculées par la suite ont été obtenues en dénombrant ces éléments vulnérables présents dans les zones inondées identifiées. Toutefois, derrière l'utilisation du mot impact, peuvent se cacher beaucoup de conséquences possibles des inondations comme les décès, les dégâts sur les infrastructures de transport, sur les bâtiments, ou encore sur les parcelles agricoles (liste non exhaustive).

La BD TOPO produite par l'IGN est une source d'information homogène, qui donne une vision d'ensemble des éléments vulnérables à l'inondation présents dans chaque zone, et qui peuvent constituer des impacts potentiels. Néanmoins, le choix a été fait de ne sélectionner que certains des éléments vulnérables identifiés via la BD TOPO. Cette sélection a été effectuée avec l'objectif de restituer une information la plus pertinente possible, dans un contexte d'utilisation opérationnelle des prévisions d'ensemble par un service de secours. Les impacts possibles ont donc été estimés à partir de la seule couche bâtiments de la BD TOPO, et le choix a été fait de ne conserver que les bâtiments à usage résidentiel et commercial. Ces bâtiments sont, en effet, ceux dans lesquels se concentre la population, et sont donc ceux sur lesquels les opérations de secours et d'évacuation se concentrent en priorité. Ainsi, les sélections suivantes ont été faites sur les différents champs de la couche bâtiments de la BD TOPO :

- Nature du bâtiment : indifférencié, les autres valeurs du champ, religieux, agricole et industriel, ayant été exclus.
- État : en service.
- Léger : non.

La Figure 6.2 illustre les différentes étapes de calcul des impacts potentiels, associés à un scénario d'inondation, sur le secteur de l'Aude à Carcassonne. La Figure 6.2 a) montre les données brutes telles que fournies dans la BD TOPO. La première étape consiste à sélectionner les bâtiments d'intérêt décrits plus haut (Figure 6.2 b)). Les bâtiments sont ensuite remplacés par leurs centroïdes (Figure 6.2 c)). Cela permet de simplifier les calculs en rendant l'intersection emprise inondée - bâtiment plus facile à mettre en place et également d'éviter d'avoir un bâtiment compté deux fois dans deux secteurs adjacents. Enfin, sur la Figure 6.2 d), seuls les centroïdes compris dans l'emprise inondée sont conservés et comptabilisés.

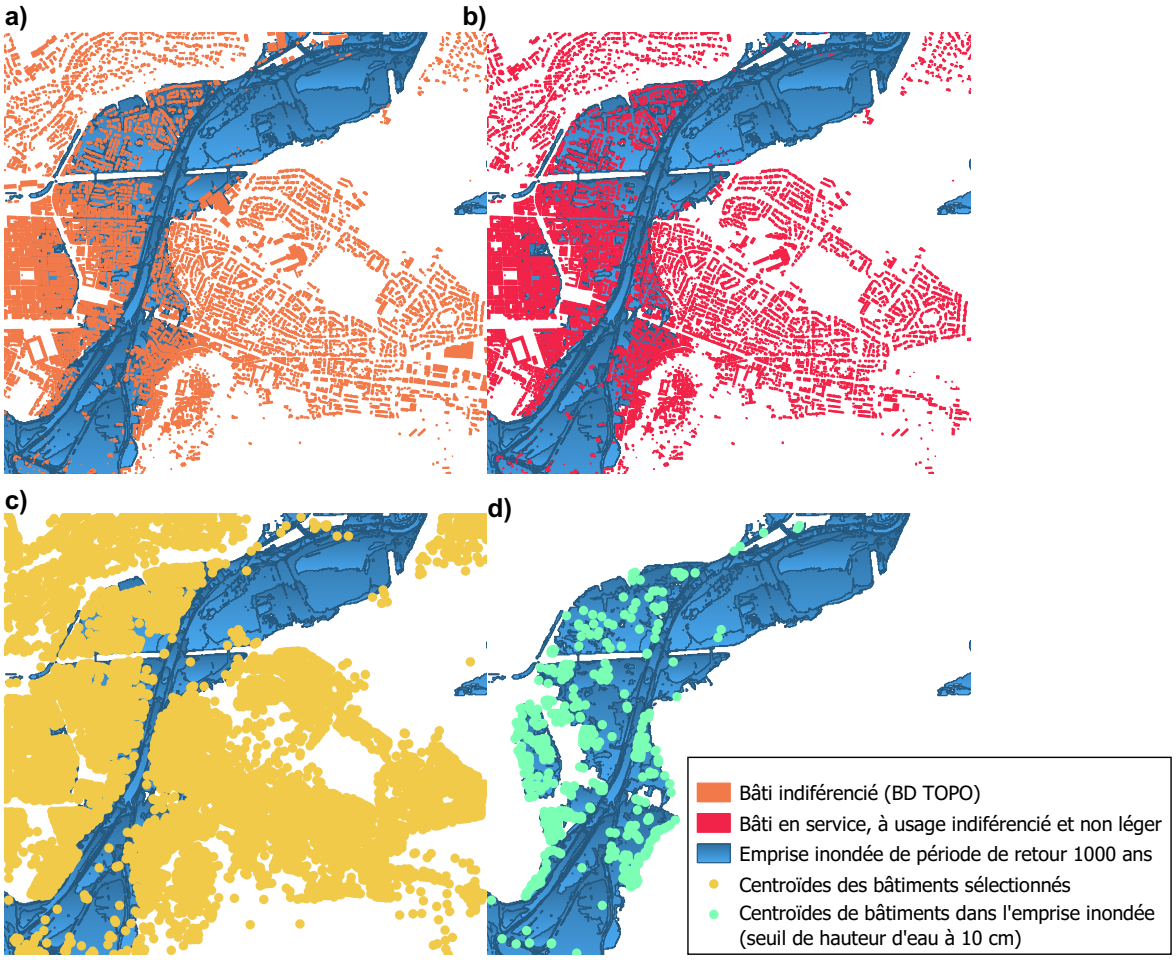

FIGURE 6.2 – *Méthodologie mise en place pour dénombrer les bâtiments situés dans une emprise d'inondation à partir de la couche bâtiments de la BD TOPO*

Il faut noter que pour la sélection des bâtiments, l'emprise inondée prise en compte a été limitée aux hauteurs d'eau supérieures à 10 cm. Cette hauteur est celle à partir de laquelle l'eau commence à entrer dans un nombre significatif de bâtiments, et peut donc générer des dégâts importants et nécessiter l'intervention des secours.

6.1.2 Obtention de courbes débit-impacts continues pour chaque tronçon de cours d'eau

À partir de la procédure de sélection décrite précédemment, il est possible d'associer à chaque emprise inondée de période de retour de 2 ans à 1000 ans un nombre de bâtiments touchés, et ce, pour chaque tronçon de cours d'eau. Les couples de valeurs débit-nombre de bâtiments sont ensuite utilisés pour construire des courbes débits-impacts continues sur chaque tronçon, par interpolation linéaire par morceau entre les valeurs disponibles.

Lors de la construction de ces courbes sur le bassin versant de l'Aude, l'observation du catalogue de scénarios d'inondations utilisé a néanmoins révélé que l'emprise correspondant à la période de retour 2 ans s'avère souvent excessive. De larges champs d'inondation sont en effet observés sur certains biefs de cours d'eau dès la période de retour de 2 ans, ce qui est

surprenant pour une période de retour aussi faible. La Figure 6.3 montre d'ailleurs qu'il n'y a presque aucune différence entre les emprises de périodes de retour 2 et 5 ans.

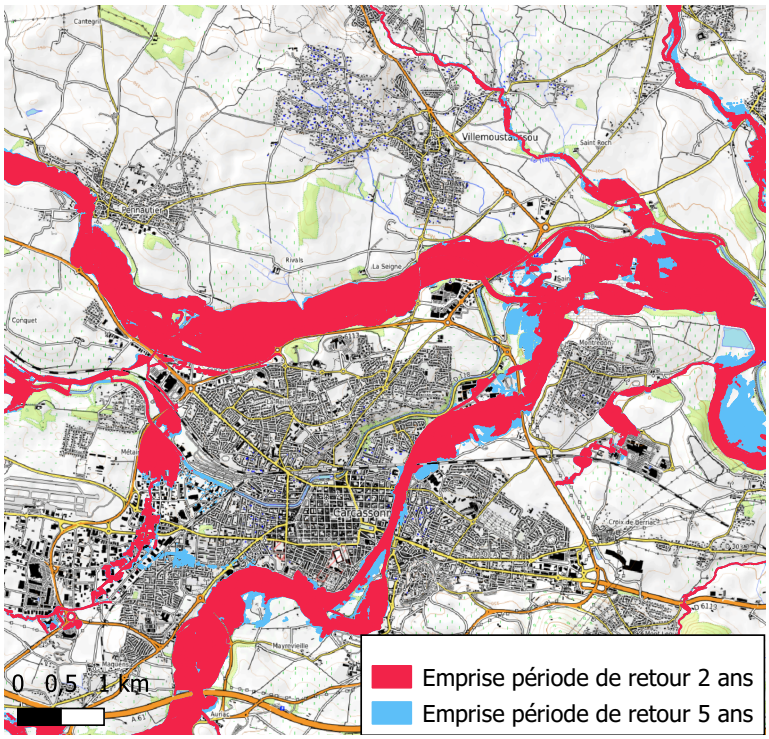

FIGURE 6.3 – *Comparaison des emprises inondées au niveau de Carcassonne pour les périodes de retour 2 ans (en rouge) et 5 ans (en bleu)*

Ces surestimations d'emprises d'inondation pour les faibles périodes de retour s'expliquent par l'absence de représentation du lit mineur de certains cours d'eau dans le modèle numérique de terrain. Le choix a donc été fait, pour le calcul des courbes d'impacts, d'ignorer la période de retour 2 ans, et de faire une interpolation entre le point (0,0) et le point correspondant à la période de retour 5 ans.

La Figure 6.4 ci-dessous illustre différentes courbes d'impacts obtenues pour plusieurs biefs de cours d'eau.

La première observation issue de cette figure est qu'il y a une grande diversité de formes pour les courbes d'impacts. Cette diversité de forme provient à la fois du nombre de bâtiments construits en lit majeur, mais également de la topographie et de la proximité des bâtiments avec le cours d'eau. Pour certains biefs, le nombre de bâtiments touchés croît de manière exponentielle avec la période de retour (Figure 6.4 a)). C'est le cas par exemple autour de Carcassonne, sur des tronçons dont le lit majeur est densément construit. Au contraire, plus en aval de l'Aude (Figure 6.4 b)) la relation débit-impacts est plus linéaire. Pour les tronçons situés très en amont comme celui de la Figure 6.4 c), la courbe atteint un maximum pour les bâtiments inondés à partir d'une certaine période de retour. Enfin, il existe également des tronçons sur lesquels il n'y a pas d'impacts observables jusqu'à la période de retour 1000 ans (Figure 6.4 d)) du fait d'une absence de bâtiments dans le secteur ou bien de bâtiments construits hors de l'emprise potentielle des inondations.

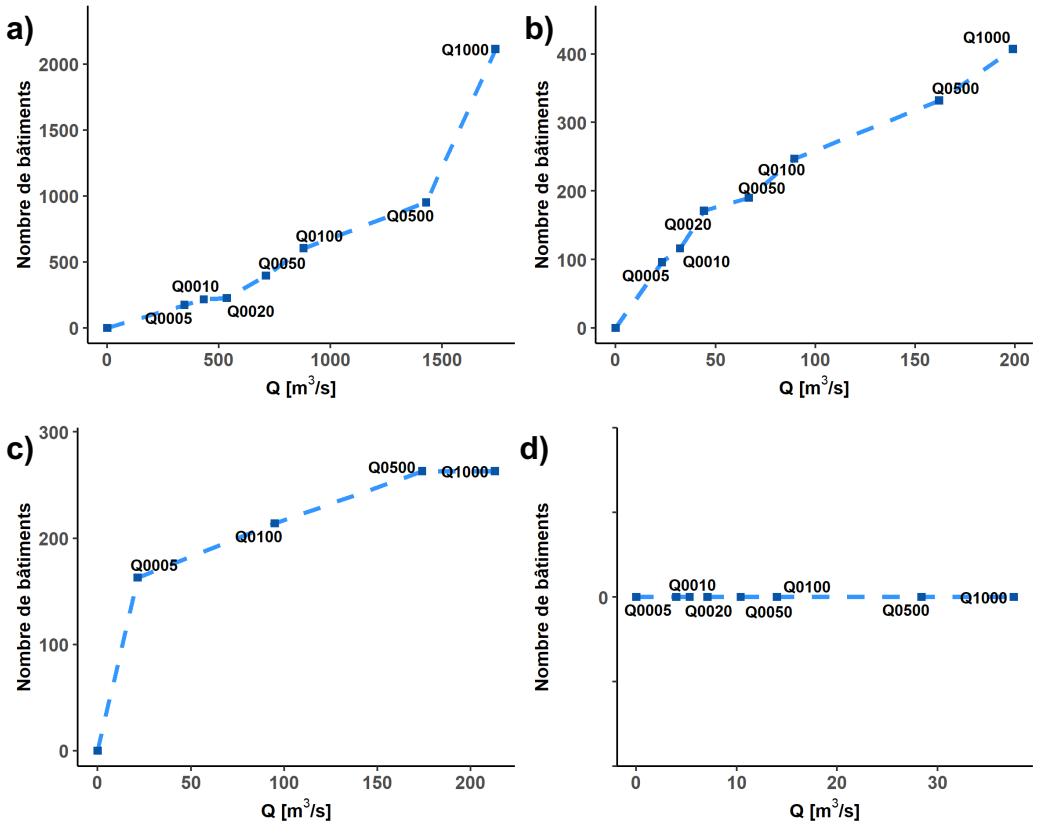

FIGURE 6.4 – *Exemple de quatre courbes d'impacts obtenues pour différents biefs de cours d'eau a) Carcassonne, b) Amont de l'Orbiel, c) Aude aval et d) Amont du Fresquel*

6.2 Évaluation de la chaîne de simulation pluie-débit-impacts obtenue

Comme cela a été précisé en introduction de ce Chapitre, les courbes débit-impacts établies peuvent être directement connectées aux sorties d'un modèle hydrologique, de façon à produire très rapidement une estimation des impacts associés à chaque crue. Cette chaîne de simulation présente néanmoins plusieurs sources d'incertitude qui peuvent s'avérer importantes. Tout d'abord, la pluie observée peut ponctuellement présenter des erreurs. Les débits peuvent aussi être mal estimés par le modèle hydrologique. Enfin, l'estimation des zones inondées sur la base d'un catalogue de scénarios d'inondation préexistant peut conduire à une estimation imparfaite de la zone réellement inondée. Ce paragraphe tente donc de donner une idée de l'incertitude globale existante sur la simulation des impacts de la crue d'octobre 2018 dans l'Aude. Pour cela, les sorties de la chaîne de simulation pluie-débits-impacts, alimentée en entrée par les observations pluviométriques Antilope J+1, sont comparées aux impacts estimés à partir de l'emprise inondée réellement observée.

6.2.1 Emprise inondée de référence pour la crue d'octobre 2018

Il existe une emprise de référence réalisée à partir des observations terrain suite à l'événement de 2018, principalement sur les secteurs à enjeux. La Figure 6.5 représente l'emprise inondée relevée pour cette crue. Il faut noter que cette emprise ne couvre pas l'ensemble du territoire d'étude sélectionné : tous les affluents n'ont pas été cartographiés, en particulier les affluents amont. Seuls les secteurs (d'amont vers l'aval) de Carcassonne, de Puichéric (confluence de l'Argent Double et de l'Aude) et de Narbonne sont bien fournis en termes d'emprises inondées. Néanmoins, l'existence de cette emprise issue d'observations de terrain permet d'évaluer la fiabilité de la simulation d'impacts réalisée à partir des pluies observées (Antilope J+1), en particulier sur les secteurs avec un grand nombre de bâtiments.

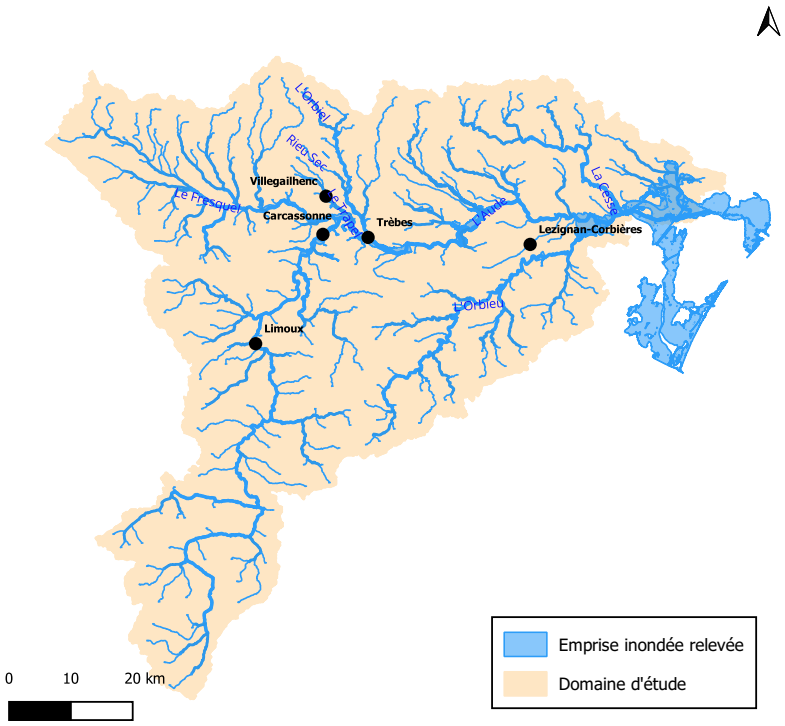

FIGURE 6.5 – *Emprise inondée relevée à la suite des crues d'octobre 2018 dans l'Aude*

6.2.2 Comparaison de la simulation pluie-débit-nombre de bâtiments avec le nombre de bâtiments situés dans l'emprise de référence

La Figure 6.6 compare les estimations du nombre de bâtiments inondés en utilisant a) l'emprise de référence de l'événement et b) la simulation pluie-débit-nombre de bâtiments réalisée à partir des observations pluviométriques Antilope J+1.

Le premier constat qui peut être fait est la différence importante du nombre de biefs concernés par des impacts, qui est nettement moindre dans le cas de l'emprise de référence (Figure 6.6 a)) par rapport à ce qui est obtenu en utilisant la chaine de simulation pluie-débit-impacts complète (Figure 6.6 b)). Ceci est essentiellement dû au fait que l'emprise de référence observée n'est pas disponible sur tous les tronçons du bassin de l'Aude, ce qui suggère que cette emprise inondée observée ne permet pas nécessairement de rendre compte de la totalité des impacts.

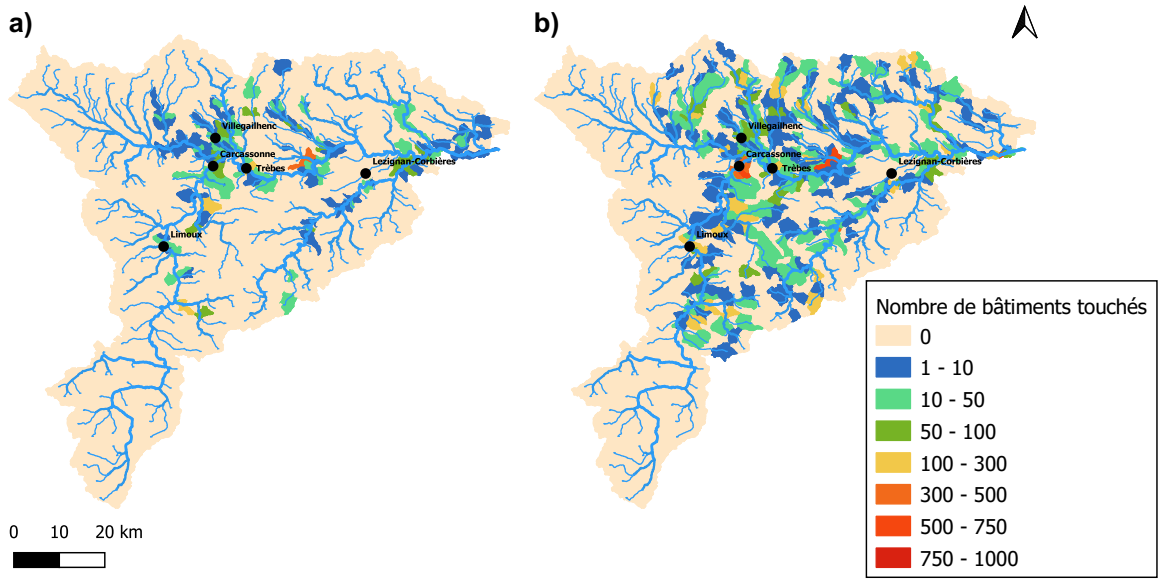

FIGURE 6.6 – *Comparaison du nombre de bâtiments inondés avec a) l'emprise de référence observée de la crue d'octobre 2018 dans l'Aude et b) l'emprise estimée à partir de la chaine de simulation pluie-débit-impacts*

Le deuxième constat, quant à lui, concerne le nombre de bâtiments touchés sur les tronçons où les deux informations (référence observée et simulation) sont disponibles. Là encore, il y a globalement moins de bâtiments touchés à partir de l'emprise de référence qu'avec la chaine de simulation mise en œuvre. Cela est nettement confirmé en comparant le nombre de bâtiments dans l'emprise simulée et la référence observée (Figure 6.7).

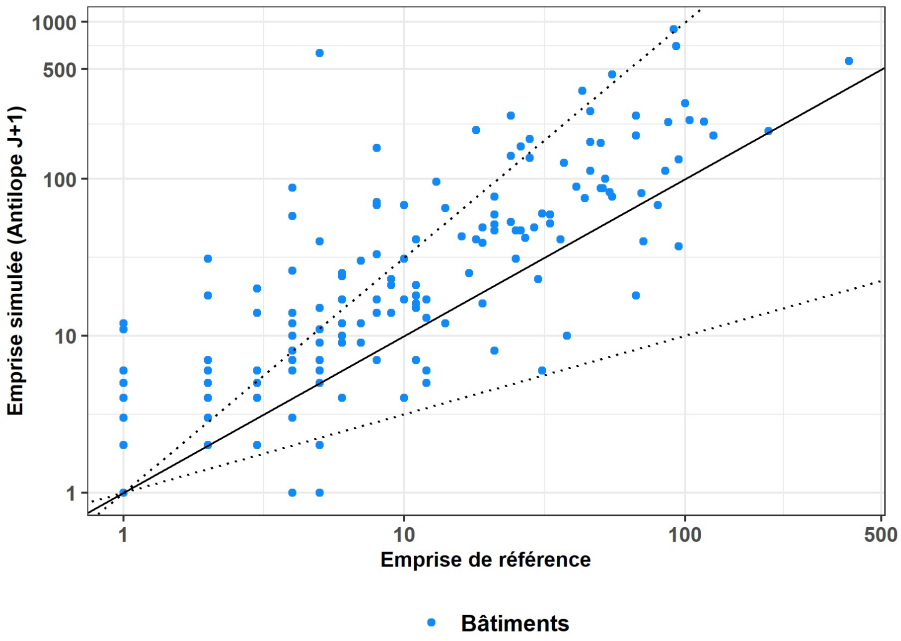

FIGURE 6.7 – *Comparaison du nombre de bâtiments au sein des emprises d'inondation entre l'emprise de référence et l'emprise obtenue à partir de l'interpolation débit/bâtiments*

Comme cela se voyait déjà sur les cartes de la Figure 6.6, le nombre de bâtiments situés dans les emprises de référence est très souvent plus faible que la valeur estimée à partir de la combinaison du modèle pluie-débit et d'une courbe d'impacts préétablie.

La conclusion principale de ces observations est que la méthode de détermination des impacts mise en place à partir des emprises du catalogue d'inondation a tendance à surestimer le nombre de bâtiments impactés par l'événement. Il est probable que les scénarios d'inondation surestiment largement les emprises sur certains secteurs. Tout d'abord, l'ajustement des débits du modèle Cinecar réalisé dans le Chapitre 3 a montré que des erreurs de +/- 20%, sur les débits simulés, sont courantes (Figure 3.9). Il a également été remarqué que pour des bas débits (autour de la période de retour 2 ans), les scénarios d'inondation ont tendance à surestimer l'extension des zones inondées. Enfin, cette surestimation peut également s'expliquer par des erreurs dues à une absence de bathymétrie ou une mauvaise représentation des digues dans le MNT (Hocini 2022).

Néanmoins, de façon à s'affranchir de ces erreurs de simulation parfois importantes, par la suite les impacts simulés par la chaine de modélisation seront pris pour référence pour l'estimation des prévisions d'ensemble des impacts. Cette approche est analogue à celle adoptée dans les Chapitres précédents pour l'évaluation des prévisions d'ensemble des débits. Elle permet de dissocier les erreurs de modélisation des erreurs associées à la prévision d'ensemble des pluies.

6.3 Évaluation des prévisions d'ensembles des impacts sur les bâtiments

Une fois la chaîne de modélisation pluie-débit-impacts mise en place, elle peut être appliquée à partir des ensembles de prévisions de pluies que nous cherchons à évaluer. Ceci va permettre d'obtenir une prévision d'ensemble des impacts sur les bâtiments, sur chaque tronçon de cours d'eau (Figure 5.1). Dans ce paragraphe, ces prévisions d'ensemble des impacts vont être évaluées en reprenant le principe de l'évaluation menée pour les prévisions hydrologiques (Chapitre 5). L'évaluation va notamment être focalisée sur la capacité des produits de prévision d'ensemble à anticiper l'occurrence d'une valeur seuil d'impacts jugée problématique. De façon analogue au Chapitre 5, l'évaluation va se baser sur les courbes ROC, les cartes d'anticipation et les délais d'anticipation, calculés en agrégeant les 6 simulations de prévision précédant le dépassement du seuil ou le maximum d'impacts dans la simulation de référence. Les seuils retenus correspondent cette fois à un nombre fixe de bâtiments dans l'emprise inondée. En effet, contrairement à l'évaluation hydrologique, il n'y a pas de valeurs adaptées à chaque tronçon comme c'était le cas avec les débits SHYREG. Des valeurs seuils de bâtiments ont donc été définies : 10 bâtiments, 20 bâtiments, 50 bâtiments et 100 bâtiments. Les plus petits seuils vont permettre d'intégrer à l'analyse l'ensemble des tronçons de cours d'eau peu densément peuplés, mais générant néanmoins quelques impacts. Au contraire, les seuils de bâtiments plus élevés vont permettre de focaliser l'analyse sur les zones plus densément peuplées. Il est important ici de remarquer que le nombre de tronçons considérés dans l'analyse et moins important que dans le cas de la prévision des débits et

dépend fortement du seuil considéré. En effet, les courbes d'impacts définies montrent qu'un nombre important de tronçons présentent un nombre maximum de bâtiments inondés soit nul, soit inférieur au seuil d'impacts à anticiper. Les tronçons concernés sont logiquement écartés de l'analyse, car ces tronçons présenteront obligatoirement un non-dépassement de seuil. Le Tableau 6.1 présente le nombre de tronçons finalement considérés pour chaque seuil d'impacts, ainsi que le nombre correspondant de tronçons pour lesquels le seuil d'impacts considéré a été dépassé lors de l'événement de 2018 (tronçons pour lesquels des succès sont recherchés avec la prévision).

TABLEAU 6.1 – *Nombre de tronçons considérés pour chaque valeur seuil de bâtiments*

Valeur seuil de bâtiments	**Nombre de tronçons conservés**
10 bâtiments	400 tronçons
20 bâtiments	271 tronçons
50 bâtiments	159 tronçons
100 bâtiments	93 tronçons

6.3.1 Courbes ROC obtenues pour l'anticipation de seuils de bâtiments inondés

La Figure 6.8 présente les courbes ROC obtenues pour les différents seuils de bâtiments inondés considérés.

Concernant le scénario de pluie future nulle RF0, la probabilité de détection a tendance à diminuer lorsque le seuil de bâtiments augmente, exception faite du seuil de 20 bâtiments. Globalement, quel que soit le seuil de bâtiments (de 10 à 100 bâtiments), les courbes ROC gardent la même forme et s'approchent de celles déjà observées pour les débits (Figure 5.6). La valeur ajoutée de pepi par rapport à l'ensemble AROME-EPS se retrouve, en particulier pour les percentiles centraux, entre 20 et 60%. Elle est cependant moins nette pour les percentiles extrêmes. Pour le percentile 75% (étoiles noires sur la Figure 6.8) le taux de fausses alarmes passe de 30 - 35%, pour les seuils de 10 et 20 bâtiments, à 20 - 25% pour les seuils plus élevés. Globalement, les courbes ROC ont tendance à s'améliorer pour les seuils élevés (50 et 100 bâtiments). Les probabilités de détection démarrent à environ 0.60 au lieu de 0.50 pour les percentiles inférieurs à 50%, et les taux de fausses alarmes diminuent pour les percentiles élevés.

La courbe correspondant à l'ensemble pertDpepi est celle qui se détache le plus des deux autres ensembles, cet ensemble apparaissant plus performant sur la base de ce critère, quel que soit le percentile considéré. La principale qualité de cet ensemble est que pour une même probabilité de détection (POD), le taux de fausse alarme (FAR) est plus faible que pour les ensembles pepi et pertDpepi. Cela rejoint ce qui avait déjà été observé pour les courbes ROC élaborées à partir des simulations de débit dans la Section 5.4.1.1.

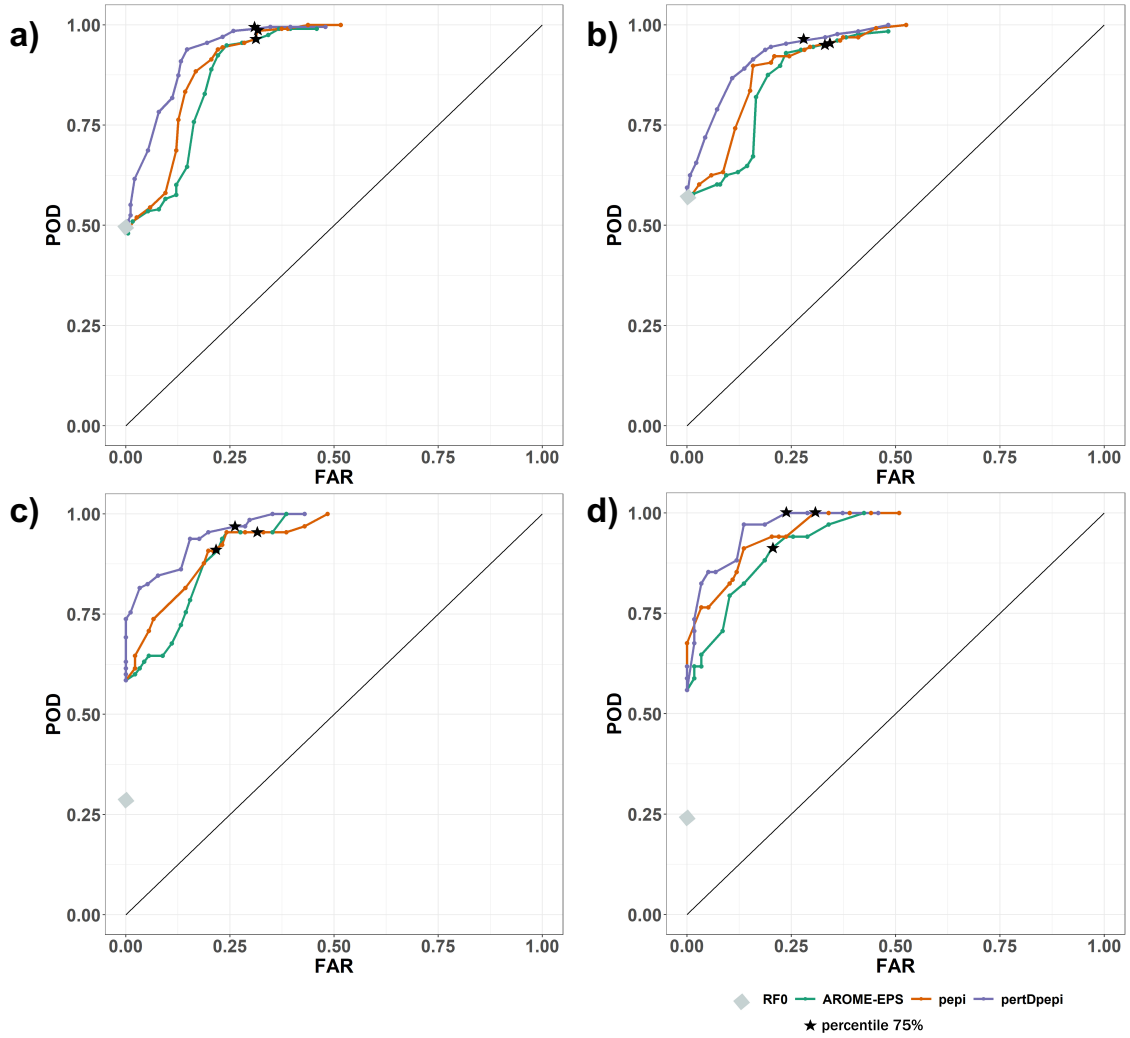

FIGURE 6.8 – *Courbes ROC obtenues pour les trois produits de prévision d'ensemble pour les quatre seuils d'impacts considérés : a) 10, b) 20, c) 50 et d) 100 bâtiments inondés. Les trois étoiles noires indiquent le percentile 75% et le losange gris, la valeur obtenue pour la prévision RF0*

6.3.2 Analyse spatiale de l'anticipation d'un seuil de bâtiments inondés pour le percentile 75%

Les distributions spatiales des succès, échecs, fausses alarmes et rejets corrects pour le percentile 75% sont présentées sur la Figure 6.9 pour un seuil de 10 bâtiments inondés, et sur la Figure 6.10 pour un seuil de 20 bâtiments inondés.

Ces cartes illustrent en premier lieu la réduction du nombre de biefs pour lesquels l'évaluation les prévisions des impacts sur les bâtiments est possibles, en comparaison de l'évaluation des prévisions de débits. Comme nous l'avons indiqué, cela s'explique par le fait que de très nombreux tronçons contiennent peu ou pas de bâtiments dans la zone maximale inondable, et ne peuvent pas atteindre les seuils d'impacts considérés, même pour les périodes de retour des débits généralement élevées observées lors de la crue de 2018 dans l'Aude (tronçons colorés en beige sur les Figures 6.9 et 6.10). Les tronçons présentant de nombreux bâtiments en zone inondable, et donc susceptibles de dépasser les seuils d'impacts considérés, sont surtout situés sur le cours principal de l'Aude.

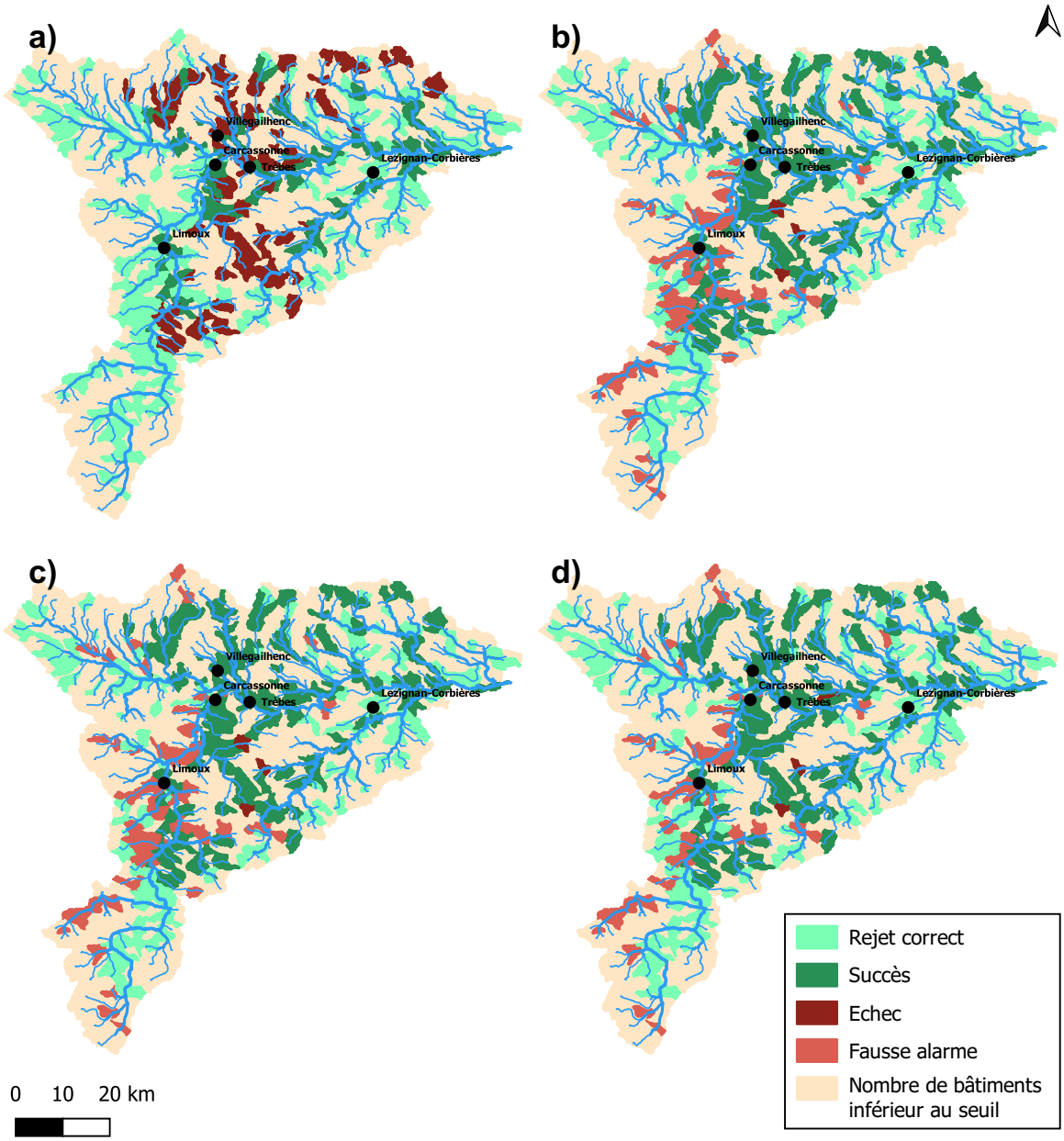

FIGURE 6.9 – *Cartes d'anticipation (0-6 h) du seuil de 10 bâtiments inondés : a) pour le scénario RF0, b) pour le percentile 75% d'AROME-EPS, c) pour le percentile 75% de pepi et d) pour le percentile 75% de pertDpepi. Les succès sont comptabilisés pour une anticipation supérieure à 15 minutes*

La comparaison avec le scénario RF0 permet d'évaluer la valeur ajoutée des prévisions d'ensembles des pluies pour la prévision des impacts. Comme pour les débits, les prévisions d'ensemble des pluies permettent de mieux anticiper de nombreux dépassements des seuils sélectionnés, dans la zone centrale, correspondant au secteur de pluie intense observée (Figure 3.1). En contrepartie, quel que soit le seuil de bâtiments considéré, un nombre important de fausses alarmes apparaît, comme pour l'évaluation des débits, à l'est du secteur d'étude. Comme cela avait déjà été observé, le nombre de fausses alarmes est plus faible pour l'ensemble pertDpepi (49 tronçons pour le seuil de 10 bâtiments inondés et 39 pour le seuil de 20 bâtiments inondés) que pour AROME-EPS (60 et 47) et pepi (60 et 51).

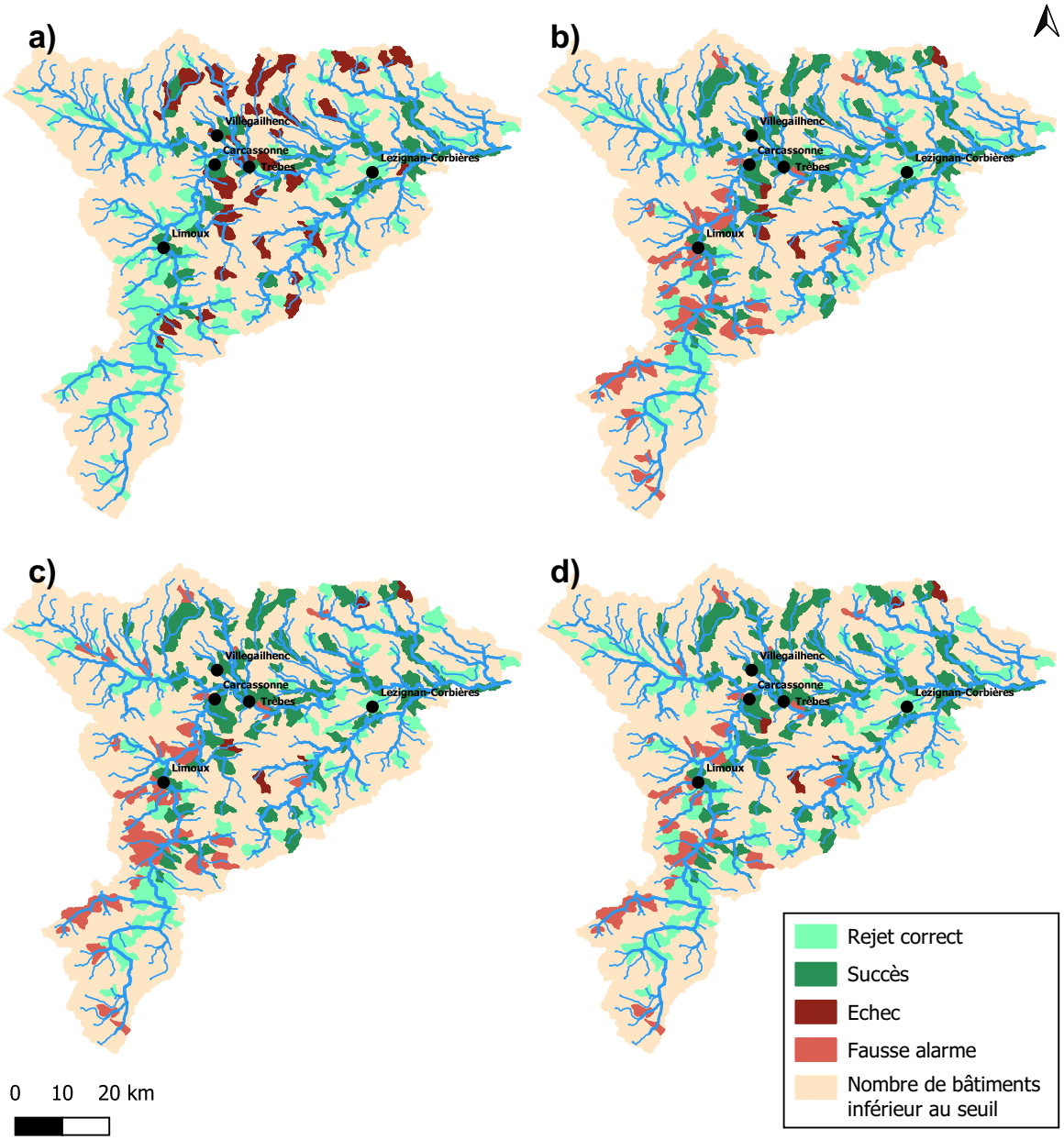

FIGURE 6.10 – *Cartes d'anticipation (0-6 h) du seuil de 20 bâtiments inondés : a) pour le scénario RF0, b) pour le percentile 75% d'AROME-EPS, c) pour le percentile 75% de pepi et d) pour le percentile 75% de pertDpepi. Les succès sont comptabilisés pour une anticipation supérieure à 15 minutes*

L'analyse spatiale de l'anticipation d'un seuil de bâtiments inondés pour le percentile 75% aboutit donc à des conclusions proches de celles obtenues pour l'évaluation de l'anticipation du dépassement d'un seuil de débits.

6.3.3 Délais d'anticipation des seuils de bâtiments inondés à partir des prévisions d'ensembles

La Figure 6.11 illustre les délais d'anticipation obtenus pour les percentiles 75 et 95%.

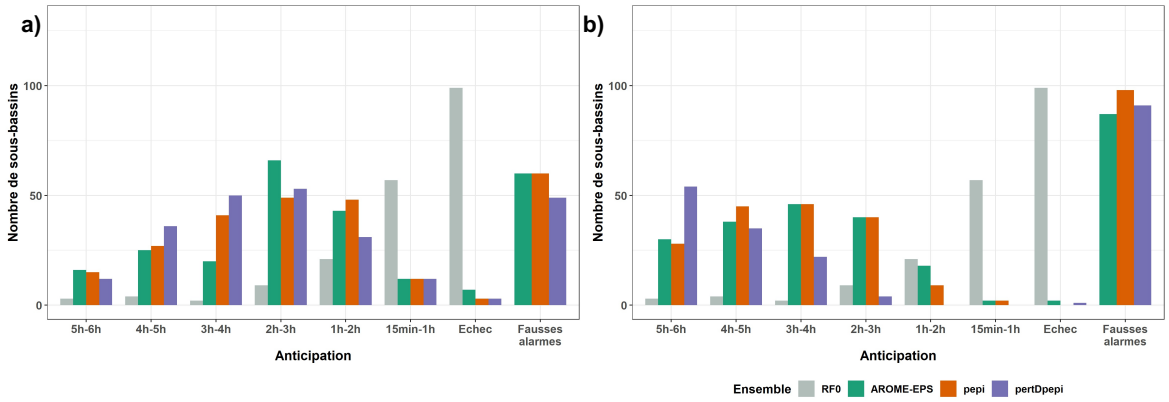

FIGURE 6.11 – *Comparaison des délais d'anticipation du dépassement de seuil de 10 bâtiments pour les percentiles 75% et 95%*

Ici aussi, les conclusions sont similaires à celles obtenues lors de l'étude des délais d'anticipation du dépassement de seuil de débits (Figures 5.8 c) et d)). L'anticipation des dépassements des seuils de bâtiments inondés se situe majoritairement autour de 2-3 h avec les prévisions d'ensemble. Les trois ensembles offrent une anticipation intéressante par comparaison au scénario RF0, qui correspond à une absence de prévision des pluies. Avec ce scénario RF0, l'anticipation ne dépasse une heure que pour une petite moitié des tronçons avec des succès.

6.4 Comparaison des résultats de l'anticipation de seuils de débits et de seuils de bâtiments inondés

Les résultats présentés à la Section 6.3 conduisent au premier abord à des conclusions très similaires à celles obtenues pour l'évaluation des prévisions de débits. Dans ce paragraphe, nous allons essayer de pousser plus loin la comparaison de ces deux évaluations, pour établir les différences éventuelles.

6.4.1 Première comparaison à partir des aires sous les courbes ROC (AUC)

Un premier critère quantitatif synthétique, l'AUC (Area Under the Curve), peut être calculé à partir des courbes ROC. L'AUC correspond à la surface sous chaque courbe ROC. La courbe ROC idéale étant la ligne horizontale passant par le point POD=1 et FAR=0, l'AUC d'un modèle de prévision parfait vaut 1. Les différentes valeurs d'AUC pour chaque courbe ROC sont présentées dans le Tableau 6.2.

Tableau 6.2 – *Aires sous les courbes ROC (AUC) calculées à partir des résultats de prévision d'ensemble des débits et des impacts*

	AUC évaluation des débits (période de retour 10 ans)	**AUC évaluation des impacts (10 bâtiments)**
AROME-EPS	0.82	0.91
pepi	0.86	0.92
pertDpepi	0.91	0.96

Comme cela avait déjà été établi visuellement, la même hiérarchie entre les trois produits de prévision AROME-EPS, pepi et pertDpepi est obtenue, que l'évaluation porte sur l'anticipation des seuils de débits ou des seuils de bâtiments inondés. Les AUC calculées pour l'anticipation de seuils de bâtiments inondés sont néanmoins plus proches de 1 pour tous les produits de prévisions. Cette différence est toujours au moins égale ou supérieure à 5%, ce qui peut être considéré comme un écart significatif entre les deux évaluations. Il apparaît donc dans le cas de la crue de l'Aude de 2018 et à l'aune du critère AUC, que la prévision d'ensemble des impacts est légèrement plus performante que la prévision des débits. Cet écart peut avoir plusieurs explications qui seront détaillées dans le paragraphe qui suit.

6.4.2 Origines possibles des différences de résultats

À première vue, il peut sembler surprenant que les deux évaluations de l'anticipation de seuils de débits et de seuils d'impacts conduisent à des résultats différents. En effet, les prévisions sont basées sur les mêmes prévisions d'ensemble de pluie et les prévisions d'impacts découlent directement des prévisions de débits, via les courbes d'impacts strictement croissantes créées pour chaque tronçon de cours d'eau.

Une première explication à ces différences peut être le fait que l'évaluation à partir de seuils de bâtiments inondés porte sur un nombre de tronçons réduit, par comparaison à l'évaluation menée sur les débits. On peut par exemple penser que les fausses alarmes constatées pour les débits concernaient proportionnellement plus des tronçons de cours d'eau amont, présentant de faibles impacts potentiels, et donc non considérés dans l'évaluation à partir de seuils d'impacts...

Une autre explication tient dans les différences entre les valeurs de seuils considérées. En effet, même si les prévisions de pluie utilisées sont identiques dans les deux cas, il n'y a pas de correspondance systématique entre le seuil en nombre de bâtiments inondés considérés, et les seuils de débits exprimés en période de retour (10 ans). Les différences entre ces deux seuils dépendent du tronçon considéré. Le tri des tronçons servant à évaluer les performances des prévisions (tronçons pour lesquels les seuils sont ou ne sont pas dépassés) ne s'effectue donc pas de la même manière selon que l'on considère les débits et les impacts.

À titre d'illustration, la Figure 6.12 présente les périodes de retour de débits associées à chaque seuil de bâtiments inondé considéré.

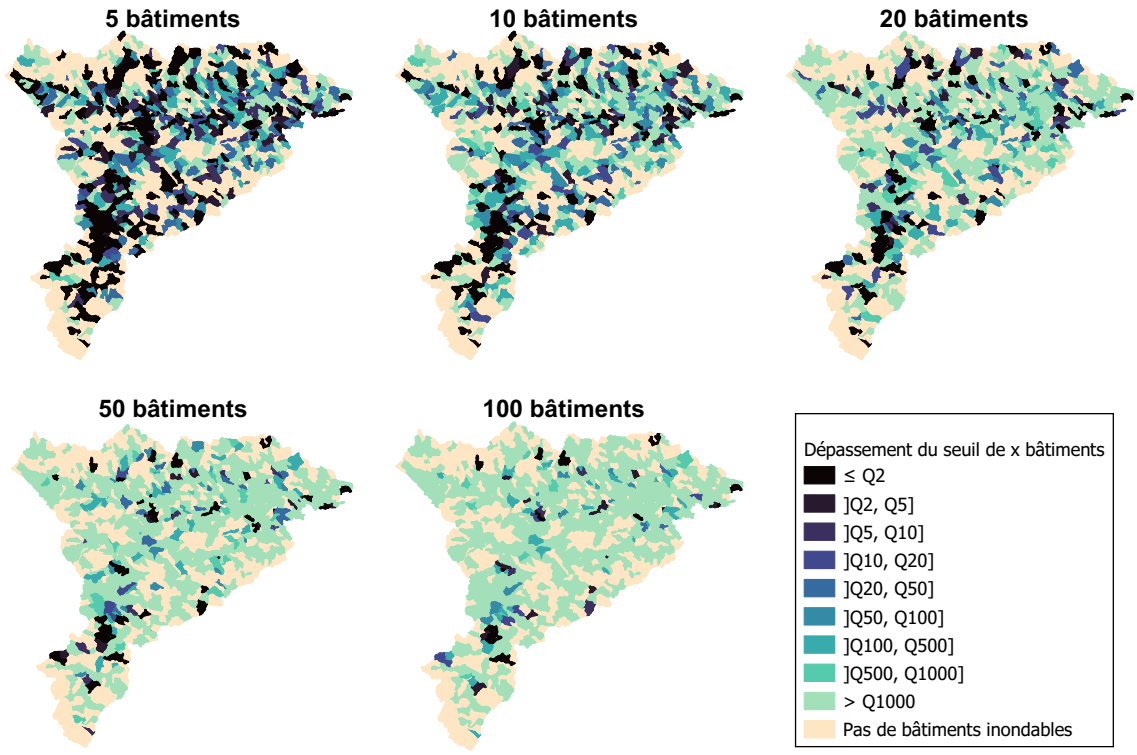

FIGURE 6.12 – *Périodes de retour des débits associées à différents seuils de bâtiments inondés*

Cette Figure illustre la forte diversité des périodes de retour des débits associées à un seuil de bâtiments, sur l'ensemble des tronçons du bassin de l'Aude. Cela est d'autant plus vrai quand le nombre de bâtiments impactés auquel on s'intéresse est peu élevé (5 ou 10 bâtiments). Pour tenter de comparer ce qui est réellement comparable, nous proposons dans le paragraphe qui suit d'examiner pour chaque période de retour de débit les seuls tronçons pour lesquels le seuil de bâtiments inondés correspond à cette période de retour.

6.4.3 Deuxième comparaison avec association des seuils de débits et des seuils d'impacts

Cette nouvelle comparaison vise à s'affranchir autant que possible des deux sources de différences identifiées au paragraphe précédent.

Tout d'abord, la comparaison est conduite après s'être assuré de la proximité entre le seuil de débit et le seuil de bâtiments inondés considérés pour chaque tronçon. Pour cela, les tronçons sont répartis en classes à partir des courbes d'impacts, en recherchant pour chaque tronçon la période de retour conduisant à dépasser pour la première fois le seuil de 10 bâtiments inondés sur la courbe d'impacts. Pour chaque classe ainsi constituée, les prévisions d'impacts sont comparées aux prévisions de débits pour la période de retour correspondante.

Par ailleurs, la comparaison proposée ici est réalisée tronçon par tronçon, ce qui permet de garantir que les tronçons pris en compte sont les mêmes dans les deux évaluations. Le critère utilisé est un décompte des tronçons pour lesquels une réelle différence entre les deux évaluations est constatée :

- Est considérée comme une amélioration de la prévision le passage d'un échec ou d'une fausse alarme pour les débits, à un rejet correct ou un succès pour l'évaluation à partir d'un seuil d'impacts.
- Est considérée comme une dégradation de la prévision le passage d'un succès ou d'un rejet correct pour les débits, à une fausse alarme ou un échec pour l'évaluation à partir d'un seuil d'impacts.
- Les autres cas de figure possibles, comme le passage d'un rejet correct à un succès, sont considérés comme neutres puisqu'ils ne changent pas la qualité de la prévision. Le passage d'une fausse alarme à un échec ou inversement est également considéré comme neutre. Ce choix est plus discutable, car les conséquences sont loin d'être identiques pour un utilisateur des prévisions.

La Figure 6.13 présente les résultats obtenus en appliquant le principe de comparaison exposé ci-dessus, pour le seuil de 10 bâtiments inondés et pour les trois produits de prévision testés, ainsi que pour RF0.

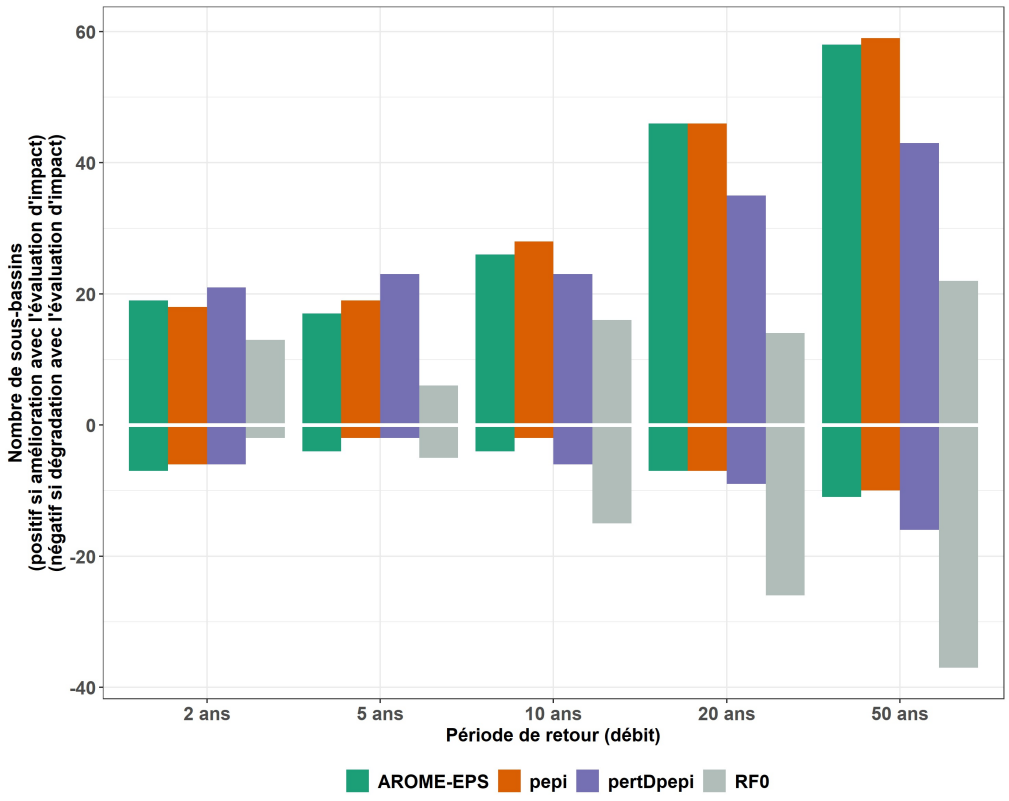

FIGURE 6.13 – *Histogramme de comparaison des deux évaluations sur les débits et les impacts (pour le seuil de 10 bâtiments)*

Le Tableau 6.3 redonne le nombre de tronçons sur lesquels aucune évolution n'apparait entre l'utilisation des débits et des impacts pour toutes les prévisions et les périodes de retour.

Tableau 6.3 – *Nombre de tronçons sans évolution avec l'utilisation des débits ou des impacts*

	AROME-EPS	**pepi**	**pertDpepi**	**RF0**
Période de retour 2 ans	106	108	105	81
Période de retour 5 ans	106	108	105	106
Période de retour 10 ans	106	108	105	115
Période de retour 20 ans	106	108	105	138
Période de retour 50 ans	198	198	208	164

La Figure 6.13, en prenant en compte le nombre de tronçons où aucun changement n'apparait entre les deux évaluations (Tableau 6.3), montre que les résultats obtenus sont bien identiques pour une large majorité de tronçons.

Toutefois, malgré les efforts réalisés pour rapprocher les deux évaluations, des différences visibles subsistent. La détection d'événement est en effet globalement améliorée pour les seuils d'impacts, par rapport aux seuils de débits correspondant, sauf pour la simulation RF0 où un plus grand nombre de tronçons sont affectés par une dégradation des résultats, en particulier pour des périodes de retour supérieures à 20 ans. Il existe aussi un certain nombre de tronçons, pour toutes les prévisions, où les résultats d'évaluation, en utilisant les impacts, se dégradent (moins de 5% des tronçons sont concernés). L'explication à ces différences toujours visibles entre les deux évaluations provient du fait que malgré l'utilisation de seuils proches, ils ne sont jamais strictement identiques. Le découpage en classe conduit à dépasser systématiquement le seuil de 10 bâtiments pour un débit sensiblement inférieur au seuil de débit correspondant à la période de retour attribué à la classe. Les deux cadres d'évaluation restent donc différents, les seuils de bâtiments considérés étant légèrement inférieurs à son équivalent en termes de débit.

Synthèse

L'objectif de ce Chapitre était de proposer une évaluation de prévisions d'ensemble représentant les impacts des crues plutôt que leurs débits, et d'identifier si cette évaluation peut conduire à des conclusions différentes que celle menée au Chapitre 5 sur les débits. L'évaluation a été réalisée uniquement sur l'événement de l'Aude, le plus intense et marquant en termes d'impacts observés par rapport aux deux autres événements de crues soudaines de l'année 2018 sur l'Ardèche et le Var. La même méthodologie que celle définie pour l'évaluation des prévisions d'ensemble de débits a été reproduite sur les impacts prévus. Les prévisions d'ensemble des impacts ont été obtenues en appliquant des courbes débits-impacts créées à partir d'un catalogue d'emprises de référence pour les périodes de retour de 2 à 1000 ans. La première remarque est que le nombre de tronçons sur lesquels l'évaluation est possible - tronçons où des impacts significatifs sont susceptibles de se produire - est nettement plus réduit que le nombre de tronçons initial. En effet, tous les secteurs ne sont pas densément construits, et de ce fait les seuils d'impacts considérés ne peuvent pas être atteints pour un nombre important de biefs. Le travail sur les prévisions d'impacts conduit donc à focaliser l'analyse sur un nombre limité de biefs.

L'évaluation via les courbes ROC, l'analyse spatiale de l'anticipation à partir des impacts et le calcul des délais d'anticipation a montré qu'il était assez difficile de tirer des conclusions sur la comparaison directe des deux évaluations menées. Les résultats obtenus pour les deux évaluations sont globalement assez comparables. La comparaison à partir des courbes ROC suggère que les prévisions d'ensemble évaluées sont sensiblement plus performantes pour l'anticipation de seuils d'impacts que de seuils de débits. Ceci peut s'expliquer par le nombre réduit de tronçons sur lesquels la prévision d'impacts est évaluée. Lorsque des différences de résultats constatées sur un même tronçon, elles sont inévitablement liées aux différences parfois importantes entre le seuil d'impacts considérés (ici 10 bâtiments) et le seuil de débits de période de retour 10 ans considéré au Chapitre 5. Une comparaison plus poussée, en particulier en retenant pour chaque tronçon des périodes de retour de débits plus proches du seuil de bâtiments inondés considéré, a permis de confirmer que pour des seuils considérés proches, les performances obtenues sont logiquement très similaires.

Néanmoins, même si les conclusions obtenues s'avèrent similaires dans le cas présenté ici, l'évaluation des prévisions d'ensemble à partir de seuils d'impacts, reste une approche complémentaire à l'évaluation basée sur les seuls débits, qui permet en particulier de cibler l'évaluation sur les zones à enjeux (zones peuplées potentiellement vulnérables) les plus importantes pour les utilisateurs potentiels des prévisions.

Chapitre 7

Intérêt des prévisions d'ensemble des impacts pour les prises de décision d'un service de secours

Objectifs du chapitre :

Ce Chapitre propose une nouvelle approche d'évaluation des prévisions d'ensemble à partir de la pertinence de décisions de secours prises à partir de ces modélisations d'impacts. Il s'agit de simuler les décisions prises par un service de secours, en particulier des décisions d'engager des moyens de secours sur le terrain, à l'aide d'un modèle mathématique multi-agents. Par rapport aux évaluations conduites dans les deux Chapitres précédents (Chapitres 5 et 6), et qui portaient sur l'anticipation de seuils de débits ou de seuils d'impacts, cette nouvelle approche d'évaluation permettra d'évaluer le poids relatif des gains associés à une meilleure anticipation, et des coûts associés aux fausses alarmes, qui conduiront ici à mobiliser inutilement des moyens de secours au détriment de secteurs réellement impactés. L'évaluation est conduite pour l'événement de l'Aude d'octobre 2018, en simulant les interventions de terrain, qui auraient pu être déclenchées par le SDIS de l'Aude à partir des prévisions d'ensemble. Les prévisions utilisées sont les prévisions d'impacts sur les bâtiments calculées dans le Chapitre 6. Ce Chapitre présentera d'abord la méthode mise en place pour simuler de façon aussi réaliste que possible les décisions et les conditions, notamment les délais, d'intervention du SDIS, puis pour évaluer l'efficacité de ces décisions. Les trois produits de prévision d'ensemble de pluie sont ensuite évalués et comparés à la référence RF0, au regard des résultats de simulation obtenus.

7.1 Étude de cas retenue : la gestion de la crue d'octobre 2018 par le SDIS de l'Aude

7.1.1 Organisation générale du SDIS de l'Aude

Il existe un schéma départemental d'analyse et de couverture des risques réalisé par le SDIS de l'Aude (SDIS11 2019). Ce document décrit en particulier les différents moyens utilisés pour chaque type d'intervention (incendies, inondations, accidents de la circulation), mais également l'ensemble des moyens de secours disponibles sur le département de l'Aude, et leur répartition sur les différents centres de secours. Nous nous sommes inspirés de ce document afin de fixer des paramètres réalistes pour les moyens d'interventions simulés.

Le schéma départemental précise notamment les moyens matériels mobilisables en cas d'inondation. Ainsi, les véhicules d'intervention les plus utilisés pour intervenir en période de crues sont les CCF, camion-citerne feu. Ces véhicules ont l'avantage d'être surélevés et tout-terrain, ce qui leur donne la possibilité d'intervenir largement malgré des conditions climatiques et de circulation difficiles. La Figure 7.1 ci-dessous montre la position des différents centres de secours de l'Aude avec le nombre de CCF pour chacun de ces centres de secours. Ces véhicules sont complétés par des engins plus spécifiques aux inondations comme le sauvetage en eaux vives (5 véhicules), l'assistance inondation (6 véhicules) ou encore les véhicules polyvalents inondation (6 véhicules) ainsi qu'une vingtaine de bateaux légers de secours. Le SDIS peut également faire appel aux hélicoptères de la sécurité civile pour le secours par les airs (22 hélicoptères sur l'arc méditerranéen). Au total, il y a 42 centres de secours comprenant 92 véhicules CCF mobilisables sur l'ensemble du département.

Pour les besoins de la modélisation, chaque véhicule CCF a été assimilé à une équipe d'intervention. L'hypothèse correspondante est que chaque centre de secours a les ressources humaines et matérielles pour mobiliser autant d'équipes de secours que de véhicules CCF possédé. 92 équipes d'intervention ont donc été localisées conformément à la répartition des véhicules CCF présentée sur la Figure 7.1. Cette Figure met en évidence le fait que la répartition des centres de secours possédant des CCF est relativement hétérogène spatialement, ce qui peut influer sur les temps de trajet et donc sur les délais d'intervention.

Le nombre de 92 équipes, même s'il correspond à la réalité des moyens disponibles dans le département de l'Aude, s'avère toutefois particulièrement faible par rapport aux lieux d'intervention potentiels, puisque la modélisation inclut jusqu'à 687 tronçons sur lesquels des interventions pourront être nécessaires. En effet, même si la modélisation hydrologique a été faite sur 1174 exutoires (découpage présenté dans le Chapitre 3), un certain nombre de tronçons ne présente pas de bâtiments compris dans l'emprise d'inondation de période de retour 1000 ans, ces tronçons ne sont donc pas comptabilisés.

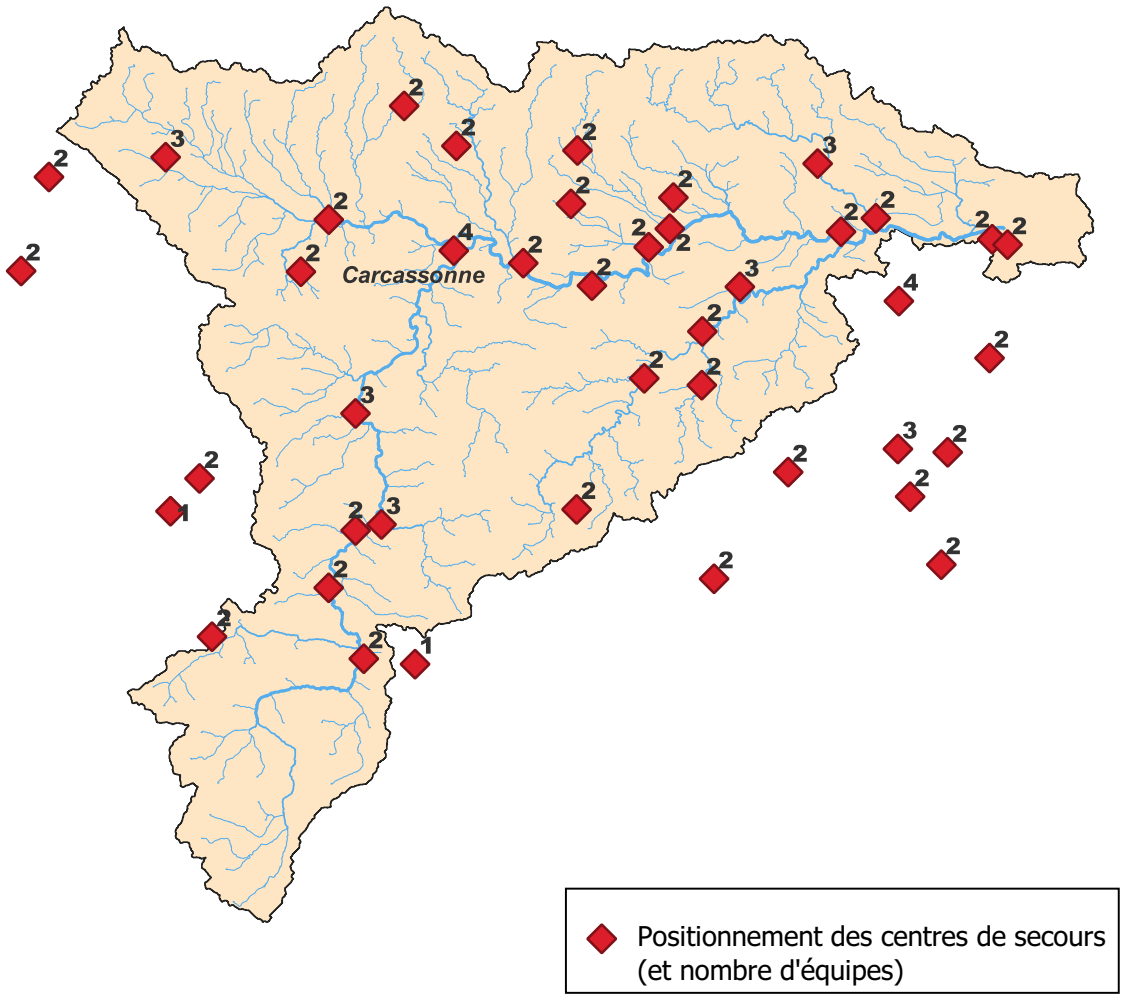

FIGURE 7.1 – *Moyens d'intervention (en nombre de véhicules de type CCF) disponibles sur le département de l'Aude par centre de secours d'après le schéma départemental d'analyse et de couverture des risques du SDIS 11*

Il faut également tenir compte du fait que durant l'événement d'octobre 2018, des équipes de secours provenant des départements voisins ont été mobilisées, ainsi que des hélicoptères. L'hypothèse de 92 équipes disponibles s'avère donc assez réductrice par comparaison à la réalité des moyens qui ont pu être mobilisés lors de la crue de 2018. Pour cette raison, une deuxième hypothèse de 184 équipes mobilisables a été retenue, en multipliant par deux les équipes disponibles dans chaque centre de secours. Ceci permet de disposer d'un scénario plus conforme à la réalité de moyens engagés en octobre 2018.

Enfin, de façon à faciliter l'analyse des résultats de modélisation, une dernière hypothèse a été retenue, dans laquelle le nombre d'équipes disponible est illimité. Ce scénario, même s'il s'avère totalement irréaliste, permet de s'affranchir de la contrainte liée à la disponibilité des moyens d'intervention.

7.1.2 Chronologie des interventions réalisées par le SDIS lors de l'événement

La Figure 7.2 présente les différents secteurs d'intervention du SDIS 11 (et de l'aide des départements extérieurs) au cours de l'événement d'octobre 2018. Une trentaine de secteurs d'interventions peut être identifiée sur le territoire de modélisation (32), ces interventions étant symbolisées par les cercles noirs de 2 km de diamètre de manière à tenir compte non seulement de l'imprécision de la localisation donnée (uniquement le nom de la ville), mais également d'une certaine mobilité de l'équipe d'intervention du SDIS.

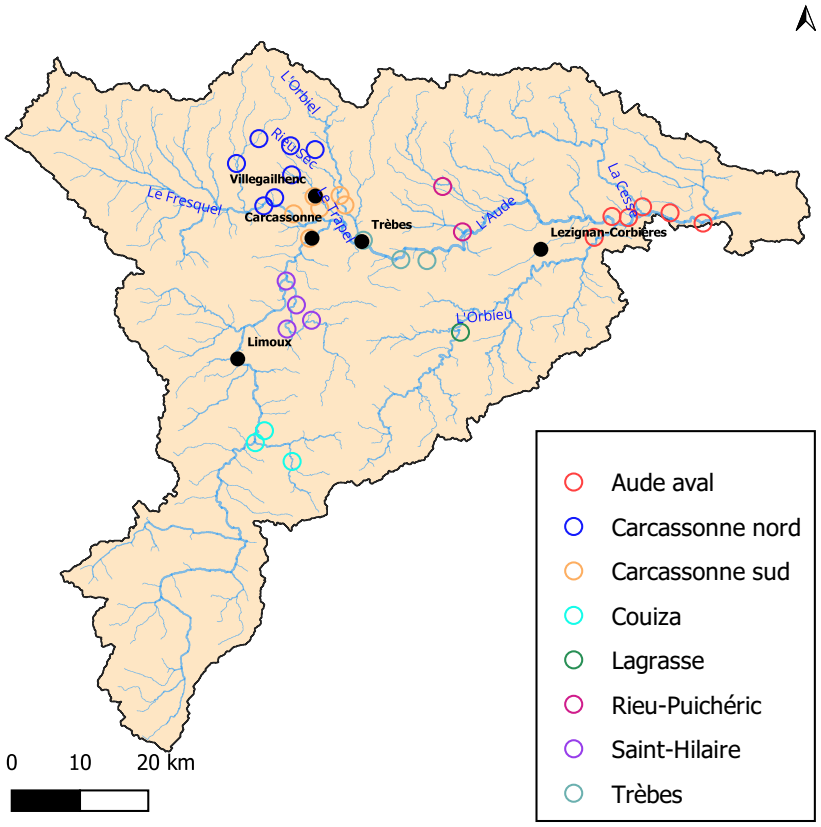

FIGURE 7.2 – *Localisation des opérations de secours menées par le SDIS 11 et les renforts extra-départementaux lors de la crue d'octobre 2018. Les interventions sont représentées par les cercles noirs.*

Pour pouvoir traiter ces interventions et les comparer plus facilement aux tronçons de modélisation, le choix a été fait de les regrouper en secteurs :

- **Carcassonne Nord** : Montolieu, Brousses, Villardonnel, Salsigne, Aragon, Pezens et Ventenac.
- **Carcassonne Sud** : Pennautier, Villegailhenc, Villemoustaussou, Conques, Viltalier et Carcassonne.
- **Saint Hilaire** : Couffoulens, Verzeilles, Ladern et Saint-Hilaire.
- **Couiza** : Lucs, Rennes Les Bains et Couiza.

- **Lagrasse**
- **Trèbes** : Trèbes, Barbeira et Capendu.
- **Rieu-Puichéric**
- **Aude Aval** : Villedaigne, Saint-Nazaire, Saint-Marcel, Sallèles d'Aude, Cuxac et Coursan.

Les premières inondations par ruissellement ont été observées dans le centre de Carcassonne dès minuit le 15 octobre. Le département est passé en vigilance orange inondations à 00 h 30. À partir d'une heure du matin, de premiers appels ont été recensés concernant le nord-Carcassonnais. Les appels se multiplient auprès des centres de secours aux alentours de Carcassonne, Villemoustier, Aragon, Villegailhenc et Barbeira entre 1 h 30 et 3 h 00 du matin avec des personnes évacuées à partir d'une heure. À 3 h, 6 secteurs sont concernés (18 CCF et 6 SAV - sauvetage en eau vive) : nord-Carcassonnais (Aude), le Trapel, l'Orbiel, le Fresquel, la Dure et le secteur de Trèbes (Aude). Entre 3 h et 4 h, le nombre de secteurs touchés évoluent peu, avec seulement un secteur d'intervention supplémentaire à Lagrasse. Néanmoins, la situation devient plus tendue avec des hauteurs d'eau qui augmentent, des appels supplémentaires et une pluie qui ne s'arrête pas, empêchant même les hélicoptères de la sécurité civile de venir en renforts. À Villegailhenc, l'accès aux maisons se révèle impossible du fait d'un courant trop fort. Peu avant 5 h, les premiers appels pour le Lauquet et Couiza arrivent, signalant des mobil-homes et des voitures pris par les eaux. À 5 h, dix secteurs étaient concernés, engageant 29 CCF et 7 SAV. A 6 h, le département était passé en vigilance rouge précipitations suivi à 6 h 15 de la vigilance rouge inondations du SCHAPI. 34 CCF et 7 SAV sont engagés sur les dix mêmes secteurs qu'à 5 h (Carcassone, Trapel, Orbiel, Fresquel, Dure, Trèbes, Orbieu et Lauquet). Peu avant 7 h, les moyens d'intervention commencent à être saturés et se retrouvent dans l'impossibilité de fermer des routes et d'intervenir sur le Lauquet. À 7 h, des renforts extérieurs au département commencent à affluer avec 6 hélicoptères. À 8 h, 46 CCF et 7 SAV sont positionnés sur douze secteurs (nouvelles interventions en cours à Narbonne-Plage). Les renforts extra-départementaux permettent d'augmenter la capacité d'intervention en arrivant sur l'Aude moyenne et l'Orbieu aval. Les premiers hélitreuillages ont lieu vers 9 h à Villegailhenc. À 11 h, l'Aude aval est également inondée avec des interventions à Narbonne et Cuxac. Les moyens sont de 7 SAV, 48 CCF et au niveau extra-départemental, 11 équipes d'interventions supplémentaires viennent en aide, ainsi que 6 embarcations militaires, les 6 hélicoptères sont toujours mobilisés. C'est le pic d'interventions réalisées simultanément. Après 15 h, la situation reste complexe, mais se détend peu à peu avec une diminution du nombre de secteurs d'intervention, l'amont du département étant désormais en décrue.

Au final, près de 211 évacuations, 91 sauvetages, 216 mises en sécurité ainsi que 141 hélitreuillages auront été réalisés le 15 octobre.

L'objectif de ce Chapitre est de montrer si un déclenchement des interventions à partir des simulations d'impacts issues des prévisions d'ensemble, plutôt que des remontées d'informations terrain, aurait pu rendre la gestion de l'événement plus efficace.

7.2 Description et mise en œuvre de la modélisation des interventions du SDIS

7.2.1 Principe général et modalités de déclenchement des interventions

La modélisation mise en œuvre a pour objectif de rejouer des décisions d'intervention lors de la crue d'octobre 2018, et de mesurer l'efficacité de la prise de décisions. Les hypothèses de modélisation sont bien entendues simplificatrices. Il s'agit ici surtout de proposer une première approche permettant de coupler les prévisions et la prise de décisions, afin d'explorer plus avant la plus-value, mais aussi les limites actuelles des prévisions hydrométéorologiques de courte échéance. Dans le modèle proposé, les décisions d'intervention ont été basées sur le nombre estimé de bâtiments inondés pour chaque tronçon de cours d'eau, nombre obtenu à partir des simulations ou prévisions présentés dans le Chapitre 6. Ce principe de modélisation s'avère relativement proche de la réalité, puisque dans la pratique les interventions des SDIS sont très souvent déclenchées à partir de remontés terrains, informations sur l'importance des impacts sur les zones habitées (appels au centre de secours). Dans cette modélisation, les décisions d'intervention à l'instant t pourront être prises soit à partir des impacts simulés (supposés ici effectifs et observés) au même instant (situation de référence qui s'avère proche du fonctionnement réel à partir de remontées d'information de terrain), soit de façon anticipée à partir de prévisions des impacts pour les pas de temps futurs, que ces prévisions soient ensemblistes (prévisions basées sur AROME-EPS, pepi et pertDpepi) ou non (prévision basée sur une pluie future nulle RF0). Le nombre de bâtiments inondés bénéficiant de l'intervention des services de secours, ainsi que le délai de cette intervention - idéalement nul si l'anticipation est parfaite et les moyens de secours suffisants - seront les critères d'évaluation de la pertinence des décisions.

La modélisation est conduite au pas de temps horaire, ce qui signifie que les décisions d'affection des équipes d'intervention sont actualisées toutes les heures. À chaque pas de temps, la situation a évolué et une nouvelle simulation de prévisions est disponible pour réévaluer la situation future (les prévisions d'ensemble étant rafraîchies au pas de temps horaire). De nouvelles décisions peuvent être prises en fonction des priorités qui se dégagent de cette actualisation de l'analyse.

La Figure 7.3 présente le logigramme suivi pour modéliser la prise de décision et l'intervention des équipes du SDIS. Pour décider de l'affectation des équipes d'intervention, les tronçons de cours d'eau sont classés par ordre décroissant du nombre de bâtiments inondés simulé ou prévu. Dans le cas des prévisions, le nombre maximum de bâtiments prévu sur la période de prévision de 6 h est alors retenu, sur la base du quantile 75% pour les prévisions d'ensemble. Les tronçons sont ensuite examinés un par un pour essayer de leur affecter une ou plusieurs équipes d'intervention en fonction de l'importance des impacts simulés ou prévus.

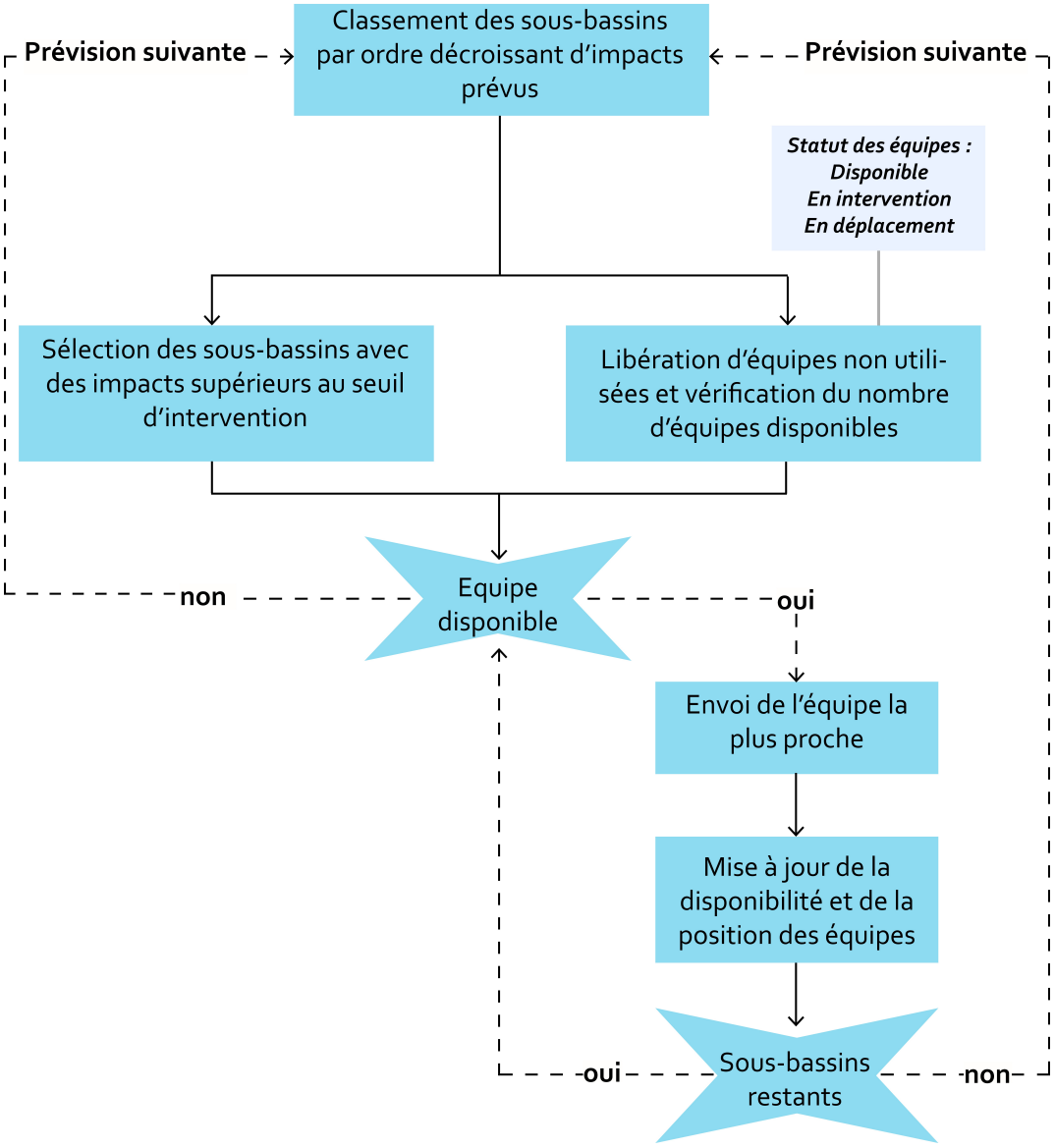

FIGURE 7.3 – *Principe de fonctionnement de la modélisation des interventions du SDIS*

La possibilité d'intervenir sur chacun des tronçons impactés par la crue dépend bien évidemment de la disponibilité d'équipes de secours en nombre suffisant. La localisation et l'activité de chacune des équipes est donc suivie et mise à jour tout au long de la modélisation. Trois statuts sont possibles pour les équipes de secours :

- **Disponible** : l'équipe est soit dans son centre de secours d'origine, soit sur le terrain, mais dans un secteur (tronçon de cours d'eau) ne nécessitant plus d'intervention.
- **En déplacement** : l'équipe se rend vers un tronçon sur lequel une intervention a été décidée et met un temps t pour y parvenir. Elle ne peut pas être réaffectée pendant son déplacement.
- **En intervention** : l'équipe de secours est en action sur un tronçon et ne peut pas être réaffectée.

La décision d'envoyer une équipe de secours sur un tronçon de cours d'eau dépend par ailleurs de trois paramètres :

- **Le seuil d'intervention** : ce paramètre correspond au nombre minimum de bâtiments inondés nécessaire pour justifier une intervention. Lorsque le nombre de bâtiments inondés n'atteint pas ce seuil, aucune équipe n'est envoyée sur place. Ce seuil est particulièrement important en cas de moyens limités, car il permet de conserver des moyens d'intervention disponibles, qui pourront être affectés ultérieurement sur des secteurs plus fortement touchés. Un seuil d'intervention élevé peut donc limiter l'effet négatif des fausses alarmes sur des tronçons présentant peu de bâtiments inondables.

- **La capacité d'intervention** : ce paramètre correspond au nombre de bâtiments inondés pouvant être secourus par une équipe d'intervention. Il permet si nécessaire de déclencher l'envoi de plusieurs équipes sur un même tronçon, lorsque le nombre de bâtiments inondés prévu dépasse la capacité d'intervention d'une seule équipe. Comme le seuil d'intervention, cette variable conditionne directement l'affectation et la répartition des équipes de secours. En effet, une capacité d'intervention faible va mobiliser un grand nombre d'équipes sur les tronçons les plus impactés, ce qui empêchera une intervention éventuelle sur d'autres tronçons par limitation des moyens disponibles.

- **La réaffectation des moyens** Ce paramètre booléen correspond à la possibilité de pouvoir réaffecter une équipe d'intervention quand elle n'a plus de bâtiments à secourir sur le tronçon où elle a été envoyée. Dans les modélisations réalisées ici, les équipes peuvent être réaffectées à partir du moment où le nombre de bâtiments touchés est repassé sous le seuil d'intervention. Alternativement, on pourrait considérer qu'une équipe envoyée en intervention ne pourra jamais être réaffectée, ou bien qu'elle pourra l'être uniquement après un temps t (correspondant à la durée des opérations de secours) à partir du moment où le pic de bâtiments inondés est passé.

Les tronçons pour lesquels le nombre de bâtiments inondés dépasse le seuil d'intervention sont examinés un par un, par ordre décroissant du nombre de bâtiments. Dans le cas où au moins une équipe d'intervention est disponible, l'équipe disponible la plus proche est sélectionnée et affectée en déplacement vers le tronçon en question. Cette boucle est répétée tant qu'il y a des besoins d'interventions et qu'il y a encore des équipes disponibles. Il est ensuite possible de passer au pas de temps suivant de la modélisation.

Les décisions d'affectation sont actualisées à chaque pas de temps en parcourant l'ensemble de l'événement. Les positions successives et disponibilités des équipes d'intervention sont archivées dans une table dédiée. La Figure 7.4 présente un exemple d'évolution de la localisation des équipes de secours au cours de l'événement de 2018 dans l'Aude, pour l'ensemble pepi.

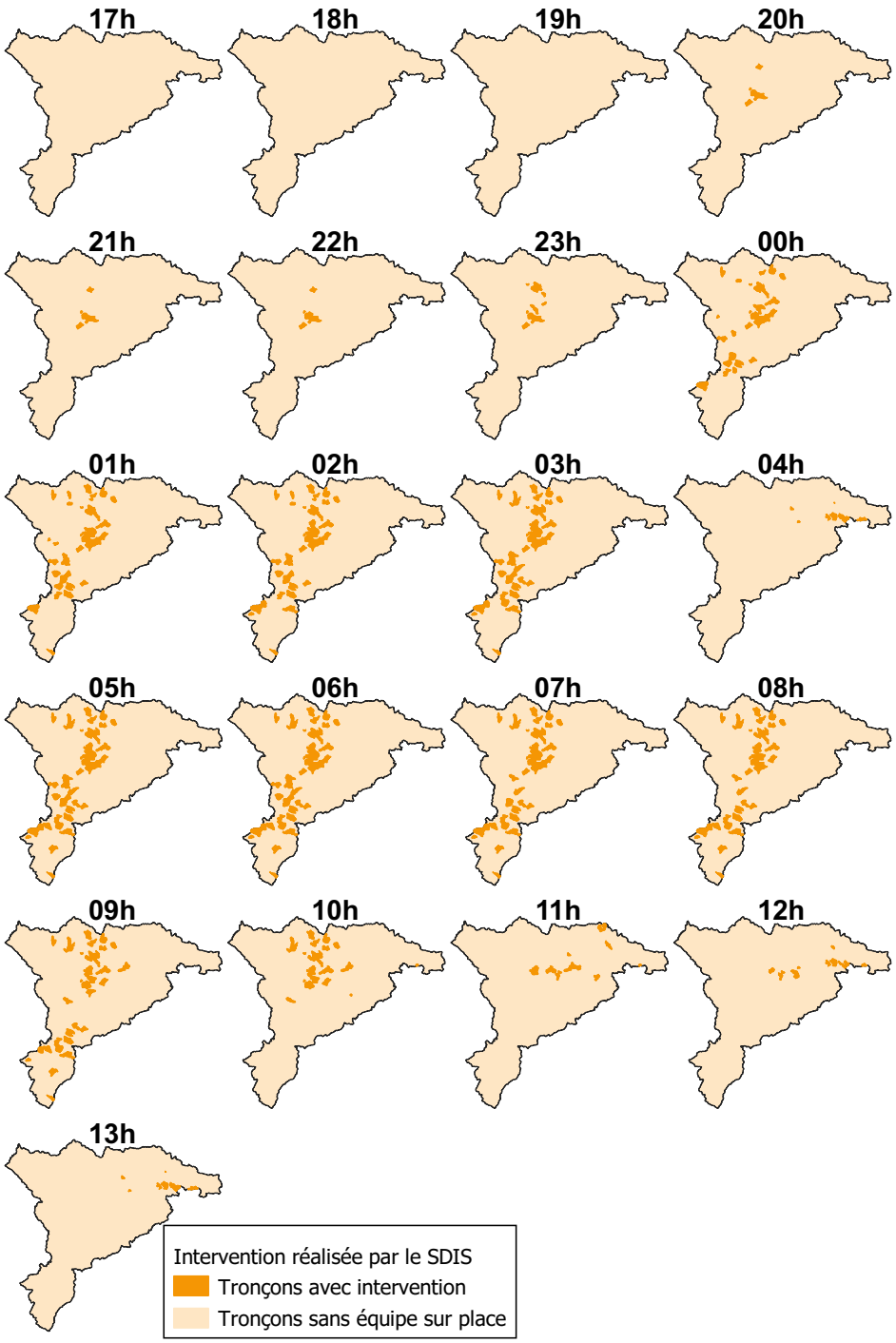

FIGURE 7.4 – *Localisation des interventions simulées pour la crue de 2018 dans l'Aude sur l'ensemble de la durée de l'événement pour l'ensemble pepi avec les paramètres de simulation suivants : 10 bâtiments pour le seuil d'intervention et 100 bâtiments/équipe pour la capacité d'intervention*

Un fichier *log* permet par ailleurs de stocker l'ensemble des décisions prises à chaque pas de temps. Un extrait de ce fichier est présenté sur la Figure 7.5.

```
Log Start Time: 2022-08-06 12:00:00
================================================================================

Ensemble de prévision évalué : pepi

Seuil d'alerte pour l'envoi d'une équipe : 10 bâtiments

Vitesse moyenne de l'équipe de secours :30 km/h

Dimensionnement des moyens de secours : 100 bâtiments par équipe de secours

Réaffectation des équipes après envoi : 1

================================================================================

================================================================================

pepi - 2018101417 : 75  Equipes disponibles restantes

================================================================================

pepi - 2018101418 : 75  Equipes disponibles restantes

================================================================================

pepi - 2018101419 : 75  Equipes disponibles restantes

pepi - 2018101419 : 1 Equipe(s) supplémentaire(s) pour le BV 2967

pepi - 2018101419 : Equipe 1201  envoyée sur BV 2967

pepi - 2018101419 : 1 Equipe(s) supplémentaire(s) pour le BV 3632

pepi - 2018101419 : Equipe 1202  envoyée sur BV 3632

pepi - 2018101419 : 1 Equipe(s) supplémentaire(s) pour le BV 3032

pepi - 2018101419 : Equipe 1203  envoyée sur BV 3032

pepi - 2018101419 : 1 Equipe(s) supplémentaire(s) pour le BV 2790
```

FIGURE 7.5 – *Extrait d'un fichier log pour un rejeu avec l'ensemble pepi avec les paramètres suivants : 10 bâtiments pour le seuil d'intervention et 100 bâtiments pour la capacité d'intervention*

À partir des informations qui précèdent, le nombre de bâtiments effectivement secourus peut être calculé sur chaque tronçon pour l'ensemble de l'événement, en prenant à chaque pas de temps le minimum entre le nombre de bâtiments effectivement inondés et le nombre de bâtiments pouvant être secourus par la ou les équipes présentes sur le tronçon en question. Un exemple d'évolution de ce nombre de bâtiments effectivement secourus est présenté sur la Figure 7.6. Dans le cas présenté, le nombre de bâtiments secourus augmente tardivement et reste assez éloigné du nombre de bâtiments inondés, du fait d'une arrivée trop tardive des équipes de secours. L'envoi d'équipes à temps n'est en effet pas possible dans ce cas, du fait du nombre d'équipes limitées et de leur mobilisation préalable sur d'autres tronçons. On note aussi sur cette Figure le retard de phase lié aux délais d'intervention. Les décisions prises mettent un certain temps à prendre effet et les moyens disponibles sur sites se retrouvent en excès à la fin de l'événement.

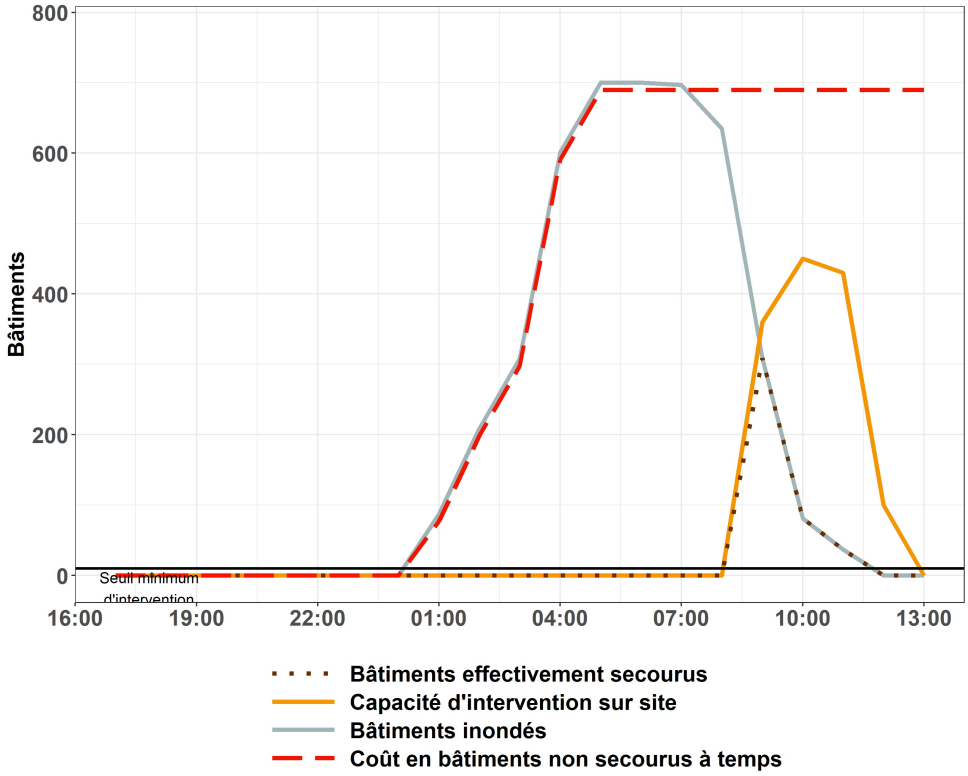

FIGURE 7.6 – *Illustration du nombre de bâtiments effectivement secourus sur le tronçon 4038 pour l'ensemble pepi avec les paramètres suivants :10 bâtiments pour le seuil d'intervention et 25 bâtiments pour la capacité d'intervention*

Une fonction de coût est finalement calculée pour chaque tronçon, afin d'évaluer et de pouvoir comparer les différentes stratégies possibles de gestion de l'événement (décisions basées sur la simulation ou la prévision). Le coût est exprimé en nombre de bâtiments non secourus à temps. Son calcul sera détaillé dans la Section 7.2.4.

7.2.2 Représentation des temps de trajet vers les sites d'intervention

Après avoir défini les moyens d'intervention disponibles et leur localisation, l'application du modèle nécessite également d'établir une matrice de temps de déplacement à l'échelle de la zone d'étude, afin de représenter le déplacement des équipes d'un centre de secours à un tronçon inondé, puis d'un tronçon inondé à un autre. Cette matrice a été construite à partir de la couche du réseau routier de la BD TOPO 2018 de l'IGN, pour relier chaque centre de secours à tous les tronçons de cours d'eau, mais également tous les tronçons entre eux. La méthodologie mise en place est détaillée ci-dessous :

- **Calcul des centroïdes des tronçons** Pour chaque tronçon de cours d'eau pour lequel les impacts ont été calculés, le centroïde est calculé pour ramener le tronçon à un unique point.
- **Création d'une unique couche de points** Une unique couche comprenant l'ensemble des points, centroïdes de tronçons et positions des centres de secours, est créée avec la nomenclature suivante : les centres de secours sont étiquetés de CS01 à CS42 et les tronçons gardent leurs identifiants déjà créés pour la modélisation hydrologique.

- **Alignement des points sur le réseau routier** Tous les points sont rapprochés du réseau routier de la BD TOPO, de manière à avoir une intersection entre le réseau et chacun des points.

- **Utilisation de l'algorithme OD-Matrix pour le calcul des distances de trajet.** Cet algorithme est déjà implémenté dans Qgis (Zanzibar, version 3.22.3). Il crée des paires origine-destination à partir d'un réseau vectoriel, avec la distance associée. La stratégie de coût retenue est celle de la distance minimale. L'augmentation des temps de parcours, liée aux coupures éventuelles d'axes routiers, n'est donc pas prise en compte dans cette première application. Ce pourrait être un développement à envisager à l'avenir, qui révélerait éventuellement un avantage du pré-positionnement de moyens de secours sur la base de prévisions hydrométéorologiques, dans des secteurs qui pourraient être isolés lors des crues.

- **Finalisation de la matrice de temps de déplacement** L'étape précédente a permis d'obtenir une matrice de distances, le choix le plus facile et relativement réaliste fait ici a été de diviser cette matrice par une vitesse moyenne de 30 km/h pour estimer les temps de trajet. La vitesse retenue s'avère cohérente avec un déplacement en conditions climatiques dégradées.

- **Passage des centres de secours aux équipes** Afin de différencier chaque équipe, un identifiant unique est attribué, en particulier pour intervenir sur différents tronçons. Une duplication des lignes de chaque centre de secours est donc faite afin d'obtenir une information par équipe (c'est-à-dire véhicule) d'intervention.

Cette matrice de déplacement permet finalement d'orienter les équipes les plus proches disponibles vers un tronçon où une intervention est requise, en fonction de leur temps de trajet estimé.

7.2.3 Choix des autres paramètres de modélisation

Le premier paramètre influant fortement sur les résultats de la modélisation est le seuil d'intervention. Dans les simulations réalisées, les différents seuils d'intervention testés ont été de 10 bâtiments, 25 bâtiments et 50 bâtiments. Ces valeurs restent à priori assez réalistes. Si le seuil d'intervention de 50 bâtiments peut paraître très élevé au premier abord, il faut noter que tous les bâtiments comptabilisés dans une zone inondée ne subissent pas forcément de dégâts significatifs. Les statistiques des assureurs montrent par exemple qu'en moyenne, moins d'un bâtiment sur deux au sein d'une zone inondée fait l'objet d'une déclaration de sinistre (Le Bihan *et al.* 2017). Le fait de comptabiliser l'ensemble des bâtiments au sein de l'emprise inondée peut donc conduire à surestimer assez largement la réalité des impacts et des opérations de secours associées.

Le deuxième paramètre à fixer, influant lui aussi fortement sur les résultats, est la capacité d'intervention, qui pour rappel correspond au nombre de bâtiments pouvant être secourus par une seule équipe. La valeur réelle de ce paramètre est relativement difficile à évaluer. Nous avons considéré ici que le nombre d'interventions de secours d'urgence à effectuer (par

exemple l'évacuation de personnes vulnérables) s'avère probablement nettement inférieur au nombre de bâtiments réellement présents dans la zone inondée. Pour cette raison, une capacité d'intervention de plusieurs dizaines de bâtiments par équipe semble réaliste. La gamme de valeurs retenue varie de 10 bâtiments à 150 bâtiments. Cette plage relativement large a été fixée en tenant compte de la capacité d'intervention globale obtenue en multipliant la capacité d'intervention par le nombre d'équipes mobilisables. Cette capacité d'intervention globale varie finalement de 750 à 22 500 bâtiments en considérant un nombre d'équipes variant de 92 à 184. Ceci est à comparer à l'estimation du nombre maximum de bâtiments inondés au cours de la crue de 2018, qui est proche de 11 000 d'après la simulation présentée au chapitre 6. La plage de variation retenue permet donc de disposer de scénarios dans lesquels la capacité d'intervention globale s'avère très limitée, ou au contraire très supérieure aux besoins.

Du fait du nombre d'équipes relativement limité, la réaffectation des moyens après une intervention sur un tronçon est également un paramètre qui peut influer significativement sur les résultats de modélisation. Le choix retenu ici est une réaffectation immédiate lorsque le nombre de bâtiments inondés repasse sous le seuil d'intervention.

7.2.4 Mesure de l'efficacité des interventions déclenchées (Fonction de coût)

Afin d'évaluer l'efficacité des décisions prises dans les différentes simulations, une fonction de coût associée à chaque simulation a été définie. Le coût est défini comme le nombre de bâtiments pour lesquels les opérations de secours n'ont pas pu être effectuées, ou l'ont été trop tardivement.

Ce coût est estimé comme un pourcentage du nombre de bâtiments inondés, qui varie en fonction de l'heure d'arrivée des secours. Le Tableau 7.1 ci-dessous présente les pourcentages appliqués, qui varient de 0% si une équipe de secours est présente sur place avant l'inondation, à 100% si aucune équipe de secours n'est arrivée sur place 3 h après le début d'inondation.

Tableau 7.1 – *Liste des différentes valeurs de variables testées*

Situations	**Coût exprimé en % du nombre de bâtiments inondés**
Équipe(s) sur place avant les impacts	0% des bâtiments inondés
Équipe(s) sur place entre 0 et 1h (inclus) après les impacts	10% des bâtiments inondés
Équipe(s) sur place entre 1h (exclus) et 3h (inclus) après les impacts	50% des bâtiments inondés
Équipe(s) sur place plus de 3h (exclus) après les impacts	100% des bâtiments inondés

Le coût est dans un premier temps calculé à l'échelle de chaque tronçon, puis une somme est réalisée pour obtenir une valeur à l'échelle de l'ensemble du territoire modélisé.

La Figure 7.7 présente la méthodologie appliquée pour le calcul de cette fonction de coût sur chaque tronçon. Le calcul nécessite en pratique de comparer temporellement la courbe de simulation des bâtiments inondés (d'après les pluies Antilope J+1, en gris sur la Figure 7.7) et la capacité d'intervention envoyée sur le tronçon concerné (en vert sur la Figure 7.7). La première étape consiste à abaisser, de la valeur du seuil d'intervention, la courbe des bâtiments inondés, en effet, les bâtiments sous ce seuil ne sont pas considérés pour l'envoi d'une équipe. Une fois cette étape réalisée, les deux courbes sont parcourues heure par heure (pas de temps de modélisation). Dès que le nombre de bâtiments inondés est strictement supérieur à la capacité d'intervention présente sur le tronçon, le pas de temps correspondant est comptabilisé.

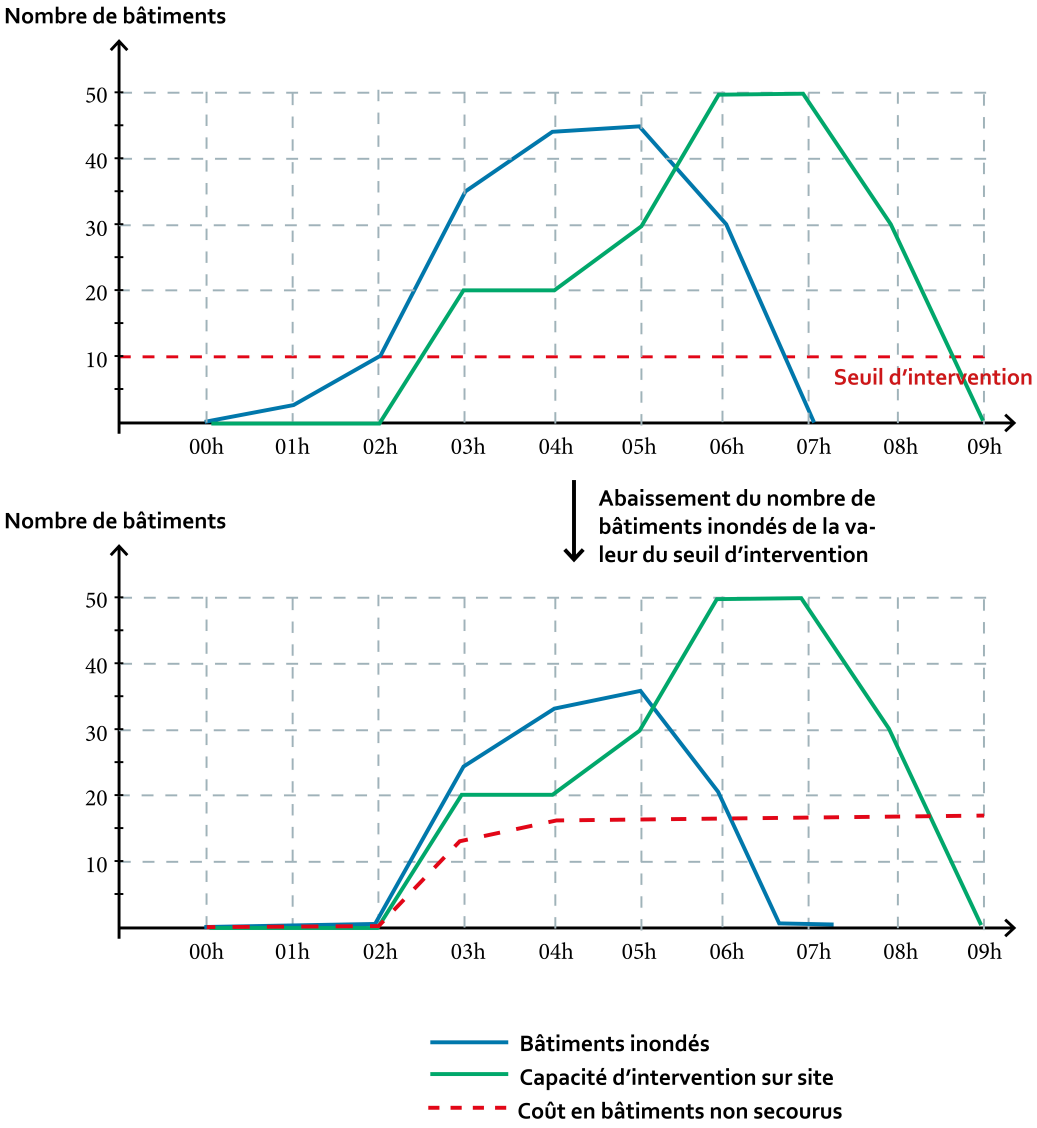

FIGURE 7.7 – *Illustration de la méthode de calcul de la fonction de coût sur un cas théorique*

L'heure à laquelle (si elle existe) la capacité d'intervention sur site devient égale au nombre de bâtiments inondés est alors recherchée. En fonction de cette heure, le coût associé en nombre de bâtiments non secourus à temps est calculé à partir des ratios présentés dans le Tableau 7.1. En repartant de l'illustration présentée sur la Figure 7.7, de 00 h à 02 h, la capacité d'intervention sur le tronçon reste supérieure aux bâtiments inondés, la fonction de coût reste donc à 0. Dès 03 h du matin, le nombre de bâtiments inondés devient supérieur à la capacité d'intervention sur site. Le nombre de bâtiments inondés est alors noté (26) et l'heure correspondante sur la courbe d'intervention est recherchée, entre 04 h et 05 h, soit moins de 3 h après. À 03 h, 5 bâtiments sont déjà secourus sur le tronçon, le coût est donc calculé sur les 26 - 5 bâtiments non secourus. La fonction de coût passe donc à 10,5 (50% de 21). Au pas de temps suivant (04 h), le nombre de bâtiments inondés est toujours supérieur à la capacité d'intervention présente sur le tronçon. Le nombre de bâtiments nouvellement inondés est de 8, et ces bâtiments sont secourus dès 05 h du matin, soit une heure plus tard. La valeur de la fonction de coût augmente donc de 0.8. À 05 h, la fonction de coût s'élève donc à 11,3. Après 05 h, le nombre de bâtiments inondés reste en dessous de la capacité d'intervention sur le tronçon, la fonction de coût reste alors constante à 11.3.

Cette fonction de coût a pour avantage de ne compter les bâtiments non secourus qu'une seule fois, puisque seuls les bâtiments nouvellement inondés sont considérés. De plus, le retard des équipes d'intervention est pris en compte de manière graduée selon la valeur du retard.

7.2.5 Choix des paramètres de modélisation

Plusieurs valeurs de variables et combinaisons de variables pour le seuil d'intervention, la capacité d'intervention et le nombre d'équipes ont été testées. Le Tableau 7.2 ci-dessous récapitule les différentes valeurs retenues.

Tableau 7.2 – *Liste des différentes valeurs de variables testées*

Nombre d'équipes de secours	92 et 184
Seuil d'intervention	10, 25 et 50
Capacité d'intervention	10, 25, 50, 75, 100, 125 et 150
Réaffectation des moyens	Réaffectation simple

7.2.6 Quelques améliorations possibles du modèle

Ce modèle est une proposition qui reste perfectible. Plusieurs améliorations peuvent être imaginées pour le rapprocher de la réalité de terrain.

En effet, au-delà des bâtiments, c'est surtout la population qui doit être secourue en priorité par un service de secours. Le choix des bâtiments a été fait pour le caractère spatial de la donnée, malgré le filtrage fait dans le Chapitre 6 pour ne garder que les bâtiments en dur et résidentiel, cela ne représente pas totalement la population sur le territoire.

Il pourrait être envisageable de pondérer le nombre de bâtiments par tronçon à la population associée. De plus, le bâtiment est compté dès qu'il est dans l'emprise interpolée, sans tenir compte de la hauteur associée. Cela est également une limite puisque la hauteur d'eau s'avère être un critère supplémentaire pour engager une intervention ou non. Le déploiement des équipes selon le classement des tronçons par ordre décroissant d'impacts est également discutable. Quand les moyens commencent à devenir limités, toutes les équipes encore disponibles sont affectées aux tronçons avec le plus d'impacts, les tronçons avec moins d'impacts prévus ne sont alors pas traités.

Enfin, une piste d'amélioration pourrait être la prise en compte des conditions de circulation, avec en particulier les ralentissements ou coupures de routes pouvant ralentir où empêcher une équipe de se rendre sur un tronçon. Il serait alors envisageable de repérer tous les sous-bassins traversés pour aller d'un point A à un point B et de pondérer le temps de déplacement par le nombre de bâtiments touchés sur chaque tronçon traversé. Néanmoins, en première approche, l'utilisation d'une vitesse constante de 30 km/h paraît raisonnable.

7.3 Première simulation avec un nombre illimité d'équipes de secours

Afin d'illustrer une première fois la modélisation mise en place, ce paragraphe présente les résultats obtenus pour des moyens d'intervention illimités en première approche. Par moyens illimités est entendu un nombre illimité d'équipes sur la zone d'étude et par centre de secours. Cela signifie qu'un nombre infini d'équipes peut partir d'un seul centre de secours sans aucune limite.

Ce choix permet de supprimer les risques associés à la prise de décision à partir de prévisions. En effet, dans cette modélisation, l'intervention d'équipes de secours suite à des fausses alarmes n'altère pas du tout les capacités d'intervention restantes. Autrement dit, l'envoi d'équipes de secours n'est jamais conditionné par la disponibilité d'au moins une équipe. Dans cette situation, les décisions prises à partir de prévisions doivent logiquement apporter une plus-value liée à une meilleure anticipation.

Les paramètres suivants ont été retenus pour la modélisation présentée :

- Seuil d'intervention : 10 bâtiments par tronçon
- Vitesse de déplacement : 30 km/h
- Capacité d'intervention : 100 bâtiments par équipe

La modélisation a été renouvelée en utilisant les prévisions de bâtiments inondés issues des trois produits de prévision d'ensemble AROME-EPS, pepi et pertDpepi, et de la prévision de pluie nulle RF0. La modélisation de référence, correspondant à une décision sans prévision, a été obtenue en utilisant les bâtiments inondés simulés pour les pluies observées Antilope J+1. Cela signifie que quand un dépassement de seuil d'intervention est observé sur un tronçon, les équipes de secours qui sont envoyées, ne peuvent intervenir qu'avec un certain retard, correspondant à leur temps de trajet.

7.3.1 Évolution spatio-temporelle des équipes d'intervention pendant l'événement

La Figure 7.8 illustre la répartition spatio-temporelle des équipes en intervention sur les différents tronçons pour la simulation de référence RS, la prévision de pluie future nulle et le produit pepi. Cette Figure permet d'avoir une vision globale de la dynamique des interventions de secours pour les différentes modélisations réalisées.

La première observation qui peut être faite concerne l'évolution temporelle des interventions. En effet, selon la modélisation observée, l'envoi des équipes vers les premiers tronçons ne se fait pas à la même heure. Ainsi, pour la modélisation de référence, les premières équipes arrivent à 23 h sur le premier tronçon secouru (portant le n°2967). Pour les modélisations basées sur des prévisions, les équipes arrivent sur ce même tronçon à 20 h pour pepi et 23h pour RF0.

Une autre observation est qu'un plus grand nombre de tronçons est globalement secouru avec la prévision (pepi et dans une moindre mesure RF0) que pour la modélisation de référence issue de la simulation (analogue à une observation de terrain). Cette observation est assez logique puisque pepi est un scénario avec une prévision de pluie non nulle, alors que RF0 qui est un scénario de pluie future nulle, et que la modélisation de référence n'utilise pas une prévision. L'utilisation des prévisions a pour conséquence de déclencher des interventions anticipées, mais également plus nombreuses en raison de la présence de faux positifs.

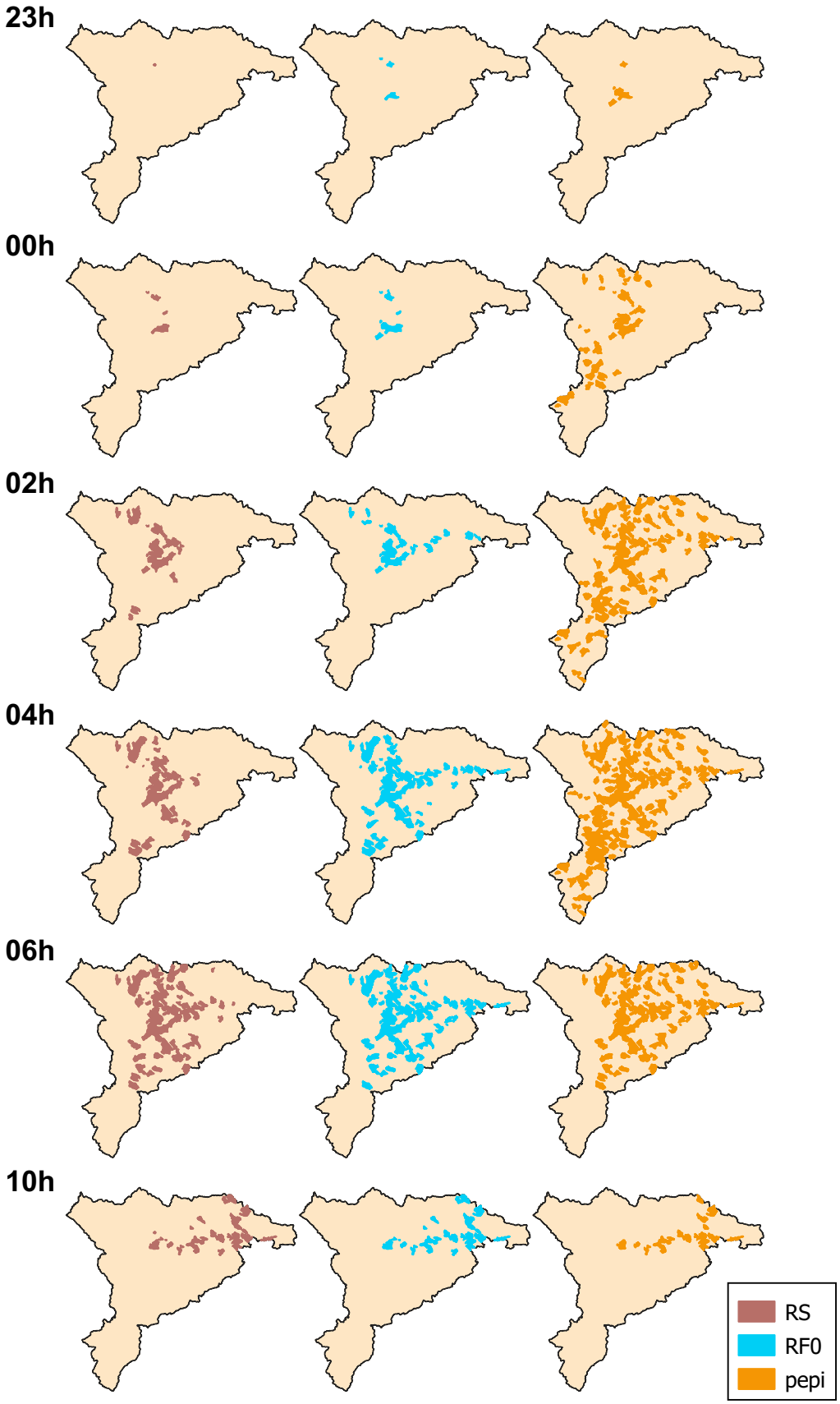

FIGURE 7.8 – *Évolution spatio-temporelle des équipes (nombre illimité) en intervention pour la simulation de référence RS (sans anticipation des impacts), pour la prévision de pluie future nulle, RF0 et pepi*

7.3.2 Évolution du nombre de bâtiments secourus selon les différents scénarios

La Figure 7.9 présente le nombre de bâtiments effectivement secourus, heure par heure, par rapport au nombre de bâtiments dans l'emprise inondée pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la simulation d'impacts à partir de la pluie observée (RS). Tous les résultats obtenus tronçon par tronçon sont ici sommés, ce qui permet d'obtenir une vue globale de l'efficacité des secours sur l'ensemble du territoire étudié.

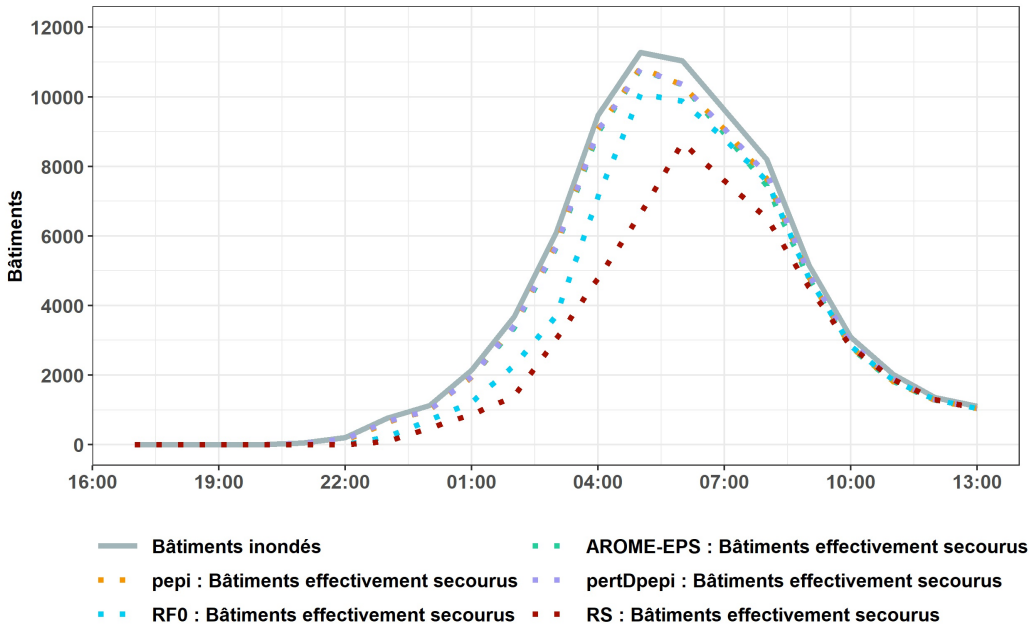

FIGURE 7.9 – *Bâtiments effectivement secourus avec un nombre d'équipes illimité pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la simulation d'impacts à partir de la pluie observée (RS)*

Le premier constat est que toutes les courbes de bâtiments effectivement secourus se trouvent en dessous de celle du nombre de bâtiments inondés (en gris). Cela signifie que même avec des moyens d'intervention illimités, tous les bâtiments ne peuvent pas être secourus. Cela peut paraître surprenant au premier abord, car il semblerait logique d'observer 100% de bâtiments effectivement secourus dans ce cas, même avec un certain retard. L'explication se trouve dans le détail des résultats tronçon par tronçon. En effet, pour certains tronçons, le nombre de bâtiments inondés augmente très brutalement, et diminue juste après le pic de manière aussi brutale. Cela entraîne donc une détection très difficile pour les prévisions et une réponse décalée dans le temps, avec comme conséquence un nombre de bâtiments effectivement secourus plus faible que le nombre de bâtiments inondés. Pour rappel, le nombre de bâtiments effectivement secourus correspond au minimum entre bâtiments inondés et la capacité d'intervention sur site. La Figure 7.10 est un bon exemple de cette situation. Sur ce tronçon, en raison du retard des interventions de secours, le nombre de bâtiments effectivement secourus (Figure 7.10 a)) est inférieur au nombre de bâtiments inondés pour la modélisation de référence (basée sur la simulation Antilope J+1). En revanche, la prévision

d'ensemble pepi (Figure 7.10 c)) offre suffisamment d'anticipation pour éviter cette situation avec un envoi d'équipe sur site une heure avant les impacts, et dans ce cas le nombre de bâtiments secourus est bien égal au nombre maximum de bâtiments inondés.

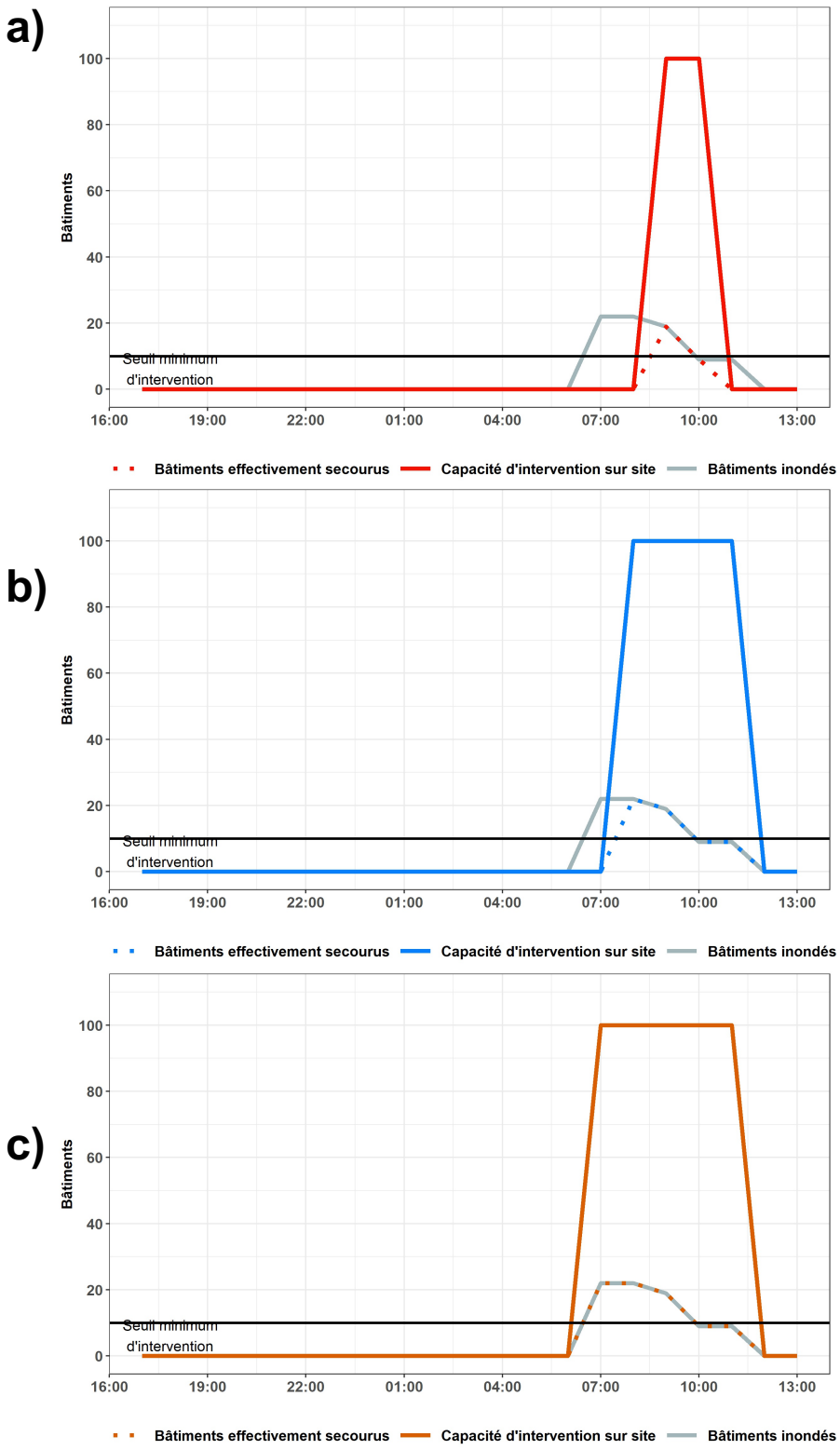

FIGURE 7.10 – *Bâtiments inondés, capacité d'intervention et bâtiments effectivement secourus pour le tronçon 716 pour a) la simulation de référence d'impacts basée sur les pluies Antilope J+1, RS, b) la prévision de pluie future nulle, RF0 et c) la prévision d'ensemble pepi*

Le deuxième constat issu de la Figure 7.9 concerne la hiérarchie observée entre les cinq courbes de bâtiments effectivement secourus. Ainsi, les courbes correspondant aux trois produits de prévision d'ensemble, sont toutes les trois analogues, et très proches de la courbe grise de bâtiments inondés recherchée. La courbe bleue correspondant à RF0 est en deçà, avec un nombre de bâtiments effectivement secourus plus réduit, en particulier en phase de montée de crue où un retard est observé. Cette observation est logique, car RF0 est un scénario avec une prévision de pluie future nulle, l'anticipation est donc moindre, et pour un certain nombre de tronçons amont, RF0 n'offre pas du tout d'anticipation. La Figure 7.10 b) est une bonne illustration de cette absence d'anticipation pour cette prévision. Enfin, de façon logique, la modélisation de référence, basée sur les bâtiments inondés simulés (équivalent à une absence de prévision) présente encore plus de retard, avec un nombre maximum de bâtiments secourus nettement plus faible (Figure 7.10 a)).

Cette première analyse se révèle assez intéressante du point de vue de la comparaison entre le scénario de pluie future nulle RF0 et la simulation de référence RS (absence de prévision). En effet, comme cela avait déjà été observé sur l'analyse des prévisions de débit faite dans le Chapitre 5, une prévision de pluie future nulle permet d'anticiper sur certains tronçons les impacts sur les bâtiments. Cette prévision à moindre frais de calculs et ne nécessitant aucune prévision de pluie offre sur un certain nombre de tronçons une anticipation suffisante pour éviter un bon nombre d'impacts.

7.3.3 Calcul du coût en bâtiments non secourus, pour toutes les modélisations

Ce paragraphe aborde les résultats du calcul de la fonction de coût, selon les modalités décrites dans la Section 7.2.4. Les coûts ont été calculés sur chaque tronçon puis sommés sur l'ensemble du domaine étudié. Les résultats obtenus sont présentés dans la Figure 7.11.

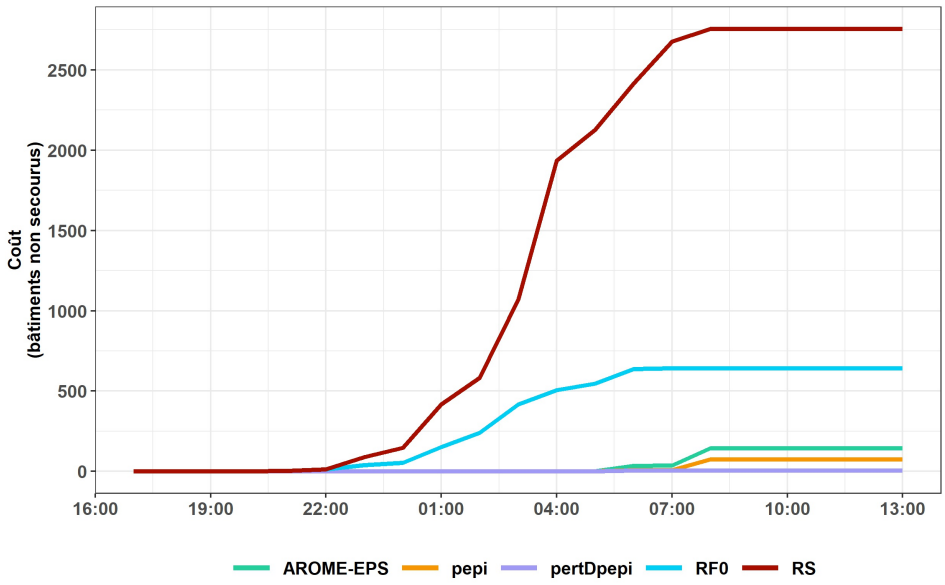

FIGURE 7.11 – *Fonctions de coût calculées sur l'ensemble des tronçons pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la simulation d'impacts à partir de la pluie observée (RS)*

Les résultats obtenus sont cohérents avec la Figure 7.9. Logiquement, la modélisation de référence (absence de prévision) présente le coût de loin le plus élevé (2830 bâtiments). La modélisation basée sur la prévision RF0 vient ensuite, avec un coût de 823,4 bâtiments. Ces valeurs viennent confirmer les observations réalisées dans la Section précédente : l'utilisation du scénario de pluie future nulle, RF0, permet un gain significatif du point de vue des secours.

Les trois produits de prévision d'ensemble aboutissent de leur côté à des coûts faibles : 145 bâtiments pour AROME-EPS, 76 pour pepi et 7 bâtiments pour pertDpepi. pertDpepi semble donc avoir un coût extrêmement réduit par rapport à AROME-EPS et pepi. Cela est cohérent avec les résultats obtenus aux chapitres précédents (évaluation des prévisions de pluie et de débits), où pertDpepi apparaissait toujours un peu meilleur en termes de CRPS ou encore d'AUC.

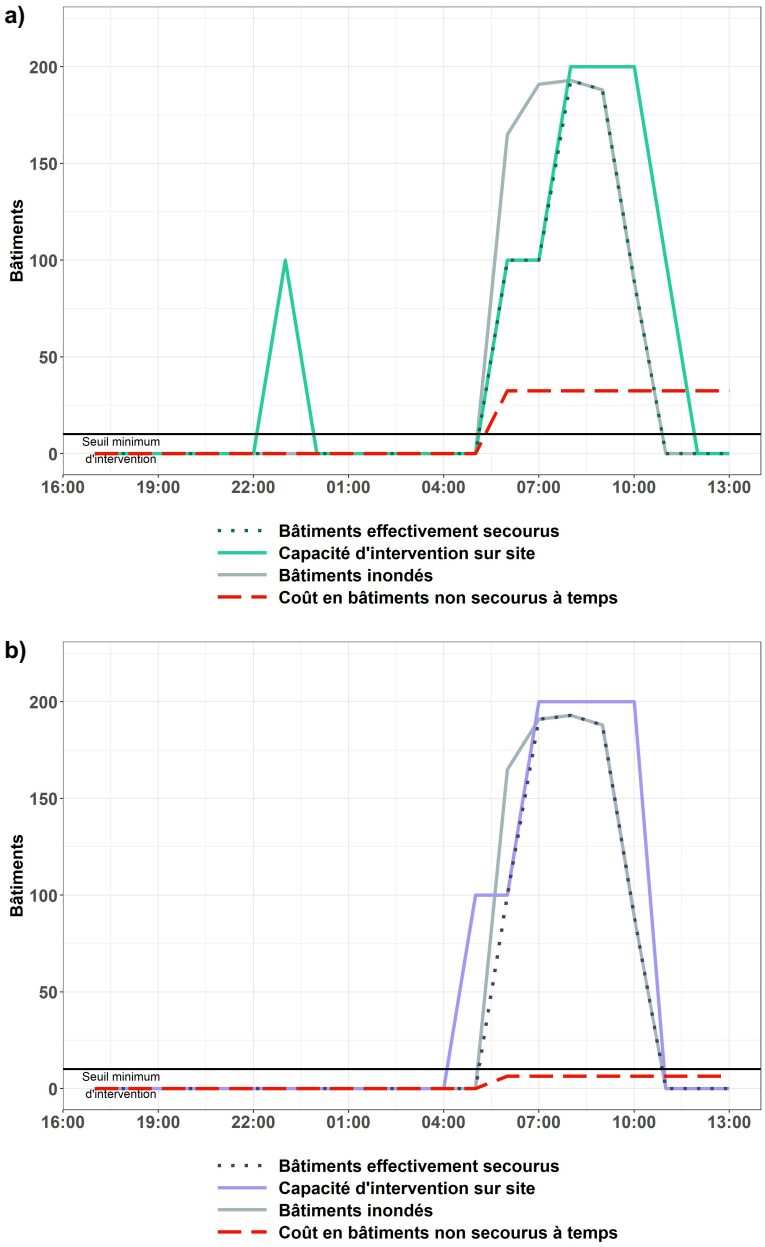

FIGURE 7.12 – *Fonctions de coût calculées pour le tronçon 697 pour a) AROME-EPS et b) pertDpepi*

Un tronçon en particulier explique une grande partie des écarts entre AROME-EPS et pertDpepi. Les résultats obtenus sur ce tronçon sont présentés sur la Figure 7.12. Sur cette Figure, il apparaît qu'AROME-EPS permet de déclencher très tôt une intervention des secours, mais pour un seul pas de temps de prévision isolé. Cette prévision n'étant pas confirmée, l'équipe d'intervention est donc démobilisée et repositionnée ailleurs. Quand les bâtiments inondés apparaissent sur le tronçon, le délai nécessaire pour renvoyer une équipe entraîne un coût élevé. Au contraire, pertDpepi anticipe moins tôt le dépassement, mais de manière cohérente et répétée (sur plusieurs pas de temps) et plus en accord avec la dynamique des impacts simulés. Cela permet d'obtenir un coût réduit.

7.3.4 Évolution temporelle du nombre d'équipes mobilisées

La Figure 7.13 présente l'évolution du nombre d'équipes mobilisées au cours de l'événement pour les différentes modélisations. Les moyens de secours étant illimités, il n'y a donc pas ici de plafonnement du nombre d'équipes mobilisées.

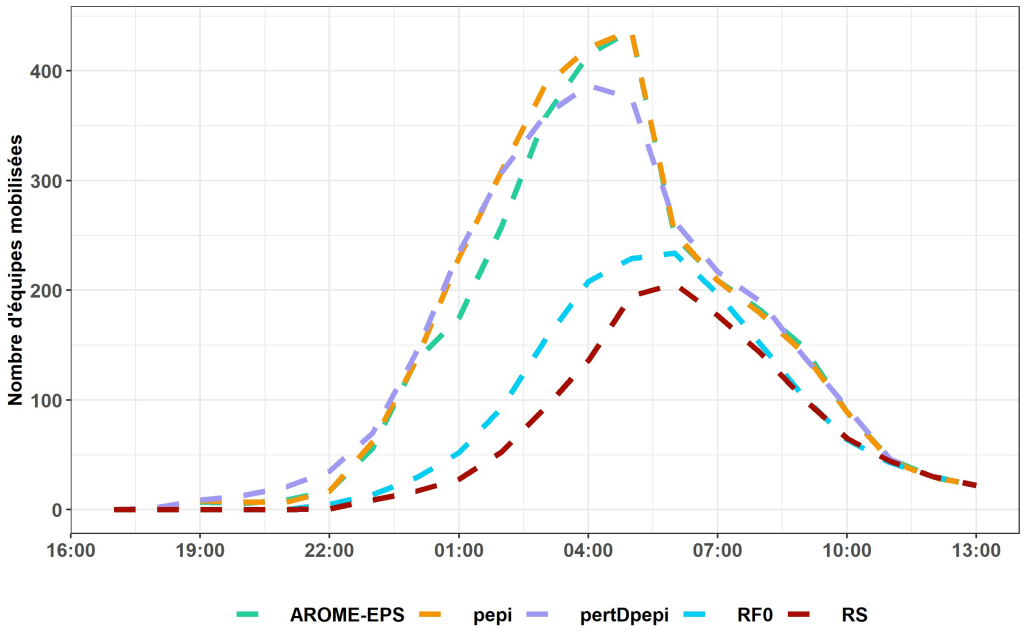

FIGURE 7.13 – *Nombre d'équipes mobilisées au cours de l'événement pour la modélisation en moyens illimités, pour les trois produits de prévision d'ensemble (AROME-EPS, pepi et pertDpepi), la prévision de pluie future nulle (RF0) et la modélisation de référence à partir du nombre de bâtiments simulé RS (sans prévision)*

Le nombre d'équipes mobilisées apparaît comme très variable selon le scénario considéré. De manière logique, le nombre d'équipes est bien plus grand pour les scénarios basés sur des prévisions de pluie (AROME-EPS, pepi et pertDpepi), par comparaison avec le scénario basé sur une pluie future nulle (RF0) ou encore l'observation des impacts (simulation sans prévision) : 518 pour AROME-EPS, 476 pour pepi, 439 pour pertDpepi, 248 pour RF0 et 214 pour RS.

Il faut également noter que le nombre d'équipes mobilisées est assez différent selon les trois produits de prévision et qu'un plus grand nombre d'équipes (comme c'est le cas pour

AROME-EPS) n'est pas forcément signe d'une meilleure performance de la prévision, comme l'a montré le calcul du coût la Section 7.3.3. En effet, la quantité d'équipes mobilisée peut être compensée par une arrivée tardive de ces équipes dans le calcul du cout.

Or un décalage temporel est aussi visible sur le maximum d'équipes envoyées. Ainsi, pour pertDpepi, le maximum d'équipes envoyées est atteint à 3 heures du matin, pour pepi à 4 heures et pour AROME-EPS à 5h. Pour RF0 et RS, le pic d'équipes en intervention est observé à 5 heures le 15 octobre. Ce décalage illustre une meilleure anticipation obtenue avec pertDpepi et pepi. En revanche, AROME-EPS offre moins d'anticipation puisque le maximum de bâtiments inondés (courbe grise de la Figure 7.9) atteint également son maximum à 5 heures du matin. Ces deux observations se pressentaient déjà au vu de l'évolution spatio-temporelle présentée sur la Figure 7.8.

7.3.5 Bilan de la première simulation avec des moyens d'intervention illimités

Cette première simulation avec un nombre d'équipes illimité conduit aux résultats auxquels on pouvait s'attendre. Elle a permis d'illustrer la mise en œuvre de la modélisation, ainsi que les différentes variables de sortie auxquelles il est important de prêter attention pour l'analyse des résultats : le nombre de bâtiments effectivement secourus (plutôt que la capacité d'intervention seulement) à comparer au nombre de bâtiments effectivement inondés, le coût engendré en termes de bâtiments non secourus à temps, et finalement l'évolution temporelle du nombre d'équipes mobilisées.

Par ailleurs, cette première modélisation confirme la nette plus-value des trois prévisions d'ensemble AROME-EPS, pepi et pertDpepi (plus particulièrement), dans l'hypothèse de moyens d'intervention illimités. Cette plus-value est directement liée à la meilleure anticipation des zones à secourir. Toutefois, cette première évaluation reste très idéalisée puisque les hypothèses de modélisation (nombre d'équipes illimité) ne pénalisent pas les faux positifs associés à la prévision. Dans la pratique, l'indisponibilité des équipes de secours envoyées en intervention, suite à des fausses alarmes, peut limiter fortement les gains associés à l'utilisation des prévisions. Ce cas de Figure est abordé dans le paragraphe qui suit.

7.4 Simulations avec des moyens d'intervention limités

Ce paragraphe présente les résultats des modélisations réalisées pour des situations plus réalistes, dans lesquelles le nombre d'équipes de secours s'avère limité. Ces modélisations permettent de représenter l'effet négatif des fausses alarmes, qui mobilisent inutilement des moyens d'intervention.

Une analyse de sensibilité est réalisée aux différents paramètres permettant de faire varier la quantité d'interventions à réaliser (seuil d'intervention), ainsi que les moyens de secours disponibles (nombre et capacité d'intervention des équipes).

7.4.1 Résultats pour une situation proche de celle d'octobre 2018

La simulation proposée ici est réalisée en prenant les paramètres jugés les plus représentatifs de la situation réellement vécue durant l'événement de 2018 dans l'Aude. Les valeurs retenues pour les différents paramètres sont présentées dans le Tableau 7.3.

TABLEAU 7.3 – *Valeurs retenues pour les différents paramètres de simulation*

Nombre d'équipes mobilisées	184
Seuil d'intervention	25
Vitesse de déplacement	30 km/h
Capacité d'intervention	75
Réaffectation des moyens	Réaffectation simple

Concernant le nombre d'équipes choisi, il est volontairement doublé par rapport aux moyens présentés sur la Figure 7.1 afin de tenir compte des équipes extérieures au département envoyées en renforts.

Les Figures 7.14, 7.15, 7.16 et 7.17 ci-dessous présentent la chronologie des interventions prévues par la simulation basée sur le produit de prévision pertDpepi (tronçons en violet) et les interventions réalisées par le SDIS (cercles noirs).

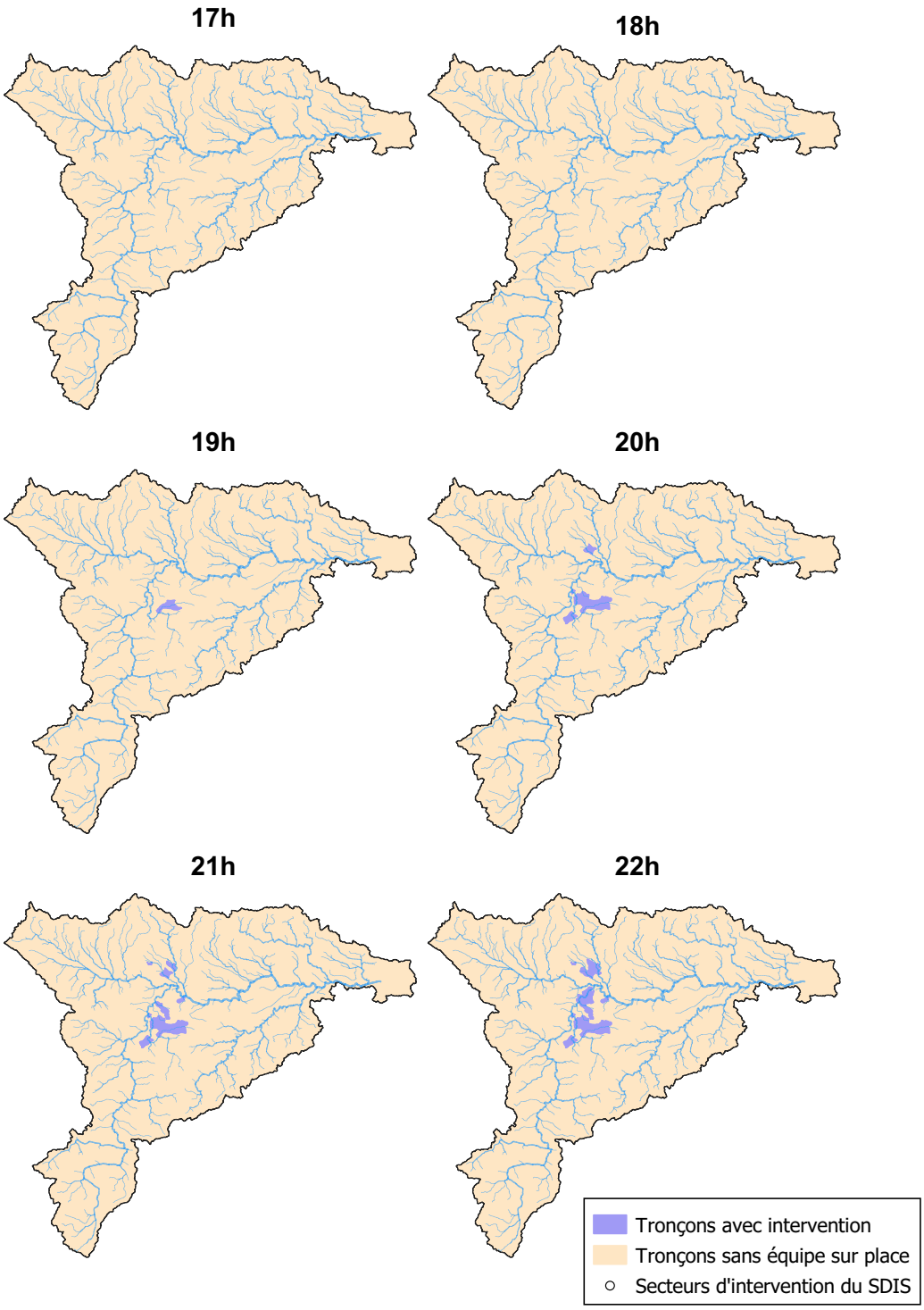

FIGURE 7.14 – *Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 17 h à 22 h*

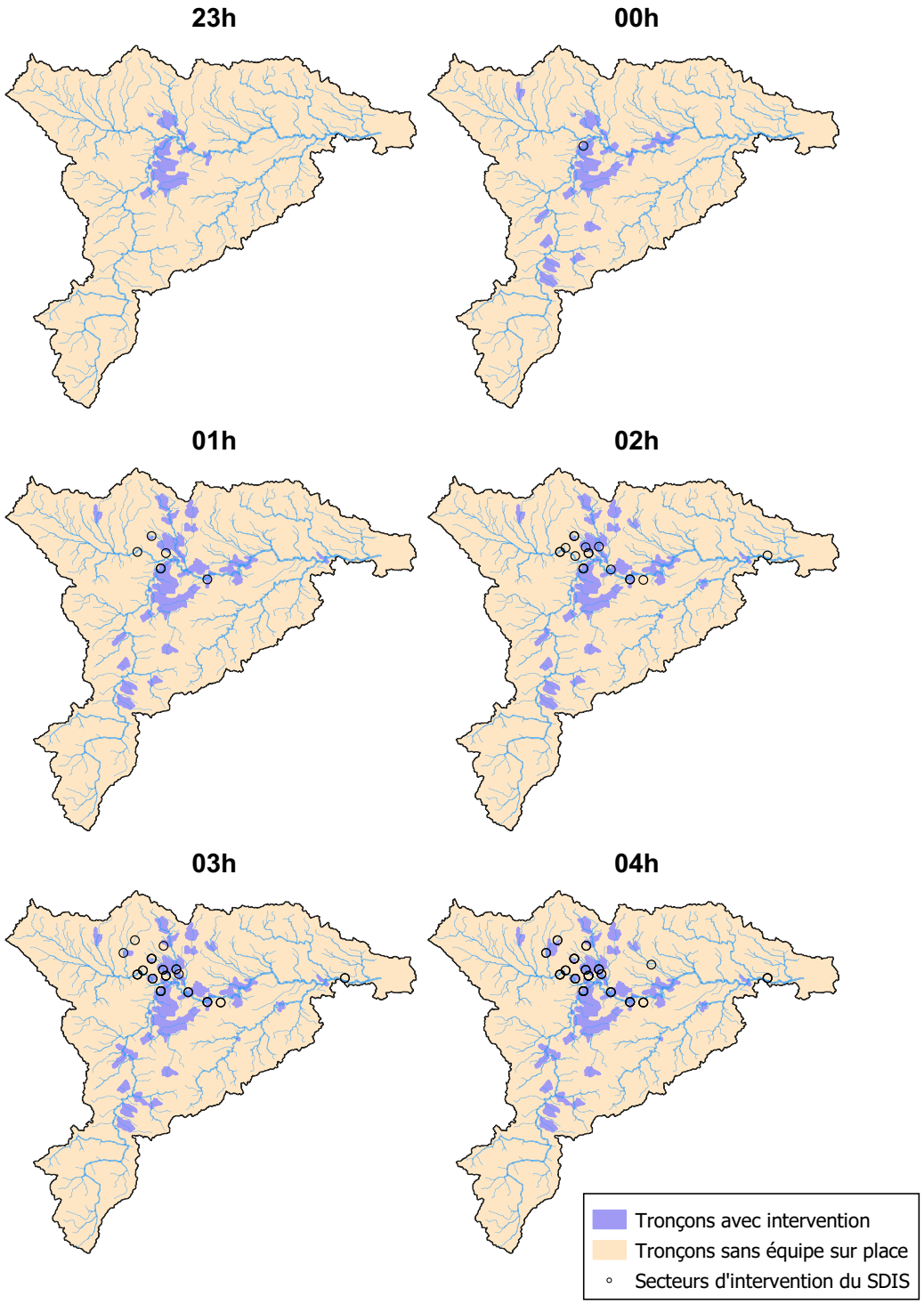

FIGURE 7.15 – *Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 23 h à 04 h*

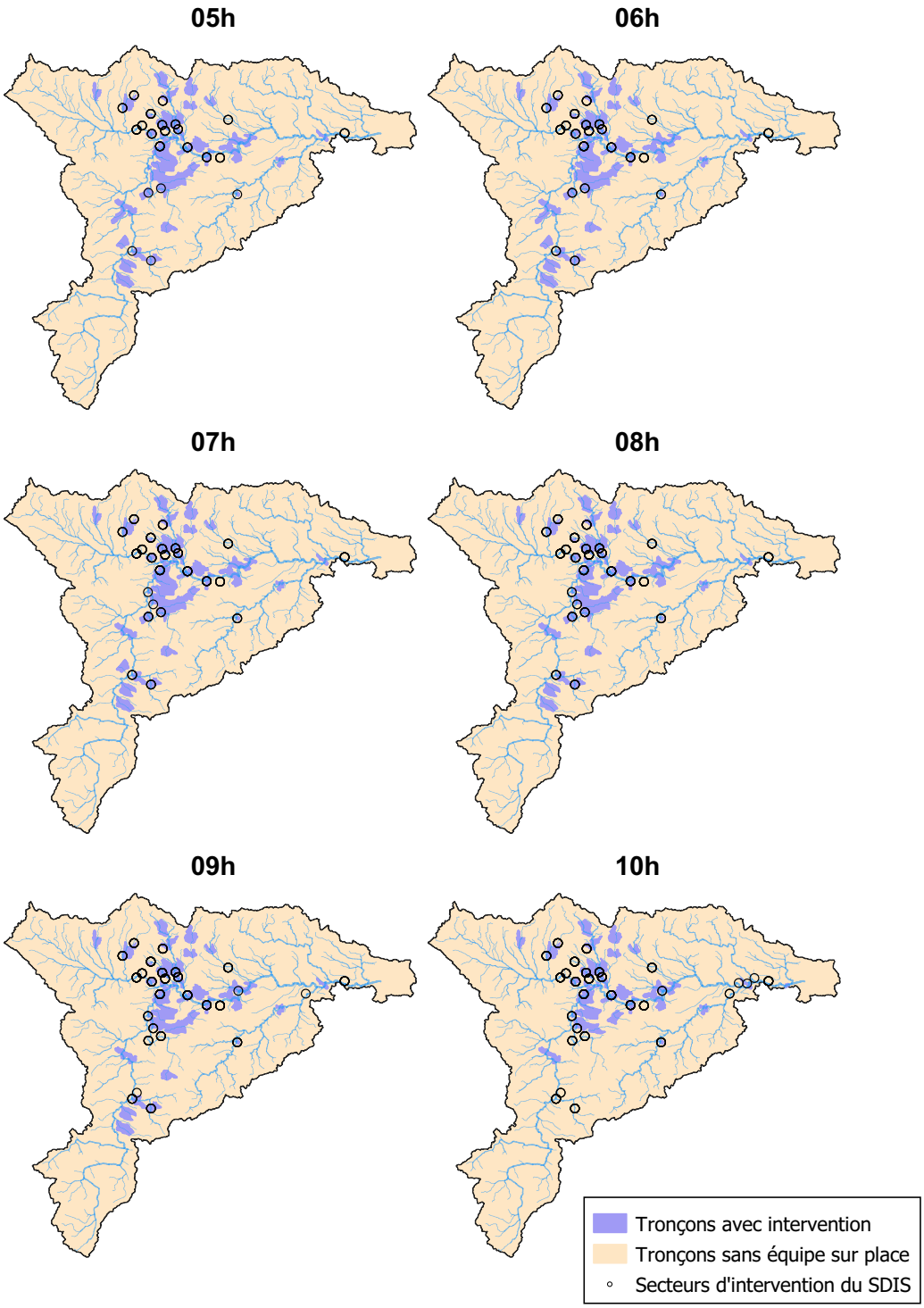

FIGURE 7.16 – *Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 05 h à 10 h*

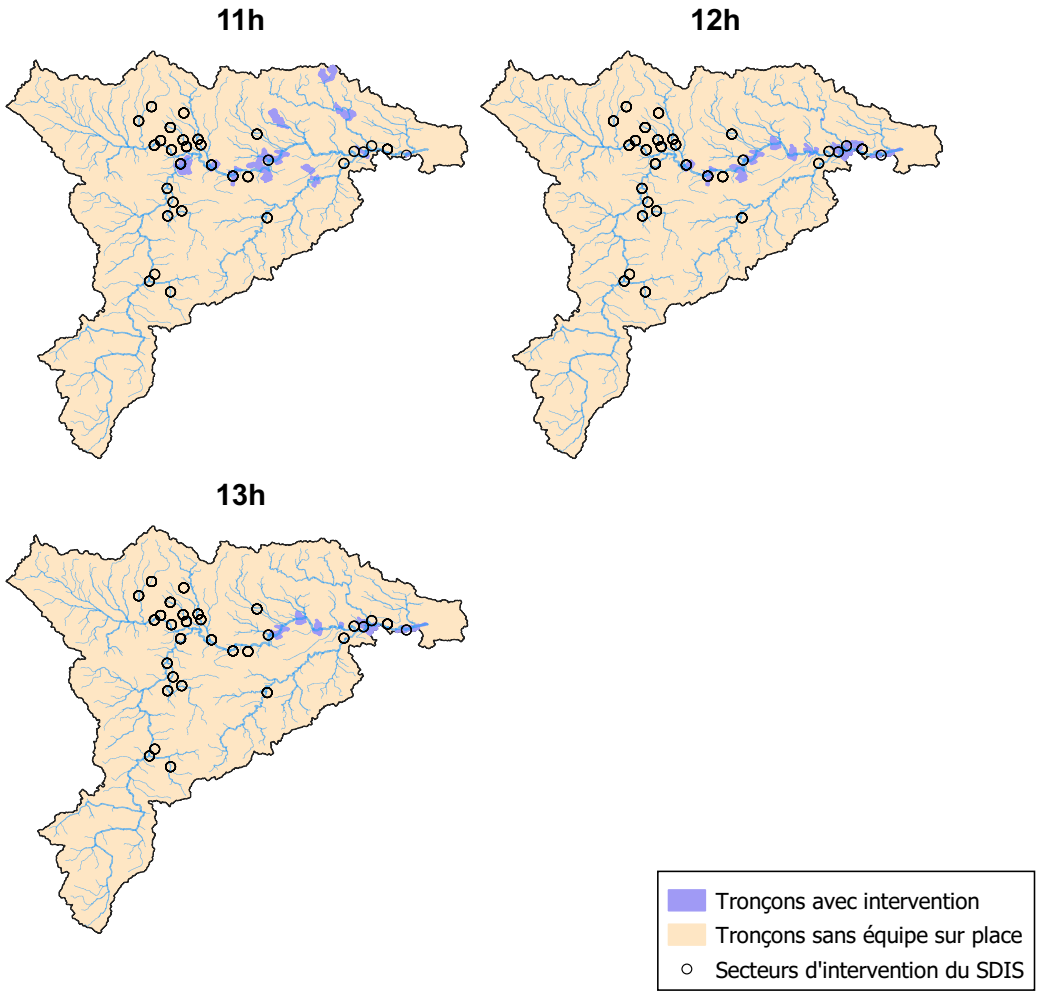

FIGURE 7.17 – *Évolution temporelle des équipes sur les différents tronçons pour l'ensemble pertDpepi et comparaison avec les données du SDIS11 de 11 h à 13 h*

7.4.2 Influence de la valeur du seuil d'intervention

Pour rappel, le seuil d'intervention correspond au nombre de bâtiments dans l'emprise inondée à partir duquel une équipe est diligentée. Ce paramètre influence directement la décision d'envoi des équipes de secours, et donc le nombre d'interventions à réaliser. La Figure 7.18 présente le nombre de bâtiments réellement secourus pour des valeurs de seuils d'intervention de 10, 25 et 50 avec une capacité d'intervention fixée à 100 bâtiments par équipe et un nombre d'équipes fixé à 92.

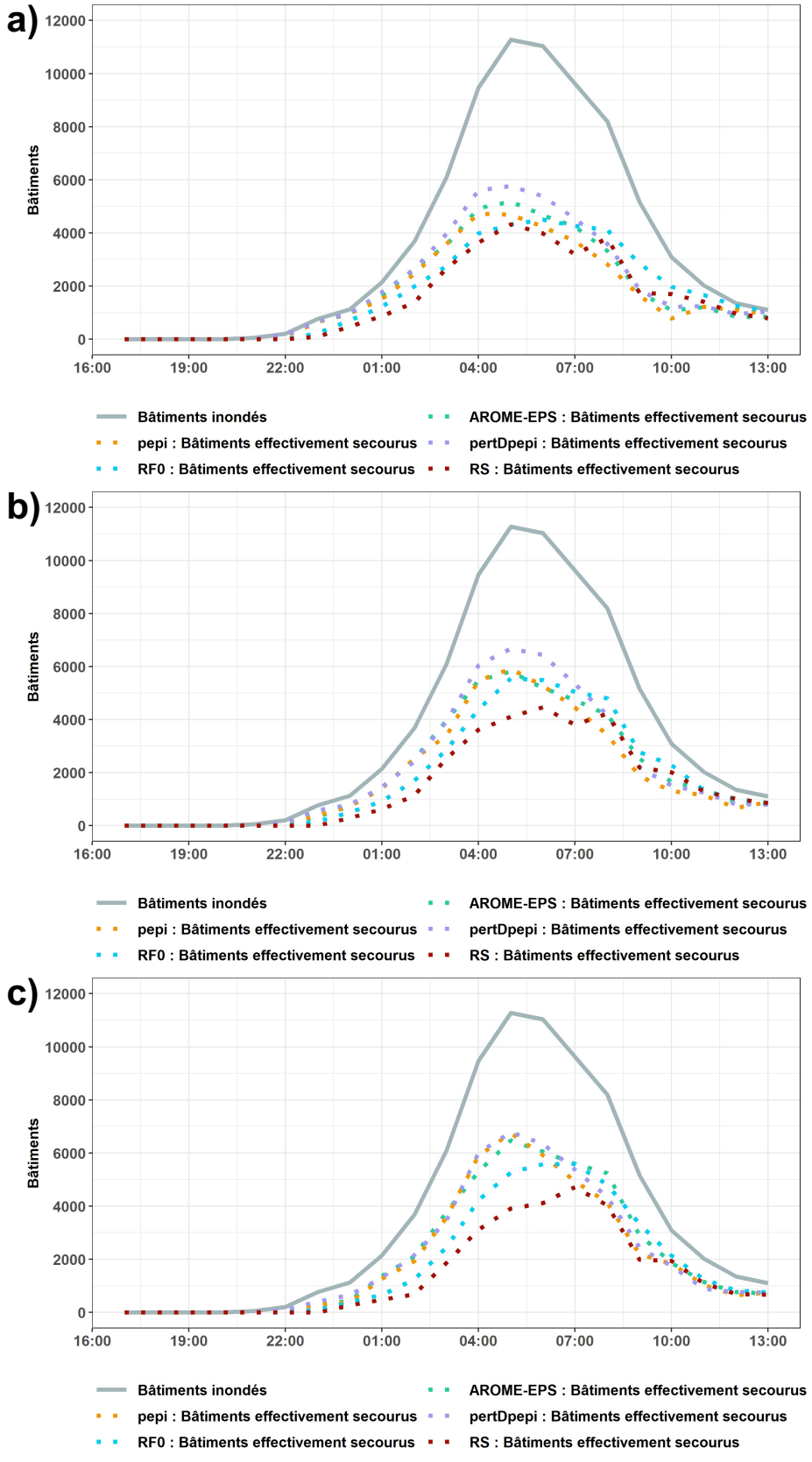

FIGURE 7.18 – *Évolution du nombre de bâtiments réellement secourus avec un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments*

Concernant le nombre de bâtiments effectivement secourus, il ressort que le nombre varie selon le seuil d'intervention fixé. Ainsi, lorsque des prévisions sont utilisées, moins de bâtiments sont secourus (4 227 pour pertDpepi) pour le seuil de 10 bâtiments que pour le seuil de 50 bâtiments (5 887 bâtiments). Ce résultat peut paraître curieux au premier abord, mais est assez logique puisqu'un seuil d'intervention bas implique la mobilisation de capacités d'interventions sur des tronçons présentant peu de bâtiments inondés, ce qui limite les moyens disponibles pour les tronçons plus impactés par la suite. Par ailleurs, un seuil d'intervention bas est susceptible d'augmenter le nombre d'interventions déclenchées suite à des fausses alarmes. Peu de différences apparaissent en revanche entre le seuil d'intervention à 25 bâtiments et 50 bâtiments. Cela peut s'expliquer par le fait que la valeur de 25 bâtiments est déjà suffisamment élevée pour limiter le nombre d'interventions à réaliser. On remarque également que les bâtiments secourus varient moins pour le scénario d'intervention de référence, basé sur la simulation des bâtiments inondés (RS). En l'absence d'anticipation, ce scénario permet en effet de mieux identifier les tronçons sur lesquels les impacts sont les plus importants, tout en évitant de déclencher des interventions suite à des fausses alarmes. Au final, les bâtiments secourus restent globalement plus nombreux lorsque des prévisions sont utilisées, mais se rapprochent du nombre obtenu avec le scénario RS lorsque le seuil d'intervention est bas (10 bâtiments).

La Figure 7.19 présente l'évolution de la mobilisation des équipes de secours pour des valeurs de seuils d'intervention de 10, 25 et 50 avec une capacité d'intervention fixée à 100 bâtiments par équipe et un nombre d'équipes fixé à 92.

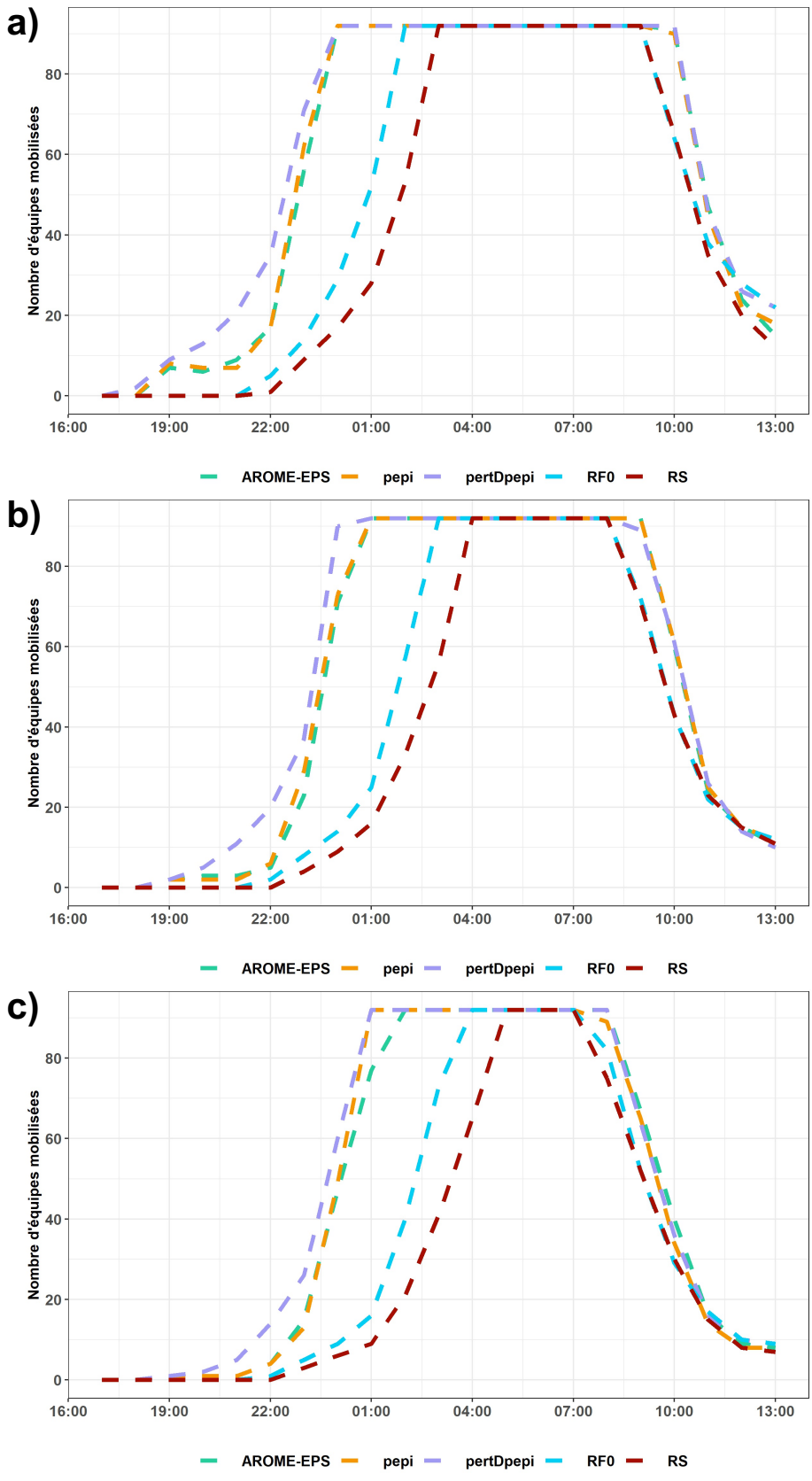

FIGURE 7.19 – *Évolution de la mobilisation des équipes de secours pour un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments*

Dans toutes les modélisations, toutes les équipes (92 équipes) sont très rapidement mobilisées au cours de l'événement. C'est un phénomène qui ne se voyait pas dans les simulations précédentes du fait du nombre illimité d'équipes. La saturation des moyens disponibles se produit plus tôt avec un seuil d'intervention à 10 bâtiments qu'avec un seuil à 50 bâtiments, puisque potentiellement plus d'interventions sont déclenchées sur des tronçons avec peu d'impacts. Ce cas est illustré avec la Figure 7.20, où avec un seuil d'intervention à 10 bâtiments (Figure 7.20 a)) une équipe est envoyée sur place sans réel besoin, alors qu'avec un seuil d'intervention à 50 bâtiments (Figure 7.20 b)), aucune équipe n'est envoyée.

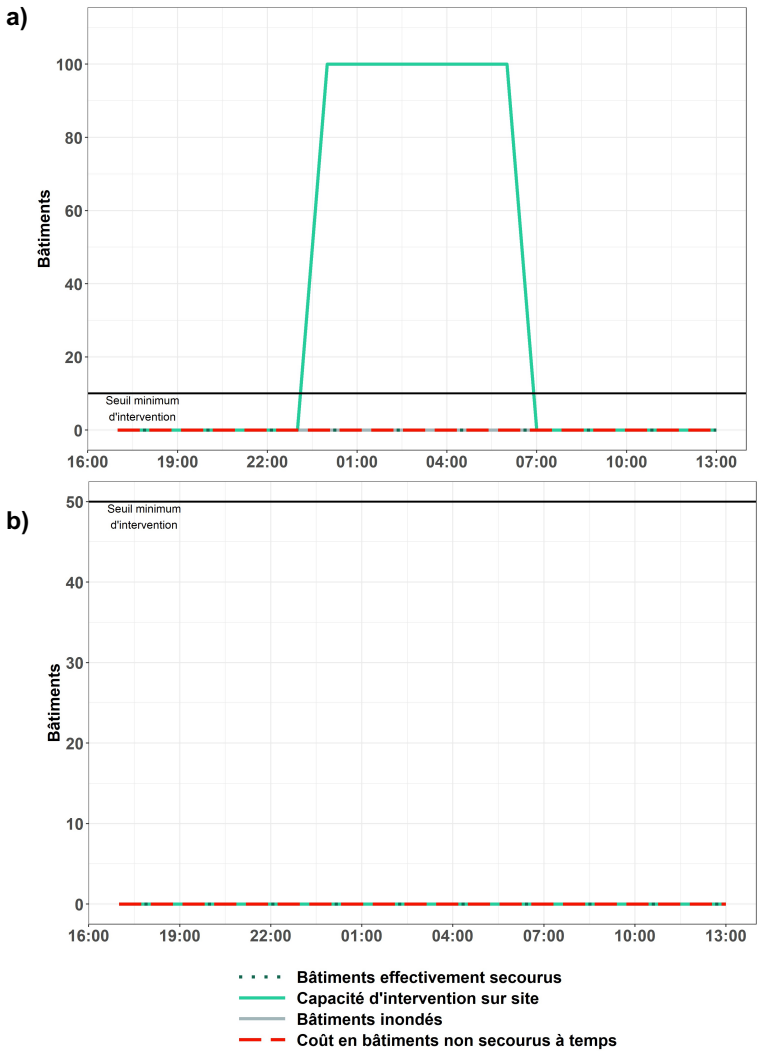

FIGURE 7.20 – *Résultats de modélisation à partir d'AROME-EPS sur le tronçon 2067 avec a) un seuil d'intervention à 10 bâtiments et b) un seuil d'intervention à 50 bâtiments*

Le nombre maximal d'équipes en intervention est plus rapidement atteint pour les prévisions d'ensemble AROME-EPS, pepi et pertDpepi que pour RF0 et RS, observation qui trouve facilement une explication dans le fait que l'anticipation est plus grande, mais également dans le fait des interventions sont déclenchées en raison de fausses alarmes. On peut également remarquer que plus le seuil d'intervention est bas, plus le temps où le nombre maximal d'équipes en intervention est long. En effet, repasser sous le seuil d'intervention est le critère retenu pour démobiliser une équipe et pouvoir éventuellement la réaffecter.

La Figure 7.21 présente l'évolution du coût des bâtiments non secourus pour des valeurs de seuils d'intervention de 10, 25 et 50 avec une capacité d'intervention fixée à 100 bâtiments par équipe et un nombre d'équipes fixé à 92.

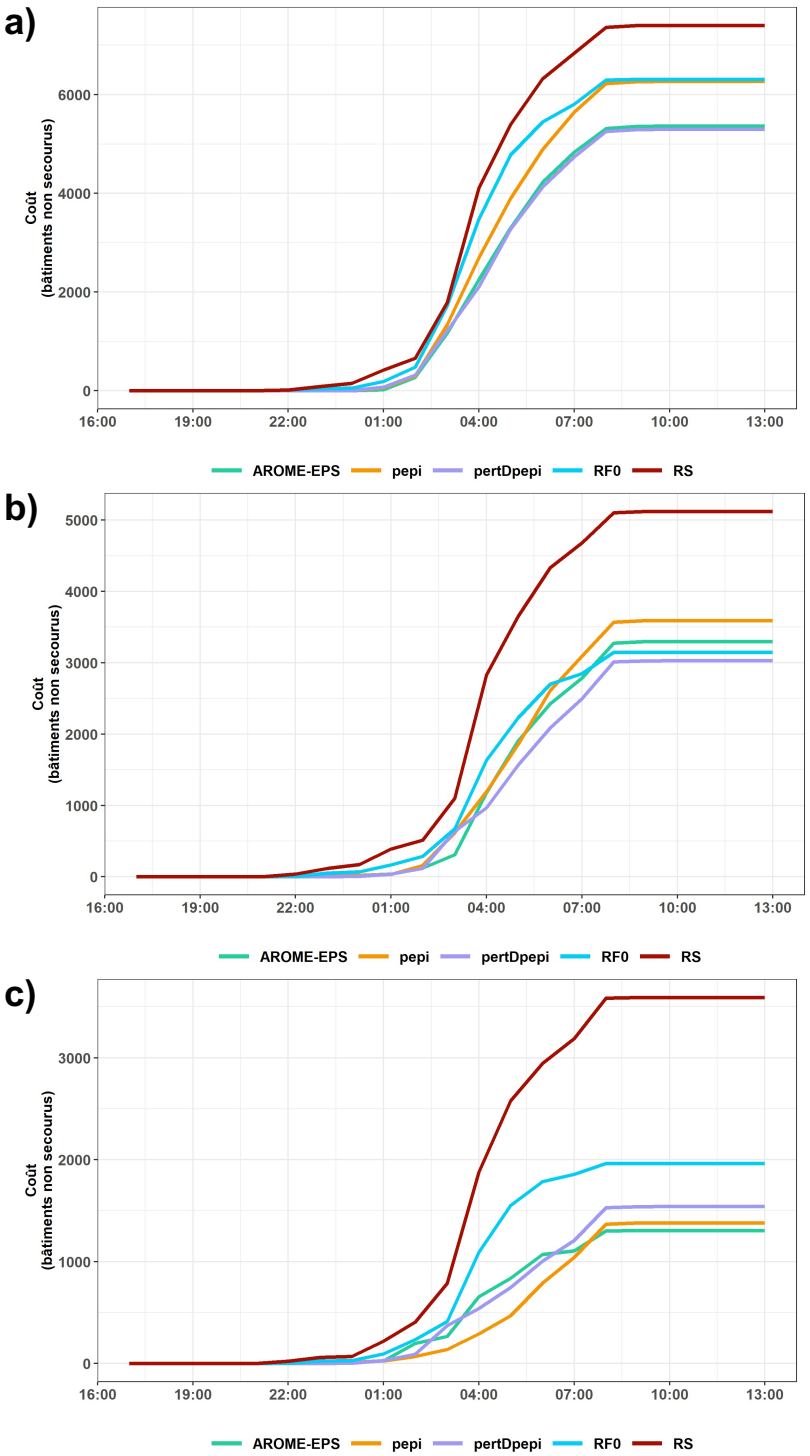

FIGURE 7.21 – *Évolution du coût en bâtiments non secourus pour un seuil d'intervention a) à 10 bâtiments, b) à 25 bâtiments et c) à 50 bâtiments*

Les coûts en bâtiments non secourus ou non secourus à temps permettent de comparer les performances globales de chaque scénario. Les coûts sont assez logiquement plus élevés avec un seuil d'intervention à 10 bâtiments (entre 6 000 et 8 000 bâtiments non secourus selon le scénario) du fait d'une mobilisation moins pertinente des moyens de secours. Ces coûts en bâtiments non secourus sont quasiment divisés par 2 pour un seuil d'intervention de 50 bâtiments (2 500 à 4 500 bâtiments non secourus).

Par ailleurs, la hiérarchie entre les différents scénarios se présente différemment selon le seuil d'intervention considéré. Ainsi, avec un seuil fixé à 10 et 25 bâtiments, pertDpepi a le coût le moins élevé (suivi par AROME-EPS, RF0 et pepi), alors que pour un seuil d'intervention fixé à 50 bâtiments, c'est AROME-EPS qui a le coût en bâtiments équivalent le plus faible. Il faut également noter que la simulation RF0 basée sur un scénario de pluie future nulle présente des résultats très honorables, en particulier par rapport à l'ensemble pepi. Cet ensemble se trouve particulièrement désavantagé dans ce contexte de moyens d'intervention limités, car c'est celui qui lors des précédentes évaluations (Chapitres 5 et 6) a conduit à un plus grand nombre de fausses alarmes.

Pour conclure sur l'influence du seuil d'intervention, ce seuil joue un rôle important sur la performance globale des opérations de secours en cas de moyens de d'intervention très limités. Un seuil trop bas ne permet pas de mobiliser les secours de façon la plus pertinente, alors qu'un seuil haut permet de cibler les secteurs les plus affectés, avec au final des secours plus efficaces. Par ailleurs, les performances relatives des scénarios de décision, basés sur les prévisions ou sur la simulation de référence (RS), sont également modifiées par ce paramètre. Un seuil d'intervention trop bas limite l'intérêt d'un recours aux prévisions, car il mobilise les équipes de façon anticipée sur des secteurs avec peu d'impacts, et augmente par ailleurs l'effet perturbateur des fausses alarmes. Pepi s'avère particulièrement sensible et défavorisé par rapport aux autres produits de prévision du fait de son plus grand nombre de fausses alarmes.

7.4.3 Influence de la capacité d'intervention choisie

Pour rappel, la capacité d'intervention correspond au nombre de bâtiments pouvant être secourus par une équipe sur un même tronçon. Ce paramètre influence directement le nombre d'équipes à envoyer sur les tronçons pour lesquels le nombre de bâtiments inondés est le plus élevé, et donc la saturation plus ou moins rapide des moyens de secours.

La Figure 7.22 présente le nombre de bâtiments réellement secourus pour des valeurs de 10, 50 et 150 bâtiments secourus par équipe, pour un seuil d'intervention fixé à 10 bâtiments par équipe et un nombre d'équipes fixé à 92.

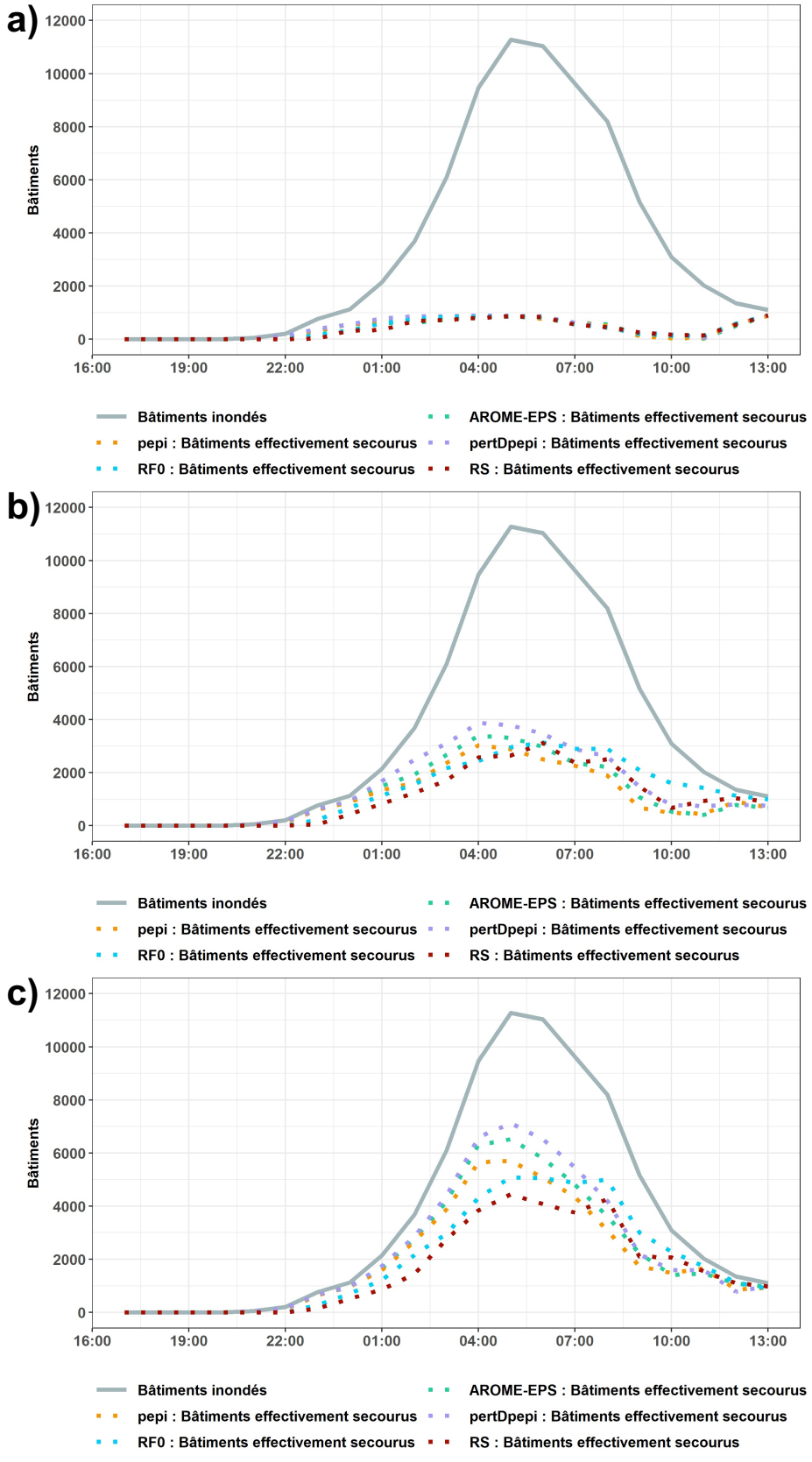

FIGURE 7.22 – *Évolution du nombre de bâtiments effectivement secourus pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments*

Pour ce qui concerne les bâtiments effectivement secourus, l'influence de la capacité d'intervention est nettement visible. Augmenter la capacité d'intervention permet de secourir un plus grand nombre de bâtiments sur un même tronçon, et aboutit logiquement à une augmentation du nombre de bâtiments secourus. Cela permet également de gagner en anticipation en cas d'évolution rapide du nombre de bâtiments inondés, à condition qu'une équipe soit envoyée sur place dès le début des impacts. Cette situation est illustrée sur le tronçon 2967 présenté sur la Figure 7.23.

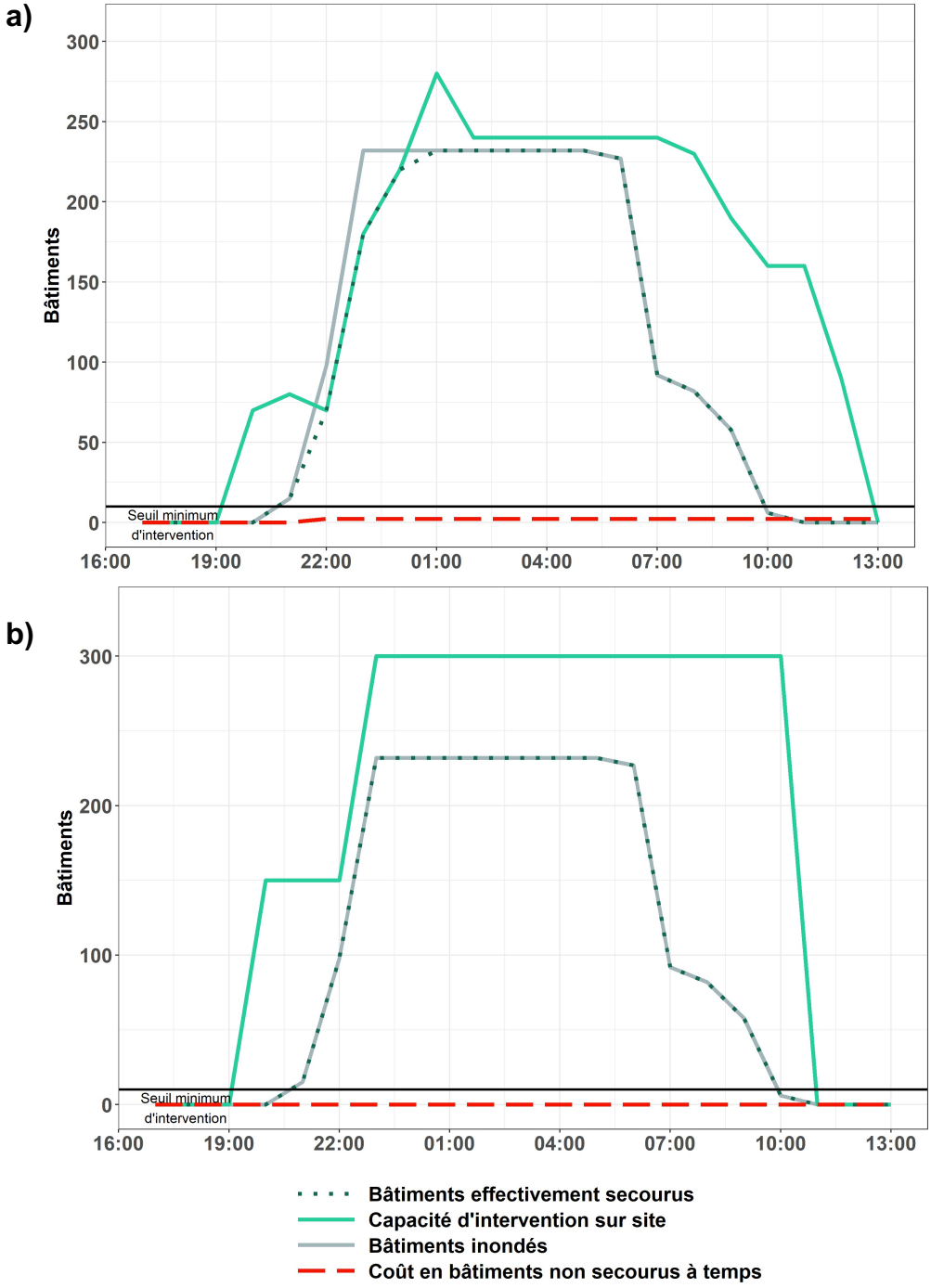

FIGURE 7.23 – *Résultats de modélisation à partir d'AROME-EPS sur le tronçon 2967 avec a) une capacité d'intervention de 10 bâtiments et b) une capacité d'intervention de 150 bâtiments*

Sur ce tronçon, l'influence de la capacité d'intervention est clairement visible avec, dans le cas d'une capacité à 10 bâtiments (Figure 7.23 a)), des envois successifs d'équipes liés à l'augmentation des impacts prévus, entraînant un léger retard de la capacité d'intervention présente sur place par rapport au nombre de bâtiments contenus dans l'emprise inondée de ce tronçon (courbe grise). Au contraire, dans le cas d'une capacité d'intervention fixée à 150 bâtiments par équipe (Figure 7.23 b)), il y a moins d'envois successifs d'équipes et au final une capacité d'intervention plus en accord avec l'évolution du nombre de bâtiments inondés pour AROME-EPS.

La Figure 7.24 présente l'évolution du nombre d'équipes mobilisées pour des capacités d'intervention de 10, 50 et 150 bâtiments, pour un seuil d'intervention fixé à 10 bâtiments par équipe et un nombre d'équipes fixé à 92.

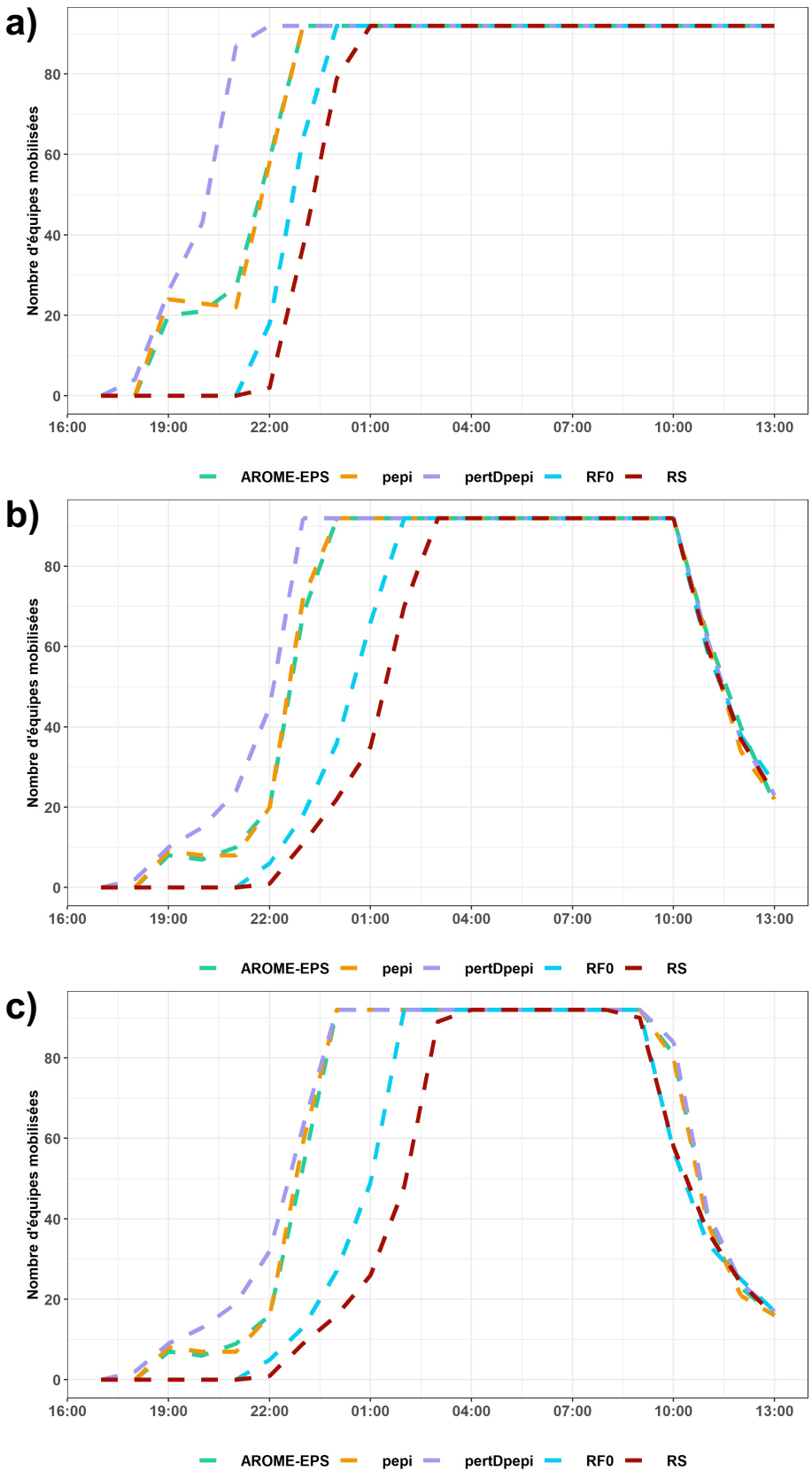

FIGURE 7.24 – *Évolution de la mobilisation des équipes de secours pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments*

Pour ce qui est de l'évolution du nombre d'équipes au cours de l'événement, le constat reste semblable à celui fait dans la Section 7.4.2 pour le seuil d'intervention. Une différence est à noter néanmoins sur la durée pendant laquelle toutes les équipes sont mobilisées sur un tronçon. En effet, dans le cas d'une capacité d'intervention réduite à 10 bâtiments, le nombre maximal d'équipes mobilisées est atteint rapidement et ne redescend pas au cours de l'événement. Cela signifie que les équipes sont en flux tendu, dès qu'une équipe se libère sur un tronçon, elle est immédiatement réaffectée à un autre tronçon. De manière évidente, en augmentant la capacité d'intervention, les équipes ne sont pas forcément toutes réaffectées. Cela signifie probablement qu'il y a moins de 92 tronçons avec des besoins d'intervention en fin d'événement, mais probablement encore quelques tronçons avec un grand nombre de bâtiments touchés.

La Figure 7.25 présente l'évolution du coût des bâtiments non secourus pour des valeurs de capacités d'intervention de 10, 50 et 150 avec un seuil d'intervention fixé à 10 bâtiments par équipe et un nombre d'équipes fixé à 92.

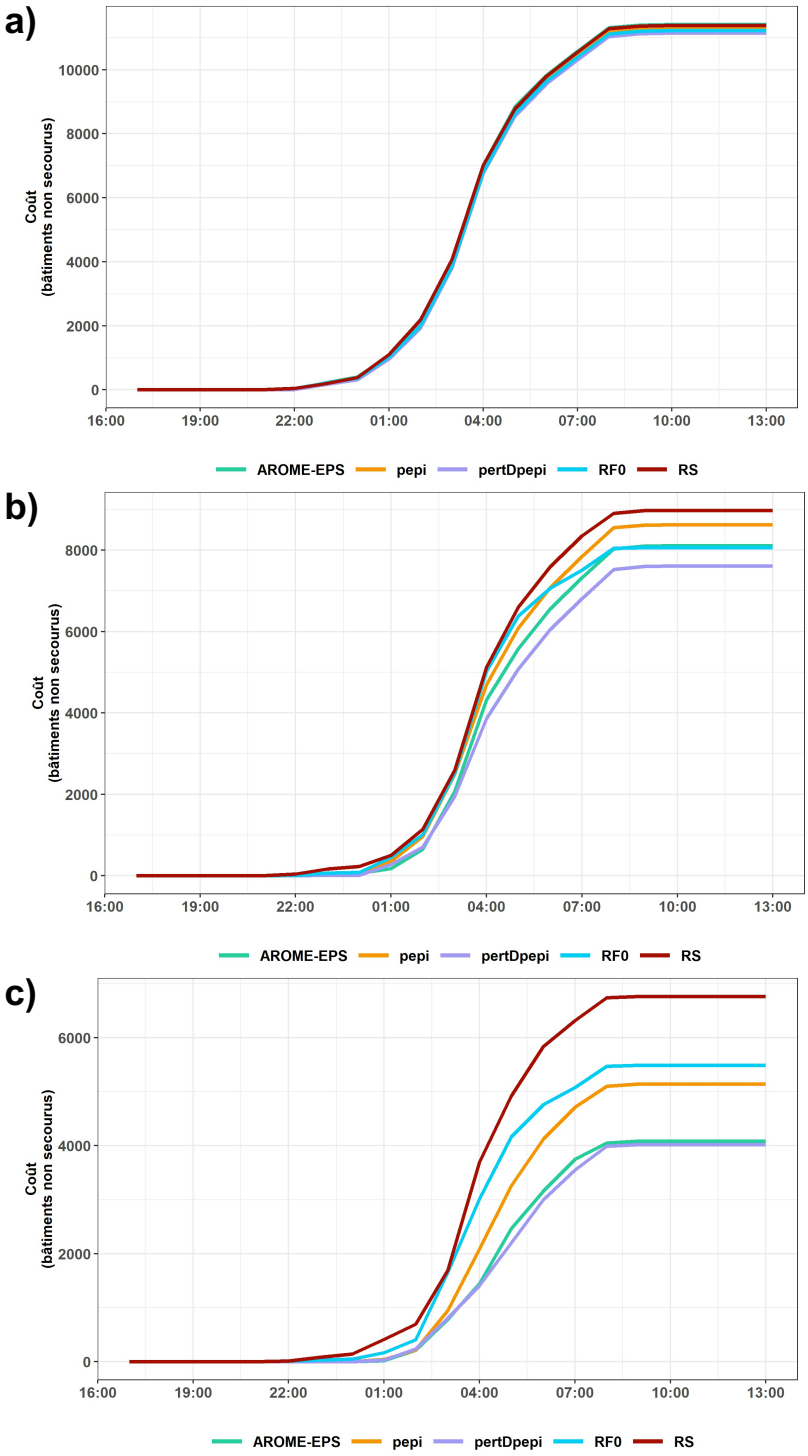

FIGURE 7.25 – *Évolution du coût en bâtiments non secourus pour une capacité d'intervention a) de 10 bâtiments, b) de 50 bâtiments et c) de 150 bâtiments*

Concernant le coût en nombre de bâtiments non secourus, on peut en premier lieu constater qu'avec une capacité d'intervention fixée à 10 bâtiments, le coût en bâtiments non secourus s'avère très élevé, alors qu'avec 50 ou 150 bâtiments, il est nettement réduit. Ceci traduit le fait qu'avec un seuil d'intervention à 10 bâtiments, les moyens d'intervention disponibles sont très inférieurs aux besoins. Il faut également noter que dans le cas de la capacité d'intervention fixée à 10 bâtiments, il y a très peu de différence de coût entre les différents scénarios

(prévisions ou référence simulée RS). Dans cette situation, les moyens d'intervention disponibles s'avèrent très fortement limitants, et ne permettent plus aux prévisions d'apporter une plus-value significative.

7.4.4 Influence du nombre d'équipes mobilisables

Le troisième paramètre influant fortement sur la performance des secours est le nombre d'équipes d'intervention disponibles. Il peut notamment être intéressant de voir comment évoluent les performances en considérant des équipes "plus petites", mais plus nombreuses. Par équipes "plus petites" est entendu une capacité d'intervention plus faible.

Les Figures 7.26, 7.27 et 7.28 présentent les résultats obtenus avec un nombre d'équipes multiplié par 2 en nombre de bâtiments effectivement secourus, nombre d'équipes mobilisées et coût des bâtiments non secourus à temps.

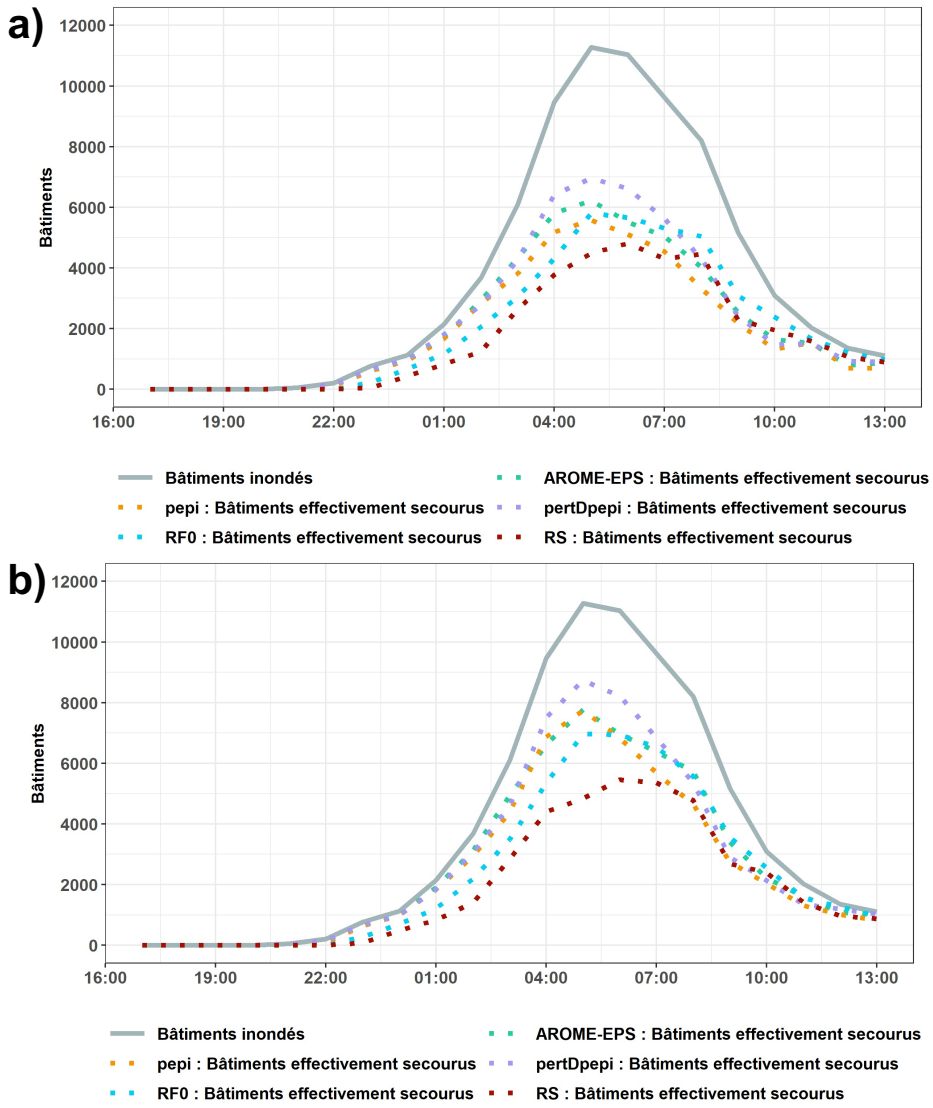

FIGURE 7.26 – *Évolution du nombre de bâtiments effectivement secourus pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe*

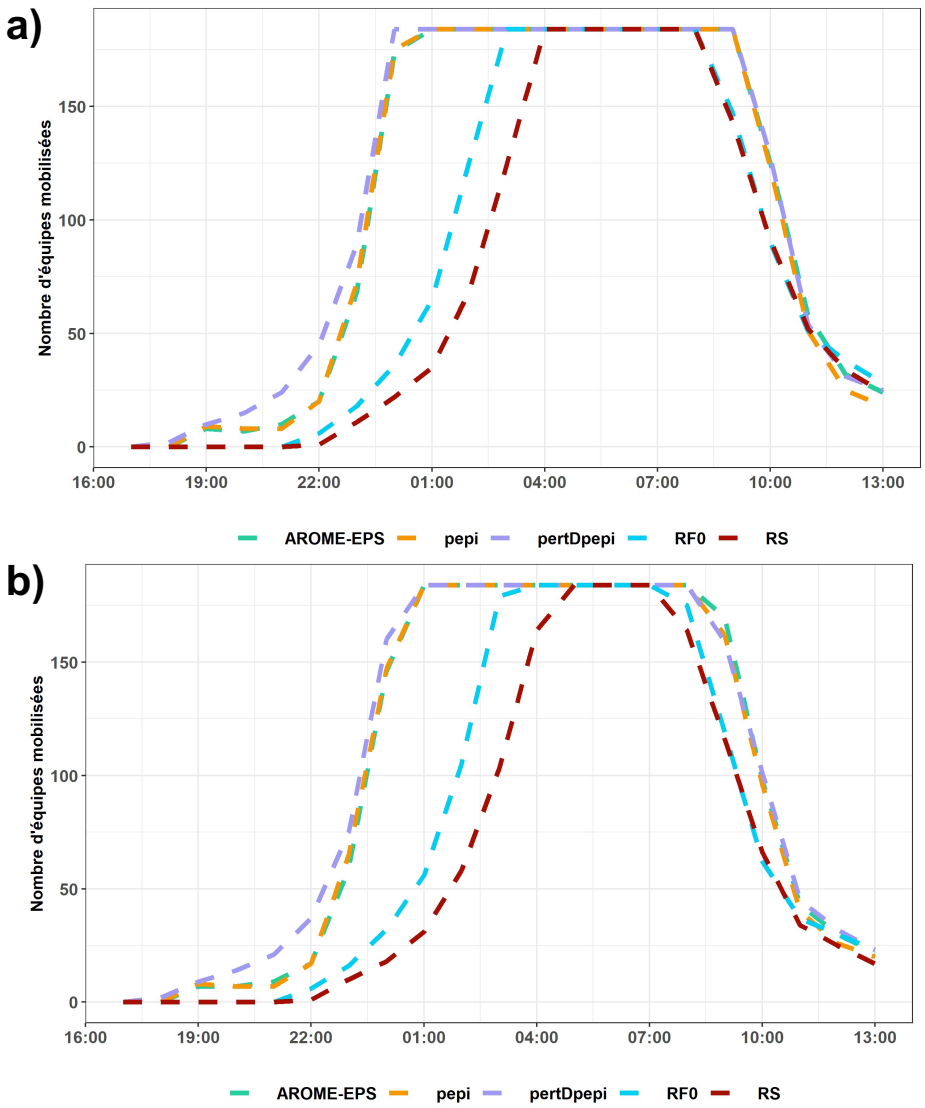

FIGURE 7.27 – *Évolution du nombre d'équipes mobilisées pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe*

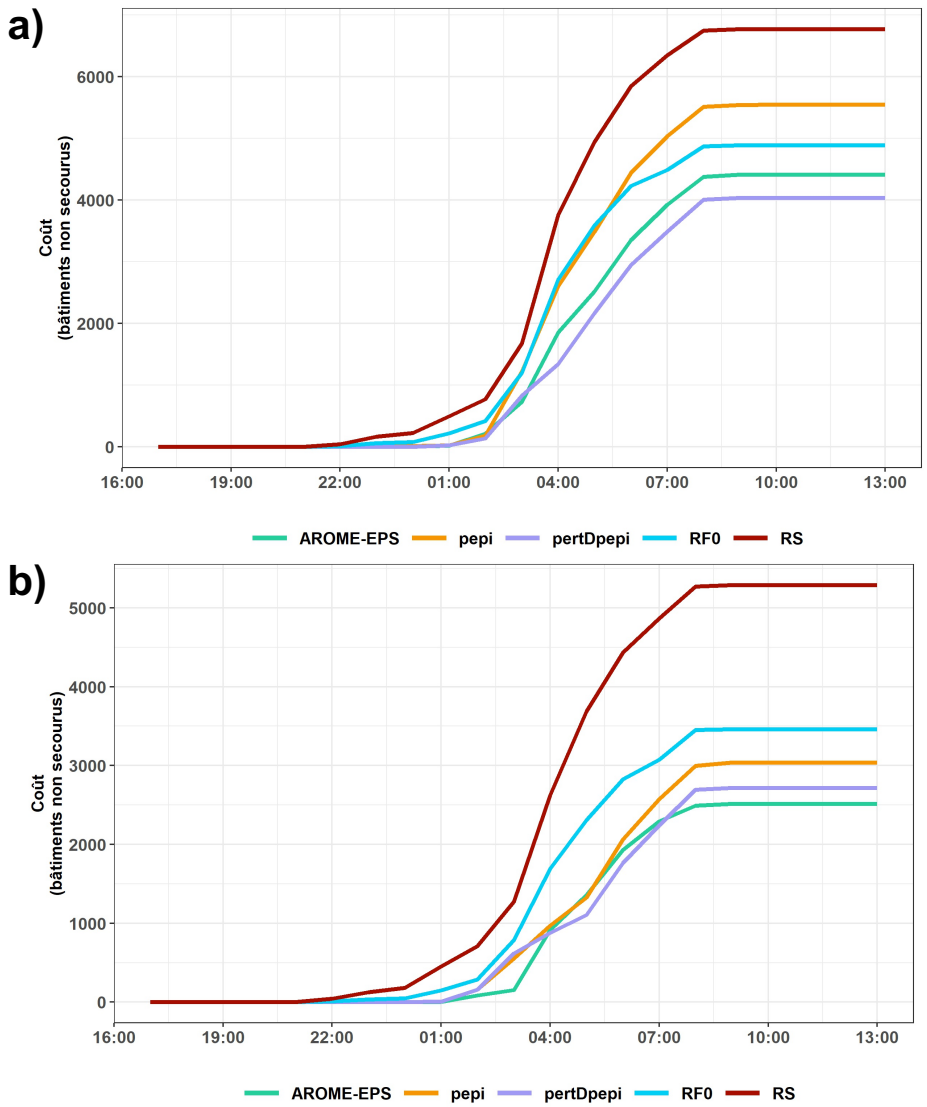

FIGURE 7.28 – *Coût des bâtiments non secourus à temps pour a) 184 équipes avec une capacité d'intervention de 50 bâtiments par équipe et b) 184 équipes avec une capacité d'intervention de 75 bâtiments par équipe*

Ces Figures peuvent être directement comparées avec les Figures 7.18 a), 7.19 a), 7.21 a) pour le nombre d'équipes fixé à 184 avec une capacité d'intervention de 50 bâtiments et avec les Figures 7.22 c), 7.24 c), 7.25 c) pour le nombre d'équipes fixé à 184 avec une capacité d'intervention de 75 bâtiments.

Les résultats présentés sur ces Figures correspondent à la même capacité d'intervention globale, en considérant le nombre d'équipes et la capacité d'intervention associée à chaque équipe.

Le constat est que les modélisations avec un nombre d'équipes doublé (et une capacité d'intervention réduite) permettent d'augmenter le nombre de bâtiments secourus, et que ces modélisations s'avèrent au final plus performantes en termes de coût en bâtiments non secourus. Cet effet se retrouve pour tous les scénarios testés, et tout particulièrement pour les trois produits de prévision d'ensemble. Cela rejoint l'observation faite lorsque l'événement était rejoué avec des moyens illimités (Section 7.3). Les trois produits de prévision d'ensemble

sont en effet fortement pénalisés par des fausses alarmes, et donc des envois d'équipes sur des tronçons sur lesquels très peu de bâtiments s'avèrent finalement inondés. L'augmentation du nombre d'équipes permet donc de limiter cet effet négatif, en conservant des équipes disponibles pour intervenir sur d'autres tronçons. Cela est particulièrement visible en sur la fonction de coût (Figures 7.25 c) et 7.28 a)) où l'on constate une forte réduction des coûts en bâtiments non secourus.

7.5 Bilan sur l'utilisation des impacts pour les prises de décision d'un service de secours

Les prévisions d'impacts ont été évaluées en se plaçant du point de vue d'un utilisateur potentiel de ces prévisions, ici un SDIS.

7.5.1 Adaptabilité de la méthode à d'autres cas d'études et d'autres utilisateurs

Le cadre de modélisation a l'avantage de pouvoir être transposé à n'importe quel territoire métropolitain. En effet, le réseau routier est disponible grâce à la BD TOPO et la vitesse choisie est une moyenne. Cette vitesse moyenne pourrait d'ailleurs être améliorée en utilisant des vitesses moyennes par type de route ou encore en tenant compte des tronçons où des interventions sont nécessaires qui seraient alors à contourner ou bien à traverser à vitesse réduite. Cette approche comporte aussi l'avantage de dépasser les limitations des évaluations fondées sur la détection des dépassements anticipés de seuils, détection unique mise en œuvre dans les Chapitres 5 et 6.

Pour ce qui est de l'adaptabilité à un autre utilisateur, cette question dépend entièrement des besoins en termes de prévisions de ce dernier, mais également de ses moyens d'intervention. Cette partie de la modélisation est donc à modifier au cas par cas pour chaque utilisateur. D'autres choix d'impacts pourraient également être fait selon le but recherché par l'utilisateur comme l'utilisation de la densité de population.

7.5.2 Questionnement sur le choix de la fonction de coût

L'efficacité des mesures d'intervention est calculée comme un pourcentage du nombre de bâtiments inondés non secourus (description faite dans la Section 7.2.4). La simulation en moyens illimités (Section 7.3) a montré des résultats particulièrement performants pour les trois produits de prévisions d'ensemble par rapport au scénario de pluie future nulle RF0. Ce résultat peut être nuancé par le fait que dans les cas d'interventions avec des moyens illimités, la fonction de coût ne pénalise pas directement l'envoi d'équipe sur des tronçons sans besoin. Il s'agit d'un cas où les ensembles se révèlent être très performants, car les fausses alarmes ne sont pas prises en compte. Néanmoins, en considérant des moyens d'intervention limités, le poids des fausses interventions est indirectement pris en compte dans le manque de moyens pour les tronçons nécessitant une intervention. Ajouter un coût pour l'envoi à

mauvais escient aurait donc eu un effet double peine sur les simulations avec des moyens limités : un coût aurait été compté pour les équipes envoyées sur les mauvais tronçons, mais un autre coût aurait été appliqué sur les tronçons avec des bâtiments non secourus. Le choix de cette fonction peu pénalisante pour la mobilisation non nécessaire joue en défaveur du scénario de pluie future nulle RF0 en cas de moyens illimités, mais ce scénario montre un réel intérêt en cas de moyens d'intervention limités par rapport aux prévisions d'ensemble qui deviennent alors fortement pénalisées par les interventions non nécessaires.

Utilisation des prévisions d'impacts dans un contexte s'approchant de l'opérationnel

Le Chapitre 7 était dédié à la mise en place d'un rejeu des crues-éclair de l'Aude d'octobre 2018, en mettant en œuvre une modélisation des interventions de secours déclenchées par le SDIS de l'Aude. Dans cette modélisation, les décisions d'intervention se sont basées soit sur les simulations et/ou les prévisions d'ensemble des bâtiments inondés évaluées dans le Chapitre 6. Cette modélisation, sur des hypothèses assez simples, mais réalistes, avait pour but de pouvoir évaluer l'intérêt des trois produits de prévision d'ensemble pour une prise de décision, en tenant compte des effets positifs d'une meilleure anticipation, mais également des effets négatifs associés aux fausses alarmes. Comme dans les chapitres précédents, l'évaluation a été effectuée par comparaison à un scénario de référence dans lequel les décisions étaient prises sur la base de la simulation des bâtiments inondés (donc sans anticipation), ce qui se rapproche du fonctionnement actuel d'interventions déclenchées suite à des remontées d'informations de terrain. Afin de comparer les performances des différents scénarios considérés, une fonction de coût a été mise en place basée sur un décompte des bâtiments non secourus à temps.

Un apport des trois produits de prévision d'ensemble est clairement visible, en particulier dans le cas d'une modélisation avec des moyens d'interventions illimités (fausses alarmes non pénalisantes), mais également dans le cas d'un nombre d'équipes limité et de paramétrages assez proches de la situation réelle vécue en 2018. L'analyse de sensibilité réalisée a également permis de simuler des situations avec des moyens d'intervention encore plus limités, qui permettent d'illustrer les limites liées aux fausses alarmes, entraînant des envois d'équipes sur des tronçons non impactés. En cas de moyens extrêmement limités, les décisions basées sur le scénario de pluie future nulle RF0 apparaissent parfois meilleures que celles utilisant des prévisions d'ensemble issues d'AROME-EPS, pepi et pertDpepi. Néanmoins, dans l'ensemble des cas simulés, les décisions prises à partir des prévisions d'ensemble ne sont jamais apparues moins efficaces que celles basées sur la simulation de référence (sans anticipation).

Les différences de performances entre les trois produits de prévision d'ensemble sont dépendantes des valeurs associées aux paramètres de modélisation (seuil d'intervention, capacité d'intervention et limitation des moyens), mais pertDpepi est la prévision qui globalement a conduit aux meilleures décisions, en raison du nombre de fausses alarmes plus limité.

Même si la modélisation réalisée se veut déjà relativement réaliste, plusieurs améliorations pourraient encore être apportées. L'utilisation des bâtiments issus

de la BD TOPO comme approximation des impacts potentiels semble assez pertinente puisque les secteurs d'intervention du SDIS sont presque tous retrouvés sur la simulation de l'événement. La fonction coût par exemple pourrait être adaptée pour donner une plus grande valeur à l'intervention anticipée des secours. Il pourrait également être intéressant de pondérer le nombre de bâtiments inondés par un facteur permettant de représenter la densité de population. En effet, lors de l'intervention sur des bâtiments inondés, c'est surtout la mise en sécurité des personnes associées à ces bâtiments qui est recherchée par les services de secours. Un travail pourrait également être mené sur la valeur réelle de la capacité d'intervention, et il serait aussi envisageable de privilégier le nombre de tronçons couverts par des interventions, plutôt que de prioriser les tronçons sur lesquels le nombre de bâtiments inondés est le plus élevé. Enfin, la matrice de déplacement pourrait être plus réaliste, celle utilisée ici restant très simplifiée, avec une vitesse moyenne appliquée sur tout le territoire.

Chapitre 8

Conclusions et perspectives

8.1 Bilan du travail de thèse

Le développement des systèmes de prévision des crues-éclair apparaît alors comme un enjeu majeur des prochaines années, pour répondre à la probable augmentation de l'occurrence de ces phénomènes avec le changement climatique. Les échelles spatiales et temporelles dans lesquelles s'inscrivent ces crues sont également un défi supplémentaire pour la mise en place d'alertes et d'avertissements. En France, si la mise en place du service Vigicrues Flash en 2017 a permis d'apporter une première réponse à ces défis. Mais ce système reste largement perfectible, en raison notamment de sa couverture spatiale encore limitée, des niveaux d'anticipation encore faibles, et de l'absence de représentation des enjeux vulnérables.

Cette thèse a été réalisée en lien étroit avec le projet de recherche ANR-PICS (Prévision Immédiate Intégrée des Impacts des Crues Soudaines), débuté en 2018 et dont l'objectif était de contribuer à des améliorations futures du service Vigicrues Flash. Ce projet s'est notamment intéressé à l'utilisation de prévisions des précipitations intenses et au développement de la prévision des impacts. Obtenir des informations directes sur les champs d'inondation et sur les impacts apparaît en effet comme essentiel pour des utilisateurs des prévisions comme les gestionnaires de crise, les gestionnaires de réseaux ou encore les sociétés d'assurances. Cette thèse a permis de combiner les outils de prévision immédiate d'ensemble des pluies, de modélisation pluie-débit et de modélisation des impacts, proposés dans le cadre du projet PICS, de façon à appliquer une chaîne complète de prévision des impacts des crues-éclair, qui a été mise en œuvre et testée sur des événements de 2018 dans l'Aude, l'Ardèche et le Var.

Cette thèse a notamment permis d'utiliser, en entrée de la chaîne de prévision, trois produits de prévision immédiate d'ensemble (échéance 6 h), dont deux sont expérimentaux et ont été proposés par le CNRM et Météo-France dans le cadre du projet PICS :

- **AROME-EPS** est composé de 12 membres. Il s'agit du seul produit de prévision immédiate d'ensemble opérationnel à Météo-France. AROME-EPS est une version ensembliste du modèle AROME-FR.

- **pepi** est produit en combinant AROME-EPS et AROME-PI, qui est une version déterministe d'AROME dédiée à la prévision immédiate (échéance 6h). AROME-PI est intégré à pepi en utilisant les six dernières simulations disponibles (une simulation effectuée toutes les heures). Pepi est donc composé de 13 à 18 membres selon l'échéance considérée du fait de l'ajout des simulations successives d'AROME-PI.

- **pertDpepi** est basé sur l'ensemble pepi, pour lequel chacun des membres a été perturbé spatialement dans les quatre directions cardinales avec un décalage de 20 km par rapport au membre initial de pepi. Le nombre de membres est donc multiplié par 5 par rapport à pepi.

La chaîne de prévision appliquée est proche dans son principe de celle appliquée par Le Bihan (2016). Elle combine un modèle hydrologique pluie-débit, et des courbes débit-impacts établies sur chaque tronçon de cours d'eau à partir d'un catalogue de scénarios d'inondation. Le modèle hydrologique Cinecar a été utilisé, non seulement pour sa capacité

vérifiée à reproduire les crues-éclair méditerranéennes, mais également, car c'est un modèle qui demande peu de paramètres et qui présente des temps de calculs courts. Les courbes débit-impacts ont été obtenues à partir d'un catalogue d'emprises inondées de référence, pour des périodes de retour de débit de 2 à 1 000 ans, réalisé dans le cadre du projet ANR-PICS avec le modèle hydraulique Floodos.

La première partie du travail de thèse a consisté à mettre en place une méthodologie d'évaluation des prévisions d'ensemble des débits, adaptée à l'évaluation sur un événement de crue unique. Préalablement à cette évaluation hydrologique, les prévisions d'ensemble de pluie ont également été évaluées (Chapitre 4) en combinant des outils de visualisation (hyétogrammes moyens horaires et répartition spatiale des cumuls sur l'événement) et des outils classiques de l'évaluation de la prévision d'ensemble (diagramme de rang et CRPS). L'enjeu était notamment de dégager un cadre spatio-temporel adapté à l'évaluation des prévisions hydrologiques pour chacun des événements étudiés. L'évaluation des prévisions de débits (Chapitre 5) a ensuite été mise en œuvre, en se focalisant sur l'anticipation des dépassements de seuils de débits sur l'ensemble du secteur d'étude (exutoire jaugé ou non jaugé).

Cette méthodologie définie pour l'évaluation des débits a aussi été mise en œuvre pour l'évaluation des prévisions d'impacts (Chapitre 6), pour l'événement le plus important de l'année 2018, celui du 15 octobre dans l'Aude. Ceci a permis d'étudier la façon dont une évaluation à partir des impacts peut modifier le regard porté sur les performances d'un modèle de prévision des débits.

Enfin, une dernière évaluation a été mise en œuvre, basée sur le rejeu des décisions d'intervention d'un service d'incendie et de secours. Ce rejeu a été réalisé à partir des prévisions d'impacts pour l'événement de l'Aude (Chapitre 7). Cette dernière évaluation a permis d'évaluer l'apport des prévisions d'ensemble dans un cadre de prise de décisions pour la gestion de crise.

Les Sections suivantes reprendront les différentes conclusions de ces évaluations successives, des prévisions de pluie au rejeu d'un événement du point de vue d'un service de secours. Dans un second temps, les perspectives d'améliorations des méthodes mises en œuvre pour des études futures seront présentées dans la Section 8.3.

8.2 Conclusions sur les résultats obtenus

8.2.1 Performances des trois produits de prévision d'ensemble testés

Les produits de prévisions d'ensemble de pluie ont été évalués avec plusieurs visions et sous plusieurs formes :

- Une évaluation visuelle et événementielle à partir des répartitions spatiales (cumul moyen) et temporelles (hyétogramme moyen horaire) des prévisions d'ensemble de précipitations par rapport à la lame d'eau observée. Cette évaluation est combinée avec des calculs de diagrammes de rang et de CRPS pour les précipitations d'ensemble de pluie.

- Une évaluation par événement de l'anticipation des dépassements de seuil de débit et de seuil d'impacts obtenue avec les prévisions d'ensemble, en utilisant une méthodologie inspirée des courbes ROC.
- Une évaluation du point de vue de l'efficacité des prises de décision pour les déclenchements d'interventions d'un SDIS.

Les évaluations révèlent que les pluies sont très bien anticipées dans l'ensemble par les produits de prévision testés, avec néanmoins une difficulté pour saisir les plus fortes intensités observées (sous-estimation des précipitations). Un décalage spatial est également observé sur les cumuls pluviométriques pour le cas d'étude du Var. Ce décalage s'avère particulièrement pénalisant pour les prévisions hydrologiques. Les prévisions hydrologiques bénéficiant de prévisions immédiates de pluie permettent d'augmenter les délais d'anticipation par rapport à une situation de prévision utilisant une pluie future nulle. Comme attendu après les résultats de l'évaluation sur les précipitations, le décalage spatial des précipitations, ainsi que la légère sous-estimation des précipitations se répercutent dans les conclusions de l'évaluation menée sur l'anticipation des débits.

Concernant la comparaison des performances des trois produits de prévision d'ensemble des pluies, le constat est que la différence porte principalement sur la représentation des précipitations intenses. Ainsi pepi apparaît plus dispersé que l'ensemble AROME-EPS et atteint des valeurs de pluies plus fortes. Pour ce qui est de pertDpepi, l'ajout des perturbations spatiales apparaît plus comme un moyen de lisser spatialement les prévisions. Ce lissage influence toutes les évaluations réalisées. Les produits pepi et pertDpepi montrent de meilleurs résultats pour les évaluations basées sur les pluies et les débits que le produit AROME-EPS. Les conclusions ont été similaires pour l'évaluation des prévisions d'impacts (événement de l'Aude).

L'utilisation d'un scénario de pluie future nulle (appelé RF0) a permis de mesurer la plus-value des trois produits de prévision par rapport à une prévision de référence reposant uniquement sur les observations de pluie. Cette plus-value en termes d'anticipation a pu être confirmée pour les trois produits de prévision évalués.

Enfin, pour ce qui est du rejeu des décisions du SDIS de l'Aude, les performances des ensembles dépendent fortement des paramètres de modélisation choisis : seuil d'intervention, capacité d'intervention ou encore moyens mobilisables. La Section 7.3 présentant des résultats de simulation avec des moyens d'intervention illimités a démontré le net intérêt de l'utilisation des prévisions d'ensemble de pluie pour bâtir une décision d'intervention. En effet, dans ce cadre de moyens illimités, l'envoi d'équipes sur des zones sans impacts n'est pas pénalisant. La fonction de coût choisie ne rend compte que du coût de la non-intervention sur les tronçons ayant des besoins. Dans le cas des moyens illimités, le coût de ces fausses interventions n'est pas pris en compte, cela explique les résultats très intéressants obtenus avec cette simulation. Il aurait pu être envisageable de mettre en place une fonction de coût prenant en compte la mobilisation à mauvais escient des moyens de secours. Néanmoins, le but étant de définir une fonction de coût unique pour l'ensemble des simulations, en moyens illimités et en moyens limités, l'utilisation d'une fonction prenant en compte ces mauvaises mobilisations aurait eu

un effet double peine. Non seulement les tronçons sans intervention, par manque de moyens, auraient été pénalisés, mais les tronçons avec une mobilisation non nécessaires l'auraient été aussi. Le choix de cette fonction peu pénalisante pour la mobilisation non nécessaire explique donc la très bonne performance des ensembles de prévision de pluies dans le cas de moyens illimités par rapport au scénario RF0. Ce scénario de pluie future nulle devient néanmoins intéressant en moyens limités, les équipes étant envoyées seulement sur des secteurs touchés.

Pour un même jeu de paramètres (seuil d'intervention, capacité d'intervention et moyens mobilisables), pertDpepi montre de meilleurs résultats en termes d'efficacité des prises de décision, par comparaison à AROME-EPS ou pepi. Cela s'explique par le taux de fausses alertes plus réduit lorsque les prévisions pertDpepi sont utilisées. Néanmoins, concernant les délais d'anticipation (et plus seulement l'anticipation), ceux-ci apparaissent légèrement réduits pour pertDpepi en comparaison avec pepi et AROME-EPS et cela quelle que soit l'évaluation considérée, débits ou impacts.

8.2.2 Intérêts et limites de la méthodologie d'évaluation des prévisions de pluie et débits

La raison principale qui a conduit à la mise en place d'une méthodologie spécifique est le besoin d'évaluer les prévisions sur de rares événements pluvieux intenses, spatialement peu étendus et relativement courts. De façon à offrir une évaluation aussi complète que possible et basée sur un jeu de données de taille suffisante, l'évaluation ne se focalise pas uniquement sur les quelques exutoires jaugés présents sur les zones sélectionnées, mais sur un ensemble de points quadrillant le territoire (exutoires des petits sous bassins non jaugés). Quelques nuances sont à apporter concernant les choix de modélisation. Les incertitudes associées à la modélisation pluie-débit ne sont pas considérées, car la comparaison est directement faite avec les débits simulés à partir des pluies observées. D'autres travaux ont été menés (Peredo-Ramirez 2021) à partir de ces produits de prévision d'ensemble, et en particulier des comparaisons aux débits de stations hydrométriques, montrant des résultats satisfaisants pour ces ensembles.

Pour ce qui concerne l'évaluation des prévisions de pluies (étape 1), la comparaison observation/prévision apparaît particulièrement efficace comme première approche pour repérer les zones de décalage des précipitations comme cela est le cas pour le Var, mais également pour mettre en évidence la difficulté des ensembles à prévoir les fortes intensités, problème qui s'est retrouvé à toutes les étapes de l'évaluation. Cela peut s'expliquer en partie par la stratification faite lors des évaluations, et cela est visible même avec des ensembles bien calibrés. Concernant le choix des scores classiques d'évaluation, autant le diagramme de rang s'est avéré informatif sur le comportement des trois produits de prévision, autant le CRPS est apparu comme un score peu pratique sans une valeur de référence à laquelle se comparer.

Les étapes deux et trois étaient consacrées à l'étude des prévisions hydrologiques avec une adaptation de la méthode des courbes ROC. L'utilisation des seuils de débits et des tableaux de contingence pour effectuer une évaluation régionale des prévisions d'ensemble hydrologiques a déjà été suggérée par plusieurs auteurs (Silvestro et Rebora 2012, Anderson

et al. 2019, Sayama *et al.* 2020). L'agrégation de toutes les échéances de prévision a permis ici de séparer l'analyse en caractérisant dans un premier temps la détection des dépassements d'un seuil de débit, puis en évaluant dans un second temps les délais d'anticipation, qu'il est particulièrement important d'évaluer dans le cas des crues-éclair. Cette approche apporte également un intérêt avec la possibilité de représenter de façon cartographique la détection des dépassements sur tous les sous-bassins considérés dans l'évaluation. Cela a permis d'avoir une vision plus détaillée des résultats ainsi que la localisation des secteurs où les crues ont pu être anticipées.

Cette analyse peut être facilement adaptée à d'autres événements de crue éclair avec des caractéristiques différentes, ou même à d'autres types d'inondations. Par exemple, différents seuils de débit peuvent être fixés pour mieux distinguer les zones où les principales réactions hydrologiques ont été observées, un niveau d'anticipation plus important qui peut être exigé pour décider d'une détection avec succès (un seuil minimum de 15 minutes de temps d'anticipation a été utilisé dans cette étude), et différents percentiles d'ensemble peuvent être sélectionnés pour mieux illustrer le rôle que joue la probabilité considérée dans l'évaluation de différents produits de prévision.

8.2.3 Plus-value de l'évaluation basée sur les impacts

Dans un premier temps, le cadre d'évaluation choisi pour les prévisions d'impacts a été identique à celui utilisé pour les débits. Des différences sont néanmoins apparues dans l'analyse des résultats. Malgré des tentatives pour faire correspondre les deux évaluations, les seuils d'impacts n'étant pas identiques aux seuils de débits, des disparités entre les deux évaluations sont inévitables. La mise en place d'une évaluation basée sur les impacts apparaît donc comme une approche complémentaire à l'évaluation sur les débits seulement, qui apporte un autre point de vue sur les performances de prévision. L'utilisation des données d'impacts permet de cibler plus particulièrement les zones présentant des enjeux, les tronçons les plus vulnérables, et s'avère donc probablement plus pertinente du point de vue de l'utilisation des prévisions.

Dans une deuxième temps, les prévisions d'impacts ont été évaluées du point de vue de leur capacité à déclencher des prises de décisions pertinentes et efficaces (Chapitre 7). Cette évaluation a permis de mesurer un coût/bénéfice de l'anticipation donnée par les prévisions d'ensembles, ainsi que l'impact des fausses alarmes sur les tronçons non secourus. Un cadre d'évaluation pensé pour un service d'intervention (exemple d'un SDIS pris ici) a été mis en œuvre afin d'évaluer l'apport des prévisions d'ensemble dans le choix d'intervention. Cette approche comporte néanmoins des défauts, les principaux étant l'adaptation nécessaire à chaque utilisateur (à ses besoins en termes de prévisions ainsi qu'à ses moyens d'intervention disponibles) ainsi que les simplifications faites pour la modélisation (vitesse constante, intervention toujours possible, durée d'intervention basée sur un retour sous le seuil d'intervention...).

8.3 Des compléments encore nécessaires pour aboutir à une véritable chaîne opérationnelle de prévision immédiate des impacts de crues soudaines

Les paragraphes qui suivent s'intéressent aux potentielles perspectives s'ouvrant à partir des conclusions de ce travail ainsi qu'aux compléments directs qui peuvent être imaginés et pourraient venir renforcer les résultats obtenus.

8.3.1 Aspects relatifs à l'évaluation des pluies

Lors de l'étape 1 de l'évaluation (Fiabilité des prévisions d'ensemble de pluie), le choix a été fait de travailler sur deux scores et indicateurs classiques de la prévision d'ensemble, le CRPS et le diagramme de rang (Wilks 2011). Ils ont été retenus afin de rendre compte de deux qualités particulièrement recherchées pour les prévisions d'ensemble, la fiabilité et la finesse. Cependant, le CRPS, en particulier, a montré des limites en termes d'interprétation du fait de l'absence de référence à laquelle comparer les valeurs obtenues pour les trois produits de prévision d'ensemble. Il pourrait donc être intéressant de réfléchir à l'utilisation d'autres scores pour compléter l'évaluation comme les diagrammes de précision proposés par Randrianasolo (2012) qui permettent d'évaluer les erreurs entre l'observation et la prévision d'ensemble avec un classement des erreurs des membres avec une centralisation sur 0 ou encore le spread-skill score utilisé par Peredo-Ramirez (2021) pour évaluer ces mêmes produits de prévision d'ensemble. Des résultats de spread-skill scores sur les exutoires dans l'Aude sont présentés dans l'Annexe A.

8.3.2 Aspects relatifs à l'évaluation des débits

La deuxième étape de l'évaluation ("Capacité d'anticipation hydrologique pendant la montée de crue") comporte, elle aussi, un choix méthodologique qui peut être discuté. En effet, l'adaptation de la méthodologie des courbes ROC engendre une stratification par les observations et seule une partie des prévisions produites est utilisée pour l'évaluation, puisque le critère de constitution de l'échantillon des prévisions se fait à partir de l'observation. Ce type de stratification n'est pas conseillé, car il peut conduire à des biais comme l'ont montré Bellier *et al.* (2017). Des travaux ont déjà été conduits pour modifier cette méthode et proposer une stratification par les prévisions, sur d'autres cas d'étude (Godet 2022). Le bilan de ce premier essai avec une stratification par les prévisions au lieu des observations a montré que les résultats ne présentaient pas de différences majeures par rapport à l'approche initiale (taux de détection plus faible et un plus grand nombre de fausses alarmes), mais il reste néanmoins souhaitable d'éviter une stratification par les observations pour limiter les risques de biais d'évaluation.

Le modèle hydrologique sélectionné ici, Cinecar, n'apparaît pas totalement adapté à cet objectif du fait de son caractère événementiel et de l'absence de procédure de calage et d'initialisation du modèle sur de longues chroniques de crues. Les CN peuvent cependant être

ajustés en fonction de l'antériorité pluviométrique. Le calage de paramètres, quel que soit le type de modélisation utilisé, reste un sujet délicat sur les bassins versants non jaugés. D'autres modèles, comme GRSDi, ont été utilisés pour évaluer ces mêmes produits de prévision d'ensemble (Peredo-Ramirez 2021) avec un calage sur une période de 10 ans (2008 - 2018), offrant ainsi une évaluation des erreurs totales de prévision aux exutoires jaugés, incluant à la fois les erreurs liées aux prévisions pluviométriques et les erreurs de modélisation de la relation pluie-débit. Cependant, cette évaluation des erreurs ne peut être réalisée qu'aux exutoires jaugés, ce qui est souvent un point limitant pour les crues-éclair, survenant dans des petits bassins versants non instrumentés.

Seules les prévisions couvrant la phase critique de l'événement (montée de crue sur l'hydrogramme et montée des bâtigrammes pour les impacts avec le dépassement de seuil ou le maximum observé) ont été considérées pour construire les tables de contingence. Les résultats obtenus pourraient être différents si d'autres prévisions et/ou d'autres phases de l'événement étaient considérées. En particulier, dans le cas d'événements avec des pics de crue multiples (respectivement d'impacts multiples), le fait que la méthode se concentre uniquement sur le premier dépassement de seuil peut être considéré comme une limitation. Dans une telle situation, chaque montée de crue (respectivement d'impacts) de l'événement considéré pourrait être examinée séparément, même si dans certains cas, les différentes phases peuvent être difficiles à séparer. Une alternative pourrait être d'analyser l'anticipation des dépassements de seuil pour chaque prévision pendant l'événement, indépendamment des moments où les seuils sont dépassés pour les hydrogrammes de la simulation de référence RS. Fournir une évaluation pour de multiples seuils pour la même crue s'avérerait également un complément intéressant.

De plus, le choix de focalisation sur le premier dépassement de seuil pour le calcul des tableaux de contingence ne tient pas compte de la cohérence des prévisions successives qui est une qualité recherchée des prévisions d'ensemble (Richardson *et al.* 2020). Des évaluations spécifiques de la cohérence des prévisions ont été proposées (Ehret et Zehe 2011, Pappenberger *et al.* 2011). Dans la méthodologie proposée ici, seule la première fois qu'une prévision dépasse la valeur seuil de débit (respectivement la valeur seuil d'impact) a été considérée pour conclure à un succès (ou une fausse alarme) et pour calculer le temps d'anticipation, sans considérer la cohérence de plusieurs prévisions successives. Il pourrait également être intéressant d'utiliser la même méthode et de considérer le nombre de cas anticipant correctement les dépassements (ou non-dépassements) de seuil pour remplir la table de contingence et tenir compte de la cohérence des prévisions successives.

8.3.3 Aspects relatifs à la prévision des impacts

Le choix a été fait d'utiliser un catalogue de scénarios d'inondation pour obtenir une relation continue débits-impacts, selon le principe proposé par Le Bihan (2016), Le Bihan *et al.* (2017). Cette approche est une source importante d'approximations. En effet, une interpolation linéaire par morceaux est faite entre chaque débit ayant un scénario d'inondation. Cependant, il n'y a pas de raison pour que la relation débit - impacts soit linéaire (tout comme la relation hauteur d'eau - débit). Deux propositions d'amélioration peuvent être faites. Une

première piste pourrait être de réaliser les simulations hydrauliques correspondant à chaque prévision. Cela permettrait de s'affranchir de l'approximation linéaire et des simplifications de régime permanent (utilisé pour le calcul hydraulique des scénarios d'inondation), mais cette approche serait beaucoup plus lourde en termes de calculs du fait de la modélisation hydraulique à réaliser. C'était d'ailleurs l'argument principal qui a encouragé l'utilisation du catalogue de scénarios d'inondation. Une deuxième possibilité pourrait être de développer le catalogue de scénarios d'inondation en ajoutant des périodes de retour supplémentaires pour avoir des points intermédiaires pour mieux approcher la relation débits - impacts.

Toujours concernant la partie impacts, il pourrait être intéressant d'utiliser d'autres types de variables pour représenter les impacts. Pour ce travail, le choix a été fait de se concentrer uniquement sur les bâtiments. Ce volet impact gagnerait à être développé en utilisant des variables pouvant être plus directement reliées à des observations, comme les coupures de réseaux (routes, électricité, téléphone, SNCF...), l'activité sur les réseaux sociaux ou encore des données de sinistralité. Cela serait plus proche des réalités d'un gestionnaire de crise et permettrait d'améliorer le réalisme du rejeu de l'événement. Des travaux antérieurs comme ceux de Saint-Martin (2018) ont déjà travaillé l'affectation d'une pondération à différentes variables permettant de représenter des catégories différentes d'impacts.

Une partie du travail réalisé ici pourrait être réutilisé dans l'optique de la construction de chaînes opérationnelles de prévision immédiate des impacts crues soudaines. En particulier, l'idée d'une évaluation des impacts possibles reliés à un débit prévu, sans passer par la modélisation hydraulique, permettant de gagner un temps crucial pour l'anticipation des crues-éclair, semble assez simple à appliquer dans un contexte de suivi d'événement en temps réel.

Par ailleurs, l'utilisation effective de produits de prévisions d'ensemble expérimentaux (pepi et pertDpepi), qui ne sont pas encore inclus dans le flux de production temps réel de Météo-France, impliquera des délais de production et de livraison des prévisions qui réduiront d'autant les délais d'anticipation. Pour obtenir une vision plus réaliste de la capacité d'anticipation, les délais de livraison des différents produits de prévision des précipitations devraient être intégrés dans le calcul des tables de contingence et des délais d'anticipation, comme proposé par Lovat *et al.* (2020). Les prévisions ne pourraient alors être considérées qu'à partir de la date de livraison des produits utilisés en données d'entrée, au lieu du début de la séquence de temps couverte par les prévisions des précipitations. Cette intégration entraînerait probablement une réduction du nombre de succès et des temps d'anticipation réduits avec les prévisions d'ensemble hydrologiques. Une évaluation des délais d'anticipation dans de véritables conditions opérationnelles reste donc nécessaire, pour obtenir une vision plus objective de la plus-value des produits de prévision de pluie étudiés ici.

Bibliographie

Alfieri, L., M. Berenguer, V. Knechtl, K. Liechti, D. Sempere-Torres, et M. Zappa (2019). *Flash Flood Forecasting Based on Rainfall Thresholds*. Berlin, Heidelberg : Springer Berlin Heidelberg. (Cité page 36.)

Alfieri, L., P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen, et F. Pappenberger (2013). Glofas - global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences 17*(3), 1161–1175. (Cité pages vi, 15 et 16.)

Alfieri, L., F. Pappenberger, et F. Wetterhall (2013, 12). The extreme runoff index for flood early warning in europe. *Nat. Hazards Earth Syst. Sci. Discuss. 1*, 7517–7548. (Cité page 17.)

Amengual, A., D. S. Carrió, G. Ravazzani, et V. Homar (2017, apr). A comparison of ensemble strategies for flash flood forecasting : The 12 october 2007 case study in valencia, spain. *Journal of Hydrometeorology 18*(4), 1143–1166. (Cité page 42.)

Anctil, F. et M.-H. Ramos (2017). *Verification Metrics for Hydrological Ensemble Forecasts*. Springer Berlin Heidelberg. (Cité page 42.)

Anderson, S. R., G. Csima, R. J. Moore, M. Mittermaier, et S. J. Cole (2019, oct). Towards operational joint river flow and precipitation ensemble verification : considerations and strategies given limited ensemble records. *Journal of Hydrology 577*, 123966. (Cité pages 43, 128 et 211.)

Arnaud, P., Y. Eglin, B. Janet, et O. Payrastre (2013). Bases de données SHYREG-débit : méthode, performances et limites. Notice utilisateurs. Research report, irstea. (Cité page 59.)

Aubert, Y., P. Arnaud, P. Ribstein, et J. A. Fine (2014). La méthode SHYREG débit-application sur 1605 bassins versants en France métropolitaine. *Hydrol. Sci. J. 59*(5), 993–1005. (Cité pages 65 et 100.)

Auger, L., O. Dupont, S. Hagelin, P. Brousseau, et P. Brovelli (2015). AROME-NWC : A new nowcasting tool based on an operational mesoscale forecasting system. *Q. J. R. Meteorol. Soc. 141*(690), 1603–1611. (Cité pages 29, 56 et 57.)

Ayphassorho, H., G. Pipien, I. De Meritens, et D. Lacroix (2019). Retour d'expérience des inondations du 14 au 17 octobre 2018 dans l'Aude. Technical report, Ministère de la transition écologique et solidaire et Ministère de l'intérieur. (Cité page 50.)

Baldauf, M. (2013). The cosmo model : towards cloud-resolving nwp. Dans *Seminar on Numerical Methods for Atmosphere and Ocean Modelling, 2-5 September 2013*, Volume III, Shinfield Park, Reading, pp. 107–122. ECMWF : ECMWF. (Cité page 29.)

Bell, V. A., A. L. Kay, R. G. Jones, et R. J. Moore (2007). Development of a high resolution grid-based river flow model for use with regional climate model output. *Hydrol. Earth Syst. Sci. 11*(1), 532–549. (Cité page 35.)

Bellier, J. (2018). *Prévisions hydrologiques probabilistes dans un cadre multivarié : quels outils pour assurer fiabilité et cohérence spatio-temporelle ?* Ph. D. thesis, Grenoble Alpes. (Cité page 83.)

Bellier, J., I. Zin, et G. Bontron (2017). Sample stratification in verification of ensemble forecasts of continuous scalar variables : Potential benefits and pitfalls. *Monthly Weather Review 145*(9), 3529 – 3544. (Cité page 213.)

Berenguer, M., C. Corral, R. Sánchez-Diezma, et D. Sempere-Torres (2005). Hydrological validation of a radar-based nowcasting technique. *J. Hydrometeorol. 6*(4), 532–549. (Cité page 35.)

Bersani, C., F. Gerard, O. Gondran, A. Helias, X. Martin, P. Puech, P. Dumas, M. Rouzeau, B. Fleury, R. Bougere, et Y. Trepos (2010). Tempête Xynthia : retour d'expérience et propositions d'action. Technical report, Direction Générale de la Prévention des Risques. (Cité page 11.)

Bevere, L. et F. Remondi (2022). Natural catastrophes in 2021 : the floodgates are open. Technical report, Swiss Re Institute. (Cité page 10.)

Bouilloud, L., K. Chancibault, B. Vincendon, V. Ducrocq, F. Habets, G. M. Saulnier, S. Anquetin, E. Martin, et J. Noilhan (2010). Coupling the ISBA land surface model and the TOPMODEL hydrological model for mediterranean flash-flood forecasting : Description, calibration, and validation. *J. Hydrometeorol. 11*(2), 315–333. (Cité page 36.)

Bourgin, F. (2014). *Comment quantifier l'incertitude prédictive en modélisation hydrologique ? Travail exploratoire sur un grand échantillon de bassins versants.* Ph. D. thesis, AgroParis Tech. (Cité page 74.)

Bouttier, F., L. Raynaud, O. Nuissier, et B. Ménétrier (2016). Sensitivity of the AROME ensemble to initial and surface perturbations during HyMeX. *Q. J. R. Meteorol. Soc. 142*(August), 390–403. (Cité page 34.)

Bouttier, F., B. Vié, O. Nuissier, et L. Raynaud (2012). Impact of stochastic physics in a convection-permitting ensemble. *Mon. Weather Rev. 140*(11), 3706–3721. (Cité pages 34 et 56.)

Brousseau, P., Y. Seity, D. Ricard, et J. Léger (2016). Improvement of the forecast of convective activity from the AROME-France system. *Q. J. R. Meteorol. Soc. 142*(699), 2231–2243. (Cité pages 29 et 56.)

Buizza, R. et T. N. Palmer (1998). Impact of ensemble size on ensemble prediction. *Mon. Weather Rev. 126*(9), 2503–2518. (Cité page 89.)

Candille, G. et O. Talagrand (2005). Evaluation of probabilistic prediction systems for a scalar variable. *Q. J. R. Meteorol. Soc. 131*(609), 2131–2150. (Cité page 41.)

Carpenter, T. M. et K. P. Georgakakos (2004, 10). Continuous streamflow simulation with the hrcdhm distributed hydrologic model. *Journal of Hydrology 298*, 61–79. (Cité page 17.)

CCR (2018a). Les inondations du 14 et 15 octobre 2018 dans le Languedoc. Technical Report 11, Caisse Centrale de Réassurance. (Cité page 50.)

CCR (2018b). Rapport d'activité 2018. Technical report, Caisse Centrale de Réassurance. (Cité pages vi, 10, 14 et 50.)

Champeaux, J.-L., P. Dupuy, O. Laurantin, I. Soulan, P. Tabary, et J.-M. Soubeyroux (2009). Rainfall measurements and quantitative precipitation estimations at météo-france : inventory and prospects. *La Houille Blanche 95*(5), 28–34. (Cité pages 26 et 27.)

Collier, C. G. (2007, jan). Flash flood forecasting : What are the limits of predictability ? *Quarterly Journal of the Royal Meteorological Society 133*(622), 3–23. (Cité page 36.)

Coppola, E., R. Nogherotto, J. M. Ciarlo', F. Giorgi, E. van Meijgaard, N. Kadygrov, C. Iles, L. Corre, M. Sandstad, S. Somot, P. Nabat, R. Vautard, G. Levavasseur, C. Schwingshackl, J. Sillmann, E. Kjellström, G. Nikulin, E. Aalbers, G. Lenderink, O. B. Christensen, F. Boberg, S. L. Sørland, M. E. Demory, K. Bülow, C. Teichmann, K. Warrach-Sagi, et V. Wulfmeyer (2021). Assessment of the European Climate Projections as Simulated by the Large EURO-CORDEX Regional and Global Climate Model Ensemble. *J. Geophys. Res. Atmos. 126*(4), 1–20. (Cité page 28.)

Courtier, P., C. Freydier, J.-F. Geleyn, F. Rabier, et M. Rochas (1991). The arpege project at meteo france. Dans *Seminar on Numerical Methods in Atmospheric Models, 9-13 September 1991*, Volume II, Shinfield Park, Reading, pp. 193–232. ECMWF : ECMWF. (Cité page 28.)

Dantec, G. et G. Pipien (2019). Retours d'expérience (REX) des inondations. Technical report, CGEDD - Ministère de la Transition Ecologique et Solidaire. (Cité page 55.)

Davolio, S., M. M. Miglietta, T. Diomede, C. Marsigli, et A. Montani (2013, jun). A flood episode in northern italy : multi-model and single-model mesoscale meteorological ensembles for hydrological predictions. *Hydrology and Earth System Sciences 17*(6), 2107–2120. (Cité page 42.)

Davolio, S., F. Silvestro, et T. Gastaldo (2017, oct). Impact of rainfall assimilation on high-resolution hydrometeorological forecasts over liguria, italy. *Journal of Hydrometeorology 18*(10), 2659–2680. (Cité pages 30 et 42.)

Davolio, S., F. Silvestro, et P. Malguzzi (2015, jul). Effects of increasing horizontal resolution in a convection-permitting model on flood forecasting : The 2011 dramatic events in liguria, italy. *Journal of Hydrometeorology 16*(4), 1843–1856. (Cité page 43.)

Davy, P., T. Croissant, et D. Lague (2017). A precipiton method to calculate river hydrodynamics, with applications to flood prediction, landscape evolution models, and braiding instabilities. *J. Geophys. Res. Earth Surf. 122*(8), 1491–1512. (Cité page 65.)

de Almeida, G. A., P. Bates, et H. Ozdemir (2018). Modelling urban floods at submetre resolution : challenges or opportunities for flood risk management ? *J. Flood Risk Manag. 11*, S855–S865. (Cité page 39.)

Demargne, J., P. Javelle, D. Organde, L. Garandeau, et B. Janet (2019). Intégration des prévisions immédiates de pluie à haute-résolution pour une meilleure anticipation des crues soudaines. *La Houille Blanche 105*(3-4), 13–21. (Cité page 35.)

Déqué, M., C. Dreveton, A. Braun, et D. Cariolle (1994, 09). The arpege/ifs atmosphere model : a contribution to the french community climate modelling. *Climate Dynamics 10*, 249–266. (Cité page 28.)

Dottori, F., M. Kalas, P. Salamon, A. Bianchi, L. Alfieri, et L. Feyen (2017). An operational procedure for rapid flood risk assessment in europe. *Natural Hazards and Earth System Sciences 17*(7), 1111–1126. (Cité pages vi et 15.)

Drobinski, P., V. Ducrocq, P. Alpert, E. Anagnostou, K. Béranger, M. Borga, I. Braud, A. Chanzy, S. Davolio, G. Delrieu, C. Estournel, N. Filali Boubrahmi, J. Font, V. Grubišić, S. Gualdi, V. Homar, B. Ivančan-Picek, C. Kottmeier, V. Kotroni, K. Lagouvardos, P. Lionello, M. C. Llasat, W. Ludwig, C. Lutoff, A. Mariotti, E. Richard, R. Romero, R. Rotunno, O. Roussot, I. Ruin, S. Somot, I. Taupier-Letage, J. Tintor, R. Uijlenhoet, et H. Wernli (2014). HYMEX : A 10-year multidisciplinary program on the mediterranean water cycle. *Bull. Am. Meteorol. Soc. 95*(7), 1063–1082. (Cité page 59.)

Edouard, S. (2016). *Prévision d'ensemble des crues rapides méditerranéennes.* Ph. D. thesis, Université Toulous 3 Paul Sabatier. (Cité pages 35 et 36.)

Ehret, U. et E. Zehe (2011). Series distance - An intuitive metric to quantify hydrograph similarity in terms of occurrence, amplitude and timing of hydrological events. *Hydrol. Earth Syst. Sci. 15*(3), 877–896. (Cité pages 129 et 214.)

Fabre, G. (1990). La catastrophe hydrologique éclair de nîmes (3 octobre 1988) (the flash flood disaster at nîmes on 3 octobre 1988). *Bulletin de l'Association de Géographes Français 67*(2), 113–122. (Cité page 13.)

Ferro, C., D. Richardson, et A. Weigelc (2007). On the effect of ensemble size on the discrete and continuous ranked probability scores. *RMetS 15*, 19–24. (Cité page 74.)

Fourrié, N., E. Bresson, M. Nuret, C. Jany, P. Brousseau, A. Doerenbecher, M. Kreitz, O. Nuissier, E. Sevault, H. Bénichou, M. Amodei, et F. Pouponneau (2015). Arome-wmed, a real-time mesoscale model designed for the hymex special observation periods. *Geoscientific Model Development 8*(7), 1919–1941. (Cité page 30.)

Gaume, E. (2002). *Eléments d 'analyse sur les crues éclair.* Ph. D. thesis, Ecole Nationale du Génie Rural, des Eaux et des Forêts. (Cité pages vii et 61.)

Gaume, E., M. Livet, M. Desbordes, et J.-P. Villeneuve (2004, 01). Hydrological analysis of the river aude, france, flash flood on 12 and 13 november 1999. *Journal of Hydrology 286*, 135–154. (Cité pages 51 et 62.)

Georgakakos, K., T. Modrick Hansen, E. Shamir, R. Campbell, Z. Cheng, R. Jubach, J. Sperfslage, C. Spencer, et R. Banks (2021, 11). The flash flood guidance system implementation worldwide : A successful multidecadal research-to-operations effort. *Bulletin of the American Meteorological Society 103*, 1–35. (Cité page 17.)

Gneiting, T., F. Balabdaoui, et A. Raftery (2007). Probabilistic and sharpness forecasts , calibration. *Jrssb 69*(2), 243–268. (Cité pages 41 et 83.)

Godet, J. (2022). Évaluation de l'apport des prévisions d'ensemble piaf dans la chaîne vigicrues flash. Technical report. (Cité page 213.)

Golding, B. (2022). *Towards the "Perfect" Weather Warning*. Springer. (Cité page 41.)

Gourley, J. J., Z. L. Flamig, H. Vergara, P.-E. Kirstetter, R. A. Clark, E. Argyle, A. Arthur, S. Martinaitis, G. Terti, J. M. Erlingis, Y. Hong, et K. W. Howard (2017, feb). The FLASH project : Improving the tools for flash flood monitoring and prediction across the united states. *Bulletin of the American Meteorological Society 98*(2), 361–372. (Cité page 17.)

Hally, A., O. Caumont, L. Garrote, E. Richard, A. Weerts, F. Delogu, E. Fiori, N. Rebora, A. Parodi, A. Mihalović, M. Ivković, L. Dekić, W. van Verseveld, O. Nuissier, V. Ducrocq, D. D'Agostino, A. Galizia, E. Danovaro, et A. Clematis (2015, mar). Hydrometeorological multi-model ensemble simulations of the 4 november 2011 flash flood event in genoa, italy, in the framework of the DRIHM project. *Natural Hazards and Earth System Sciences 15*(3), 537–555. (Cité page 42.)

Hamill, T. M. (2001). Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Weather Rev. 129*(3), 550–560. (Cité page 72.)

Hapuarachchi, H. A. P., Q. J. Wang, et T. C. Pagano (2011, mar). A review of advances in flash flood forecasting. *Hydrological Processes 25*(18), 2771–2784. (Cité page 35.)

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather Forecast. 15*(5), 559–570. (Cité page 73.)

Heuvelink, D., M. Berenguer, C. C. Brauer, et R. Uijlenhoet (2020). Hydrological application of radar rainfall nowcasting in the Netherlands. *Environ. Int. 136*(August 2019), 105431. (Cité page 35.)

Hocini, N. (2022). *Evaluation de méthodes automatisées de cartographie des zones inondables adaptées à la prévision des crues soudaines*. Ph. D. thesis, Nantes Université. Thèse de doctorat dirigée par Gaume, Eric et Payrastre, Olivier Renaud Sciences de la Terre et de l'environnement Nantes 2022. (Cité pages 37, 39 et 144.)

Hocini, N., O. Payrastre, F. Bourgin, E. Gaume, P. Davy, D. Lague, L. Poinsignon, et F. Pons (2021, jun). Performance of automated methods for flash flood inundation mapping : a

comparison of a digital terrain model (DTM) filling and two hydrodynamic methods. *Hydrology and Earth System Sciences 25*(6), 2979–2995. (Cité page 63.)

Houdré, F. (2001). L'annonce des crues : Histoire et évolution des services de 1847 à nos jours. Technical report, Ministère de l'aménagement du territoire et de l'environnement, Grenoble. (Cité page 16.)

Hu, P., Q. Zhang, P. Shi, B. Chen, et J. Fang (2018). Flood-induced mortality across the globe : Spatiotemporal pattern and influencing factors. *Science of the Total Environment 643*(3), 171–182. (Cité page 10.)

Huet, P., X. Martin, J.-L. Prime, P. Foin, C. Laurain, et P. Cannard (2003). Retour d'expérience des crues de septembre 2002 dans les départements du Gard, de l'Hérault des Bouches du Rhône, de l'Ardèche et de la Drôme. Technical report, Inspection générale de l'administration ; Conseil général des ponts et chaussées ; Conseil général du génie rural des eaux et de forêts ; Inspection générale de l'environnement, Paris. (Cité page 13.)

IAHS (1974). Flash Floods Symposium. Paris, pp. 1 – 139. (Cité page 12.)

James, E. P., C. R. Alexander, D. C. Dowell, S. S. Weygandt, S. G. Benjamin, G. S. Manikin, J. M. Brown, J. B. Olson, M. Hu, T. G. Smirnova, T. Ladwig, J. S. Kenyon, et D. D. Turner (2022). The High-Resolution Rapid Refresh (HRRR) : An Hourly Updating Convection-Allowing Forecast Model. Part 2 : Forecast Performance. *Weather Forecast. 37*(7), 1397–1417. (Cité page 31.)

Jauffret, C. (2016). La prévision immédiate : un aperçu des outils et produits. Technical report, Météo-France. (Cité page 31.)

Javelle, P., C. Fouchier, P. Arnaud, et J. Lavabre (2010). Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *J. Hydrol. 394*(1-2), 267–274. (Cité page 35.)

Jay-Allemand, M. (2021). *Estimation variationnelle des paramètres d ' un modèle hydrologique distribué*. Ph. D. thesis, Aix-Marseille Université. (Cité page 36.)

Jolliffe, I. et D. Stephenson (2012). *Jolliffe et Stephenson edition 2012_Intro.pdf*. John Wiley & Sons, Ltd. (Cité pages 40 et 100.)

Lagasio, M., F. Silvestro, L. Campo, et A. Parodi (2019, jul). Predictive capability of a high-resolution hydrometeorological forecasting framework coupling WRF cycling 3dvar and continuum. *Journal of Hydrometeorology 20*(7), 1307–1337. (Cité pages 30 et 42.)

Laurantin, O. (2008). Hourly rainfall analysis merging radar and rain gauge data. Dans *Proc. Int. Symp. Weather Radar Hydrol.*, Grenoble, pp. 2–8. (Cité page 56.)

Lavabre, J., C. Fouchier, N. Folton, et Y. Gregoris (2003). SHYREG : une méthode pour l'estimation régionale des débits de crue. Application aux régions méditerranéennes françaises. *Ingénieries Spécial*, 97–111. (Cité page 59.)

Le Bihan, G. (2016). *Modèles hydrologiques régionaux pour la prévision distribuée des crues rapides : vers une estimation des impacts et des dommages potentiels.* Ph. D. thesis, Université Bretagne Loire. (Cité pages 37, 38, 39, 61, 62, 137, 208 et 214.)

Le Bihan, G., O. Payrastre, E. Gaume, D. Moncoulon, et F. Pons (2017). The challenge of forecasting impacts of flash floods : Test of a simplified hydraulic approach and validation based on insurance claim data. *Hydrol. Earth Syst. Sci. 21*(11), 5911–5928. (Cité pages vii, 38, 168 et 214.)

Lebouc, L., O. Payrastre, et F. Bourgin (2019a). Reconstitution des débits de pointe des crues du 15 octobre 2018 dans le bassin de l'Aude. Technical report, IFSTTAR, Nantes. (Cité pages 59 et 63.)

Lebouc, L., O. Payrastre, et F. Bourgin (2019b). Reconstitution des débits de pointe des crues du 9 août 2018 dans les bassins de l'Ardèche et la Cèze. Technical report, IFSTTAR. (Cité pages 59 et 63.)

Lefrou, C., X. Martin, J.-P. Labarthe, J. Varret, B. Maziere, et R. Feunteun (2000). Les crues des 12, 13 et 14 novembre 1999 dans les départements de l'Aude, de l'Hérault, des Pyrénées-Orientales et du Tarn. Technical report, Conseil Général des Ponts et Chaussées ; Inspection générale de l'Environnement, Paris. (Cité page 13.)

Leleu, I., I. Tonnelier, R. Puechberty, P. Gouin, I. Viquendi, L. Cobos, A. Foray, M. Baillon, et P.-O. Ndima (2014). La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques. *La Houille Blanche 1*, 25–32. (Cité page 59.)

Levine, R. A. et D. S. Wilks (2000). *Statistical Methods in the Atmospheric Sciences*, Volume 95. Academic press. (Cité page 100.)

Ligda, M. G. H. (1953). The horizontal motion of small precipitation areas as observed by radar. Technical report, M.I.T. Dept. of Meteorology, Cambridge. (Cité page 31.)

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of Atmospheric Sciences 20*(2), 130 – 141. (Cité page 34.)

Lovat, A. (2019). *Prévision à très courte échéance des crues rapides méditérranéennes.* Ph. D. thesis, INP Toulouse. (Cité page 36.)

Lovat, A., B. Vincendon, et V. Ducrocq (2020). Hydrometeorological evaluation of two nowcasting systems for Mediterranean heavy precipitation events with operational considerations. *Hydrol. Earth Syst. Sci. Discuss. 26*(December), 1–32. (Cité pages 32 et 215.)

Lu, C., H. Yuan, B. E. Schwartz, et S. G. Benjamin (2007). Short-range numerical weather prediction using time-lagged ensembles. *Weather Forecast. 22*(3), 580–595. (Cité page 57.)

Lumbroso, D. et E. Gaume (2012). Reducing the uncertainty in indirect estimates of extreme flash flood discharges. *J. Hydrol. 414-415*, 16–30. (Cité page 60.)

Marshall, G. (2007). *Statistical methods in the atmospheric sciences, second edition D. S. Wilks. 1995. International Geophysics Series, Vol 59, Academic Press, 464pp. ISBN-10 : 0127519653. ISBN-13 : 978-0127519654. 59.99.*, Volume 14. Royal Meteorological Society. (Cité page 40.)

Mason, I. (1982). A model for assessment of weather forecasts. *Aust. Meteor. Mag 30*, 291–303. (Cité page 100.)

Melun, G., F. Liébault, G. Piton, P. Passy, et C. Martins (2022). Crues exceptionnelles de la Vésubie et de la Roya (octobre 2020) : caractérisation hydrogéomorphologique et perspectives de gestion. Technical Report octobre 2020, Office Français de la Biodiversité, INRAE. (Cité pages 11 et 13.)

Météo-France (2020). Descriptif AROME-PI. Technical report, Météo-France, Direction des Opérations pour la Prévision. (Cité pages vi, 30 et 31.)

Météo-France (2021). Descriptif PIAF-Lame d'eau. (Cité page 27.)

Michel, C., V. Andréassian, et C. Perrin (2005, 2). Soil conservation service curve number method : How to mend a wrong soil moisture accounting procedure ? *Water Resources Research 41*, 6. (Cité page 62.)

Mottet, G., V. Malhaire, et C. Coste (1993). Fortes pluies et crues de septembre 1992 sur le bassin cristallin de l'ardèche en amont d'aubenas : causes et conséquences / the heavy rainfall and floods of september 1992 in the crystalline river basin of the ardèche above aubenas : causes and consequences. *Géocarrefour 68*(2), 159–169. Included in a thematic issue : Risques naturels dans le couloir rhodanien : les excès pluviométriques. (Cité page 55.)

Murphy, A. (1993). What is a good forecast ? An essay on the Nature of Goodness in Weather Forecasting. *Am. Meteorol. Soc. 8*(2), 13. (Cité page 41.)

Murphy, A. H. (1995). A coherent method of stratification within a general framework for forecast verification. *Monthly Weather Review 123*(5), 1582 – 1588. (Cité page 83.)

Météo-France (2015). Evolution modèle arome. Technical report, Météo-France CNRM. (Cité page 56.)

Nardo-Caseri, A. (2017). *Apport de la simulation conditionnelle géostatistique pour la prévision immédiate de pluie et l'alerte aux crues rapides.* Ph. D. thesis, AgroParis Tech. (Cité pages 36 et 37.)

National Centers for Environmental Prediction (U.S.) (2004). The GFS atmospheric model. (Cité page 28.)

Naulin, J. P., O. Payrastre, et E. Gaume (2013). Spatially distributed flood forecasting in flash flood prone areas : Application to road network supervision in southern france. *Journal of Hydrology 486*, 88–99. (Cité pages 37, 39, 61, 63 et 97.)

Neal, J., G. Schumann, T. Fewtrell, M. Budimir, P. Bates, et D. Mason (2011). Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. *J. Flood Risk Manag. 4*(2), 88–95. (Cité page 39.)

Nicolle, P. (2021). Production d'un catalogue d'emprises inondées mono-fréquence sur les petits cours d'eau de l'arc méditerranéen. Technical report, Université Gustave Eiffel. (Cité pages 65 et 137.)

Nuissier, O., C. Marsigli, B. Vincendon, A. Hally, F. Bouttier, A. Montani, et T. Paccagnella (2016, aug). Evaluation of two convection-permitting ensemble systems in the HyMeX special observation period (SOP1) framework. *Quarterly Journal of the Royal Meteorological Society 142*(S1), 404–418. (Cité page 42.)

Organde, D., P. Arnaud, J. A. Fine, C. Fouchier, N. Folton, et J. Lavabre (2013, apr). Régionalisation d'une méthode de prédétermination de crue sur l'ensemble du territoire français : La méthode SHYREG. *Rev. des Sci. l'Eau 26*(1), 65–78. (Cité page 59.)

Osinski, R. et F. Bouttier (2018). Short-range probabilistic forecasting of convective risks for aviation based on a lagged-average-forecast ensemble approach. *Meteorol. Appl. 25*(1), 105–118. (Cité page 57.)

Pailleux, J., J.-F. Geleyn, et E. Legrand (2011, 01). La prévision numérique du temps avec les modèles arpÈge et aladin - bilan et perspectives. *La Météorologie [ISSN 0026-1181], 2000, Série 8, N° 30 ; p. 32-60 30*, 32–60. (Cité page 28.)

Papagiannaki, K., O. Petrucci, M. Diakakis, V. Kotroni, L. Aceto, C. Bianchi, R. Brázdil, M. G. Gelabert, M. Inbar, A. Kahraman, Ö. Kılıç, A. Krahn, H. Kreibich, M. C. Llasat, M. Llasat-Botija, N. Macdonald, M. M. de Brito, M. Mercuri, S. Pereira, J. Řehoř, J. R. Geli, P. Salvati, F. Vinet, et J. L. Zêzere (2022). Developing a large-scale dataset of flood fatalities for territories in the Euro-Mediterranean region, FFEM-DB. *Sci. Data 9*(1), 1–14. (Cité page 39.)

Pappenberger, F., H. L. Cloke, A. Persson, et D. Demeritt (2011). HESS Opinions "on forecast (in)consistency in a hydro-meteorological chain : Curse or blessing ?". *Hydrol. Earth Syst. Sci. 15*(7), 2391–2400. (Cité pages 39, 129 et 214.)

Park, S., M. Berenguer, et D. Sempere-Torres (2019, jun). Long-term analysis of gauge-adjusted radar rainfall accumulations at european scale. *Journal of Hydrology 573*, 768–777. (Cité page 31.)

Pedrozo-Acuña, A., J. P. Rodríguez-Rincón, M. Arganis-Juárez, R. Domínguez-Mora, et F. J. González Villareal (2015). Estimation of probabilistic flood inundation maps for an extreme event : Pánuco River, México. *J. Flood Risk Manag. 8*(2), 177–192. (Cité page 39.)

Peredo-Ramirez, D. (2021). *Quels gains d une modélisation hydrologique adaptée et d'une approche d'ensemble pour la prévision des crues*. Ph. D. thesis, Sorbonne Université. (Cité pages 211, 213 et 214.)

Perrin, C., C. Michel, et V. Andréassian (2001). Does a large number of parameters enhance model performance ? Comparative assessment of common catchment model structures on 429 catchments. *J. Hydrol. 242*, 275–301. (Cité page 36.)

Perrin, F., P. Sauzey, B. Menoret, et P.-A. Roche (2017). Inondations de mai et juin 2016 dans les bassins moyens de la Seine et de la Loire. Rapport CGEDD n° 010743-01 et IGA n° 16080-R. Technical report, Ministère de l'Environnement de l'Energie et de la Mer, Ministère de l'Intérieur. (Cité page 11.)

Petrucci, O., K. Papagiannaki, L. Aceto, L. Boissier, V. Kotroni, M. Grimalt, M. C. Llasat, M. Llasat-Botija, J. Rosselló, A. A. Pasqua, et F. Vinet (2019). MEFF : The database of MEditerranean Flood Fatalities (1980 to 2015). *J. Flood Risk Manag. 12*(2), 1–17. (Cité page 39.)

Poletti, M. L., F. Silvestro, S. Davolio, F. Pignone, et N. Rebora (2019, sep). Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts. *Hydrology and Earth System Sciences 23*(9), 3823–3841. (Cité pages 30 et 42.)

Ponce, V. M. et R. H. Hawkins (1996, 1). Runoff curve number : Has it reached maturity ? *Journal of Hydrologic Engineering 1*, 11–19. (Cité page 62.)

Price, J. D., S. Lane, I. A. Boutle, D. K. E. Smith, T. Bergot, C. Lac, L. Duconge, J. McGregor, A. Kerr-Munslow, M. Pickering, et R. Clark (2018). Lanfex : A field and modeling study to improve our understanding and forecasting of radiation fog. *Bulletin of the American Meteorological Society 99*(10), 2061 – 2077. (Cité page 27.)

Randrianasolo, A. (2012). *Généralisation de l 'approche d 'ensemble à la prévision hydrologique dans les bassins versants non jaugés.* Ph. D. thesis, AgroParis Tech. (Cité page 213.)

Randrianasolo, R. A. (2009). *Evaluation de la qualité des prévisions pour l ' alerte aux crues.* Ph. D. thesis, Université Pierre et Marie Curie, Ecoles des Mines de Paris et Ecole Nationale du Génie Rural des Eaux et des Forêts. (Cité page 37.)

Raymond, F., P. Camberlin, et A. Ullmann (2016). Précipitations intenses sur le Bassin Méditerranéen : quelles tendances entre 1950 et 2013 ? *Cybergeo : Eureopean journal of geography* (760). (Cité page 14.)

Raynaud, L. et F. Bouttier (2016). Comparison of initial perturbation methods for ensemble prediction at convective scale. *Q. J. R. Meteorol. Soc. 142*(695), 854–866. (Cité page 56.)

Richardson, D. S., H. L. Cloke, et F. Pappenberger (2020). Evaluation of the Consistency of ECMWF Ensemble Forecasts. *Geophys. Res. Lett. 47*(11), 1–8. (Cité pages 129 et 214.)

Rollet, A.-J., S. Dufour, et C. Morhange (2012). La rupture du barrage de Malpasset 50 ans après : un observatoire privilégié pour l'étude des événements extrêmes ? *Revue géographique des pays méditerranéens 118*, 75–83. (Cité page 11.)

Saint-Martin, C. (2018). *Prise en compte des vulnérabilités territoriales dans l 'avertissement des crues rapides : vers une amélioration de la méthode AIGA.* Ph. D. thesis, Université de Montpellier - Paul Valery. (Cité pages 39 et 215.)

Saltikoff, E., G. Haase, L. Delobbe, N. Gaussiat, M. Martet, D. Idziorek, H. Leijnse, P. Novák, M. Lukach, et K. Stephan (2019, 06). Opera the radar project. *Atmosphere 10*, 320. (Cité page 27.)

Sayama, T., M. Yamada, Y. Sugawara, et D. Yamazaki (2020, dec). Ensemble flash flood predictions using a high-resolution nationwide distributed rainfall-runoff model : case study of the heavy rain event of july 2018 and typhoon hagibis in 2019. *Progress in Earth and Planetary Science 7*(1), 123966. (Cité pages 42, 128 et 212.)

Schaake, J. C., T. M. Hamill, R. Buizza, et M. Clark (2007). HEPEX : The hydrological ensemble prediction experiment. *Bull. Am. Meteorol. Soc. 88*(10), 1541–1547. (Cité page 36.)

SCHAPI (2018). Vigicrues flash en quelques mots. Technical report, Service central d'hydrométéorologie et d'appui à la prévision des inondations. (Cité pages 17 et 35.)

Schwartz, C. S. et R. A. Sobash (2017). Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches : A review and recommendations. *Mon. Weather Rev. 145*(9), 3397–3418. (Cité page 101.)

SDIS11 (2019). Schéma départemental d'analyse et de couverture des risques (sdacr), risques courants et risques complexes. Technical report, Service Départemental d'Intervention et de Secours de l'Aude. (Cité page 158.)

Seity, Y., S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, et V. Masson (2011). The AROME-France convective-scale operational model. *Mon. Weather Rev. 139*(3), 976–991. (Cité pages 29 et 56.)

Sénat (2002). Rapport de la Commission d'enquête sur les inondations de la Somme. Technical Report 1, Sénat. (Cité page 11.)

Shen, X., K. Meng, L. Zhang, et X. Zuo (2022). A method of radar echo extrapolation based on dilated convolution and attention convolution. *Sci. Rep. 12*(1), 1–12. (Cité page 31.)

Silvestro, F. et N. Rebora (2012, mar). Operational verification of a framework for the probabilistic nowcasting of river discharge in small and medium size basins. *Natural Hazards and Earth System Sciences 12*(3), 763–776. (Cité pages 43 et 211.)

Silvestro, F., N. Rebora, et L. Ferraris (2011, dec). Quantitative flood forecasting on small- and medium-sized basins : A probabilistic approach for operational purposes. *Journal of Hydrometeorology 12*(6), 1432–1446. (Cité page 128.)

Sorbet, C., C. De Saint-Aubin, L. Cortes, V. Pourret, et A. Marchandise (2014). The chrome project : the building of an operational hydrometeorological ensemble to better anticipate flash-floods on french mediterranean rivers. Dans *8th HyMeX workshop, 15-18 September 2014, Valletta, Malta.* (Cité page 36.)

Soubeyroux, J.-M. (2021). LES FUTURS DU CLIMAT ET DE L'EAU EN FRANCE NOUVELLES SIMULATIONS CLIMATIQUES DRIAS-2020. Technical report, Météo-France, Chinon. (Cité pages vi et 14.)

Tabary, P. (2007). The new french operational radar rainfall product. part i : Methodology. *Weather and Forecasting 22*(3), 393 – 408. (Cité page 26.)

Tabary, P., C. Augros, J.-L. Champeaux, J.-L. Chèze, D. Faure, D. Idziorek, R. Lorandel, B. Urban, et V. Vogt (2013, 11). Le réseau et les produits radars de météo-france. *La Météorologie 8ème série*, 15–27. (Cité page 26.)

Tabary, P., J. Desplats, K. D. Khac, F. Eideliman, C. Gueguen, et J.-C. Heinrich (2007). The new french operational radar rainfall product. part ii : Validation. *Weather and Forecasting 22*(3), 409 – 427. (Cité page 26.)

Talagrand, O. et R. Vautard (1997). Evaluation of probabilistic prediction systems for a scalar variable. (Cité page 72.)

Thieken, A. H., P. Bubeck, A. Heidenreich, J. von Keyserlingk, L. Dillenardt, et A. Otto (2022). Performance of the flood warning system in germany in july 2021 – insights from affected residents. *EGUsphere 2022*, 1–26. (Cité page 10.)

Thielen, J., J. Bartholmes, M. H. Ramos, et A. De Roo (2009). The European flood alert system ĝ€" part 1 : Concept and development. *Hydrol. Earth Syst. Sci. 13*(2), 125–140. (Cité pages 15 et 36.)

Versini, P.-A. (2007). *Surveillance du risque hydrologique diffus le long des itinéraires routiers - Eléments pour la construction d'un outil d'annonce de coupure.* Ph. D. thesis, Université Bretagne Loire. (Cité pages 37 et 61.)

Vié, B. (2013). *Méthodes de prévision d'ensemble pour l'étude de la prévisibilité à l'échelle convective des épisodes de pluies intenses en Méditerranée.* Ph. D. thesis, Université Paris Est. (Cité page 32.)

Vié, B., G. Molinié, O. Nuissier, B. Vincendon, V. Ducrocq, F. Bouttier, et E. Richard (2012, aug). Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for mediterranean heavy precipitating events. *Natural Hazards and Earth System Sciences 12*(8), 2631–2645. (Cité page 42.)

Vincendon, B., V. Ducrocq, O. Nuissier, et B. Vié (2011). Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Nat. Hazards Earth Syst. Sci. 11*(5), 1529–1544. (Cité pages 32, 36, 42 et 58.)

Vinet, F., J. P. Cherel, K. Weiss, M. Lewandowski, et L. Boissier (2022). La mortalité liée aux inondations en région méditerranéenne française (1980–2020). *LHB Hydrosci. J. 108*(1), 1–14. (Cité page 39.)

Wilks, D. (1995). *Chapter 7 Forecast verification.* Elsevier. (Cité page 41.)

Wilks, D. S. (2011). *Forecast Verification*, Volume 100. Elsevier. (Cité pages 41, 45 et 213.)

Winsemius, H. C., L. P. Van Beek, B. Jongman, P. J. Ward, et A. Bouwman (2013). A framework for global river flood risk assessments. *Hydrol. Earth Syst. Sci. 17*(5), 1871–1892. (Cité page 39.)

Wood, E. F., J. K. Roundy, T. J. Troy, L. P. H. Van Beek, M. F. P. Bierkens, E. Blyth, A. De Roo, P. Döll, M. Ek, J. Famiglietti, D. Gochis, N. Van de Giesen, P. Houser, P. R. Jaffé, S. Kollet, B. Lehner, D. P. Lettenmaier, C. Peters-Lidard, M. Sivapalan, J. Sheffield, A. Wade, et P. Whitehead (2011). Hyperresolution global land surface modeling : Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res. 48*(1), 1–10. (Cité page 39.)

Yıldırım, A. A., D. Watson, D. Tarboton, et R. M. Wallace (2015, sep). A virtual tile approach to raster-based calculations of large digital elevation models in a shared-memory system. *Comput. Geosci. 82*, 78–88. (Cité page 63.)

Zanchetta, A. et P. Coulibaly (2020, feb). Recent advances in real-time pluvial flash flood forecasting. *Water 12*(2), 570. (Cité page 35.)

Annexes

A

Un cadre méthodologique pour l'évaluation événementielle des prévisions hydrométéorologiques des crues éclair à courte échéance

A methodological framework for the evaluation of short-range flash-flood hydrometeorological forecasts at the event scale

Maryse Charpentier–Noyer[1], Daniela Peredo[2,3], Axelle Fleury[4], Hugo Marchal[4], François Bouttier[4], Eric Gaume[1], Pierre Nicolle[1], Olivier Payrastre[1], and Maria-Helena Ramos[3]

[1]GERS/LEE, Univ Gustave Eiffel, Nantes, 44344, France
[2]UMR Metis, Sorbonne Université, Paris, France
[3]Université Paris-Saclay, INRAE, UR HYCAR Antony, 92160, France
[4]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Correspondence: Maryse Charpentier–Noyer (maryse.charpentier-noyer@univ-eiffel.fr)

Abstract.

This paper presents a methodological framework designed for the event-based evaluation of short-range hydro-meteorological ensemble forecasts, in the specific context of an intense flash-flood event characterized by high spatio-temporal variability. The proposed evaluation adopts the point of view of end-users in charge of the organization of evacuations and rescue operations at a regional scale. Therefore, the local exceedance of discharge thresholds should be anticipated in time and accurately localized. A step-by-step approach is proposed, including first an evaluation of the rainfall forecasts. This first step helps to define appropriate spatial and temporal scales for the evaluation of flood forecasts . The anticipation of the flood rising limb (discharge thresholds) is then analyzed at a large number of ungauged sub-catchments, using simulated flows and zero-future rainfall forecasts as references. Based on this second step, several gauged sub-catchments are selected, at which a detailed evaluation of the forecast hydrographs is finally achieved.

This methodology is tested and illustrated on the October 2018 flash-flood, which affected part of the Aude River basin (south-eastern France). Three ensemble rainfall now casting research products recently proposed by Météo-France are evaluated and compared. The results show that, provided that the larger ensemble percentiles are considered (75% percentile for instance), these products correctly retrieve the area where the larger rainfall accumulations were observed, but have a tendency to overestimate its spatial extent. The hydrological evaluation indicates that the discharge threshold exceedances are better localized and anticipated if compared to a naive zero-future rainfall scenario, but at the price of a significant increase of false alarms. Some differences in the performances between the three ensemble rainfall forecast products are also identified

Finally, even if the evaluation of ensemble hydro-meteorological forecasts based on a low number of documented flood events remains challenging due to the limited statistical representation of the available data, the evaluation framework proposed herein should contribute to draw first conclusions about the usefulness of newly developed rainfall forecast ensembles for flash-flood forecasting purpose, and about their limits and possible improvements.

1 Introduction

Flash floods contribute in a significant proportion to flood-related damages and fatalities in Europe, particularly in the Mediterranean countries (Barredo, 2006; Llasat et al., 2010, 2013; Petrucci et al., 2019). As an illustration, over the period 1989-2018, four of the eight most damaging floods in France were flash floods, each having caused insurance losses that exceeded 500 million euros according to the French Central Reinsurance Fund (CCR, 2020). These floods are characterized by fast dynamics and high specific discharges (Gaume et al., 2009; Marchi et al., 2010), which largely explain their destructive power. They generally result from heavy precipitation events, typically exceeding hundreds of millimeters of rainfall totals in less than 6 hours, falling over river basins of less than 1000 km^2 of drainage area. They are also characterized by a high spatio-temporal variability and limited predictability (Georgakakos, 1986; Borga et al., 2008). Improving the capacity of flood monitoring and forecasting systems to anticipate such events is a key factor to limit their impacts and improve flood risk management. Although several operational flash-flood monitoring services based on weather radar rainfall are already implemented worldwide (Price et al., 2012; Clark et al., 2014; de Saint-Aubin et al., 2016; Javelle et al., 2016; Gourley et al., 2017; Park et al., 2019), these systems can only offer limited anticipation due to the short response times of the affected catchments.The integration of short-range, high resolution rainfall forecasts in flash-flood monitoring services is required to increase anticipation times beyond current levels (Collier, 2007; Hapuarachchi et al., 2011; Zanchetta and Coulibaly, 2020).

Weather forecasting systems that are well suited to capture heavy precipitation events have been developed in the last decade with the emergence of high resolution and convection-permitting numerical weather prediction (NWP) models (Clark et al., 2016). The spatial and temporal resolution of such models (typically 1 km and 1 min) are more relevant to the scales of the (semi-)distributed hydrological models that are commonly used in flash-flood monitoring and forecasting systems. Convection-permitting NWP models may also be combined with radar measurements through assimilation and/or blending techniques to obtain improved and seamless short-range rainfall forecasts (Davolio et al., 2017; Poletti et al., 2019; Lagasio et al., 2019). Despite these advances, the use of high resolution rainfall forecasts to issue flash-flood warnings still faces numerous challenges, mainly due to the uncertainties in the temporal distribution and the spatial location of the high rainfall accumulations cells over small areas (Silvestro et al., 2011; Addor et al., 2011; Vincendon et al., 2011; Hally et al., 2015; Clark et al., 2016; Armon et al., 2020; Furnari et al., 2020). Accurate forecasts and a reliable representation of forecast uncertainties are thus necessary to provide useful flash-flood warnings based on outputs of NWP models. The question of quantifying uncertainties in hydrometeorological forecasting systems has been increasingly addressed through ensemble forecasting approaches (e.g. Valdez et al., 2022; Bellier et al., 2021; Thiboult et al., 2017). Several ensemble flash-flood forecasting chains have been proposed in the literature, involving either convection-permitting NWP models for early warnings (Silvestro et al., 2011; Addor et al., 2011; Vié et al., 2012; Alfieri and Thielen, 2012; Davolio et al., 2013, 2015; Hally et al., 2015; Nuissier et al., 2016; Amengual et al., 2017; Sayama et al., 2020; Amengual et al., 2021), or radar advection approaches and/or radar data assimilation in NWP models for very short-range forecasting (Berenguer et al., 2011; Vincendon et al., 2011; Silvestro and Rebora, 2012; Davolio et al., 2017; Poletti et al., 2019; Lagasio et al., 2019).

New approaches to flash-flood forecasting need to be appropriately evaluated, and cannot rely only on the evaluation of the high resolution rainfall forecasts used as input. In one sense, flood forecasting verification can be seen as a form of fuzzy verification of rainfall forecasts (Ebert, 2008; Roberts and Lean, 2008), accounting for the averaging effect and the non-linearity of the rainfall-runoff process, and also for the postitions of the watershed limits. Flood forecast evaluation is generally based on long time series of observed and forecast data. However, in the case of flash-flood forecast evaluation, working on long time series is often not possible. At first, high resolution ensemble rainfall forecasts from convection-permitting NWP models are rarely available for long periods of re-forecasts. This is due to the fast evolution of input data (type, availability) and the frequent updates brought to the NWP models or the data assimilation approaches (Anderson et al., 2019). Secondly, because of the limited frequency of occurrence of heavy precipitation events triggering flash floods, the data-sets available for evaluation are very often limited to some specific severe events, which highly limits the possibility of satisfying the typical requirements for a robust evaluation of hydrological forecasts (Addor et al., 2011; Davolio et al., 2013). Event-based evaluations often rely on the visual inspection of forecasts against observed hydrographs or on the assessment of the anticipation of exceedances over pre-defined discharge thresholds, and such often at a few gauged outlets, where the main hydrological responses to the high rainfall accumulations were observed during the event (Vincendon et al., 2011; Vié et al., 2012; Davolio et al., 2013; Hally et al., 2015; Nuissier et al., 2016; Amengual et al., 2017; Lagasio et al., 2019; Sayama et al., 2020). When different forecast runs are available for the same event and along its duration (e.g., short-range forecasts generated by NWP models from different forecast initialization cycles) and/or when different basins are affected by the same event, statistical scores and frequency analyses, such as the RMSE, the CRPS, contingency tables or ROC curves, may also be used to provide a synthetic evaluation of the performance of the forecasts for the event being evaluated (Davolio et al., 2017; Poletti et al., 2019; Sayama et al., 2020). Although widely used in the scientific literature and in post-event reports, these evaluation frameworks raise several methodological questions: i) without a general guidance, the focus on one event or a few typical severe events may generate an event-specific evaluation, which might not be reproducible across different events or might not be statistically representative of forecast performance for other future events; ii) scores that offer a synthesis of performance over spatial and temporal scales might conceal the internal (in space and in time) variability of forecast performance (over different forecast initialization times and along lead times); iii) forecast evaluations that focus on gauged outlets that display the main hydrological responses to rainfall only offers a partial view of the forecasting system's performance, notably when impacts are also observed at ungauged sites and/or when significant spatial shifts exist between observed and forecast rainfalls. Therefore, the evaluation of short-range flash-flood forecasts at the event scale requires considering specific forecast quality attributes evaluated at gauged sites (where observations are available) but also a more regional-scale evaluation at ungauged sites, in order to achieve a more robust evaluation of forecast performance. Several authors already pointed out the interest of providing such regional scale hydrological evaluations (Silvestro and Rebora, 2012; Davolio et al., 2015; Anderson et al., 2019; Sayama et al., 2020).

In this paper, an evaluation of three new ensemble rainfall forecast products is presented in the perspective of their use for flash flood forecasting. The evaluated products have been specifically developed by the French meteorological service (Météo-France) to generate short-range rainfall forecasts (1 to 6 hours of lead time) that can potentially better capture Mediterranean heavy precipitation events. The products comprise the French AROME-EPS reference ensemble forecast (Bouttier et al., 2012;

Raynaud and Bouttier, 2016), and two experimental products merging AROME-EPS and another convection permitting NWP model-AROME-NWC (Auger et al., 2015), with optional incorporation of spatial perturbations as post-processing (Vincendon et al., 2011). Since the two experimental products have been released only for the autumn 2018 period in France, the evaluation can only be based on a single event, i.e. the heavy flood that occurred in the Aude River basin on October 15th, 2018. Therefore, a new framework is proposed for the evaluation of flash-flood hydro-meteorological ensemble forecasts at the event scale. In this approach, the evaluation is mainly focused on the capacity of the hydrometeorological forecasts to anticipate the exceedance of predefined discharge thresholds and to accurately localize the affected streams within the region of interest. These are two essential qualities of hydrometeorological forecasts that are needed to plan rescue operations in real time. The forecast-based financing approach developed for humanitarian actions adopts a similar pragmatic approach, but with the aim to release funding and trigger short-term actions in disaster-prone areas worldwide (Coughlan de Perez et al., 2015). Others methods with a specific interest for operational considerations have also been recently proposed for the case of deterministic forecasts (Lovat et al., 2020).

In the following, Section 2 presents the step-by-step evaluation framework proposed for the event-scale evaluation of ensemble forecasts. Section 3 presents the case study, the data and models used to produce discharge forecasts. In Section 4, the obtained results are presented and evaluated. Section 5 discusses the main outcomes, while Section 6 summarizes the conclusions and draws the perspectives of this study.

2 Methodology for an event-scale evaluation of hydro-meteorological ensemble forecasts

The proposed evaluation framework aims at determining if the magnitude of the floods generated by heavy precipitation events can be correctly anticipated based on ensemble rainfall forecasting products. It is considered that such products might not perfectly capture the complex spatial and temporal patterns of the observed rainfall, although they might still be useful to inform flood risk decision-making. More precisely, the question of anticipating high discharges that might exceed predefined discharge thresholds is addressed. The evaluation should not focus only on selected river sections, but offer a comprehensive view of anticipation capacities for the whole river network, including ungauged rivers.

Another challenge of the event-scale evaluation is to select a limited number of aggregated criteria that help drawing sound conclusions, owing to the possible high spatial and temporal variability of rainfall and runoff values and of model performance, as well as to the many possible combinations of time steps, forecast lead times and locations along the river network, that need to be considered in the evaluation. For this reason, a step-by-step approach is proposed (Figure 1). It is first based on an initial assessment of the rainfall forecasts, with a focus on the time and space windows of observed or forecast high rainfall accumulations. Then, a geographical analysis of the anticipation capacities of the ensemble flood forecasts is performed at a large number of ungauged outlets, focusing on the most critical phase of the floods (hydrograph rising limbs). Based on this second step, a detailed evaluation of the performance of the flood forecasts is conducted at some selected representative catchment outlets. These different steps are described below.

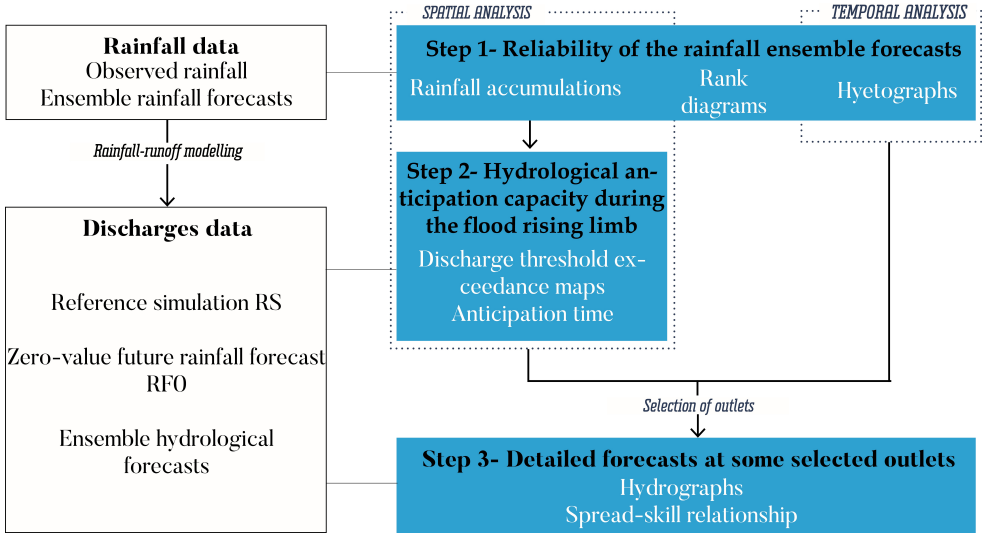

Figure 1. Overall principle of the proposed evaluation framework with its three steps of evaluation along the hydrometeorological forecasting chain

2.1 Step 1: Reliability of the rainfall ensemble forecasts

The twofold objective of this initial phase is to analyze the quality of the rainfall forecasts and to define the relevant spatial and temporal scales to be used in the subsequent analyses. Three different aspects are considered for a comparison of observed and forecast rainfall values:

- The hyetographs of the average rainfall intensities over the studied area are first plotted for each rainfall ensemble product and the different forecast lead times. They are used to assessing if, on average, the time sequence and the magnitude of the rainfall intensities are well captured by the products. A reduced time window, where significant intensities are forecast or measured, is selected for the next steps of the analysis. This time window is hereafter called "hydrological focus time" (HFT).
- Maps of the sum of forecast and observed hourly rainfall amounts during the HFT are then generated to assess if the areas where high rainfall accumulations were predicted and actually occurred coincide. One map is produced per forecast product, forecast lead time and ensemble percentile. These maps compare the spatial distribution of accumulated rainfalls from the aggregation of all forecast runs that were delivered during the event. It is thus the rainfall totals that are first

evaluated. These maps also help to delineate the areas affected by high totals of measured or forecast rainfalls. The next evaluation steps will focus on these areas, which are hereafter designated as "hydrological focus area" (HFA).

– Classical rank diagrams (Talagrand and Vautard, 1997; Hamill, 2001) are computed for the entire HFA domain and HFT period, for each forecast product and specific lead times. The aim is to quickly detect or confirm possible systematic biases or lack of variability in the forecast ensemble rainfall products. The diagrams are calculated considering all rainfall pixels over the entire HFA, and also for particular high rainfall intensities of interest (in mm/h). Rank diagrams show the frequencies at which the observation falls in each of the ranks of the ensemble members for each forecast product, when members are sorted from lowest to highest. They are used to determine the reliability of ensemble forecasts and to diagnose errors in its mean and spread (Hamill, 2001). Typically, sloped diagrams will indicate consistent biases in the ensemble forecasts (under- or over-estimation of the observations); U-shaped or concave diagrams are a signal of a lack of variability in the distribution given by the ensemble forecasts; an excess of variability will result in a rank diagram where the middle ranks are overpopulated.

2.2 Step 2: Hydrological anticipation capacity during the flood rising limb

The objective of this second step is to characterize the anticipation capacity of pre-selected discharge threshold exceedances for the whole HFA (including a large number of ungauged outlets) and during the most critical phase of the event (i.e. the flood rising limb), based on the ensemble discharge forecasts. The evaluation is essentially based on a classical contingency table approach (Wilks, 2011), with some important adaptations aiming to focus the analysis on the most critical time window from a user perspective (runs of forecasts preceeding the threshold exceedance), and to aggegate the forecats issued during this time window, independently of the lead-times (i.e., a hit is considered if at least one of the forecasts has exceeded the threshold at any lead time). Based on this framework, forecasting anticipation times are also computed (see Appendix A for a detailed description of the implemented method).

To ensure a certain homogeneity over the focus area (HFA), the same return period is used to define the discharge thresholds at the different outlets. A 10-year return (Aubert et al., 2014) was considered appropriate in this study given the magnitude of the flood event investigated (see Section 3), but it can be adapted to the intensity of any evaluated flood event. In addition, maps can be drawn for each forecasting system, each threshold level and ensemble percentile, to show the spatial distribution of the outlets that display hits, misses, false alarms or correct rejections. The corresponding histograms of misses, false alarms and hits, sorted by categories of anticipation time, can also be drawn to provide an overall performance visualization for the comparison of the different systems when considering all the HFA focus area. Finally, ROC curves based on the above definitions of hits, misses and false alarms are drawn to help to rank the methods independently of a specific percentile (see Appendix A).

To enable the integration of ungauged outlets in the analysis, the discharges simulated with observed radar rainfall are considered as reference values for the computation of the evaluation criteria (reference simulation, RS hereafter). An additional run is generated to help to interpret the results: it corresponds to a zero-value future rainfall scenario (RF0 hereafter), i.e.

forecasts are based on the propagation along the river network of past generated runoff only. This scenario helps to distinguish the part of the anticipation that is actually attributable to the rainfall forecasts and the part that can be explained by the propagation delays of the generated runoff flows along the river network.

2.3 Step 3: Detailed forecast analysis at some selected outlets

The main objective of this last step is to make the connection between the discharge thresholds anticipation results (step 2), and the detailed features of the ensemble hydrological forecasts. This is achieved by an in depth analysis of forecast hydrographs at different outlets, covering the whole HFT. The outlets are selected according to the results appearing on the maps elaborated in step 2, with the objective to cover the various situations (hit, miss or false alarm). The evaluation is based on the visual analysis of the forecast hydrographs for fixed forecast lead times (Berenguer et al., 2005) and on the spread-skill relationship, which evaluates the consistency between ensemble spread and ensemble mean error for the different lead times (Fortin et al., 2014; Anctil and Ramos, 2017). The spread-skill score for each lead time is obtained by comparing the RMSE of the ensemble mean (the skill) and the average of the standard deviations of the ensemble forecasts (the spread), as suggested by Fortin et al., 2014. The advantage of this score is that it can be easily calculated from the forecast outputs and provides also a measure of reliability of the ensembles (Christensen, 2015; Hopson, 2014). Finally, gauged outlets can also be examined in this phase, allowing the evaluation framework to also incorporate the hydrological modeling errors in the analysis.

3 Case study, data and models

3.1 The October 2018 flash-flood event in the Aude River basin

The Aude River basin is located in southern France (Figure 2). It extends from the Pyrenees mountains, in its South upstream edge, to the Mediterranean Sea. Its drainage area is 6,074 km^2. The climate is Mediterranean, with hot and dry summers and cool and wet winters. The mean annual precipitation over the basin is about 850 mm. High discharges are observed in winter and spring, but the major floods generally occur in autumn, and result from particularly intense convective rainfall events.

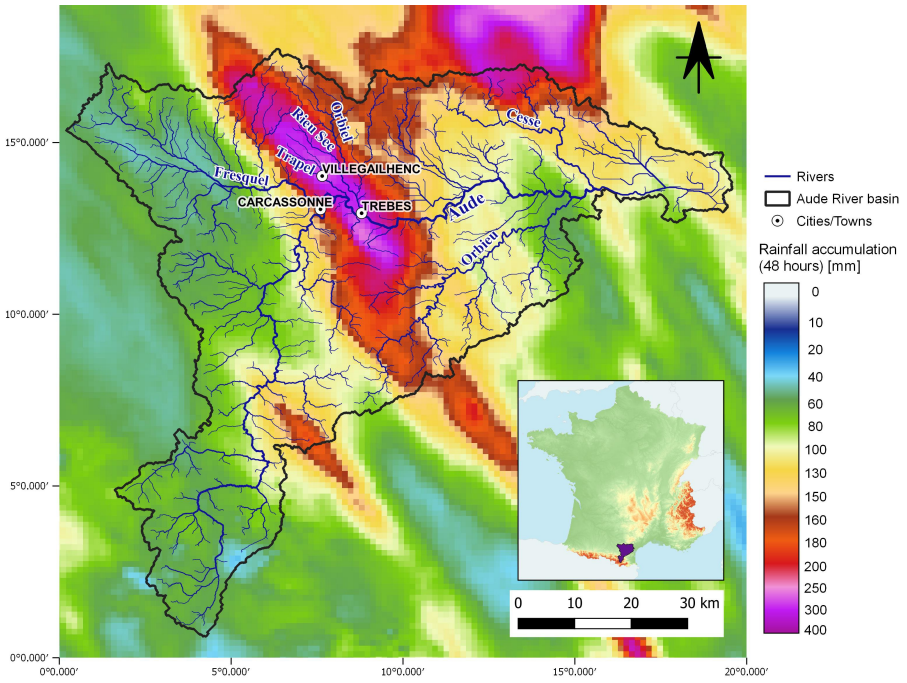

Figure 2. The Aude River basin, its river network, and the rainfall accumulations observed from14 October 2018 00:00 to 15 October 2018 23:00, according to the ANTILOPE J+1 quantitative precipitation estimates (see Section 3.2).

A major precipitation event occurred in this area from the 14th to the 16th October 2018, with particularly high rainfall accumulations during the night of 14th to 15th October. The maximum rainfall accumulations hit the intermediate part of the Aude River basin, just downstream the city of Carcassonne (Figure 2). Up to 300 mm of point rainfall accumulation was observed (Caumont et al., 2021). The maximum accumulated rainfall amounts over short durations were also extreme: up to 60 mm in one hour and 213 mm in six hours recorded at Villegailhenc (Figure 2), while the local 100-year rainfall accumulation is 200 mm in six hours (Ayphassorho et al., 2019). By its intensity and spatial extent, the 2018 event nears the record storm and flood event that hit the Aude River basin in November 1999 (Gaume et al., 2004).

The October 2018 floods of the Aude River and its tributaries led to the activation of the highest (red) level of the French flood warning system "Vigicrues" (vigicrues.gouv.fr). They caused 14 deaths, about 100 injuries and between 130 and 180 million euros of insured damages (CCR, 2018). The floods were particularly severe on the tributaries affected by the most

intense rains in the intermediate part of the Aude basin (Trapel, Rieu Sec, Orbiel, Lauquet), and also on the upstream parts of the Cesse and Orbieu tributaries (see Figure 2). The peak discharge on the main Aude River at Trèbes reached the 100-year value (Ayphassorho et al., 2019) only 9 hours after the onset of the rainfall event. In contrast, the peak discharges were high but not exceptional and caused limited inundation in the upstream and downstream parts of the Aude River basin. The high observed spatial and temporal rainfall and runoff variability makes this flash-flood event a challenging and interesting case study, for the event-scale evaluation framework of short range flood forecasts proposed in this study.

3.2 Observed hydrometeorological data

High resolution quantitative precipitation estimates (QPEs) for the event were obtained from the Météo-France's ANTILOPE algorithm (Laurantin, 2008), which merges operational weather radar (30 radars operating in October 2018) and rain gauge observations, including not only real time observations, but all observations available one day after the event. These QPEs are called ANTILOPE J+1 hereafter. A comprehensive reanalysis of this product is available for the period from 1 January 2008 to 18 October 2018, at the hourly time step and 1 km by 1 km spatial resolution. This product was used in this study to calibrate and run the hydrological models.

Discharge series were retrieved from the French Hydro database (Leleu et al., 2014; Delaigue et al., 2020) for the 31 stream gauges located in the Aude River basin, over the period 2008-2018. Additionally, peak discharges of the October 2018 flood event at ungauged locations were estimated during a post-flood field campaign (Lebouc et al., 2019) organized within the Hydrological cycle in the Mediterranean eXperiment (HyMeX; Drobinski et al., 2014) research program.

Evapotranspiration values, necessary to run the continuous rainfall-runoff model, were estimated using the Oudin formula (Oudin et al., 2005), based on temperature data extracted from the SAFRAN meteorological reanalysis produced by Météo-France on an 8 km x 8 km square grid (Vidal et al., 2010).

3.3 AROME-based short-range rainfall ensemble forecast products

Rainfall forecasts are based on Météo-France's AROME-France NWP model (Seity et al., 2011; Brousseau et al., 2016). AROME-France is an operational limited area model that provides deterministic weather forecasts up to two days ahead. Its high horizontal resolution allows to explicitly resolve deep convection, which is well suited to forecast heavy precipitations. Three different AROME-based short-range rainfall ensemble forecast products were evaluated in this work:

- **AROME-EPS** is the ensemble version of AROME-France (Bouttier et al., 2012; Raynaud and Bouttier, 2016). AROME-EPS results from perturbations of model equations, initial conditions and large-scale coupling of the NWP model. In 2018, AROME-EPS was a 12-member ensemble forecast, updated every 6 hours (four times per day: at 3h, 9h, 15h and 21h UTC), and providing forecasts up to two days ahead.

- **"pepi"** is a combination of AROME-EPS and AROME-NWC (Auger et al., 2015). AROME-NWC is a configuration of AROME-France designed for now-casting purposes; it is updated every hour and provides forecasts up to 6 hours ahead. In order to take into account sudden weather changes, AROME-NWC is used with time lagging (Osinski and

Bouttier, 2018; Lu et al., 2007), which consisted here in using the last 6 successive runs of AROME-NWC instead of using only the most recent run. The resulting "pepi" product provides forecasts for a maximum lead time of 6 hours. It combines 12 members from the last available AROME-EPS run, and 1 to 6 members from AROME-NWC, depending on the considered lead time. The resulting number of members varies between 13 (for a 6-hour lead time) and 18 (for a 1-hour lead time).

– **"pertDpepi"** is obtained by shifting the original "pepi" members 20 km in the four cardinal directions on top of the unperturbed pepi members, to account for uncertainties in the forecast rainfall location. The number of members is five times the number of the "pepi" ensemble, i.e. varies from 65 to 90 members depending on the lead time. This product is based on the concepts proposed by Vincendon et al. (2011), but uses a simpler framework to derive the test-product specifically designed for this study. The shift scale of 20 km represents a typical forecast location error scale: according to Vincendon et al. (2011), 80% of location errors are less than 50 km. The value of 20 km has been empirically tuned to produce the largest possible ensemble spread on a set of similarly intense precipitation cases, without noticeably degrading the ensemble predictive value as measured by user-oriented scores such as the area under the ROC curves..

These three rainfall forecast products have the same spatial (0.025° by 0.025°) and temporal (1 hour) resolutions, and cover the same saptial window as AROME (metropolitan territory of France). For the comparison with rainfall observations, they were disaggregated on the corresponding 1 km by 1km grid. In order to issue hydrological forecasts every hour, the last available runs of AROME-NWC and AROME-EPS (and the resulting pepi and pertDpepi ensembles) were systematically used, according to the products updates. The AROME-NWC and AROME-EPS forecasts were supposed to be immediately available for each update (i.e. the computation delays were not considered).

3.4 Rainfall-runoff models

The ensemble rainfall forecasts were used as input of two rainfall-runoff models, which differ in their resolution and structure. The GRSDi model is a semi-distributed continuous hydrological model adapted from the GRSD model (Le Moine et al., 2008; Lobligeois, 2014; De Lavenne et al., 2016) to better simulate autumn Mediterranean floods that typically occur after long dry summer periods (Peredo et al., 2022). The Cinecar model is an event-based distributed model, specifically developed to simulate flash-floods of small ungauged headwater catchments, with limited calibration needs (Versini et al., 2010; Naulin et al., 2013; Le Bihan, 2016). Both models were calibrated against observations and presented good performance for the 2018 flood event, as presented below, where we provide more information about the models and their implementations.

The objective of this study is not to compare the rainfall-runoff models. Since the RS hydrographs (hydrographs simulated with ANTILOPE J+1 rainfall observations) are systematically used as reference for the evaluation of the flood forecasts, the evaluation results should not be directly dependent on the rainfall-runoff model but rather on the nature of the rainfall forecasts used as input. The interest of using two models here is mainly to strengthen the evaluation, by involving two complementary models in terms of resolution and calibration approach: a) because of its high spatial resolution, the Cinecar models helps to extend the evaluation of discharge threshold anticipation to small ungauged catchments, b) because it was not specifically

calibrated on the 2018 event (calibration on the whole 2008-2018 period), the GRSDi models offers an evaluation of the total forecast errors at gauged outlets, including both the rainfall forecasts errors and the rainfall-runoff modeling errors. This is achieved by the comparison of flood forecasts with both RS hydrographs and observed hydrographs. However, the proposed evaluation framework could also be applied by using one unique rainfall runoff model.

GRSDi model

GRSDi is a soil moisture accounting, reservoir-based hydrological model that runs at an hourly time step with rainfall and evapotranspiration as input data. The model has 7 parameters to be calibrated against observed discharges. For its implementation, the Aude River basin was divided into 123 modelling units (gray contours in Figure 3a) of approximately 50 km^2. The model calibration and validation were performed for the same period (2008-2018), including the October 2018 flood event: 16 gauged outlets were selected for the calibration and 15 for the validation. The averaged KGE (Kling-Gupta Efficiency - Gupta et al., 2009) values were of 0.80 (0.71) for the 16 calibration outlets (15 validation outlets), which indicates good model performance, except for one validation outlet, where a low KGE value of 0.1 was obtained (Figure 3a). After visual inspection of the simulated discharges, this low performance was explained by an overestimation of the base flow, with however limited impact on the peak discharges during the high flows and flood events. The model and its performance evaluation is presented in details in Peredo et al. (2022).

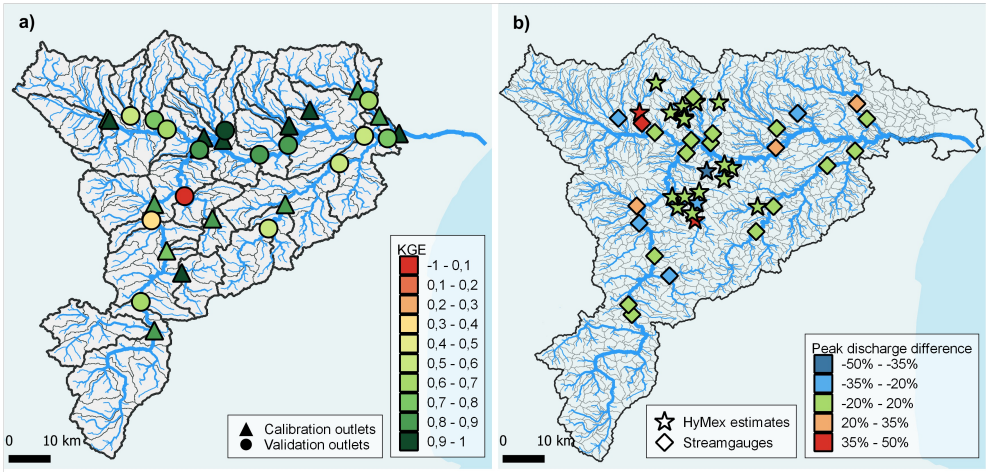

Figure 3. Calibration-validation results for the two rainfall-runoff models: a) KGE values at calibration and validation outlets (stream gauges) for the GRSDi model and the period 01/10/2008 at 0h to 15/10/2018 at 23h, b) Peak discharge difference (in %) between CINECAR simulated discharges and HyMex (HYdrological cycle in the Mediterranean EXperiment) estimates (stars - Lebouc et al., 2019) or stream gauges (diamonds) when the model is calibrated over the Aude 2018 flood event.

CINECAR model

The Cinecar model combines a SCS-CN (Soil Conservation Service - Curve Number, CN) model for the generation of effective runoff on the hill slopes and a kinematic wave propagation model on the hill slopes and in the stream network. The model runs at a 15 min time-step and was specifically developed to simulate fast runoff during flash floods, with a lesser focus on reproducing delayed recession limbs. The main parameter of the model requiring calibration is the Curve Number (Naulin et al., 2013). For the implementation of the model, the Aude River basin was divided into 1174 sub-basins with an average area of 5 km^2 (gray contours in Figure 3b). The shapes of the river reaches and hill slopes are simplified in the model but their main geometric features (slopes, areas, length, width) are directly extracted from the Digital Terrain Model (DTM). The CN values, first fixed on the basis of soil types and antecedent conditions, were further tuned to reach a better agreement between simulated and observed peak discharges for the Aude 2018 flood (Hocini et al., 2021). The resulting model was overall consistent with field observations, with errors on peak discharges generally comprised in the $\pm 20\%$ range (Figure 3b).

4 Results

4.1 Overall performance of the rainfall ensemble forecast products

Figure 4 shows the temporal evolution of the hourly observed rainfall (ANTILOPE J+1) and the hourly forecast rainfall for the three ensemble products (AROME-EPS, pepi and pertDpepi) at a) 1 hour and b) 6 hours of lead time from the 14 th October 07:00 to the 15 th October 19:00. Rainfall intensities (mm/h) are averaged over the entire Aude River basin area. Overall, the observed rainfalls are well captured by the forecast products. The ensemble forecast distributions are similar among the three products, except for some time steps corresponding to the most intense rainfall period, where the added value of the AROME-NWC model used in the pepi and pertDpepi products can be noticed. As expected, the spread of the ensemble forecasts increases from AROME-EPS to pepi and pertDpepi, with some rare exceptions. This is in agreement with the increase in the number of ensemble members.

In general, the observed average rainfall rates are contained in the ensemble forecast ranges, except at the end of the rainfall event, on the 15th October between 7:00 and 11:00 UTC, where all ensemble forecasts overestimate the rainfall rates, particularly for the 1-hour lead time forecast. This is expected to have limited impact on the anticipation of floods, since it affects only the end of the rainfall event. The shape and magnitude of the average observed rainfall hyetograph is well anticipated by the 1-hour lead time forecasts, while for the 6-hour lead time forecasts a time-shift of 2 hours is observed during the rising phase of the hyetograph, and lower average rainfalls are observed at the rainfall peak . Even if not totally satisfactory, the time-shift remains significantly lower than the lead time of 6 hours. This means that the rainfall forecasts are still helpful to anticipate the actually observed intense rainfall period, even if it is not perfectly positioned in time. This should result in an added value of the hydro-meteorological ensemble forecasts to anticipate the flood rising limbs, for this specific area and rainfall event.

As mentioned in section 2, the objective of this first step is not only to evaluate the overall quality of the rainfall forecast products, but also to define relevant space and time frames (HFA and HFT, respectively), which will help in illustrating the quality and usefulness of these products for flood forecasting. The selected space and time frames must include observed as

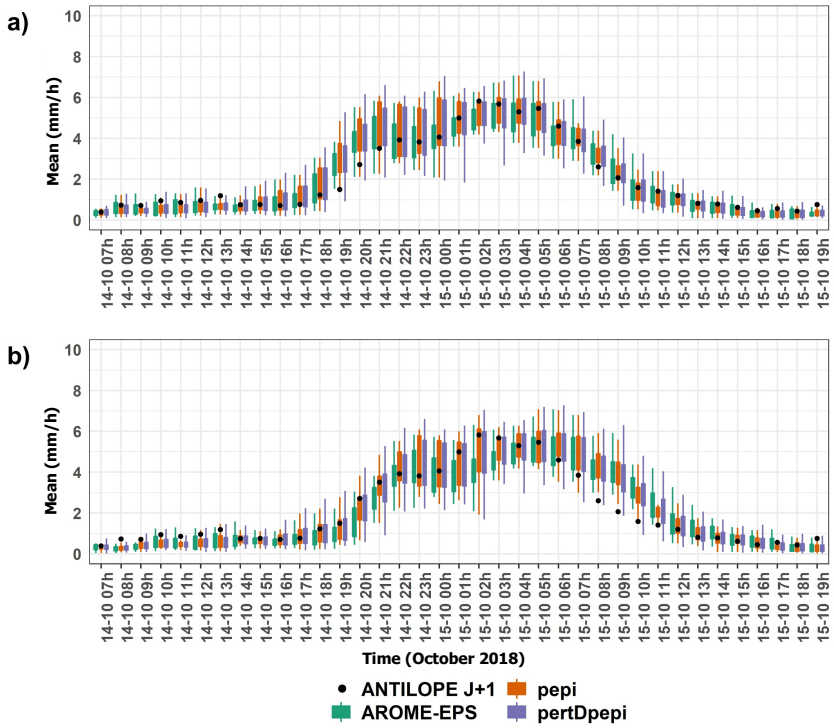

Figure 4. Temporal evolution of observed hourly rainfall rates (black dots), and a) 1 hour and b) 6 hours lead time ensemble forecasts (boxplots), during the October 2018 flood event. Rainfall rates are averaged over the Aude River basin (6074 km^2). The boxplots correspond to AROME-EPS (green), pepi (orange) and pertDpepi (purple). Whiskers reflect the min-max range and boxes the inter-percentile (25%-75%) for the forecasts.

well as forecast high intensities, but not many areas or time steps with low intensities in both observation and forecasts. These areas are of little interest to flood forecasting, and they may mask the main features of the forecasts when carrying out an event-based forecast quality analysis. Based on the hyetographs presented on Figure 4, the HFT was set hereafter from the 14th of October at 20h00 to the 15th at 10h00 UTC. This corresponds to a 15-hour period over which hourly observed and forecast average rainfall intensities on the Aude basin were larger than 2 mm/h. The choice of this threshold is relatively subjective and

just corresponds to a significant average rainfall intensity. Other threshold values could also have been selected. The choice of the HFA will be commented hereafter.

Figure 5 shows the spatial distribution of accumulated observed and forecast rainfall over the 15-hour HFT time window. For the forecast products, the ensemble mean, the 75% percentile and the 95% percentile at 1-h lead time are plotted. Note

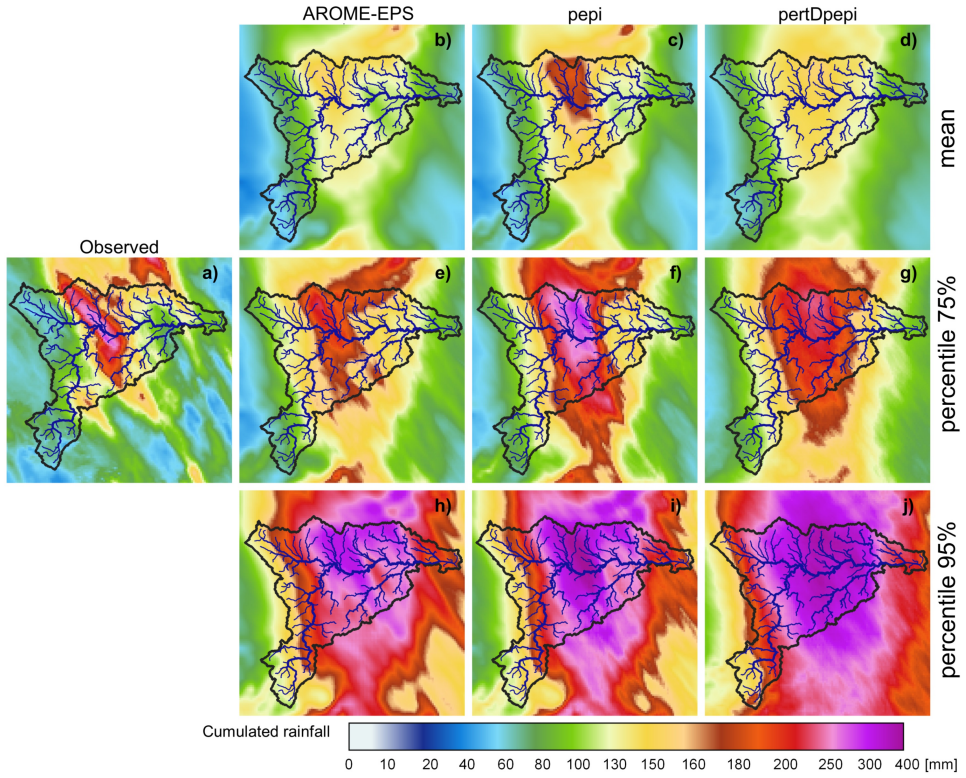

Figure 5. Comparison of observed and forecast (one hour lead time) rainfall amounts over 15 hours (from 14th of October 22 h to 15th of October 13 h) : a) observed rainfall, b) e) h) AROME-EPS ensemble mean, 75% and 95% percentiles, c) f) i) pepi ensemble mean, 75% and 95% percentiles, d) g) j) pertDpepi ensemble mean, 75% and 95% percentiles

320

that the forecast pannels do not correspond to rainfall accumulations for one unique run of forecast, but rather to a cumulated representation of the areas affected with high forecast rainfall intensities, for the successive forecasts issued during the event.

The figure shows that the area affected by the high rainfall accumulations (over 280 mm within 15 hours) is, in general, captured by the rainfall forecast products. The added value of the AROME-NWC model, generating locally more intense rainfall rates in the forecasts, can be seen by comparing the AROME-EPS and the pepi fields. The influence of the spatial shifts introduced in the pertDpepi product is also visible, when comparing it with the pepi product. It is also interesting to note that the ensembles means are not able to capture the magnitude of the highest observed rainfall accumulations. Only the tail of the ensemble distributions at each pixel (75% and 95% percentiles) can approach the observed intensities. To produce hydrological forecasts based on a good estimate of the rainfall rates over the high rainfall accumulation period and area, it may be necessary to work based on a high ensemble percentile value (a least the 75% percentile in the present case study; Figure 5.e-g), rather than on the ensemble mean value. However, for these percentiles, the area of high intensities spreads and becomes larger than the area seen in the observed field of rainfall accumulations.This may be attributed to the location errors of some members in the successive runs of forecasts. This behavior is particularly marked in the pertDpepi product, which is probably a combined effect with the spatial shift of the perturbations introduced by this product. Finally, Figure 5 also provides the required information to set the HFA. Even if not entirely hit by the observed heavy precipitation event, the Aude River basin is almost entirely covered with repeated high forecast rainfall intensities during the event, at least for the larger quantiles. This led to the choice of keeping the entire Aude River basin as HFA. Considering this whole area will help in evaluating the risks of false alarms attributed to rainfall forecast location errors when forecasting floods. But as for the HFT, the choice of the exact limits of the HFA remains relatively subjective, and other extents of the HFA could also have been selected.

The overall spread of the rainfall ensemble products can be further analyzed based on the rank diagrams presented on Figure 6. The diagrams are plotted for a lead time of one hour and pull together all the time steps within the HFT and the pixels of the HFA. Note also that they have been divided (i.e., stratified) into three sub-samples according to the observed hourly rainfall intensities. Overall, the U-shape of the rank histograms indicates a lack of spread of the three ensemble products (under-dispersive ensembles), although still moderate, when it comes to capture the observations within the ensemble spread at the right time and location. The histograms obtained for each sub-sample reveal a shift towards negative (resp. positive) bias, when the highest (resp. lowest) observed rainfall intensities are considered. This bias appearing in the rank diagrams when data sets are stratified based on observed values, even for perfectly calibrated ensembles, is a well-documented phenomenon (Bellier et al., 2017). Observation-based stratification should therefore be considered with caution in forecast quality evaluation. It is nevertheless interesting to note the influence of stratification in the bias revealed by the rank diagrams. It can be a consequence of the limited spread of the ensemble rainfall forecasts. But it may also reflect the shifts and mismatches in time and space of the ensemble forecasts, already illustrated on Figures 4 and 5.

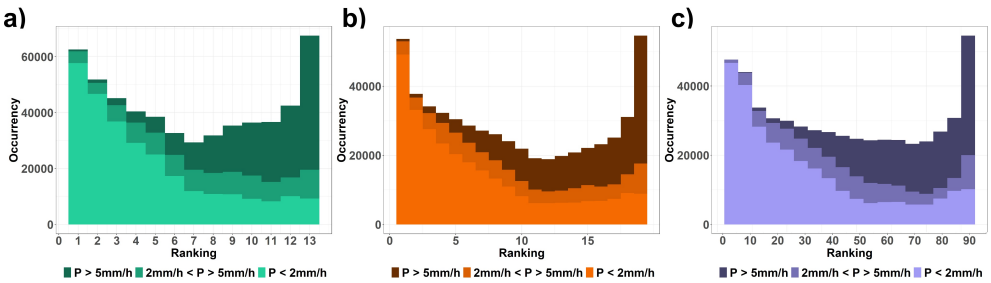

Figure 6. Rank diagrams of the three ensemble rainfall forecast products for rainfall rates under 2 mm/h, between 2 and 5 mm/h and above 5 mm/h and for a lead time of one-hour : a) AROME-EPS, b) pepi, c) pertDpepi.

4.2 Hydrological anticipation capacity

As mentioned in section 2, the hydrological forecasts are first evaluated on their ability to detect, with anticipation, the exceedance of pre-defined discharge thresholds. The objective is to provide a detailed and comprehensive view of the anticipation capacity for ungauged streams and main gauged rivers. The criteria and maps presented were computed based on the CINECAR hydrological model simulations and forecasts, and include the entire sample of 1174 sub-basins outlets. The hydrological simulations based on observed ANTILOPE J+1 rainfall are considered as the reference simulation (RS). The model is run from the 14 th October 07:00 and hourly rainfall accumulations are uniformly disaggregated to fit the 15-min time resolution of the model. The forecasts are issued every hour, by using the ANTILOPE J+1 rainfall up to the time of forecast, and one of the 3 rainfall forecast ensembles, or a zero future rainf scenario (RF0), for the 6 next hours. The hydrological model is first run for each member of the rainfall forecast to generate a hydrological forecast ensemble. From this ensemble, at each hydrological outlet in the HFA time series of forecast discharges are obtained for several probability thresholds.

The discharge thresholds correspond to the 10-year discharge return period, estimated by the SHYREG method (Aubert et al., 2014). This method provides flood quantiles estimates for all ungauged outlets with a drainage area exceeding 5 km^2. The choice of a 10-year return period was motivated by two main reasons: i) this is a discharge level for which significant river overflows and damages are likely to be observed on many streams ; ii) according to the RS scenario, during the October 2018 flood event about half of the 1174 sub-basins in the HFA were hit by floods with peak discharge exceeding this threshold. It is also important to remind that six runs of forecasts and all lead times are combined at each outlet for the computation of the scores presented in this section (see Appendix A): i.e. one exceedance detected in advance, for at least one of the 6 forecast runs issued just before the exceedance, is considered as a detection (hit). This means that one unique result (either a hit, a miss, a false alarm or a correct rejection) is obtained for each of the 1174 sub-basins and for each ensemble percentile.

Figure 7 illustrates the resulting ROC curves obtained for the hydrological ensemble forecasts based on the three rainfall products.

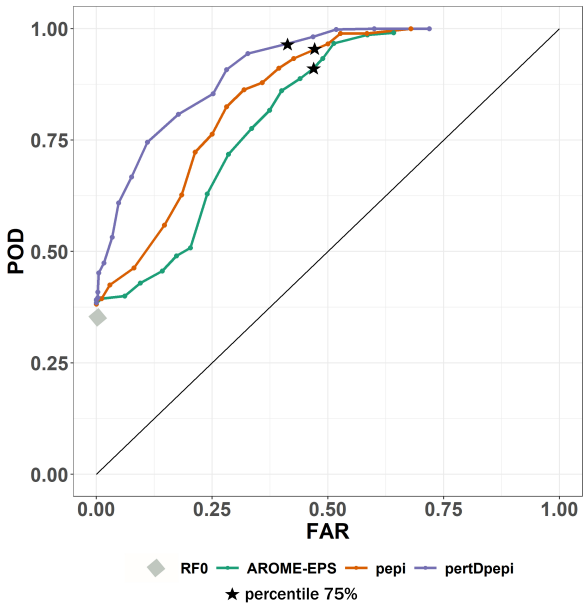

Figure 7. ROC curves summarizing the anticipation of exceedances of the 10-year discharge threshold, for the October 2018 flood event in the Aude River basin. The hydrological forecasts presented are based on AROME-EPS, pepi and pertDpepi rainfall ensemble products, and the CINECAR hydrological model. The black stars indicate the scores obtained for the 75% percentile of the hydrological ensemble forecasts. The points represent the scores obtained for the other percentiles, from 5% to 95%. The gray diamond shows the POD for the RF0 forecasts (zero-future rainfall forecast).

In the figure, the points related to the 75% percentiles of the ensembles are highlighted on each curve. The point obtained for forecasts using a zero rainfall scenario (RFO) is also presented. The added value of the pepi ensemble, compared to the AROME-EPS ensemble, was already observed in the forecast rainfall analysis (Section 4.1). It is also clearly visible here : for a selected ensemble percentile, both ensembles provide similar false alarm ratios, but the ensemble based on the pepi rainfall product has a higher probability of detection. For a spatial overview of the results, the maps of hits, misses, false alarms, and correct rejections based on the 75% percentiles are presented on Figure 8. In the central area of the Aude River basin, the missed detections for pepi ensemble are lower (26, Figures 8.c), when compared to the AROME-EPS ensemble (50,

Figures 8.b). This is in line with the better capacity of the pepi ensemble forecast product to capture the observed high rainfall accumulations (Figure 5). Overall, the use of the pertDpepi ensemble product leads to a reduction of the number of false alarms (255, Figure 8.d), when compared to the two other ensemble products (290 for AROME-EPS and 292 for pepi). In practice, the effects of the spatial perturbation introduced by the pertDpepi ensemble differ depending on the area and ensemble percentile considered : it leads to an increase of the area covered by high accumulated rainfalls when considering the highest percentile values of the rainfall ensembles, as clearly visible on Figure 5 for the 95% percentile, but it may also have a smoothing effect of the high rainfall values when considering intermediate ensemble percentiles. This may be an explanation for the reduction of false alarms in some areas for the 75% percentile of the flood forecasts based on the pertDpepi product (Figure 8).

The comparison of the ROC results of the ensemble hydrological forecasts with those of the RF0 scenario - i.e. future rainfall set equal to zero - helps to further evaluate the added value of the rainfall forecast products. All ensemble forecasts lead to an increase of the number of hits (Figure 8) and of the probability of detection (Figure 7). However, this gain is obtained at the cost of an increase in the number of false alarms. When we look at the 75% percentile of the ensembles (black stars), close to 50% of the sub-basins of the Aude River basin where the 10-year discharge return period is not exceeded are associated with a false alarm.

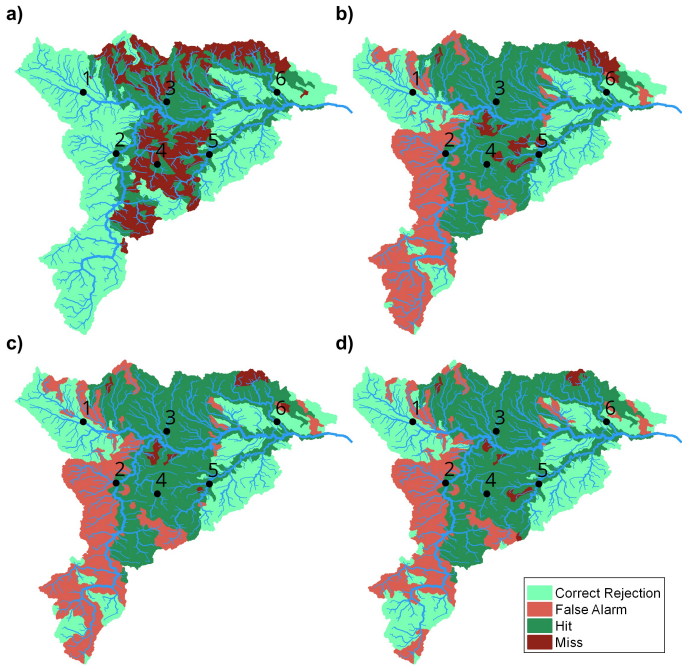

Figure 8. Maps illustrating the detailed anticipation results (hits - misses - false alarms - correct rejections) of the 10-year return period discharge threshold, for the hydrological forecasts based on : a) RF0 scenario, b) AROME-EPS ensemble (75% percentile), c) pepi ensemble (75% percentile), d) pertDpepi ensemble (75% percentile). Hits are displayed for anticipation times exceeding 15 min. Black dots represent the outlets where a detailed analysis of the forecasts is proposed in Section 4.3.

395 Figure 9 shows the distribution of anticipation times, within categories ranging from 15 min to 6 hours, for sub-basins with hits in Figure 8. Note that the anticipation time is defined here as the difference between the time of exceedance of the discharge threshold by the RS hydrograph, and the time of the first run of forecast that detects the threshold exceedance (see Appendix A). The Figure also shows the number of sub-basins where either no anticipation (i.e., those with misses) or false alarms are observed, and confirms the observations already made on Figure 8: lower number of false alarms for the ensemble 400 forecasts based on the pertDpepi rainfall product, when compared to AROME-EPS and pepi ; and higher number of sub-basins associated with misses for forecasts of the RF0 reference scenario, as expected. But the results presented on Figure 9 also confirm the added value of using the rainfall ensemble products to increase anticipation times of the exceedance of the 10-year

discharge threshold: the anticipation times remain mostly between 15 min and 1 hour for RF0 forecasts, whereas they exceed 2 hours for more than a half of the sub-basins in hit with the three ensemble rainfall products. The use of the rainfall ensemble forecasts is key in the present case study to extend anticipation times beyond 2 hours.

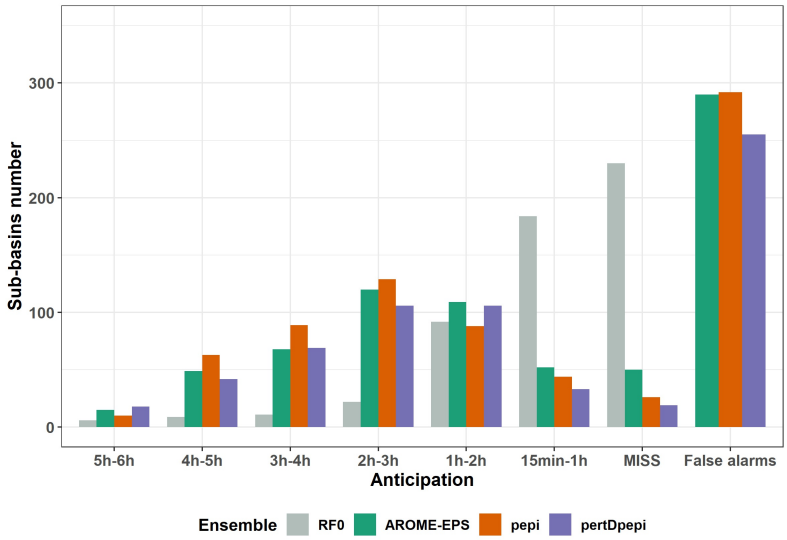

Figure 9. Comparison of 10-year threshold anticipation times for the hydrological forecasts based on the RF0 scenario and the three rainfall ensemble products. For the hydrological forecasts ensembles, the 75% percentile is considered.

Figure 9 illustrates the necessary trade-off between gaining hits and larger anticipation times, and increasing false alarms as a consequence, when using ensemble hydrological forecasts to anticipate discharge threshold exceedances. Is it worth when compared to the RF0 forecasts (no future rainfall), where false alarms are avoided, but the number of misses is increased, and the anticipation times are lower, as in the case of the 2018 event? The answer to this question depends on the end-users of the forecasts and their capabilities to respond to flood warnings. Further considerations on the cost and benefits associated with the hits and increased forecasting anticipation times, but also with the misses and false alarms, would be needed to fully address the question. For instance, in our case study, if we summary the contingency tables used to draw figures 8 and 9 by using the Percent Correct (PC) score (Wilks, 2011) - i.e. the fraction of the N forecasts for which a forecast correctly anticipated the event (hit) or the no-event (correct rejection) - then the conclusion would be that the RF0 forecast outperforms the three ensemble hydrometeorological forecasts, at least when the decision is based on their 75% percentile. This is because the PC scores are of

80% for RF0, 71% for AROME-EPS, 73% for pepi and 77% for pertDpepi. However, this score gives the same cost or benefit to misses, false alarms, hits and correct rejections. Additionally, it does not consider the added value of the increased anticipation of the threshold exceedances. This shows that an in-depth evaluation of hydrometeorological forecasts could strongly benefit from scores and metrics taking into account the end-user's constraints and needs, as much as possible. This would help to measure the actual cost/benefit balance of using hydrometeorological forecats in real-time operations. Nevertheless, from the evaluation framework proposed here, we were able to identify that the hydrometeorological ensemble forecasts can offer better anticipation times than the RF0 forecasts, although this anticipation has a cost, i.e. an increase in the number of false alarms, which the user would need to be able to accept.

4.3 Detailed hydrological forecasts evaluation at selected outlets

To complement the assessment results, a detailed analysis of the hydrometeorological forecasts is carried out at six selected gauged outlets, covering a variety of situations (see Figure 8 for their location).

The results presented hereafter are based on the GRSDi model simulations and forecasts. The choice of this model is driven by the possibility to compare the forecast hydrographs to both the simulated (RS scenario) and the observed hydrographs at gauging stations, with a hydrological model which was not specifically calibrated to the 2018 flood. This allows us to compare errors related to the rainfall forecasts to those related to the modelling errors. Figures 10 to 15 present, for each of the six selected outlets, three panels of results. The first one on the left shows the spread/skill score of discharge forecasts for all lead times (one to six hours) . The spread is the average of the standard deviation of ensemble forecasts. It is divided by the skill, which is the RMSE of the ensemble mean, to obtain the spread/skill score. A value of 1 means that spread and skill are equivalent. The other panels allow us to visually inspect the shape and spread of the forecast hydrographs (mean and quantiles of the ensemble forecasts) for an intermediate lead time of three hours, and to compare the forecasts to the simulated (RS) and observed discharges. Only pepi (middle panel) and pertDpepi (right panel) forecast hydrographs are presented (the AROME-EPS hydrographs are provided in the Appendix B). It should be noted that the "forecast hydrographs" are represented for a fixed three hours lead time, and are not continuous discharges series in time as in the case of the observed and simulated hydrographs. They rather represent what could be predicted by the forecasts issued three hours before each time step t of the flood event.

Outlets 1 and 2 (Figures 10 and 11) correspond to weak hydrological reactions - i.e. the peak discharge of the reference simulation (RS) remains largely below the 10-year discharge threshold in both cases. At outlet 1, the previous step, i.e. the evaluation of discharge threshold anticipation capacity during the flood rising limb, showed a correct rejection, whereas at outlet 2 a false alarm was obtained with the three ensemble forecast products. Outlets 3 and 4 (Figures 12 and 13) correspond to small size watersheds, located in the most intense observed rainfall area. For both outlets, the 10-year threshold exceedance could not be detected with significant anticipation with the RF0 forecast scenario (15 min of anticipation time for outlet 3, and 0 min -i.e. miss - for outlet 4), whereas anticipation times of 3h15min were obtained with the 3 ensemble forecast products. Outlets 5 and 6 (Figures 14 and 15) represent larger watersheds. At these outlets, the 10-year threshold exceedance could be anticipated with the RF0 forecast scenario (anticipation times of 45 and 60 min respectively). However, these outlets also

show an increase of the anticipation capacity with the ensemble rainfall forecasts: anticipation times ranging from 2h45min to 4h45min at outlet 5 depending on the rainfall forecast product, and of 6 hours at outlet 6 for the three products.

For outlet 1, the simulated (RS) hydrograph is overall correctly retrieved by the forecasts (Figure 10). For outlet 2, the forecast hydrographs show a larger dispersion and a significant overestimation, in particular for the upper percentiles (Figure 11). This overestimation is present in the forecasts based on the three ensemble rainfall products. It explains the false alarm observed on Figure 8, when the percentile 75% is considered. It is directly related to the overestimation of rainfall in areas surrounding the actually observed high rainfall accumulations area, as identified on Figure 5. The spread/skill relationship scores (Figures 10 a) and 11 a)) show that the hydrological ensemble forecasts tend to have a higher spread/skill score for the first lead times in all rainfall products. It becomes close to one after 2 to 4 hours of lead time for the hydrological forecasts based on AROME-EPS and pepi rainfall products (i.e., the spread correlates well with the errors). For the forecasts based on pertDpepi, the score tends to stabilize around 1.5 (i.e. there is 50% more spread than ensemble mean skill in these forecasts), highlighting a shortcoming of potentially over - or unnecessary - wide spread in this ensemble product. It might be caused by an excessive spatial shift with respect to the geographical size of the investigated catchment. The larger dispersion of the pertDpepi forecast ensemble is confirmed by the hydrographs, and is in particular visible for the 5%-95% percentiles inter-quantile range, while the range defined by the 25% and 75% percentiles show limited difference with the respective inter-quantile range in the pepi ensemble product. This confirms the tendency of pertDpepi rainfall product to modify differently the extreme percentiles and the intermediate ones when compared to pepi. In practice, the over-dispersion observed for extreme percentiles could have limited the performance of the pertDpepi ensemble if these extreme percentiles had been used by a user. This was also visible on Figure 7: a large increase of false alarm rates with the pertDpepi ensemble product in the upper part of the ROC curve.

Lastly, the comparison of observed and simulated hydrographs shows a different performance of the hydrological model between outlets 1 and 2. The modelling errors are limited in the case of outlet 1, whereas the observed discharges are overestimated by the model in the case of outlet 2. In this second case, the comparison between observed, simulated and forecast hydrographs shows that a large part of the total forecast error can be attributed to modelling errors.

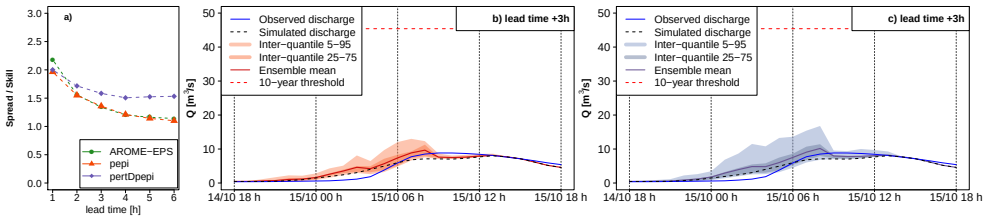

Figure 10. Detailed hydrological forecasts evaluation at outlet 1 (216 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging station.

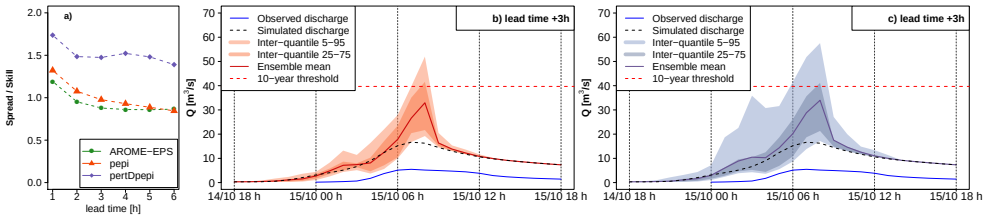

Figure 11. Detailed hydrological forecasts evaluation at outlet 2 (197 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging stations.

In the cases of outlets 3 and 4 (small size watersheds in the most intense observed rainfall area - Figures 12 and 13), the RS hydrograph largely exceeds the 10-year discharge threshold at the very beginning of the flood rising limb. The 3-hour forecasts also exceed the thresholds and anticipate well the initial increase of the river discharges. This results in the hits presented on Figure 8, and in the good anticipation times, exceeding 3 hours for both outlets. A forecast user could be satisfied with such information (anticipation of threshold exceedance) to start a flood response. For both outlets, the forecast hydrographs show a delay in the flood rising limb and flood peak, comparatively to the simulated hydrographs. This means that shortcomings are present in the rainfall forecast products in terms of the dynamics of the time evolution of the rainfall event. These shortcomings could not be directly captured in the previous evaluation scores, but might be an important feature for forecast users. The flood rising limb is overall better represented in the case of outlet 3. At outlet 4, the 3-hour simulated hydrograph presents a large delay comparatively to the simulated hydrographs (3 hours between observed and 75% percentile 3-hour forecast peak discharges). This delay comes with a significant underestimation of the flood peak magnitude by the ensemble mean and 75% percentile, although the 95% percentile is closer to the simulated discharge peaks. This goes along with the general tendency of the rainfall ensemble means to underestimate the higher rainfall accumulations (Figures 5 and 6).

For both outlets, the spread/skill relationship scores reflect the influence of the temporal shifts between the simulated and forecast hydrographs. At outlet 3, where forecasts resemble better the simulated hydrographs, the skill-spread relationship is close to one for the ensembles based on AROME-EPS and pertDpepi, which can be interpreted as a sufficient variance of the forecast spread to cover the errors. The ensemble based on pepi, on the contrary, has 1.5 times more spread than skill at 3 hours of lead time. At outlet 4, the disparity between forecast and simulated hydrographs is too high and, the large forecast spread, as seen in the hydrographs, is not enough to cover the timing errors. At both outlets, and comparatively to outlets 1 and 2, we also note that the spread-skill relationship score almost does not change with forecast lead time, except for pepi-based forecast at the longer lead times.

The comparison of observed hydrographs with simulated and forecast hydrographs also provides interesting information about the relative importance of modelling and forecasting errors. At outlet 3, the observed and simulated hydrographs have a different shape, whereas the simulated and forecast hydrographs appear similar despite a time delay. At outlet 4, observed and simulated discharges resemble (despite a time delay), but both are very different from the forecast hydrograph. This indicates

that modelling errors are a major source of uncertainty at outlet 3 (if we consider that observations represent well the true flow values), comparatively to rainfall forecast errors, while the opposite is seen at outlet 4.

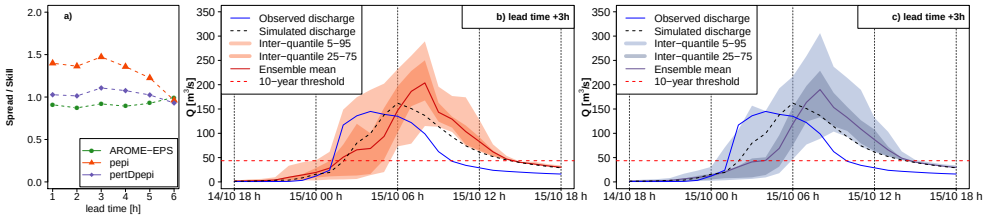

Figure 12. Detailed hydrological forecasts evaluation at outlet 3 (85 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging station.

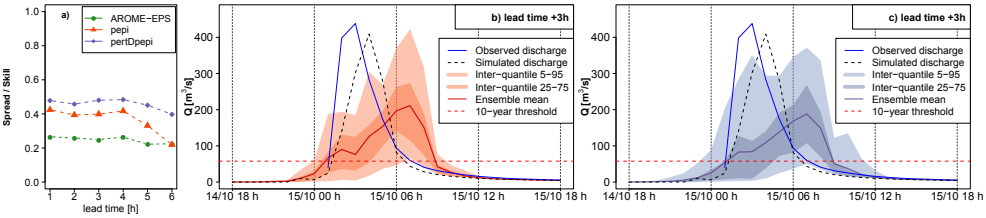

Figure 13. Detailed hydrological forecasts evaluation at outlet 4 (173 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging station.

Finally, Figures 14 and 15 present the forecast results obtained for outlets 5 and 6, corresponding to outlets at which significant anticipation is possible even without rainfall forecasts (RF0 scenario, see Figure 8.a). It can be seen that the hydrological reactions at these outlets are important, as the 10-year discharge threshold is exceeded by the the RS hydrograph in the intermediate phase of the flood rising limb, although not as much as for the outlets 3 and 4, where the discharges largely exceeded this threshold. This is explained by the location of these catchments, at the limits of the most intense observed rainfall area. Overall, the forecast hydrographs are less dispersed than in the preceding cases, and the differences between the three rainfall forecast products is less evident, with only a slightly larger spread for the forecasts based on the pertDpepi ensembles. This lower difference between products can be a direct result of the larger influence of flood propagation in the hydrological forecast, limiting the influence of rainfall forecast uncertainties in comparison with the former cases. However, some shortcomings of the rainfall forecasts are still visible in both cases. At outlet 5, the flood rising limb of the simulated hydrograph is well captured by the 25%-75% inter-quantile range of the 3-hour ahead forecasts, but the forecasts display a delay of about 2 hours in terms of timing of the peak discharge. This time shift of forecasts may explain the misses that appear in the upstream parts

of the catchment (Figure 8). At outlet 6, the forecasts tends to slightly overestimate the peak of the simulated hydrograph, although the 5% quantiles remain close to the simulated peak.

The influence of modelling errors also clearly appears here. At outlet 5, simulated and observed hydrographs are close to each other during the main peak, but the first peak of the observed hydrograph is not represented in the simulation. At outlet 6, the simulated discharges follow the same pattern as the observed discharges, but the magnitude of the peak discharge is underestimated by the model. In both cases, the modelling errors are partially compensated in the forecasts, under the condition that a user looks at the higher quantiles of the ensemble forecasts, as the 5%-95% inter-quantile range of the 3-hour ahead forecasts gets closer to the observed peak discharges. Overall, these two outlets mainly illustrate how the evaluation of hydrological forecasts can be influenced also by the ability of the rainfall-runoff model to correctly represent the flood dynamics, when the evaluation is achieved based on actual flow observations.

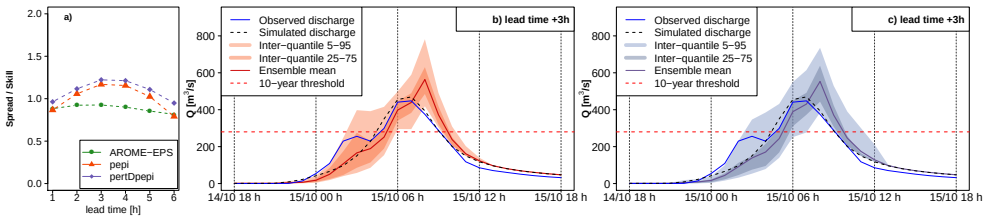

Figure 14. Detailed hydrological forecasts evaluation at outlet 5 (263 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging station.

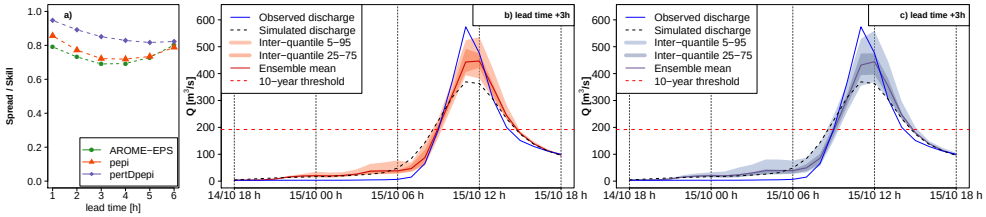

Figure 15. Detailed hydrological forecasts evaluation at outlet 6 (257 km^2): a) spread/skill relationship, b) hydrographs of 3 hours lead time forecasts with pepi, c) hydrographs of 3 hours lead time forecasts with pertDpepi. In blue, the observed hydrograph at the gauging station.

5 Discussion

5.1 Added value and limitations of the implemented evaluation framework

After an initial analysis of the performance of the three rainfall forecast products, the evaluation procedure proposed in this paper focused on the capacity of the hydrological forecasts to anticipate the exceedance of selected discharge thresholds. Several authors have suggested using thresholds and contingency tables to perform a robust regional evaluation of hydrological ensemble forecasts (Silvestro and Rebora, 2012; Anderson et al., 2019; Sayama et al., 2020). In this study, a step further is proposed towards (i) focusing only on the rising limb phase of the flood hydrograph - i.e. the most critical phase in terms of anticipation needs for preparedness and emergency response, as already suggested by Anderson et al. (2019), and (ii) splitting the analysis into a geographical point of view (discharge threshold anticipation maps) and a temporal point of view (histogram of anticipation times). The illustrative example of the October 2018 flood in the Aude river basin in France showed that this approach provides a synthetic and meaningful information about the possible gains in anticipation compared with the RF0 reference forecast where future rainfalls are considered as null. It also informs about the relative performance of the evaluated rainfall forecast ensemble products, even if the products appeared relatively similar at first sight in the example showed here. This analysis can be easily adapted to other flash flood events with different characteristics, or even to other types of floods. For instance, different discharge thresholds can be set to better distinguish the areas where the main flood reactions were observed, a larger threshold can be applied to the anticipation time to decide for a hit (a minimum threshold of 15 minutes of anticipation time was used in this study), and different ensemble percentiles can be selected to better illustrate the role that the probability threshold plays when evaluating different forecast products.

The threshold anticipation maps are also helpful for the selection of outlets where hydrographs should be analyzed in more details in the last phase of the evaluation. They avoid focusing only on outlets where large anticipation is possible without rainfall forecasts (RF0 scenario). The differences between the hydrological forecasts based on rainfall ensemble products may be more difficult to identify at these outlets, because of the large influence of flood propagation in the hydrological forecasts. The threshold anticipation maps also allow selecting outlets corresponding to a variety of situations (i.e. hit/miss/false alarm/correct rejection), which may help in bringing up details about the complex features of the hydrological ensemble forecasts.

If the proposed approach has several advantages, it also relies on several methodological choices which could be adapted or improved. It can be noted that only the runs of forecasts covering the most critical phase of the event (i.e. the time of the threshold exceedance, or the maximum of the RS hydrograph) were considered to build the contingency tables. The results obtained could differ if other runs of forecasts and/or other phases of the event were considered. Particularly, in case of events with multiple flood peaks, the fact that the method focuses only on the first threshold exceedance can be seen as a limitation. In such a situation, each rising phase of the flood event could be examined separately, even if in some cases the different phases of the floods may be difficult to separate. An alternative could be to analyse the anticipation of threshold exceedances for each run of forecast during the event, independently of the times the thresholds are exceeded for the RS hydrographs. Providing an evaluation for multiple flood events, or for multiple thresholds for the same flood, may also be interesting complements to the

results presented here. This may nevertheless cause some difficulties in setting the HFT and HFA, which are event specific. The HFA may also require to be adapted to the considered threshold. To avoid changing the limits of the HFA, an option could be to use a score that does not account for Correct Rejections, and therefore would be less sensitive to the extent of the HFA, such as the Critical Success Index. In addition, the procedure used for the computation of the contingency tables and the anticipation times can be considered as optimistic, at least in some situations. At first, as explained by Richardson et al., 2020, having a sequence of several consistent forecasts is often a desired quality in forecasting systems. Specific tests of the consistency of forecasts have been proposed (Ehret and Zehe, 2011; Pappenberger et al., 2011). In this study, only the first time a forecast exceeds the discharge threshold value was considered to conclude to a hit (or a false alarm) and to compute the anticipation time, without considering the consistency of several successive forecasts. This can be seen pertinent when dealing with flash floods, when it is often not possible to wait for forecast consistency in real-time to activate flood response actions. However, it could also be interesting to use the same method and to consider the number of runs correctly anticipating the threshold exceedances (or non-exceedances) to fill the contingency table. Moreover, since two of the three rainfall products used in this study are still experimental (pepi and pertDpepi), and are not yet included in the real time production workflow of Météo-France, it was not possible to integrate the actual delivery times of the forecasts in this study. To obtain a more realistic view of the actual anticipation capacity, the delivery times of the different rainfall forecast products need to be integrated in the computation of the contingency tables and anticipation times, as proposed by Lovat et al. (2020). This means that the hydrological ensemble forecasts should be considered only from the delivery time of the rainfall forecast products used as input, instead of from the beginning of the sequence of time covered by the rainfall forecasts. This integration would probably lead to a reduction of the number of hits and of the anticipation times obtained with the hydrological ensemble forecasts in this study, whereas it would not affect the results obtained for the RF0 forecast, where no future rainfall scenario is used.

Lastly, it should also be noted that the last step of the evaluation analysis proposed in this study, which is based on the whole hydrographs and spread and skill scores, remains essential to bring a detailed view of the strengths and limits of the different ensemble products evaluated. Particularly, the shortcomings of an ensemble with a too large dispersion (which is the case of the pertDpepi product) and the impact of temporal shifts on the quality of the forecasts would have been more difficult to identify without this last step of the analysis. The analysis of hydrographs also provides detailed explanations for the hits/misses/false alarms/correct rejections observed on the discharge threshold anticipation maps.

5.2 Performance of the three ensemble forecast products

The contribution of the rainfall forecast ensemble products to the anticipation of discharge thresholds, with respect to the scenario where no future rainfall forecast is available (RF0 scenario) was clearly illustrated here. This contribution lies mainly in the high reduction of the number of misses on the threshold anticipation maps (Figure 8), and in a significant gain in the anticipation times for almost all sub-basins (Figure 9). It also depends on the ensemble probability (or percentile) that is considered to define a forecast event (in our case, the 75% percentile). The use of ensemble forecasts was essential to obtain anticipation times greater than one hour when a hit was observed, even if this positive effect was counterbalanced by the occurrence of a high number of false alarms in other areas of the catchment. This implies that, in order to conclude about

the added value of ensemble forecasts, it is necessary to consider the balance between the actual gains and costs for a given user of the forecasts (Verkade and Werner, 2011). Hydrometeorological forecasting is no exception to the well-known decision principle: the larger the uncertainties (i.e. the lower the sharpness) affecting the predictions of the variable(s) on which the decision is based, the larger the safety margin will be to guard against the risk of unwanted consequences. Particularly in the case of flash floods, the user is often interested in reducing misses and increasing hits and anticipation times. This means that users need to consider how much of a false alarm rate they can handle, and how much risk they can take (i.e., what percentile of the ensemble distribution they should consider). These are, however, difficult questions to answer. More flood event evaluations are needed to enhance our understanding of these trade-offs.

In the present case study, the pepi rainfall ensemble product was built from the AROME-EPS product by adding one to six additional members (depending on the lead time). This strategy had globally a positive influence on the quality of the ensemble forecasts, with a significant decrease in the number of misses for an equivalent number of false alarms (Figure 9). This can be directly related to the increased reliability of the pepi forecasts compared to the AROME-EPS forecasts, and therefore the better capture of the high rainfall accumulations periods (Figure 4) and extension (Figure 5) when considering the 75% percentile. Finally, in the case of this event, the pepi product provides an added value for the characterization of the main intense rain cell, without a significant degradation of the ensemble performance in the surrounding areas, where the rainfall accumulations were less intense. However, this overall improvement does not result in a significant increase of the anticipation times obtained with the pepi product compared to the AROME-EPS product (Figure 9).

Adding spatial perturbations in the four cardinal directions to the members of the pepi ensemble led to an increase of the number of members in the pertDpepi ensemble product, which varies from 65 to 90 members depending on the lead time. Overall, this resulted in a larger dispersion of the ensemble, particularly visible on the extreme percentiles (5%-95%), and led to positive as well as negative effects, both visible in the evaluation results. The high rainfall accumulation period is still well captured (Figure 4), but the maximum rainfall area is not as well localized, comparatively to the observations, as for the pepi ensemble (Figure 5). Regarding the discharge threshold anticipation maps (Figures 8 c) and d)), the reduction of the number of false alarms with pertDpepi is mainly due to the choice of the 75% percentile, which is not modified in the same way as higher percentiles when compared to the pepi ensemble. This is confirmed by the hydrographs analyses (Figures 10 to 15), which overall confirm the over-dispersion of pertDpepi for higher percentiles of the ensemble distribution.

Moreover, the anticipation times are slightly reduced with pertDpepi when compared to the two other ensemble products (Figure 9): 221 (over 1174) sub-basins show an anticipation time of more than two hours, against 245 for AROME-EPS and 259 for pepi. Again, the detailed hydrographs analyses (Figures 10 to 15) confirm that pertDpepi does not efficiently compensate the time shifts in the representation of hydrographs. The addition of spatial perturbations leads to a significant increase in the number of members, which significantly increases the computing times with the pertDpepi rainfall ensemble product (up to 72 additional members compared to pepi), but there is no clear benefit in adding these members in the case evaluated here.

6 Conclusions

In this paper, we proposed and tested a methodological framework for the evaluation of short-range flash-flood hydrometeorological ensemble forecasts at the event scale. Focusing on the evaluation of a single flood event is necessary because of i) the relatively low frequency of flash floods occurrences, and ii) the limited length of available re-forecast periods for newly developed ensemble rainfall forecast products based on convection permitting NWP models. In order to enhance the value of a single-event evaluation, the point of view of an end-user of the forecasts is adopted in the proposed methodological approach. This approach combines different steps. Firstly, it comprises a preliminary analysis of rainfall observations and forecasts, which provides a first overview of the quality of the ensemble rainfall products and information needed to select the most relevant ensemble percentile and spatial and temporal scales for the analysis. Secondly, a hydrological forecast analysis evaluates the detection of discharge threshold exceedances and the corresponding anticipation times. It considers the flood rising limb of the hydrographs, and illustrates the gains in anticipation comparatively to a reference scenario where zero rainfall is forecast for the future. Thirdly, a detailed analysis of forecast hydrographs and spread/skill scores at outlets selected based on the former step is carried out. It aims to better illustrate the skills and the limits of the evaluated hydrological ensembles. This methodology was tested and illustrated on the major flash-flood event that hit the Aude River basin in South-Eastern France in October 2018. Three ensemble rainfall forecast products recently developed by Météo-France were compared, i.e. the AROME-EPS ensemble (12 members), the pepi ensemble combining AROME-EPS and AROME-NWC members (13 to 18 members), and the pertDpepi ensemble adding spatial perturbations to the pepi ensemble (65 to 90 members).

The study illustrated the multifaceted and complex issue of evaluating single flood events. It confirmed the interest of the proposed evaluation procedure to help forecast users to efficiently define beforehand the space and time windows to be used for the forecast performance evaluation. This was achieved by comparing the observed and forecast rainfall hyetographs and accumulations over the entire catchment area (step 1). The analysis of discharge threshold exceedance detection (step 2) contributed to illustrate the geographical features of the flood event and the anticipation capacities offered by each of the rainfall ensemble products. The comparison with a zero rainfall future scenario confirmed the gains in anticipation offered by the ensembles. From the threshold anticipation maps, it was also possible to select outlets that correspond to a variety of situations of interest to a forecast user (hits, misses, false alarms, correct rejections). These outlets could then be investigated in more details, based on observed, simulated (with observed rainfall as input to the hydrological model) and forecast hydrographs (step 3). This last phase illustrated further strengths and limitations of the different rainfall forecast products, including the shifts in timing that affect forecast skill and the main sources of forecast errors (rainfall forecasts or hydrological model).

The evaluation finally led to balanced conclusions about the three products of rainfall ensemble forecasts. These products clearly enhance the anticipation times when compared to a zero-rainfall future scenario, but tend to overestimate the spatial extent of the area of the highest rainfall accumulations. This results in a higher risk of overestimating flood threshold exceedances in the surroundings of this area (i.e., an increase of false alarms). The actual added value of the evaluated forecasts for an end-user therefore depends on the relative benefits of the increased anticipation and costs of the false alarms. However, the results also showed a hierarchy between the three evaluated products, and a clear added value of the pepi ensemble product could be

identified. According to the significant time shifts of forecasts hydrographs in several sub-basins, which seem to significantly limit the anticipation times, the introduction of temporal perturbations could probably be an interesting alternative to be tested

Finally, even if evaluating ensemble hydrological forecasts based on a single flood event remains a very challenging issue due to the limited statistical representation of the available data, single event evaluations are needed and important to advance operational flood forecasting systems. The evaluation framework proposed in this study could be helpful in drawing rapid and meaningful analyses about the interest of newly developed rainfall ensemble forecast approaches.

Data availability

Author contributions The initial idea was proposed by MCN, OP, DP and MHR. The initial version of the paper was written by MCN and DP with a general contribution of EG, MHR, and OP, and specific contributions of HM and FB for the description of the AROME-based short-range rainfall ensemble forecast products, and of PN for the rainfall-runoff modelling . AF generated the pepi and pertDpepi short-range rainfall ensemble forecast products. MCN performed the Cinecar rainfall-runoff modelling, and used the calibration results obtained by PN and OP for the Aude 2018 event. DP performed the GRSDi rainfall-runoff modelling.

Competing interests The authors declare that they have no conflict of interest.

Acknowledgements Peak discharge data for the Aude 2018 flood event were obtained as part of the HyMeX research program (http://hymex.org), with financial support from the MISTRALS program of the CNRS and the Ministry of Ecological and Solidarity Transition (DGPR/SCHAPI). Rainfall observation and forecasts data were provided by Météo-France.

Financial support This research is part of the PICS research project (https://pics.ifsttar.fr) and has been supported by the Agence Nationale de la Recherche (grant no. ANR-17-CE03-0011).

675 **Appendix A: Methodology used for the evaluation of discharge thresholds anticipation with hydrological ensemble forecasts**

The methodology used to evaluate the anticipation of the exceedances of selected discharge thresholds was directly inspired by the principle of the ROC curves method developed by Mason (1982); Wilks (2011). The ROC curve (Relative Operating Characteristics diagram) is a criterion evaluating the capacity of a forecasting system to detect the occurrence (reciprocally 680 the non-occurrence) of an event. In hydrology, the event generally corresponds to the exceedance of a discharge threshold, and the ROC curve is generally built to summarize long time series of forecasts at one single basin outlet. Comparing at each time step the respective position of the forecast (for a given probability and a fixed lead time), the reference observation or simulation, and the considered discharge threshold, permits to calculate the probability of detection (POD) and the false alarm rate (FAR) (Wilks, 2011; Jolliffe and Stephenson, 2012). The ROC curve is then plotted by using the pairs of points (FAR; 685 POD) corresponding to different forecast probabilities.

In this study, the conventional computation of a ROC curve was adapted following the idea to focus the analysis on the first exceedance of a selected alert discharge threshold, at the event temporal and geographical scales. This led to introduce the following evolutions:

- all the ungauged outlets located in the HFA (1174 in this study) are considered to build the contingency table. For this 690 purpose, hydrographs simulated with observed rainfall are used as reference hydrographs (RS hydrographs hereafter), and the discharge thresholds are adapted at each outlet based on a given return period.

- only one unique time step for each outlet is considered, i.e. the first time at which the RS hydrograph exceeds the considered threshold,or the time of peak of the RS hydrograph, in case of no threshold exceedance. The use of many outlets (1174 in this study) compensates the small number of considered time steps for the elaboration of the ROC curves.

695 – all the forecasts delivered before and covering the considered time step are merged to evaluate the anticipation/detection of the threshold exceedance, without considering a unique lead time. In practice, we consider that a threshold exceedance is forecasted if at least one of the merged forecasts hydrographs exceeds the threshold. This neighborhood method is in the line of the methods discussed in Schwartz and Sobash (2017), the difference being that here it is a temporal neighborhood by max function that is considered.

700 According to these general principles, for a given forecast probability, the contingency table is obtained in the following way:

- If the threshold is exceeded by the RS hydrograph (fig A1), the date (day and hour) of the first exceedance is identified. All the runs of forecasts issued before and covering this date are then selected (i.e. 6 runs here according to the maximum forecast lead time of 6 hours and the 1 hour time step between successive runs) . A hit is counted in the contingency 705 table if at least one of the six runs exceeds the discharge threshold at any lead time (fig A1 a)), and a miss is counted if none of the six hydrological forecasts exceed the threshold at any lead time (fig A1 b)).

– In case the discharge threshold is not exceeded by the RS hydrograph (fig A2), the date of the peak discharge of the RS hydrograph is first identified. The runs of forecasts issued before and covering this date are then selected. Again, a false alarm is counted if at least one of the selected runs exceeds the discharge threshold at any lead time (fig A2 a)), and a correct rejection if none of the selected runs exceed the discharge threshold at any lead time (fig A2 b))

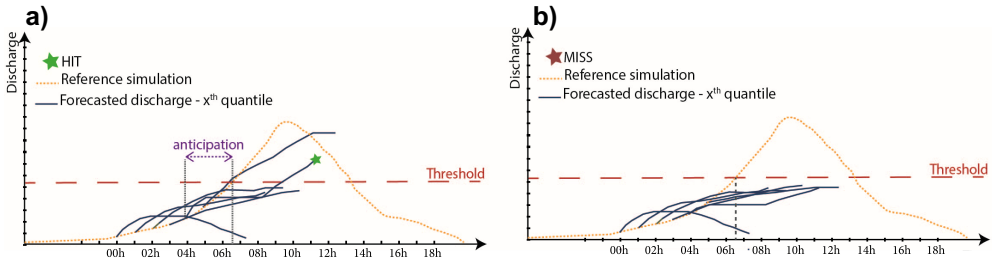

Figure A1. Illustration of the discharge threshold exceedance detection method: a) Hit case and b) miss. The way the anticipation is computed is also illustrated (time difference between the time the discharge threshold is exceeded for the RS hydrograph, and the time of the first run which detects the exceedance).

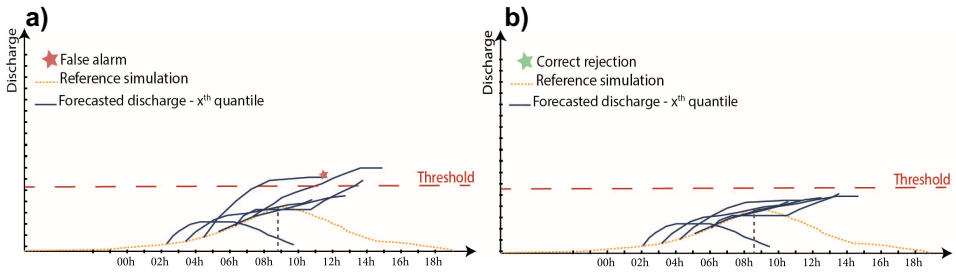

Figure A2. Illustration of the discharge threshold exceedance detection method: a) False alarm case and b) correct rejection.

Each contingency table finally contains as many values (hit, misses, false alarms, correct rejection) as the number of outlets in the HFA. The ROC curve can then be drawn by deriving the POD and FAR scores from the contingency tables, following the traditional way described above. An advantage of this way of building contingency tables, is that each table can be presented on a map, by drawing the evaluation obtained at each outlet or sub-basin (hit, miss, false alarm or correct rejection). This permits to easily analyze how the ensemble forecasts performed geographically according to these criteria.

Additionally, for each hit entry of the contingency tables, the anticipation time can also be evaluated (fig A1 a)). It corresponds to the difference between (i) the time of exceedance of the discharge threshold by the RS hydrograph and (ii) the time of the first forecast run that detects a threshold exceedance. We can thus draw the distribution of the anticipation times obtained over the considered outlets, for the different percentiles and ensemble forecast.

Appendix B: Forecast hydrographs obtained for AROME-EPS

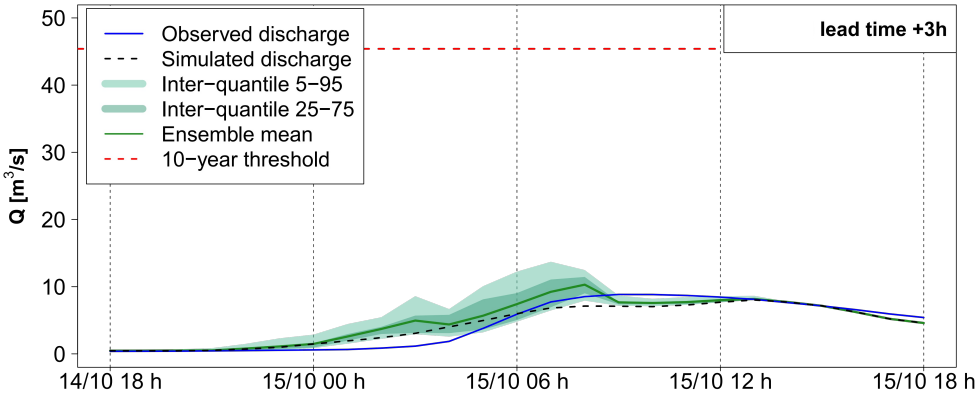

Figure B1. Detailed hydrological 3 hours lead time forecasts evaluation at outlet 1 for AROME-EPS. In blue, the observed discharge at the gauging station

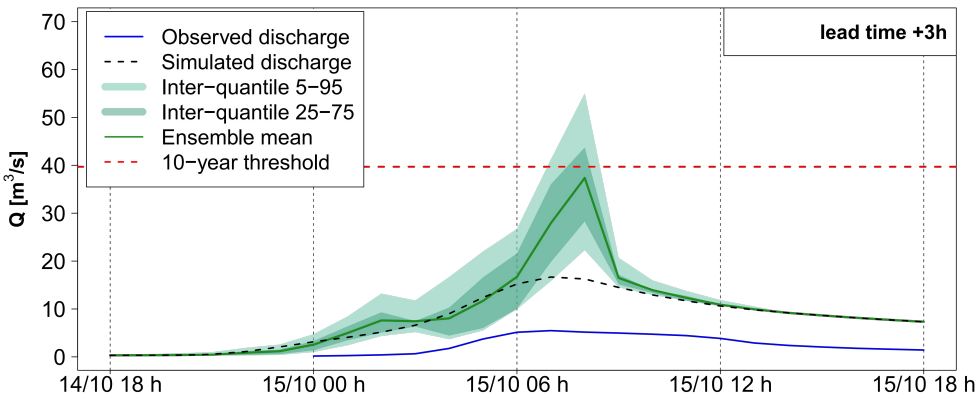

Figure B2. Detailed hydrograph for 3 hours lead time forecasts evaluation at outlet 2 for AROME-EPS. In blue, the observed discharge at the gauging station

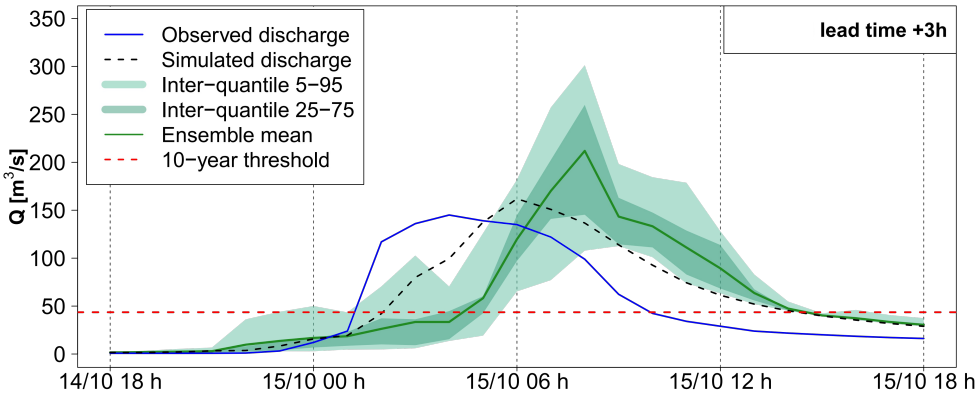

Figure B3. Detailed hydrograph for 3 hours lead time forecasts evaluation at outlet 3 for AROME-EPS. In blue, the observed discharge at the gauging station

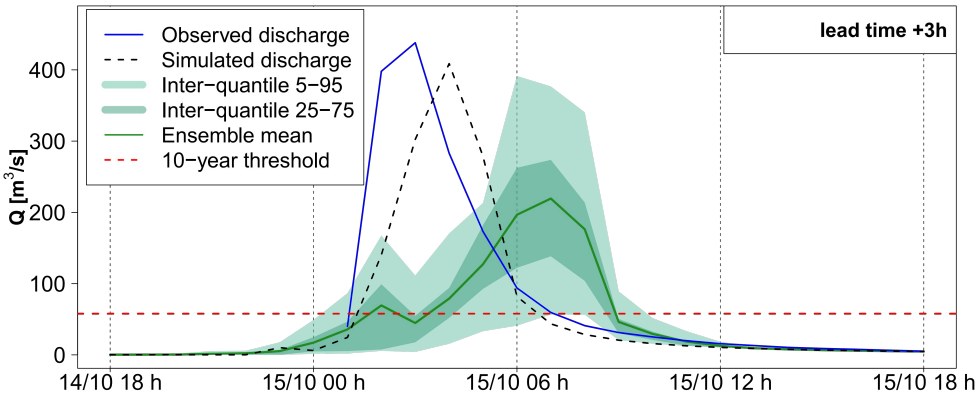

Figure B4. Detailed hydrograph for 3 hours lead time forecasts evaluation at outlet 4 for AROME-EPS. In blue, the observed discharge at the gauging station

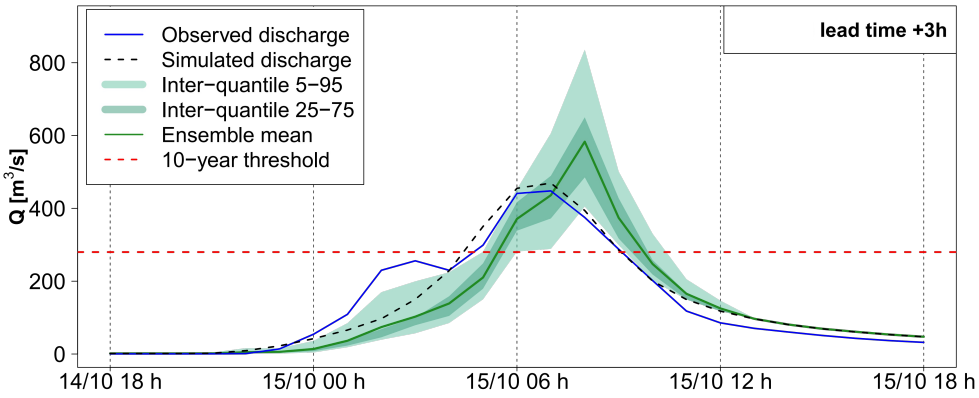

Figure B5. Detailed hydrograph for 3 hours lead time forecasts evaluation at outlet 5 for AROME-EPS. In blue, the observed discharge at the gauging station

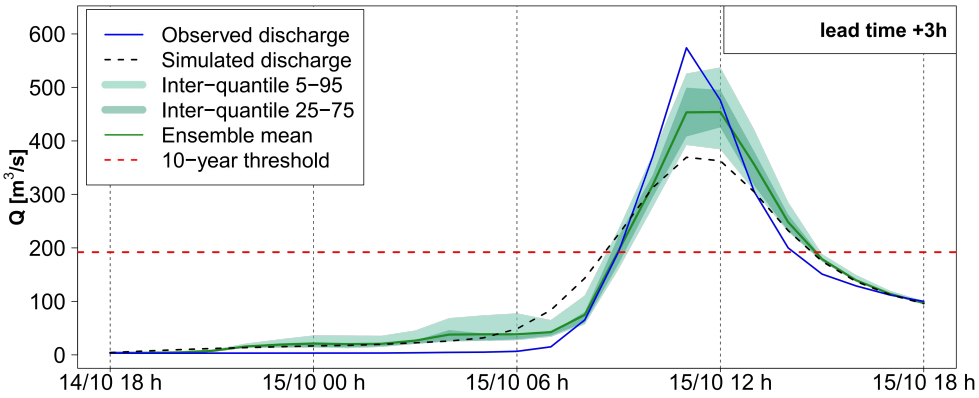

Figure B6. Detailed hydrograph for 3 hours lead time forecasts evaluation at outlet 6 for AROME-EPS. In blue, the observed discharge at the gauging station

References

Addor, N., Jaun, S., Fundel, F., and Zappa, M.: An operational hydrological ensemble prediction system for the city of Zurich (Switzerland): skill, case studies and scenarios, Hydrology and Earth System Sciences, 15, 2327–2347, https://doi.org/10.5194/hess-15-2327-2011, 2011.

Alfieri, L. and Thielen, J.: A European precipitation index for extreme rain-storm and flash flood early warning, Meteorological Applications, 22, 3–13, https://doi.org/10.1002/met.1328, 2012.

Amengual, A., Carrió, D. S., Ravazzani, G., and Homar, V.: A Comparison of Ensemble Strategies for Flash Flood Forecasting: The 12 October 2007 Case Study in Valencia, Spain, Journal of Hydrometeorology, 18, 1143–1166, https://doi.org/10.1175/JHM-D-16-0281.1, 2017.

Amengual, A., Hermoso, A., Carrió, D. S., and Homar, V.: The sequence of heavy precipitation and flash flooding of 12 and 13 September 2019 in eastern Spain. Part II: A hydro-meteorological predictability analysis based on convection-permitting ensemble strategies, Journal of Hydrometeorology, https://doi.org/10.1175/jhm-d-20-0181.1, 2021.

Anctil, F. and Ramos, M.-H.: Verification Metrics for Hydrological Ensemble Forecasts, pp. 1–30, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-40457-3_3-1, 2017.

Anderson, S. R., Csima, G., Moore, R. J., Mittermaier, M., and Cole, S. J.: Towards operational joint river flow and precipitation ensemble verification: considerations and strategies given limited ensemble records, Journal of Hydrology, 577, 123 966, https://doi.org/10.1016/j.jhydrol.2019.123966, 2019.

Armon, M., Marra, F., Enzel, Y., Rostkier-Edelstein, D., and Morin, E.: Radar-based characterisation of heavy precipitation in the eastern Mediterranean and its representation in a convection-permitting model, Hydrology and Earth System Sciences, 24, 1227–1249, https://doi.org/10.5194/hess-24-1227-2020, 2020.

Aubert, Y., Arnaud, P., Ribstein, P., and Fine, J. A.: La méthode SHYREG débit-application sur 1605 bassins versants en France métropolitaine, Hydrological Sciences Journal, 59, 993–1005, https://doi.org/10.1080/02626667.2014.902061, 2014.

Auger, L., Dupont, O., Hagelin, S., Brousseau, P., and Brovelli, P.: AROME-NWC: A new nowcasting tool based on an operational mesoscale forecasting system, Quarterly Journal of the Royal Meteorological Society, 141, 1603–1611, https://doi.org/10.1002/qj.2463, 2015.

Ayphassorho, H., Pipien, G., De Meritens, I., and Lacroix, D.: Retour d'expérience des inondations du 14 au 17 octobre 2018 dans l'Aude, Tech. rep., Ministère de la transition écologique et solidaire et Ministère de l'intérieur, http://www.cgedd.developpement-durable.gouv.fr/IMG/pdf/012561-p{_}rapport{_}cle1e876a.pdf, 2019.

Barredo, J. I.: Major flood disasters in Europe: 1950–2005, Natural Hazards, 42, 125–148, https://doi.org/10.1007/s11069-006-9065-2, 2006.

Bellier, J., Zin, I., and Bontron, G.: Sample Stratification in Verification of Ensemble Forecasts of Continuous Scalar Variables: Potential Benefits and Pitfalls, Monthly Weather Review, 145, 3529–3544, https://doi.org/10.1175/MWR-D-16-0487.1, 2017.

Bellier, J., Bontron, G., and Zin, I.: Selecting components in a probabilistic hydrological forecasting chain: the benefits of an integrated evaluation, LHB Hydrosci. J., 107, https://doi.org/10.1080/27678490.2021.1936825, 2021.

Berenguer, M., Corral, C., Sánchez-Diezma, R., and Sempere-Torres, D.: Hydrological validation of a Radar-Based nowcasting technique, Journal of Hydrometeorology, 6, 532–549, https://doi.org/10.1175/JHM433.1, 2005.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G.: SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, Journal of Hydrology, 404, 226–240, https://doi.org/10.1016/j.jhydrol.2011.04.033, 2011.

Borga, M., Gaume, E., Creutin, J. D., and Marchi, L.: Surveying flash floods: gauging the ungauged extremes, Hydrological Processes, 22, 3883–3885, https://doi.org/10.1002/hyp.7111, 2008.

Bouttier, F., Vié, B., Nuissier, O., and Raynaud, L.: Impact of stochastic physics in a convection-permitting ensemble, Monthly Weather Review, 140, 3706–3721, https://doi.org/10.1175/MWR-D-12-00031.1, 2012.

Brousseau, P., Seity, Y., Ricard, D., and Léger, J.: Improvement of the forecast of convective activity from the AROME-France system, Quarterly Journal of the Royal Meteorological Society, 142, 2231–2243, https://doi.org/10.1002/qj.2822, 2016.

Caumont, O., Mandement, M., Bouttier, F., Eeckman, J., Brossier, C. L., Lovat, A., Nuissier, O., and Laurantin, O.: The heavy precipitation event of 14-15 October 2018 in the Aude catchment: A meteorological study based on operational numerical weather prediction systems and standard and personal observations, Natural Hazards and Earth System Sciences, 21, 1135–1157, https://doi.org/10.5194/nhess-21-1135-2021, 2021.

CCR: Les inondations du 14 et 15 octobre 2018 dans le Languedoc, Tech. Rep. 11, Caisse Centrale de Réassurance, 2018.

CCR: Les catastrophes naturelles en France - Bilan 1982-2019, Tech. rep., Caisse Centrale de Réassurance (CCR), https://catastrophes-naturelles.ccr.fr/-/bilan-cat-nat-1982-2019, 2020.

Christensen, H. M.: Decomposition of a New Proper Score for Verification of Ensemble Forecasts, Monthly Weather Review, 143, 1517–1532, https://doi.org/10.1175/MWR-D-14-00150.1, 2015.

Clark, P., Roberts, N., Lean, H., Ballard, S. P., and Charlton-Perez, C.: Convection-permitting models: a step-change in rainfall forecasting, Meteorological Applications, 23, 165–181, https://doi.org/10.1002/met.1538, 2016.

Clark, R. A., Gourley, J. J., Flamig, Z. L., Hong, Y., and Clark, E.: CONUS-Wide Evaluation of National Weather Service Flash Flood Guidance Products, Weather and Forecasting, 29, 377–392, https://doi.org/10.1175/WAF-D-12-00124.1, 2014.

Collier, C. G.: Flash flood forecasting: What are the limits of predictability?, Quarterly Journal of the Royal Meteorological Society, 133, 3–23, https://doi.org/10.1002/qj.29, 2007.

Coughlan de Perez, E., van den Hurk, B., van Aalst, M. K., Jongman, B., Klose, T., and Suarez, P.: Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts, Natural Hazards and Earth System Sciences, 15, 895–904, https://doi.org/10.5194/nhess-15-895-2015, 2015.

Davolio, S., Miglietta, M. M., Diomede, T., Marsigli, C., and Montani, A.: A flood episode in northern Italy: multi-model and single-model mesoscale meteorological ensembles for hydrological predictions, Hydrology and Earth System Sciences, 17, 2107–2120, https://doi.org/10.5194/hess-17-2107-2013, 2013.

Davolio, S., Silvestro, F., and Malguzzi, P.: Effects of Increasing Horizontal Resolution in a Convection-Permitting Model on Flood Forecasting: The 2011 Dramatic Events in Liguria, Italy, Journal of Hydrometeorology, 16, 1843–1856, https://doi.org/10.1175/JHM-D-14-0094.1, 2015.

Davolio, S., Silvestro, F., and Gastaldo, T.: Impact of Rainfall Assimilation on High-Resolution Hydrometeorological Forecasts over Liguria, Italy, Journal of Hydrometeorology, 18, 2659–2680, https://doi.org/10.1175/JHM-D-17-0073.1, 2017.

De Lavenne, A., Thirel, G., Andréassian, V., Perrin, C., and Ramos, M. H.: Spatial variability of the parameters of a semi-distributed hydrological model, IAHS-AISH Proceedings and Reports, 373, 87–94, https://doi.org/10.5194/piahs-373-87-2016, 2016.

de Saint-Aubin, C., Garandeau, L., Janet, B., and Javelle, P.: A new French flash flood warning service, E3S Web of Conferences, 7, 18 024, https://doi.org/10.1051/e3sconf/20160718024, 2016.

Delaigue, O., Brigode, P., Génot, B., and Lobligeois, F.: Summary sheets of watershed-scale hydroclimatic observed data for France, Tech. rep., INRAE, https://doi.org/10.15454/UV01P1, 2020.

Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., Estournel, C., Boubrahmi, N. F., Font, J., Grubišić, V., Gualdi, S., Homar, V., Ivančan-Picek, B., Kottmeier, C., Kotroni, V., Lagouvardos, K., Lionello,

P., Llasat, M. C., Ludwig, W., Lutoff, C., Mariotti, A., Richard, E., Romero, R., Rotunno, R., Roussot, O., Ruin, I., Somot, S., Taupier-Letage, I., Tintore, J., Uijlenhoet, R., and Wernli, H.: HyMeX: A 10-Year Multidisciplinary Program on the Mediterranean Water Cycle, Bulletin of the American Meteorological Society, 95, 1063–1082, https://doi.org/10.1175/BAMS-D-12-00242.1, 2014.

Ebert, E. E.: Fuzzy verification of high-resolution gridded forecasts: A review and proposed framework, vol. 15, pp. 51–64, John Wiley and Sons Ltd, https://doi.org/10.1002/met.25, 2008.

Ehret, U. and Zehe, E.: Series distance - An intuitive metric to quantify hydrograph similarity in terms of occurrence, amplitude and timing of hydrological events, Hydrol. Earth Syst. Sci., 15, 877–896, https://doi.org/10.5194/hess-15-877-2011, 2011.

Fortin, V., Albaza, M., Anctil, F., and Turcotte, R.: Why should ensemble spread match the RMSE of the ensemble mean?, Journal of Hydrometeorology, 15, 1708–1713, 2014.

Furnari, L., Mendicino, G., and Senatore, A.: Hydrometeorological Ensemble Forecast of a Highly Localized Convective Event in the Mediterranean, Water, 12, 1545, https://doi.org/10.3390/w12061545, 2020.

Gaume, E., Livet, M., Desbordes, M., and Villeneuve, J.-P.: Hydrological analysis of the river Aude, France, flash flood on 12 and 13 November 1999, Journal of Hydrology, 286, 135–154, https://doi.org/10.1016/j.jhydrol.2003.09.015, 2004.

Gaume, E., Bain, V., Bernardara, P., Newinger, O., Barbuc, M., Bateman, A., Blaškovičová, L., Blöschl, 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, Journal of Hydrology, 367, 70–78, https://doi.org/10.1016/j.jhydrol.2008.12.028, 2009.

Georgakakos, K. P.: On the Design of National, Real-Time Warning Systems with Capability for Site-Specific, Flash-Flood Forecasts, Bulletin of the American Meteorological Society, 67, 1233–1239, https://doi.org/10.1175/1520-0477(1986)067<1233:OTDONR>2.0.CO;2, 1986.

Gourley, J. J., Flamig, Z. L., Vergara, H., Kirstetter, P.-E., Clark, R. A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J. M., Hong, Y., and Howard, K. W.: The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States, Bulletin of the American Meteorological Society, 98, 361–372, https://doi.org/10.1175/bams-d-15-00247.1, 2017.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hally, A., Caumont, O., Garrote, L., Richard, E., Weerts, A., Delogu, F., Fiori, E., Rebora, N., Parodi, A., Mihalović, A., Ivković, M., Dekić, L., van Verseveld, W., Nuissier, O., Ducrocq, V., D'Agostino, D., Galizia, A., Danovaro, E., and Clematis, A.: Hydrometeorological multi-model ensemble simulations of the 4 November 2011 flash flood event in Genoa, Italy, in the framework of the DRIHM project, Natural Hazards and Earth System Sciences, 15, 537–555, https://doi.org/10.5194/nhess-15-537-2015, 2015.

Hamill, T. M.: Interpretation of rank histograms for verifying ensemble forecasts, Mon. Weather Rev., 129, 550–560, https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2, 2001.

Hapuarachchi, H. A. P., Wang, Q. J., and Pagano, T. C.: A review of advances in flash flood forecasting, Hydrological Processes, 25, 2771–2784, https://doi.org/10.1002/hyp.8040, 2011.

Hocini, N., Payrastre, O., Bourgin, F., Gaume, E., Davy, P., Lague, D., Poinsignon, L., Pons, F., Bouguenais, F., Rennes, G., and Rennes, U.: Performance of automated methods for flash flood inundation mapping : a comparison of a digital terrain model (DTM) filling and two hydrodynamic methods, Hydrology and Earth System Sciences, 25, 2979–2995, https://doi.org/10.5194/hess-25-2979-2021, 2021.

Hopson, T. M.: Assessing the ensemble spread-error relationship, Mon. Weather Rev., 142, 1125–1142, https://doi.org/10.1175/MWR-D-12-00111.1, 2014.

Javelle, P., Organde, D., Demargne, J., Saint-Martin, C., de Saint-Aubin, C., Garandeau, L., and Janet, B.: Setting up a French national flash flood warning system for ungauged catchments based on the AIGA method, E3S Web of Conferences, 7, 18 010, https://doi.org/10.1051/e3sconf/20160718010, 2016.

Jolliffe, I. and Stephenson, D.: Forecast Verification : A Practitionner's Guide in Atmospheric Science, John Wiley & Sons, Ltd, 2012.

Lagasio, M., Silvestro, F., Campo, L., and Parodi, A.: Predictive Capability of a High-Resolution Hydrometeorological Forecasting Framework Coupling WRF Cycling 3DVAR and Continuum, Journal of Hydrometeorology, 20, 1307–1337, https://doi.org/10.1175/jhm-d-18-0219.1, 2019.

Laurantin, O.: Hourly rainfall analysis merging radar and rain gauge data., Proceedings of the International Symposium on Weather Radar and Hydrology, pp. 2–8, 2008.

Le Bihan, G.: Modèles hydrologiques régionaux pour la prévision distribuée des crues rapides : vers une estimation des impacts et des dommages potentiel, Ph.D. thesis, Université Bretagne Loire, 2016.

Le Moine, N., Andréassian, V., and Mathevet, T.: Confronting surface- and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France), Water Resources Research, 44, 1–10, https://doi.org/10.1029/2007WR005984, 2008.

Lebouc, L., Payrastre, O., and Bourgin, F.: Reconstitution des débits de pointe des crues du 15 octobre 2018 dans le bassin de l'Aude, Tech. rep., IFSTTAR, Nantes, 2019.

Leleu, I., Tonnelier, I., Puechberty, R., Gouin, P., Viquendi, I., Cobos, L., Foray, A., Baillon, M., and Ndima, P.-O.: La refonte du système d'information national pour la gestion et la mise à disposition des données hydrométriques, La Houille Blanche, 1, 25–32, https://doi.org/10.1051/lhb/2014004, 2014.

Llasat, M. C., Llasat-Botija, M., Prat, M. A., Porcú, F., Price, C., Mugnai, A., Lagouvardos, K., Kotroni, V., Katsanos, D., Michaelides, S., Yair, Y., Savvidou, K., and Nicolaides, K.: High-impact floods and flash floods in Mediterranean countries: the FLASH preliminary database, Advances in Geosciences, 23, 47–55, https://doi.org/10.5194/adgeo-23-47-2010, 2010.

Llasat, M. C., Llasat-Botija, M., Petrucci, O., Pasqua, A. A., Rosselló, J., Vinet, F., and Boissier, L.: Towards a database on societal impact of Mediterranean floods within the framework of the HYMEX project, Natural Hazards and Earth System Sciences, 13, 1337–1350, https://doi.org/10.5194/nhess-13-1337-2013, 2013.

Lobligeois, F.: Mieux connaître la distribution spatiale des pluies améliore-t-il la modélisation des crues ?, Ph.D. thesis, AgroParis Tech, 2014.

Lovat, A., Vincendon, B., and Ducrocq, V.: Hydrometeorological evaluation of two nowcasting systems for Mediterranean heavy precipitation events with operational considerations, Hydrol. Earth Syst. Sci. Discuss., pp. 1–32, https://doi.org/10.5194/hess-2020-629, 2020.

Lu, C., Yuan, H., Schwartz, B. E., and Benjamin, S. G.: Short-Range Numerical Weather Prediction Using Time-Lagged Ensembles, Weather and Forecasting, 22, 580 – 595, https://doi.org/10.1175/WAF999.1, 2007.

Marchi, L., Borga, M., Preciso, E., and Gaume, E.: Characterisation of selected extreme flash floods in Europe and implications for flood risk management, Journal of Hydrology, 394, 118–133, https://doi.org/10.1016/j.jhydrol.2010.07.017, 2010.

Mason, I.: A model for assessment of weather forecasts, Aust. Meteor. Mag, 30, 291–303, 1982.

Naulin, J. P., Payrastre, O., and Gaume, E.: Spatially distributed flood forecasting in flash flood prone areas: Application to road network supervision in Southern France, Journal of Hydrology, 486, 88–99, https://doi.org/10.1016/j.jhydrol.2013.01.044, 2013.

Nuissier, O., Marsigli, C., Vincendon, B., Hally, A., Bouttier, F., Montani, A., and Paccagnella, T.: Evaluation of two convection-permitting ensemble systems in the HyMeX Special Observation Period (SOP1) framework, Quarterly Journal of the Royal Meteorological Society, 142, 404–418, https://doi.org/10.1002/qj.2859, 2016.

Osinski, R. and Bouttier, F.: Short-range probabilistic forecasting of convective risks for aviation based on a lagged-average-forecast ensemble approach, Meteorological Applications, 25, 105–118, https://doi.org/https://doi.org/10.1002/met.1674, 2018.

Oudin, L., Hervieu, F., Michel, C., and Perrin, C.: Which Potential Evapotranspiration Input for a Lumped Rainfall-Runoff Model? Part 2—Towards a Simple and Efficient Potential Evapotranspiration Model for Rainfall-Runoff Modeling, Journal of Hydrology, 303, 290–306, https://doi.org/10.1016/j.jhydrol.2004.08.026, 2005.

Pappenberger, F., Cloke, H. L., Persson, A., and Demeritt, D.: HESS Opinions "on forecast (in)consistency in a hydro-meteorological chain: Curse or blessing?", Hydrol. Earth Syst. Sci., 15, 2391–2400, https://doi.org/10.5194/hess-15-2391-2011, 2011.

Park, S., Berenguer, M., and Sempere-Torres, D.: Long-term analysis of gauge-adjusted radar rainfall accumulations at European scale, Journal of Hydrology, 573, 768–777, https://doi.org/10.1016/j.jhydrol.2019.03.093, 2019.

Peredo, D., Ramos, M.-H., Andréassian, V., and Oudin, L.: Investigating hydrological model versatility to simulate extreme flood events, Hydrological Sciences Journal, 0, 1–18, https://doi.org/10.1080/02626667.2022.2030864, 2022.

Petrucci, O., Aceto, L., Bianchi, C., Bigot, V., Brázdil, R., Pereira, S., Kahraman, A., Özgenur Kılıç, Kotroni, V., Llasat, M. C., Llasat-Botija, M., Papagiannaki, K., Pasqua, A. A., Řehoř, J., Geli, J. R., Salvati, P., Vinet, F., and Zêzere, J. L.: Flood Fatalities in Europe, 1980–2018: Variability, Features, and Lessons to Learn, Water, 11, 1682, https://doi.org/10.3390/w11081682, 2019.

Poletti, M. L., Silvestro, F., Davolio, S., Pignone, F., and Rebora, N.: Using nowcasting technique and data assimilation in a meteorological model to improve very short range hydrological forecasts, Hydrology and Earth System Sciences, 23, 3823–3841, https://doi.org/10.5194/hess-23-3823-2019, 2019.

Price, D., Hudson, K., Boyce, G., Schellekens, J., Moore, R. J., Clark, P., Harrison, T., Connolly, E., and Pilling, C.: Operational use of a grid-based model for flood forecasting, Proceedings of the Institution of Civil Engineers - Water Management, 165, 65–77, https://doi.org/https://doi.org/10.1680/wama.2012.165.2.65, 2012.

Raynaud, L. and Bouttier, F.: Comparison of initial perturbation methods for ensemble prediction at convective scale, Quarterly Journal of the Royal Meteorological Society, 142, 854–866, https://doi.org/10.1002/qj.2686, 2016.

Richardson, D. S., Cloke, H. L., and Pappenberger, F.: Evaluation of the Consistency of ECMWF Ensemble Forecasts, Geophys. Res. Lett., 47, 1–8, https://doi.org/10.1029/2020GL087934, 2020.

Roberts, N. M. and Lean, H. W.: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, Monthly Weather Review, 136, 78–97, https://doi.org/10.1175/2007MWR2123.1, 2008.

Sayama, T., Yamada, M., Sugawara, Y., and Yamazaki, D.: Ensemble flash flood predictions using a high-resolution nationwide distributed rainfall-runoff model: case study of the heavy rain event of July 2018 and Typhoon Hagibis in 2019, Progress in Earth and Planetary Science, 7, https://doi.org/10.1186/s40645-020-00391-7, 2020.

Schwartz, C. S. and Sobash, R. A.: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations, Mon. Weather Rev., 145, 3397–3418, https://doi.org/10.1175/MWR-D-16-0400.1, 2017.

Seity, Y., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective-scale operational model, Monthly Weather Review, 139, 976–991, https://doi.org/10.1175/2010MWR3425.1, 2011.

Silvestro, F. and Rebora, N.: Operational verification of a framework for the probabilistic nowcasting of river discharge in small and medium size basins, Natural Hazards and Earth System Sciences, 12, 763–776, https://doi.org/10.5194/nhess-12-763-2012, 2012.

Silvestro, F., Rebora, N., and Ferraris, L.: Quantitative Flood Forecasting on Small- and Medium-Sized Basins: A Probabilistic Approach for Operational Purposes, Journal of Hydrometeorology, 12, 1432–1446, https://doi.org/https://doi.org/10.1175/JHM-D-10-05022.1, 2011.

Talagrand, O. and Vautard, R.: Evaluation of probabilistic prediction systems for a scalar variable, 1997.

Thiboult, A., Anctil, F., and Ramos, M. H.: How does the quantification of uncertainties affect the quality and value of flood early warning systems?, J. Hydrol., 551, 365–373, https://doi.org/10.1016/j.jhydrol.2017.05.014, 2017.

Valdez, E. S., Anctil, F., and Ramos, M. H.: Choosing between post-processing precipitation forecasts or chaining several uncertainty quantification tools in hydrological forecasting systems, Hydrol. Earth Syst. Sci., 26, 197–220, https://doi.org/10.5194/hess-26-197-2022, 2022.

Verkade, J. S. and Werner, M. G.: Estimating the benefits of single value and probability forecasting for flood warning, Hydrology and Earth System Sciences, 15, 3751–3765, https://doi.org/10.5194/hess-15-3751-2011, 2011.

Versini, P.-A., Gaume, E., and Andrieu, H.: Application of a distributed hydrological model to the design of a road inundation warning system for flash flood prone areas, Natural Hazards and Earth System Sciences, 10, 805–817, https://doi.org/10.5194/nhess-10-805-2010, 2010.

Vidal, J.-P., Martin, E., Franchistéguy, L., Baillon, M., and Soubeyroux, J.-M.: A 50-year high-resolution atmospheric reanalysis over France with the Safran system, International Journal of Climatology, 30, 1627–1644, 2010.

Vié, B., Molinié, G., Nuissier, O., Vincendon, B., Ducrocq, V., Bouttier, F., and Richard, E.: Hydro-meteorological evaluation of a convection-permitting ensemble prediction system for Mediterranean heavy precipitating events, Natural Hazards and Earth System Sciences, 12, 2631–2645, https://doi.org/10.5194/nhess-12-2631-2012, 2012.

Vincendon, B., Ducrocq, V., Nuissier, O., and Vié, B.: Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting, Natural Hazards and Earth System Sciences, 11, 1529–1544, https://doi.org/10.5194/nhess-11-1529-2011, 2011.

Wilks, D. S.: Forecast Verification, vol. 100, Elsevier, https://doi.org/10.1016/B978-0-12-385022-5.00008-7, 2011.

Zanchetta, A. and Coulibaly, P.: Recent Advances in Real-Time Pluvial Flash Flood Forecasting, Water, 12, 570, https://doi.org/10.3390/w12020570, 2020.

DOCTORAT ECOLOGIE BRETAGNE GEOSCIENCES LOIRE AGRONOMIE ALIMENTATION

Nantes Université

Titre : Évaluation de prévisions hydrométéorologiques d'ensemble des crues soudaines à partir de données sur les impacts

Mots clés : Crue rapide, prévision d'ensemble, impacts

Résumé : Les crues soudaines (ou crues-éclair) sont responsables chaque année de décès et d'importants dégâts matériels. En France, ce type d'évènement concerne principalement l'arc méditerranéen et ses petits cours d'eau. Depuis plusieurs années, d'importantes évolutions sur les différents modèles et méthodes pour la prévision immédiate des crues-éclair ont été réalisées.

Dans ce travail, trois produits météorologiques de prévisions d'ensemble ont été utilisés comme données d'entrée pour une modélisation hydrologique à partir du modèle Cinecar et évalués à partir d'un cadre méthodologique pensé pour l'évaluation événementielle. Cette méthodologie a été appliquée sur trois cas de crues soudaines survenues en 2018 dans l'Aude, l'Ardèche et le Var et repose sur une adaptation de la méthodologie de construction des courbes ROC. Cette approche permet en particulier de représenter spatialement les anticipations de dépassement de seuil de débits.

À partir des simulations de débits et d'un catalogue de scénario d'inondation, des prévisions d'ensemble d'impacts ont été réalisées et à leur tour évaluées avec le même cadre méthodologique basé sur les courbes ROC. Les prévisions d'impacts ont également été évaluées du point de vue de leur capacité à déclencher des prises de décisions pertinentes et efficaces via un rejeu de l'événement de l'Aude 2018 du point de vue du SDIS.

L'ensemble de ce travail a permis de mettre en évidence les performances de ces trois ensembles pour la prévision des crues-éclair, mais aussi de proposer un cadre méthodologique intéressant à exploiter dans le cadre d'évaluations événementielles, ce cadre restant encore perfectible. Enfin, le passage à des simulations d'impacts a montré l'intérêt de l'utilisation de ce genre de données pour l'anticipation et la gestion de crise de ce type d'événement.

Title: Evaluation of hydrometeorological ensemble flash flood forecasts from impact data

Keywords: flash floods, ensemble forecasts, impacts

Abstract: Flash floods are responsible each year for deaths and important material damages. In France, these events mainly concern the Mediterranean arc and its small rivers. Since several years, important evolutions on the models and methods for the immediate forecasting of flash floods have been realized.

In this work, three meteorological products of ensemble forecasts have been used as input data for a hydrological simulation based on the Cinecar model and evaluated from a methodological framework designed for the event evaluation. This methodology was applied to three flash floods that occurred in 2018 in the Aude, Ardèche and Var. The methodology is based on an adaptation of the ROC curve construction. Specifically, this approach is designed for spatial representation of the thresholds exceedances anticipations.

From the discharge simulations and a catalog of flooding scenarios, ensemble impact forecasts were made and evaluated with the same methodological framework based on the ROC curves. The impact forecasts were also evaluated from an emergency service point of view, with a specific attention to the ability of the impact forecasts to be relevant and decision-making.

This work allowed to highlight the performance of these three ensembles for flash flood forecasting, but also to propose an interesting methodological framework which can be useful for event-based evaluations, this framework still remaining perfectible. Finally, the transition to impact simulations has shown the interest of using this kind of data for the anticipation and crisis management of this type of event.