



HAL
open science

Intégration des pluies prévues du système COSMO-DE-EPS dans le système AIGA

J. Demargne, P. Javelle

► **To cite this version:**

J. Demargne, P. Javelle. Intégration des pluies prévues du système COSMO-DE-EPS dans le système AIGA. [Rapport de recherche] irstea. 2014, pp.56. hal-02606076

HAL Id: hal-02606076

<https://hal.inrae.fr/hal-02606076>

Submitted on 16 May 2020

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

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



SCHAPI
Service Central
d'Hydrométéorologie
et d'Appui à la
Prévision des
Inondations



Marché « Mise en œuvre de la méthode AIGA » n°2014-01-16

**Action « prévision des inondations »
AXE III : « Crues Soudaines»
A. Recherches et développements**

**Intégration des pluies prévues
du système COSMO-DE-EPS
dans le système AIGA**

Julie Demargne (HYDRIS-Hydrologie)

Février 2015

HYDRIS
hydrologie

Table des Matières

Résumé.....	3
1. Objectifs.....	5
a) Prise en compte des incertitudes dans le système d'anticipation AIGA.....	5
b) Prévisions d'ensemble de pluie COSMO-DE-EPS.....	7
c) Objectifs de l'étude.....	9
2. Données et Modélisation Hydrologique.....	10
a) Bassins versants.....	10
b) Modélisation hydrologique et données d'observations.....	11
c) Sélection des évènements.....	14
3. Méthodologies.....	15
a) Alertes AIGA probabilisées.....	15
b) Evaluation de la qualité des alertes.....	21
c) Vérification des prévisions d'ensemble.....	23
4. Résultats.....	26
a) Qualité des prévisions d'ensembles de pluie de bassin.....	26
b) Qualité des prévisions d'ensembles de débit.....	30
c) Qualité des alertes probabilisées.....	39
5. Conclusions et Perspectives.....	46
6. Références.....	50
Appendice A : Liste des Bassins Versants.....	52
Appendice B : Liste des Figures et Tableaux.....	54

Résumé

Cette étude a été réalisée pour le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) pour analyser l'amélioration du système d'avertissement aux crues rapides AIGA en prenant en compte les incertitudes des pluies futures et des paramètres du modèle hydrologique calibré et régionalisé. Le SCHAPI développe actuellement un service d'avertissement automatisé valable sur l'ensemble de la France, qui sera complété par un service expertisé d'avertissement sur des bassins dits prioritaires. Ce service expertisé concerne les bassins avec des temps de réaction courts compris entre 2h et 6h et avec des enjeux particuliers.

Pour améliorer la caractérisation et l'anticipation des crues soudaines, l'objectif est d'intégrer des prévisions d'ensemble de pluie à haute résolution spatio-temporelle, notamment les prévisions d'ensemble multi-modèles COSMO-DE-EPS. Ces prévisions sont produites par l'agence allemande Deutscher Wetterdienst (DWD) sur le tiers Nord-Est de la France. Le modèle atmosphérique COSMO-DE permet une modélisation physique des processus convectifs à la résolution de 2,8 km. Il quantifie l'incertitude des prévisions de pluie en produisant, toutes les 3 heures et 8 fois par jour, 20 membres pour une échéance maximale de 21 heures. Les grilles de prévision d'ensemble de pluie ont été mises à disposition par le SPC (Service de Prévision des Crues) Meuse-Moselle sur la zone du SPC au format et dans le système de projection tels qu'utilisés sur la plateforme opérationnelle LARSIM. L'intégration de ces grilles de prévision a nécessité leur reprojection dans le système Lambert 2 étendu et leur interpolation sur une grille régulière à la résolution de 1 km² (cette étape est décrite dans l'étude de Demargne 2014a).

L'analyse de l'apport des prévisions d'ensemble de débit a porté sur 7 événements de crue entre décembre 2010 et décembre 2012 pour des bassins de la Meuse et de la Moselle de surfaces inférieures à 1000 km². Les 20 ensembles de pluie COSMO-DE-EPS sont intégrés dans la modélisation hydrologique pour obtenir 20 prévisions d'ensemble de débit et produire ainsi des alertes probabilisées de dépassement de seuil de débit. L'incertitude paramétrique du modèle horaire semi-distribué a également été considéré en utilisant 5 jeux de paramètres différents, qui ont été testés lors de la régionalisation par Organde (2014). Ces 5 jeux de paramètres sont considérés comme équiprobables, même si seul le jeu de paramètres finaux a été optimisé pour la détection des dépassements de quantiles de débit. Les 5 jeux de paramètres sont couplés aux 20 ensembles de pluie pour produire 100 ensembles de débit qui prennent en compte l'incertitude de pluie future et l'incertitude paramétrique du modèle semi-distribué.

L'apport des prévisions d'ensemble COSMO-DE-EPS est analysé vis-à-vis de prévisions de référence : la prévision de débit générée avec la pluie persistante issue de la dernière heure de pluie observée au moment de l'initialisation de la prévision, la prévision probabiliste climatologique définie par la distribution empirique des observations de 2009-2012, ainsi que la prévision de débit issue de la pluie « parfaite » (i.e. pluie observée) pour montrer l'impact des erreurs et des incertitudes des prévisions d'ensemble de pluie. La vérification des prévisions d'ensemble de pluie et de débit est réalisée avec le système de vérification EVS (Brown et al. 2010) en sélectionnant des critères pour décrire le biais relatif et la corrélation des moyennes des ensembles avec les observations, ainsi que le gain en termes d'erreur en probabilité et de discrimination des dépassements de quantiles observés. L'analyse de la qualité des alertes inclut différents critères pour décrire notamment la fiabilité, le biais en fréquence et l'anticipation des alertes. Les cartes d'alerte probabilisées issues des prévisions d'ensemble de débit sont présentées et complétées avec les cartes de performance des alertes pour offrir une visualisation spatiale de la qualité des alertes sur les bassins jaugés. De telles

cartes peuvent renseigner sur la qualité potentielle des alertes sur les bassins non jaugés voisins.

L'évaluation des ensembles de débit a montré une bonne corrélation des moyennes des ensembles avec les observations mais une sous-estimation significative pour les débits les plus forts, à cause du biais négatif du modèle hydrologique et, dans une moindre mesure, du biais des ensembles de pluie COSMO-DE-EPS. Cependant, certaines surestimations des pluies moyennes peuvent compenser les erreurs de sous-estimation du modèle hydrologique. Les ensembles de pluie et de débit ont une bonne capacité pour discriminer les dépassements des quantiles observés 75% et 90%. En comparaison de la prévision probabiliste climatologique et de la prévision persistante, les ensembles ont de meilleures performances, notamment en termes d'erreur en probabilité.

L'évaluation pour les alertes de dépassement du quantile de débit biennal a montré une amélioration significative en termes de détection, fiabilité et anticipation en comparaison au système AIGA actuel (qui n'intègre pas de prévision de pluie) et aux alertes issues de la prévision de pluie persistante. L'utilisation de 5 jeux de paramètres pour le modèle hydrologique semi-distribué pour prendre en compte l'incertitude paramétrique n'a pas permis d'améliorer les performances. Ceci s'explique par le fait qu'un seul des 5 jeux de paramètres a été optimisé pour augmenter la réactivité du modèle hydrologique et mieux différencier les dépassements des différents quantiles de débit.

Des pistes d'amélioration pour l'intégration de prévisions immédiates (nowcasts), la prise en compte des incertitudes de source hydrologique et l'utilisation des informations post-crue de dégâts notamment pour les bassins non jaugés, sont également présentées.

Le présent rapport constitue le livrable 00050 de la commande n°4500253337 du 16/09/2014 du marché « Mise en œuvre de la méthode AIGA » n°2014-01-16, conclu entre IRSTEA et Hydris Hydrologie.

1. Objectifs

Cette étude a pour objectif d'étudier l'apport des prévisions d'ensemble de pluie du système COSMO-DE-EPS pour le système d'avertissement des crues rapides AIGA. Ce travail est réalisé pour le SCHAPI qui développe actuellement un nouveau service d'avertissements expertisés Crues Soudaines sur des bassins dits prioritaires, dont le temps de réaction varie entre 2h et 6h, et pour lesquels des enjeux ont été identifiés concernant le risque de crue soudaine.

L'étude réalisée pour le SCHAPI en 2013 a montré l'apport des prévisions de pluie à haute résolution issues du modèle AROME (Application of Research to Operations at Mesoscale) de Météo-France pour améliorer la détection et l'anticipation des dépassements du débit biennal pour des bassins de la zone du SPC de Meuse-Moselle (Demargne 2014b). Les prévisions AROME de type ensembliste étant en cours de développement à Météo-France, les prévisions déterministes AROME disponibles toutes les 6 heures ont été utilisées pour construire des prévisions d'ensemble par décalage temporel de prévisions déterministes successives (prévisions dites d'ensembles timelag). Cette approche pragmatique permet de prendre en compte les incertitudes sur les prévisions de pluie à une échelle spatio-temporelle fine (2,5 km) et pour des événements de pluie intense. Or, le SPC Meuse-Moselle ainsi que le SPC Rhin-Sarre ont accès aux prévisions d'ensemble de pluie à haute résolution issues du modèle COSMO-DE-EPS de l'agence Deutscher Wetterdienst (DWD) sur la plateforme opérationnelle LARSIM. L'objectif est donc d'intégrer les prévisions d'ensemble de pluie COSMO-DE-EPS dans le système AIGA pour analyser leur apport potentiel pour les avertissements des crues soudaines.

a) Prise en compte des incertitudes dans le système d'anticipation AIGA

Pour l'alerte aux crues soudaines en milieu non jaugé, Irstea et Météo-France ont développé le système d'anticipation AIGA (Javelle et al. 2010, 2014), qui fonctionne en temps réel sur le pourtour méditerranéen depuis 2005. Dans ce système, un modèle hydrologique distribué intègre les pluies horaires produites à la résolution de 1km^2 par Météo-France, à partir des radars météorologiques et des stations pluviométriques, pour estimer le débit en tout point des cours d'eau. Ces débits sont alors comparés en temps réel aux quantiles de débit préalablement estimés par la méthode régionalisée SHYREG (Arnaud et Lavabre 2002, Arnaud et al. 2008). Cette méthode a permis de fournir une base de quantiles de débit de pointe et de débits pour les durées de 1h à 72h pour les périodes de retour de 2 ans à 100 ans sur la France métropolitaine pour tout bassin versant de superficie comprise entre 10 et 2000 km^2 .

Pour caractériser la gravité des événements en cours, le système AIGA affiche, toutes les 15 minutes, la carte des périodes de retour estimées le long des cours d'eau pour les catégories suivantes : entre 2 et 10 ans, entre 10 et 50 ans, et supérieur à 50 ans. Ces cartes sont disponibles pour les services du SCHAPI et des SPC des régions du Sud-Est pour leur permettre d'émettre, si nécessaire, les bulletins d'alerte aux crues.

La calibration et la régionalisation du modèle pluie-débit ont été récemment améliorées par Organde (2013, 2014) pour l'implémentation d'ici 2016 par le SCHAPI d'un système national d'alerte Crues Soudaines. En particulier, les travaux d'Organde de 2013 ont permis de tester différentes options de modélisation hydrologique, événementielle ou continue,

distribuée, semi-distribuée ou globale, ainsi que différents critères de calibration et différentes méthodes de régionalisation, afin de déterminer la ou les modélisations hydrologiques les plus pertinentes pour un système d'alerte national. Les procédures de calibration et d'optimisation de la régionalisation pour la première mise en place du système national ont été finalisées dans l'étude d'Organde (2014).

La caractérisation et l'anticipation des crues soudaines peut également être significativement améliorée en intégrant des prévisions de pluie, notamment grâce aux développements et à la plus grande disponibilité de prévisions de pluie à plus fine échelle spatio-temporelle, issues d'une meilleure modélisation des phénomènes de pluie intense par les modèles atmosphériques numériques. Cependant, tout système de prévision est impacté par de nombreuses sources d'incertitude. Ces incertitudes sont issues de la variabilité (naturelle et causée par l'homme) des phénomènes modélisés et des variables, des modélisations imparfaites de ces phénomènes, ainsi que des limitations en termes de mesures et données disponibles. Comme l'a souligné notamment le rapport du NRC (2006), ces incertitudes de prévisions doivent être prises en compte et quantifiées pour communiquer aux utilisateurs des prévisions avec leurs incertitudes associées et permettre ainsi leur utilisation optimale.

Pour un système de prévision hydrométéorologique, on différencie généralement (Demargne et al. 2014):

- les incertitudes météorologiques liées aux observations et prévisions des variables atmosphériques (ex. pluie et température) qui sont utilisées en entrée des modèles hydrologiques,
- les incertitudes hydrologiques, provenant des états initiaux des modèles hydrologiques, des paramètres de ces modèles, de la structure des modèles, et des régulations potentielles de débit.

Une des approches pour quantifier ces différentes sources d'incertitude est de produire des prévisions d'ensemble qui décrivent des scénarios équiprobables avec les différents membres de l'ensemble. L'approche ensembliste est mise en œuvre depuis plus de 20 ans dans les systèmes de prévisions météorologiques pour les prévisions à court, moyen et long terme, ce qui a conduit plus récemment au développement de systèmes de prévision d'ensemble hydrométéorologique. Les articles de Cloke et Pappenberger (2009) et de Zappa et al. (2010) donnent de nombreux exemples de systèmes opérationnels ou pré-opérationnels de prévision d'ensemble pour l'hydrologie.

Le système AIGA peut également bénéficier de cette approche ensembliste pour intégrer des prévisions d'ensemble de pluie et quantifier les principales incertitudes liées à l'estimation des prévisions de débit et des dépassements de quantiles de crue. La Figure 1 décrit les principales sources d'incertitude pour le système AIGA. Avec la modélisation hydrologique actuellement utilisée dans AIGA, les incertitudes météorologiques sont liées aux observations de pluie, issues des données des radars et des pluviomètres, et aux prévisions de pluie. Pour le modèle hydrologique, les incertitudes concernent la paramétrisation du modèle, son initialisation, ainsi que la structure du modèle. Lorsque des prévisions d'ensemble de pluie sont utilisées avec un ou, potentiellement, plusieurs jeux équiprobables de pluie observée, de conditions initiales et de paramètres hydrologiques, le système produit des prévisions d'ensemble de débit qui sont également équiprobables. Les quantiles de crue de la base de données SHYREG sont également issus d'une modélisation hydrométéorologique et donc entachées d'incertitude. La comparaison des prévisions d'ensemble de débit équiprobables avec les quantiles de crue (dont les incertitudes sont idéalement caractérisées) conduit à

estimer la probabilité de dépasser un certain quantile de crue. Ainsi, le prévisionniste peut donner un avertissement de crue associé à un niveau de probabilité. En donnant de manière objective des niveaux de confiance aux avertissements, les prévisionnistes et utilisateurs peuvent considérer des niveaux de risque acceptable pour différentes applications et leur prise de décision peut donc être optimisée.

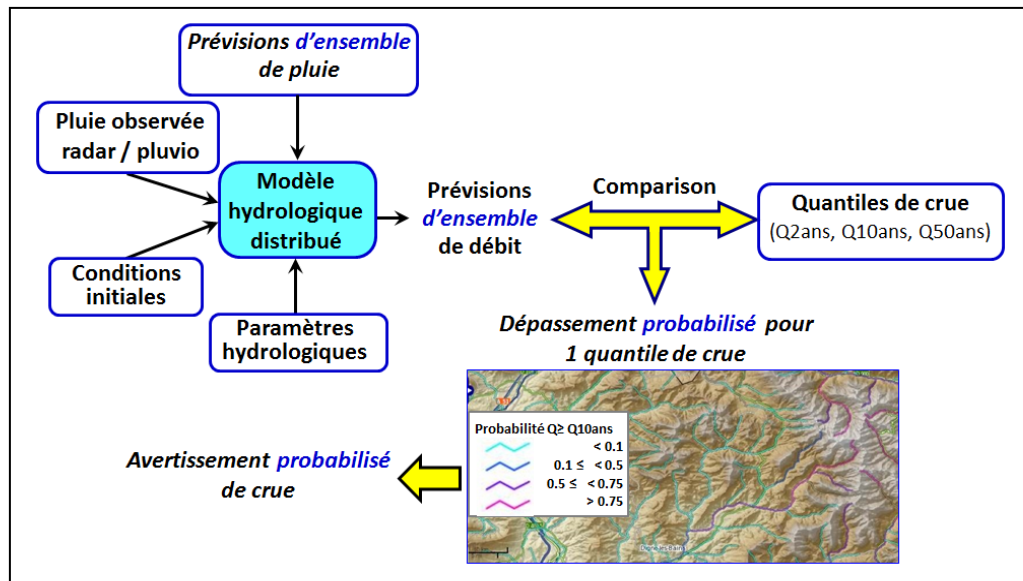


Figure 1 : Identification et prise en compte des principales sources d'incertitude dans la chaîne d'alerte aux crues soudaines du système AIGA

Cette étude considère les incertitudes des prévisions de pluie grâce à l'intégration de prévisions d'ensemble de pluie. De plus, le travail de régionalisation réalisé par Organde en 2014 a permis d'estimer différents jeux de paramètres hydrologiques issus de différentes contraintes sur les valeurs de 2 paramètres. Pour une première prise en compte de l'incertitude de paramétrisation du modèle hydrologique, nous avons choisi d'utiliser ces différents jeux de paramètres hydrologiques. En faisant l'hypothèse que ces différents jeux de paramètres sont équiprobables, ils sont intégrés au système pour générer, à partir des prévisions d'ensemble de pluie, des prévisions d'ensemble de débit caractérisant les incertitudes des pluies prévues et des paramètres hydrologiques.

b) Prévisions d'ensemble de pluie COSMO-DE-EPS

Cette étude utilise une archive des prévisions d'ensemble de pluie COSMO-DE-EPS mise à disposition par le SPC Meuse-Moselle sur la période de décembre 2010 à décembre 2013. La première étape du travail a consisté à extraire les grilles de prévisions de pluie sur le domaine d'étude dans le système de projection et à la résolution spatiale utilisés dans le système AIGA. Cette partie du travail est décrite dans l'étude (Demargne 2014a). Nous rappelons ici les informations principales de ce rapport.

L'agence allemande Deutscher Wetterdienst (DWD) a développé un modèle déterministe COSMO-DE à la résolution de 2,8 km permettant de modéliser les processus convectifs. A partir du modèle déterministe COSMO-DE, des prévisions d'ensemble sont produites en prenant en compte les incertitudes de l'initialisation, des conditions aux limites et de paramétrisation physique du modèle.

Comme le montre la Figure 2, le domaine du modèle COSMO-DE couvre le tiers Nord-Est de la France, et donc en particulier la zone d'étude du SPC Meuse-Moselle. Météo-France, l'agence DWD et l'agence anglaise UK Met ont actuellement un projet de développement de prévisions dites super ensembles, issues de leurs 3 systèmes de prévisions ensemblistes, calibrées sur une grille commune, pour le transport aérien (projet Single European Sky Air Traffic Management Research, Gebhardt et al. 2012). Ces prévisions dites super ensembles pourront également être pertinentes pour améliorer la description probabiliste des événements de pluie convective.

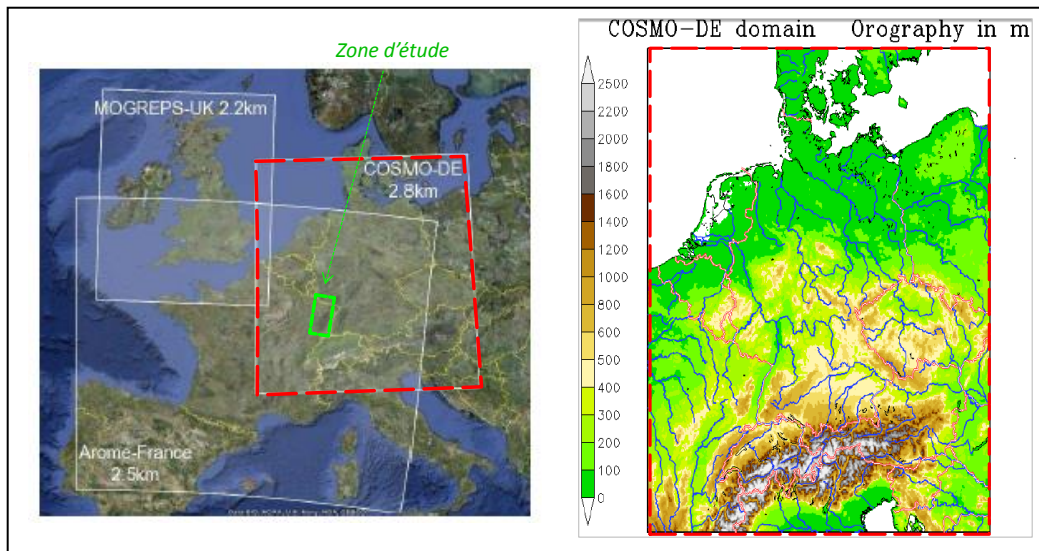


Figure 2 : Domaine des prévisions d'ensemble COSMO-DE-eps vis-à-vis de la zone d'étude Meuse-Moselle et complémentarité avec les prévisions AROME de Météo-France et prévisions MOGREPS de l'agence UK Met (d'après Gebhardt et al. 2012)

Les caractéristiques des prévisions d'ensemble COSMO-DE-eps sont les suivantes :

- prévisions d'ensemble multi-modèles basées sur l'utilisation de 4 modèles atmosphériques globaux et du modèle déterministe COSMO-DE qui inclut une modélisation physique des processus convectifs ;
- des prévisions à la résolution spatiale de 2,8 km, au pas de temps horaire, pour une échéance maximale de 21 heures ;
- des prévisions disponibles 8 fois par jour pour les réseaux suivants : 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC et 21 UTC.

La figure 3 présente le schéma de construction des 20 prévisions d'ensemble multi-modèles COSMO-DE-eps prenant en compte :

- les incertitudes de l'initialisation et des conditions aux limites pour le modèle déterministe COSMO-DE en utilisant les sorties de 4 modèles atmosphériques globaux qui sont le modèle global GME de l'agence allemande DWD, le modèle Integrated Forecast System de l'agence européenne ECMWF, le Global Forecast System de l'agence américaine NCEP, et le Global Spectral Model de l'agence météorologique japonaise,
- les incertitudes de paramétrisation physique de COSMO-DE en utilisant 5 configurations différentes de paramètres physiques.

On a observé sur les événements de l'étude que les 20 membres de la prévision d'ensemble de pluie se regroupaient par groupe de 5 membres, correspondant aux 5 membres issus d'un

même modèle atmosphérique global. Ceci indique que les sources d'incertitude qui semblent les plus importantes et/ou les mieux décrites (si les ensembles sont fiables) par les prévisions COSMO-DE-EPS sont les incertitudes de l'initialisation et des conditions aux limites, et non les incertitudes de la paramétrisation du modèle COSMO-DE.

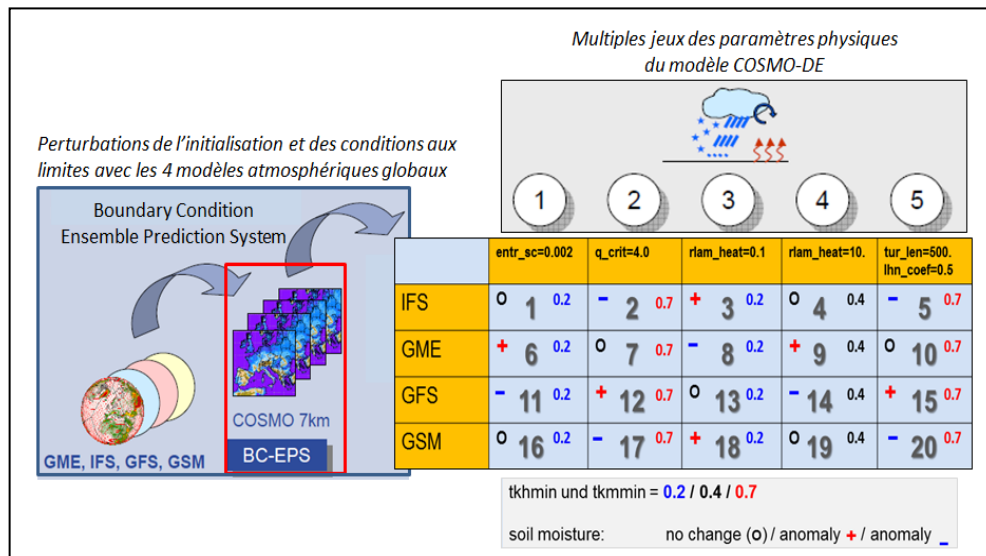


Figure 3 : Schéma du système de prévision d'ensemble COSMO-DE-EPS considérant l'incertitude d'initialisation et des conditions aux limites ainsi que l'incertitude de paramétrisation du modèle déterministe COSMO-DE (d'après Gebhardt et al. 2012 et Theis et al. 2014)

L'intégration des informations globales multi-modèles utilise une résolution intermédiaire de 7 km du modèle COSMO pour le passage de l'échelle synoptique à l'échelle convective. La méthode dite de Boundary Condition Ensemble Prediction System (BC-EPS) permet de produire 4 jeux de conditions aux limites et de perturbations des conditions initiales pour le modèle COSMO-DE. A l'échelle de 2,8 km du modèle COSMO-DE, les informations de pluie issues des radars sont également assimilées à la date d'initialisation. En utilisant 5 jeux de paramétrisation physique équiprobable et les informations issues de 4 modèles atmosphériques globaux, le système COSMO-DE-EPS produit 20 ensembles équiprobables de prévision de pluie.

Dans cette étude, le délai de mise à disposition des prévisions par l'agence DWD n'était pas connu et donc n'a pas été considéré. Mais ce délai devra être pris en compte pour une utilisation opérationnelle de ces prévisions de pluie.

c) Objectifs de l'étude

Le travail proposé répond aux trois objectifs suivants :

- 1) prendre en compte l'incertitude des prévisions de pluie en intégrant les prévisions d'ensemble de pluie COSMO-DE-EPS dans la modélisation hydrologique pour produire des prévisions d'ensemble de débit qui reflètent l'incertitude de pluie future et générer des alertes probabilisées de dépassement de seuil de débit ;
- 2) prendre en compte l'incertitude paramétrique du modèle hydrologique en complément de l'incertitude de pluie pour juger de l'amélioration potentielle des alertes avec l'utilisation de différents jeux de paramètres hydrologiques régionalisés ;

- 3) évaluer la qualité des prévisions d'ensemble de pluie et de débit, ainsi que la qualité des alertes vis-à-vis du système AIGA actuel (sans pluie future) et de prévisions déterministes de référence.

2. Données et Modélisation Hydrologique

Le travail de cette étude utilise les données de débit et la modélisation hydrologique proposée dans l'étude d'Organde (2014) pour l'implémentation du système national AIGA. La zone d'étude de Meuse-Moselle est une des 3 zones d'étude choisies en collaboration avec le SCHAPI pour l'évaluation de l'apport des prévisions AROME de l'étude de Demargne (2014b), zone pour laquelle les prévisions COSMO-DE-EPS ont été mises à disposition par le SPC Meuse-Moselle.

a) Bassins versants

Les bassins versants sont issus de la base de bassins versants du SCHAPI, et plus particulièrement de la sélection des bassins réalisée par Organde (études de 2013 et 2014) pour finaliser la calibration et la régionalisation de la modélisation hydrologique. Les critères de sélection sont les suivants :

- les données de débit observées sont disponibles sur la période de calibration et régionalisation de la modélisation hydrologique 2001-2006 avec un taux de lacune inférieur jusqu'à 10% et dont la qualité en hautes eaux a été jugée acceptable par le réseau des SCP;
- les stations hydrométriques sont à une distance maximale de 100 km d'un radar opérationnel pour assurer une bonne qualité de la lame d'eau observée issue du signal radar ;
- les surfaces de bassin versant sont inférieures à 1 000 km² et les valeurs de superficie issues de la base de bassins versants du SCHAPI et de la banque HYDRO sont similaires ;
- les bassins versant sont « peu influencés » par les ouvrages hydrauliques ; la sélection est basée sur l'indice de confiance de Caruso (2012), estimé pour la méthode SHYREG en tout pixel de 50 m x 50 m et défini comme le ratio entre la superficie drainée par le(s) barrage(s) situé(s) sur le bassin versant et la surface du bassin versant à l'exutoire ; les bassins pour lesquelles l'indice de confiance est inférieur à 0,2 (influence dite faible selon Caruso (2012)) ont été retenus.

La sélection finale réalisée par Organde (2014) inclut 719 bassins versants sur le territoire national, dont 59 bassins versants pour la zone Meuse-Moselle. Sur ces 59 bassins, seuls 29 bassins sont totalement inclus dans le domaine des prévisions d'ensemble COSMO-DE-EPS mises à disposition par le SPC. La Figure 4 donne la carte des bassins versants, présentés par classes de surface drainée, ainsi que la distribution des surfaces. Les surfaces des 29 bassins versants varient de 8 km² à 944 km², avec une valeur médiane de 225 km². Par la suite, nous étudions l'impact de la surface de bassin sur la qualité des prévisions et des alertes en différenciant les bassins de surface inférieure ou supérieure à 240 km². Les caractéristiques des bassins versants, incluant les identifiants, noms, quantiles de débit de la base de données SHYREG version 1, sont données en Appendice A.

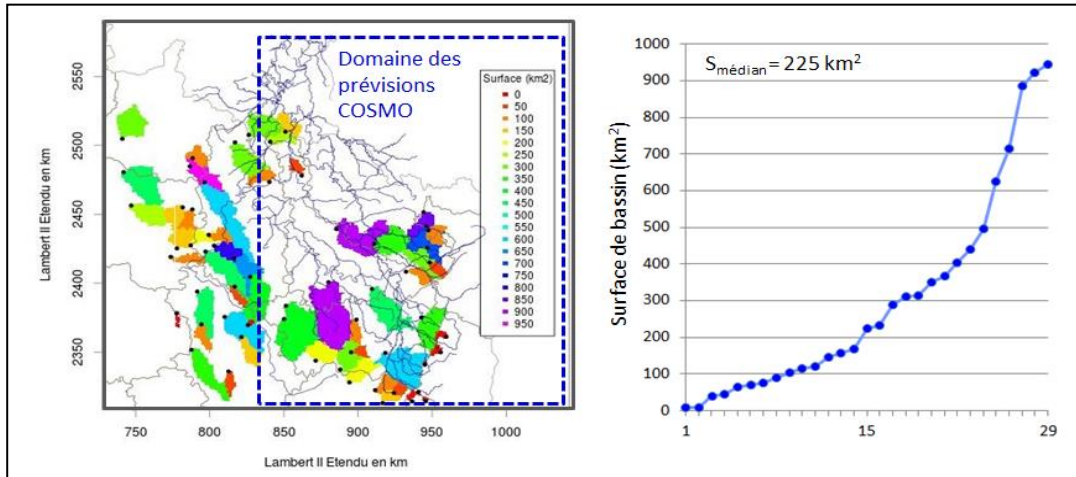


Figure 4 : Carte des 59 bassins versants calibrés de la zone d'étude Meuse-Moselle (par classes de surface de bassin), incluant les 29 bassins versants avec des prévisions COSMO et leur distribution de surface

b) Modélisation hydrologique et données d'observations

Comme le montre le schéma de la Figure 5, la modélisation hydrologique s'appuie sur le couplage de trois modèles (cf. Organde 2014 pour plus de détails) :

- le modèle événementiel horaire semi-distribué, noté GR-SD, basé sur une adaptation du modèle GR distribué du système AIGA actuel pour utiliser une production distribuée, calculée avec un réservoir de production en chaque pixel i du bassin, de capacité A_i , et un transfert global de bassin, estimé avec un unique réservoir de transfert, de capacité B .
- le modèle continu global journalier GR4J à quatre paramètres (Edijatno et al. 1999, Perrin 2002, et Perrin et al. 2003), qui fournit le débit journalier pour l'initialisation du niveau du réservoir de transfert du modèle horaire au début de l'événement ; le taux de remplissage initial du réservoir de transfert de GR-SD (R_0/B) est ajusté quotidiennement si la pluie journalière de bassin de la veille est inférieure à 10 mm afin de restituer le débit issu du modèle journalier global GR4J.
- le modèle continu distribuée journalier GR3J (Edijatno et Michel 1989) qui fournit le taux de remplissage du réservoir $S_i A_i(J)$ au pixel i du bassin; les valeurs de $S_i A_i(J)$ sont estimées à partir des grilles de pluies journalières et des grilles journalières d'ETP calculées à partir de la formule d'Oudin (Oudin et al., 2005) ; le taux de remplissage initial du réservoir de production du modèle GR-SD (S_{i0}/A_i) est ajusté quotidiennement si la pluie journalière de la veille au pixel est inférieure à 10 mm avec la relation suivante : $S_{i0}/A_i = \mathbf{a}_i \cdot S_i A_i(J)$

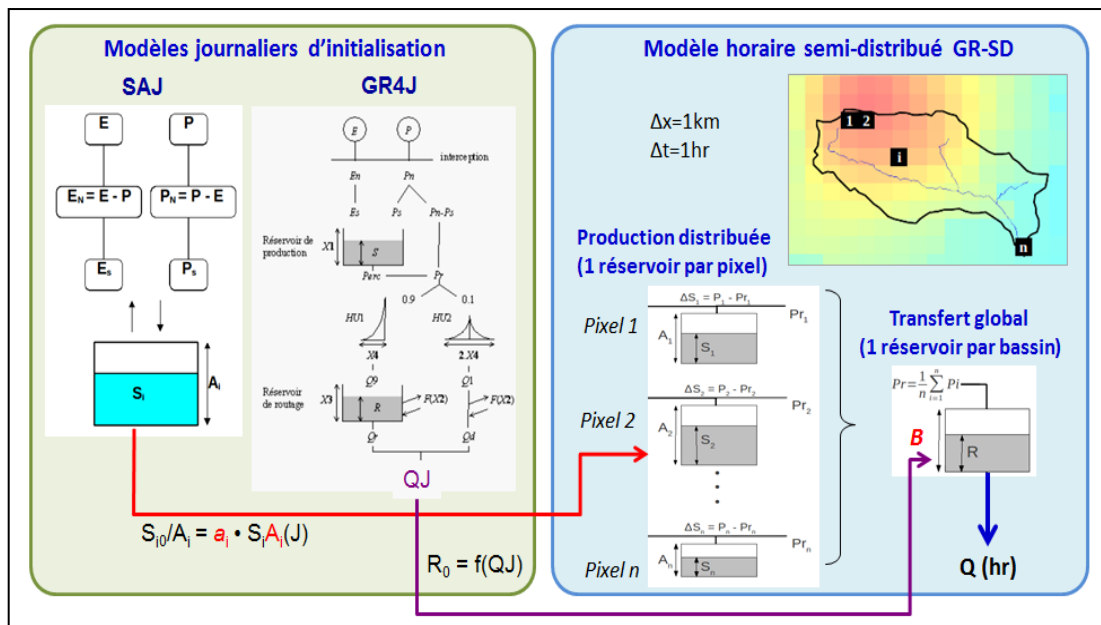


Figure 5 : Schéma du couplage des modèles journaliers continus SAJ et GR4J et du modèle horaire semi-distribué GR-SD pour le système AIGA national

Les paramètres des trois modèles GR-SD, SAJ et GR4J qui doivent être calibrés et régionalisés sont les suivants :

- capacité A_i du réservoir de production de GR-SD et du modèle SAJ, défini en chaque pixel i du territoire,
- paramètre a_i pour l'initialisation du modèle GR-SD, défini en chaque pixel i du territoire,
- capacité B du réservoir de production de GR-SD défini pour chaque bassin versant,
- les 4 paramètres du modèle GR4J définis en chaque bassin versant.

Le calage des modèles GR-SD et GR4J a été réalisé à partir des données de pluie radar ré-analysées de Météo-France (COMEPHORE) sur la période 2001-2006. Les données hydrométriques proviennent de la banque HYDRO et de l'actualisation effectuée par le réseau de SPC à l'automne 2013.

Les procédures et résultats de calibration et de régionalisation sont décrites dans les études Organde 2013 et 2014. La méthode de régionalisation a utilisé la carte des hydro-écorégions définies par Maire et Javelle (2012) à partir de critères climatiques, écologiques et hydrologiques ; cette carte des hydro-écorégions est présentée en Figure 6. Les paramètres régionaux du modèle GR4J sont estimés comme les moyennes des paramètres calibrés des 3 voisins « donneurs » les plus proches (avec la distance euclidienne entre les barycentres des bassins versants). Les paramètres régionaux de GR-SD sont estimés par la méthode de régression en utilisant des critères morphométriques et climatiques.

La Figure 6 montre les résultats de calibration et de régionalisation du modèle semi-distribué GR-SD avec le critère de Nash Pseudo-Continu Décalé. Ce critère correspond au critère de Nash borné calculé sur les débits simulés et observés en mettant les différents évènements de crues bout à bout et avec un décalage de l'hydrogramme simulé pour chaque évènement compris entre 1h et 14h, la valeur retenue du NPCD étant le plus fort Nash borné calculé.

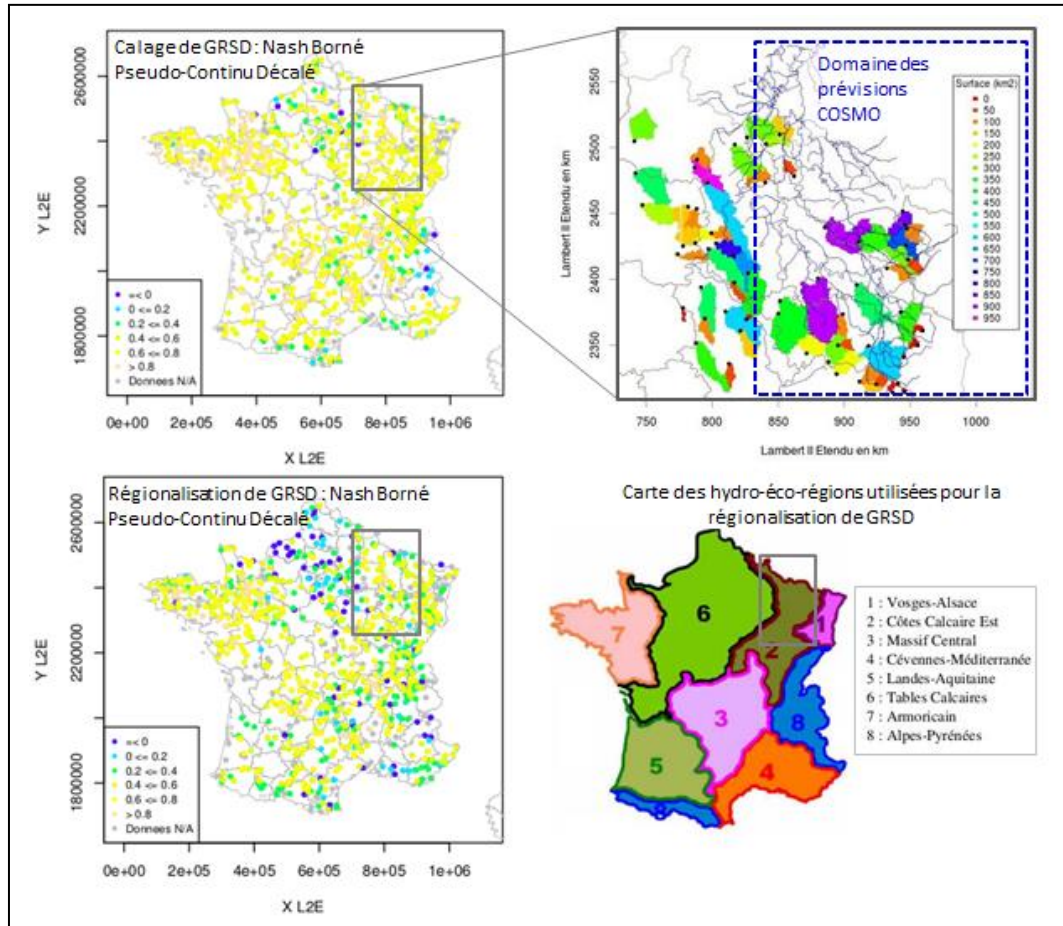


Figure 6 : Résultats du calage et de la régionalisation du modèle horaire semi-distribué GR-SD avec la carte des bassins de la zone Meuse-Moselle et la carte des hydro-écorégions utilisées pour la régionalisation de GR-SD

Pour améliorer la réactivité du modèle GRSD, il a été choisi de fixer la valeur de 2 paramètres avant d'estimer les paramètres \mathbf{a}_i et \mathbf{B} :

- capacité \mathbf{A}_i du réservoir de production de GR-SD: $\mathbf{A}_i = 200$ ou 300 mm,
- valeur minimale du niveau de remplissage initial du réservoir de production : $\mathbf{S}_{i0}/\mathbf{A}_i\text{min} = 0.1$ ou 0.2 .

Ceci permet d'estimer 4 jeux de paramètres issus de ces choix : ($\mathbf{A}_i=200$; $\mathbf{S}_{i0}/\mathbf{A}_i\text{min}=0.1$), ($\mathbf{A}_i=200$; $\mathbf{S}_{i0}/\mathbf{A}_i\text{min}=0.2$), ($\mathbf{A}_i=300$; $\mathbf{S}_{i0}/\mathbf{A}_i\text{min}=0.1$), ($\mathbf{A}_i=300$; $\mathbf{S}_{i0}/\mathbf{A}_i\text{min}=0.2$).

De plus, une optimisation régionale a été réalisée par hydro-écorégion avec le choix de paramètres : $\mathbf{A}_i=200$ mm et $\mathbf{S}_{i0}/\mathbf{A}_i\text{min}=0.2$. L'optimisation utilise le critère de Heidke skill score (HSS) basé sur la détection des dépassements pour les seuils de quantiles de débit de période de retour de 2, 10 et 50 ans (en considérant le tirage aléatoire comme référence).

Par conséquent, les prévisions de débit sont produites avec le jeu de paramètres optimisé régionalement. Pour prendre en considération l'incertitude de paramétrisation du modèle hydrologique, on utilise ce jeu optimisé régionalement, ainsi que les 4 autres jeux de paramètres. Cette analyse préliminaire fait l'hypothèse que ces 5 jeux de paramètres sont équiprobables et peuvent potentiellement permettre de mieux caractériser les incertitudes des

alertes de débit en prenant en compte les incertitudes des pluies futures et des paramètres hydrologiques.

Ce travail d'amélioration de la réactivité du modèle GRSD a été réalisé avec les données d'observations de pluie du fonctionnement de la chaîne d'alerte en temps réel. Ces données spatialisées opérationnelles sont produites par Météo-France à la résolution spatiale de 1 km² à partir des systèmes radars et des stations pluviométriques (Tabary, 2007; Tabary et al., 2007):

- les données « Antilope », qui permettent d'estimer les pluies journalières pour les 2 modèles journaliers SAJ et GR4J ;
- les données « Panthère » au pas de temps de 15 minutes, qui sont utilisées pour le modèle horaire GRSD.

En collaboration avec le SCHAPI, il a été décidé qu'il était plus pertinent d'utiliser les pluies journalières Antilope pour les modèles journaliers car ces pluies Antilope bénéficient des informations issues des postes pluviométriques pour corriger les pluies issues de radar. Les données Panthère accordent un poids supérieur aux lames d'eau radar mais sont les seules données disponibles en temps réel toutes les 5 minutes. Pour Antilope, il y a une différenciation de traitement suivant le type de pluie : les lames d'eau radar sont utilisées telles quelles pour la partie convective des précipitations. Les données de pluviomètres sont spatialisées par krigeage pour la partie stratiforme. A partir d'août 2009, une correction des quantités de pluie issues des lames d'eau radar est introduite : une spatialisation en 1/r² des résidus pluvio-radar au niveau des cumuls convectifs est faite. Le résultat de cette spatialisation est ensuite ajouté à la partie convective radar.

Cette étude utilise également les données journalières Antilope et les données horaires Panthère pour produire les prévisions d'ensemble de débit et les alertes probabilisées.

c) Sélection des événements

La sélection des événements est réalisée à partir des événements de l'étude avec les prévisions AROME (20 événements entre décembre 2009 et décembre 2012), pour lesquels les prévisions COSMO-DE-EPS sont disponibles. Le choix initial des 20 événements, qui est décrit dans l'étude de Demargne (2014b), a été réalisé pour retenir les 20 plus forts événements en termes de pluie observée et/ou débit observé (décrit en termes de nombre de dépassements des quantiles de débit SHYREG pour les périodes de retour de 2 et 10 ans).

Le Tableau 1 donne les caractéristiques des 7 événements retenus pour les 29 bassins versants de l'étude. Sont inclus le nombre de bassins versants pour lesquels le débit observé (Qobs) et le débit simulé (Qsim) dépassent les quantiles de débit biennal et décennal (Q2ans et Q10ans) de la base de quantiles de débit SHYREG (version1).

ID Eve	Date de début	Date de fin	NB[Qobs> Q2ans]	NB[Qsim> Q2ans]	NB[Qobs> Q10ans]	NB[Qsim> Q10ans]
1	2010/12/05 - 03	2010/12/09 - 21	16	11	3	4
2	2011/01/01 - 09	2011/01/09 - 21	5	0	2	0
3	2011/08/04 - 03	2011/08/07 - 21	0	0	0	0
4	2011/12/01 - 06	2011/12/18 - 06	19	23	6	12
5	2012/01/01 - 03	2012/01/06 - 21	10	3	0	0
6	2012/10/03 - 06	2012/10/15 - 21	0	0	0	0
7	2012/12/22 - 03	2012/12/28 - 21	3	2	0	0
Nombre total de dépassements			53	39	14	16

Pas de dépassements obs.
 Dépassements Qobs > Q2ans
 Dépassements Qobs > Q10ans

Tableau 1 : Caractéristiques des 7 événements sélectionnés avec le nombre de dépassements observés et de dépassements du débit simulé pour les quantiles de débit biennal et décennal de la base de données SHYREG (version 1) pour les 29 bassins versants de l'étude

3. Méthodologies

On décrit ici la méthodologie pour prendre en compte l'incertitude des pluies futures et l'incertitude de paramétrisation du modèle hydrologique pour produire des alertes probabilisées, ainsi que la méthodologie d'évaluation de la qualité des alertes et des prévisions d'ensemble.

a) Alertes AIGA probabilisées

La méthodologie pour quantifier les principales sources d'incertitude du système AIGA est de produire des prévisions d'ensemble de débit décrivant des scénarios équiprobables de débit avec les différents membres de l'ensemble. Ces ensembles de débit permettent alors de produire des alertes de dépassement de quantile de débit qui sont associées à une probabilité. Dans cette partie, on décrit comment sont générées les prévisions d'ensemble de débit et les alertes probabilisées.

Estimation des prévisions d'ensemble de débit

Les 2 sources d'incertitude considérées dans cette étude sont :

- l'incertitude des prévisions de pluie, grâce à l'intégration des 20 ensembles de pluie COSMO-DE-EPS,
- l'incertitude de paramétrisation du modèle hydrologique semi-distribué, en utilisant le jeu de paramètres optimisé régionalement, ainsi que 4 autres jeux de paramètres, en faisant l'hypothèse que ces jeux de paramètres sont équiprobables.

La prévision de débit est obtenue en intégrant la pluie observée jusqu'à la date d'initialisation de prévision T0 et ensuite la prévision de pluie dans le modèle hydrologique. Les prévisions d'ensemble de débit sont obtenues à partir des 20 membres de la prévision d'ensemble COSMO-DE-EPS, chaque ensemble de pluie générant un ensemble de débit. On obtient ainsi une prévision d'ensemble de débit avec 20 membres décrivant l'impact de l'incertitude de pluie future. Lorsque l'incertitude de paramétrisation hydrologique est également prise en compte, chacun des 5 jeux de paramètres est utilisé pour produire 20 ensembles de débit. La prévision d'ensemble de débit inclut alors un total de 100 membres décrivant l'impact des incertitudes de pluie future et de paramétrisation hydrologique.

Pour montrer l'apport de la prévision d'ensemble de débit, nous utilisons de prévisions de référence :

- le débit simulé, sans pluie future, qui produit les alertes du système AIGA actuel ; cette « prévision sans échéance » (notée Q_{simu}) correspond au débit simulé obtenu à la date d'initialisation T_0 ,
- le débit simulé obtenu avec la pluie future « parfaite » (en négligeant les erreurs de la pluie observée), connue *a posteriori* ; la prévision de débit issue de la pluie observée est produite pour la même échéance que la prévision d'ensemble (prévision notée $Q[pobs(21hr)]$);
- la prévision de débit issue de la pluie persistante, définie en maintenant la valeur de pluie observée de la dernière heure avant l'initialisation des prévisions ; cette prévision persistante est évaluée pour une échéance de 12 heures ou 3 heures, plus courte que l'échéance des prévisions d'ensemble pour éviter de générer trop de fausses alertes (cette prévision étant notée $Q[pers(12hr)]$ ou $Q[pers(3hr)]$ respectivement);
- et la moyenne des prévisions d'ensemble de débit (notée Q_{moy}), qui permet de montrer l'intérêt d'utiliser une prévision déterministe issue des ensembles de débit.

Les prévisions d'ensemble de débit sont générées toutes les 3 heures lorsque les prévisions de pluie COSMO-DE-EPS sont disponibles. Toutes les prévisions de débit sont produites pour ces mêmes dates d'initialisation.

La Figure 7 présente un exemple des hydrogrammes de débit issus des 20 ensembles de pluie COSMO-DE-EPS, du débit simulé et du débit observé avec une échéance de prévision de 21 heures. Les ensembles de pluie sont représentés par la moyenne et le maximum des 20 ensembles. Les 20 ensembles de débit correspondants sont représentées par leur enveloppe et la moyenne en calculant, pour chacune des 21 heures d'échéance, la valeur minimale, la valeur maximale et la valeur moyenne des 20 membres. Dans cet exemple, les prévisions de pluie ont tendance à surestimer la pluie observée ; les prévisions COSMO de débit dépassent alors le débit biennal en anticipation vis-à-vis du débit simulé et du débit observé. La prévision persistante de débit reste très inférieure au débit simulé et aux ensembles de débit COSMO car pour cette date d'initialisation en avance des fortes pluies, la dernière pluie observée est minimale.

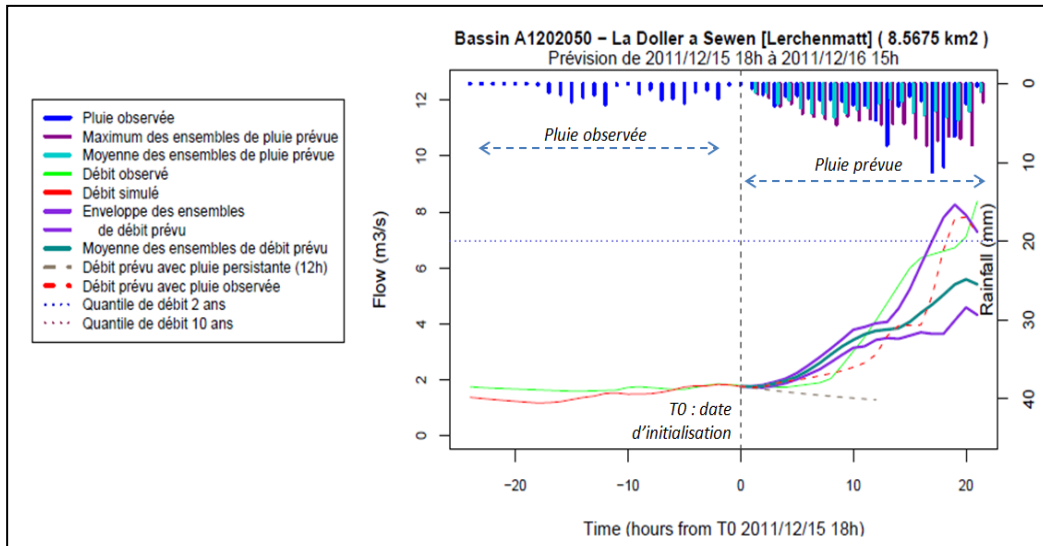


Figure 7 : Exemple d’hydrogrammes avec le débit simulé, l’enveloppe et la moyenne des 20 ensembles de débit COSMO, la prévision de débit issu de la pluie persistante pour la date d’initialisation du 15 décembre 2011 à 18h (TU) en comparaison avec l’hydrogramme de débit observé pour le bassin de La Doller à Sewen

Les Figures 8 et 9 donnent des exemples d’hydrogrammes pour différents bassins et différentes dates d’initialisation pour illustrer la qualité des alertes basées sur les prévisions d’ensemble de débit en comparaison des dépassements observés et des dépassements prévus avec les autres prévisions de référence.

En Figure 8, pour l’exemple du bassin de La Chiers à Longlaville pour la date d’initialisation du 6/01/2011 à 12hr, la pluie observée reste relativement faible alors que le débit observé dépasse le débit biennal. Les prévisions COSMO-DE-EPS surestiment la pluie observée, ce qui permet au pic des ensembles de débit COSMO de dépasser le débit biennal. Même si la date de dépassement est en retard vis-à-vis du dépassement observé, une alerte peut être donnée à 12hr au moment de l’initialisation, donc en avance du dépassement observé. L’exemple de la Semouse pour la date d’initialisation du 7/01/2011 à 6hr montre que la surestimation de la pluie conduit à une surestimation significative du débit, avec la moyenne des ensembles COSMO atteignant le seuil de débit biennal, ce qui produirait une fausse alerte. Dans ce cas, c’est la valeur minimale des ensembles COSMO qui est la plus proche des débits observés. L’exemple du Breuchin à la Poiselere-et-Langle montre l’évolution des prévisions de débit sur 12 heures pendant l’événement de décembre 2011. Le débit simulé surestime le débit observé alors que les ensembles de débit COSMO restent inférieures au débit simulé. On observe de fausses alertes de dépassement du seuil de débit biennal pour la prévision de débit issue de la pluie observée et la prévision persistante de débit, alors que les prévisions de débit COSMO restent inférieures au débit biennal.

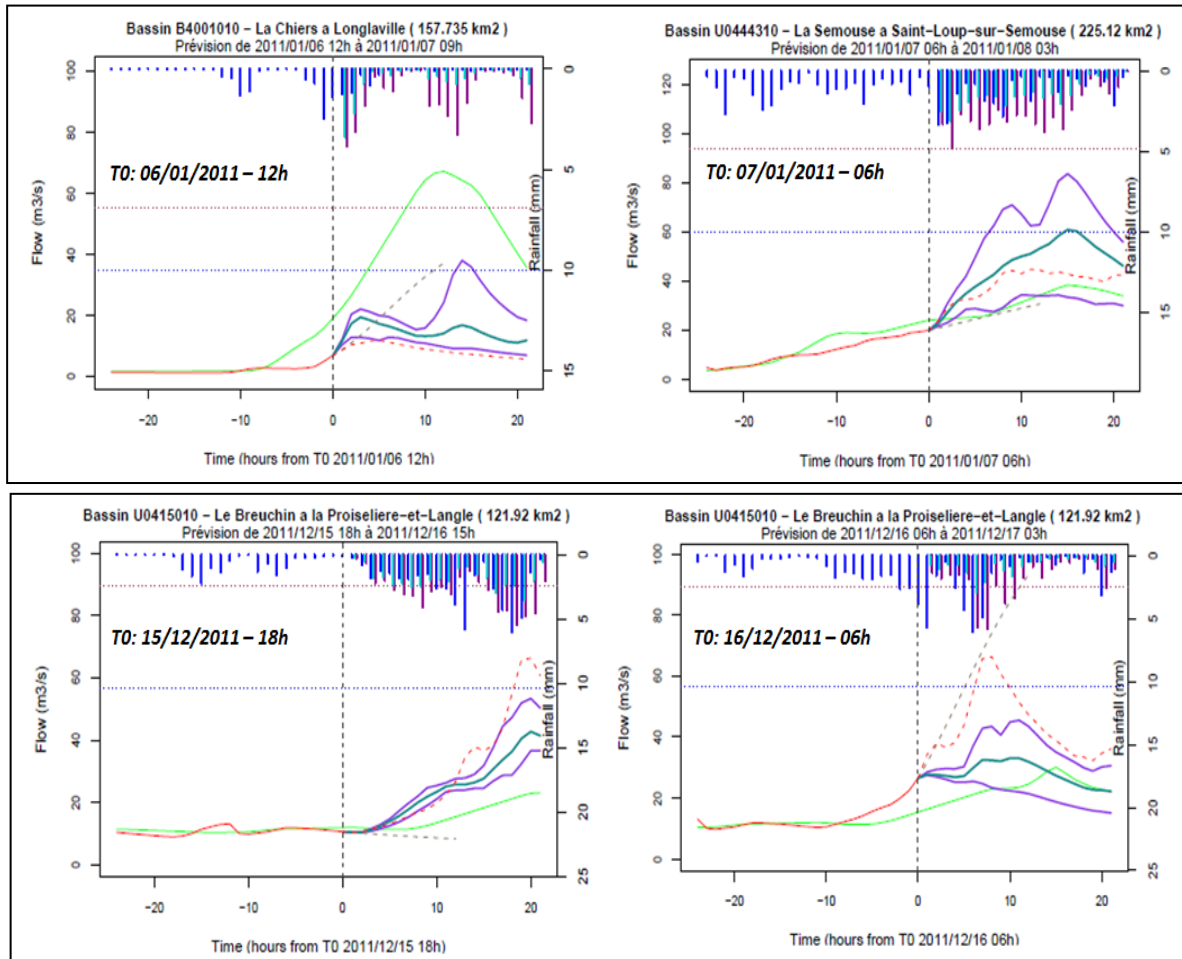


Figure 8 : Exemples d'hydrogrammes pour les événements de janvier 2011 et décembre 2011

La Figure 9 donne des exemples pour l'événement de décembre 2012. L'exemple pour le bassin de La Chiers à Longlaville montre une anticipation du dépassement du débit biennal avec les ensembles de débit COSMO grâce à des prévisions fortes de pluie. Ces erreurs de pluie permettent une bonne alerte basée sur le pic des ensembles COSMO pour la date d'initialisation du 5/01/2012 à 3hr, alors que le dépassement n'est pas confirmé avec les prévisions COSMO de 9hr. Le bénéfice de surestimation des pluies COSMO est également illustré pour le bassin de La Saône à Monthureux-sur-Saône. Les prévisions COSMO permettent d'anticiper la montée des débits, même si les pics prévus de débit restent inférieurs au pic du débit observé. Mais ce dépassement du débit biennal n'est pas maintenu avec les prévisions COSMO de 9hr ; à cette date, seule la prévision persistante dépasse le débit biennal en cohérence vis-à-vis du débit observé.

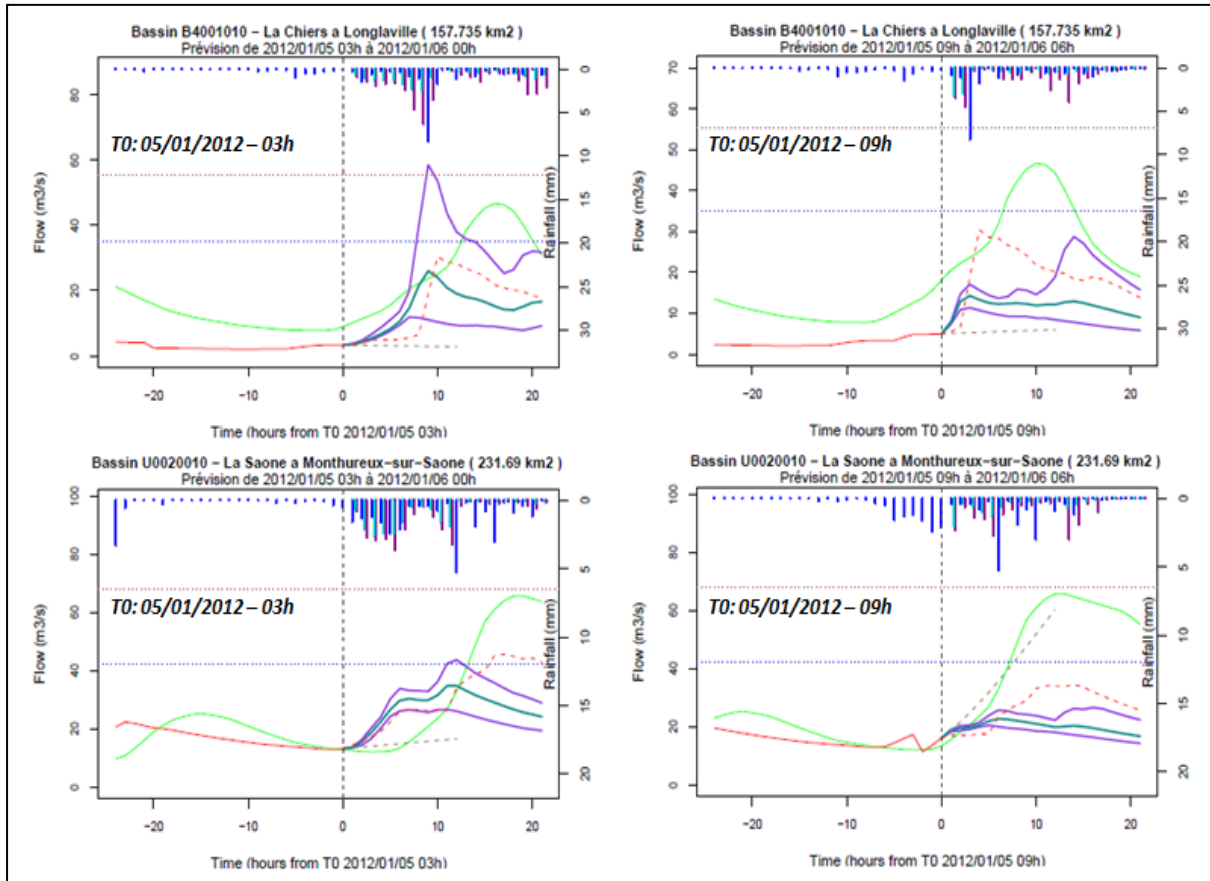


Figure 9 : Exemples d'hydrogrammes pour l'événement de décembre 2012

Alertes probabilisées avec les prévisions d'ensemble de débit

Pour les alertes de dépassement d'un seuil de débit, les prévisions d'ensemble de débit conduisent à estimer une probabilité de dépassement du seuil en fonction du nombre de membres dépassant le seuil. Les membres de l'ensemble sont considérés comme tous équiprobables et nous utilisons la loi de probabilité empirique pour définir la probabilité pour chaque membre de la façon suivante :

$$P[Q \leq Q_{i_{ens}}] = i_{ens} / (nb_{ens} + 1)$$

avec nb_{ens} le nombre total des membres de l'ensemble

Ainsi, la distribution empirique de probabilité pour la prévision d'ensemble de débit de dépasser un seuil de débit Q_{seuil} est estimée avec la formule suivante :

$$P[Q \geq Q_{seuil}] = i_{ens} / (nb_{ens} + 1)$$

avec i_{ens} le nombre de membres de l'ensemble dépassant le seuil de débit et nb_{ens} le nombre total des membres de l'ensemble

La Figure 10 donne un exemple de l’alerte probabilisée issue de la prévision d’ensemble avec 20 membres pour le dépassement du débit biennal pour la date du 15/12/2011 à 16hr. Pour cette carte, nous avons choisi de sélectionner 5 niveaux de probabilité correspondant aux dépassements prévus pour :

- 1 membre, la probabilité de dépasser le seuil étant de $P[Q > Q_{seuil}] = 1/21 = 0.05$,
- 5 membres ($P[Q > Q_{seuil}] = 5/21 = 0.24$),
- 10 membres ($P[Q > Q_{seuil}] = 10/21 = 0.48$),
- 15 membres ($P[Q > Q_{seuil}] = 15/21 = 0.71$),
- 19 membres ($P[Q > Q_{seuil}] = 19/21 = 0.9$).

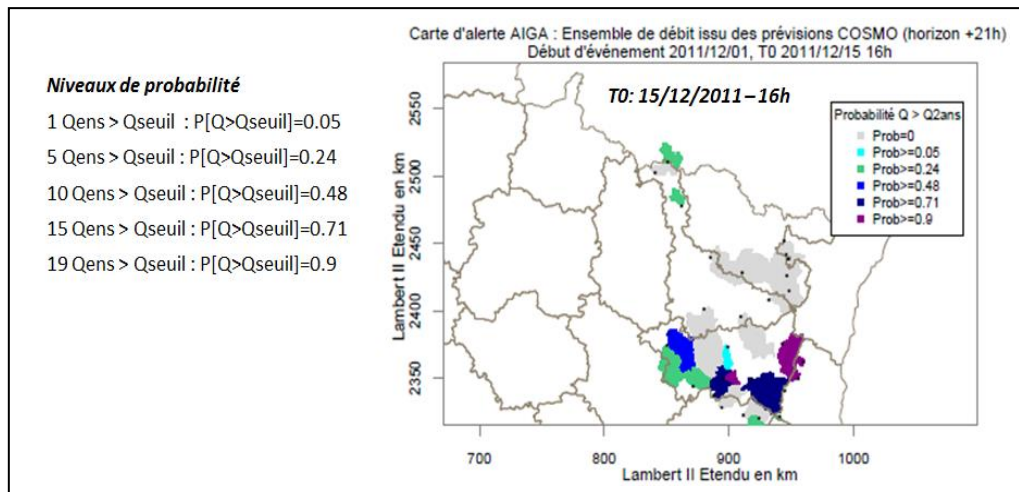


Figure 10 : Exemple de carte d’alerte issue des 20 ensembles de débit COSMO pour le dépassement du débit biennal, en utilisant 5 niveaux de probabilité définis à partir de la distribution de probabilité empirique, pour la date du 15/12/2011 16hr (TU)

Pour l’utilisation des prévisions d’ensemble dans la chaîne d’alerte AIGA, les prévisionnistes doivent pouvoir choisir les niveaux de probabilité qu’ils veulent représenter. Dans cette étude, comme les ensembles COSMO peuvent être généralement regroupés en fonction du modèle atmosphérique global utilisé dans le système COSMO-DE-EPS, nous avons choisi les seuils de 5 membres, 10 membres, 15 membres et 19 membres. Le choix d’un seul ensemble permet de décrire les valeurs maximales issues des 20 membres.

Les prévisions de débit sont produites toutes les 3 heures lorsque les prévisions d’ensemble de pluie COSMO-DE-EPS sont disponibles. Mais les alertes probabilisées sont estimées toutes les heures en considérant les échéances disponibles de la prévision la plus récente. Ainsi, les alertes probabilisées sont produites pour les prévisions suivantes :

- pour la date d’initialisation T0, la prévision d’ensemble de débit sur les 21 heures prochaines, de T0+1hr à T0+21hr,
- pour T0+1hr, la prévision des 20 heures d’échéance de T0+2hr à T0+21hr,
- pour T0+2hr, la prévision des 19 heures d’échéance de T0+3hr à T0+21hr.

Il faut noter que, pour la prévision sans pluie future et la prévision issue de la pluie observée, elles sont toutes les deux issues des valeurs de débit simulé. Mais la prévision issue de la pluie observée ayant la même échéance que les prévisions d’ensembles, elle inclut, pour la date d’initialisation T0, 21 valeurs de débit simulé de T0+1hr jusqu’à T0+21hr. La prévision sans pluie future est la valeur du débit simulé à T0.

La Figure 11 donne des exemples des cartes d’alerte probabilisées pour les 2 événements les plus forts de décembre 2010 et 2011. Ces cartes montrent l’évolution de la probabilité de dépasser le seuil de débit biennal pour les 29 bassins versants sélectionnés en fonction de la date de l’alerte T0.

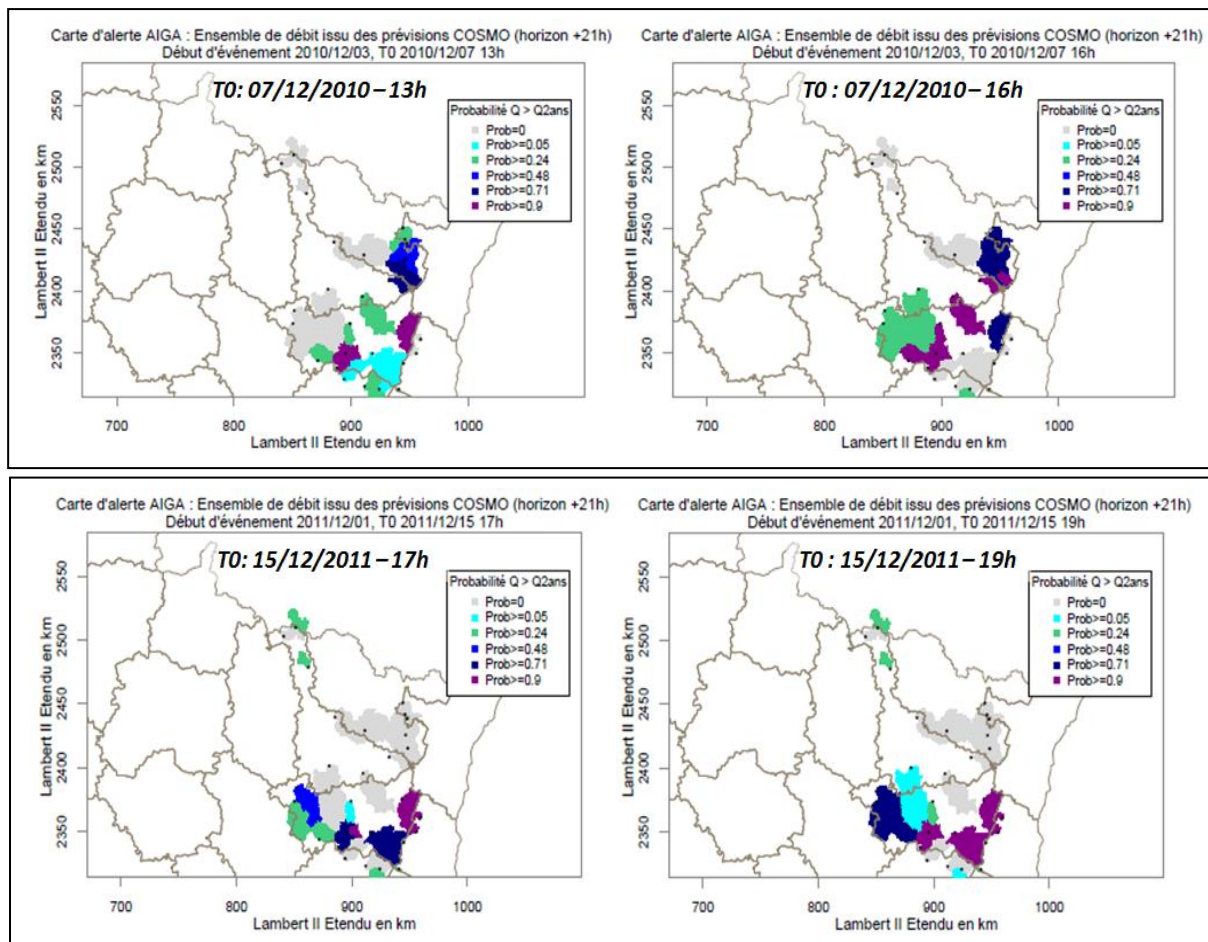


Figure 11 : Exemples de carte d’alerte issue des 20 ensembles de débit COSMO pour le dépassement du débit biennal pour les événements de décembre 2010 et décembre 2011

Les cartes d’alerte sont aussi produites pour les prévisions de référence : débit simulé correspondant aux alertes du système AIGA actuel sans pluie future, débit issu de la pluie parfaite, et prévision persistante de débit. L’évaluation a pour objectif de comparer la qualité des avertissements du système AIGA actuel avec celle des avertissements issus des prévisions d’ensemble de débit et des prévisions de référence.

b) Evaluation de la qualité des alertes

L’analyse de la qualité des alertes nécessite d’intégrer plusieurs critères de vérification pour quantifier les différents aspects de leur qualité, en termes notamment de fiabilité, justesse, et anticipation (Jolliffe et Stephenson 2003, Wilks 2011, WMO 2013). La qualité des alertes utilise les scores de contingence des dépassements de seuil par le débit observé et par les différentes prévisions de débit. Les scores sont calculés pour chaque type de prévision pour le jeu d’évènements étudiés et les différents bassins. L’évaluation concerne également le degré d’anticipation des alertes vis-à-vis des dépassements de seuil observés.

Dans cette étude, on évalue les alertes issues des prévisions de débit, ainsi que les alertes du système AIGA actuel (sans pluie future), par comparaison avec les dépassements observés en utilisant une unique valeur de seuil de débit : le débit biennal issu de la base de données SHYREG version 1.

Cette évaluation des alertes avec les dépassements de débit observés permet de considérer l'impact de l'ensemble des sources d'incertitude et d'erreur sur les débits futurs. On peut remarquer qu'une évaluation réalisée avec les alertes issues du débit simulé (à partir de la pluie observée) permet de ne pas prendre en compte les sources d'incertitude et d'erreur liées au modèle hydrologique puisque le débit simulé comme les prévisions d'ensembles sont impactés par les erreurs du modèle hydrologique.

Il faut également noter que les alertes de dépassement d'un seuil de débit pour les prévisions peuvent être définies non pas avec le quantile de débit, mais avec une valeur de seuil issue des simulations de débit avec le modèle hydrologique. Cette approche permet de prendre en compte le biais potentiel du modèle ; dans ce cas, la valeur de seuil de débit pour définir un dépassement diffère entre les débits simulés ou prévus et les débits observés. Mais cette approche n'est pas actuellement mise en œuvre dans la chaîne d'alerte AIGA.

Scores de contingence des dépassements de seuil de débit

Le tableau de contingence se définit pour un événement donné, c'est-à-dire pour un seuil fixé (et unique) qui définit si oui ou non l'observation ou la prévision a dépassé le seuil considéré. A noter que la prévision dépasse le seuil si, pour une des échéances, il y a dépassement. De plus, on comptabilise les dépassements du débit observé et du débit prévu pour chaque événement dans son ensemble (i.e. pour toutes les dates d'initialisation de prévision de l'évènement). Sur le jeu d'évènements étudiés, on comptabilise les dépassements de seuil issus des observations et des prévisions de la façon suivante :

Evènement : dépassement d'un seuil donné		Evènement observé	
		Oui	Non
Evènement prévu	Oui	a (succès/hit)	b (fausse alerte/false alarm)
	Non	c (manqué/miss)	d (rejet correct/ true negative)

A partir du tableau de contingence, on définit les scores d'évaluation suivants (WMO 2013):

- **Biais en fréquence** ou **Frequency Bias** (FB), qui décrit si les prévisions ont tendance à produire trop de fausses alertes ou trop d'alertes manquées : $FB = (a+b)/(a+c)$; score optimal $FB=1$.
- **Taux de bonne détection** ou **Probability Of Detection** (POD) (encore appelé Hit Rate), qui est conditionné par les dépassements observés : $POD = a/(a+c)$; le score varie entre 0 (toutes les alertes observées sont manquées) et 1 (toutes les alertes observées sont détectées), avec un score optimal égal à 1.
- **Probability Of False Detection** (POFD) (ou False Alarm Rate), qui est conditionné par les non-dépassements observés : $POFD = b/(b+d)$; le score varie entre 0 (aucune fausse alerte) et 1 (des alertes prévues pour tous les non-dépassements observés), avec un score optimal égal à 0.

- **Taux d'alerte correcte** ou **Success Ratio (SR)**, qui est conditionné par les alertes prévues pour mesurer la fiabilité des prévisions : $SR=a/(a+b)$; le score varie entre 0 (aucune alerte prévue qui soit correcte) et 1 (aucune fausse alerte), avec un score optimal égal à 1.
- **Critical Success Index (CSI)** qui permet de considérer les alertes manquées et les fausses alertes (mais sans différencier ces erreurs); $CSI=a/(a+b+c)$; le score varie entre 0 et 1, avec un score optimal égal à 1.

Prise en compte de l'anticipation des alertes

L'analyse de la qualité des alertes se base également sur l'anticipation des alertes issues des prévisions. L'anticipation est définie vis-à-vis des alertes observées, i.e. par rapport à la date à laquelle le débit observé dépasse le seuil de débit biennal. L'anticipation est positive lorsque la prévision à une date d'initialisation T_0 produit un dépassement du seuil de débit avant la date de dépassement effectif par le débit observé. Pour cette analyse relative à l'ensemble d'un évènement, l'anticipation est définie par la valeur la plus grande relative à la première date de dépassement par la prévision de débit. L'anticipation est négative lorsque la première date du dépassement prévu T_0 intervient après le dépassement effectif du débit observé. On peut donc comparer sur un même jeu de bassins l'anticipation des alertes issues des différents types de prévisions.

Comme on cherche à optimiser l'anticipation des alertes vis-à-vis des dépassements observés, les critères de contingence sont calculés en prenant en compte l'anticipation du dépassement de seuil par la prévision. Ainsi, un dépassement de seuil est prévu avec retard si la première date T_0 de dépassement prévu est plus tardive que la date du dépassement effectif du débit observé ; il est alors considéré comme une alerte manquée. Le dépassement de seuil est comptabilisé comme une bonne alerte si la prévision de dépassement a une anticipation égale ou supérieure à 0.

c) Vérification des prévisions d'ensemble

Pour compléter l'analyse de la qualité des alertes, les prévisions d'ensemble de pluie de bassin et de débit sont vérifiées pour quantifier la similarité entre les observations et les prévisions. Cette évaluation de la qualité des prévisions d'ensemble est réalisée avec le système de vérification Ensemble Verification System (EVS) (Brown et al. 2010, téléchargeable sur le site www.nws.noaa.gov/oh/evs.html).

L'outil de vérification EVS permet d'évaluer les variables hydrométéorologiques (ex. pluie, température, débit) pour différentes échéances et à de multiples échelles temporelles (ex. horaire, journalier). La qualité d'une prévision peut être analysée en fonction de différents facteurs tels que l'échéance de prévision, la saisonnalité, l'agrégation temporelle, la magnitude de l'évènement (ex. dépassement d'un seuil), ou la valeur d'une autre variable (ex. qualité de la prévision de débit en fonction de la quantité de pluie observée ou prévue). EVS inclut un grand nombre de métriques de vérification pour analyser les différents aspects de fiabilité, justesse, discrimination, et gain vis-à-vis d'une prévision de référence. Ces différentes métriques permettent d'analyser la qualité des prévisions grâce à des scores et des graphiques plus ou moins détaillés. Il est également possible d'agréger les prévisions à de multiples échelles temporelles (ex. du pas de temps horaire au pas de temps tri-horaire). On peut également agréger les performances sur un jeu de bassins, soit en calculant les scores de

vérification à partir de toutes les paires observations-prévisions de l'ensemble des bassins, soit en calculant une moyenne pondérée des scores de vérification obtenus pour chacun des bassins.

Concernant le choix des scores de vérification, les paires observations-prévisions pour les 7 événements de cette étude constituent un échantillon de taille modeste. C'est pourquoi on a sélectionné des critères d'évaluation globaux, qui ne demandent pas un sous-échantillonnage des paires observations-prévisions (comme par exemple avec le critère de reliability diagram). De plus, comme les bassins versants ont des surfaces très variables, on privilégie des mesures relatives de vérification ou des scores de gain de performance par rapport à une référence, ce qui permet d'agréger les résultats sur l'ensemble des bassins pour obtenir des résultats plus robustes. Pour analyser l'apport des prévisions d'ensemble vis-à-vis des prévisions déterministes de référence (la prévision de débit issue de la pluie observée et la prévision persistante de débit), les scores choisis peuvent être estimés pour des prévisions de type ensembliste ou déterministe.

Pour évaluer les différents aspects de la qualité des prévisions avec un petit nombre de scores, on a sélectionné les critères suivants :

- **l'erreur relative moyenne** (estimée pour la prévision déterministe ou pour la moyenne des ensembles), qui permet de quantifier le biais inconditionnel ou biais relatif des prévisions en mesurant la différence moyenne entre les prévisions et les observations, exprimée en termes de fraction de la moyenne des observations ; plus le critère est proche de 0, moins la prévision est biaisée ; une erreur relative moyenne négative / positive indique une sous-estimation / surestimation respectivement de la prévision déterministe ou de la moyenne des ensembles ;
- **le coefficient de corrélation de Pearson**, estimée pour la prévision déterministe ou pour la moyenne des ensembles, qui décrit la relation linéaire entre les prévisions et les observations ; plus le critère est proche de 1, meilleure est la relation linéaire entre les prévisions et les observations (mais les prévisions peuvent être biaisées vis-à-vis des observations) ;
- **le Continuous Rank Probability Skill Score**, score de gain en performance pour le critère de Continuous Rank Probability Score (CRPS) ; le CRPS permet de mesurer l'erreur moyenne en probabilité de la prévision d'ensemble et correspond à l'erreur moyenne absolue pour une prévision déterministe ; la Figure 12 donne une représentation graphique de l'estimation du CRPS pour une prévision donnée par comparaison des distributions cumulées de probabilité de la prévision et de l'observation ; le CRPS est ensuite moyenné sur l'ensemble des paires observations-prévisions ; le score CRPSS est alors calculé en comparant le CRPS de la prévision avec le CRPS d'une référence avec la formule suivante :
$$\text{CRPSS} = 1 - [\text{CRPS}_{\text{prévi}} / \text{CRPS}_{\text{ref}}]$$
 ; une valeur négative / positive de CRPSS indique que la prévision a une erreur en probabilité moins / plus forte que la référence respectivement ;
- **le Relative Operating Characteristic Score (ROCS)**, qui quantifie la capacité de discrimination des prévisions pour un événement donné ; ce critère est estimé à partir de la courbe ROC de la probabilité de détection (POD) en fonction de la probabilité de fausse détection (POFD), qui est représentée en Figure 12 pour un événement donné (ici, le dépassement d'un seuil) ; pour une prévision ensembliste, on doit définir différents niveaux de probabilité pour obtenir une prévision binaire (dépassement prévu ou non) qui est alors comparée à l'observation (dépassement

observé ou non) ; une prévision parfaite est définie par $POD=1$ et $POFD=0$; la prévision doit avoir une courbe ROC au-dessus de la diagonale (pour laquelle $POD=POFD$) pour pouvoir correctement prévoir les dépassements en évitant les fausses alarmes pour les non-dépassements ; le critère ROCS correspond à 2 fois l'aire entre la courbe ROC et la diagonale ; plus le ROCS est proche de 1, plus la prévision différencie les dépassements et les non-dépassements.

Pour définir des seuils de dépassements cohérents vis-à-vis des échantillons de prévisions-observations sur les différents bassins, on définit les seuils par les quantiles des observations à partir de la distribution empirique sur la période de vérification 01/01/2009 – 31/12/2012. Dans cette étude, on s'intéresse aux quantiles observés 75% et 90% pour la pluie de bassin et le débit, ce qui permet d'assurer une même taille d'échantillon pour chacun des bassins. A noter que l'utilisation du seuil de débit biennal aurait nécessité un échantillon de paires prévisions-observations plus large pour obtenir des résultats statistiques robustes.

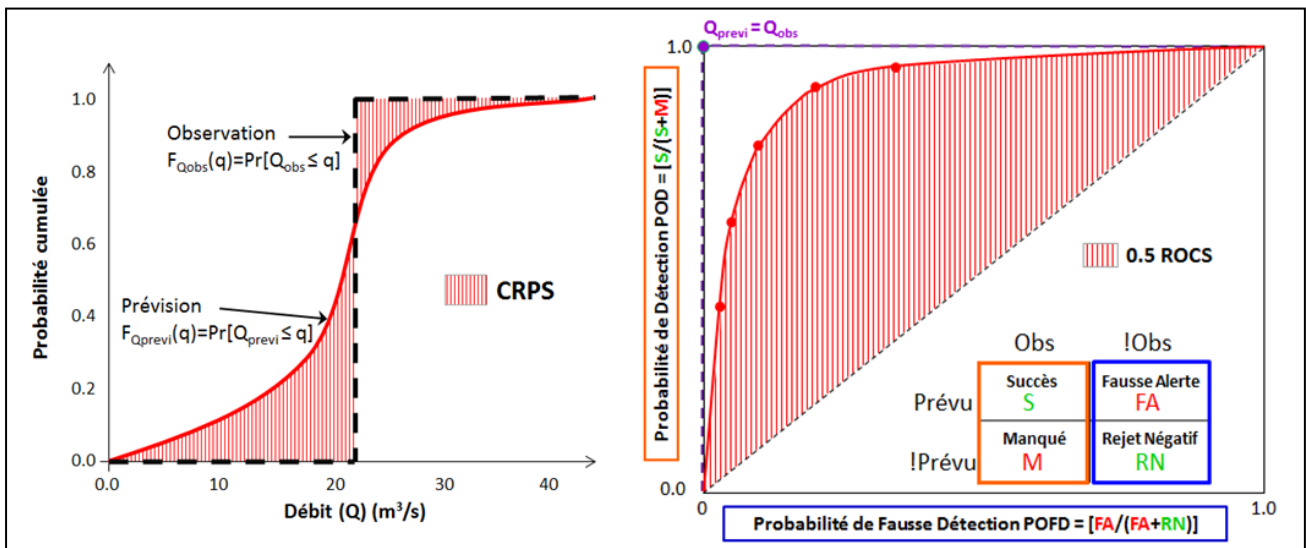


Figure 12 : Estimation des scores de Continuous Rank Probability Score (CRPS) et de Relative Operating Curve Score (ROCS) pour une prévision probabiliste

Les quantiles observés pour le score de discrimination sont également utilisés pour présenter les scores d'évaluation pour les sous-échantillons des paires prévisions-observations dont les observations sont supérieures aux quantiles. Ceci permet de mieux analyser la qualité des prévisions pour les valeurs observées les plus fortes, ce qui est le focus du système d'alerte aux crues.

Pour le score de performance CRPSS, 2 prévisions de référence ont été utilisées :

- la prévision probabiliste climatologique, définie par la distribution de probabilité empirique de toutes les observations sur la période de vérification 01/01/2009 – 31/12/2012 ; cette référence permet de quantifier la performance des prévisions d'ensemble pour les situations les moins « communes » de l'échantillon ;
- la prévision déterministe persistante de débit, qui est issue de la prévision de pluie persistante (définie par la dernière valeur observée de pluie) ; cette prévision sera généralement performante en cours d'événement lorsque l'on s'approche d'un potentiel dépassement ; pour éviter un trop grand nombre de fausses alertes, la prévision persistante est utilisée pour une échéance maximale de 12 heures.

4. Résultats

Les résultats de l'évaluation sont présentés tout d'abord pour la qualité des prévisions d'ensemble de pluie de bassin et de débit et, ensuite, pour la qualité des alertes de dépassement du seuil de débit biennal. Les résultats d'évaluation des prévisions d'ensemble ont été obtenus pour les 7 événements et pour les 27 bassins pour lesquels les échantillons de paires observations-prévisions concernent les mêmes dates (les résultats des bassins A9091060 et U0444310 n'ont pu être inclus aux résultats). Les statistiques présentées ci-dessous sont donc basées sur des échantillons réduits de bassins et de paires observations-prévisions. En faisant l'hypothèse que les scores d'évaluation ont des caractéristiques homogènes pour tous les bassins de la zone d'étude, les statistiques sont présentées après agrégation des résultats en considérant la moyenne des statistiques sur les 27 bassins (un même poids étant donné à chaque bassin). Les résultats de la qualité des alertes sont basés sur les dépassements du débit biennal pour les 29 bassins versants et les 7 événements. Tous ces résultats mériteraient d'inclure un échantillon plus large de bassins et d'événements pour améliorer leur robustesse mais sont indicatifs de l'apport potentiel des prévisions d'ensemble de pluie COSMO pour les alertes aux crues soudaines dans la zone concernée.

a) Qualité des prévisions d'ensembles de pluie de bassin

Les prévisions d'ensembles COSMO-DE-EPS permettent de produire des grilles de prévisions de pluie pour chaque bassin versant. Ces prévisions de pluie sont vérifiées en pluie de bassin pour les échéances de 1hr à 21hr et pour leurs cumuls sur 3hr, 6hr, 12hr et 21hr. Les observations de pluie sont issues des données Panthère qui sont utilisées en entrée du modèle horaire semi-distribué GR-SD. Les résultats présentés incluent les histogrammes des erreurs des prévisions d'ensemble en fonction de la pluie observée et les scores de vérification. Les scores sont calculés pour chaque bassin pour toutes les paires observations-prévisions, ainsi que pour les paires dont les observations sont supérieures aux quantiles observés 75% et 90% ; les scores sont ensuite moyennés sur les 27 bassins. Les échantillons pour chaque bassin sont en moyenne de 462 paires au total, 121 paires pour le quantile observé 75%, et 49 paires pour le quantile observé 90%.

Erreurs des ensembles de pluie en fonction de la pluie observée

La Figure 13 présente les histogrammes des erreurs des prévisions d'ensemble en fonction de la pluie observée pour identifier les éventuels biais ou valeurs aberrantes. Pour chaque paire d'observation-prévision (qui inclut 20 membres), on calcule les erreurs des 20 membres de la prévision ; la distribution de ces 20 erreurs est alors représentée par une boîte à moustache avec les valeurs des quartiles, le minimum et le maximum. Toutes ces distributions d'erreur sont représentées pour chaque paire d'observation-prévision en fonction de la valeur observée. Les graphes d'erreur ci-dessous incluent les résultats des 27 bassins. Idéalement, les erreurs des prévisions doivent être faibles et capturer la valeur nulle. Les erreurs sont présentées pour différentes agrégations (cumuls de pluie sur 1hr, 3hr, 6hr, 12 et 21hr) et en fonction de l'horizon de prévision.

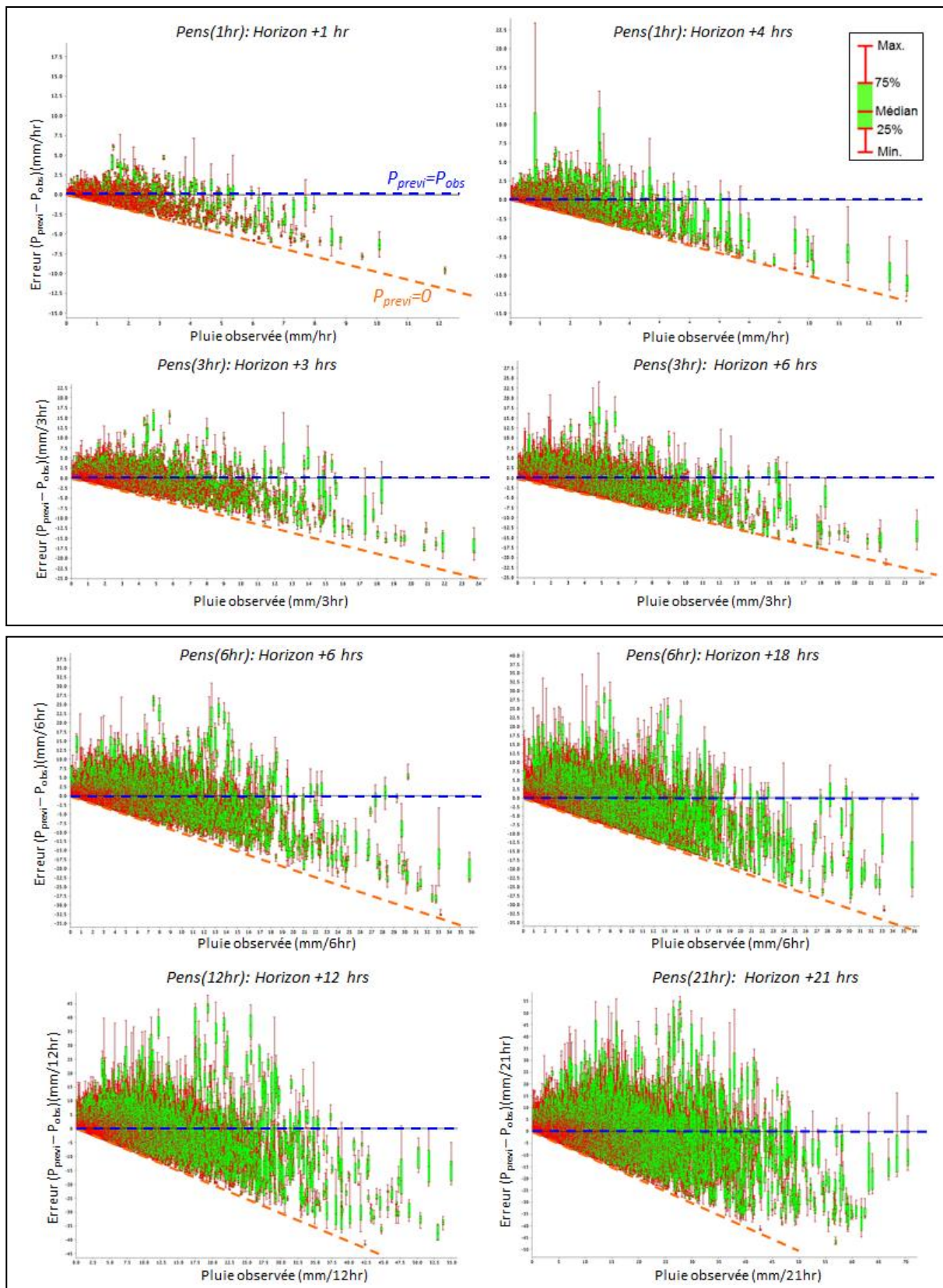


Figure 13 : Histogrammes des erreurs des prévisions d'ensemble de pluie de bassin en fonction de la pluie observée pour différentes agrégations (cumuls de pluie sur 1hr, 3hr, 6hr, 12 et 21hr) et en fonction de l'horizon de prévision

Le premier graphe de la Figure 13, qui concerne la pluie horaire pour la première échéance, montre une sous-estimation des fortes valeurs observées et une surestimation des faibles précipitations, ce qui démontre un biais conditionnel des prévisions de pluie (conditionnel car dépendant de la valeur observée). Les ensembles souffrent également d'un manque de dispersion (la valeur observée n'est pas incluse dans l'intervalle [minimum, maximum] des 20 membres). Pour l'échéance de +4hrs du deuxième graphe, la dispersion des ensembles permet de mieux approcher ou inclure la valeur observée dans la distribution des 20 membres mais certaines erreurs sont également bien plus larges. Pour les autres graphes qui concernent les cumuls de pluie, on peut observer une tendance à la dégradation des erreurs, la dispersion des membres étant cependant meilleure pour les plus longues échéances.

Les figures suivantes concernent l'impact de l'échéance de la prévision ainsi que l'impact de la période d'accumulation des pluies sur les scores de vérification des ensembles de pluie.

Impact de l'échéance sur la qualité des ensembles de pluie de bassin

Les scores d'évaluation sont représentés en fonction de l'échéance de prévision dans les graphes ci-dessous. On calcule également les scores pour les sous-échantillons pour lesquels les observations sont supérieures aux quantiles observés 75% et 90%. Ces quantiles observés sont les valeurs de seuil pour définir les dépassements pour le score de discrimination ROCS.

L'erreur relative moyenne en Figure 14 de la prévision horaire montre une sous-estimation générale des pluies observées, avec une dégradation avec l'échéance de prévision, notamment une dégradation après les premières 4 ou 5 échéances. La dégradation est plus forte pour les observations les plus fortes. On peut également observer une oscillation cyclique toutes les 3 heures, cette oscillation provenant sans doute du système COSMO-DE-EPS (peut-être dû à une désagrégation de certaines informations). Les coefficients de corrélation se dégradent également avec l'échéance et avec le seuil de quantile observé. Les scores de CRPSS en référence à la prévision probabiliste climatologique montrent une performance moins bonne des prévisions d'ensemble lorsque l'on considère tout l'échantillon de prévisions-observations, mais une meilleure performance pour les observations les plus fortes. Le score de discrimination ROCS pour les 2 valeurs de seuil (définies par les quantiles observés 75% et 90%) montre une bonne discrimination par les ensembles de pluie des dépassements des seuils de pluie de bassin, même si cette capacité de discrimination est dégradée pour le seuil le plus élevé.

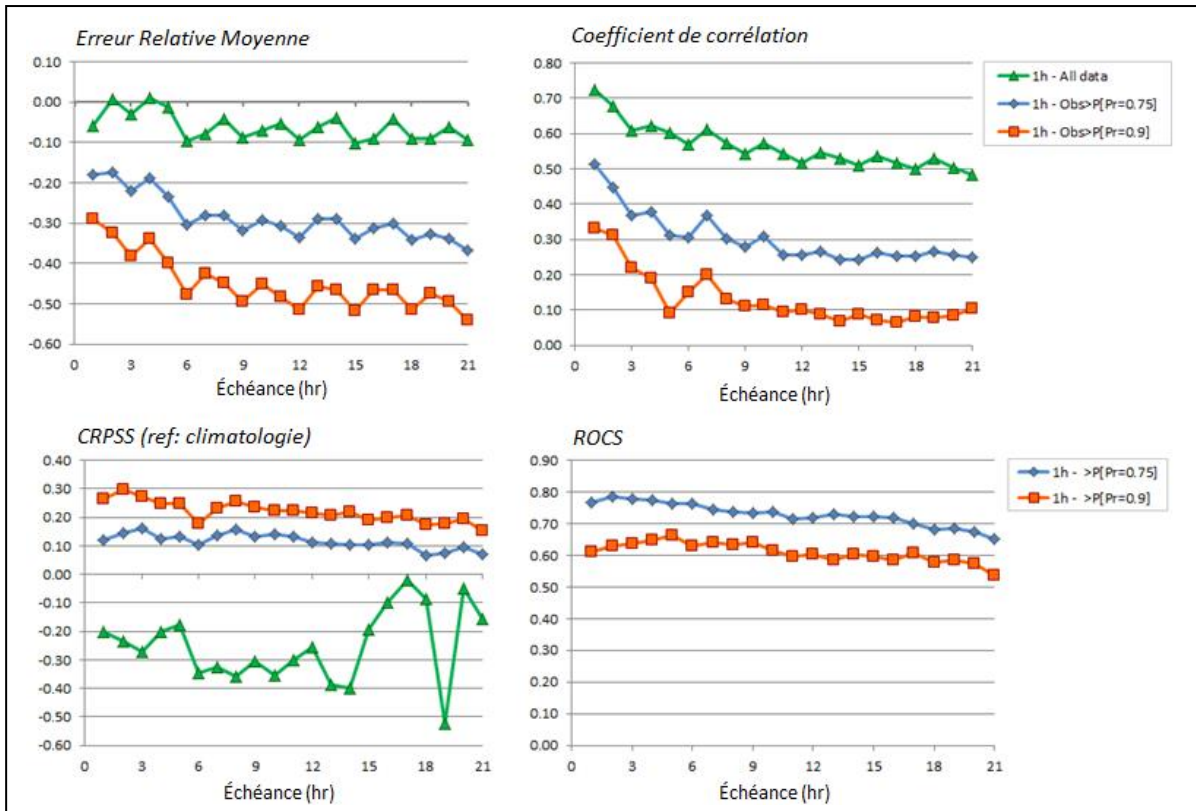


Figure 14 : Impact de l'échéance sur les scores de vérification des prévisions d'ensemble de pluie de bassin pour tout l'échantillon et pour les seuils des quantiles observés 75% et 90%

Impact de la période d'agrégation sur la qualité des ensembles de pluie de bassin

La Figure 15 compare les scores d'évaluation en fonction de la période de cumul de la pluie de bassin pour seules les observations supérieures au quantile observé 90%. Il faut noter que généralement l'agrégation améliore les scores d'évaluation par effet de lissage (les erreurs pouvant se compenser sur les périodes plus longues).

Les erreurs moyennes relatives montrent que la sous-estimation demeure pour tous les cumuls pour les observations supérieures au quantile observé 90%. Cependant les erreurs pour les différents cumuls de 3 heures à 21 heures pour l'échéance égale à la période d'agrégation sont toutes de l'ordre de -25% et donc sont légèrement meilleures que l'erreur de -29% de la première prévision horaire. [A noter que les erreurs moyennes relatives pour le cumul sur 21 heures est de -7% pour toutes les paires prévisions-observations et de -18% pour le quantile observé 75%, c'est-à-dire similaires aux erreurs de la pluie horaire de la première échéance de la Figure 14]. On observe des résultats similaires pour le coefficient de corrélation (à l'exception du cumul sur 6 heures dont le score est dégradé). Le gain en CRPS est significativement amélioré avec le cumul des pluies sur 3 à 21 heures, avec des gains supérieurs à 35% pour les différents cumuls pour l'échéance égale à la période d'agrégation. La discrimination pour les dépassements du quantile observé 90% est légèrement améliorée avec les cumuls des pluies, avec un score ROCS supérieur à 62%.

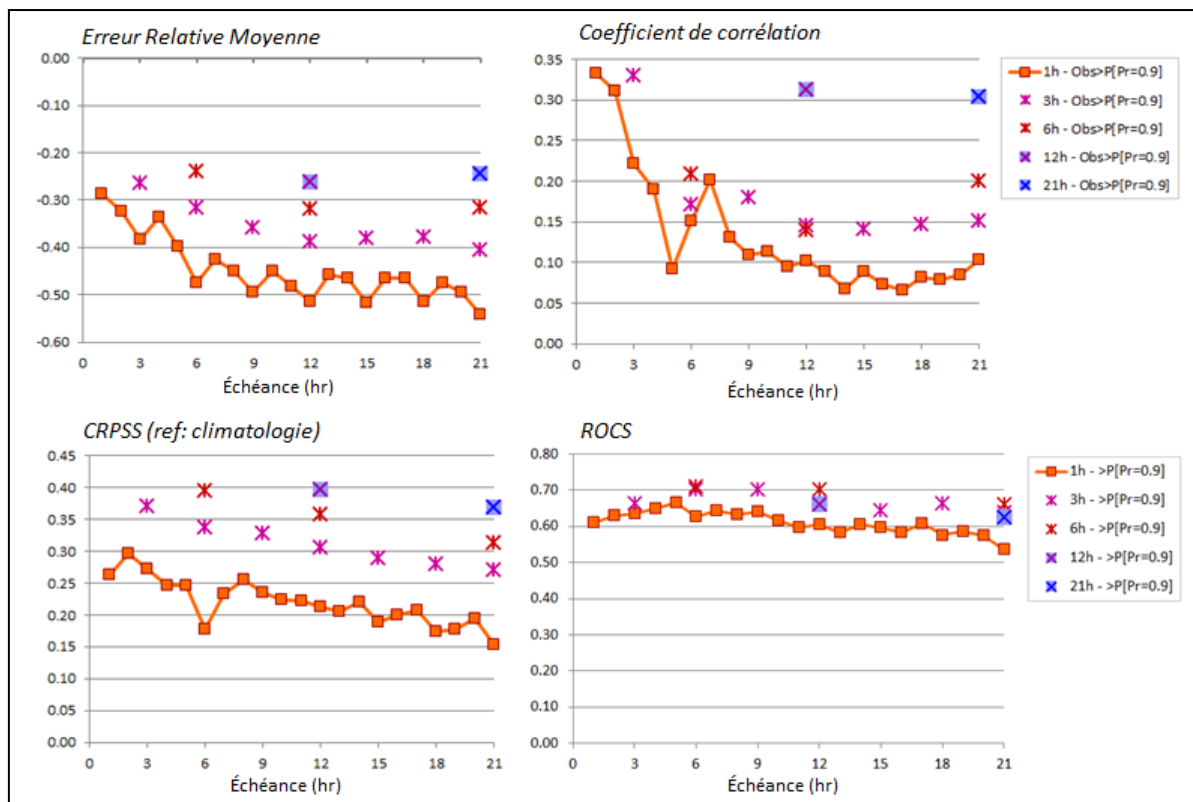


Figure 15 : Impact de l'agrégation sur les scores de vérification des prévisions d'ensemble du cumul de pluie de bassin sur 1hr, 3hr, 6hr, 12hr et 21hr pour le seuil de quantile observé 90%

Les résultats de l'évaluation des prévisions de pluie de bassin ont permis d'identifier le biais conditionnel des moyennes des ensembles de pluie, qui ont tendance à surestimer les faibles pluies et à significativement sous-estimer les fortes pluies, ainsi que le manque de dispersion des ensembles pour les premières échéances. Mais les ensembles montrent une bonne capacité à discriminer les dépassements des seuils de quantiles observés 75% et 90% et de meilleurs scores d'erreur en probabilité en comparaison à la prévision climatologique pour les observations les plus fortes.

b) Qualité des prévisions d'ensembles de débit

Comme l'alerte est basée sur le dépassement d'un seuil donné par le pic de débit, les prévisions horaires de débit sont vérifiées pour la variable de débit maximal sur les périodes de 1h, 3h, 6h, 12h et 21h (cette agrégation temporelle s'effectue dans l'outil EVS pour toutes les prévisions et observations).

Comme pour les pluies de bassin, les scores de vérification sont calculés pour chaque bassin pour toutes les paires observations-prévisions, ainsi que pour les paires dont les observations sont supérieures aux quantiles de débit observés 75% et 90% ; les scores sont ensuite moyennés sur les 27 bassins. Les échantillons sont en moyenne pour chaque bassin de 467 paires au total, 109 paires pour le quantile observé 75%, et 42 paires pour le quantile observé 90%. Il faut noter que le quantile observé 90% défini à partir des observations de 2009-2012 a une valeur inférieure au débit biennal qui est le seuil utilisé pour l'analyse de la qualité des alertes. Mais la taille réduite de l'échantillon pour ce quantile n'a pas permis d'analyser les résultats pour un quantile plus élevé.

Les prévisions de débit sont vérifiées avec les débits observés et les débits simulés. Cela permet de comparer l'impact des erreurs de sources hydrologique et météorologique (en vérifiant les prévisions avec le débit observé) avec l'impact des erreurs de source météorologique (en vérifiant les prévisions avec le débit simulé).

Erreurs des ensembles de débit en fonction du débit observé

Les Figures 16 et 17 présentent les histogrammes des erreurs des prévisions d'ensemble en fonction du débit observé et du débit simulé pour identifier les éventuels biais ou valeurs aberrantes des prévisions. Les erreurs sont calculées pour chacun des 20 membres de la prévision et sont représentées avec les valeurs des quartiles, le minimum et le maximum. Les graphes d'erreur ci-dessous incluent les résultats des 27 bassins.

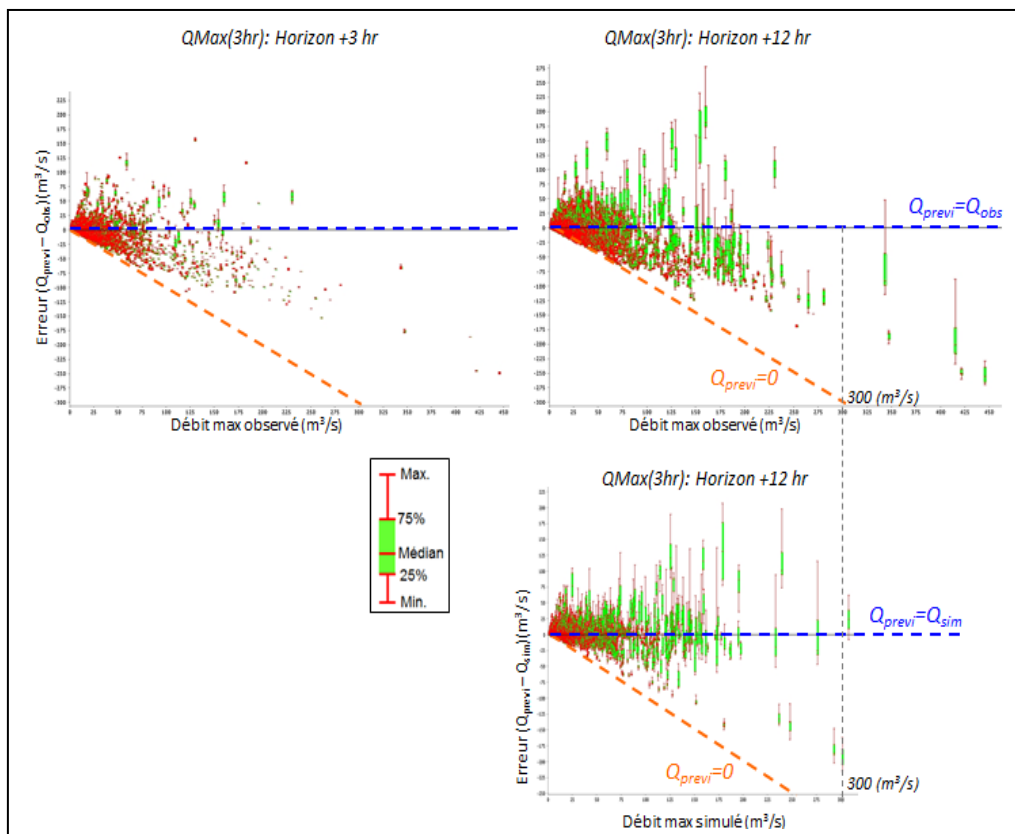


Figure 16 : Histogrammes des erreurs des ensembles de débit COSMO en fonction du débit observé et du débit simulé pour le débit maximal sur 3hr pour les horizons de 3hr et 12hr

Le premier graphe de la Figure 16, qui concerne le débit maximal tri-horaire pour la première échéance de +3hrs, montre un fort manque de dispersion des ensembles et une sous-estimation marquée des valeurs observées les plus fortes. Ceci s'explique par le manque de dispersion des prévisions d'ensemble de pluie, notamment dans les premières échéances, et par le biais conditionnel des pluies (cf. Figure 13). Pour l'échéance de +12hrs du deuxième graphe, la dispersion des ensembles est améliorée (elle permet de mieux approcher ou inclure la valeur observée dans la distribution des 20 membres) mais avec de plus fortes erreurs. Pour comparer l'impact des erreurs du modèle hydrologique et l'impact des erreurs de pluie sur le biais des prévisions de débit, les prévisions sont également vérifiées avec les valeurs de débit simulé (ce qui permet de décrire l'impact des erreurs de pluie uniquement). La comparaison des 2 graphes de droite en Figure 16 montre l'impact du modèle hydrologique lui-même, qui sous-estime les débits les plus forts (les erreurs vis-à-vis du débit simulé sont toutes

inférieures à $300 \text{ m}^3/\text{s}$). On peut observer que les ensembles de débit incluent plus fréquemment les valeurs de débit simulé et dans certains cas, surestiment les débits simulés, ce qui serait dû à la surestimation des pluies ; dans le même temps, certains ensembles montrent un fort manque de dispersion et une large sous-estimation, ces mêmes caractéristiques ayant été observées avec les ensembles de pluie.

Comme les 27 bassins versants ont des surfaces variables, les résultats des erreurs sont également présentés en Figure 17 pour le débit spécifique (défini par le débit divisé par la surface du bassin). Les graphes de la Figure 17 conduisent à la même analyse que les graphes de la Figure 16. Pour le troisième graphe de vérification avec les débits simulés, la tendance à la surestimation des débits spécifiques pour les faibles valeurs de 0.15 à $0.45 \text{ m}^3/\text{s.km}^2$ apparaît plus clairement. Ainsi les erreurs de surestimation des prévisions de pluie pourront, dans certains cas, compenser les erreurs de sous-estimation du modèle hydrologique.

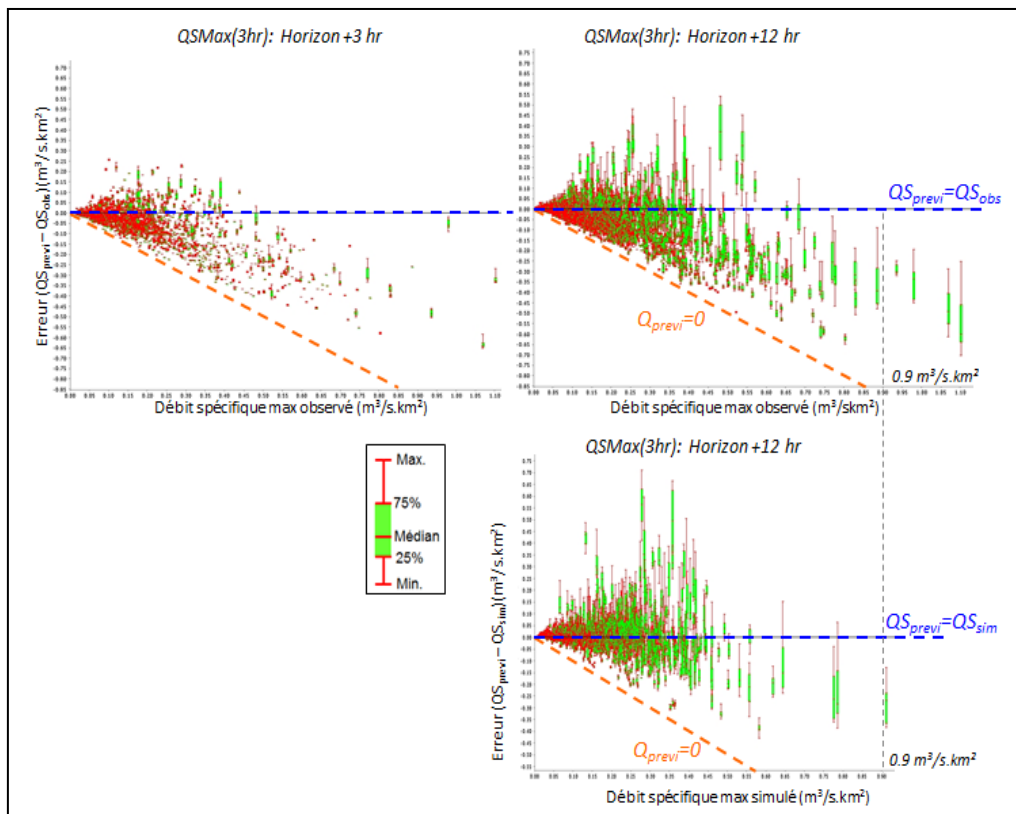


Figure 17 : Histogrammes des erreurs des ensembles COSMO de débit spécifique en fonction du débit spécifique observé et du débit spécifique simulé pour le débit maximal sur 3hr pour les horizons de 3hr et 12hr

Impact de l'échéance sur la qualité des ensembles de débit

La Figure 18 présente les scores d'évaluation en fonction de l'échéance de prévision dans les graphes ci-dessous. On calcule également les scores relatifs aux quantiles observés 75% et 90%.

L'erreur relative moyenne en Figure 18 de la moyenne des ensembles de débit horaire montre un biais négatif qui s'améliore avec l'échéance : les prévisions horaires sur de plus longues échéances sous-estiment moins les débits observés. Les biais négatifs sont plus forts pour les observations supérieures aux quantiles observés 75% et 90%, atteignant jusqu'à -30%. Les

coefficients de corrélation restent stables sur toute l'échéance, même s'ils sont plus modestes pour les observations les plus fortes. Les scores de CRPSS en référence à la prévision probabiliste climatologique de débit montrent des erreurs en probabilité plus faibles pour les prévisions d'ensemble (à l'exception des premières échéances pour le quantile observé 75%), alors que le score de CRPSS des ensembles de pluie était négatif sur toute l'échéance. Les gains en CRPS s'améliorent avec l'échéance, atteignant jusqu'à 20% pour le quantile observé 90%. Le score de discrimination ROCS pour les seuils des quantiles observés 75% et 90% montre également une qualité de discrimination qui s'améliore avec l'échéance.

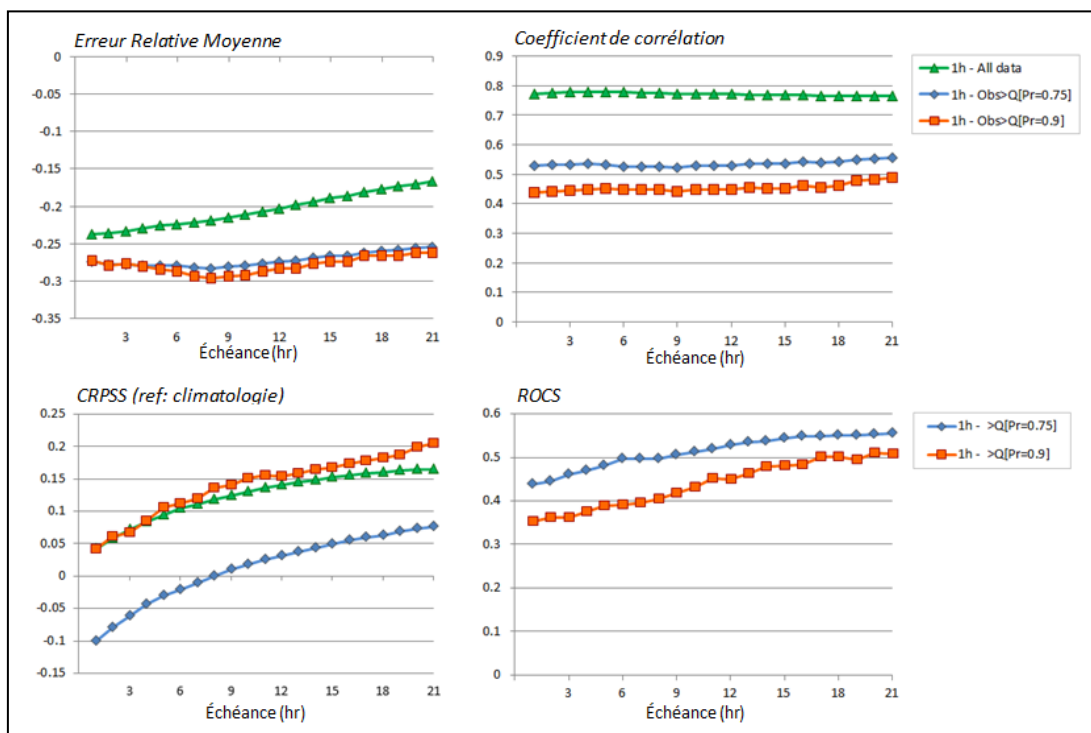


Figure 18 : Impact de l'échéance sur les scores de vérification des prévisions d'ensemble de débit pour tout l'échantillon et pour les seuils des quantiles observés 75% et 90%

Impact de la période d'agrégation sur la qualité des ensembles de débit

En Figure 19, les scores d'évaluation sont comparés pour les différentes périodes de calcul du pic de débit de 3 à 21 heures. Les scores concernent les observations supérieures aux quantiles de débit observés 75% et 90%. Il faut noter que les alertes basées sur les ensembles de débit considèrent toute l'échéance de prévision et donc elles sont basées sur la prévision de débit maximal sur 21 heures.

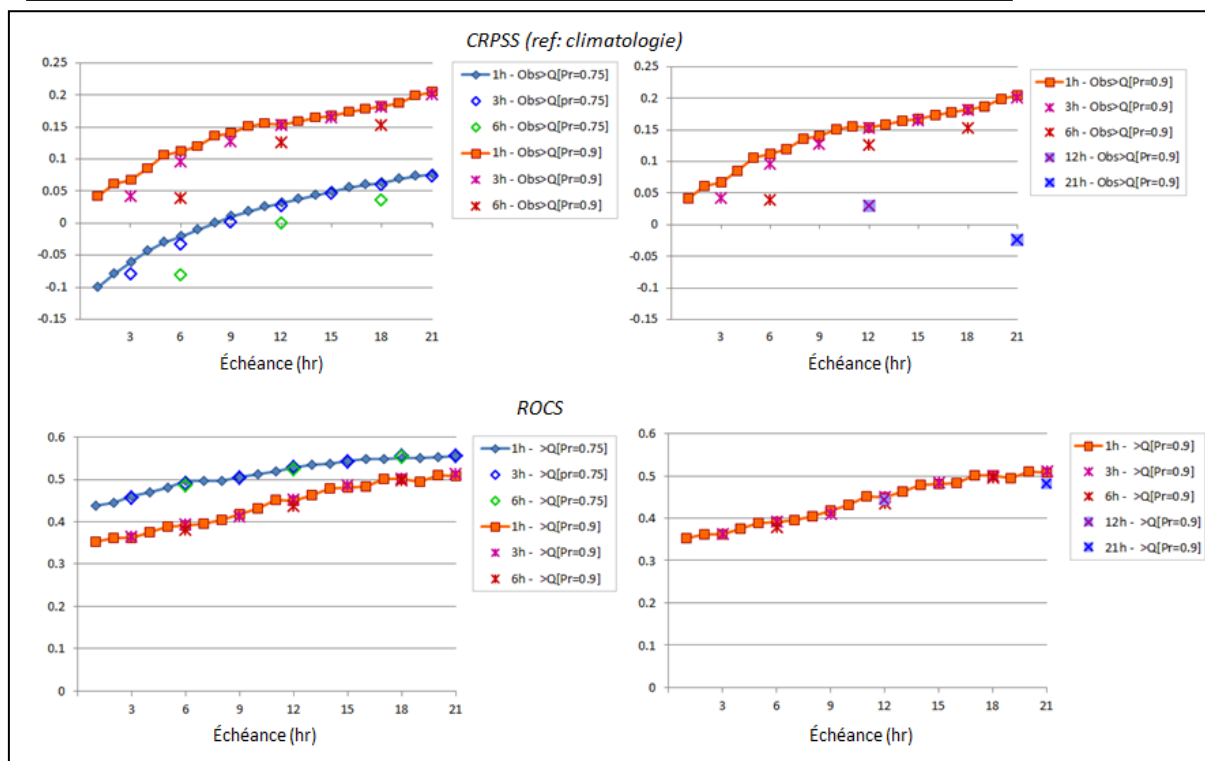
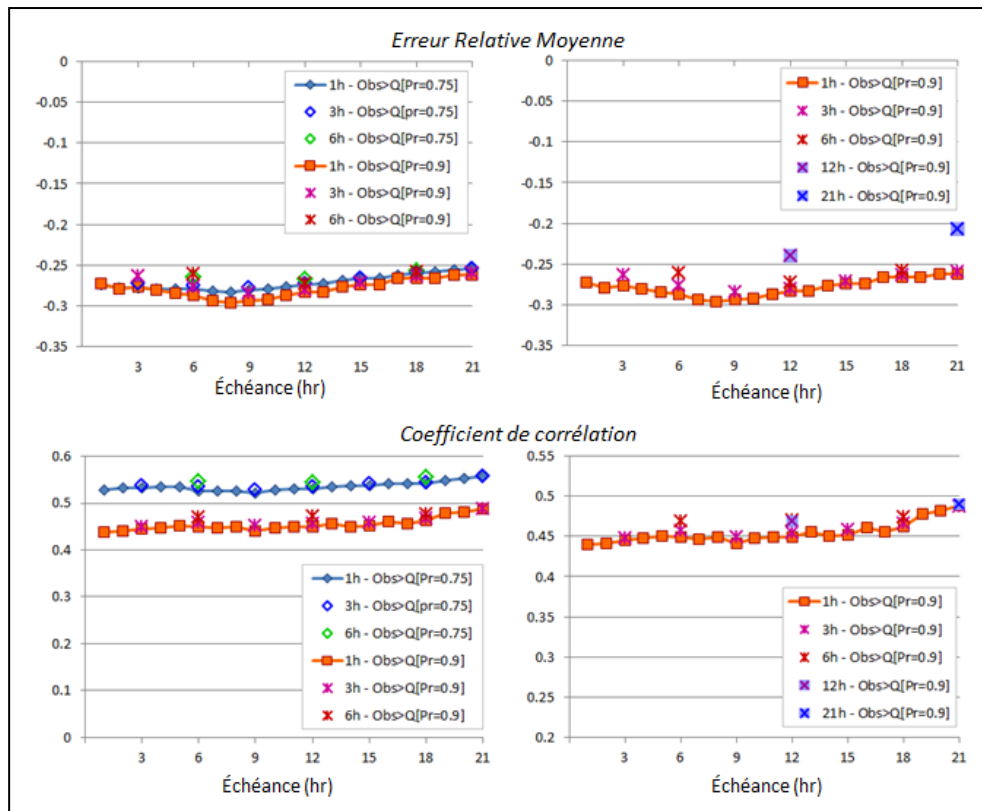


Figure 19 : Impact de l'agrégation sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour les seuils des quantiles observés 75% et 90%

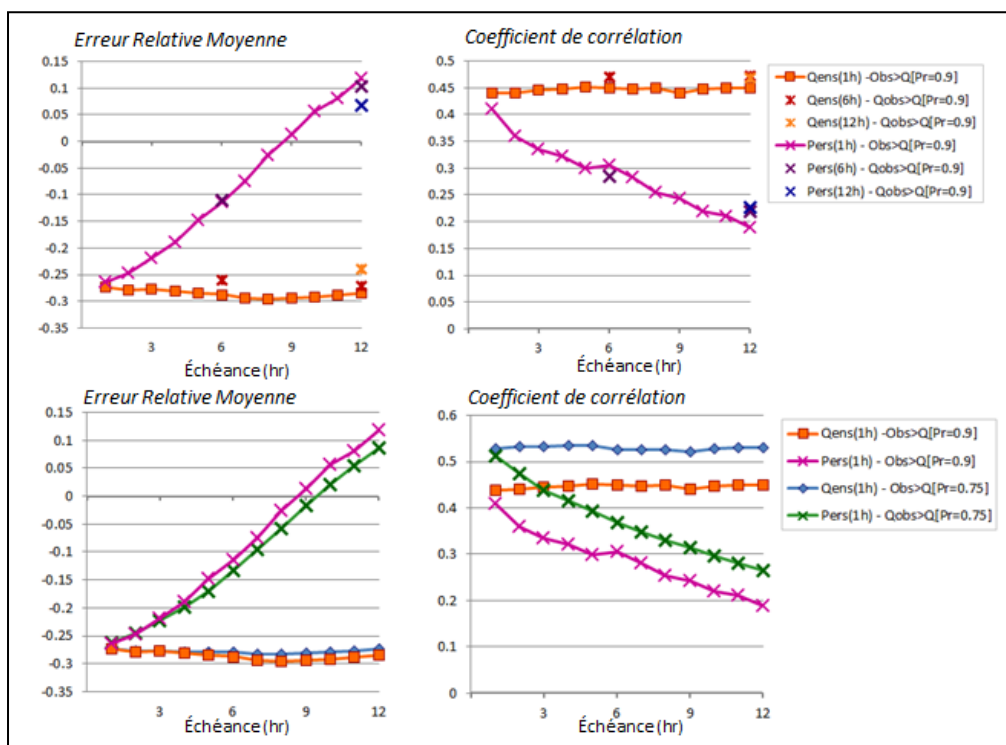
Comme les variables de débit maximal sur des échéances de plus en plus longues sont corrélées, les scores d'évaluation sont relativement similaires pour les variables agrégées sur ces différentes échéances, le plus souvent en amélioration avec l'échéance. Pour le quantile observé 90%, l'erreur relative moyenne sur 21 heures est de -21% alors que, pour le débit

horaire, cette erreur varie entre -26% et -30%. L'agrégation n'a quasiment aucun impact sur les coefficients de corrélation et sur le score de discrimination ROCS. Par contre, le gain en CRPS des ensembles par rapport à la prévision probabiliste climatologique se détériore avec l'agrégation sur des échéances plus longues (car l'agrégation permet aussi d'améliorer la qualité de la prévision probabiliste climatologique). Notamment pour le quantile observé 90%, les erreurs en probabilité de la prévision du pic de débit sur 21 heures sont similaires à celles de la prévision probabiliste climatologique (le CRPS étant de -2%).

Comparaison de la qualité des ensembles de débit avec la prévision persistante

La Figure 20 permet de comparer les scores des ensembles de débit COSMO avec la prévision issue de la pluie persistante (définie par dernière valeur de pluie observée) pour l'échéance maximale de 12 heures, en fonction de l'échéance et de la période d'agrégation. Les scores sont relatifs aux observations supérieures aux quantiles observés 75% et 90%, pour lesquelles la prévision persistante présente a priori le plus d'intérêt pour les alertes.

Le biais de la prévision persistante est négatif pour la première échéance (comme le biais de la moyenne des ensembles COSMO) mais ce biais s'améliore avec l'échéance pour devenir positif après l'échéance de 9 heures. Par contre, le coefficient de corrélation est fortement dégradé avec la prévision persistante par rapport à la moyenne des ensembles COSMO, avec des valeurs divisées par 2 à l'échéance de 12 heures. Les gains en CRPS des ensembles par rapport à la prévision persistante augmentent avec l'échéance, pour atteindre une valeur de plus de 40% à l'échéance de 12 heures. La capacité de discrimination pour les 2 valeurs de seuil est également plus forte avec les ensembles COSMO.



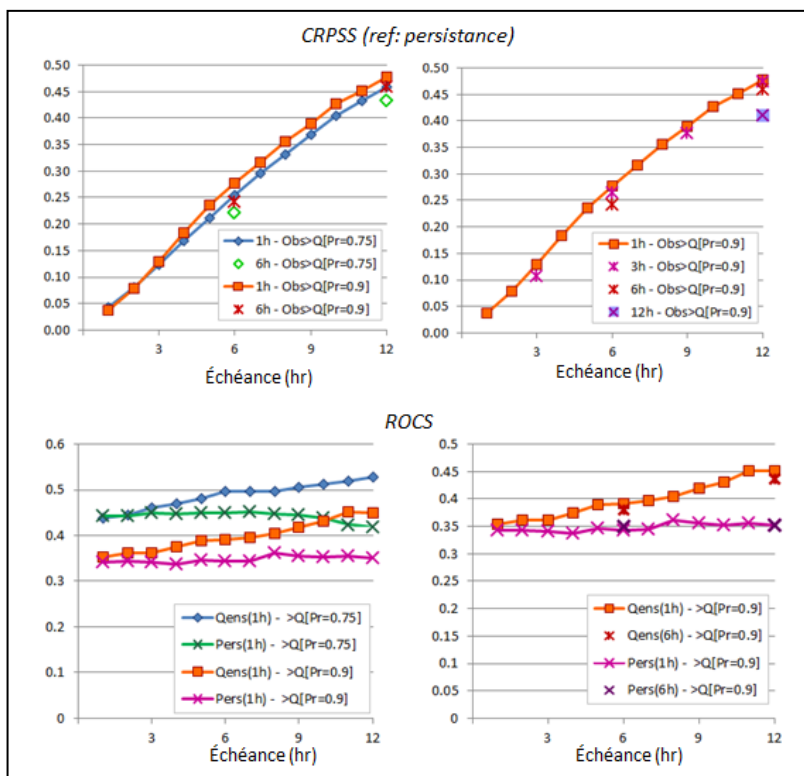


Figure 20 : Comparaison des scores de vérification des prévisions d'ensemble COSMO de débit maximal avec la prévision de débit issu de la pluie persistante sur 12 hrs pour les seuils des quantiles observés 75% et 90%

Comparaison de l'impact des erreurs de sources hydrologique et météorologique

En Figure 21, les scores sont présentés en vérifiant les prévisions COSMO avec le débit observé, ce qui traduit l'impact des erreurs de sources hydrologique et météorologique sur les prévisions de débit, et avec le débit simulé, pour quantifier l'impact des erreurs de source météorologique. Les résultats sont relatifs aux prévisions de débit horaires et aux prévisions de pic de débit sur 21 heures ou 12 heures (pour la comparaison avec la prévision persistante de débit). Il faut noter que, pour la première échéance, les prévisions de débit sont très similaires au débit simulé, alors que les différences entre les prévisions de débit et le débit observe peuvent être importantes, dues aux erreurs d'initialisation et de simulation du modèle hydrologique.

Alors que la moyenne des ensembles COSMO sous-estime significativement le débit observé, elle a tendance à surestimer le débit simulé, ce biais positif se dégradant avec l'échéance. Cependant, pour le pic de débit sur 21 heures, le biais par rapport au débit simulé est seulement de 3% alors que le biais par rapport au débit observé est de -20%. On note donc que les 2 types de biais vont en partie se compenser mais que l'impact des erreurs de source hydrologique est en moyenne plus important. Le coefficient de corrélation pour la moyenne des ensembles COSMO montre un fort impact des erreurs hydrologiques pour les premières échéances (dégradation de plus de 20% pour la première échéance), cette dégradation diminuant avec l'échéance. Le coefficient de corrélation est dégradé d'environ 8% pour le pic de débit sur 21 heures. Pour les erreurs en probabilité des ensembles, lorsque les ensembles COSMO sont comparés à la prévision probabiliste climatologique, les erreurs hydrologiques dégradent fortement les gains, même si cette dégradation s'atténue avec l'échéance ; pour le

pic de débit sur 21 heures, les erreurs hydrologiques réduisent le gain de plus de 40% avec une valeur de CRPSS de 9%. Avec la prévision persistante de débit, la dégradation des gains en CRPSS s'accroît pour les 3 premières échéances pour se stabiliser ensuite (pour la première échéance, la prévision persistante est relativement similaire au débit simulé, tout comme la prévision d'ensemble COSMO). Pour le pic de débit sur 12 heures, les erreurs hydrologiques réduisent le gain de plus de 30% pour une valeur de CRPSS de 39%.

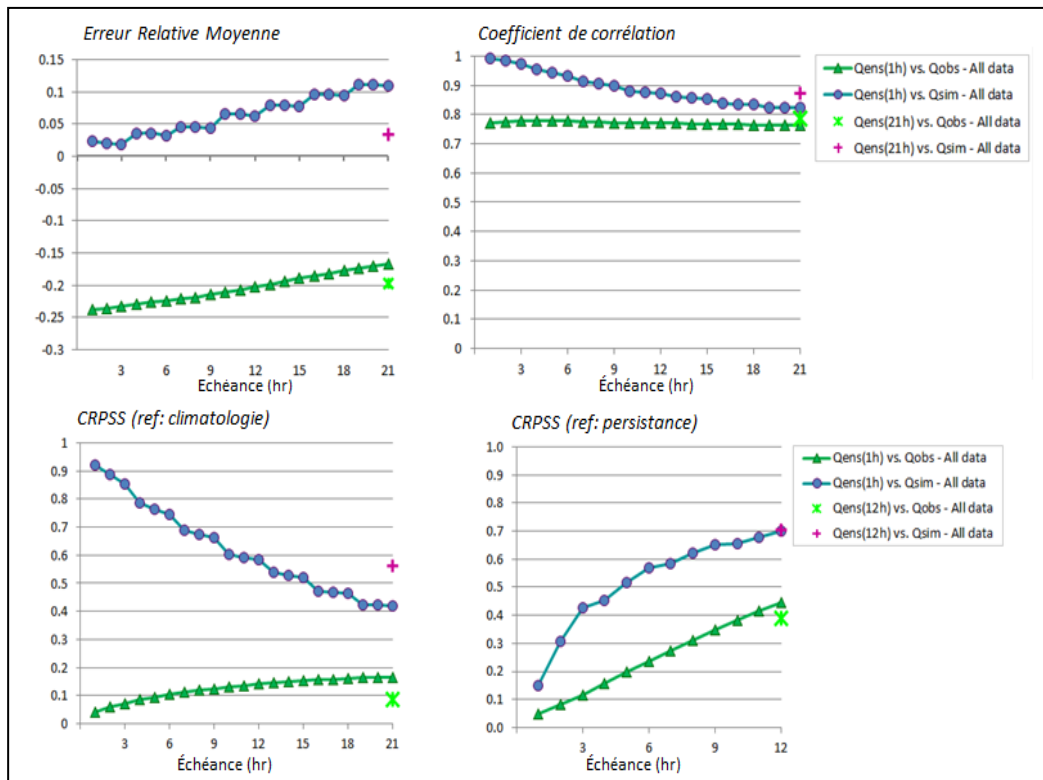


Figure 21 : Comparaison de l'impact des erreurs de sources hydrologique et météorologique (en vérifiant avec le débit observé) et de l'impact des erreurs de source météorologique (en vérifiant avec le débit simulé) sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour tout l'échantillon

Cette analyse montre l'intérêt d'améliorer par le futur les performances du modèle hydrologique puisque l'impact des erreurs hydrologiques semble plus fort que l'impact des erreurs météorologiques sur les prévisions d'ensemble de débit COSMO.

Impact de la surface de bassin sur la qualité des ensembles de débit

Les résultats en Figure 22 permettent d'analyser l'impact de la surface de bassin sur les scores de vérification des ensembles COSMO. Même si les scores d'évaluation choisis sont des scores relatifs, ils sont calculés ici avec les débits spécifiques pour les 2 catégories suivantes de surface de bassin:

- surface comprise entre 8 km² et 232 km²,
- surface comprise entre de 288 km² à 944 km².

Les scores d'évaluation sont présentés pour les observations supérieures aux quantiles observés 75% et 90% pour se concentrer sur les plus forts débits. Cependant ce double sous-échantillonnage (par la surface de bassin et par la valeur du débit observé) augmente l'incertitude des valeurs des scores d'évaluation, qui sont obtenus sur des échantillons plus

réduits. Les résultats sont donc indicatifs de l'impact potentiel de la surface de bassin sur la qualité des prévisions d'ensemble de débit.

Pour les plus grands bassins, les erreurs relatives moyennes pour les observations supérieures aux quantiles observés 75% et 90% sont plus fortes et se détériorent avec l'échéance (alors que ces erreurs s'améliorent avec l'échéance pour les plus petits bassins). Les coefficients de corrélation sont relativement similaires pour le quantile observé 90% mais pour le quantile observé 75%, la corrélation est moins forte pour les plus grands bassins. Pour le score de CRPSS, les gains en termes d'erreur en probabilité par rapport à la prévision de débit climatologique se dégradent pour les plus grands bassins (dégradation progressive pour le quantile 90%). Il en est de même pour la discrimination des dépassements des quantiles observés 75% et 90%, qui est plus faible pour les grands bassins. Ces résultats semblent indiquer que les prévisions d'ensemble de débit COSMO pour les plus fortes observations sont de meilleure qualité pour les petits bassins, de surface jusqu'à 232 km².

Figure 22 : Impact de la surface de bassin sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour les seuils des quantiles observés 75% et 90%

En conclusion, l'évaluation des ensembles de débit a montré une bonne corrélation des moyennes des ensembles avec les observations mais une sous-estimation significative pour les débits les plus forts, à cause du biais négatif du modèle hydrologique et, dans une moindre mesure, du biais des ensembles de pluie COSMO-DE-EPS. Cependant, certaines surestimations des pluies moyennes peuvent compenser les erreurs de sous-estimation du modèle hydrologique. Les ensembles de pluie et de débit montrent également un manque de dispersion pour les courtes échéances. Mais ces ensembles ont une bonne capacité pour discriminer les dépassements des quantiles observés 75% et 90%, ce qui semble prometteur pour la qualité des alertes. En comparaison de la prévision probabiliste climatologique et de la prévision persistante, les ensembles ont de meilleures performances, notamment en termes

d'erreur en probabilité. A cause du biais négatif du modèle hydrologique et du manque de dispersion pour les échéances courtes, les gains en performance des prévisions d'ensemble de débit COSMO vis-à-vis des prévisions de référence s'améliorent avec l'échéance. Il semble également que ces gains en performance soient légèrement plus importants pour les plus petits bassins de surface jusqu'à 232 km².

c) Qualité des alertes probabilisées

Cette section présente les résultats de l'analyse de la qualité des alertes pour le dépassement du seuil de débit biennal en se basant sur les critères de contingence et l'anticipation des alertes.

Cartes de vérification des alertes

L'analyse est réalisée pour chacun des 7 événements et sur l'ensemble des 29 bassins versants. Pour un événement donné, on peut représenter sur une carte la qualité de l'alerte pour chaque bassin versant en se basant sur les 4 catégories du tableau de contingence : succès (hit), manqué (miss), fausse alerte (false alarm), rejet négatif (true negative).

La Figure 23 donne des exemples de carte de vérification des alertes pour les 2 événements de décembre 2010 et décembre 2011 pour les prévisions suivantes : 2 prévisions probabilistes, notées $Q[\text{Prob}=p]$ avec p égal à 0.05 et 0.71, la prévision de débit simulé sans pluie future (Q_{simu}), correspondant au système AIGA actuel, ainsi que la prévision issue de la pluie persistante maintenue sur 3 heures (notée $Q[\text{pers}(3\text{hr})]$). Pour les alertes probabilisées, en élevant le niveau de probabilité, le système produit davantage d'alertes manquées et un nombre plus réduit de fausses alertes. Les exemples de la Figure 23 montrent une légère amélioration avec les alertes probabilisées en comparaison du système actuel pour quelques bassins.

Pour une prévision donnée, cette carte de vérification des alertes ne peut être réalisée que lorsque le débit peut être observé (ou que l'éventuel dépassement du seuil de débit peut être estimé, par exemple à partir de dégâts sur le terrain). Par conséquent, cette carte de vérification peut être générée pour caractériser la qualité des dépassements pour les bassins jaugés et ainsi permettre d'analyser la qualité potentielle des alertes sur les bassins non jaugés voisins. Elle peut également permettre de comparer la qualité pour différentes prévisions, notamment pour déterminer quel(s) serait(ent) le(s) meilleur(s) niveau(x) de probabilité à considérer pour les alertes probabilisées.

Pour caractériser spatialement la qualité des alertes de façon plus robuste, la carte de vérification peut être réalisée avec les statistiques de contingence agrégées par bassin sur différents événements passés. Par exemple, une carte du biais en fréquence des alertes peut permettre aux prévisionnistes d'analyser spatialement la qualité des alertes et d'identifier potentiellement des zones pour lesquelles le système a tendance à bien détecter les dépassements, ou bien à produire une majorité de fausses alertes ou d'alertes manquées.

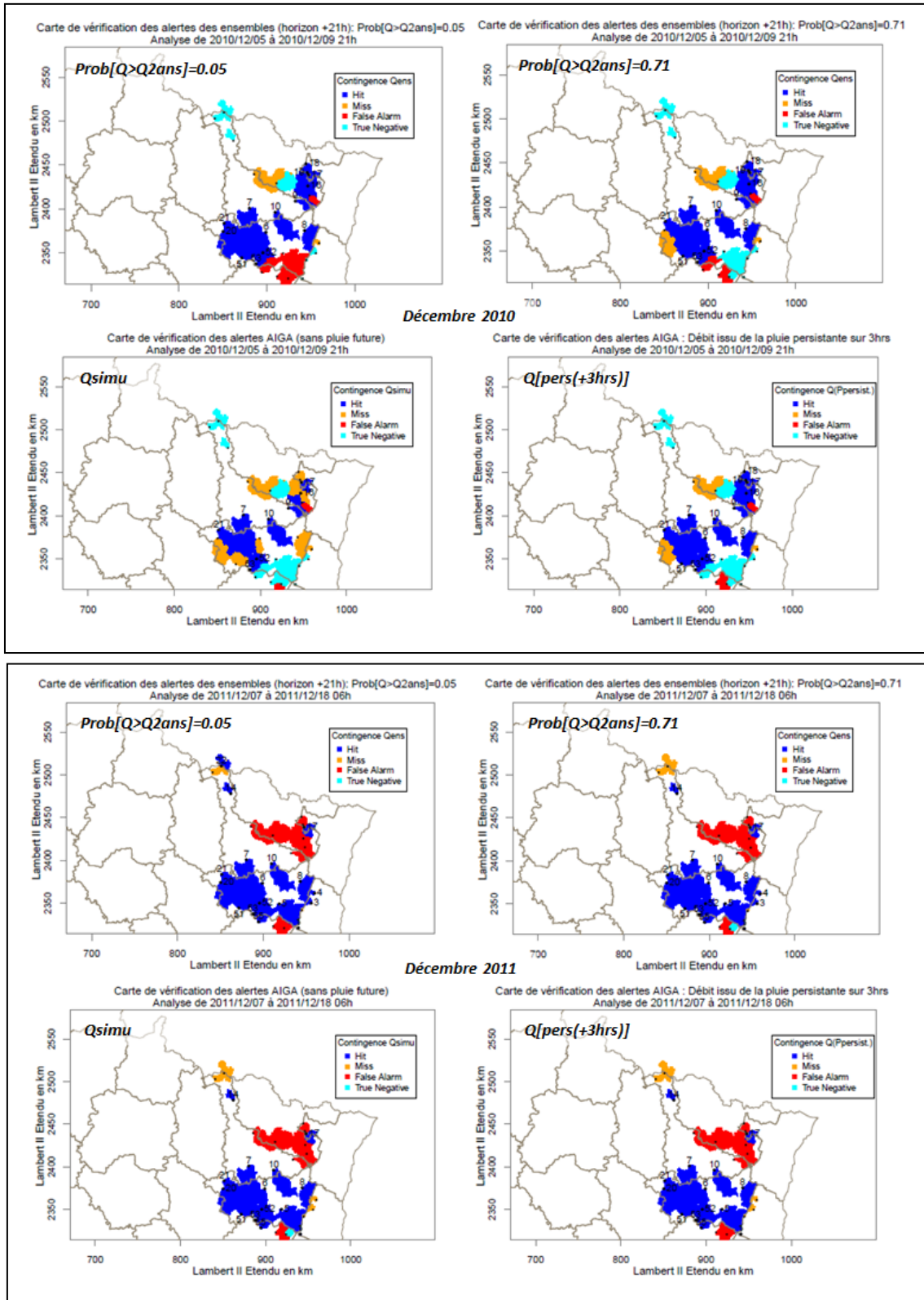


Figure 23 : Exemples de cartes de vérification des alertes pour les événements de décembre 2010 et décembre 2011 ; les identifiants des bassins sont données lorsque l’alerte est correcte

Critères de contingence pour les alertes

Pour analyser la qualité des alertes avec les différentes prévisions, la Figure 24 montre les scores de contingence (en prenant en compte l'anticipation) obtenus pour les 7 événements étudiés avec les prévisions suivantes :

- les prévisions probabilistes, notées $Q[\text{Prob}=p]$ avec p égal à 0.05, 0.48, 0.71 et 0.9, issues des 20 ensembles de débit sur le graphe de gauche;
- les prévisions probabilistes obtenues à partir des 100 ensembles (basés sur les 5 jeux de paramètres différents) pour 2 niveaux de probabilité, p égal à 0.05 et 0.24, pour le graphe de droite,
- la prévision de débit simulé sans pluie future (Q_{simu}), correspondant au système AIGA actuel,
- les 2 prévisions de référence, la prévision persistante (notée $Q[\text{pers}(3\text{hr})]$ pour une échéance maximale de 3 heures) et la prévision avec la pluie observée (Q_{pobs}) sur le graphe de gauche.

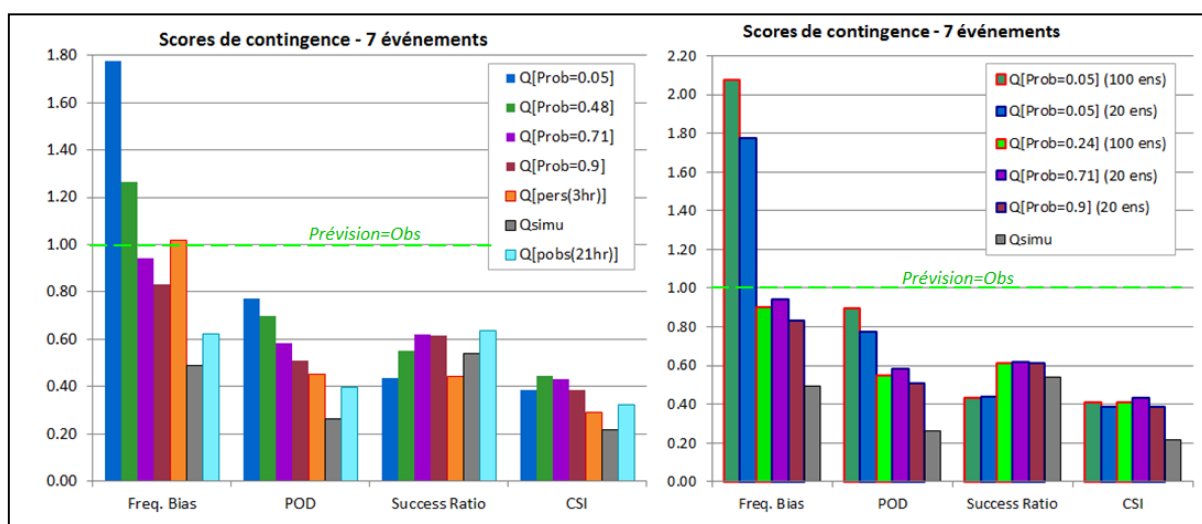


Figure 24 : Scores de contingence des alertes de dépassement du débit biennal avec les prévisions probabilistes de débit ($Q[\text{Prob}=p]$ avec p égal à 0.05, 0.48, 0.71 et 0.9) issus de 20 ensembles et 100 ensembles, la prévision de débit issu de la pluie persistante ($Q[\text{pers}(3\text{hr})]$), la prévision de débit sans pluie future (Q_{simu}) et la prévision de débit issue de la pluie observée ($Q[\text{pobs}(21\text{hr})]$)

Lorsque l'on considère uniquement les prévisions probabilistes issues des 20 ensembles de débit (graphe de gauche), les résultats montrent que les prévisions probabilistes sont très généralement supérieures au système actuel sans pluie future. Par exemple, lorsque l'on compare les scores du système AIGA actuel avec ceux des prévisions pour le niveau de probabilité de 71%, le gain en POD est de 32%, celui en fiabilité (score Success Ratio) de 8%, et le gain en CSI est de 22%. La prévision probabiliste de 5% a tendance à générer trop de fausses alertes en comparaison des alertes manquées (biais en fréquence proche de 1.8), ce qui permet d'améliorer la probabilité de détection (POD proche de 80%) mais ce qui dégrade le score de fiabilité (Success Ratio de 40% inférieur à celui du système AIGA actuel). Les prévisions pour les niveaux de probabilité de 71% et 90% ont une légère tendance à sous-estimer les dépassements mais obtiennent, pour les 3 autres scores, des résultats meilleurs ou similaires aux 2 prévisions de référence issues de la pluie persistante et de la pluie observée. En s'aidant de ces statistiques, le prévisionniste pourra estimer quel(s) est(sont) le seuil de probabilité le(s) plus adapté(s) pour produire un nombre acceptable de fausses alertes et d'alertes manquées, tout en assurant un système d'alerte suffisamment fiable et précis.

On peut noter que la prévision issue de la pluie observée améliore les résultats vis-à-vis du système actuel sans pluie future, mais du fait des erreurs de simulation (le modèle hydrologique ayant tendance à sous-estimer les débits), le gain en qualité reste limité. En comparaison, la prise en compte des prévisions d'ensemble de pluie permet de compenser en partie ces erreurs du modèle et d'améliorer ainsi les scores de POD et CSI.

Le graphe de droite de la Figure 24 permet de comparer les scores obtenus avec les prévisions probabilistes intégrant uniquement les incertitudes de pluie (prévisions issues des 20 ensembles) avec ceux des prévisions intégrant les incertitudes de pluie et des paramètres hydrologiques (prévisions issues de 100 ensembles). En considérant les deux sources d'incertitude, la prévision de probabilité 5% issue des 100 ensembles conduit à des résultats plus biaisés, favorisant fortement les fausses alertes, ce qui dégrade le score de fiabilité. La prévision de probabilité 24% issue des 100 ensembles produit des scores similaires à la prévision de probabilité de 71% issue des 20 ensembles. Ces résultats ne permettent pas de montrer un apport significatif des autres jeux de paramètres pour améliorer la qualité des alertes. Ceci s'explique notamment par le fait que le jeu optimisé de paramètres a eu pour objectif d'augmenter la réactivité du modèle hydrologique et de mieux différencier les dépassements des quantiles de débit de période de retour de 2, 10 et 50 ans. Les ensembles issus des 4 autres jeux de paramètres n'ont pas été optimisés et, pour l'échantillon relativement réduit de bassins et d'événements de cette analyse, n'améliorent pas la précision de détection et de fiabilité des alertes probabilisées.

Cette première évaluation serait à compléter par l'analyse d'autres événements sur un plus grand nombre de bassins pour obtenir des statistiques plus robustes. Elle est complétée ci-dessous par l'analyse de l'amélioration potentielle du degré d'anticipation des dépassements observés. Par la suite, les prévisions probabilistes sont uniquement issues des 20 ensembles de débit prenant en compte l'incertitude de pluie future.

Anticipation des alertes

Les résultats d'anticipation sont tout d'abord présentés par bassin pour les 2 événements les plus forts. Ils sont basés sur les dépassements observés du débit biennal ; ces dépassements sont prévus pour 13 bassins en décembre 2010 et 16 bassins en décembre 2011 (se reporter à l'Appendice A pour la liste des bassins avec leurs identifiants)

La Figure 25 permet de comparer l'anticipation des alertes pour les prévisions probabilistes de 5%, 48% et 90%, avec l'anticipation de la prévision sans pluie future (Q_{simu}), correspondant au système AIGA actuel, et la prévision avec la pluie persistante ($Q_{pers(3hr)}$). Les résultats sont représentés par valeurs décroissantes de l'anticipation avec la prévision probabiliste de 5%. Pour les prévisions probabilistes de 5% et 48%, l'anticipation est positive pour tous les bassins des 2 événements. En augmentant le seuil de probabilité pour produire une alerte, l'anticipation se dégrade ; ceci conduit notamment à des retards pour 5 des 13 bassins avec la prévision probabiliste de 90% pour l'événement de décembre 2010. Comme remarqué lors de l'analyse des résultats de contingence, on note une diminution du nombre d'alertes manquées avec les prévisions probabilistes, notamment en comparaison avec le système AIGA actuel (Q_{simu}).

Il faut noter que les très fortes valeurs d'anticipation du débit simulé sans pluie future correspondent à une large surestimation du débit observé par le débit simulé. Cette

surestimation peut être liée à un décalage temporel de l'hydrogramme simulé puisque les procédures de calibration cherchent un modèle hydrologique qui soit le plus réactif. De plus, cette surestimation du débit peut être renforcée par une surestimation de la prévision de pluie et conduire ainsi à une augmentation des fausses alertes. C'est pourquoi il est important de contrebalancer l'amélioration de l'anticipation par les scores de fiabilité et de CSI, scores qui incluent le nombre de fausses alertes.

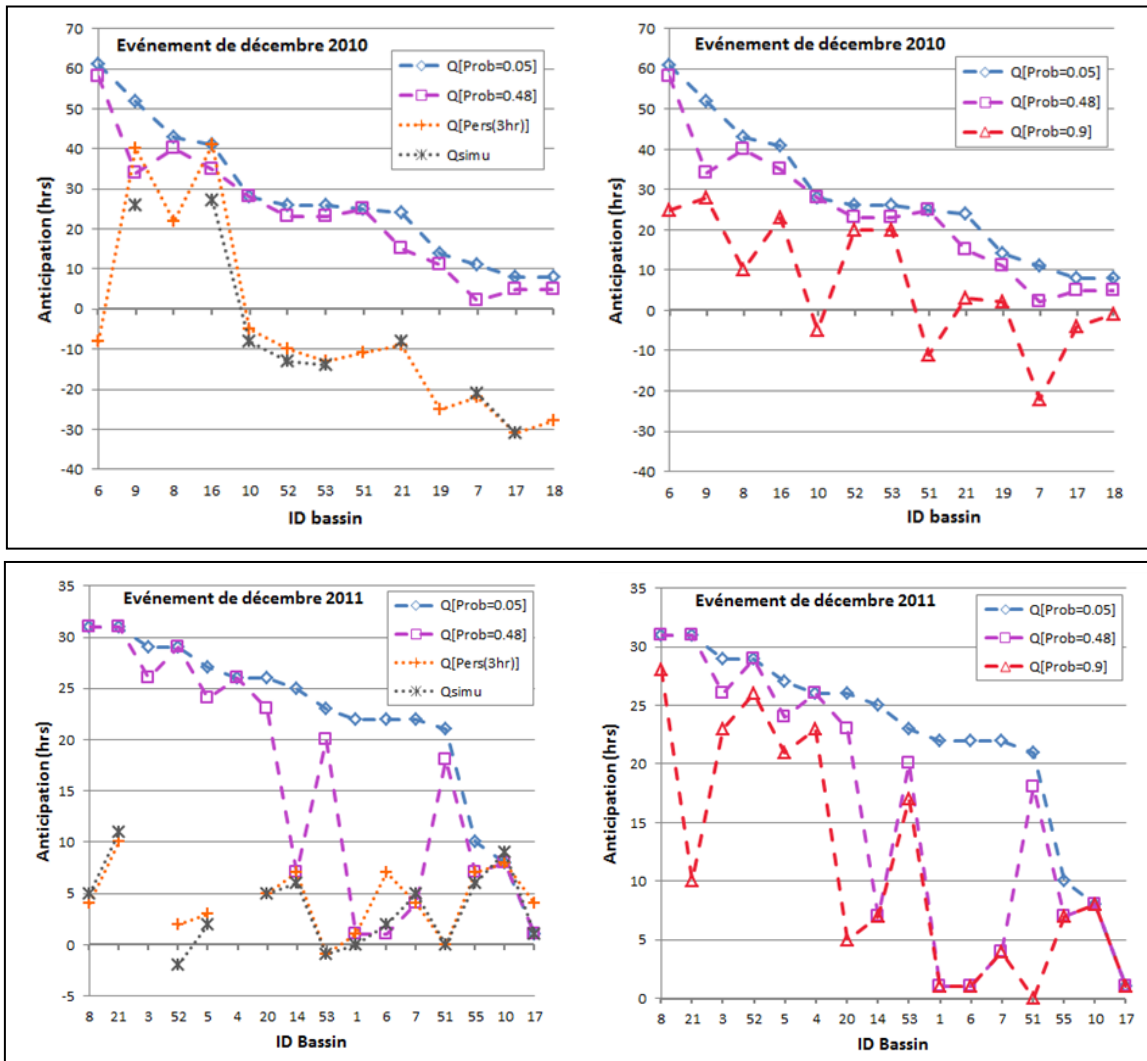


Figure 25 : Anticipation des alertes de dépassement du débit biennal avec les prévisions probabilistes de débit ($Q[Prob=p]$ avec p égal à 0.05, 0.48 et 0.9), la prévision de débit issu de la pluie persistante ($Q[pers(3hr)]$) et la prévision de débit sans pluie future ($Qsimu$) pour les événement de décembre 2010 et décembre 2011

Lorsque l'on considère uniquement les alertes correctes du système AIGA actuel (i.e. en excluant les alertes manquées, ce qui correspond à limiter l'analyse aux 8 des 13 bassins pour décembre 2010 et aux 14 des 16 bassins pour décembre 2011), on peut calculer la moyenne du gain en anticipation pour la prévision probabiliste et la prévision persistante.

Gain moyen vis-à-vis d'AIGA actuel	$Q[Prob=0.48]$	$Q[pers(3hr)]$
Décembre 2010 (8 bassins)	11,1 hr	0,8 hr
Décembre 2011 (14 bassins)	26,3 hr	4,8 hr

Ces résultats pour les 2 événements les plus forts montrent les gains significatifs en anticipation lorsque les prévisions de pluie COSMO-DE-EPS sont prises en compte. La comparaison entre la prévision issue de la pluie persistante et le système AIGA actuel montre une amélioration relativement modeste. Mais il faut noter que cette prévision issue de la pluie persistante est estimée toutes les 3 heures uniquement (lorsque les prévisions de pluie COSMO-DE-EPS sont disponibles pour générer les prévisions d'ensemble de débit), alors que le système AIGA actuel (avec le débit simulé) génère des alertes toutes les heures. C'est pourquoi pour quelques bassins (tels que les bassins 8 et 15 pour l'événement de décembre 2011), l'anticipation avec le système AIGA actuel est meilleure d'1 heure vis-à-vis de la prévision persistante de débit.

Comparaison du degré d'anticipation vis-à-vis des prévisions déterministes

L'analyse du degré d'anticipation des prévisions probabilistes est réalisée à partir de l'échantillon de 38 dépassements observés pour les 4 événements les plus forts de décembre 2010, janvier 2011, décembre 2011 et janvier 2012. Pour calculer le gain en anticipation de la prévision probabiliste vis-à-vis d'une prévision de référence, seuls les dépassements prévus par les 2 types de prévisions sont inclus. Du fait des alertes manquées pour certaines prévisions de référence, les gains en anticipation sont estimés à partir des échantillons suivants :

- pour le gain vis-à-vis du système AIGA actuel, 24 dépassements prévus,
- pour le gain vis-à-vis de la prévision persistante, 36 dépassements prévus,
- pour le gain vis-à-vis de la prévision de pluie observée, 25 dépassements prévus.

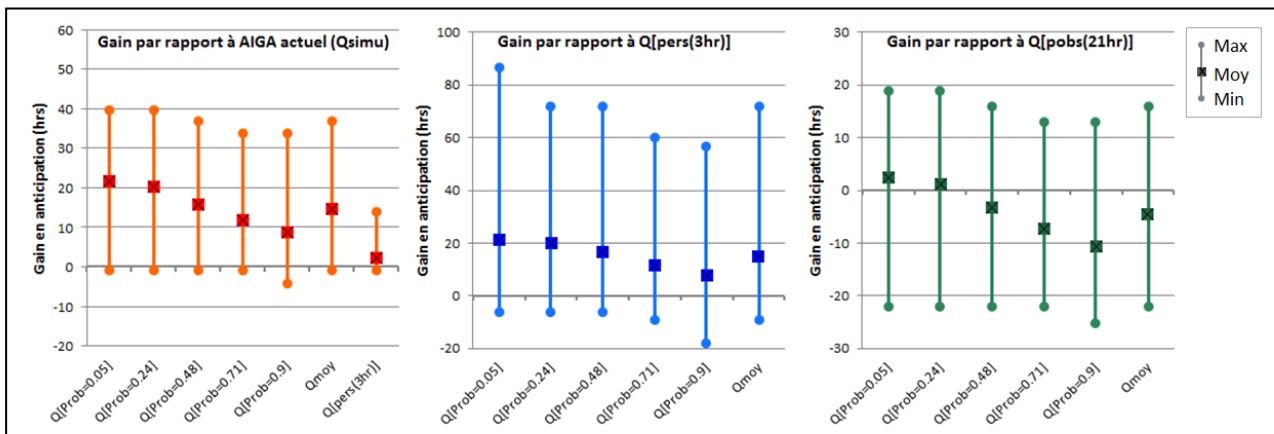


Figure 26 : Statistiques des gains en anticipation des alertes de dépassement du débit biennal avec les différentes prévisions probabilistes (Q[Prob=p]) et la prévision moyenne des ensembles de débit (Qmoy) par rapport à la prévision sans pluie future (Qsimu), la prévision persistante de débit (Q[pers(3hr)]) et la prévision de débit issu de la pluie observée (Q[pobs(21hr)])

La Figure 26 représente les statistiques de minimum, maximum et moyenne des gains en anticipation vis-à-vis du système AIGA actuel, de la prévision de pluie persistante et de la prévision issue de la pluie observée. Pour la comparaison avec le système AIGA actuel, la moyenne du gain en anticipation avec toutes les prévisions probabilistes varie de 21 heures (pour le niveau de probabilité de 5%) à 8.7 heures (pour le niveau de probabilité de 90%) sur les 24 dépassements prévus. On peut noter que l'amélioration en anticipation pour la moyenne des ensembles de débit est également significative, avec un gain moyen de 14.5 heures sur les 24 dépassements. Malgré l'échantillon réduit utilisé pour cette étude, ceci montre clairement l'intérêt d'intégrer les prévisions de pluie COSMO-DE-EPS.

Les statistiques du degré d'anticipation vis-à-vis de la prévision persistante montrent des gains moyens positifs pour toutes les prévisions issues des ensembles COSMO-DE-EPS. Ceci démontre l'intérêt d'intégrer les prévisions de pluie COSMO-DE-EPS qui décrivent la variabilité de la pluie future, d'autant plus que ces prévisions COSMO sont utilisées pour leur échéance maximale de 21 heures alors que les valeurs de pluie persistante ne sont maintenues que sur les 3 premières heures d'échéance (pour éviter de générer trop de fausses alertes). Les valeurs négatives du gain potentiel minimum concernent 3 des 36 dépassements. A noter que les très fortes valeurs maximales du degré d'anticipation s'expliquent par les surestimations des prévisions d'ensemble de débit conjuguées aux retards pour les alertes de la prévision persistante.

Pour la comparaison avec la prévision issue de la pluie observée, il permet d'évaluer l'impact des erreurs des prévisions de pluie sur le degré d'anticipation. Pour les prévisions probabilistes de 5% et 24%, le gain moyen en anticipation est très légèrement positif (de 2.5 hrs et 1hr respectivement), alors qu'il est négatif pour les autres niveaux de probabilité. Ceci montre une compensation significative des erreurs de la prévision d'ensemble de pluie avec les erreurs du modèle hydrologique, du moins pour un certain nombre de membres de la prévision d'ensemble, ce qui permet d'améliorer l'anticipation pour les prévisions probabilistes de 5% et 24%. Mais, pour les prévisions probabilistes de 48% à 90%, les erreurs des prévisions de pluie entraînent une dégradation du degré d'anticipation moyen.

Impact de la surface de bassin sur le degré d'anticipation

Cette analyse se base sur la comparaison du degré d'anticipation en fonction de la surface de bassin. Pour les 38 dépassements observés des 4 événements les plus forts, on différencie les résultats pour 2 catégories de surface de bassin de la façon suivante :

- 19 dépassements observés et prévus pour les bassins de surface comprise entre 8 km² et 232 km²,
- 19 dépassements observés et prévus pour les bassins de surface de 288 km² à 944 km².

La moyenne du degré d'anticipation pour chacune des prévisions est déterminée pour ces 2 sous-groupes de bassins et représentée en Figure 27. A noter que les résultats présentés ci-dessous sont seulement indicatifs car la différenciation par la surface de bassin réduit à nouveau les échantillons de calcul.

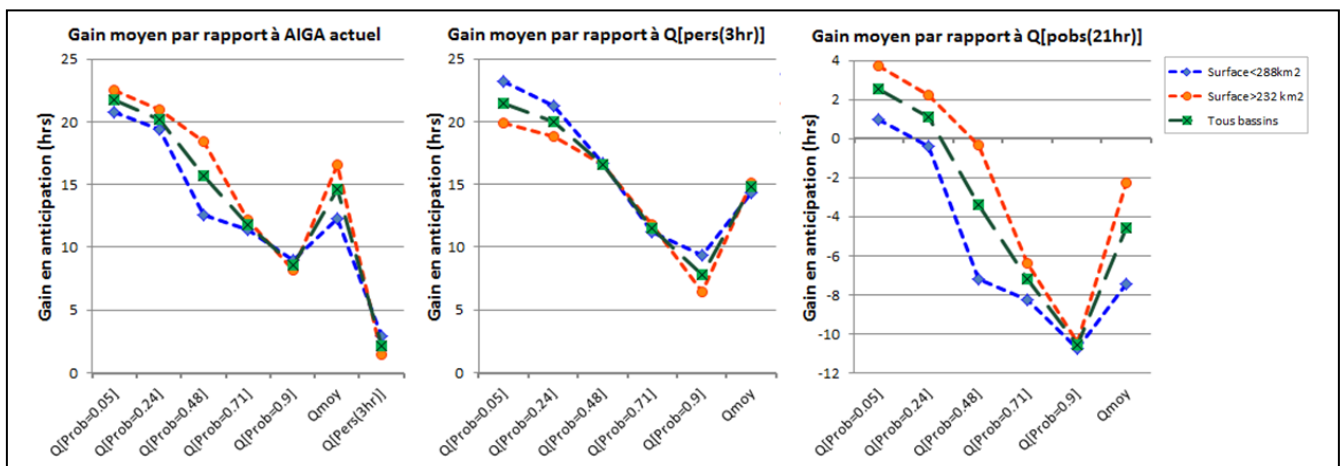


Figure 27 : Impact de la surface de bassin sur le gain moyen en anticipation des alertes de dépassement du débit biennal avec les différentes prévisions probabilistes (Q[Prob=*p*]) et la prévision moyenne des ensembles de débit (Qmoy) par rapport à la prévision sans pluie future (AIGA actuel), la prévision persistante de débit (Q[pers(3hr)]) et la prévision de débit issu de la pluie observée (Q[pobs(21hr)])

La Figure 27 montre une légère différenciation des gains moyens d'anticipation avec les prévisions probabilistes et la moyenne des ensembles de débit suivant la surface de bassin. Pour le gain moyen en anticipation vis-à-vis du système AIGA actuel et de la prévision de pluie observée, le gain est légèrement plus important pour les plus grands bassins pour les prévisions probabilistes de 5% à 48% et pour la moyenne des ensembles. Ainsi, pour le dépassement du débit biennal, les prévisions d'ensemble de pluie ont un impact plus positif sur l'anticipation des alertes pour les plus grands bassins. En particulier, pour ces valeurs fortes de débit, si certains membres de la prévision de pluie surestiment les pluies (s'agissant des prévisions probabilistes de 5% à 48%), ils peuvent compenser en partie les erreurs du modèle hydrologique et conduire à des alertes mieux anticipées.

Il faut noter que ces résultats sur le gain moyen en anticipation semblent différents des résultats obtenus avec les scores de vérification des prévisions d'ensemble de débit représentés dans la Figure 22, pour lesquels les scores de vérification des prévisions étaient meilleurs pour les petits bassins. Cependant, l'analyse de la qualité des alertes valorise toute anticipation des pics de débit, alors que la vérification des prévisions, qui est effectuée pour chaque échéance, en comparant la variable observée et la prévision correspondante, pénalise le décalage temporel. Ainsi une prévision qui anticipe le pic de crue obtiendra un bon degré d'anticipation mais cette anticipation va dégrader les scores de vérification du fait du décalage temporel. Il faut également noter que les différences de résultats s'expliquent également par les différences d'échantillons. En effet, les scores de vérification sont calculés avec les quantiles observés 75% et 90%, ce qui permet d'inclure, en moyenne sur les 27 bassins de l'analyse, 109 et 42 paires d'observations-prévisions de débit respectivement.

Pour le gain moyen vis-à-vis de la prévision de pluie persistante, le gain est légèrement accru pour les petits bassins avec les prévisions probabilistes. En effet, les prévisions d'ensemble de pluie permettent de mieux caractériser la variabilité des pluies futures sur les petits bassins. Cette variabilité de pluie étant moindre pour les plus grands bassins, l'apport des prévisions d'ensemble vis-à-vis de la pluie persistante est plus limité.

En conclusion, cette analyse sur l'amélioration du degré d'anticipation montre à nouveau l'intérêt d'inclure, dans la modélisation hydrologique étudiée, des prévisions de pluie COSMO-DE-EPS, même imparfaites, pour améliorer l'anticipation des détections de dépassement de seuil vis-à-vis des détections du système actuel sans pluie future. Les gains en anticipation sont également significatifs en comparaison de la prévision persistante de débit sur 3 heures d'échéance. Cependant, les erreurs des prévisions de pluie entraînent une dégradation du degré d'anticipation vis-à-vis des prévisions issues de la pluie observée, dégradation plus ou moins importante suivant les bassins et l'évènement étudiés.

5. Conclusions et Perspectives

Cette étude a permis d'analyser l'apport de prévisions d'ensemble de débit dans la chaîne d'alerte AIGA pour 7 événements de crue entre décembre 2010 et décembre 2012 et pour 29 bassins de la Meuse et de la Moselle de surfaces inférieures à 1000 km². L'incertitude des pluies futures est prise en compte en intégrant les 20 ensembles de pluie COSMO-DE-EPS dans la modélisation hydrologique. Ces prévisions d'ensemble sont produites 8 fois par jour à la résolution de 2,8 km (pour permettre de modéliser les processus convectifs) à partir de 4 modèles atmosphériques globaux et de 5 jeux de paramétrisation physique différente du modèle COSMO-DE. Les informations de pluie issues des radars sont également assimilées à

la date d'initialisation des prévisions. Les 20 ensembles équiprobables de prévision de pluie permettent de générer 20 prévisions d'ensemble de débit équiprobables et produire ainsi des alertes probabilisées de dépassement de seuil de débit. L'incertitude paramétrique du modèle hydrologique horaire semi-distribué régionalisé GR-SD a également été considérée en utilisant 5 jeux de paramètres différents, qui ont été testés lors de la régionalisation du modèle par Organde. Ces 5 jeux de paramètres sont considérés comme équiprobables, même si seul le jeu de paramètres finaux a été optimisé pour la détection des dépassements de quantiles de débit. Les 5 jeux de paramètres sont couplés aux 20 ensembles de pluie pour produire 100 ensembles de débit qui prennent en compte l'incertitude de pluie future et l'incertitude paramétrique du modèle semi-distribué.

L'apport des prévisions d'ensemble COSMO-DE-EPS est analysé vis-à-vis de prévisions de référence : la prévision de débit générée avec la pluie persistante issue de la dernière heure de pluie observée au moment de l'initialisation de la prévision, la prévision probabiliste climatologique définie par la distribution empirique des observations de 2009-2012, ainsi que la prévision de débit issue de la pluie « parfaite » (i.e. pluie observée) pour montrer l'impact des erreurs et des incertitudes des prévisions d'ensemble de pluie. La vérification des prévisions d'ensemble de pluie et de débit est réalisée avec l'outil de vérification EVS (Ensemble Verification System) en sélectionnant des critères pour décrire le biais relatif et la corrélation des moyennes des ensembles avec les observations, ainsi que le gain en termes d'erreur en probabilité et de discrimination des dépassements de quantiles observés. L'analyse de la qualité des alertes inclut différents critères pour décrire notamment la fiabilité, le biais en fréquence et l'anticipation des alertes.

Grâce aux prévisions d'ensemble de débit, on présente, pour un quantile de débit donné, des cartes d'alerte probabilisées décrivant pour chaque bassin la probabilité de dépasser le quantile. Ces cartes d'alertes probabilisées sont complétées avec les cartes de performance des alertes pour offrir une visualisation spatiale de la qualité des alertes sur les bassins jaugés. De telles cartes peuvent renseigner sur la qualité potentielle des alertes sur les bassins non jaugés voisins.

Les principaux résultats de l'étude peuvent être résumés ainsi :

- Les moyennes des ensembles de débit montrent une bonne corrélation avec les observations mais une sous-estimation significative pour les débits les plus forts, à cause du biais négatif du modèle hydrologique et, dans une moindre mesure, du biais des ensembles de pluie COSMO-DE-EPS. Par exemple, l'erreur relative moyenne du pic de débit sur 21 heures par rapport au débit observé (ce qui inclut l'impact des erreurs hydrologiques et météorologiques) est de -20%, alors que le biais par rapport au débit simulé (incluant uniquement l'impact des erreurs des prévisions de pluie) est de +3%. Cependant, certaines surestimations des pluies moyennes peuvent compenser les erreurs de sous-estimation du modèle hydrologique.
- Les ensembles de pluie et de débit ont une bonne capacité pour discriminer les dépassements des quantiles observés 75% et 90%. En comparaison de la prévision probabiliste climatologique et de la prévision persistante, les ensembles ont de meilleures performances, notamment en termes d'erreur en probabilité. Pour les erreurs en probabilité des ensembles mesurées avec le score CRPS, le gain en termes de CRPS lorsque les ensembles de débit COSMO sont comparés à la prévision probabiliste climatologique pour le pic de débit sur 21 heures est de 9%. Ce gain avec la prévision persistante de débit pour le pic de débit sur 12 heures, est de 39%. Il

semble également que ces gains en performance soient légèrement plus importants pour les plus petits bassins de surface (jusqu'à 232 km²).

- Pour les alertes de dépassement du quantile de débit biennal, les ensembles de débit COSMO ont permis une amélioration significative en termes de détection, fiabilité et anticipation en comparaison au système AIGA actuel (qui n'intègre pas de prévision de pluie) et aux alertes issues de la prévision de pluie persistante. Par exemple, pour la prévision de probabilité de 71%, la probabilité de détection est de 58% (contre 26% pour le système AIGA actuel) et le Critical Success Index est de 43% (contre 22% pour le système AIGA actuel) ; la moyenne du gain en anticipation vis-à-vis du système AIGA actuel est d'environ 11 heures sur les 24 dépassements prévus. L'utilisation de 5 jeux de paramètres pour le modèle hydrologique semi-distribué pour prendre en compte l'incertitude paramétrique n'a pas permis d'améliorer les performances de fiabilité et de détection des alertes. Ceci s'explique par le fait qu'un seul des 5 jeux de paramètres a été optimisé pour augmenter la réactivité du modèle hydrologique et mieux différencier les dépassements des différents quantiles de débit.

Il faut noter que l'évaluation des ensembles de pluie de bassin et de débit et l'évaluation de la qualité des alertes et de l'anticipation sont complémentaires. En effet, la vérification des prévisions, qui est effectuée pour chaque échéance, compare la variable observée et la prévision correspondante et va, de ce fait, pénaliser les éventuels décalages temporels et donc l'anticipation (cette pénalisation peut être limitée en vérifiant, par exemple, le débit maximal ou le cumul de pluie sur l'échéance maximale de prévision). Ainsi une prévision qui anticipe le pic de crue obtiendra un bon degré d'anticipation mais cette anticipation peut dégrader les scores de vérification de la prévision du fait du décalage temporel.

Des pistes d'amélioration peuvent être proposées pour compléter cette analyse et pour les améliorations futures de la chaîne d'alerte AIGA :

- il serait nécessaire de poursuivre l'évaluation des performances avec les données plus récentes de 2013-2014, et sur un plus grand nombre de bassins pour travailler avec des échantillons plus importants ; cela pourrait éventuellement permettre d'étudier les dépassements du débit décennal ; il est cependant nécessaire de s'assurer que les observations de débit sont également critiquées par le réseau de SPC pour intégrer des informations de qualité satisfaisante pour les hautes eaux ; l'évaluation de la qualité des alertes devrait aussi utiliser des retours post-crue sur les dégâts observés sur le terrain pour ne pas limiter l'analyse aux résultats en bassins jaugés ;
- les seuils de débit utilisés dans cette étude (issus des quantiles de débit SHYREG de la base de données version 1) peuvent être redéfinis à partir des débits simulés pour prendre en compte les biais du modèle pluie-débit ; les seuils peuvent également être modulés en fonction des enjeux, pour une prise en compte de la vulnérabilité et fournir ainsi des avertissements sur le risque (par croisement entre les informations d'aléa et de vulnérabilité) ;
- les alertes basées sur les prévisions de débit pourraient être réactualisées toutes les heures en intégrant les dernières observations de pluie et valeurs de simulation (dans cette étude, les prévisions sont produites toutes les 3 heures) pour être plus cohérent vis-à-vis des alertes du système AIGA ;
- la prise en compte des incertitudes de prévisions de pluie peut être améliorée, notamment en testant l'apport de prévisions immédiates issues des données radar des dernières heures (grâce aux techniques de nowcasting) ; ces prévisions ont vocation à être utilisées en complément des prévisions produites par les modèles atmosphériques

tels qu'AROME ou COSMO-DE ; de plus, les ensembles issus du système COSMO-DE-EPS sont en cours d'amélioration, notamment avec un post-traitement statistique pour une meilleure fiabilité ; idéalement la chaîne d'alerte AIGA devrait bénéficier de la combinaison des différentes prévisions de type déterministe (prévisions pouvant définir des ensembles timelag par décalage temporel), ensembliste ou probabiliste ; les incertitudes sur les pluies observées pourraient également être prises en compte (cf. le travail de thèse d'Angelica Caseri, Caseri et al. 2014), ce qui permettrait d'améliorer la dispersion (et donc la fiabilité) des ensembles de débit pour les toutes premières échéances ;

- les incertitudes hydrologiques ont un impact très significatif sur les résultats, notamment du fait de l'application d'un modèle semi-distribué horaire en milieu non-jaugé après régionalisation des paramètres ; les améliorations du modèle pluie-débit concernent notamment l'initialisation du modèle, grâce à l'intégration des informations sur les derniers débits observés pour des bassins jaugés voisins ; l'assimilation de données peut également permettre d'initialiser le modèle de manière ensembliste ; la prise en compte des incertitudes de paramétrisation du modèle doit également être poursuivie pour travailler avec des jeux de paramètres équiprobables ; par ailleurs, Irstea Aix-en-Provence travaille actuellement sur l'estimation des incertitudes des quantiles de débit SHYREG ; plus généralement, toute amélioration de la modélisation hydrologique (par exemple, prise en compte des processus de neige, modélisation au pas de temps infra-horaire) et de la chaîne d'alerte AIGA pourra également conduire à des améliorations pour la prise en compte des incertitudes.

6. Références

- Arnaud P., Y. Aubert, D. Organde, P. Cantet, C. Fouchier, N. Folton, 2014. Estimation de l'aléa hydro-météorologique par une méthode par simulation: La méthode SHYREG : Présentation – Performances – Base de données, *La Houille Blanche* n°2, 20-26.
- Aubert Y., P. Arnaud, P. Ribstein et J. A. Fine, 2014. La méthode SHYREG débit, application sur 1605 bassins versants en France Métropolitaine. *Hydrological Sciences Journal*, 59 (5), 1-13.
- Ben Bouallègue Z., Theis S. E., Gebhardt C. 2013. Enhancing COSMO-DE-EPS ensemble forecasts by inexpensive techniques. *Meteorologische Zeitschrift*, 22(1), 49-59.
- Brown J. D., Demargne J., Seo D.-J., et Y. Liu, 2010. The Ensemble Verification System (EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations. *Environmental Modelling Software*, 25, 854-872.
- Casari A., Javelle P., Ramos M.H., Leblois E., 2014. Generating precipitation ensembles for flood alert and risk management. *Journal of Flood Risk Management* (soumis).
- Cloke, H.L., et F. Pappenberger, 2009. Ensemble flood forecasting: a review. *Journal of Hydrology*, 375, 613-626.
- Demargne J., 2013. Application de la méthode de Flash Flood Guidance avec AIGA pour la prévision des crues éclairs en milieu non jaugé. Convention DGPR/SCHAPI – IRSTEA 2012, *Rapport d'étude HYDRIS Hydrologie*, 63 p.
- Demargne J. 2014a. Traitement des grilles de pluies prévues du système COSMO-DE-EPS pour leur intégration dans le système AIGA. Convention DGPR/SCHAPI – IRSTEA 2014, *Rapport d'étude HYDRIS Hydrologie*, 29 p.
- Demargne J. 2014b. Amélioration de l'approche AIGA par l'intégration des pluies prévues. Convention DGPR/SCHAPI – IRSTEA 2013, *Rapport d'étude HYDRIS Hydrologie*, 71 p.
- Demargne, J., Wu, L., Regonda, S., Brown, J.D., Lee, H., He, M., Seo, D.-J., Hartman, R., Fresch, M., Schaake, J., et Zhu, Y. 2014. The science of NOAA's operational Hydrologic Ensemble Forecast Service. *Bulletin of American Meteorological Society*, 20 p. doi: 10.1175/BAMS-D-12-00081.1
- Edijatno et Michel C., 1989. Un modèle pluie-débit à trois paramètres. *La Houille Blanche*, 2, 113-121.
- Edijatno, Nascimento N.O., Yang X., Makhlof Z. et Michel C., 1999. GR3J : a daily watershed model with three free parameters. *Hydrological Sciences Journal*, 44(2), 263-278.
- Gebhardt C., Theis S. E., Paulat M., Ben Bouallègue Z., 2011. Uncertainties in COSMO-DE precipitation forecasts introduced by model perturbations and variation of lateral boundaries. *Atmospheric Resources*, 100(2-3), 168–177.
- Gebhardt C., Theis S. E., Ben Bouallègue Z., Ropnack A., Schuhen N., Buchhold M., 2012. COSMO-DE-EPS: an operational convection permitting ensemble prediction system for the atmosphere. *International Conference on Ensemble Methods in Geophysical Sciences*, 12-16 novembre 2012, Toulouse, 48 p.

- Javelle P., Fouchier C., Arnaud P. et Lavabre J., 2010. Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *Journal of Hydrology*, 394(1-2), 267-274.
- Javelle P., Demargne J., Defrance D., Arnaud P., 2014. Evaluating flash flood warnings at ungauged locations using post-event surveys: a case study with the AIGA warning system. *Hydrological Sciences Journal*. doi: 10.1080/02626667.2014.923970
- Jolliffe I.T., Stephenson, D.B., 2003. Forecast Verification: A Practitioner's Guide in Atmospheric Science, Wiley and Sons, Chichester, England, 2003, 254 p.
- Maire A. et Javelle P., 2012. Extension nationale de la méthode AIGA : Comparaison à des avertissements basés sur des dépassements de seuil de pluie. *Rapport d'étude Irstea*, 44 p.
- National Research Council of the National Academies (NRC), 2006. Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts. The National Academies Press, 124 pp.
- Organde D., Arnaud P., Fine J.-A., Fouchier C., Folton N. et Lavabre J., 2013. Régionalisation d'une méthode de prédétermination de crue sur l'ensemble du territoire français : la méthode SHYREG. *Revue des Sciences de l'Eau*, 26 (1) : 65-78.
- Organde D., 2013. Extension nationale de la méthode AIGA, Évaluation de AIGA débit sur un échantillon de bassins versants du territoire national. *Rapport d'étude HYDRIS Hydrologie*, 39 p.
- Organde D., 2014. Extension nationale de l'approche AIGA, Amélioration de la réactivité du modèle. *Rapport d'étude HYDRIS Hydrologie*, 20 p.
- Oudin L., Hervieu F., Michel C., Perrin C., Andréassian V., Anctil, F. et Loumagne, C., 2005. Which potential evapotranspiration input for a rainfall-runoff model? Part 2 - Towards a simple and efficient PE model for rainfall-runoff modelling. *Journal of Hydrology*, 303(1-4), 290-306.
- Perrin C., 2002. Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative. *La Houille Blanche*, 6-7, 84-91.
- Perrin C., Michel C., et Andréassian V., 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1-4), 275-289.
- Tabary P. 2007. The New French operational radar rainfall product. Part I: Methodology. *Weather and Forecasting*, 22(3), 393-408.
- Tabary P. et al. 2011. Evaluation of two integrated polarimetric Quantitative Precipitation Estimation (QPE) algorithms at C-band. *Journal of Hydrology*, 405(3-4), 248-260.
- Theis S., Gebhardt C. et Buchhold M., 2014. The convection-permitting ensemble COSMO-DE-EPS: from development to applications. The World Weather Open Science Conference, 16-21 août 2014, Montréal, Canada, 60 p.
- Wilks D.S., 2006. Statistical Methods in Atmospheric Sciences, Academic Press, San Diego, California, 627 p.
- World Meteorological Organization (WMO), 2013. Forecast Verification – Issues, Methods and FAQ, WWRP/WGNE Joint Working Group on Verification, <http://www.cawcr.gov.au/projects/verification/>
- Zappa, M., et co-auteurs, 2010. Propagation of uncertainty from observing systems and NWP into hydrological models: COST-731 Working Group 2. *Atmospheric Science Letters*, 11 (2), 83-91.

Appendice A : Liste des Bassins Versants

Tableau 2 : Liste des 29 bassins calibrés avec des prévisions COSMO-DE-EPS disponibles pour la zone Meuse-Moselle
Sont inclus les attributs suivants : identifiant du bassin (ID bassin), code de la base de données SCHAPI (Code), surface de bassin (de la base de données SCHAPI), nom du bassin (Nom), identifiant de l'exutoire du bassin (ID Exutoire), indice d'influence de la méthode SHYREG (Influence SHYREG), quantiles de la base de données SHYREG version 1.0 pour les périodes de retour de 2 ans, 10 ans et 50 ans.

ID bassin	Code	Surface (km ²)	Nom	ID Exutoire	Influence SHYREG	Q2ans (m ³ /s)	Q10ans (m ³ /s)	Q50ans (m ³ /s)
1	A1202050	8.6	La Doller à Sewen [Lerchenmatt]	RS_12184	0.00	6.98	13.54	23.46
2	A1402020	7.7	La Thur à Wildenstein	RS_10861	0.00	6.58	11.99	19.66
3	A2023030	44.1	La Petite Fecht à Stosswihr [Village]	RS_11193	0.01	13.7	26.86	46.68
4	A2113010	38.6	La Behine à Lapoutroie	RS_2785	0.00	8.36	16.28	27.1
5	A4200630	626.0	La Moselle à Saint-Nabord [Noir Gueux]	MEMO_2640	0.00	268.13	426.91	647.33
6	A4632010	104.9	L_Aviere à Frizon [Basse Frizon]	MEMO_2448	0.15	25.47	42.11	64.57
7	A5431010	943.7	Le Madon à Pulligny	MEMO_890	0.00	162.48	254.67	368.21
8	A6051020	368.1	La Meurthe à Saint-Die	MEMO_730	0.00	54.73	93.68	140.91
9	A6431120	113.8	La Vezouze à Blamont	MEMO_3118	0.00	16.01	28.82	44.75
10	A6731220	496.2	La Mortagne à Gerbeviller	MEMO_908	0.00	62.31	102.91	152.85
11	A7581020	350.1	La Seille à Moyenvic	MEMO_2190	0.00	27.31	48.37	76.26
12	A7821010	923.3	La Seille à Nomeny	MEMO_1688	0.00	61.6	107.96	166.87
14	A8322010	74.7	Le Woigot à Briey	MEMO_5601	0.00	16.22	27.26	41.63
15	A9013050	89.7	La Sarre Rouge à Vasperviller	RS_6636	0.00	11.62	21.96	34.57
16	A9021010	310.1	La Sarre à Sarrebourg	RS_4209	0.00	37.5	65.88	99.88
17	A9072050	147.1	L'Isch à Postroff	RS_4252	0.00	18.37	33.28	52.15
18	A9091050	884.5	La Sarre à Keskastel	RS_3805	0.07	85.74	148.33	222.41
19	A9091060	715.5	La Sarre à Diedendorf	RS_3790	0.05	72.14	125.51	188.73

20	B1092010	402.7	Le Mouzon à Circourt-sur-Mouzon [Villars]	MEMO_2116	0.00	74.31	115.88	167.39
21	B1282010	440.8	Le Vair à Soulosse-sous-Saint-elophe	MEMO_429	0.00	71.49	113.06	163.52
22	B4001010	157.7	La Chiers à Longlaville	MEMO_3828	0.00	34.97	55.47	80.08
23	B4031010	288.3	La Chiers à Montigny-sur-Chiers	MEMO_2842	0.00	48.28	77.78	112.06
51	U0020010	231.7	La Saone à Monthureux-sur-Saone	Rhone_11280	0.00	42.17	67.69	100.31
52	U0104010	63.4	Le Coney à Xertigny	Rhone_16410	0.00	7.09	13.66	22.09
53	U0124010	313.0	Le Coney à Fontenoy-le-Chateau	Rhone_9056	0.00	42.38	71.89	108.38
54	U0415010	121.9	Le Breuchin à la Proiseliere-et-Langle	Rhone_5022	0.00	56.45	89.36	135.15
55	U0444310	225.1	La Semouse à Saint-Loup-sur-Semouse	Rhone_5162	0.00	60.14	93.98	138.02
56	U1004010	71.5	L_Ognon à Servance [Fourguenons]	Rhone_5047	0.00	47.39	79.62	125.58
57	U1014020	168.3	L_Ognon à Montessaux	Rhone_5059	0.00	65.81	108.92	168.57

Appendice B : Liste des Figures et Tableaux

Liste des Figures

Figure 1 : Identification et prise en compte des principales sources d'incertitude dans la chaîne d'alerte aux crues soudaines du système AIGA ..	7
Figure 2 : Domaine des prévisions d'ensemble COSMO-DE-EPS vis-à-vis de la zone d'étude Meuse-Moselle et complémentarité avec les prévisions AROME de Météo-France et prévisions MOGREPS de l'agence UK Met (d'après Gebhardt et al. 2012).....	8
Figure 3 : Schéma du système de prévision d'ensemble COSMO-DE-EPS considérant l'incertitude d'initialisation et des conditions aux limites ainsi que l'incertitude de paramétrisation du modèle déterministe COSMO-DE (d'après Gebhardt et al. 2012 et Theis et al. 2014).....	9
Figure 4 : Carte des 59 bassins versants calibrés de la zone d'étude Meuse-Moselle (par classes de surface de bassin), incluant les 29 bassins versants avec des prévisions COSMO et leur distribution de surface	11
Figure 5 : Schéma du couplage des modèles journaliers continus SAJ et GR4J et du modèle horaire semi-distribué GR-SD pour le système AIGA national.....	12
Figure 6 : Résultats du calage et de la régionalisation du modèle horaire semi-distribué GR-SD avec la carte des bassins de la zone Meuse-Moselle et la carte des hydro-écorégions utilisées pour la régionalisation de GR-SD.....	13
Figure 7 : Exemple d'hydrogrammes avec le débit simulé, l'enveloppe et la moyenne des 20 ensembles de débit COSMO, la prévision de débit issu de la pluie persistante pour la date d'initialisation du 15 décembre 2011 à 18h (TU) en comparaison avec l'hydrogramme de débit observé pour le bassin de La Doller à Sewen.....	17
Figure 8 : Exemples d'hydrogrammes pour les événements de janvier 2011 et décembre 2011	18
Figure 9 : Exemples d'hydrogrammes pour l'événement de décembre 2012	19
Figure 10 : Exemple de carte d'alerte issue des 20 ensembles de débit COSMO pour le dépassement du débit biennal, en utilisant 5 niveaux de probabilité définis à partir de la distribution de probabilité empirique, pour la date du 15/12/2011 16hr (TU).....	20
Figure 11 : Exemples de carte d'alerte issue des 20 ensembles de débit COSMO pour le dépassement du débit biennal pour les événements de décembre 2010 et décembre 2011.....	21
Figure 12 : Estimation des scores de Continuous Rank Probability Score (CRPS) et de Relative Operating Curve Score (ROCS) pour une prévision probabiliste.....	25

Figure 13 : Histogrammes des erreurs des prévisions d'ensemble de pluie de bassin en fonction de la pluie observée pour différentes agrégations (cumuls de pluie sur 1hr, 3hr, 6hr, 12 et 21hr) et en fonction de l'horizon de prévision	27
Figure 14 : Impact de l'échéance sur les scores de vérification des prévisions d'ensemble de pluie de bassin pour tout l'échantillon et pour les seuils des quantiles observés 75% et 90%	29
Figure 15 : Impact de l'agrégation sur les scores de vérification des prévisions d'ensemble du cumul de pluie de bassin sur 1hr, 3hr, 6hr, 12hr et 21hr pour le seuil de quantile observé 90%	30
Figure 16 : Histogrammes des erreurs des ensembles de débit COSMO en fonction du débit observé et du débit simulé pour le débit maximal sur 3hr pour les horizons de 3hr et 12hr	31
Figure 17 : Histogrammes des erreurs des ensembles COSMO de débit spécifique en fonction du débit spécifique observé et du débit spécifique simulé pour le débit maximal sur 3hr pour les horizons de 3hr et 12hr.....	32
Figure 18 : Impact de l'échéance sur les scores de vérification des prévisions d'ensemble de débit pour tout l'échantillon et pour les seuils des quantiles observés 75% et 90%.....	33
Figure 19 : Impact de l'agrégation sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour les seuils des quantiles observés 75% et 90%.....	34
Figure 20 : Comparaison des scores de vérification des prévisions d'ensemble COSMO de débit maximal avec la prévision de débit issu de la pluie persistante sur 12 hrs pour les seuils des quantiles observés 75% et 90%	36
Figure 21 : Comparaison de l'impact des erreurs de sources hydrologique et météorologique (en vérifiant avec le débit observé) et de l'impact des erreurs de source météorologique (en vérifiant avec le débit simulé) sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour tout l'échantillon.....	37
Figure 22 : Impact de la surface de bassin sur les scores de vérification des prévisions d'ensemble COSMO de débit maximal pour les seuils des quantiles observés 75% et 90%.....	38
Figure 23 : Exemples de cartes de vérification des alertes pour les événements de décembre 2010 et décembre 2011 ; les identifiants des bassins sont données lorsque l'alerte est correcte	40
Figure 24 : Scores de contingence des alertes de dépassement du débit biennal avec les prévisions probabilistes de débit ($Q[Prob=p]$ avec p égal à 0.05, 0.48, 0.71 et 0.9) issus de 20 ensembles et 100 ensembles, la prévision de débit issu de la pluie persistante ($Q[pers(3hr)]$), la prévision de débit sans pluie future ($Q[simu]$) et la prévision de débit issue de la pluie observée ($Q[pobs(21hr)]$).....	41

Figure 25 : Anticipation des alertes de dépassement du débit biennal avec les prévisions probabilistes de débit ($Q[\text{Prob}=p]$ avec p égal à 0.05, 0.48 et 0.9), la prévision de débit issu de la pluie persistante ($Q[\text{pers}(3\text{hr})]$) et la prévision de débit sans pluie future (Q_{simu}) pour les événements de décembre 2010 et décembre 2011	43
Figure 26 : Statistiques des gains en anticipation des alertes de dépassement du débit biennal avec les différentes prévisions probabilistes ($Q[\text{Prob}=p]$) et la prévision moyenne des ensembles de débit (Q_{moy}) par rapport à la prévision sans pluie future (Q_{simu}), la prévision persistante de débit ($Q[\text{pers}(3\text{hr})]$) et la prévision de débit issu de la pluie observée ($Q[\text{pobs}(21\text{hr})]$)	44
Figure 27 : Impact de la surface de bassin sur le gain moyen en anticipation des alertes de dépassement du débit biennal avec les différentes prévisions probabilistes ($Q[\text{Prob}=p]$) et la prévision moyenne des ensembles de débit (Q_{moy}) par rapport à la prévision sans pluie future (AIGA actuel), la prévision persistante de débit ($Q[\text{pers}(3\text{hr})]$) et la prévision de débit issu de la pluie observée ($Q[\text{pobs}(21\text{hr})]$)	45

Liste des Tableaux

Tableau 1 : Caractéristiques des 7 événements sélectionnés avec le nombre de dépassements observés et de dépassements du débit simulé pour les quantiles de débit biennal et décennal de la base de données SHYREG (version 1) pour les 29 bassins versants de l'étude.....	15
Tableau 2 : Liste des 29 bassins calibrés avec des prévisions COSMO-DE-EPS disponibles pour la zone Meuse-Moselle	52