

Utilisation des prévisions saisonnières de pluie produites par Météo France (modèle ARPEGE Système 5) pour la prévision du volume naturel de crue du fleuve Sénégal.

Jean-Claude Bader

UMR G-EAU

Institut de Recherche pour le Développement, BP 64501, 34394 Montpellier Cedex 5, France

Le 2 octobre 2015

RÉSUMÉ ET PERSPECTIVES

Les index pluviométriques reconstitués en mode prévision par le modèle Arpege Système 5 sur l'Afrique de l'ouest pour la période 1991-2014 sont utilisés pour mettre au point deux modèles de prévision du débit naturel de crue du fleuve Sénégal à Bakel, calés sur 16 années et validés sur 8 années d'observations.

- La prévision du débit de septembre-octobre utilise les index prévus début août pour les mois de septembre et octobre sur une zone L8a8 de 1200X1200 km comprise entre les longitudes 10.54 et 21.79 et les latitudes -1.40 et 9.81.
- La prévision du débit d'août-septembre-octobre utilise les index prévus début juillet pour les mois d'août, septembre et octobre sur une zone B8a8 de 1200X1200 km comprise entre les longitudes -17.58 et -6.33 et les latitudes -1.40 et 9.81.

L'utilisation de ces modèles en temps réel donne les débits prévus suivants pour l'année 2015 :

- Débit moyen de septembre-octobre: 1668 m³/s (valeur dépassée 11 fois sur 24 sur la période 1991-2012 par les débits observés)
- Débit moyen de août-septembre-octobre : 1337 m³/s (valeur dépassée 20 fois sur 24 sur la période 1991-2012 par les débits observés)

Ces prévisions doivent être considérées avec beaucoup de prudence, car les résultats obtenus pour la mise au point des modèles montrent qu'il existe un risque non négligeable que les bonnes performances de ceux-ci en calage et en validation ne résultent que du seul hasard (faibles échantillons de calage et validation ; nombreux modèles testés ; non validation des modèles les mieux calés). Seule l'utilisation en temps réel des modèles pendant plusieurs années permettrait d'en vérifier la robustesse.

Une nouvelle piste mériterait d'être explorée pour réaliser l'adaptation statistique des index pour la prévision des débits. Elle consisterait à les pondérer en fonction de la latitude des mailles pour chercher des relations plus réalistes, qui tiendraient compte de la diminution globale de pluviométrie observée du sud vers le nord sur l'Afrique de l'ouest.

OBJECTIF

L'étude fait suite aux travaux précédents concernant la prévision saisonnière de crue du fleuve Sénégal, réalisés depuis 2000 grâce à la collaboration entre l'IRD, Météo France et l'OMVS. Elle vise à mettre au point de nouveaux modèles de prévision, basés sur l'adaptation statistique des prévisions saisonnières de pluie produites par le modèle le plus récent de Météo France (Arpege Système 5). Les nouveaux modèles sont utilisés pour prévoir le débit moyen de la crue de 2015.

DONNÉES

Prévisions saisonnières de pluie

Pour les années 1991 à 2015, on dispose d'index de précipitations centrées réduites (période 1991-2011) sur un maillage de 150X150 km couvrant l'Afrique de l'Ouest entre les latitudes -0.7 et 14.708 et les longitudes -19.688 et 29.531 (centres des mailles). Sur chacune de ces 432 mailles (fig.1), les index sont produits en mode prévision par Arpege Système 5 aux dates suivantes :

- Début juillet : un index par mois pour la période de juillet à octobre
- Début août : un index par mois pour la période d'août à novembre

Débit naturel du fleuve Sénégal à la station de Bakel

Le débit naturel du fleuve à Bakel est connu au pas de temps journalier depuis l'origine des données à la station (1904) jusqu'à fin 2014. Ces débits, qui correspondent au régime d'écoulement observé jusqu'en 1986 et à un régime reconstitué ensuite (influence du barrage de Manantali), ont été réévalués récemment par l'IRD à l'occasion d'une étude réalisée pour l'OMVS (actualisation de la monographie hydrologique du fleuve Sénégal, 2013).

MISE AU POINT DES MODÈLES

Deux modèles sont mis au point par l'étude :

- Le premier modèle vise à prévoir le débit moyen naturel du fleuve de la période septembre-octobre à partir des index de septembre et d'octobre calculés début août.
- Le second modèle vise à prévoir le débit moyen naturel du fleuve de août à octobre à partir des index d'août, de septembre et d'octobre calculés début juillet

Pour l'essentiel, la mise au point de chacun de ces modèles repose sur la même méthode de calcul que les modèles antérieurs, rappelée brièvement ci-dessous.

Méthode

Regroupement des mailles d'Arpege par zones élémentaires.

Exploitant successivement les résultats d'Arpege 3 forcé, Arpege 4.6 forcé et Arpege 4.5 couplé, les études antérieures ont chacune identifié des zones carrées de 1200X1200km parmi différentes formes testées, où les index pluviométriques s'avéraient les plus prédictifs pour le débit du Sénégal. Seul ce type de zone élémentaire est testé ici pour exploiter les résultats d'Arpege Système 5, comme il avait été procédé en 2013 pour ceux d'Arpege Système 4.

La nomenclature de zone est adaptée à la grille des résultats d'Arpege Système 4 et Système 5 (fig. 1). Une lettre comprise entre A et O représente la limite ouest de chaque zone, alors qu'une lettre comprise entre a et c représente la limite sud. Ainsi, une zone G16c8 a son angle

sud-ouest situé à la longitude -3.52 et la latitude 4.20. Elle s'étend sur 16 mailles (2400 km) vers l'est et 8 mailles (1200 km) vers le nord.

Les 45 zones élémentaires testées ici sont toutes du type X8y8. Elles s'étendent de la zone A8a8 pour la plus au sud-ouest, jusqu'à O8c8 pour la plus au nord-est.

Séparation des données en échantillons de calage et de validation

Sur la période 1991-2014, on dispose à la fois des données d'index et de débit. Parmi ces 24 années, seize sont utilisées pour le calage des modèles et les huit autres pour leur validation. Les 8 années de validation, choisies pour représenter au mieux la variabilité interannuelle du débit naturel observé sur les 24 années, correspondent aux rangs suivants des débits observés classés par ordre croissant : 2, 5, 8, 11, 14, 17, 20, 23. Ce sont les années 2002, 1992, 2011, 2001, 2008, 2007, 2003 et 1999 pour le débit de septembre-octobre et 1993, 1996, 1997, 1991, 2008, 2009, 1994 et 1999 pour le débit d'août-septembre-octobre.

Analyse en composantes principales des index sur l'échantillon de calage

Pour chaque zone élémentaire testée, comportant 64 mailles, les index pluviométriques constituent des variables de base dont l'effectif très important (128 pour la prévision SO et 192 pour la prévision ASO) dépasse largement la taille de l'échantillon de calage (16). L'ACP détermine des vecteurs propres (VP) obtenus par combinaisons linéaires de ces variables de base, indépendants entre eux et classés par variance décroissante. On retient les 8 premiers vecteurs propres, qui agrègent la majeure partie de l'information contenue dans les index.

Recherche de relation entre les vecteurs propres d'index et les débits sur l'échantillon de calage

Pour chaque zone élémentaire de 64 mailles, une première relation est établie sur les années de calage par régression multiple entre les 5 premiers vecteurs propres d'index (V_1 à V_5) et le débit Q . Cinq autres relations sont ensuite établies par régression multiple descendante entre les n vecteurs propres les plus prédictifs (M_1 à M_n , déterminés parmi V_1 à V_8) et Q , pour n compris entre 5 et 1.

Sélection de modèles en fonction des résultats obtenus en calage et en validation

Parmi les 270 modèles testés (45 zones X 6 types de modèle), on ne conserve que ceux donnant à la fois de bons résultats sur les années de calage ($R^2 = NSE > 0.48$) et des résultats acceptables sur les années de validation ($NSE > 0.2$, correspondant à une erreur moyenne quadratique inférieure à 479 m³/s pour septembre-octobre et 461 m³/s pour août-septembre-octobre).

Parmi les modèles conservés, on sélectionne sur chaque zone concernée celui qui donne le plus fort NSE en validation.

Choix du modèle à utiliser pour la prévision en temps réel

Parmi les modèles sélectionnés, on écarte ensuite tous ceux pour lesquels il existe au moins un autre modèle présentant des valeurs inférieures à la fois pour le NSE de calage et pour le NSE de validation.

A l'issue de ce processus ne restent que quelques modèles en lice, parmi lesquels on choisit le modèle à utiliser en tenant compte des NSE et de l'éloignement des zones concernées par rapport au bassin du Sénégal.

Résultats obtenus

Prévision début août du débit moyen de septembre-octobre

Les résultats obtenus montrent que 10 modèles concernant 4 zones sont à la fois satisfaisants en calage et assez satisfaisants en validation (tableau 1, surlignage jaune). En sélectionnant sur chaque zone concernée le meilleur de ces modèles en validation (tableau 1, valeurs en gras), on obtient finalement 4 modèles parmi lesquels deux sont encore rejetés, du fait que leurs résultats sont moins bons à la fois en calage et en validation que ceux des deux modèles suivants, qui restent en lice :

- régression à partir des vecteurs propres d'index M1 à M3 sur la zone L8a8
- régression à partir des vecteurs propres M1 à M4 sur la zone O8c8

Par rapport au premier modèle, le second est légèrement meilleur en calage mais moins bon en validation. Il utilise par ailleurs les index d'une zone beaucoup plus éloignée du bassin du Sénégal, ce qui nous conduit à l'écartier.

C'est donc la régression établie à partir des vecteurs M1 à M3 des index de la zone L8a8 qui est retenue comme modèle pour prévoir le débit moyen de septembre-octobre du fleuve Sénégal (fig. 2, tab. 3). Ce modèle présente une erreur type de calage de 406 m³/s et une erreur moyenne quadratique valant respectivement 392 et 365 m³/s sur les 8 années de validation et les 24 années de l'ensemble des données. Son coefficient d'efficacité NSE s'élève respectivement à 0.539, 0.463 et 0.512 en calage, en validation et sur la période globale. Les coefficients multiplicateurs d'index du modèle ne présentent pas d'organisation spatiale évidente (fig.3).

Prévision début juillet du débit moyen d'août-septembre-octobre

Après analyse des résultats obtenus en calage et en validation, on ne retient que 37 des 270 modèles testés (tableau 2, surlignage jaune), parmi lesquels sont sélectionnés les 16 modèles donnant les meilleurs résultats en validation sur les zones concernées (tableau 2, valeurs en gras). Parmi ces 16 modèles, 13 sont éliminés car ils donnent de moins bons résultats à la fois en calage et en validation que les trois modèles restants, qui utilisent les index des zones B8a8, C8a8 et A8b8.

Du fait qu'elle donne les meilleurs résultats sur la période globale parmi les 3 modèles encore en lice, c'est finalement la régression établie à partir des vecteurs M1 à M4 des index de la zone B8a8 qui est choisie comme modèle pour prévoir le débit moyen de août-septembre-octobre du fleuve Sénégal (fig. 4, tab. 3). Ce modèle présente une erreur type de 372 m³/s sur la période de calage et une erreur quadratique moyenne valant respectivement 384 et 335 m³/s sur la période de validation et la période globale. Sur ces trois périodes, son coefficient d'efficacité NSE s'élève respectivement à 0.661, 0.447 et 0.593. Les coefficients multiplicateurs d'index du modèle ne présentent pas d'organisation spatiale évidente (fig. 5).

Remarques concernant la robustesse des modèles mis au point

Les résultats présentés dans les tableaux 1 et 2 montrent que pour chaque période de prévision, les modèles les mieux classés sur la période de calage ne donnent pas de bons résultats en validation. Ainsi, pour les modèles utilisant les 5 meilleurs vecteurs propres d'index M1 à M5 parmi V1 à V8 on remarque que :

- pour la prévision SO, les 11 meilleurs modèles sur 45 en calage affichent des NSE supérieurs à 0.698 en calage et inférieurs à -0.075 en validation ;
- pour la prévision ASO, les 8 meilleurs modèles sur 45 en calage affichent des NSE supérieurs à 0.721 en calage et inférieurs à 0.033 en validation.

Ceci montre que le bon calage des modèles ne prouve aucunement leur robustesse. Or les modèles satisfaisants en calage (avec NSE supérieur à 0.48, seuil arbitraire) sont très

nombreux : 120 pour la prévision SO et 109 pour la prévision ASO. Vu l'importance de ces effectifs, les chances que certains de ces modèles puissent être validés par pur hasard sont non négligeables. On peut donc légitimement se demander si les performances des modèles sélectionnés pour le temps réel –bons en calage et assez bons en validation-, ne résultent pas simplement du seul hasard. Seule l'utilisation de ces modèles en temps réel pendant plusieurs années permettra de connaître leur validité.

Le problème de robustesse évoqué ci-dessus s'explique en partie par la petitesse des échantillons de calage et de validation (16 et 8 années), comparée au grand nombre de modèles testés (270). Mais il conduit aussi à s'interroger sur la pertinence des modèles qu'on cherche à établir, en reliant des débits à des valeurs de précipitations centrées réduites.

En effet, les vecteurs propres d'index, qui permettent d'agréger en peu de variables l'essentiel de l'information contenue dans les index d'une zone donnée, considèrent tous ces index avec le même poids. En essayant de relier le débit à ces vecteurs propres, on suppose implicitement que le débit à l'exutoire du bassin présente la même sensibilité vis-à-vis des variations relatives de précipitation sur des zones différemment arrosées du bassin. Or dans le cas du Sénégal, il est évident qu'une augmentation d'un écart type des précipitations par rapport à la moyenne aura beaucoup plus d'importance si elle intervient sur la partie sud du bassin, très arrosée, que sur la partie nord, très peu arrosée.

L'idéal serait de pouvoir traduire les index en valeurs de précipitation, avant d'en faire l'adaptation statistique pour la prévision des débits. A défaut, une simple pondération des index basée sur la latitude des mailles pourrait être envisagée, avec un poids décroissant du sud vers le nord pour prendre en compte la diminution assez générale des précipitations observée suivant cette direction en Afrique de l'ouest..

PRÉVISION DES DÉBITS DE CRUE POUR L'ANNÉE 2015

L'utilisation des modèles retenus donne les prévisions suivantes pour l'année 2015 :

- Débit moyen de septembre-octobre prévu à partir des index calculés début août sur la zone L8a8 : 1668 m³/s (probabilité de non dépassement sur les observations de la période 1991-2014 : 0.522)
- Débit moyen de août-septembre-octobre prévu à partir des index calculés début juillet sur la zone B8a8 : 1337 m³/s (probabilité de non dépassement sur les observations de la période 1991-2014 : 0.182)

Ces prévisions sont à prendre avec prudence compte tenu des réserves formulées plus haut sur la robustesse des modèles utilisés.

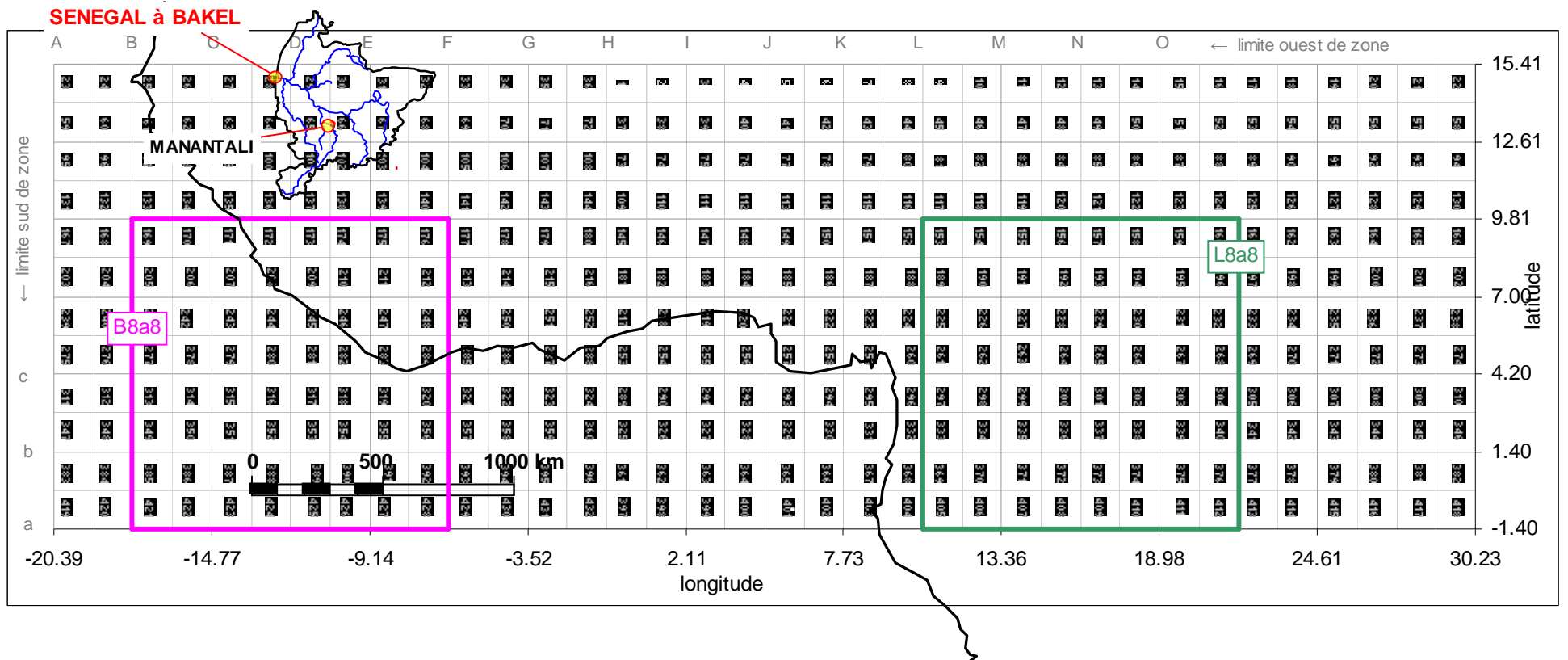


Figure 1 : grille de sortie des résultats d'Arpege Système 5 sur l'Afrique de l'ouest et zones d'index utilisés pour la prévision du débit naturel du

Sénégal

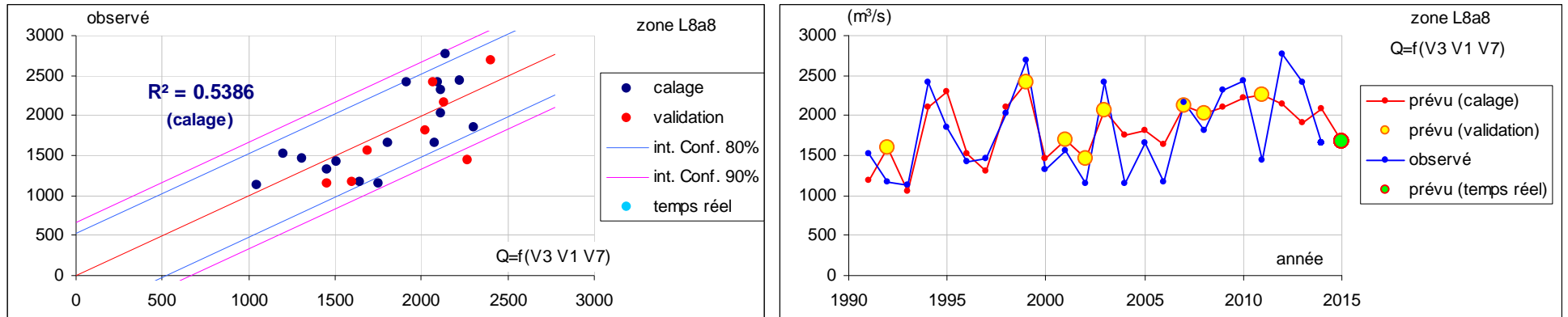


Figure 2 : débits observés et prévus pour la période septembre-octobre, avec le modèle exploitant les index de la zone L8a8

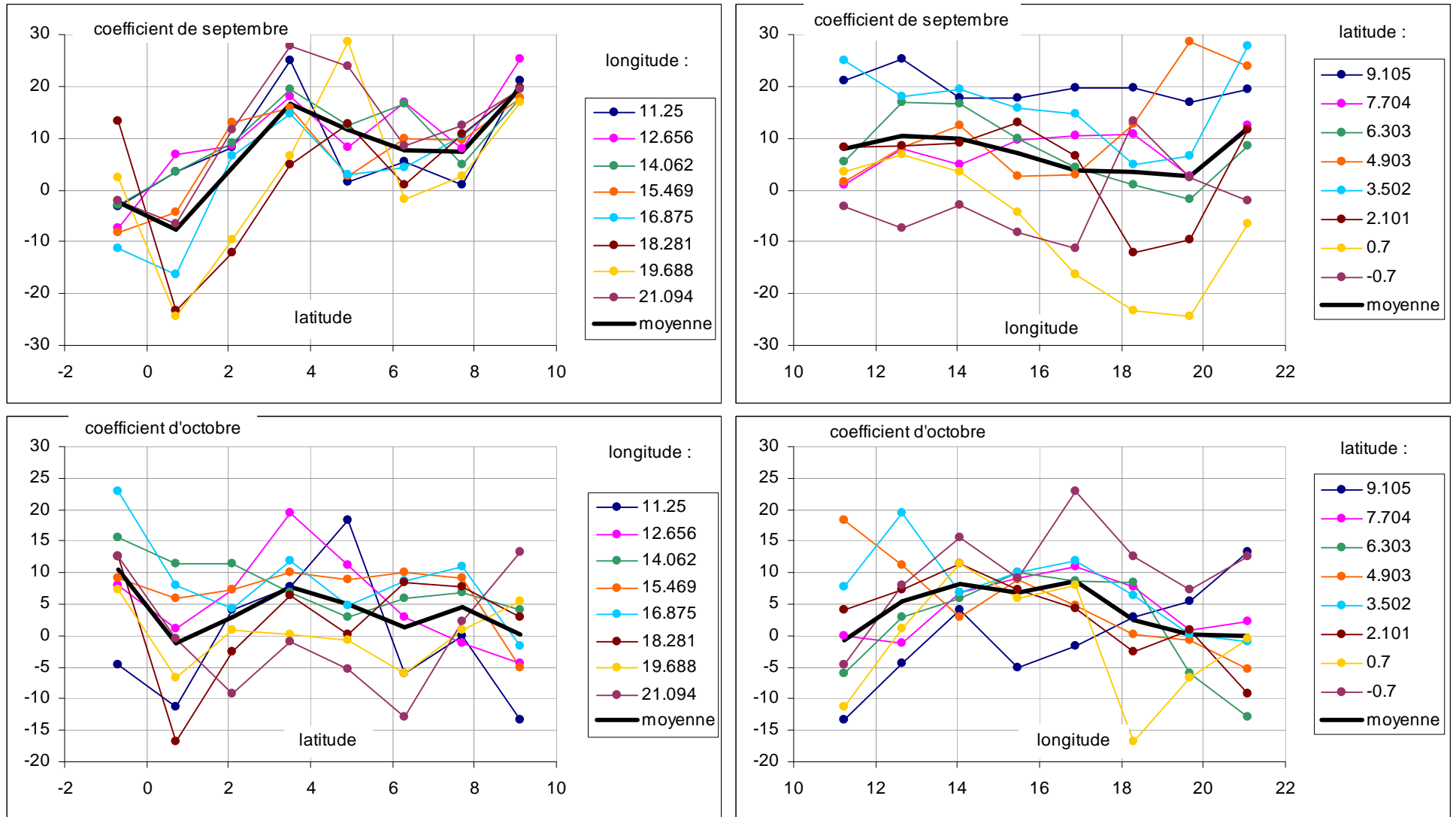


Figure 3 : coefficients multiplicateurs des index pluviométriques de la zone L8a8 dans le modèle prévoyant début août le débit moyen naturel de septembre-octobre

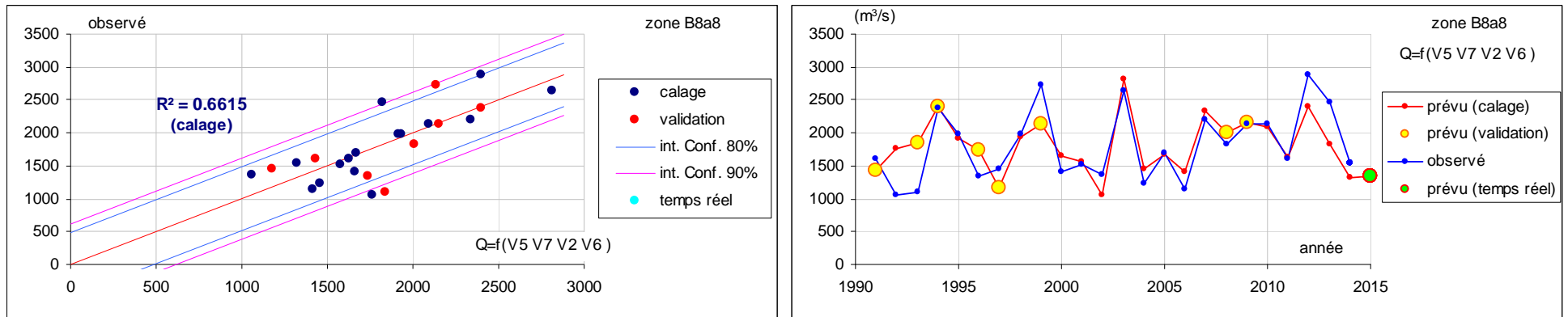


Figure 4 : débits observés et prévus pour la période août-septembre-octobre, avec le modèle exploitant les index de la zone B8a8

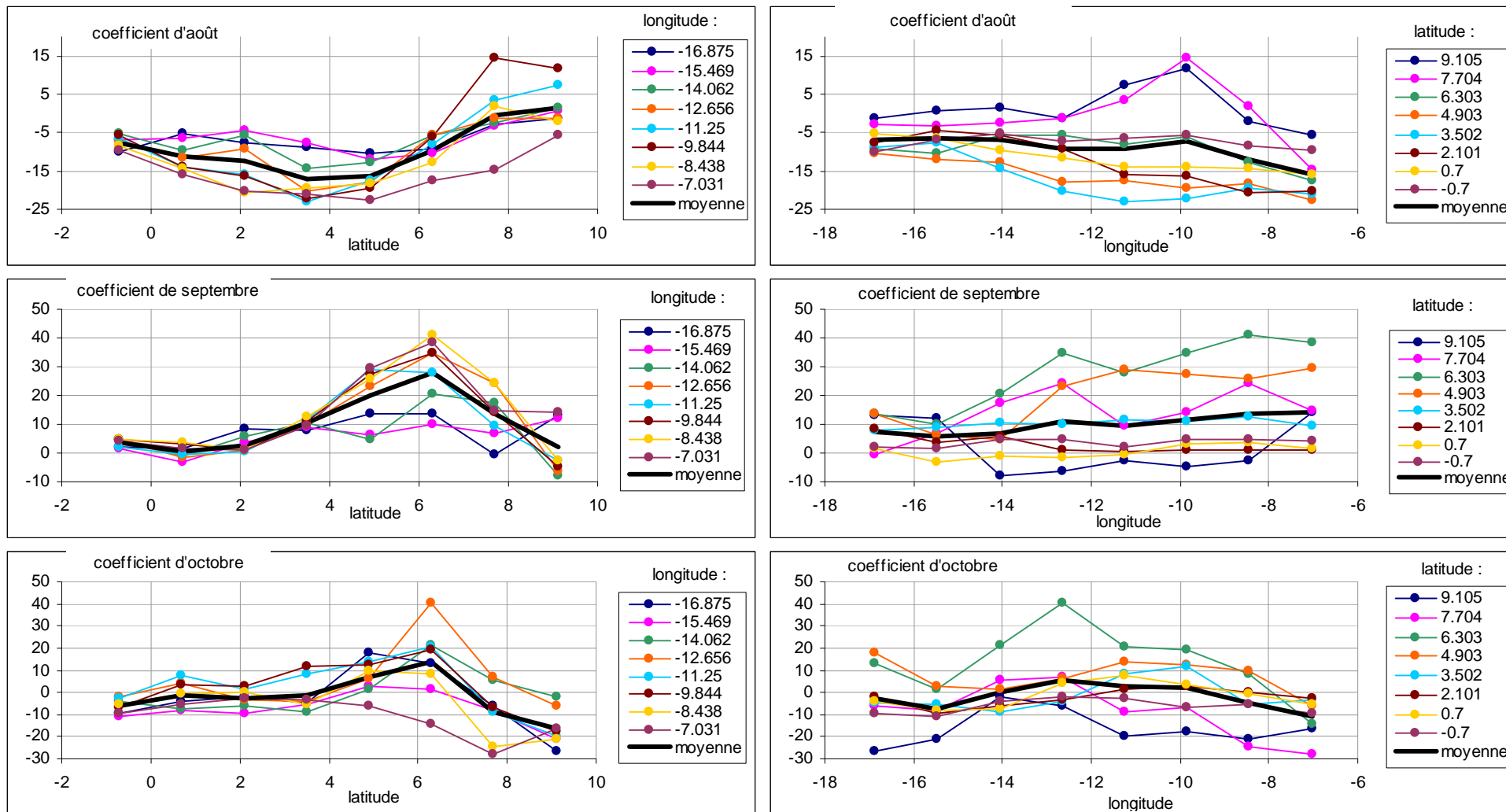


Figure 5 : coefficients multiplicateurs des index pluviométriques de la zone B8a8 dans le modèle prévoyant début juillet le débit moyen naturel de août-septembre-octobre

Tableau 3 : débit naturel du fleuve Sénégal à Bakel sur les périodes septembre-octobre et août-septembre-octobre. Valeurs observées et valeurs prévues (sans surlignage : calage ; jaune : validation ; vert : temps réel) à partir des index produits par Arpege Système 5 début août sur la zone L8a8 (prévision SO) et début juillet sur la zone B8a8 (prévision ASO).

année	septembre-octobre		août-septembre-octobre	
	observé	prévu	observé	prévu
1991	1513	1198	1601	1434
1992	1164	1597	1059	1763
1993	1136	1053	1108	1840
1994	2415	2101	2368	2400
1995	1859	2303	1984	1913
1996	1421	1513	1352	1739
1997	1459	1315	1451	1174
1998	2020	2113	1984	1936
1999	2679	2407	2727	2137
2000	1333	1454	1418	1659
2001	1555	1694	1513	1572
2002	1148	1452	1360	1060
2003	2408	2068	2643	2817
2004	1155	1753	1236	1454
2005	1656	1809	1698	1667
2006	1160	1643	1138	1412
2007	2153	2128	2190	2342
2008	1808	2022	1820	2012
2009	2312	2110	2126	2151
2010	2434	2219	2140	2095
2011	1435	2268	1616	1631
2012	2775	2140	2884	2399
2013	2407	1914	2457	1823
2014	1663	2078	1545	1323
2015		1668		1337