

Estimation des pluies au Sahel : utilisation d'un modèle d'erreur pour évaluer réseaux sol et produits satellitaires

Abdou Ali^{1,2}
Abou Amani¹
Thierry Lebel²

¹ Centre régional de formation et
d'application en agrométéorologie,
hydrologie et météorologie (AGRHYMET),
BP 11011,

Niamey
Niger
<ali@hmg.inpg.fr>
<Abdou.Ali@inpg.fr>

² Institut de recherche pour le développement
(IRD),
Laboratoire d'étude des transferts en
hydrologie et en environnement (LTHE),
UMR 5564,
BP 53
38041 Grenoble cedex 09
France
<thierry.lebel@hmg.inpg.fr>

Résumé

La sécheresse se poursuit au Sahel, les réseaux pluviométriques se sont dégradés et cela a pour effet d'augmenter les erreurs d'estimation au sol. Pour évaluer la capacité des réseaux de mesure opérationnels à fournir des valeurs de référence de la pluie au sol et pour évaluer à quel point sont significatifs ou non les écarts entre pluies au sol et produits satellitaires, on présente ici une approche par modèle d'erreur. Le modèle d'erreur utilisé est étalonné sur une zone test densément instrumentée. Une première application consiste à déterminer des réseaux optimaux (au sens d'un seuil d'erreur admissible) types pour toute la zone sahélienne. On constate que le réseau opérationnel dont les données sont collectées par le Centre Régional AGRHYMET (CRA) est loin d'être optimal en ce sens qu'il ne garantit qu'en de rares endroits une erreur d'estimation inférieure à 10 % de la pluie mensuelle sur des pavés de 1° x 1°. Le modèle d'erreur proposé permet également de mener une inter-comparaison des produits satellitaires CMAP, GPI, GPCC et GPCP sur la zone des pays du CILSS et de montrer que tous ont une distribution plus resserrée que la distribution des valeurs de référence calculées à partir du réseau sol. Les résultats dépendent fortement de la latitude. Le produit GPI semble particulièrement mauvais, alors que le produit CMAP, qui incorpore une quantité importante de données sols, est le meilleur en moyenne.

Mots clés : Pluviométrie, Modélisation, Sahel

Summary

Rainfall estimation in the Sahel: Using an error model to evaluate rain gauge and satellite products

Drought is continuing in the Sahel. Lower rainfall and the decrease of the number of rain gauges are two factors that lead to larger errors in the estimation of rainfall from ground observations. This paper consequently investigates the quality of rainfall estimation from the rain gauge operational network and the differences existing between ground-based rainfall estimates and various satellite-based rainfall estimates. To provide an objective base for this evaluation, an error model approach is presented. The error model is calibrated on a densely instrumented test area. A first application of the model is to study the concept of standard optimal networks, optimal referring here to a given level of average expected error for a given space-time resolution. Taking 1° x 1° monthly resolutions as the lowest acceptable values for many water resources and agronomical applications, it is shown that the AGRHYMET Regional Center (CRA) operational network is far from being optimal, since it guarantees in rare places only an average estimation error lower than 10%. The proposed error model also makes it possible to carry out an intercomparison of the most commonly used satellite products, that is CMAP, GPI, GPCC, and GPCP. All these products have a distribution of their estimated rainfall values which is significantly less dispersed than the distribution of the CRA ground-based reference values (computed at the same resolution as the satellite rainfall products). The differences

between the reference values and the satellite products and the differences between the various satellite products are strongly dependent on the latitude. The performance of the GPI product is the worst, whereas the CMAP product, which in fact includes a large quantity of rain gauge data, is the best on average.

Keywords: Rainfall, Modelling, Sahel

Les difficultés rencontrées pour estimer correctement les précipitations au Sahel font l'objet depuis quelques années d'une attention renouvelée. La dégradation des réseaux (figure 1), la difficulté pour accéder aux données, les préoccupations croissantes sur une possible modification durable du régime pluviométrique [1], les performances encore modestes des algorithmes [2, 3] sont autant de facteurs qui ont motivé des études récentes sur le sujet [4]. La contribution présentée ici concerne le développement d'une méthode objective de calcul d'erreurs d'estimation, méthode qui est appliquée, d'une part, à l'évaluation du potentiel des réseaux sol actuels pour estimer les pluies au Sahel et, d'autre part, à l'évaluation de produits pluviométriques élaborés à partir de données satellitaires. Par-delà la généralité de l'approche proposée, deux points d'ordre méthodologique sont discutés plus en détail. D'une part, plus la pluie diminue et devient irrégulière, plus les erreurs d'estimation augmentent. Il faut donc paradoxalement plus de stations pour une estimation de qualité comparable au nord du Sahel qu'au sud, alors que sur le terrain, les réseaux sont plus denses au sud qu'au nord. D'autre part, on doit se poser la question de ce que vaut la précision réelle d'un produit comparé à une valeur de référence elle-même entachée d'erreur.

Données utilisées et conditions actuelles de la climatologie au Sahel

Données

Les données utilisées couvrent deux échelles différentes. Au niveau local, nous utilisons la base de données de l'expérience EPSAT-Niger (Estimation des précipitations par Satellite) qui a débuté en 1990 dans la région de Niamey au Niger et qui se poursuit encore [5]. Le réseau de pluviographes est composé de 30 appareils à acquisition numérique couvrant une superficie de 16 000 km², ce qui permet d'estimer la vérité sol avec des erreurs très faibles, même au pas de temps journalier [6]. Au niveau régional, on a accès aux données du réseau pluviométrique du Centre régional AGRHYMET (CRA), à Niamey, qui regroupe les données des réseaux opérationnels des pays du Comité inter-États de lutte contre la sécheresse au Sahel (CILSS). Ce réseau couvre environ trois millions de km² et est constitué de près de 1 200 stations, mais 750 seulement en moyenne font parvenir leurs données au CRA chaque année. Nous reviendrons sur l'évolution temporelle et la variabilité spatiale de ce réseau. Notons simplement que chacun de ces deux réseaux représente à son échelle, le

meilleur de la région. La figure 2 montre leur configuration spatiale.

Conditions actuelles de la climatologie par rapport aux décennies précédentes

Après deux décennies particulièrement sèches [1], la période 1990-2002 n'a pas vu le retour aux conditions humides des années 1950-1970 — mais des différences notables avec la période précédente sont tout de même à signaler. Tout d'abord, on a observé deux années très pluvieuses (1994 et 1999) — proches des records observés et une année bien arrosée (1998). Ensuite, comme le montre la figure 3, une opposition est apparue entre le Sahel ouest — qui reste très sec — et le Sahel est qui a vu, pour ce qui est de la moyenne interannuelle, un retour à une meilleure pluviométrie.

Modèle d'erreur et calcul de valeurs pluviométriques de référence

Qu'est-ce qu'une valeur de référence ?

La problématique de l'estimation quantitative des précipitations est bien cernée. Elle a fait l'objet de nombreux travaux au cours des 20 dernières années et on pourra notamment se référer aux synthèses sur :

- les méthodes disponibles [7] ;
- le statut des satellites dédiés à l'estimation des pluies [8] ;
- les perspectives qui s'offrent aujourd'hui [9].

On pourra aussi se reporter au travail de Ramge [3] qui aborde ces différents points. De façon générale, on dispose, d'une part, de données pluviométriques ponctuelles — considérées comme étant de « vraies » valeurs de pluie, après évaluation d'un certain nombre d'erreurs possibles [10, 11] ; d'autre part, on dispose de mesures obtenues par télédétection, reliées à la pluie de manière plus ou moins directe. L'utilisateur est, lui, intéressé par une valeur moyenne sur un domaine, valeur que ne fournissent directement ni les réseaux de pluviomètres, ni les mesures par télédétection. Pour obtenir une valeur

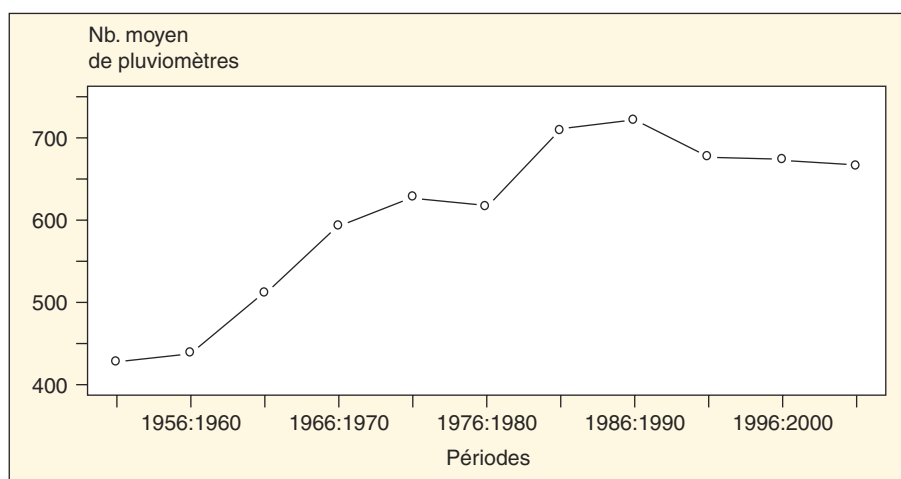


Figure 1. Évolution du nombre moyen de stations pluviométriques (période de cinq ans) de la base de données du Centre AGRHYMET.

Réseau mensuel (pour qu'une station soit prise en compte il faut qu'il y ait au minimum un cumul mensuel non manquant). La tendance est à la baisse à partir des années 1985.

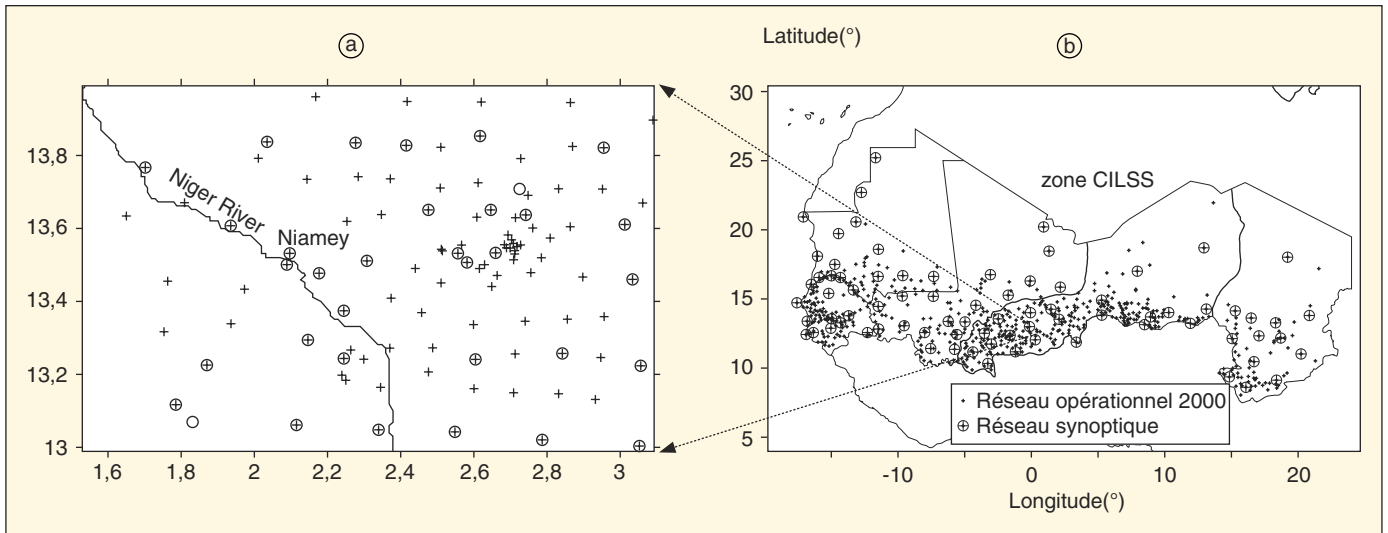


Figure 2. Configuration spatiale des réseaux. a) échelle locale : réseau Epsat-Niger (coordonnées en km avec origine à 2° E-13° N). Le réseau dense (points) comprend 107 pluviographes à acquisition numérique et le réseau de base (cercles) en compte 30. b) échelle régionale : réseau AGRHYMET.

Le réseau opérationnel pour l'année 2000 comprend 680 stations pluviométriques et le réseau synoptique en compte 87.

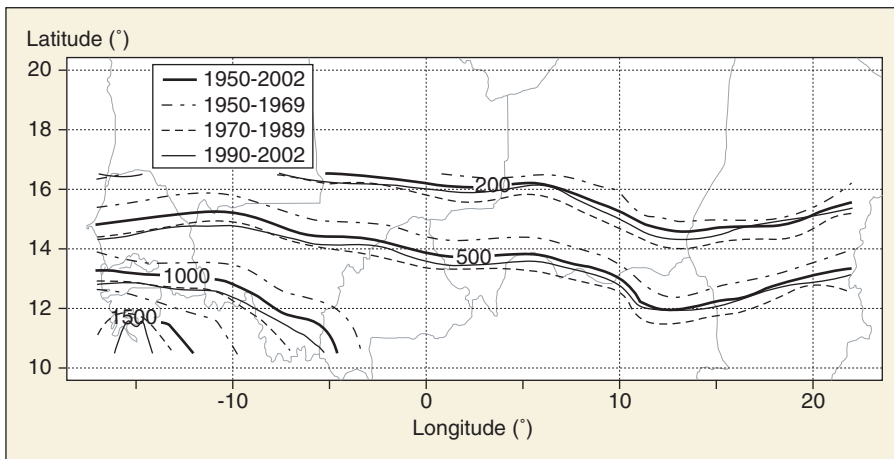


Figure 3. Comparaison des isohyètes moyennes (200 mm, 500 mm, 1 000 mm) pour les périodes humide (1950-1969), sèche (1970-1989) et récente (1990-2002).

On note l'apparition de deux comportements différents entre le Sahel est et le Sahel ouest. Sur le Sahel ouest, la dernière décennie est très proche de 1970-1989, alors que sur le Sahel est, elle est souvent plus humide que la période 1970-1989, avec des valeurs proches de la moyenne long terme (1950-2002). Entre la période humide 1950-1969 et la période sèche 1970-1989, on assiste à une descente d'environ 1° en latitude des isohyètes vers le sud. Le décalage est assez important sur l'ensemble de la zone. En revanche, à compter du début des années 1990, on note une remontée des isohyètes de la partie est par rapport à la partie sud.

moyenne sur un domaine à partir d'un réseau de pluviomètres, il faut interpolier des valeurs ponctuelles, c'est-à-dire reconstituer de l'information manquante, ce qui implique des erreurs. Pour faire de même à partir de données satellitales, il faut convertir une réflectance ou une température de brillance en pluie, via des modèles plus ou moins complexes dont l'utilisation entraîne également des erreurs, parfois importantes. Dans un tel contexte une question devient centrale pour départager les estimations entre elles : peut-on disposer de chiffres qui, à défaut d'être exacts, soient utilisables comme valeurs de référence ?

Lorsque l'on s'intéresse à des échelles relativement grandes, dans le temps comme dans l'espace, les réseaux nationaux de pluviomètres et pluviographes sont généralement considérés comme fournissant de telles valeurs de référence, notamment dans les pays dotés de bonnes infrastructures météorologiques. Ces valeurs de référence restent malgré tout des valeurs approchées et il faut évaluer leur degré de précision, qui va dépendre de :

- la nature des champs de pluie à interpoler ;
- la densité du réseau disponible ;
- la mise en œuvre d'une méthodologie adéquate d'interpolation.

Qu'en est-il de ces trois facteurs dans le cas du Sahel ?

• Nature des champs de pluie

Deux points sont à souligner :

1. Les systèmes pluvieux sahéliens, étant essentiellement de nature convective, présentent une grande intermittence spatiale qui complique l'interpolation ;
2. Les champs pluvieux présentent des dérives importantes (non-stationnarité dans l'espace de la moyenne et de l'écart type) dès lors que l'on travaille sur des durées de cumul supérieures à la journée : Ali *et al.* [12] ont montré que cette dérive était avant tout liée à la probabilité d'occurrence de la pluie, c'est-à-dire au nombre moyen d'événements pluvieux observés en un lieu donné pour une période donnée. Cette non-stationnarité est associée à une anisotropie, qui implique que la fonction d'interpolation utilisée ne sera pas la même dans toutes les directions (en pratique on utilise une fonction unique avec une métrique non euclidienne). Il résulte de tout cela que l'erreur d'estimation relative décroît fortement lorsque le nombre d'événements augmente (les utilisateurs savent empiriquement qu'on commet moins d'erreur pour interpoler un champ de pluie annuel que pour interpoler un champ décadaire).

• Densité du réseau

La densité du réseau sahélien s'est dégradée après un optimum atteint dans les années 1982-1983. Par ailleurs, comme le montre la figure 4, le réseau est hétérogène dans l'espace. Le nombre de stations par maille de 2,5° x 2,5° est en moyenne de 9,6 – soit une densité

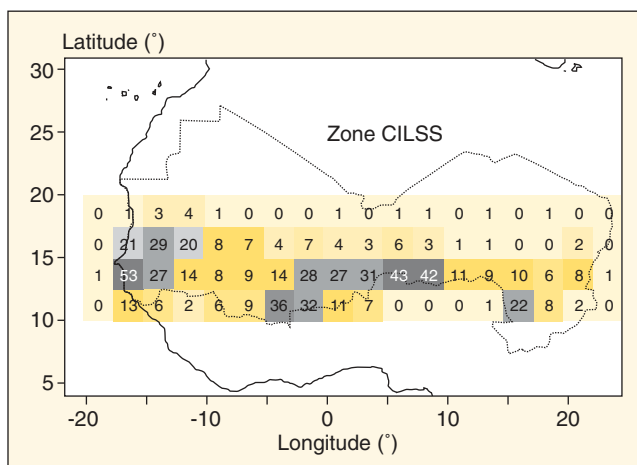


Figure 4. Exemple de variabilité spatiale du réseau correspondant à l'année 2002.

Les chiffres indiquent le nombre de stations par maille de 2,5° x 2,5°. La moyenne est de 9,6 stations par maille soit une densité moyenne de 1,5 station par degré carré, mais avec des écarts très importants d'une maille à l'autre.

moyenne de 1,5 station par 10 000 km² avec un coefficient de variation de 1,31. Localement en revanche, il existe des données haute résolution, comme c'est le cas sur le degré carré de Niamey (jusqu'à 107 pluviographes à acquisition numérique entre 1990 et 1993).

• Choix des méthodes d'interpolation

Seules les méthodes d'interpolation construites dans un cadre stochastique permettent de calculer conjointement une pluie moyenne sur une surface et l'erreur probable associée [13, 14]. Le mot « probable » a ici toute son importance. En effet c'est le respect, par le phénomène étudié, des hypothèses du modèle statistique utilisé pour l'interpolation qui garantit que l'espérance mathématique de l'erreur sera correctement évaluée. Par exemple, dans le cas de la géostatistique, l'utilisation d'un variogramme inadapté peut conduire à de très mauvaises évaluations des erreurs (voir ci-après). Le fait de disposer d'un réseau extrêmement dense sur le degré carré de Niamey a permis d'étudier en détail ces questions pour les pluies sahéliennes, de valider une fonction de structure adaptée et de construire des valeurs de référence dotées d'un écart type d'estimation qui en mesure la qualité [6].

Modèle d'erreur

Commençons par montrer, à l'aide d'un exemple simple, à quel degré les évaluations d'erreur peuvent être incorrectes si une mauvaise représentation de la structure des champs de pluie est utilisée. L'exercice a consisté à étudier les erreurs commises en reconstituant des valeurs connues par une procédure de validation croisée : on recalcule en chaque point x_i d'un

réseau comprenant n points de mesure, une valeur interpolée $Z^*(x_i)$, à partir des $n-1$ autres observations. On peut alors calculer l'écart, e_i , entre la valeur observée, $Z(x_i)$, et la valeur reconstituée $Z^*(x_i)$. Ce calcul est fait en utilisant deux interpolateurs différents : l'un reposant sur une fonction de structure optimisée à partir des données EPSAT-Niger, et l'autre sur une fonction de structure ne prenant en compte ni l'anisotropie ni la dérive constatée sur les observations EPSAT-Niger.

$$e_i = Z(x_i) - Z^*(x_i) \quad (1)$$

Si la fonction de structure du processus a été bien identifiée, l'erreur quadratique moyenne¹ :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2)$$

¹ RMSE (root mean square error) : écart quadratique moyen des erreurs.

Tableau 1. Comparaison entre erreurs théoriques pour deux méthodes de krigeage et erreurs expérimentales déduites d'une procédure de validation croisée (moyennes sur 8 années).

Pas de temps	Krigeage « optimal »			Krigeage « non optimal »		
	RMSE (mm)	Théorique (mm)	I	RMSE (mm)	Théorique (mm)	I
Journalier	10,1	9,6	1,01	11,08	7,75	1,78
Pentade	20,3	20,6	0,94	20,46	19,4	1,2
Décade	27,82	26,98	1,01	29,38	18,92	3,1
Mois	53,92	50,93	1,1	54,14	32,68	2,69
Année	99,4	106,75	0,86	99,66	77,06	1,73

RMSE (root mean square error) : écart quadratique moyen des erreurs ; I : Moyenne quadratique de l'écart type de krigeage normé par l'écart expérimental entre valeur estimée et valeur mesurée.

doit être proche de la moyenne des écarts types de krigeage, notés σ_i . La similitude entre la série des $\{e_i\}$ et celle des $\{\sigma_i\}$, est mesurée par la statistique I suivante :

$$I = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\sigma_i}{e_i} \right)^2} \quad (3)$$

Le tableau 1 montre que les écarts empiriques (RMSE) ne sont pas nécessairement très différents entre krigeage « optimal » et krigeage « non optimal », mais que, en revanche, la méthode optimisée fournit des écarts théoriques bien plus réalistes, la statistique I étant très proche de 1 pour tous les cas. La méthode non optimisée prévoit des erreurs très sous-évaluées.

La fonction de structure utilisée pour le krigeage « optimal » dont on a donné les résultats dans le tableau 1, conduit au calcul d'un modèle d'erreur qui a la forme suivante [6] :

$$E_k(A, N, K, M_k) = \frac{1.05}{\sqrt{N} \sqrt{K}} \left(\frac{M_k}{K} \right)^{-0.2} \quad (4)$$

$$\left[0.283 + 0.168 \log \left(\frac{A}{N} \right) \right]$$

où $E_k(C)$ est l'erreur relative commise sur l'estimation d'un cumul M_k (en mm) de K événements, pour un réseau C de N stations sur une maille de surface A (en km²).

L'équation (4) fait explicitement apparaître les trois facteurs dont dépend l'erreur d'estimation, à savoir : la taille de la zone pour laquelle on cherche à estimer la pluie moyenne spatiale, le pas de temps (via le nombre d'événements pluvieux K) et le nombre de stations disponibles. Pour appliquer ce modèle sur l'ensemble du Sahel, sans données pluviographiques permettant de déterminer le nombre d'événements pluvieux, on a assimilé l'événement pluvieux au jour pluvieux. Les tra-

vaut récents de Balme² montrent que cette approximation est acceptable.

Réseau optimal pour une résolution de 1° x 1°

La notion de réseau optimal est contingente d'une échelle de travail (temps K et espace A). Le critère d'optimalité correspond à un seuil d'erreur que l'on se fixe. Pour une erreur fixée, l'inversion de la relation (4) permet de déterminer le nombre optimal N^* de stations. On va considérer ici un seuil d'erreur E de 10 % sur les cumuls mensuels, pour une résolution de 1° x 1° (soit $A = 12\ 000\ \text{km}^2$). Cette résolution est le minimum requis pour obtenir des résultats qui aient un sens avec des modèles hydrologiques ou de bilan hydrique (pour des résolutions plus grossières, du type de 2,5° x 2,5° examiné plus loin, la modélisation hydrologique n'a plus de sens, vu l'hétérogénéité climatique et écologique internes à la maille). Le calcul de N^* a été effectué pour chaque degré carré de la zone CILSS et pour deux périodes : la période humide - 1950-1969 - et la période sèche - 1970-2000. Pour obtenir des valeurs s'appliquant à l'ensemble d'une période, on a calculé des valeurs moyennes de K sur cette période et pour les trois mois de juillet-août-septembre, que l'on notera \bar{K} . En période humide, on a $\bar{K} = 12^3$ sur l'ensemble de la zone ($\bar{K} = 16$ à 11° N et $\bar{K} = 10$ à 17° N). En période sèche, on a $\bar{K} = 10$ sur l'ensemble de la zone ($\bar{K} = 13$ à 11° N et $\bar{K} = 7$ à 17° N). Les deux cartes ainsi obtenues sont reproduites dans les figure 5a et b. On voit que la baisse de la pluie en période sèche implique une augmentation du nombre de postes nécessaires de 1 à 3 unités par degré carré selon que l'on se trouve au sud ou au nord de la zone (plus la latitude augmente, plus le nombre de stations doit augmenter pour garantir le même seuil d'erreur). Au total, un réseau avec exactement le bon nombre de postes par degré carré devrait actuellement comporter 1 356 pluviomètres (contre 824 lors de la période humide), cela afin de garantir partout une erreur d'estimation des cumuls mensuels (JAS) infé-

² Communication personnelle.

³ \bar{K} est une valeur mensuelle que l'on considère identique sur les trois mois concernés alors qu'en réalité elle ne l'est pas (par exemple, \bar{K} est plus faible en septembre qu'en août). Les cartes de la figure 5 sont donc une moyenne sur 3 mois de cartes mensuelles (les erreurs sont plus fortes pour septembre que pour août).

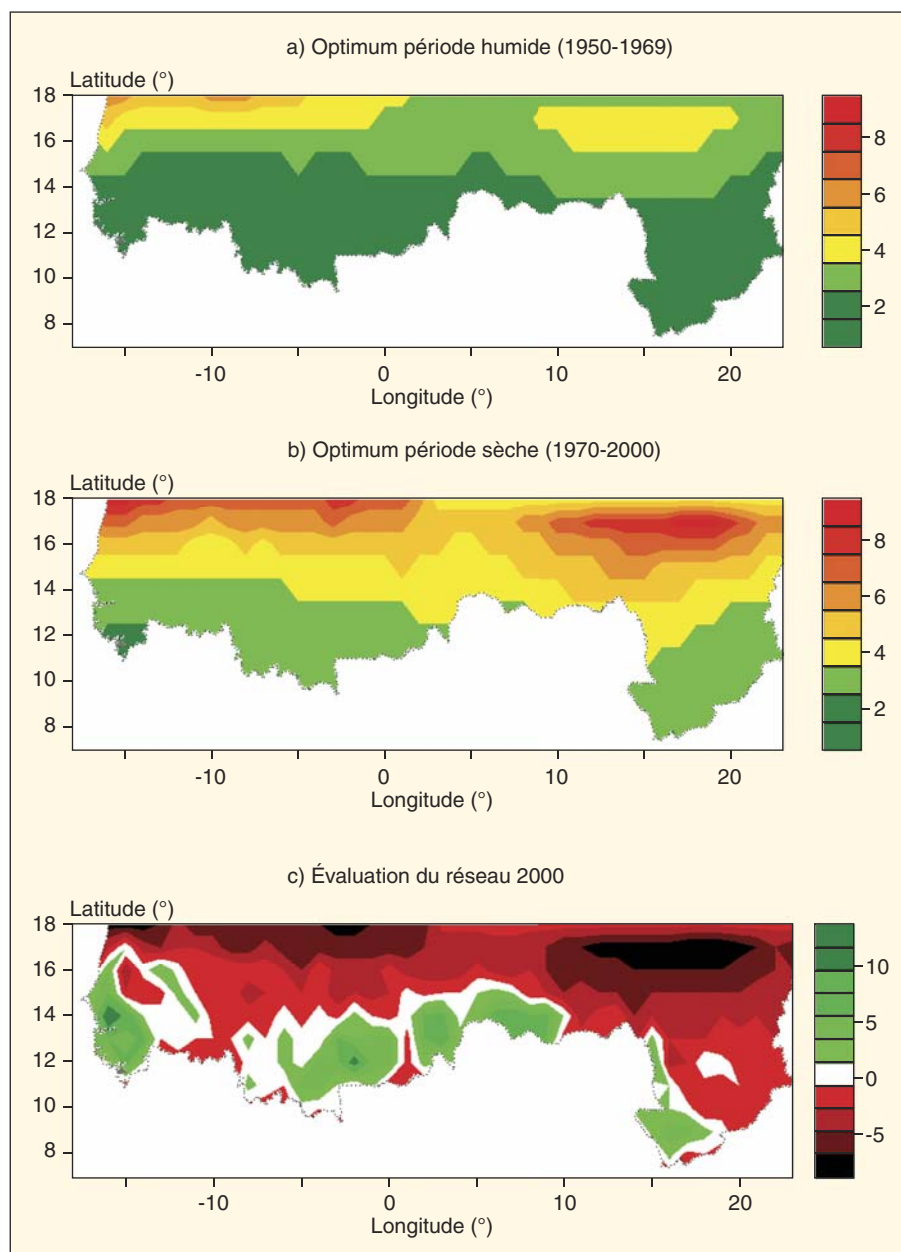


Figure 5. Réseaux optimaux au seuil de 10 % d'erreur sur une grille de 1°x1° pour le cumul mensuel de la période de juillet, août, septembre (JAS) et évaluation du réseau opérationnel de l'année 2000. a) le réseau optimal climatologique de la période humide (1950-1969) nécessite moins de stations que b) celui de la période sèche (1970-2000). La partie nord en demande plus que la partie sud. c) surplus (chiffres positifs) ou déficit (chiffres négatifs) du réseau 2000 par rapport au réseau optimal.

On constate que sur la majeure partie de la zone d'étude, il n'y a pas assez de stations pour garantir une estimation de la pluie mensuelle avec moins de 10 % d'erreur.

rieurs à 10 %. On a également représenté en figure 5c une évaluation du réseau de l'année 2000 par rapport au critère d'optimalité fixé. On voit que ce réseau, qui compte un total de 680 stations, est presque partout déficitaire. Au seuil de 20 % et pour un mois, il faut 6 stations sur un degré carré situé à 17° N, alors qu'il ne faut que 1 station à 10° N (valeur moyenne pour l'ensemble de la période 1950-2000). On peut éga-

lement retenir comme ordre de grandeur qu'en moyenne un réseau garantissant 20 % d'erreur à l'échelle du mois, fournit des estimations avec des erreurs moyennes de 30 % à l'échelle de la décennie et de 10 % à celle de l'année. D'un point de vue agronomique ou hydrologique, le critère d'optimalité à viser serait plutôt moins de 10 % d'erreur à l'échelle décadaire sur 1° x 1°. Le nombre de stations requises serait alors de 3 074. Pour respecter ce

critère sur la zone située au sud de 15° N seulement, il faudrait 1 725 stations.

Évaluation de quelques produits pluviométriques à la résolution 2,5° x 2,5°

La plupart des produits pluviométriques mensuels incorporant de l'information satellitaire sont calculés sur des grilles de résolution 2,5° x 2,5°. On a donc recalculé, avec la même résolution de 2,5° x 2,5°, des grilles pluviométriques mensuelles à partir des données sol archivées au CRA, pour pouvoir les comparer ensuite aux autres produits pluviométriques.

Produit pluviométrique du CRA

Pour créer ce produit on a utilisé le krigeage « optimal » dont les erreurs théoriques ont été validées pour la résolution 1° x 1° à la section précédente (tableau I). La zone considérée s'étend de 17,5° O à 17,5° E en longitude et comprend trois bandes de 2,5° : la bande sud, entre 10° N et 12,5° N (centre : 11,25°), la bande centrale, entre 12,5° N et 15° N (centre : 13,75°), la bande nord, entre 15° N et 17,5° N (centre : 16,25°). Logiquement, les erreurs moyennes du produit sol CRA sont plus faibles à cette résolution de 2,5° x 2,5° qu'elles n'étaient pour la résolution de 1° x 1°. Ainsi, les erreurs relatives sont de l'ordre de 34 % en moyenne sur le nord du Sahel (contre 52 % pour 1° x 1°), de 8 % au centre (contre 11 % pour 1° x 1°) et de 10 % sur le sud (contre 15 % pour 1° x 1°). En d'autres termes, le réseau CRA est optimal, au sens où il garantit une erreur inférieure à 10 %, pour l'estimation des cumuls mensuels en dessous de 15° N à la résolution 2,5° x 2,5°. On a va donc considérer que l'on dispose là de valeurs de référence pour évaluer les produits satellitaires (sauf sur les mailles de la bande nord).

Évaluation de quelques produits sur la zone CILSS

Les produits évalués sont les produits mensuels CMAP, GPI, GPCC et GPCP. Pour une représentation de ces produits, on peut se référer à différentes publications [15-18]. Notons que si, par commodité, tous ces produits sont ici qualifiés de « satellitaires », ils incluent tous, sauf GPI, des données sol. Dans un premier temps, on opère une évaluation globale en comparant les distributions d'ensemble des valeurs de chaque produit. La taille des échantillons est de 1 260, correspondant à 42 mailles pour les 3 mois sur 10 ans.

Ces distributions sont comparées en figure 6 sous deux formes. En figure 6a, on a une représentation synthétique sous forme de *box plots*. Ces *box plots* sont assez semblables, sauf pour GPI, dont le quartile inférieur (quantile 25 %) est plus grand que les médianes des trois autres produits et de la référence CRA. La dispersion interquartile des produits GPCC et GPCP est plus faible que celle de CMAP et de CRA, qui sont très proches l'une de l'autre. La moyenne GPCC (143,40 mm) et de CMAP (144 mm) sont les plus proches de celle du CRA (138 mm), celle de GPCP étant légèrement supérieure (147 mm pour GPCP) ; la moyenne GPI est beaucoup trop forte (200,8 mm). En ce qui concerne les valeurs extrêmes (moyennes mensuelles supérieures à 320 mm, qui est la valeur du quantile CRA 95 % sur une maille de 2,5° x 2,5°), CRA, CMAP, GPCC et GPCP sont assez proches avec respectivement 5 %, 4 %, 3,73 % et 3,33 % de valeurs supérieures à 320 mm. GPI continue à être biaisé vers le haut avec 11 % de valeurs supérieures à 320 mm. Globalement, les distributions de CMAP, GPCC et GPCP sont donc très proches de la référence CRA, ce qui peut être attribué au fait qu'elles incluent des données sol. La figure 6b montre toutefois une plus forte accumulation de valeurs centrales pour ces trois produits, dont la distribution est nettement décalée vers la droite par rapport à celle du CRA. Les distributions des produits satellitaires sont plus resserrées, avec un décalage particulièrement net dans la partie basse des distributions, où seule une faible proportion des valeurs CMAP (17 %), GPCC (13 %), GPCP (13 %) et GPI (6 %) sont inférieures au quartile 25 % du CRA. En réalité, ce décalage provient principalement du comportement particulier de la distribution des valeurs sur la bande nord, valeurs dont on sait qu'elles sont estimées avec peu de précision par CRA (34 % de RMSE). On remarquera aussi que sur cette bande nord, la distribution GPI est très irréaliste, ce qui s'explique assez bien par le fait que les nuages sont moins efficaces à ces latitudes et que la relation standard GPI devrait donc être corrigée pour éviter de fortes surestimations. On peut résumer ce décalage en disant que les produits satellitaires sous-estiment fortement l'effet de l'intermittence des pluies qui implique une probabilité non négligeable de faibles valeurs – y compris à l'échelle mensuelle – sur de grandes mailles. Sur les bandes sud et centrale, CRA, CMAP, GPCC et GPCP ont des modes et des médianes très proches et les probabilités de non-dépassement des quartiles CRA sont voisines de la valeur théorique (tableau III). Sous l'angle des RMSE, la qualité des produits est égale-

ment fonction de la latitude. Les erreurs (RMSE) sont minimums à 13,75° : 16 % pour CMAP, 17 % pour GPCC, 21 % pour GPCP et 53 % pour GPI, contre respectivement 24 %, 22 %, 23 % et 44 % à 11,25°. À 16,25° elles sont trop fortes : 55 % pour CMAP, 69 % pour GPCC et 80 % pour GPCP, avec une valeur complètement irréaliste pour GPI (130 %). En moyenne sur l'ensemble de la zone d'étude, CMAP apparaît comme étant le meilleur produit avec 2 % et 25 % comme biais et RMSE, suivi par GPCC (4 % et 27 %), GPCP (5 %, 28 %) et GPI (43 % et 55 %).

Conclusion

Un modèle d'erreur d'utilisation simple a été proposé pour évaluer les réseaux de mesure au sol et des produits pluviométriques disponibles sur le Sahel utilisés pour des études climatiques ou de validation de modèles. On constate que la baisse combinée de la pluviométrie (par rapport aux années 1950 et 1960) et de la densité des réseaux opérationnels (par rapport au début des années 1980) a eu pour effet d'augmenter les erreurs d'estimation de la pluie. Le réseau pluviométrique opérationnel des pays du CILSS ne permet qu'en de rares endroits de garantir une erreur de moins de 10 % sur l'estimation des pluies mensuelles de la période juillet-septembre, lorsque l'on travaille à une résolution de 1° x 1°. On est donc très loin d'un réseau optimal pour les applications hydrologiques ou de gestion des ressources en eau si l'on considère qu'une taille de maille de 1° x 1° constitue un seuil au-delà duquel il n'y a plus de sens à faire fonctionner des modèles hydrologiques ou de bilan hydrique. A une résolution plus grossière de 2,5° x 2,5°, acceptable pour les applications climatologiques, les erreurs relatives sont de l'ordre de 34 % en moyenne sur le nord du Sahel, de 8 % au centre et de 10 % sur le sud. Ces valeurs sont à comparer avec les erreurs associées aux produits satellitaires. Par exemple, le produit CMAP, qui est en moyenne le meilleur produit pour la période juillet-août-septembre (tant sur le plan du biais que sur celui du RMSE ou des distributions), a un RMSE deux fois plus grand que le RMSE du produit sol CRA : 25 % contre 10 % à 11,25°, 16 % contre 8 % à 13,75° et 55 % contre 34 % à 16,25°. La mauvaise performance des produits au nord du Sahel s'explique par la conjugaison d'une mauvaise qualité des estimations sol et aussi par le fait que les nuages y sont moins pluvio-gènes. On peut légitimement espérer que les estimations satellitaires s'amélioreront dans le futur, surtout

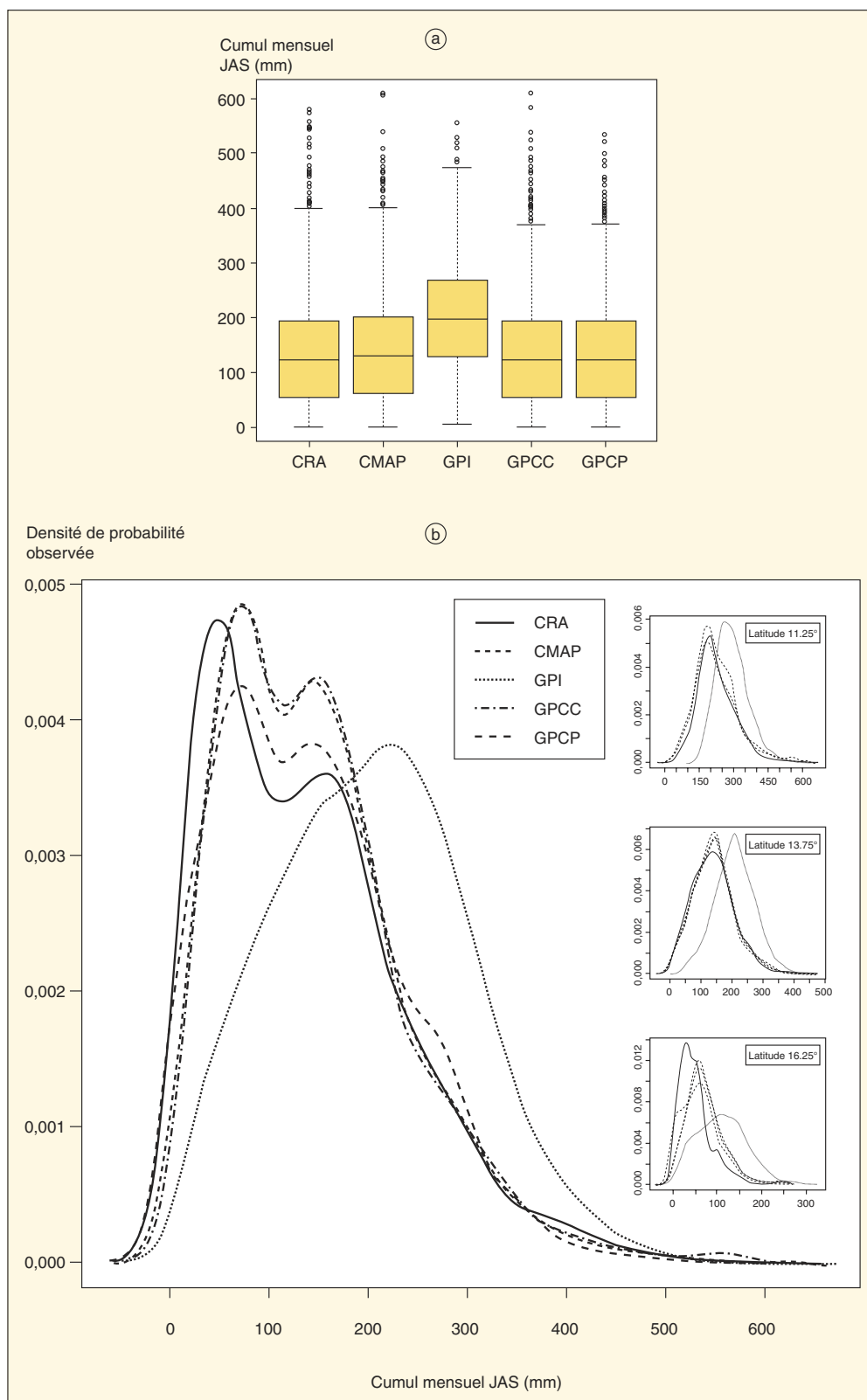


Figure 6. a) les boîtes à moustaches (box plot) montrent qu'en dehors de GPI, les quantiles 25 %, 50 % et 75 % sont très proches pour les différents produits. L'intervalle interquartile 25 %-75 % du CMAP est le plus proche de celui du CRA, ceux des GPCC et GPCP sont moins étalés. Ces boîtes à moustaches montrent également que GPI rate les fortes valeurs. b) les fonctions de densité expérimentales montrent des différences en fonction de la latitude : les plus faibles valeurs sont sous-estimées et les valeurs médianes surestimées.

L'analyse latitudinale montre également que le décalage entre les distributions globales sur la zone est principalement dû au comportement irréaliste dans la zone nord du Sahel.

Tableau II. Évaluation de quatre produits pluviométriques mensuels sur la période 1990-1999.

Latitudes		Quantiles de la distribution des valeurs calculées à partir du réseau sol archivé au CRA				
		Min	Quantile 25 %	Médiane	Quantile 75 %	Max
11,25° 14 mailles sur 3 mois et 10 ans : 420 valeurs	CMAF	0	20	45	76	100
	GPI	0	1	11	46	100
	GPCC	0	30	56	80	100
	GPCP	0	28	53	79	100
13,75° 14 mailles sur 3 mois et 10 ans : 420 valeurs	CMAF	0	23	50	76	100
	GPI	0	4	12	28	99
	GPCC	0	21	48	75	100
	GPCP	0	22	45	72	100
16,25° 14 mailles sur 3 mois et 10 ans : 420 valeurs	CMAF	4	23	36	53	99
	GPI	0	5	15	24	90
	GPCC	0	10	25	45	98
	GPCP	0	10	23	46	97

Les chiffres correspondent à la probabilité de non-dépassement du quantile CRA (Centre régional AGRHY-MET), pour le produit concerné (exemple : pour la latitude 13,75°, 28 % seulement des valeurs de GPI sont inférieures au quantile 75 % de la distribution CRA, ce qui indique un fort décalage vers la droite de la distribution GPI que l'on constate bien sur la figure 6 ; de même à 16,25°, 4 % des valeurs CMAF sont inférieures au minimum de la série CRA, ce qui indique un décalage à gauche pour les faibles valeurs, alors que pour les autres quantiles, le décalage est à droite : 36 % seulement de valeurs CMAF sont inférieures à la médiane CRA et 53 % seulement sont inférieures au quantile 75).

aux échelles examinées dans cet article, du fait de la combinaison de nouveaux capteurs et missions, du type MSG et GPM. Néanmoins, il sera toujours nécessaire de recourir à la notion de réseaux optimaux – l'optimalité étant une notion relative aux objectifs de précision que l'on se fixe et aux échelles de travail. Le réseau optimum peut constituer un objectif à atteindre pour couvrir des besoins opérationnels que les satellites ont du mal à satisfaire. Il permet également de dimensionner des systèmes locaux de validation, du type du réseau EPSAT-Niger, utilisé pour étalonner les algorithmes d'estimation et fournir des valeurs de référence aux exercices de validation. La méthode présentée ici constitue une approche intégrée qui permet de calculer à la fois la densité de réseaux optimaux et des valeurs de référence utilisables pour la validation des produits satellitaires. ■

Références

1. Lebel T, Diedhiou A, Laurent H. Seasonal cycle and interannual variability of the Sahelian rainfall at hydrological scales. *J Geophys Res* 2003 ; 108 (D8) : 8389-92.
2. Jobard I. Status of satellite retrieval of rainfall at different scales using multi-source data. Megha-Tropiques 2nd Scientific Workshop 2-6 July 2001, Paris (France), 2001, 9 p.
3. Ramge K. Contribution à l'estimation des précipitations et à la caractérisation des nuages en Afrique de l'Ouest à partir des observations des satellites METEOSAT et TRMM. Thèse de doctorat, université Paris 7, 2002, 210 p.
4. Nicholson SE, Some B, McCollum J, et al. Validation of TRMM and other rainfall estimates with high-density gauge dataset for West Africa. Part I: Validation of GPCC rainfall product and Pre-TRMM satellite and Blended products. *J Applied Meteor* 2003 ; 42 : 1337-54.

5. Balme M, Galle S, Lebel T. Démarrage de la saison des pluies au Sahel : variabilité à des échelles hydrologique et agronomique analysée à partir des données EPSAT-Niger. *Sécheresse (article accepté)*.
6. Lebel T, Amani A. Rainfall estimation in the Sahel: What is the ground truth? *J Applied Meteor* 1999 ; 38 : 555-68.
7. Kidder SQ, Vonder H. *Satellite meteorology: An Introduction*. New York: Academic Press, 1995 ; 340 p.
8. Petty GW. The status of satellite-based rainfall estimation over land. *Remote Sensing Environ* 1995 ; 51 : 125-37.
9. Levizzani V, Bauer P, Buzzi A, et al. Looking into the future of satellite rainfall estimations. Proc « The 2001 EUMESAT meteorological satellite data users' conf » Antalya, 2001 : 375-84.
10. Sevruk B. *Methods of correction for systematic error in point precipitation measurement for operational use*. WMO Report n°589. Geneva: World Meteorological Organization 1982 ; 91 p.
11. World Meteorological Organization (WMO). Guide to hydrological practices. WMO Publication No. 168. Geneva: WMO, 1994 ; 735 p.
12. Ali A, Lebel T, Amani A. Invariance in the spatial structure of sahelian rain fields at climatological scales. *J Hydrometeor* 2003 ; 4 : 996-1011.
13. Delhomme JP. *Application de la théorie des variables régionalisées dans le sciences de l'eau*. Thèse de doctorat, université Pierre et Marie Curie, 1976, 152 p.
14. Creutin JD. *Méthodes d'interpolation optimales des champs hydrométéorologiques*. Thèse de docteur-ingénieur, université de Grenoble, 1979, 174 p.
15. Xie P, Arkin PA. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull Amer Meteor Soc* 1997 ; 78 : 2539-58.
16. Arkin PA, Janowiak J. Analysis of the global distribution of precipitation. *Dyn Atmos Oceans* 1991 ; 16 : 5-16.
17. Rudolf B. *Management and analysis of precipitation data on a routine basis*. Proc Int WMO/IAHS/ETH Symp on precipitation and evaporation. Bratislava (Slovakia): Slovak Hydro-meteorology Institute, 1993 : 69-76.
18. Adler RF, Huffman GJ, Chang A, et al. The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979-Present). *J Applied Meteor* 2003 ; 4 : 1147-67.