



HAL
open science

Contribution à l'estimation des précipitations tropicales : préparation aux missions Megha-Tropiques et Global Precipitation Measurement

Philippe Chambon

► **To cite this version:**

Philippe Chambon. Contribution à l'estimation des précipitations tropicales : préparation aux missions Megha-Tropiques et Global Precipitation Measurement. Sciences de la Terre. Université Paris-Est, 2011. Français. NNT : 2011PEST1072 . tel-00659418v2

HAL Id: tel-00659418

<https://pastel.hal.science/tel-00659418v2>

Submitted on 30 Jan 2012

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.

THÈSE de DOCTORAT de l'UNIVERSITÉ PARIS-EST

pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ PARIS-EST

SPÉCIALITÉ : Sciences et Techniques de l'Environnement

présentée par

Philippe CHAMBON

CONTRIBUTION À L'ESTIMATION DES PRÉCIPITATIONS TROPICALES : PRÉPARATION AUX MISSIONS MEGHA-TROPIQUES ET GLOBAL PRECIPITATION MEASUREMENT

Soutenue le 18 novembre 2011 devant le jury composé de :

Mme. Laurence Picon	Présidente
M. Ziad Haddad	Rapporteur
M. Guy Delrieu	Rapporteur
M. Rémy Roca	Directeur de Thèse
Mme. Isabelle Jobard	Co-directrice de Thèse
M. Luc Musson-Génon	Co-directeur de Thèse
M. Serge Janicot	Examineur
M. Jérôme Lafeuille	Invité

Thèse préparée au Laboratoire de Météorologie Dynamique
Ecole Polytechnique

REMERCIEMENTS

Voilà le moment d'ajouter les dernières lignes à ce manuscrit. L'écriture de cette thèse marque pour moi l'aboutissement de presque quatre années de travail au Laboratoire de Météorologie Dynamique dans l'équipe "Cycle de l'Eau et de l'Énergie dans les Tropiques", depuis le stage de Master en février 2008, jusqu'à ma soutenance de doctorat.

Mes premiers remerciements vont évidemment à mes encadrants, Rémy Roca et Isabelle Jobard. J'ai rencontré Rémy à mon début de Master, alors que j'étais alors à la recherche d'un stage et d'un sujet de thèse. Rémy m'a proposé de travailler avec lui sur les précipitations estimées par satellite dans le cadre de la mission Megha-Tropiques. J'avais très peu de connaissances dans ce domaine, mais l'engouement avec lequel Rémy m'a décrit le projet m'a tout de suite motivé. Grâce à son énergie, sans limite je crois, Rémy m'a réellement transmis son enthousiasme pour la recherche. J'ai beaucoup appris en travaillant à ses côtés, tant, grâce à son expertise dans les domaines des observations satellitales et de la météorologie tropicale, que grâce à la rigueur de sa démarche scientifique. Rémy a su orienter et développer ma motivation, et m'encourager à aller au bout des choses et ceci grâce à son recul et sa vision scientifique très large du domaine de la physique de l'atmosphère. Rémy m'a également aidé à faire les bons choix pour cette nouvelle page qui se tourne et je l'en remercie.

Je voudrais remercier ma codirectrice Isabelle Jobard. Cela a été un plaisir de travailler avec Isabelle qui a été très disponible pour moi durant ces trois années. Isabelle n'a jamais hésité à rester très tard le soir au laboratoire pour répondre à mes questions et m'aider à analyser mes résultats. Son expertise dans le domaine de la télédétection des précipitations et la qualité de ses remarques m'ont permis d'avancer dans mes travaux plus rapidement. Isabelle a su également avec patience améliorer mes écrits grâce à son "œil de lynx", de mon premier extrait de conférence jusqu'à ce manuscrit. Son regard critique m'a donné le goût d'une réelle exigence par rapport à mon travail.

Je remercie également Luc Musson-Génon pour avoir accepté la codirection "administrative" de cette thèse.

Merci à mes rapporteurs Ziad Haddad et Guy Delrieu ainsi qu'aux membres de mon jury Serge Janicot, Jérôme Lafeuille et Laurence Picon qui ont évalué mon travail avec intérêt. En particulier, je tiens à remercier Laurence qui a su me donner les conseils dont j'avais besoin depuis mon arrivée dans le master Méthodes Physiques en Télédétection en 2007.

J'ai effectué mes travaux de thèse au Laboratoire de Météorologie Dynamique, sur le site de l'École Polytechnique, mais également sur le site de l'Université Pierre et Marie Curie pour la phase de rédaction de ce manuscrit. Je voudrais dire ma gratitude aux directeurs successifs,

Hervé Le Treut et Vincent Cassé, qui m'ont accueilli au sein de leur laboratoire et m'ont permis de mener mes travaux dans de bonnes conditions.

Quand je pense au LMD X, je pense évidemment au "préfa" qui a été ma seconde maison pendant cette thèse. Je garderai longtemps le souvenir de cette atmosphère toute particulière au préfa et qui m'aura permis de travailler dans la perpétuelle bonne humeur. Merci à Olivier Chomette, Patrick Raberanto, Karim Ramage, André Szantaï, Nicolas Gif, Michel Capderou, Sophie Cloché, Thomas Fiolleau, Julien Lenseigne, Jérémie Aublanc, Chantal Claud, Paul-Étienne Mallet et Bahjat Alhammoud.

Je tiens à remercier plus personnellement Michel Capderou qui m'a fait partager son expertise à la fois en orbitographie, mais aussi en cuisine du canard ! Sa passion pour les voyages m'a transporté de nombreuses fois loin du préfa ou du "Villejuif express" grâce à des récits plus extraordinaires les uns que les autres. Merci à Sophie Cloché qui a su répondre à bon nombre de mes questions dans le cadre de Megha-Tropiques, merci d'avoir également partagé le café du matin avec moi et m'avoir ainsi permis de toujours commencer la journée de bonne humeur !

J'en serais sans doute encore au tout début de mes travaux de thèse sans le soutien de Karim Ramage et Julien Lenseigne. C'est une réelle chance d'avoir pu travailler dans de telles conditions en matière de soutien informatique ; Karim et Julien ont toujours su répondre à mes nombreuses questions avec patience et grande disponibilité (même pour répondre aux questions les plus désespérantes et dont la réponse était souvent "-R"). Merci également à Christophe Boitel et Robert Franchisseur dont la présence tardive au LMD Jussieu est toujours réconfortante pour un thésard en pleine rédaction !

Merci également à Jérémie Aublanc pour son soutien et sa gentillesse ; sans son aide je n'aurais pu faire face à la fois à ma thèse et au développement du produit TAPEER-BRAIN pour la mission Megha-Tropiques. Je tiens aussi à remercier Nicolas Gif qui m'a aidé de nombreuses fois pour des réalisations graphiques ou autres ; ses connaissances en informatique ont éclairé nombre de mes problèmes.

Je remercie Nicolas Viltard pour toutes les discussions que nous avons pu avoir sur les micro-ondes ainsi que son équipe BRAIN, Sandrine Baudrand, Christophe Dufour et Pierre-Emmanuel Kirstetter. Pierre-Emmanuel m'a initié aux géostatistiques et m'a fait partager ses connaissances sur les données pluviométriques, je tiens à le remercier pour toutes les discussions que nous avons pu avoir sur le sujet et qui m'ont beaucoup apporté. Merci à Marielle Gosset pour les nombreux échanges que nous avons pu avoir sur les données sols et la pluie, je remercie également Marielle pour tous ses commentaires qui m'ont permis de bien préparer ma soutenance de thèse.

Je n'oublie pas Stéphane Sportouch, Martine Roux, Eliane Rier, Mariame Seydi et en particulier Martine Mahérou qui m'ont aidé avec gentillesse dans toutes les démarches administratives qui jalonnent la vie d'un doctorant.

Mon travail n'aurait sans doute pas abouti sans l'énergie insufflée par mes collègues et amis du LMD. Je tiens à remercier mon camarade thésard Thomas Fiolleau qui a partagé avec moi ces années de travail pour la mission MT. Merci pour tous ces bons moments passés à discuter et rire, du plateau de Palaiseau jusqu'à la pointe de l'Inde, en passant par notre quartier général ou "Yalta bis", la pizzeria "San Giovanni". Merci également à Nicolas Rochetin, Jessica Vial, Nathalie Bertrand, Virginie Guémas, Ara Arakelian et Romain Pilon pour tous ces fous rires dans les couloirs du LMD et toutes les soirées passées ensemble qui me manqueront beaucoup à Washington. Merci à Christelle Castet d'avoir été une super coloc en amenant chez moi sa bonne humeur et ses cours d'anglais personnalisés pendant les quelques mois d'absence de mon colocataire et ami Aurélien. Merci à Anthony Guignard, Paul-Étienne Mallet, Rodrigo Guzman, Isabelle Tobin, Alberto Cassado, Abdoul-Khadre Traore, Youssouph Sané, Clémentine Junquas, Solange Fermepin, Benjamin Charnay, Aymeric Spiga, Joséphine Ghattas et Romain Roehrig. Merci également à Véronique Fabart pour sa gentillesse et pour avoir partagé son bureau à Jussieu avec moi pendant la phase de rédaction. Je risque d'avoir oublié certains d'entre vous, mais sachez que votre rencontre aura marqué mes années au LMD.

Enfin, je tiens à remercier mes frères et soeur François, Christine et Romain, mon père et ma mère, qui m'ont tous soutenu pendant ces années de thèse, pendant les bons moments, comme dans les plus difficiles. Je souhaite également remercier Céline qui m'a soutenu (et supporté!) pendant la fameuse période de rédaction. Je terminerai en remerciant mes cousins et amis, Aurélien, Adrien, Anne-Laure, Alice, Bénédicte, Charlotte, Chrystèle, Florent, François-Éric, Gaëlle, Hélène, Laudine, Marie-Delphine, Mylène, Paul, Rachel et Samy pour avoir été présents le jour de ma soutenance; j'ai vraiment été touché par votre venue.

RÉSUMÉ

CONTRIBUTION À L'ESTIMATION DES PRÉCIPITATIONS TROPICALES : PRÉPARATION AUX MISSIONS MEGHA-TROPIQUES ET GLOBAL PRECIPITATION MEASUREMENT

Les précipitations résultent d'un phénomène atmosphérique caractérisé par une variabilité spatiale et temporelle forte. Cette variabilité dans la distribution des pluies et des événements intenses a des impacts en hydrologie de surface variés selon les régions du monde. Toute modification du climat tropical est associée à une modification du cycle de l'eau et de l'énergie dans ces régions. Dans un contexte de changement climatique, il est donc important de développer des outils permettant d'estimer quantitativement les précipitations, à l'échelle du globe, à la fois sur les surfaces continentales et les surfaces océaniques.

Les travaux présentés dans cette thèse s'intéressent à l'observation des précipitations depuis l'espace. En effet, la mesure des pluies nécessite une densité d'observations élevée qui, sur l'ensemble des Tropiques, n'est accessible qu'à partir d'observations spatiales. Depuis plusieurs décades, les moyens satellitaires à disposition ont beaucoup évolué et offrent aujourd'hui une densité d'observations de plus en plus fortes. Grâce aux nouvelles missions déployées telles que Megha-Tropiques au sein de la future constellation GPM (Global Precipitation Measurement), on a accès à un ensemble de systèmes d'observations qui amène à une densité accrue d'observations spatiales. L'estimation quantitative des précipitations n'était possible qu'à l'échelle mensuelle, il est maintenant envisageable d'estimer la pluie par satellite à des échelles de temps de plus en plus fines. Cette thèse s'intéresse aux échelles $1^\circ/1$ -jour, échelle clé pour les études météorologiques et hydrologiques.

Il existe un large spectre de méthodes d'estimation de précipitations par satellite, de qualité inégale. Dans un premier temps, une analyse des produits issus des développements les plus récents montre que leur qualité a atteint un degré suffisant pour être utilisé de manière quantitative aux échelles de temps pertinentes en météorologie. Il apparaît également qu'à ces échelles de temps, il est nécessaire d'utiliser les estimations de cumuls de précipitations conjointement avec leurs barres d'erreurs. Une nouvelle méthode d'estimations de précipitations sur l'ensemble de la ceinture tropicale, appelée TAPEER (Tropical Amount of Precipitation with an Estimate of ERrors), est donc développée dans le but d'estimer des cumuls de pluie et leurs erreurs associées à l'échelle $1^\circ/1$ -jour. Cette approche est fondée sur une méthode de fusion de données de l'imagerie Infrarouge d'une constellation de satellites géostationnaires et d'estimations de taux de pluie issues de radiomètres Micro-ondes d'une constellation de satellites défilant. Des techniques de modélisations sont mises en oeuvre afin d'associer une erreur aux cumuls de pluie produits. Une investigation détaillée du bilan d'erreur de la méthode TAPEER montre que les sources principales d'incertitudes sont liées à l'échantillonnage et aux biais systématiques sur les taux de pluie d'intensité moyenne.

Une étude sur l'été 2009 révèle l'importance de l'utilisation de la barre d'erreur dans l'analyse de la distribution des pluies, en particulier pour les plus forts cumuls sur la ceinture tropicale.

ABSTRACT

CONTRIBUTION TO THE ESTIMATION OF TROPICAL PRECIPITATION : PREPARATION TO THE MEGHA-TROPIQUES AND GLOBAL PRECIPITATION MEASUREMENT MISSIONS

Precipitation results from atmospheric phenomena, which are characterized by a large space and time variability. The distribution of rainfall, in particular of strong rainy events, has various impacts in surface hydrology over the different regions in the world (e.g. floods). Any change in the Tropical climate is associated with a modification of the water and energy cycle over those regions. Therefore, in a context of climate change, it is important to develop new tools able to provide quantitative precipitation measurements, both over land and over the open oceans. The work presented hereafter deals with precipitation estimation from space. Indeed, measuring rainfall requires a high density of observations, which, over the whole tropical belt, can only be provided from space. For several decades, the availability of satellite observations has greatly increased and offers an increasing number of measurements. Thanks to newly implemented missions like the Megha-Tropiques mission and the forthcoming GPM constellation (Global Precipitation Measurement mission), measurements from space become available from a set of observing systems. Quantitative precipitation estimation were only available at the monthly scale, it is now possible to estimate rainfall from space at increasingly fine scale. In this work, we focus on the $1^\circ/1$ -day scale, key scale of meteorological and hydrological studies.

Various methods exist to estimate rainfall from space but they provide estimates of unequal quality. First, a meteorological benchmark is set up with ground-based observations from the African Monsoon Multidisciplinary Analysis (AMMA) program. The analysis shows that the last generation of combined infrared-microwave products is describing the variability of rainfall similarly to ground measurements at meteorologically relevant scales. It also appeared that at these scales, rain accumulation estimations should be used taking into account their uncertainties. A novel methodology for quantitative precipitation estimation is introduced; its name is TAPEER (Tropical Amount of Precipitation with an Estimate of ERrors) and it aims to provide $1^\circ/1$ -day rain accumulations and associated errors over the whole Tropical belt. This approach is based on a combination of infrared imagery from a fleet of geostationary satellite and passive microwave derived rain rates from a constellation of low earth orbiting satellites. Modelling techniques are developed in order to associate an error with the individual rain accumulations. An investigation of the error budget of the TAPEER method shows that the two main contributions to the total error are related to sampling and systematic errors on rain rates of medium intensity.

A study on the summer 2009 period reveals the importance of using error bars when analyzing the distribution of rainfall, especially for the most important rain accumulations of the tropics.

Table des matières

1	Introduction à l'estimation quantitative des précipitations tropicales depuis l'espace	1
1.1	Le cycle de l'eau et de l'énergie dans les Tropiques	1
1.2	Enjeux liés à l'estimation de l'intensité des précipitations	3
1.3	L'observation de la pluie	5
1.4	Erreurs sur l'estimation des précipitations par satellite	12
1.4.1	Estimation quantitative des précipitations (QPE)	12
1.4.2	Bilan des erreurs associées à l'estimation des précipitations	18
1.4.3	État de l'art sur la modélisation des erreurs des QPE estimées par satellite	20
1.5	Objectif et organisation de la thèse	25
1.5.1	Objectif	25
1.5.2	Organisation de la thèse	26
2	Une nouvelle méthode d'évaluation des produits satellites	29
2.1	Introduction	30
2.2	Data	32
2.2.1	Rationale	32
2.2.2	The ground-based rain products	33
2.2.2.1	Rain gauge networks	33
2.2.2.2	Products and associated error computation	35
2.2.3	The satellite-based rain products	36
2.2.3.1	The satellite rain estimates	36
2.2.3.2	The error model	37
2.2.4	Error computation results	39
2.3	Methodology of comparison using error estimates	41
2.3.1	Approach	41
2.3.2	Score indices	42
2.4	Results	43
2.4.1	Monsoon preonset and postonset and seasonal scales	43
2.4.2	Synoptic scales	43
2.4.2.1	Daily time series	43
2.4.2.2	3-5 Daily filtered time series	46
2.4.3	Diurnal scale	48
2.5	Summary and discussion	51

2.6	Conclusion du chapitre	53
3	Étude du bilan d’erreurs des précipitations estimées par satellite	55
3.1	Introduction	56
3.2	Data	58
3.2.1	Thermal infrared geostationary images	58
3.2.2	Passive Microwave instantaneous rain estimates	58
3.2.3	Validation datasets from ground-based rain measurements	59
3.3	Methodology : TAPEER	60
3.3.1	Rain accumulation estimation	60
3.3.2	Error estimation	61
3.3.3	Error propagation method	62
3.4	Results of the error budget analysis	63
3.4.1	Propagating errors from PM-derived rain rates to accumulated-rain	63
3.4.2	Propagating errors from TIR geostationary data to accumulated-rain	66
3.4.3	Comparing error sources	67
3.5	Performance of TAPEER-BRAIN rain estimations	70
3.5.1	Validation over two African sites	70
3.5.2	A first global comparison	72
3.6	Summary and perspectives	74
3.6.1	Summary of the error budget analysis	74
3.6.2	Perspectives	74
3.7	Appendix : Variogram computation	75
3.8	Conclusion du chapitre	80
4	Simulations de la constellation GPM dans les Tropiques	83
4.1	Introduction	83
4.2	Méthodologie	89
4.2.1	Simulation de produits de Niveau-2	89
4.2.2	Degrés de liberté de la méthode TAPEER : configuration de la constellation de satellites et volume spatio-temporel d’apprentissage	92
4.3	Étude de sensibilité des estimations TAPEER-GSMaP pour des constellations idéalisées	96
4.3.1	Sensibilité à l’heure locale de passage d’un satellite héliosynchrone	96
4.3.2	Sensibilité au nombre de satellites héliosynchrones dans une constellation	99
4.3.3	Sensibilité à la présence d’un satellite placé sur une orbite tropicale dans une constellation	101
4.4	Simulation de l’apport de Megha-Tropiques : constellation réaliste	106

4.5	Conclusion	108
5	Conclusion, discussion et perspectives	111
5.1	Résumé des travaux présentés	111
5.2	Discussion	114
5.3	Perspectives	118
A	Radiométrie micro-ondes pour la télédétection spatiale des précipitations	121
A.1	Rappels théoriques sur le transfert radiatif dans le domaine des micro-ondes . . .	121
A.1.1	Interaction entre le rayonnement et l'atmosphère	121
A.1.2	Relation entre les données satellite Micro-ondes et la pluie	123
A.2	Exemple de l'algorithme BRAIN pour l'estimation des pluies instantannées . . .	128
A.3	Résolution spatiale des radiomètres micro-ondes	131
B	The TAPEER-BRAIN product	137
B.1	TAPEER : Algorithm Theoretical Basis Document	138
B.1.1	Tropical Amount of Precipitation	138
B.1.2	Estimate of ERRors	143
B.2	TAPEER implementation for Megha-Tropiques and GPM missions : the TAPEER- BRAIN product	145
B.2.1	Observing systems	145
B.2.1.1	IR Fleet	145
B.2.1.2	Microwave Constellation	146
B.2.1.3	Space and Time GEO/LEO Collocation files	146
B.2.1.4	General Pre-processing	147
B.2.2	Algorithm processing	147
B.3	Preliminary validation of the TAPEER-BRAIN product	150
B.3.1	Comparison with TRMM-3B42 over the whole Tropical belt	150
B.3.2	Comparison with surface rainfall products over West Africa at Meteorolo- gically relevant scales	150
	Bibliographie	161
	Acronymes	177
	Liste des figures	181
	Liste des tableaux	190

Introduction à l'estimation quantitative des précipitations tropicales depuis l'espace

1.1 Le cycle de l'eau et de l'énergie dans les Tropiques

Le cycle de l'eau atmosphérique et les échanges d'énergie du système Terre-Océan-Atmosphère sont très actifs dans les Tropiques, car ces régions reçoivent en moyenne plus d'énergie du soleil qu'elles n'en renvoient vers l'espace (figure 1.1). De ce fait, ce sont les circulations océaniques et atmosphériques de grande échelle qui redistribuent vers nos latitudes cet excédent d'énergie. Toute modification du bilan énergétique des zones tropicales a des conséquences sur le climat de la planète toute entière.

Dans les régions subtropicales, l'évaporation est importante par rapport aux précipitations (figure 1.1); c'est la circulation de Hadley/Walker qui transporte cet excédent d'énergie sous forme de chaleur latente vers la zone de convergence des alizés où les basses couches de l'atmosphère se chargent en vapeur d'eau. L'énergie apportée par ce transport est évacuée dans la troposphère libre sous forme de chaleur latente grâce aux phénomènes convectifs produisant des précipitations.

Près de 60% des précipitations globales ont lieu dans les régions tropicales. Ce volume d'eau n'est pourtant fourni que de manière saisonnière à des régions situés dans les zones tropicales ou équatoriale, laissant place à des périodes de sécheresse parfois longues.

Les précipitations résultent en effet d'un phénomène atmosphérique fortement variable dans l'espace et dans le temps. Elles conditionnent en partie les ressources en eau disponibles à l'échelle régionale. Par conséquent, les précipitations ont un impact direct sur les populations et leur environnement. Un excès de pluie entraîne des inondations et met en danger les personnes et les biens. Un manque de pluie a des conséquences sur les rendements des productions agricoles et les sécheresses peuvent entraîner des déplacements de population.

À la fois source de vie et de catastrophes naturelles, les précipitations ont donc une valeur sociétale et économique forte. La mesure des pluies présente ainsi un intérêt crucial pour la gestion des ressources en eau sur l'ensemble des Tropiques. En particulier, il est important de caractériser

les hauteurs d'eau précipitées, la fréquence des pluies, et leurs répartitions géographiques.

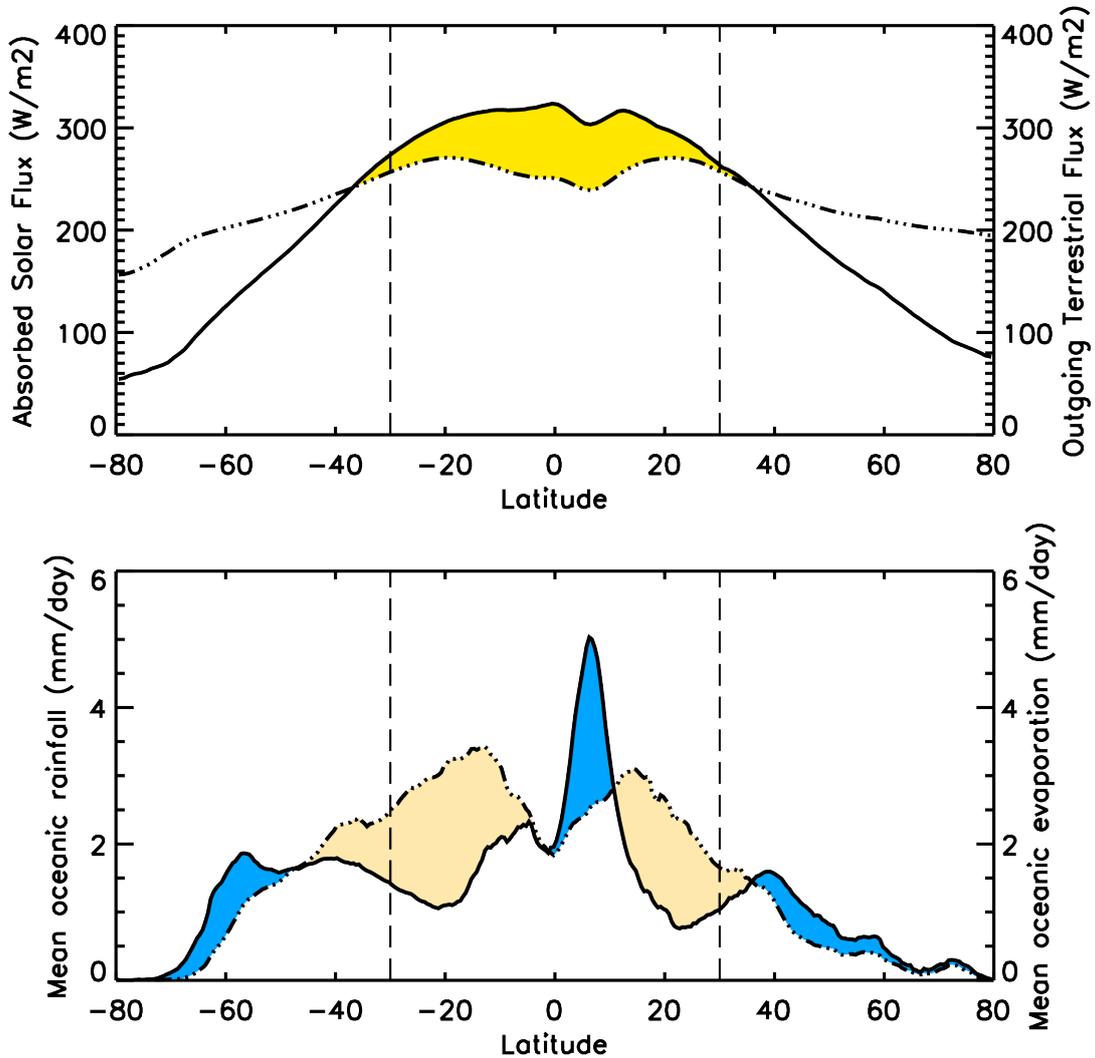


Figure 1.1 – Figure du haut : moyennes zonales du rayonnement solaire absorbé (trait plein), et du rayonnement émis vers l'espace (trait pointillé) par le système Terre-Océan-Atmosphère. Ces moyennes sont réalisées à partir des données de flux sortant au sommet de l'atmosphère CERES-EBAF pour la période 2000 à 2010 (Loeb *et al.*, 2009). La zone orange met en évidence la bande de latitude excédentaire en énergie. Figure du bas : moyennes zonales des précipitations océaniques (trait plein) et de l'évaporation océanique (trait pointillé). Ces moyennes sont réalisées à partir des produits Évaporation et Précipitations HOAPS-3 pour la période 1987 à 2005 (Andersson *et al.*, 2010). Les zones bleues (resp. beiges) mettent en évidence les bandes de latitude où le bilan [Évaporation - Précipitation] est négatif (resp. positif)

1.2 Enjeux liés à l'estimation de l'intensité des précipitations

Des changements substantiels dans le cycle de l'eau sont attendus du fait du réchauffement climatique (Allan *et al.*, 2010). En effet, l'augmentation de la température moyenne à l'échelle du globe entraîne une augmentation du contenu potentiel en vapeur d'eau des basses couches de l'atmosphère, suivant la relation de Clausius-Clapeyron :

$$\frac{d \ln(e_S)}{dT} = \frac{L}{R.T^2} = \begin{cases} 0,14K^{-1} & T = 200K \\ 0,07K^{-1} & T = 273K \\ 0,06K^{-1} & T = 300K \end{cases} \quad (1.1)$$

où e_S est la pression de vapeur saturante, T la température, L la chaleur latente de vaporisation de l'eau et R la constante des gaz parfaits. À des températures typiques de la basse troposphère (273 à 300 K), le rapport $\frac{L}{R.T^2}$ vaut de 6 à 7 $\%.K^{-1}$. L'analyse des simulations réalisées à partir de modèle de circulation générale montre que l'on assiste en effet à une humidification des basses couches de l'atmosphère à un taux proche de celui-ci (Held and Soden, 2006; Pall *et al.*, 2007).

Le bilan [Évaporation - Précipitation] est nul pour le système fermé Terre-Océan-Atmosphère dans son ensemble, mis à part un terme de stockage négligeable. Les précipitations globales devraient donc augmenter d'un taux proche de 6 à 7 $\%.K^{-1}$ suivant l'augmentation du contenu en vapeur d'eau de l'atmosphère. La quantité de pluie n'est cependant pas uniquement reliée à la quantité de vapeur d'eau disponible, mais également à d'autres considérations énergétiques. En revanche, les échanges d'énergie des Tropiques vers les plus hautes latitudes aboutissent à un bilan [Évaporation - Précipitation] non nul à l'échelle de ces régions (figure 1.1). Il est donc difficile de formuler des hypothèses quant à un changement dans la distribution des pluies dans la zone intertropicale. Ce changement est pourtant un élément d'analyse essentiel à la compréhension de l'impact de l'augmentation des températures sur le climat global.

L'évolution des pluies pourrait se caractériser de deux manières distinctes : soit les pluies moyennes augmentent sur l'ensemble des Tropiques soit la distribution des pluies évolue à moyenne constante, en particulier l'intensité et l'occurrence des événements pluvieux intenses. En effet, l'augmentation du contenu potentiel en vapeur d'eau conduisant à une humidification des basses couches de l'atmosphère ne se ferait pas en permanence, mais de manière sporadique dans l'espace et dans le temps (échelle journalière) (Beucher, 2010). Le surplus d'énergie ainsi disponible localement serait alors évacué dans la troposphère libre par des précipitations plus intenses (Allen and Ingram, 2002).

Des études récentes utilisant des bases de données d'observations satellitaires ont conduit à différents éléments qui font débat. Wentz *et al.* (2007) a mis en évidence une augmentation des pluies moyennes sur les océans Tropicaux, à un taux proche de celui dérivé de la relation de Clausius Clapeyron. Une analyse sur l'ensemble des Tropiques (continents et océans) a pourtant

permis d'observer une tendance quasi nulle à l'augmentation des pluies moyennes, une fois isolée la variabilité des pluies liée au phénomène El Niño et aux éruptions volcaniques (Gu *et al.*, 2007). La figure 1.2 (c) montre cette tendance quasi nulle de 0,0154 mm/jour/10 ans, détectés sur la période 1979-2010. La variabilité interannuelle des précipitations est très forte, ce qui rend l'analyse de tendances complexe.

En revanche, une analyse de la distribution des pluies sur l'ensemble des Tropiques montre une augmentation de l'occurrence des cumuls de pluie les plus faibles et les plus intenses (Lau and Wu, 2007).

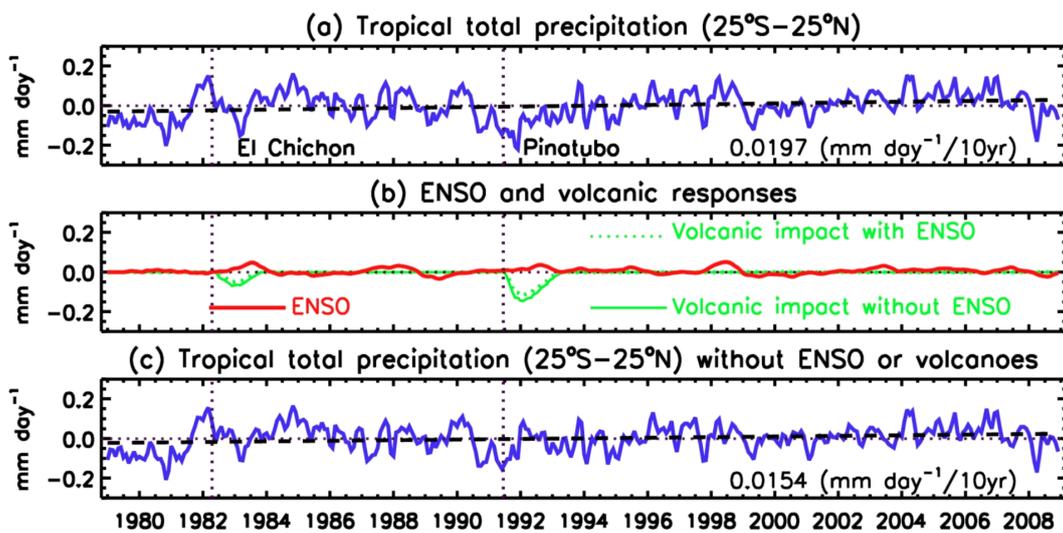


Figure 1.2 – (a) Anomalie des précipitations tropicales en mm/jour (moyenne glissante de 3 mois) pour la période 1979-2010, (b) réponse linéaire des précipitations tropicales à l'ENSO et aux éruptions volcaniques, (c) anomalie des précipitations tropicales une fois isolée la variabilité des pluies liée au phénomène ENSO et aux éruptions volcaniques (Source : C. Kummerow)

Cette tendance est également observée sur les surfaces océaniques pour les cumuls journaliers les plus forts (Allan *et al.*, 2010). La figure 1.3 montre en effet une augmentation de l'occurrence de ces cumuls en fonction de la température de surface des océans tropicaux. La tendance dérivée de l'analyse d'observations satellitaires est cependant deux fois plus forte que celle déduite des simulations numériques.

L'évolution des pluies pourrait impacter la disponibilité des ressources en eau dans les Tropiques (IPCC, 2007), en particulier pour les productions agricoles. Dans le cas d'une augmentation de l'intensité et/ou de l'occurrence des événements pluvieux intenses, les inondations pourraient, elles aussi, être plus fréquentes et/ou causer des dommages plus importants. Il est donc crucial d'effectuer des estimations quantitatives des précipitations afin de mieux caractériser ces

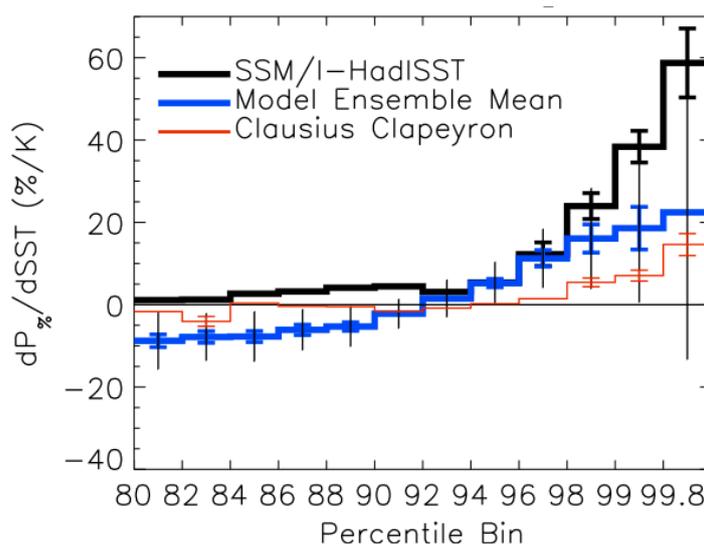


Figure 1.3 – Sensibilité linéaire de la fréquence des intensités journalières des précipitations à un changement de température de l’océan pour : des estimations par satellite (noté SSM/I), une moyenne d’ensemble de résultats de simulations de modèles, et une expérience contrôlée par la relation de Clausius-Clapeyron. (Allan *et al.*, 2010)

évolutions en analysant les séries temporelles d’observations disponibles. Il faut aussi améliorer la représentation du cycle hydrologique dans les modèles de prévision du temps/du climat grâce à la comparaison avec les observations (Stephens *et al.*, 2010). La prévision des récoltes et des crues nécessite également des estimations quantitatives des précipitations en forçage des modèles (Teo, 2006; Hong *et al.*, 2006).

1.3 L’observation de la pluie

Les pluviomètres et les radars sont des systèmes d’observation conventionnels de la pluie :

- Les premières mesures effectuées grâce à des pluviomètres qui ont été répertoriées en Europe remontent au 17^{ième} siècle (New *et al.*, 2001). Plusieurs institutions centralisent les données de cumuls mensuels récoltées grâce à des réseaux regroupant plusieurs dizaines de milliers de pluviomètres à travers le monde. Elles reconstituent des archives historiques de plusieurs siècles. Parmi ces archives de données, on peut citer le GHCN (Global Historical Climatology Network, Peterson and Vose, 1997), le GPCC (Global Precipitation Climatology Center, Rudolf *et al.*, 2010) et le CRU (Climatic Research Unit, New *et al.*, 2000). Ces systèmes d’observation sont très inégalement répartis à la surface du globe : sur les surfaces continentales, seules certaines régions du monde comme l’Amérique du Nord et

l'Europe sont équipées de réseaux denses de pluviomètres. Sur les surfaces océaniques les données mesurées depuis quelques atolls sont les seules disponibles (New *et al.*, 2001).

- Les radars météorologiques sont des systèmes d'observation de la pluie par télédétection. Ils sont dérivés des radars utilisés à partir des années 1930 pour le repérage des avions. Les échos mesurés par ces radars sont parasités par les précipitations : ces signaux parasites sont devenus l'objet de mesures pour la pluviométrie. Les réseaux de radars existants, tel le réseau ARAMIS en France (Parent du Châtelet, 2003), sont répartis uniquement sur certaines régions du monde et donc tout aussi inégalement que les réseaux de pluviomètres.

Plus récemment, les satellites ont permis d'observer depuis l'espace les continents et océans, et en particulier toute la ceinture intertropicale.

Les trois systèmes d'observation permettent d'estimer les précipitations de manière quantitative. Néanmoins, seuls les satellites offrent la couverture spatiale nécessaire pour l'estimation des pluies sur l'ensemble des Tropiques.

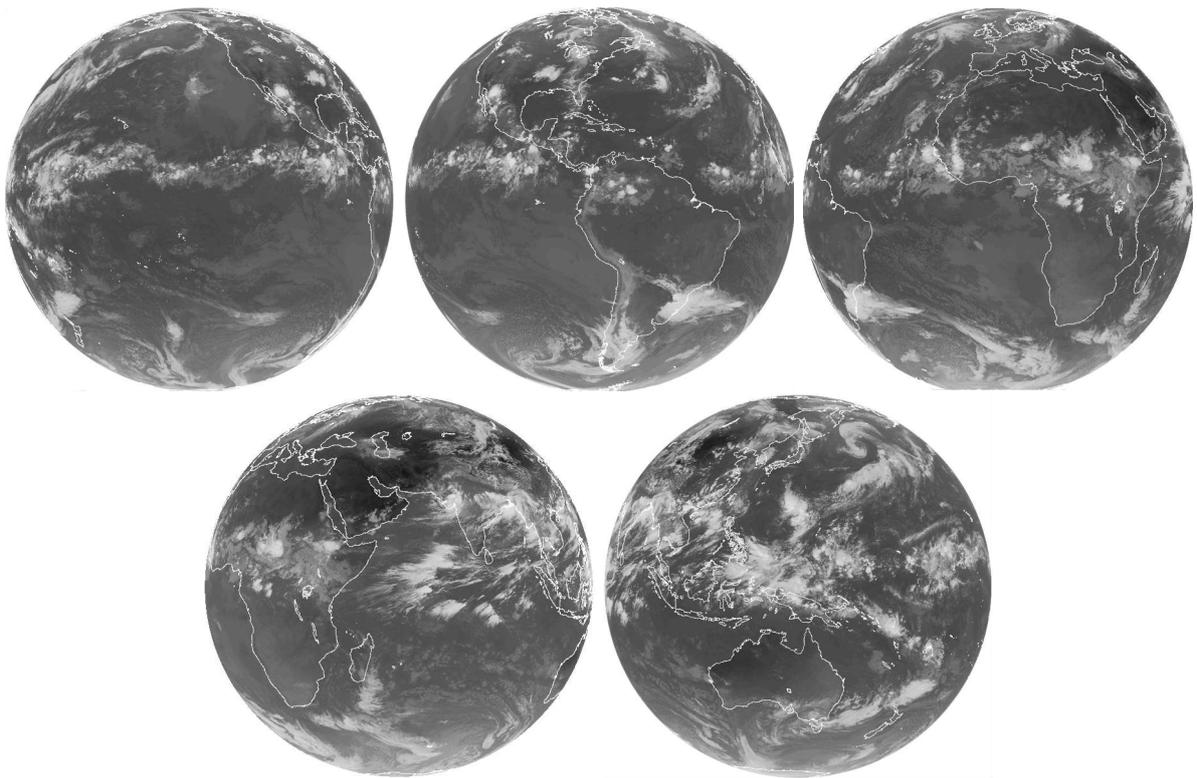


Figure 1.4 – Images acquises dans le domaine IR thermique à 06h00-UTC le 9 août 2011 depuis 5 satellites en orbite géostationnaire. De gauche à droite, première ligne : GOES-W (NOAA), GOES-E (NOAA), Meteosat-9 (EUMETSAT), et deuxième ligne : Meteosat-7 (EUMETSAT), MTSAT-1 (JMA). (Source : goes.noaa.gov)

Un **système d'observation** satellitaire, noté "SO" dans la suite de ce manuscrit, se définit par plusieurs caractéristiques : (i) le domaine des fréquences auxquelles les mesures sont effectuées, (ii) la résolution spatiale, (iii) la résolution temporelle. Ces trois éléments dépendent de l'instrument du SO ainsi que de l'orbite et de l'altitude du satellite sur lequel l'instrument est placé. Un SO sera donc désigné par la combinaison "Instrument/Satellite" dans la suite du manuscrit (exemple : TMI/TRMM voir ci-dessous).

Les domaines de fréquences utiles pour l'estimation des précipitations sont principalement le domaine des micro-ondes (MO) et le domaine de l'infrarouge (IR) (voir annexe A et annexe B).

Les imageurs IR à bord de satellites en orbite géostationnaire permettent de produire des mesures de température du sommet des nuages de façon opérationnelle depuis les années 1970. Grâce au large champ de vue des géostationnaires, cinq satellites suffisent pour observer l'ensemble des Tropiques (figure 1.4) à des résolutions temporelles de 15 à 30 minutes (sur 90% de la ceinture tropicale). La température du sommet d'un nuage est toutefois très indirectement liée aux précipitations qu'il engendre (Arkin, 1979; Arkin and Meisner, 1987) contrairement aux mesures dans le domaine des MO (voir Annexe A).

Le premier radiomètre passif MO, EMSR (Electrically Scanning Microwave Radiometer) a été embarqué à bord des satellites Nimbus-5 et Nimbus-6 en 1972 et 1975. Après le SSMR (Scanning Multichannel Microwave Radiometer) lancé à bord de Nimbus-7 en 1978, c'est le radiomètre SSMI (Special Sensor Microwave/Imager) qui a été lancé à bord de la série des satellites héliosynchrones du programme DMSP (Defense Meteorological Satellite Program) à partir de 1987. Ces SO ne permettent d'acquérir qu'une seule observation par jour pour un lieu donné dans la bande de latitude [20°S - 20°N]. En 1997, la mission TRMM (Tropical Rainfall Measuring Mission) est la première mission spatiale dédiée à l'étude des précipitations tropicales à être lancée (Kummerow *et al.*, 1998). Ce satellite est le premier à être équipé d'un radar pour les précipitations (bande K - 13,8 GHz), en plus du radiomètre à micro-ondes passif TMI (TRMM Microwave Imager) et de deux autres instruments. La conception de TMI est proche de celle des radiomètres SSMI dont les fréquences d'observation sont comprises entre 18 et 85 GHz, avec un canal supplémentaire à 10 GHz.

L'orbite de TRMM, inclinée à 35 degrés par rapport à l'équateur, permet d'observer un lieu donné à des heures locales variables, contrairement à un satellite sur une orbite héliosynchrone. En revanche, la zone d'observation de TMI/TRMM est plus étroite que celle de SSMI/DMSP. Il y a donc un peu moins d'un passage par jour en moyenne pour un lieu donné dans la bande de latitude [20°S - 20°N]. En 2002, le satellite Aqua en orbite héliosynchrone est lancé avec à son bord le radiomètre AMSRE (Advanced Microwave Scanning Radiometer - Earth Observing System) (Kawanishi *et al.*, 2003) suivi par d'autres satellites équipées de sondes dans le domaine des micro-ondes (Kidd and Huffman, 2011).

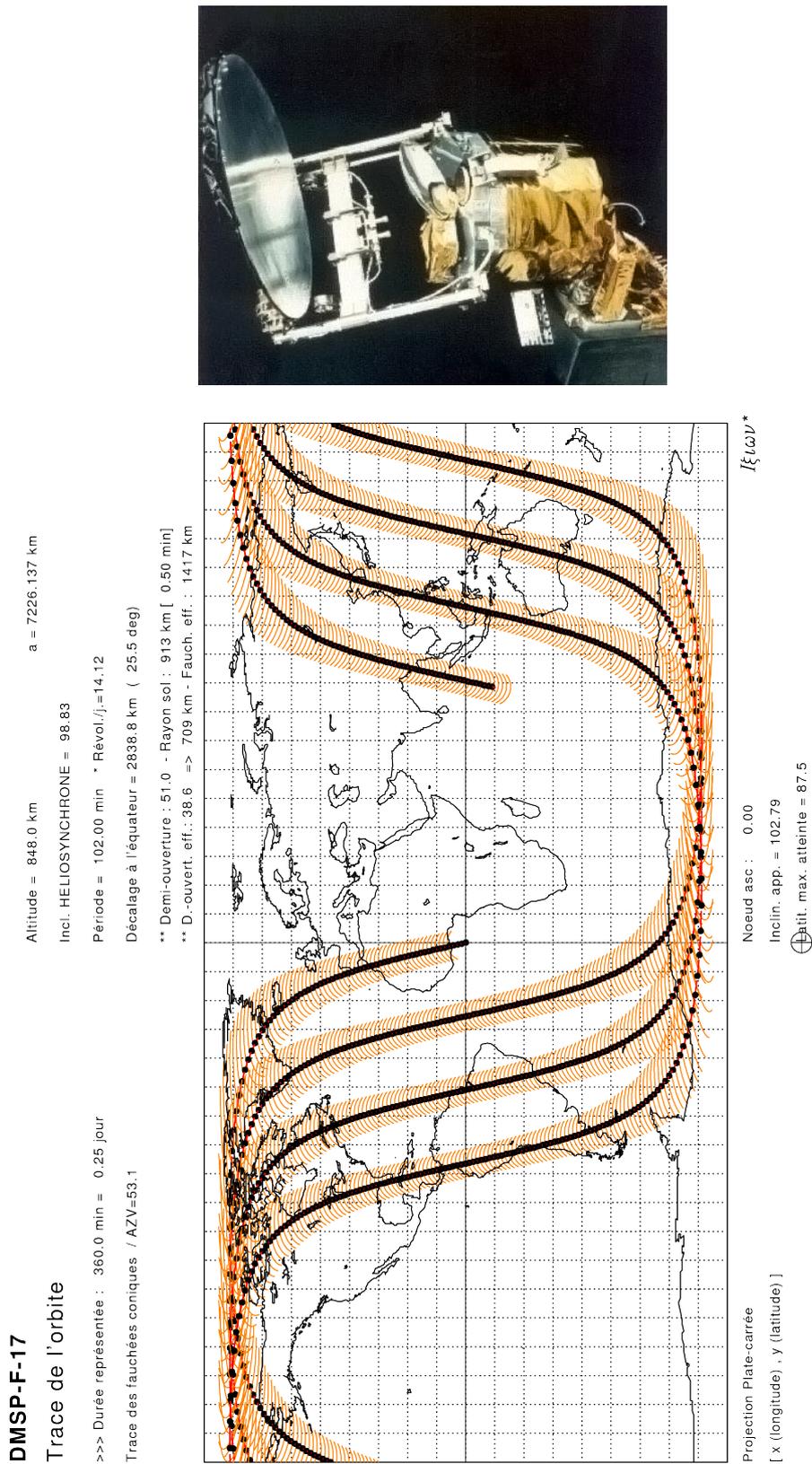


Figure 1.5 – Figure du haut : photographie de l'instrument. Figure du bas : Trace au sol de l'orbite du satellite en noir et fauchées de l'instrument en orange pour une période de 6 h consécutives (calculées grâce au logiciel d'orbitographie IXION (Capderou, 2005, 2009). Instrument : SSMI, Satellite : DMSP-F17.

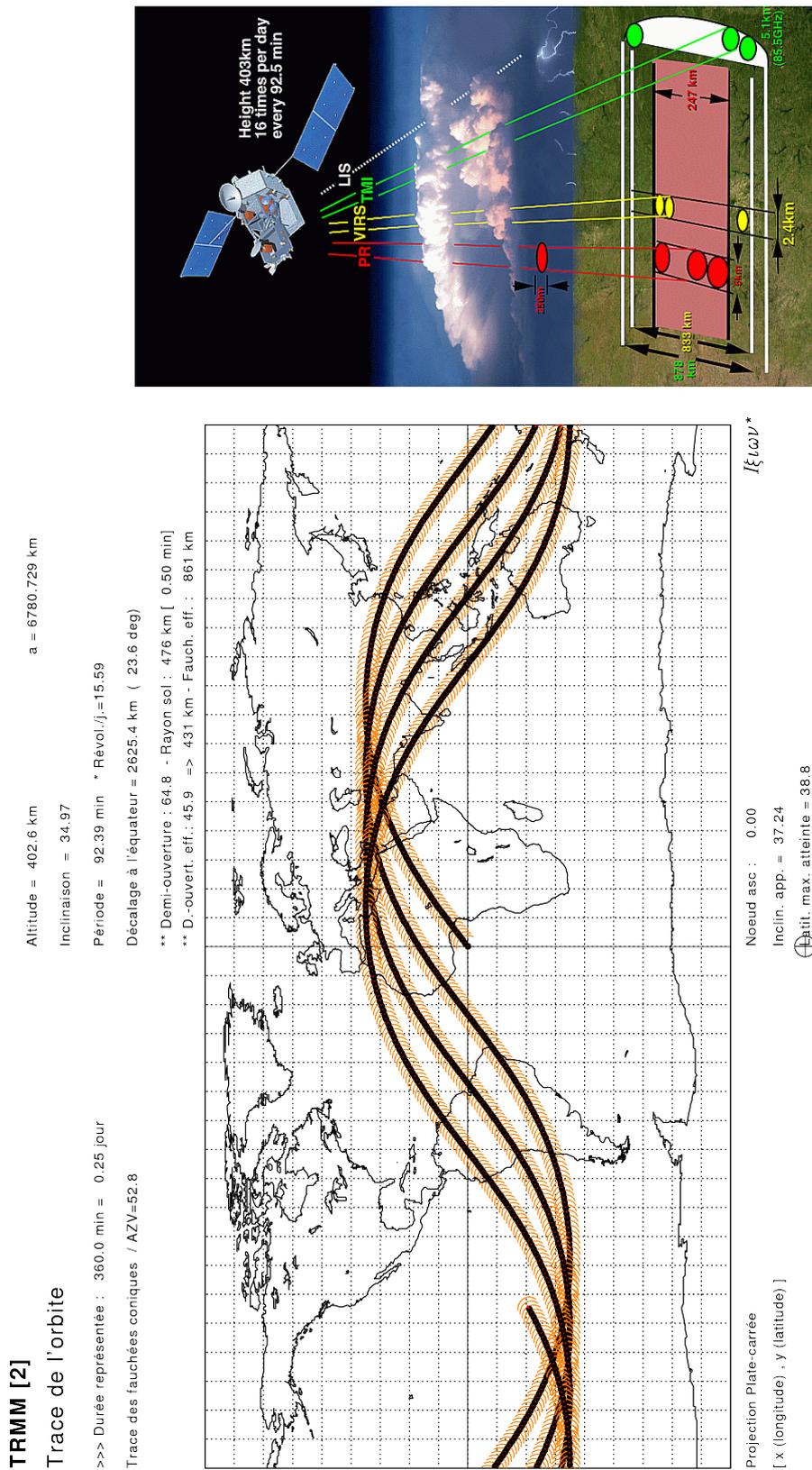


Figure 1.6 – Figure du haut : vue d'artiste du satellite. Figure du bas : idem que figure 1.5. Instrument : TMI, Satellite : TRMM

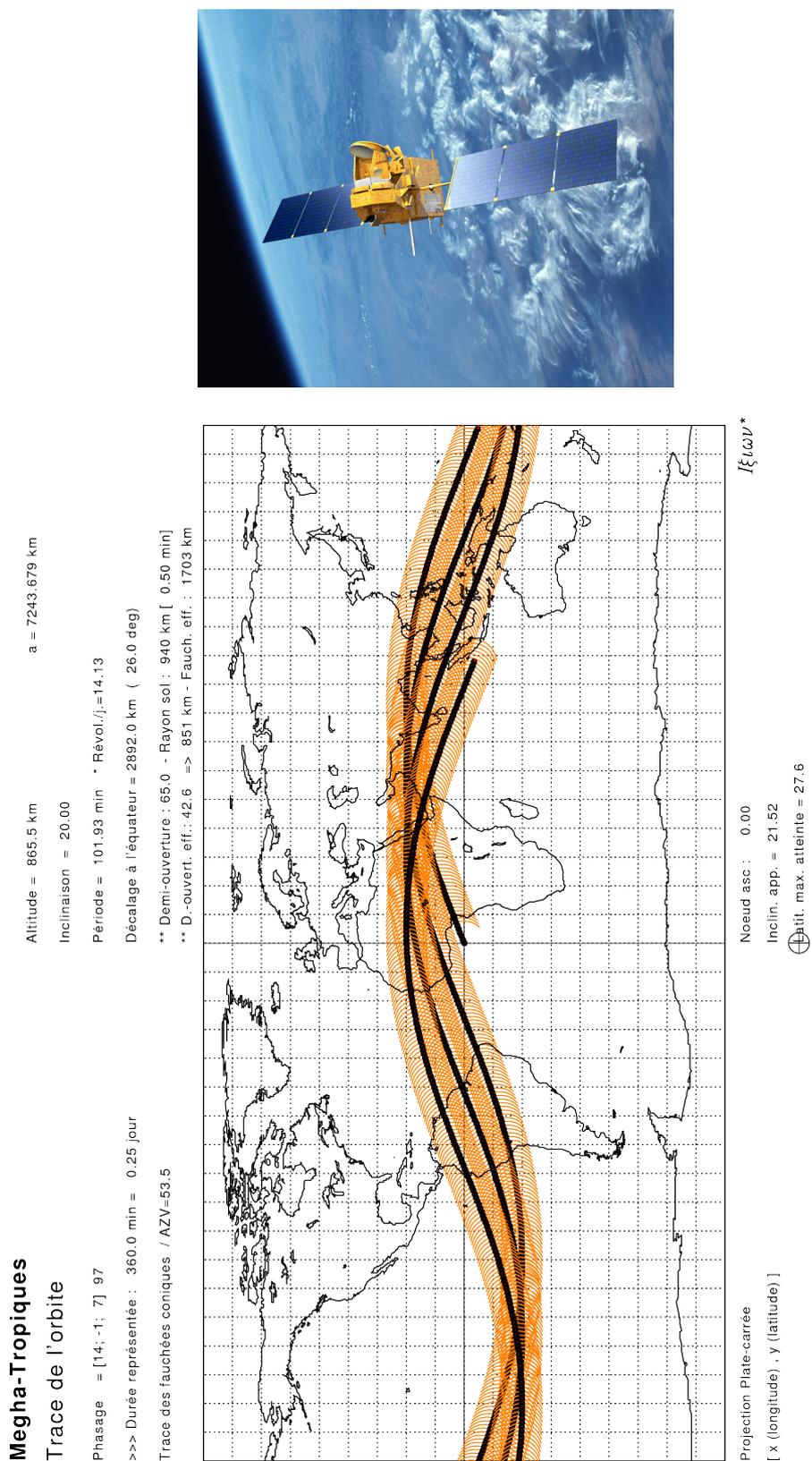


Figure 1.7 – Figure du haut : vue d’artiste du satellite. Figure du bas : idem que figure 1.5. Instrument : MADRAS, Satellite : Megha-Tropiques

Le 12 octobre 2011, le satellite franco-indien Megha-Tropiques (Roca *et al.*, 2011) est lancé, avec entre autres à son bord le radiomètre à micro-ondes passif MADRAS (fréquences d'observations comprises entre 18 et 157 GHz). La combinaison de la faible inclinaison de son orbite (20°) et de son altitude (866 km) permet entre 3 et 6 passages par jour pour un lieu donné dans la bande de latitude $[20^\circ\text{S} - 20^\circ\text{N}]$.

Les figures 1.5, 1.6 et 1.7 montrent les surfaces couvertes pour les SO SSMI/DMSP F17, TMI/TRMM et MADRAS/Megha-Tropiques pendant 6 heures consécutives.

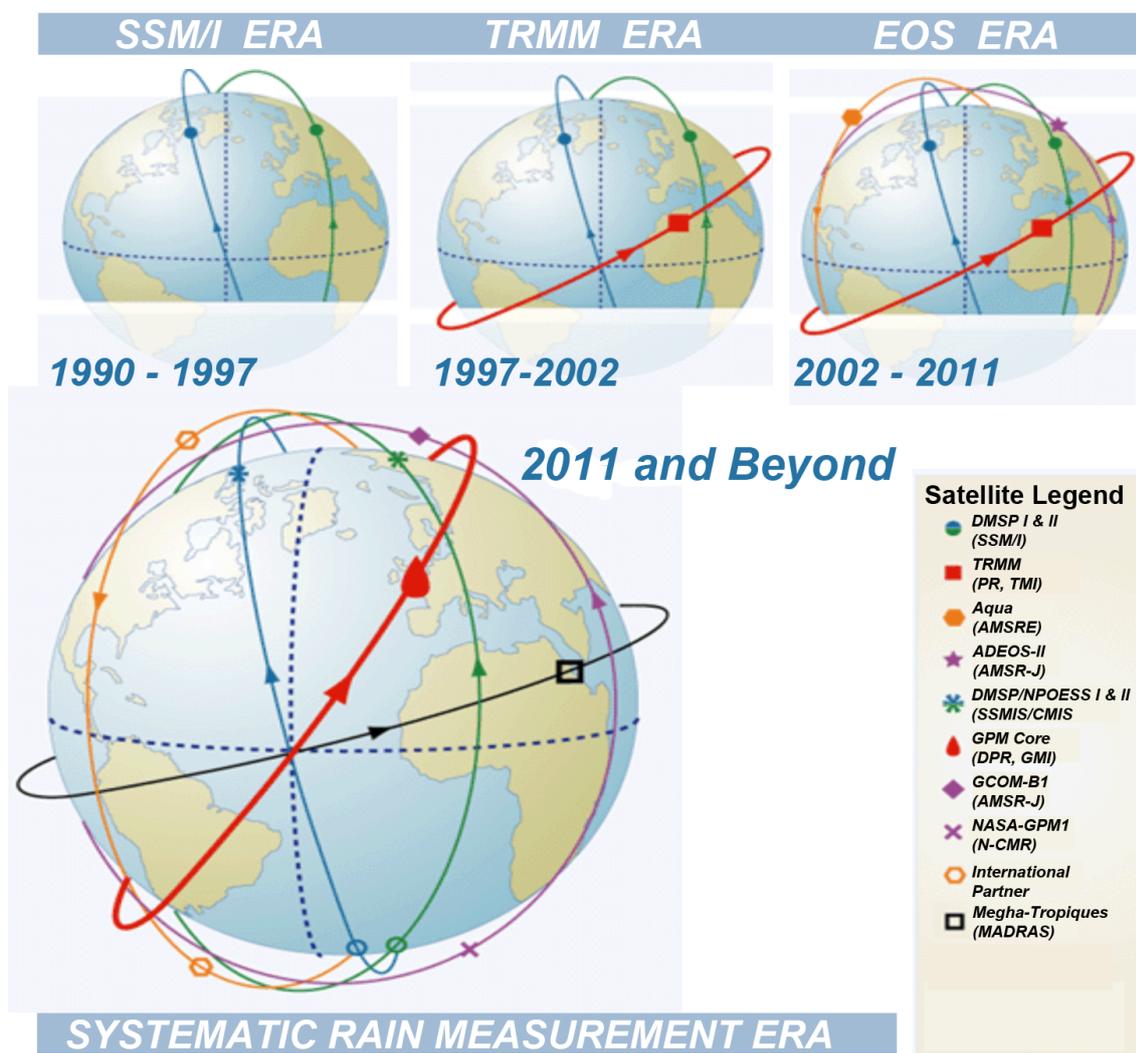


Figure 1.8 – Évolution des moyens d'observation des précipitations par radiométrie micro-ondes depuis 1990 (Source : d'après <http://pmm.nasa.gov/>)

La mise en orbite d'un nombre grandissant de systèmes d'observation permet de systématiser les mesures dans le domaine des micro-ondes sur l'ensemble des Tropiques. La figure 1.8 illustre en effet l'importante évolution des observations spatiales depuis 1990 vers une constellation de

SO dans le cadre du programme Global Precipitation Measurement (GPM) (Hou *et al.*, 2011).

1.4 Erreurs sur l'estimation des précipitations par satellite

Les méthodes d'estimation des précipitations par satellite peuvent bénéficier de chaque mise en orbite d'un nouveau SO. Cette dernière génération d'algorithmes combine en effet les mesures multispectrales (IR+MO) d'une constellation de satellites défilants et d'une flotte de satellites géostationnaires pour utiliser le maximum d'observations à disposition (Exemple : TMPA, Huffman *et al.*, 2007).

On peut citer les techniques assimilant par filtre de Kalman des observations micro-ondes issues de constellations de satellites défilants dans des systèmes de suivi d'amas nuageux. Ces systèmes de suivi d'amas nuageux utilisent eux-mêmes les mesures d'une flotte de satellites géostationnaires (Exemple : les nouvelles générations de CMORPH et GSMaP ; Joyce and Xie, 2011; Aonashi *et al.*, 2009; Ushio *et al.*, 2009). Dans certaines méthodes, ces amas nuageux sont classifiés pour un apprentissage spécifique dans l'assimilation des micro-ondes (Exemple : PERSIANN Cloud Classification System, Hong *et al.*, 2004). Plusieurs revues de ces méthodes variées ont été publiées (Kidder and Haar, 1995; Levizzani *et al.*, 2007; Kidd *et al.*, 2009; Kidd and Huffman, 2011).

La sophistication apportée à l'estimation des cumuls de pluie par les produits de dernière génération se fait souvent au détriment de la caractérisation des erreurs. En effet, plus la méthode d'estimation de cumuls se complexifie, plus le bilan d'erreurs associées est difficile à évaluer.

Les développements algorithmiques dans le cadre de production opérationnelle d'estimations de pluie ont principalement été dédiés à la minimisation des erreurs plus qu'à leur caractérisation. Cependant, il importe d'assortir une barre d'erreur à chaque cumul estimé (Kirstetter, 2008).

1.4.1 Estimation quantitative des précipitations (QPE)

On peut définir une estimation quantitative des précipitations comme étant la combinaison d'une hauteur d'eau précipitée et d'une erreur d'estimation, pour une région et pour une période donnée (QPE/QPF, 2010). Dans la suite de la thèse, la notation QPE sera utilisée pour l'acronyme anglais Quantitative Precipitation Estimation. Ainsi par exemple, on parlera de QPE satellite pour l'estimation quantitative des précipitations à partir de mesures satellites.

L'erreur quadratique moyenne d'une QPE, notée E , est composée d'un biais B et d'une incertitude ou écart type S :

$$E^2 = B^2 + S^2 \tag{1.2}$$

Les erreurs B et S peuvent être caractérisées pour un ensemble de QPE (exemple : 120 cumuls journaliers pendant quatre mois de saison des pluies). Ce sont deux erreurs de nature

indépendante : le biais d'un ensemble de QPE peut être faible mais son incertitude importante, de même que l'incertitude d'un ensemble de QPE peut être faible et son biais important (figure 1.9).

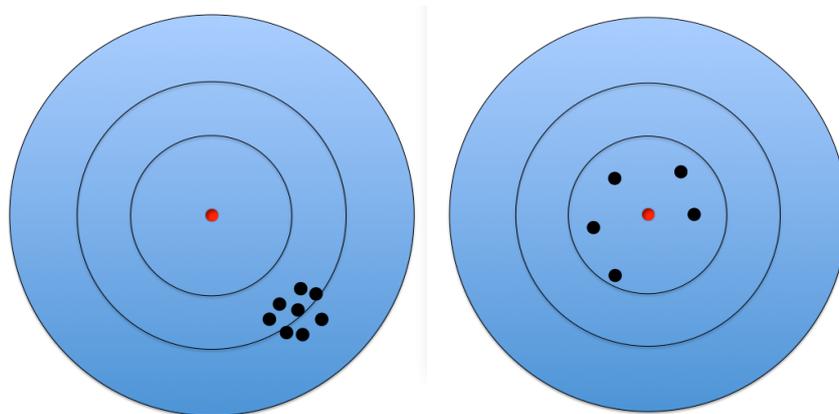


Figure 1.9 – Illustration des notions de biais et d'incertitudes d'un système de mesure pour un ensemble de QPE, par analogie avec une cible. La pluie réelle pour chacune des mesures est représentée au centre de la cible (point rouge). Les points noirs représentent des QPE associées. La figure de gauche illustre un système de mesure caractérisé par une incertitude faible mais un biais important. La figure de droite illustre un système de mesure dont le biais est faible mais l'incertitude importante.

La caractérisation des erreurs d'une QPE est complexe, quelque soit le système d'observation utilisé, car les précipitations résultent d'un phénomène atmosphérique fortement variable dans l'espace et dans le temps. Néanmoins, elle est essentielle pour leurs utilisations dans divers domaines d'applications tels que l'hydrologie, la climatologie, la météorologie, l'agrométéorologie.

En effet, les résultats de tendances climatiques issus d'analyses de QPE satellite font débat, car ces QPE sont utilisées sans prendre en compte les erreurs d'estimations (exemple : Arkin *et al.*, 2010). L'analyse de Lau and Wu (2007) montre une tendance climatique à l'augmentation de l'occurrence des cumuls faibles de pluie, cependant les QPE de faibles intensités présentent de larges incertitudes (Bell *et al.*, 2001). De même, l'analyse de Allan *et al.* (2010), mettant en évidence une augmentation de l'occurrence des cumuls de pluie les plus forts, utilise des QPE à l'échelle journalière dont la méthode d'estimation mène à de larges incertitudes et biais (Bell and Kundu, 2003). L'utilisation des QPE dans leur ensemble (cumuls et erreurs) remettrait-elle en cause ces deux tendances sur la distribution des pluies et son évolution ?

La prévision des crues nécessite l'utilisation de QPE pour le forçage de modèles hydrologiques. Pour les régions du monde équipées de radars météorologiques, les QPE sont utilisées de façon opérationnelle pour la prévision du débit des rivières, des crues, etc. Afin d'étendre ces prévisions

à des régions du globe dont les réseaux d'observation des précipitations au sol sont peu denses, un système opérationnel de prévision global des risques d'inondations a été mis en place. Ce système utilise les estimations de pluie par satellite du produit TRMM-3B42 (noté TMPA) (Huffman *et al.*, 2007) et le modèle hydrologique CREST (Coupled Routing and Excess Storage) (Wang *et al.*, 2011). Des études ont montré la nécessité de caractériser les erreurs d'estimation d'une QPE pour ces prévisions (Nijssen and Lettenmaier, 2004; Hong *et al.*, 2006; Moradkhani and Meskele, 2010). Des méthodes ont ainsi été développées afin d'exploiter les barres d'erreur des QPE dans le forçage des modèles pour réaliser des prévisions d'ensemble. Dans ces cas, la hauteur d'eau d'une QPE est interprétée comme étant la moyenne d'un ensemble de réalisations du champ de pluie et l'erreur de la QPE comme étant l'écart type de cet ensemble.

La vérification de prévisions numériques des précipitations nécessite également la prise en compte d'une barre d'erreur associée aux QPE de référence (Bowler, 2006; Candille and Talagrand, 2008; Ghelli and Santos, 2010).

Depuis 30 ans, le développement des systèmes d'observation satellitaires a permis d'établir des techniques d'estimation de QPE variées. Une approche classique consiste à estimer les hauteurs d'eau précipitées à partir de mesures satellitaires (voir annexe A). Une comparaison (ou validation) de ces estimations de cumuls de pluie avec une QPE de référence permet alors de caractériser les biais et incertitudes associés.

Ces vingt dernières années, des programmes internationaux de validation ont eu lieu grâce à des données acquises lors de campagnes de mesures pour la série des Precipitation Intercomparison Project (PIP) et des Algorithm Intercomparison Programme (AIP) (Barrett *et al.*, 1994; Ebert *et al.*, 1996; Smith *et al.*, 1998; Adler *et al.*, 2001). Depuis, l'IPWG (International Precipitation Working Group) diffuse pour les États-Unis, l'Australie, le Brésil et l'Europe, les caractéristiques d'erreurs d'estimations des précipitations de diverses méthodes à partir de mesures de référence (issues de pluviomètres et de radars).

Les biais et incertitudes d'une QPE dépendent en effet non seulement des échelles d'accumulation temporelle et d'intégration spatiale mais également des régimes de pluie, donc des régions et saisons (<http://www.isac.cnr.it/~ipwg/>). Plus la durée de la période d'accumulation est longue et la surface géographique considérée est grande (exemple : cumuls mensuels et moyennes spatiales sur des zones de $2,5^\circ \times 2,5^\circ$), plus les incertitudes sont faibles (Gebremichael and Krajewski, 2004). Pour ces échelles spatio-temporelles, on peut alors caractériser l'erreur d'une QPE à partir de comparaisons avec une autre QPE prise comme référence.

L'intercomparaison menée par Jobard *et al.* (2011) pour 3 années de saison des pluies en Afrique de l'Ouest et pour l'échelle d'accumulation de 10 jours a ainsi permis d'estimer les erreurs de 10 produits différents d'estimations des précipitations par satellite (dans la suite de ce manuscrit, on parlera également de produits satellites). Pour chacun des produits, ces erreurs sont différentes. On peut voir sur la figure 1.10 que les principales caractéristiques du champ

de référence sont reproduites par tous les produits, mais aussi que ce sont les produits qui sont calculés quasiment en temps réel qui présentent les plus larges différences avec la référence (versions "Real Time" de PERSIANN (Hsu *et al.*, 1997), CMORPH (Joyce *et al.*, 2004) et TMPA (Huffman *et al.*, 2007)). Cette étude de validation a conduit aux conclusions suivantes :

- les derniers produits présentent un très bon accord avec les QPE issues de réseaux de pluviomètres à l'échelle d'accumulation de 10 jours
- les meilleurs produits sont ceux qui combinent trois sources d'informations dans le domaine de l'Infrarouge, des Micro-ondes et des estimations de pluie issues de mesures de pluviomètres (voir annexe A et annexe B).
- les produits calculés quasiment en temps réel présentent des performances très inférieures aux meilleurs produits

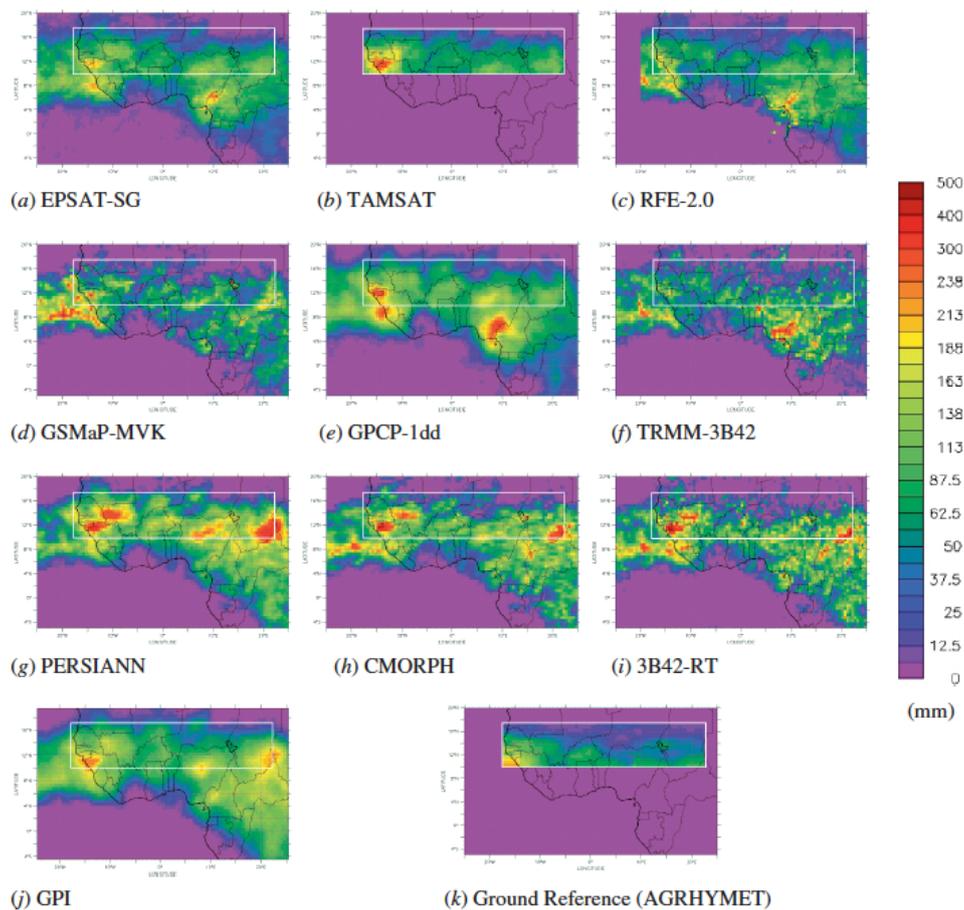


Figure 1.10 – Cumuls des précipitations pour la troisième décennie d'août 2005 sur l'Afrique de l'Ouest fournis par 10 produits satellites (Jobard *et al.*, 2011). La zone encadrée en blanc correspond à la zone du réseau AGRHYMET pour laquelle la comparaison statistique a été effectuée.

Dans le cas où chaque QPE satellite et la QPE de référence possèdent une incertitude importante, la caractérisation des erreurs nécessite une autre approche que la validation classique.

La figure 1.11 montre que la variance observée peut être forte lors d'une comparaison de deux QPE, l'une obtenue à partir d'une observation d'un SO satellitaire et l'autre obtenue à partir d'un pluviomètre. En effet, la contribution des incertitudes respectives de chaque QPE à la variance totale est supérieure à 100% du cumul, pour toutes les échelles d'accumulations spatio-temporelles présentées.

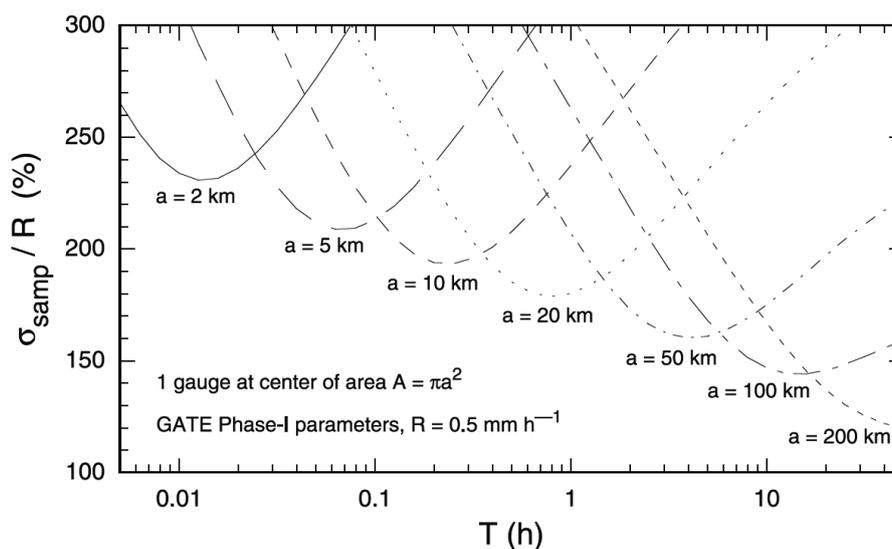


Figure 1.11 – Estimation de $\frac{\sigma_{s\text{amp}}}{R}$ l'amplitude relative de l'écart type observée lors de la comparaison de deux QPE à cause des incertitudes de chaque QPE. L'une des QPE est obtenue à partir du passage d'un SO satellitaire pendant la période d'accumulation et l'autre à partir d'un pluviomètre dans la zone géographique considérée. Les comparaisons sont effectuées pour des durées d'accumulations de 0,005 h à 50 h et pour des surfaces d'intégration spatiales de rayon a variant de 2 à 200 km. (Bell and Kundu, 2003)

Morrissey and Wang (1995) ont également montré qu'afin de distinguer les caractéristiques d'erreurs de deux produits d'estimations des précipitations mensuelles, il est nécessaire de traiter les données de référence en minimisant leurs incertitudes (méthode Non Contiguous Raingauge, Morrissey and Greene, 1993).

Des méthodes ont donc été développées afin de traiter partiellement les problèmes liés aux comparaisons de QPE dont les estimations de cumuls présentent de larges incertitudes (Morrissey and Wang, 1995; Ciach and Krajewski, 1999a,b; Bell and Kundu, 2003; Ali *et al.*, 2005b). Les erreurs des données de référence servant à la validation sont en général minimisées. Dans le cas des pluviomètres, les données sont ponctuelles. Elles sont alors spatialisées grâce à des méthodes

de block-krigeage (Journel and Huijbregts, 1978).

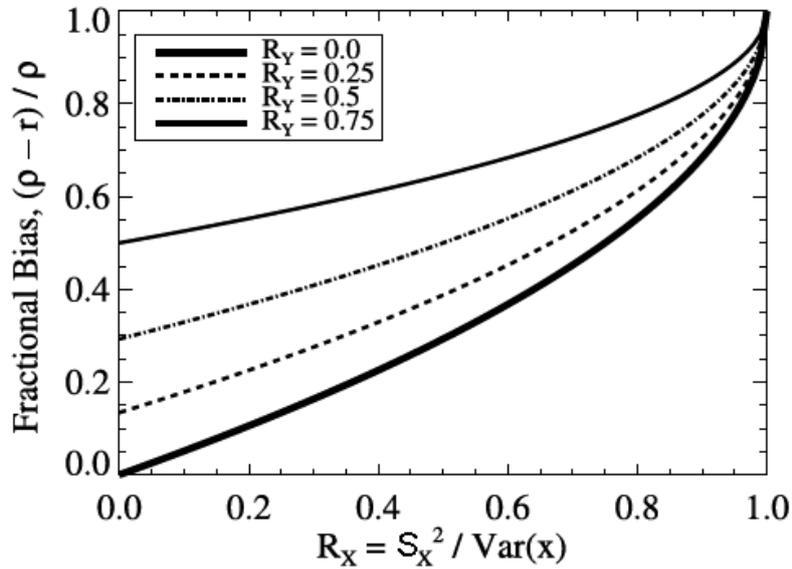


Figure 1.12 – Écart relatif entre une corrélation classique r entre deux séries d'estimations de pluie (x_i) et (y_i) et une corrélation ρ prenant en compte les erreurs associées à chaque estimation, en fonction R_x et R_y . S_x est la moyenne des incertitudes associées à chaque estimation de pluie x_i . Idem pour S_y . R_x est le rapport entre S_x^2 et la variance des estimations (x_i). Idem pour R_y . (Kelly, 2007)

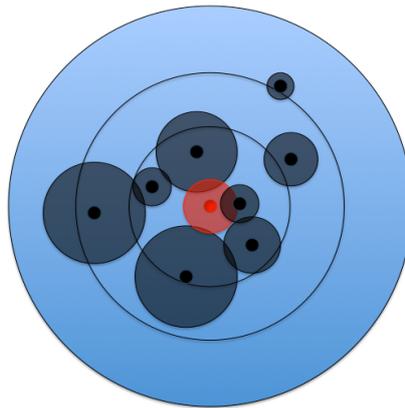


Figure 1.13 – Illustration d'une comparaison d'un ensemble de QPE dont chaque estimation est caractérisée par une incertitude importante. La QPE de référence est représentée au centre de la cible (point rouge) avec son incertitude (cercle rouge). Les QPE à valider par rapport à cette référence sont représentées par les points noirs, avec leurs propres incertitudes (cercles noirs).

Kelly (2007) montre qu'en effet il est nécessaire de prendre en compte les incertitudes de chaque estimation lors d'une comparaison si celles-ci sont importantes, sous peine de voir les résultats de la comparaison biaisés. Par exemple, la corrélation entre deux séries d'estimations peut être diminuée de près de 50 % par rapport à la corrélation réelle si R_x et R_y sont de l'ordre de 50 % (figure 1.12). Dans ce cas de figure, la caractérisation des erreurs d'un ensemble de QPE par un processus de validation ne se représente alors plus par la figure 1.9 mais par la figure 1.13. Une telle comparaison requiert un processus mathématique dédié.

Afin de répondre aux besoins d'estimations des précipitations pour l'étude du cycle hydrologique sur l'ensemble des Tropiques, il est nécessaire de produire des QPE à des échelles de temps courtes (cumuls journaliers ou sub-journaliers), dont l'erreur est caractérisée et faible (exemple pour les applications en agrométéorologie : le seuil d'utilité d'une QPE établi par la WMO est une QPE pourvue d'un biais maximum de 2 mm sur un cumul journalier (WMO, 2003)). À ces échelles de temps, les incertitudes de QPE sont suffisamment fortes pour que les processus de validation deviennent inappropriés et qu'ils nécessitent la caractérisation des incertitudes de chaque estimation afin de déduire B et S sur un ensemble de mesures. La modélisation d'erreur paraît donc être une solution en adéquation avec cette difficulté.

1.4.2 Bilan des erreurs associées à l'estimation des précipitations

Le cumul des précipitations, noté $\langle R \rangle_{A,T}$, représente l'intégration d'un champ de pluie instantanée $R(\vec{x}, t)$ sur une surface A et pendant une période T (Eq. 1.3). L'estimation de ce cumul par un SO est noté $\langle R_{Est} \rangle_{A,T,SO}$.

$$\langle R \rangle_{A,T} = \frac{1}{A} \cdot \frac{1}{T} \cdot \int_A \int_T R(\vec{x}, t) \cdot d\vec{x} \cdot dt \quad (1.3)$$

L'erreur ε associée à cette estimation vaut $\langle R \rangle_{A,T} - \langle R_{Est} \rangle_{A,T,SO}$. Cette erreur totale provient de plusieurs sources :

- l'échantillonnage (mesures ponctuelles dans l'espace pour les pluviomètres, ponctuelles dans le temps pour les instruments satellitaires, voir figure 1.14)
- l'algorithmie (méthode permettant d'estimer un taux de pluie instantané à partir des mesures de températures de brillance MO)
- la calibration des instruments

L'erreur ε peut se réécrire :

$$\begin{aligned} \varepsilon &= \langle R \rangle_{A,T} - \langle R_{Est} \rangle_{A,T,SO} \\ &= \langle R \rangle_{A,T} - (\langle R \rangle_{A,T,SO} - \langle R \rangle_{A,T,SO}) - \langle R_{Est} \rangle_{A,T,SO} \\ &= [\langle R \rangle_{A,T} - \langle R \rangle_{A,T,SO}] + [\langle R \rangle_{A,T,SO} - \langle R_{Est} \rangle_{A,T,SO}] \end{aligned} \quad (1.4)$$

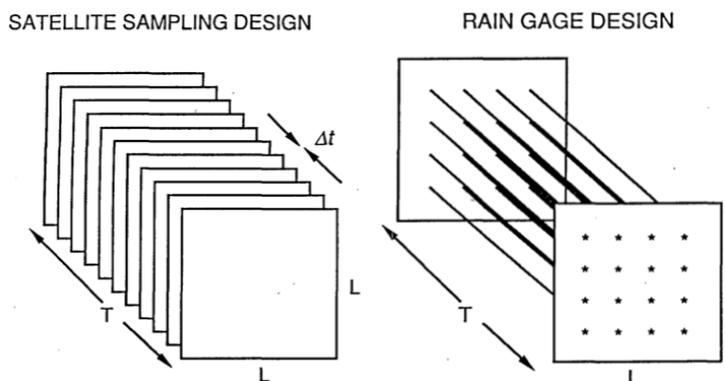


Figure 1.14 – Schéma illustrant la différence par nature des mesures pluviométriques intégrant la pluie dans le temps et les estimations par satellite intégrant la pluie dans l'espace (North and Nakamoto, 1989)

Le premier terme entre crochets de l'équation 1.4 correspond à une erreur liée à l'échantillonnage. Le deuxième terme entre crochets correspond à une erreur liée à l'algorithmie et la calibration. La variance de ε sur un ensemble de QPE, noté S^2 précédemment, est une information sur la distribution d'erreurs associée aux estimations des précipitations. Si l'on suppose les trois sources d'erreurs comme non corrélées (hypothèse dite EVS pour Error Variance Separation, Ciach and Krajewski, 1999a; Krajewski *et al.*, 2000; Gebremichael *et al.*, 2003), on peut alors écrire S^2 comme la somme de trois écarts quadratiques moyens (Eq. 1.5).

$$S^2 = S_{Echantillonnage}^2 + S_{Algorithmie}^2 + S_{Calibration}^2 \quad (1.5)$$

Sous cette hypothèse, il devient alors possible d'étudier le bilan d'erreurs d'une QPE en examinant chacun des termes avec des approches différentes, des études expérimentales ou de modélisations d'erreur.

La figure 1.15 illustre un exemple de dispositif expérimental réalisé afin d'estimer S^2 pour une QPE obtenue à partir de mesures de pluviomètres. On peut constater que la dispersion des mesures est large pour les faibles intensités de pluie, et se réduit pour les plus fortes intensités.

La modélisation des erreurs des QPE estimées par satellite est fondée sur l'utilisation d'outils géostatistiques dédiés à l'étude des pluies, hérités de l'hydrologie de surface (Grimes, 2009a). Il existe plusieurs approches statistiques pour cette modélisation. La première famille de méthodes est celle des méthodes paramétriques pour lesquelles on propose un modèle géostatistique pour les champs de pluie. Les paramètres de ce modèle sont adaptés grâce à des mesures expérimentales. La seconde famille est celle des méthodes non paramétriques basées sur des techniques purement empiriques, utilisant en général des bases de données de mesures haute résolution (mesures

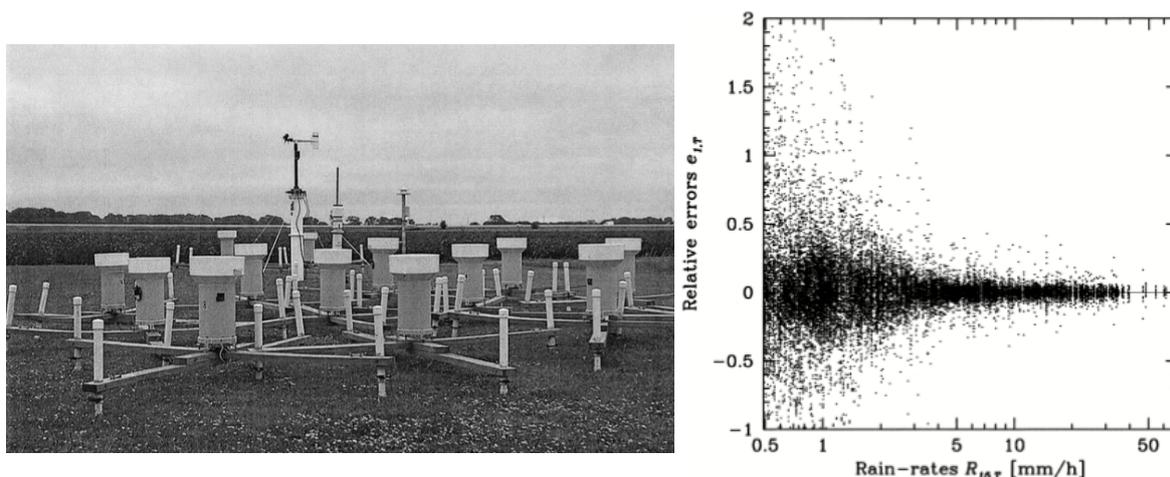


Figure 1.15 – Photographie d'un dispositif expérimental (à gauche) d'une série de 15 pluviomètres très proches les uns des autres afin d'estimer le terme S^2 pour une QPE obtenue à partir de mesures de pluviomètres. Nuage de points (à droite) montrant la dispersion des mesures des 15 pluviomètres pour un temps d'intégration de 5 minutes, en fonction de l'intensité du cumul moyen mesuré. (Ciach, 2003)

radars par exemple) qui sont ensuite sous-échantillonnées selon le mode d'échantillonnage d'un SO donné.

1.4.3 État de l'art sur la modélisation des erreurs des QPE estimées par satellite

(a) **Les modèles paramétriques** permettant d'estimer le premier terme d'échantillonnage du bilan d'erreurs, reposent souvent sur la réécriture d'une expression analytique (modèle de *Laughlin and Bell*). Cette expression décrit l'erreur d'échantillonnage temporel comme une loi de puissance, fonction de $\langle R_{Est} \rangle_{A,T,SO}$ sous des hypothèses d'homogénéité spatiale et de stationnarité temporelle des champs de pluie :

$$S_{Echantillonnage-Temporel} = a \cdot \left(\frac{1}{A}\right)^b \cdot \left(\frac{\Delta t}{T}\right)^c \cdot (\langle R_{Est} \rangle_{A,T,SO})^d \quad (1.6)$$

où a , b , c et d sont des paramètres à ajuster, A et T , sont les échelles spatiales et temporelles d'accumulation de la pluie, Δt l'intervalle de temps entre les mesures successives d'un SO et $\langle \hat{R} \rangle_{A,T,SO}$ l'intensité du cumul pour lequel on calcule l'erreur d'échantillonnage.

Dans un objectif similaire, des solutions statistiques alternatives comme la modélisation paramétrique de la covariance des champs de pluie, ont aussi été avancées, en s'appuyant sur les propriétés géostatistiques des précipitations observées pendant la campagne GATE (GARP Atlantic Tropical Experiment, <http://www.ametsoc.org/sloan/gate/index.html>). Ce modèle a permis la génération stochastique de champs de pluie à échantillonner par la méthode Monte

Carlo selon différents modes d'échantillonnage de SO (Bell *et al.*, 1990). Il a également permis d'exprimer de manière analytique l'erreur d'échantillonnage dans des conditions variées et illustrées par la figure 1.11 (Bell and Kundu, 1996, 2000, 2003).

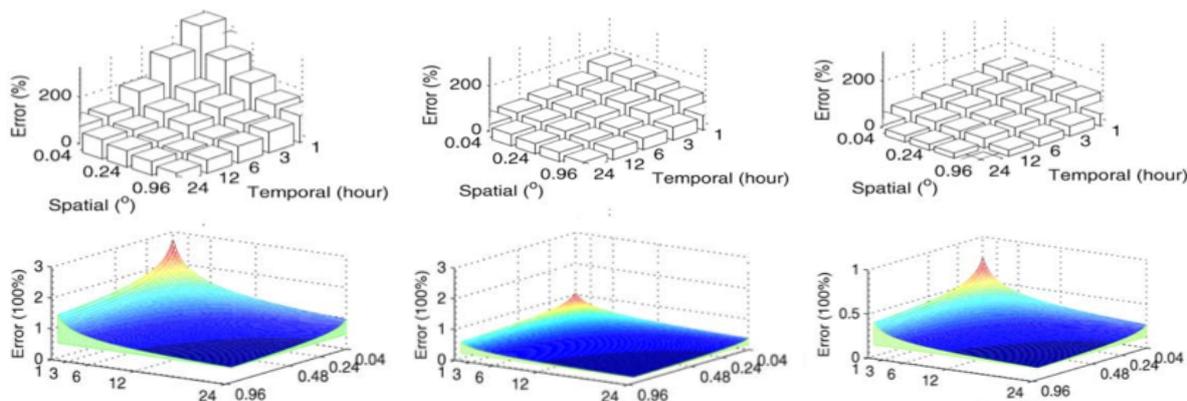


Figure 1.16 – Distributions empiriques d'erreurs d'échantillonnage relatives (ligne du haut) de QPE issues de l'algorithme PERSIANN CCS (Hong *et al.*, 2004), obtenues par comparaison avec des QPE estimées par radar, pour différentes échelles d'accumulations spatio-temporelles. Distributions modélisées d'erreurs d'échantillonnage relatives (ligne du bas) par optimisation des paramètres a , b , c et d de l'équation 1.6 afin de reproduire les distributions empiriques. Trois classes de distributions sont construites pour des taux de pluie de 0,1 mm/h (colonne de gauche), 2,0 mm/h (colonne du milieu), 4 mm/h (colonne de droite) intégrés sur différentes échelles d'accumulations (d'après : Hong *et al.*, 2006).

Le modèle *Laughlin and Bell* a été largement utilisé depuis les années 1980/1990 afin d'estimer dans un premier temps l'impact de l'orbite du satellite TRMM sur les estimations de moyenne mensuelle des précipitations (Laughlin, 1981; Steiner, 1996). Plus récemment, ce modèle a été utilisé pour quantifier l'apport de la constellation de radiomètres micro-ondes GPM (Global Precipitation Measurement mission) et pour le forçage de modèles hydrologiques (Steiner *et al.*, 2003; Gebremichael and Krajewski, 2004; Iida *et al.*, 2005; Hong *et al.*, 2006).

La mise en oeuvre de ce modèle conduit à des estimations d'erreur toujours supérieures à 100% pour les faibles valeurs de pluie ($0,1 \text{ mm.h}^{-1}$) et toujours inférieures à 50% pour les pluies les plus fortes ($4,0 \text{ mm.h}^{-1}$) (figure 1.16). La contribution de $S_{Echantillonnage}^2$ à l'erreur totale, en amplitude relative, est donc décroissante avec l'intensité du cumul de pluie.

Si les modèles paramétriques d'erreurs mentionnés ci-dessus ont permis de quantifier le terme $S_{Echantillonnage}^2$, d'autres modèles paramétriques permettent d'évaluer le terme $S_{Algorithmie}^2$ du bilan d'erreurs. Grâce à des outils géostatistiques similaires (de type modélisation du champ de covariance des champs de pluie estimés par satellite) il est possible de générer un ensemble de

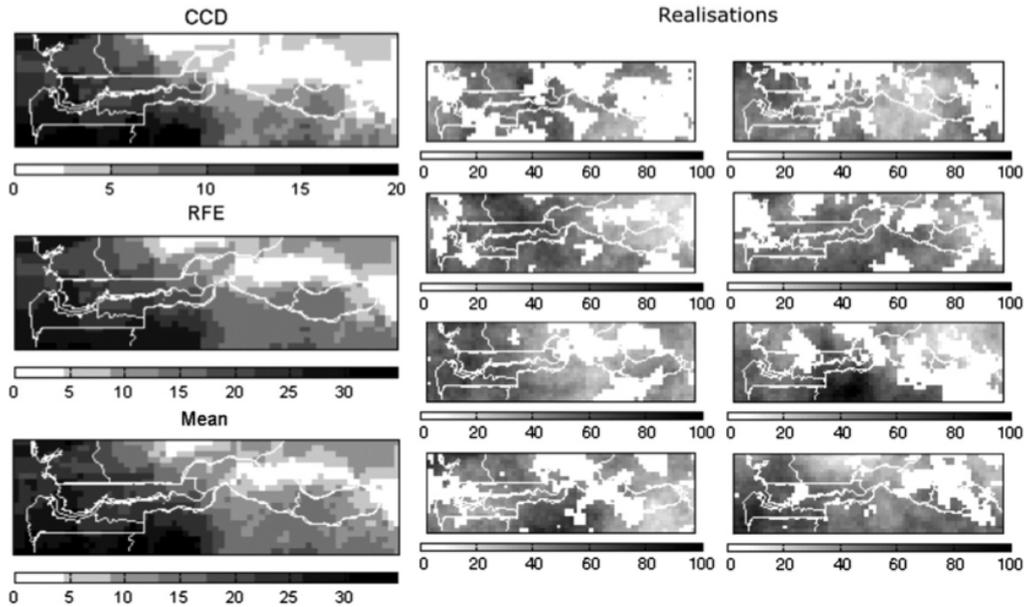


Figure 1.17 – Illustration d'un ensemble de 8 réalisations du champ de pluie TAMSIM au dessus de la Gambie pour le 7 août 2001. CCD : Cold Cloud Duration ; RFE : TAMSAT Rainfall Estimation (estimation déterministe) ; Mean : moyenne de l'ensemble des réalisations. (Teo and Grimes, 2007)

réalisations de champs de pluie (Bellerby and Sun, 2005; Teo and Grimes, 2007; Grimes, 2009a). Parmi les réalisations d'ensemble de la figure 1.17, on peut constater de l'importante variété des champs générés pour les QPE issues de la méthode TAMSAT (Grimes *et al.*, 1999; Thorne *et al.*, 2001). La position géographique du maximum de pluie se situe au niveau de la zone de nuages froids pour le champ original (Cold Cloud Duration élevé) contrairement aux 8 membres de la réalisation. Ces réalisations sont caractérisées par les mêmes propriétés géostatistiques que l'original : c'est la dispersion des réalisations qui permet d'estimer $S_{Algorithmie}^2$.

(b) Les modèles non paramétriques utilisent de larges bases de données de mesures radar ou satellitaire. Ces modèles ne font pas d'hypothèse a priori sur la distribution des erreurs. La méthode *Moving Block Bootstrap*, qui est une technique de sous-échantillonnage de champs de pluie à haute résolution, a par exemple permis d'estimer le terme $S_{Echantillonnage}^2$ spatio-temporel à des échelles d'accumulations variées.

On peut voir sur la figure 1.18 la forte dépendance de la distribution de $\varepsilon_{Echantillonnage}$ aux échelles d'accumulation spatiale et temporelle ainsi qu'à l'intervalle de temps entre deux mesures et au régime de pluie (été/hiver) (Gebremichael and Krajewski, 2004, 2005). Par exemple, dans le cas de cumuls journaliers sur une zone de 32×32 km estimés grâce à des mesures effectuées toutes les douze heures, la distribution de $\varepsilon_{Echantillonnage}$ s'étend de 300% à 500% avec une médiane à

400% en été, et de 300% à 700% avec une médiane à 450% en hiver.

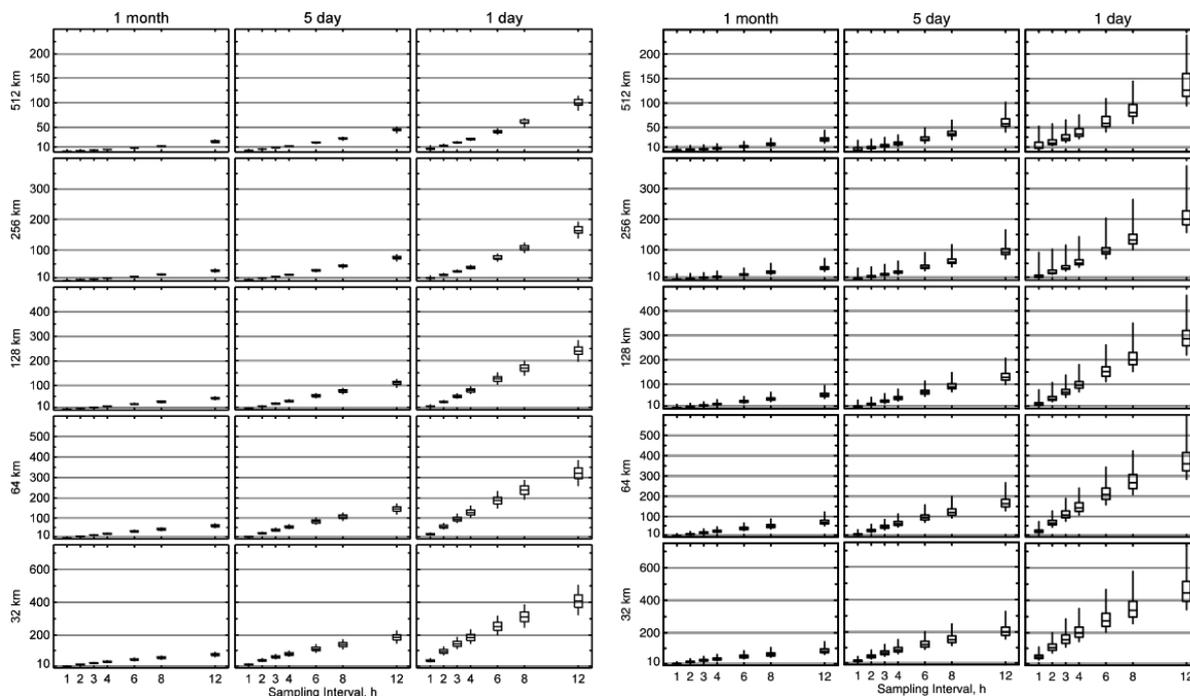


Figure 1.18 – Distributions d’erreurs d’échantillonnage relatives (en %), calculées à partir d’une méthode empirique utilisant des données radars, en fonction des intervalles de temps entre deux mesures allant de 1 heure à 12 heure. Ces distributions sont représentés pour des pluies cumulées sur des périodes de 1mois, 5 jours et 1 jour (colonnes) et pour des zones d’extension géographique de 512 x 512 km à 32 x 32 km (lignes), sur le bassin du Mississippi pour les mois de juin à août (à gauche) et Décembre à Février (à droite). (Gebremichael and Krajewski, 2004)

Plus récemment, un modèle non paramétrique a été développé dans le but de modéliser directement ε . L’objectif est de fournir pour chaque QPE estimée par satellite une distribution de valeurs probables pour la pluie réellement tombée au sol (Gebremichael *et al.*, 2011). Ce modèle prend en considération les problèmes de détection de la pluie ce qui n’est pas le cas des autres modèles cités permettant d’estimer une erreur si un cumul de pluie est détecté. La figure 1.19 montre un exemple de validation du modèle, dans le sud des États-Unis, dans le cas où l’algorithme CMORPH présente des problèmes de détection. Si les résultats de ce modèle sont très prometteurs, sa mise en oeuvre est cependant complexe car il nécessite une base de données de référence importante. En effet, les résultats de l’apprentissage du modèle à partir des données radars est difficilement généralisable à d’autres régions du globe car ils présentent déjà une variabilité importante sur les 25 blocs de $2,25^\circ \times 2,25^\circ$ étudiés.

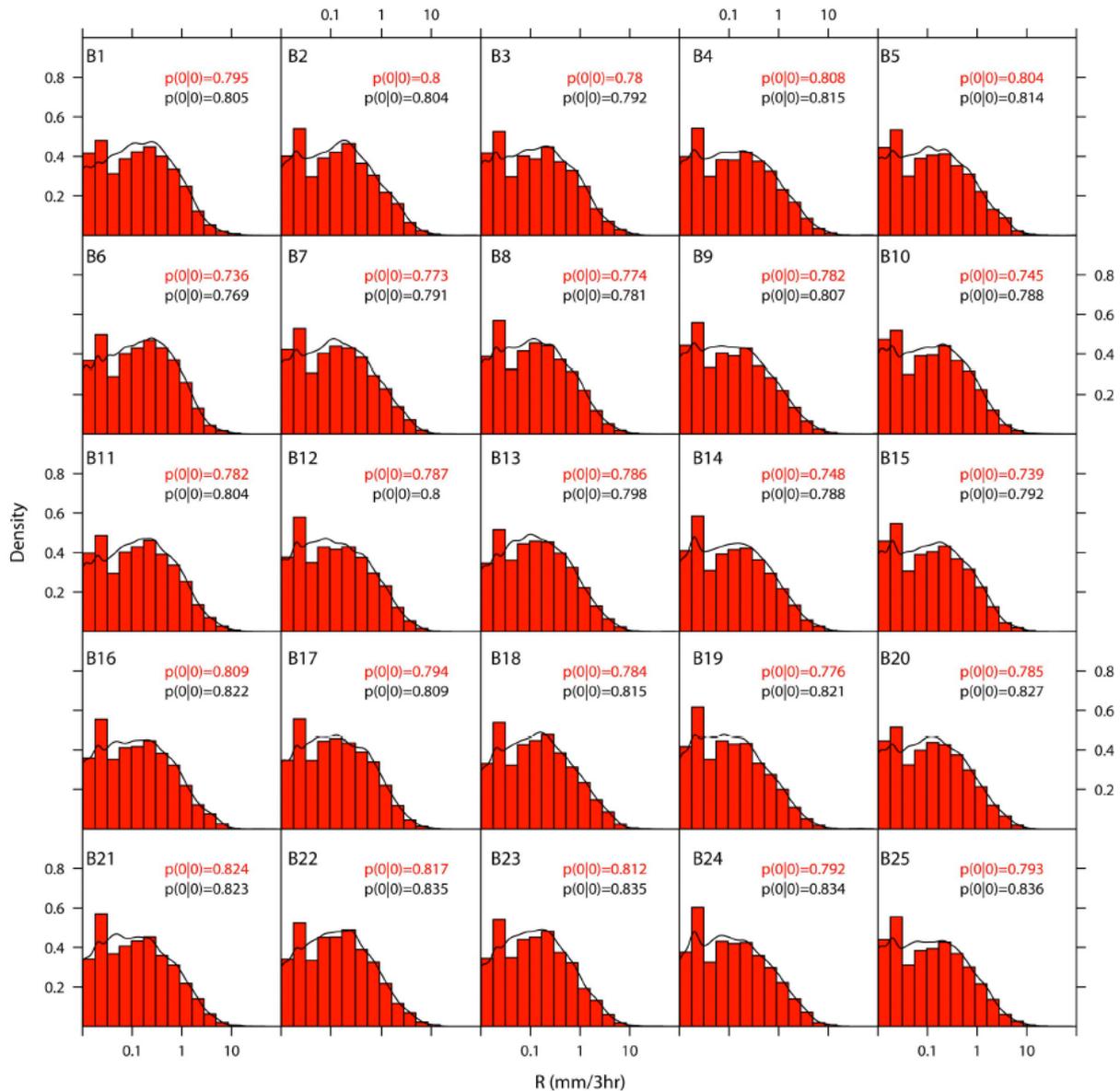


Figure 1.19 – Validation d'un modèle d'erreur non paramétrique pour l'algorithme CMORPH fournissant une estimation de $S^2_{\text{echantillonnage}} + S^2_{\text{algorithmie}} + S^2_{\text{calibration}}$. Les distributions B1 à B25 correspondent aux probabilités d'intensités de la pluie estimée par radar dans le cas où l'estimation des précipitations de CMORPH est nulle : $P(\text{Pluie Ref} / \text{Pluie CMORPH} = 0)$. Chaque distribution de B1 à B25, correspond à des zones de validation différentes de $2,25^\circ \times 2,25^\circ$ dans le sud des États-Unis, le résultat du modèle (courbe noire) est superposé aux distributions empiriques (histogrammes rouges) pour validation du modèle. Les valeurs des probabilités $P(\text{Pluie Ref} = 0 / \text{Pluie CMORPH} = 0)$ pour le modèle d'erreur (en noir) et les valeurs des probabilités empiriques (en rouge) sont aussi indiquées pour chaque zone géographique. (Gebremichael *et al.*, 2011)

1.5 Objectif et organisation de la thèse

1.5.1 Objectif

Un des rares exemples de produit d'estimation des précipitations proposant une approche intégrée aux QPE est le produit d'estimation de cumuls mensuels à la résolution de 2,5° GPCP (Global Precipitation Climatology Project) (Huffman *et al.*, 1997, 2009). Le calcul des estimations de pluie pour ce produit consiste à (i) minimiser les incertitudes grâce à une combinaison de QPE pondérées par leurs variances, (ii) minimiser le biais en assimilant les données sols du réseau de pluviomètres GPCC (Rudolf *et al.*, 2010), (iii) caractériser des incertitudes grâce à un modèle d'erreur (Huffman, 1997).

Certains produits de dernière génération utilisent l'assimilation de données sols dans le même but que le produit GPCP mais sur des périodes plus courtes que l'échelle mensuelle (Xie *et al.*, 2010). En revanche, la caractérisation des erreurs de ces derniers produits reste fondée sur des méthodes de validation classique, pourtant critique aux échelles de temps considérées. Les dernières techniques de validation proposées ne permettent de traiter qu'une partie du problème car elles se basent principalement sur la minimisation des erreurs des QPE de référence. Cependant, l'impact de ces erreurs, aussi minimisées soient-elles, est en général négligé dans le processus de validation (Ebert *et al.*, 2007). Ces méthodes de validation ne se substituent donc qu'en partie à une étude du bilan d'erreurs complet des QPE. Ce bilan comporte plusieurs sources dont il est possible d'évaluer les contributions à l'erreur totale grâce à un modèle d'erreur.

L'objectif de cette thèse est de construire une nouvelle méthode intégrée à l'estimation quantitative des précipitations sur l'ensemble de la ceinture tropicale afin de répondre aux besoins de la communauté en hydrométéorologie. Les spécifications scientifiques de cette méthode sont les suivantes :

- estimer une QPE sur l'ensemble des Tropiques à une résolution spatio-temporelle de 1°/1-jour, échelle clé pour les études météorologiques et hydrologiques. Cette échelle est intermédiaire à celle proposée pour les produits très haute résolution et celle des produits d'estimations mensuelles existants.
- caractériser les erreurs des QPE en associant une barre d'erreur à chaque estimation de cumul de pluie.

L'estimation de cumuls des précipitations à la résolution de 1°/1-jour est réalisée grâce à une méthode de combinaison d'informations IR et MO issues des mesures d'une constellation de satellites défilants et d'une flotte de satellites géostationnaires. Jobard *et al.* (2011) a montré que cette combinaison IR et MO, associée à l'assimilation de données de pluviomètres, offre les

erreurs d'estimation les plus faibles pour des échelles d'accumulation de 10 jours en Afrique de l'Ouest. L'assimilation de données de pluviomètres permet en effet de réduire les erreurs des QPE en Afrique de l'Ouest. Pour d'autres régions, Bitew and Gebremichael (2011) montrent que les erreurs de QPE satellite peuvent aussi être accentuées suite à cette assimilation. Cette approche pour la minimisation du biais dépend donc fortement de la densité du réseau de pluviomètres utilisé. De plus, des données de pluviomètres à la résolution journalière seraient nécessaires à la minimisation des biais de cumuls de pluie à la résolution de $1^\circ/1$ -jour (Joyce and Xie, 2011). De telles données n'existent pas sur l'ensemble des Tropiques.

L'utilisation d'estimations de taux de pluie par micro-ondes, caractérisé par de faibles biais, permet de minimiser les biais lors de la combinaison de données IR/MO et de s'affranchir de l'utilisation de données sols. L'algorithme BRAIN (Viltard *et al.*, 2006, 2011) est ainsi choisi pour les études menées dans cette thèse. Les premières validations de cet algorithme indiquent en effet que ses caractéristiques d'erreurs seraient compatibles avec l'estimation de QPE à la résolution $1^\circ/1$ -jour (Kirstetter *et al.*, 2011). Cet algorithme bénéficie d'améliorations grâce aux recherches menées pendant différentes campagnes de mesures (AMMA 2006, Niamey 2010, CINDY DYNAMO 2011).

La caractérisation des biais et incertitudes de QPE nécessite une étude complète du bilan d'erreurs associées. Une investigation est donc menée afin de proposer une approche intégrée à l'estimation des précipitations par satellite (cumul+erreur). Cette démarche se concrétise en la réalisation d'un des produits pluie pour la mission Megha-Tropiques.

1.5.2 Organisation de la thèse

Ce manuscrit est organisé selon le plan suivant :

- Le **Chapitre 2** propose une évaluation des produits d'estimation des précipitations par satellite issus des développements algorithmiques les plus récents. Cette évaluation est réalisée à différentes échelles pertinentes en météorologie tropicale : l'échelle saisonnière, l'échelle synoptique et l'échelle diurne. À ces échelles de temps, les méthodes de comparaison classiques présentent des limitations dues aux incertitudes sur chacune des estimations de pluie, y compris les estimations de pluie choisies comme référence. Un modèle est proposé afin d'évaluer les erreurs d'échantillonnage associées à toute estimation quantitative de cumul de pluie par satellite. Une nouvelle méthodologie de validation permet alors d'exploiter ces barres d'erreurs afin de dépasser les limites des études de validation classique.
- Le **Chapitre 3** présente une étude du bilan des erreurs associées aux estimations des précipitations par satellite. Ce bilan comporte trois termes dont une contribution à l'erreur totale est liée à l'échantillonnage, une deuxième contribution découle de l'algorithme utilisé pour l'estimation des pluies et une troisième contribution dépend de la calibration

des instruments satellitaires. La méthode Tropical Amount of Precipitation with an Estimate of ERRors (TAPEER), exposée dans ce chapitre, propose une technique d'estimation des termes d'algorithmie et de calibration du bilan d'erreurs en complément du modèle d'erreur d'échantillonnage présenté dans le chapitre 2. Dans le cadre de la mission Megha-Tropiques, la mise en oeuvre de cette approche amène au développement d'un nouveau produit d'estimation des précipitations à la résolution $1^\circ/1$ -jour appelé TAPEER-BRAIN.

- Le **Chapitre 4** a pour objectif l'étude de la sensibilité de la méthode TAPEER à l'échantillonnage des pluies par une constellation de systèmes d'observation. Des simulations sont réalisées grâce à une méthode synthétique afin de quantifier l'apport des missions nouvellement déployées telles que la mission Megha-Tropiques et la future constellation GPM aux estimations des précipitations $1^\circ/1$ -jour TAPEER-BRAIN.
- Le **Chapitre 5** présente un résumé des travaux entrepris. Une discussion est ensuite proposée sur l'importance de l'utilisation des incertitudes dans les analyses effectuées à partir d'estimations des précipitations par satellite. Cette discussion est illustrée par une première analyse de la répartition géographique des cumuls de pluie intenses sur l'ensemble des Tropiques à partir des estimations TAPEER-BRAIN. Des perspectives à ce travail de thèse sont ensuite proposées.

Deux annexes complètent ce document :

- L'annexe A présente des éléments de compréhension de la radiométrie micro-ondes pour la télédétection spatiale des précipitations. Après des rappels théoriques sur le transfert radiatif dans le domaine des micro-ondes, les caractéristiques de l'algorithme BRAIN sont présentées. Puis les difficultés liées à la résolution spatiale qui diffère pour chaque radiomètre micro-ondes d'une constellation de satellites sont montrées.
- L'annexe B est consacrée à la description de l'algorithme TAPEER dont les bases théoriques ainsi que sa réalisation technique et sa validation préliminaire sont détaillées.

Une nouvelle méthode d'évaluation des produits satellites aux échelles pertinentes en météorologie

La mousson est un élément central du climat en Afrique de l'Ouest. L'eau précipitée pendant cette période apporte une contribution majeure aux ressources annuelles disponibles dans ces régions. La mousson africaine est caractérisée par des phases actives, et des pauses dans les précipitations. Il est essentiel de comprendre la variabilité intra-saisonnière des pluies car l'intensité de la mousson représente une source de préoccupations importantes pour les populations ouest-africaines.

Comme il a été décrit précédemment, l'observation des pluies nécessite une densité de mesures élevée qui n'est disponible qu'à partir d'observations spatiales pour l'ensemble de l'Afrique de l'Ouest. L'utilisation des produits d'estimations des précipitations par satellite se révèle être pertinente pour une meilleure compréhension de la variabilité des précipitations pendant les périodes de mousson.

Dans ce but, il est nécessaire d'évaluer la qualité des estimations de cumuls de pluie par satellite en caractérisant leurs erreurs dans ces régions. Différents produits d'estimations des précipitations par satellite sont validés de manière régulière afin de caractériser leurs erreurs sur différentes régions du globe. En revanche, peu d'efforts de validation ont été menés en Afrique de l'Ouest.

L'objectif de ce chapitre est de valider (ou d'invalider) l'usage des derniers produits d'estimations de pluie par satellite en Afrique de l'Ouest, à des échelles pertinentes en météorologie tropicale : l'échelle saisonnière, l'échelle synoptique et l'échelle diurne. Dans ce chapitre, une investigation est menée à partir de quatre réseaux de pluviomètres dont les données ont été récoltées en 2006 pendant la campagne AMMA (Analyse Multidisciplinaire de la Mousson Africaine).

Aux échelles considérées dans ce chapitre, la caractérisation des erreurs des estimations de pluie par satellite nécessite une méthodologie dédiée, différente des validations classiques comme introduit dans le chapitre 1. En effet, les estimations des précipitations, par satellite ou mesurées grâce à des réseaux de pluviomètres, présentent des incertitudes fortes à ces échelles. Dans ce cas, la comparaison d'estimations de pluie requiert la connaissance préalable de leurs propres

incertitudes. Les incertitudes associées à la spatialisation des mesures ponctuelles de pluviomètres peuvent être calculées à partir d'une technique de block krigage. Les incertitudes associées aux estimations des précipitations par satellite sont évaluées grâce à un nouveau modèle d'erreur introduit dans ce chapitre. Une méthode de validation dédiée est alors développée afin de prendre en compte ces incertitudes dans le processus de comparaison.

Le travail présenté ci-dessous a fait l'objet d'une publication dans *Journal of Applied Meteorology and Climatology* pour la Special Issue de l'International Precipitation Working Group :

Roca, R., Chambon, P., Jobard, I., Kirstetter, P.-E., Gosset, M., and Bergès, J.-C. (2010). **Comparing satellite and surface rainfall products over west africa at meteorologically relevant scales during the amma campaign using error estimates.** *J. Appl. Meteorol. and Climatol.*, **49**(4), 715-731. doi : 10.1175/2009JAMC2318.1

2.1 Introduction

The West African monsoon (WAM) brings the main part of the yearly rainfall over West Africa, including Sahel (Hastenrath, 1991), and is at the heart of the water resources of this part of the continent (Conway *et al.*, 2009). The functioning of the monsoon and of the rainfall variability has hence received a lot of attention, especially since the advent of a long-term drought over the Sahel in the 1970s to the 1990s (Barbé *et al.*, 2002; Ali and Lebel, 2008; Giannini *et al.*, 2008b). Its future evolution is also of concern in the context of the global climate change (e.g., Giannini *et al.*, 2008a). The need for a deeper understanding and forecasting capability of the WAM prompted the community to devote a vast observational program over the region, the African Monsoon Multidisciplinary Analysis (AMMA ; Redelsperger *et al.*, 2006) ; the data from the AMMA campaign are used in this study.

The main feature of the seasonal march of the monsoon is the rapid onset occurring in late June and characterized by a northward jump of the ITCZ (Sultan and Janicot, 2003) showing a strong modification of the rain regimes and synoptic-scale variability (Gu and Adler, 2004). The day-to-day variability of rainfall alternates between active and inactive phases at the intraseasonal scale with modes at 10-25 days and around 40-50 days, respectively (Sultan *et al.*, 2003). At shorter synoptic time scales, variability of the rainfall is also observed and linked to tropical wave dynamics (e.g., Gu *et al.*, 2003; Mounier *et al.*, 2007) and extratropical intrusions of dry air (Roca *et al.*, 2005). The most documented mode of synoptic variability is associated with the African easterly waves (AEW) and was recently fully redocumented (Kiladis *et al.*, 2006). These 3-5-day waves are modulating strongly the rainfall during the monsoon through their relationship to organized convective systems (Machado *et al.*, 1993). Finally, as observed on most tropical continents, the diurnal cycle of insolation drives a significant portion of the overall meteorological variability over the WAM (e.g., Desbois *et al.*, 1988). Redelsperger *et al.* (2002) systematically

investigated the 1992 wet season in West Africa and detailed the previously mentioned various scales for that very year. They underscored that such a multiscale approach was an encouraging method to deepen our understanding of the interaction of the various processes occurring in the WAM. Hence, further understanding of the WAM requires in-depth analysis of this multiscale variability of rainfall.

Satellite observations are a powerful tool to cover these scales and to be used for these much needed meteorological investigations over the WAM where the pluviograph network is scarce. The recent generation of combined infrared (IR) and microwave (MW) products (Hsu *et al.*, 1997; Herman *et al.*, 1997; Huffman *et al.*, 2001; Joyce *et al.*, 2004; Ushio *et al.*, 2009; Huffman *et al.*, 2007; Levizzani *et al.*, 2007; Bergès *et al.*, 2010) especially fits very well to such endeavor, but their hard-to-tell ability to provide useful information is often underscored. Although a number of studies investigated the quality of these satellite estimates in various regions of the world (e.g., Negri *et al.*, 1995; Ebert *et al.*, 1996; Ebert and Manton, 1998; Joyce *et al.*, 2004; Huffman *et al.*, 2007; Xie *et al.*, 2007; Ebert *et al.*, 2007; Zeweldi and Gebremichael, 2009b; Sapiano and Arkin, 2009, ; see also the International Precipitation Working Group (IPWG) Web site, which is available online at <http://www.isac.cnr.it/~ipwg/IPWG.html>), few addressed the West African climate. These rare studies focused on accumulated time scales such as 10-day and monthly scales (Laurent *et al.*, 1998; Nicholson *et al.*, 2003a,b; Ali *et al.*, 2005b; Lamptey, 2008). Jobard *et al.* (2011) recently provided 10-day scale intercomparisons of 10 products using 3 yrs of data over the Sahel. Their results indicate good enough agreements at this scale to conduct further investigation at shorter time scales in link with the meteorology of the region.

Here, we provide a comparison of a representative subset of IR-MW satellite and surface rainfall products over West Africa, at meteorologically relevant scales, during the AMMA campaign (2006 rainy season). Our aim is to validate (or not) the use of the satellite rainfall estimates for physical investigation of the monsoon. The 10-day scale before and after the onset, the daily scale over the full course of the monsoon, the 3-5-day spectral band before and after the onset, and the diurnal scale before and after the onset are documented as a meteorological benchmark to establish the usefulness of the satellite rain products. A dedicated statistical approach is proposed that relies on the use in the comparisons of the errors associated to the areal mean rainfall estimates for both gauges and satellite products. This novel methodology is applied at the previously mentioned scales over the WAM region and over three dedicated sites that were adequately instrumented during the AMMA campaign (Lebel *et al.*, 2009, 2010).

The paper is organized as follows : first, ground-based and satellite data are both introduced along with their respective error budget in section 2.2. Section 2.3 quickly presents the statistical procedure, taking into account the errors used for the comparison. Section 2.4 details the results of the confrontation for each scale of our meteorological benchmark. A summary and discussion are reported in section 2.5.

2.2 Data

2.2.1 Rationale

Over the mentioned range of scales, classical (without errors) evaluation techniques (for an insightful review, see Ebert, 2007) fall short as the errors in both the gauges, and the satellite products can become significant and can corrupt the evaluation. For instance, the measurement errors significantly attenuate the observed correlation, making the observed correlation between two series weaker than if one does not account for the measurement errors. Kelly (2007) shows that, when the measurements errors make up to 50% of the variance on each series, the coefficient of correlation is reduced by roughly 50%. Such an example calls for a careful handling of the errors in running the comparisons. The estimation of the errors in both the gauges and the satellite rainfall estimates is a complex task. It is therefore usually left aside by the investigators in intercomparison exercises (Ebert, 2007). Such efforts have nevertheless been undertaken at different time and space scales for both satellite measurements over different parts of the world (e.g., Huffman, 1997; Ali *et al.*, 2005b) and gauge measurements (e.g., Lebel *et al.*, 1987; Morrissey *et al.*, 1995; Ali *et al.*, 2005a; Grimes *et al.*, 1999; Pardo-Iguzquiza *et al.*, 2006; Teo and Grimes, 2007). In short, the error budget of the satellite and gauges products can be written as

$$S^2 = S_{calibration}^2 + S_{algorithm}^2 + S_{sampling}^2 \quad (2.1)$$

where the calibration term is associated with systematic errors and relevant to the accuracy of the products. The algorithm term is associated with the intrinsic errors of the retrieval, and the sampling term refers purely to the error resulting from the discrete nature (in space or time) of observations; both terms are relevant to the accuracy and the precision of the products. It is difficult to separate both latter terms, because the precision and accuracy of retrievals depends on the scale of interest and hence the sampling term is sometimes considered as a mixture of the two (Huffman, 1997).

Rainfall estimation by gauges is subject to a very low measurement error (e.g., Ciach and Krajewski, 1999a) because of the direct nature of this measurement. On the other hand, because of the punctual nature of gauge measurement, the sampling error when computing areal mean rainfall cannot be neglected. The algorithm-based errors can be assumed to be small compared to the sampling errors of the gauges (Habib *et al.*, 2001), even if uncertainty can be significant at small rain rates and accumulation scales (Ciach, 2003). Satellite products are not affected by a calibration uncertainty of a few Kelvin's for the geostationary IR and low earth orbit satellite (LEO) microwave brightness temperatures (Jobard and Desbois, 1994; Viltard *et al.*, 2006), and we hence neglect it in the error budget. The algorithm term is difficult to estimate and is usually considered to be small compared to the sampling uncertainty. Gebremichael and

Krajewski (2005), based on previous studies, assumed that sampling errors indeed dominate the error budget of satellite products and proposed a parametric model of the sampling error distribution of rainfall depending on the mean rainfall from radar measurements from monthly down to daily scale over a variety of space scales. The error budget for both ground-based and satellite-based rain products hence finally reduces to

$$S^2 \sim S_{\text{sampling}}^2 \quad (2.2)$$

The sampling term error computations are performed using two different approaches for the gauges and satellite rain estimates, even if the techniques are closely related to each other. Errors are evaluated for each areal-averaged estimate at various scales.

2.2.2 The ground-based rain products

2.2.2.1 Rain gauge networks

Four rain gauge networks are used for this study. The characteristics of the areas and localization of the networks sites are presented in Table 2.1, and their gauge distributions are shown in Figs. 2.1 and 2.2. The Comité Permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel (CILSS) rain gauge network, spread over the Sahelian band (Fig. 2.1), included during the 2006 rainy season about 570 rain gauge stations and 75 synoptic stations of the Global Telecommunication System (GTS). The GTS stations are not used for this study because they are used in some of the satellite products. The daily rain data are aggregated at a 10-day time resolution for regional comparison with satellite estimations.

Name	Localization	Area	Gauges
CILSS	10.0°-17.5°N, 17.5°W-22.5°E	4 000 000 km ²	570
Niamey	13.0°-14.2°N, 1.4°-3.0°E	25 000 km ²	54
Ouémé	9.0°-10.0°N, 1.5°-2.8°E	15 400 km ²	52
Dakar	14.4°-14.7°N, 16.9°-16.6°W	1200 km ²	21

Tableau 2.1 – Characteristics of the rain gauge networks.

The three other networks cover smaller areas located in Niger, Benin and Senegal. The Benin and Niger networks are part of the AMMA Coupling the Tropical Atmosphere and the Hydrological Cycle (AMMA-CATCH) observing system (Lebel *et al.*, 2009). The distribution of gauges of the network around Niamey was optimized (Lebel and Amani, 1999) to obtain a good accuracy for rainfall estimation over the area at the rain event scale (typically a few hours); the number of tipping-bucket gauges was increased up to 54 for the AMMA enhanced observing period (EOP) in 2006. The upper Ouémé catchment in Benin was instrumented with 52 tipping-bucket gauges

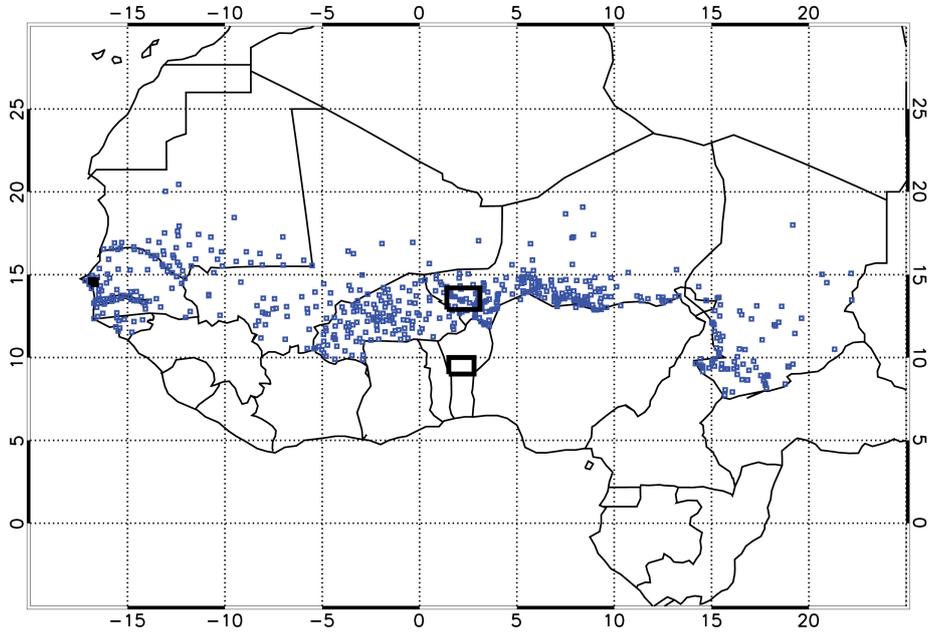


Figure 2.1 – CILSS area rain gauge network distribution (Ali *et al.*, 2005a)

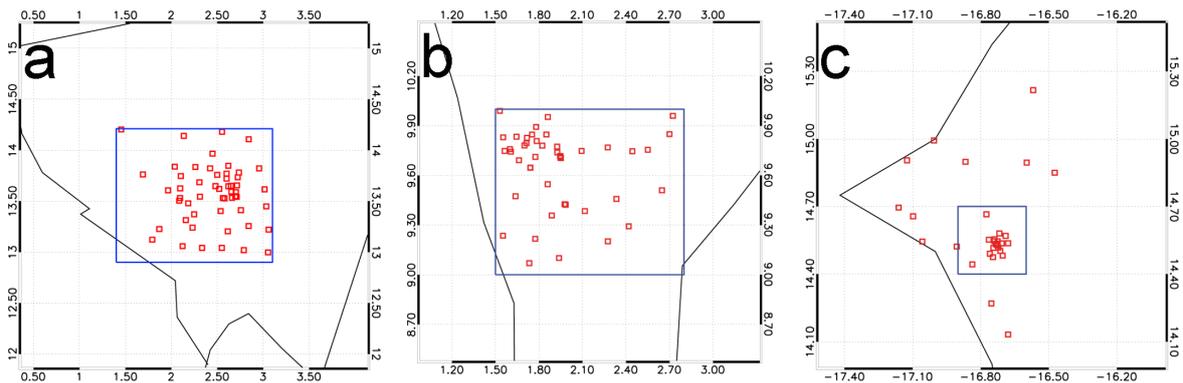


Figure 2.2 – Rain gauge distributions of the three dense networks of (a) Niamey, (b) Ouémé, and (c) Dakar. The squares represent the area selected for the kriging interpolation. Black lines are the borders of Niger, Benin and Senegal.

at stations that are evenly distributed on a surface of about 15 000 km². This configuration is suitable for the mesoscale analysis of the rainfall associated with tropical convective systems. Both the Ouémé and Niamey networks have a density of about 1 gauge per 200 km² over a 1° × 1° area. The Dakar network consisted of about 35 stations unevenly distributed over a 1° × 1° area (Jenkins et al. 2010, unpublished manuscript ; Jenkins and Gaye, 2010) ; only a subsample of the network on a small, denser area (0.3° × 0.3°) that includes 21 gauges (see Fig. 2.2c) available for two months (August-September 2006), a density of about 1 gauge per 50 km². The gauge data from these three networks are computed at 3-h and daily time steps and also aggregated into mean seasonal diurnal cycle.

2.2.2.2 Products and associated error computation

Various techniques of rain gauge point value interpolation have been developed to estimate not only area-averaged ground rainfall accumulation but also the sampling errors associated to the interpolation of the point values over their surrounding spatial domain. Among those techniques, linear estimators are widespread used interpolators, and their general expression for $R_{ref}(A, T)$ area rainfall estimate over a spatial domain A and a period of accumulation T is

$$R_{ref}(A, T) = \sum_{i=1}^{N_g} \lambda_i R_g(x_i, T) \quad (2.3)$$

where λ_i represents weights on the accumulations of the i th rain gauges $R_g(x_i, T)$ and N_g represents the total number of rain gauges available. Morrissey and Wang (1995) detail the computation of an unweighted linear interpolation of rain gauge accumulations and compare sampling errors associated to various geometries of networks. Here, the ordinary block-kriging technique (Journel and Huijbregts, 1978) is used to estimate areal rainfall estimates and their respective sampling errors, first for 10-day accumulations over 18 grid boxes in the Sahel (Ali *et al.*, 2005a) and then for daily and 3-hourly accumulations and for a 3-hourly seasonally cumulated diurnal cycle over the three sites of Niamey, Ouémé, and Dakar. By weighting the rain gauges individually, the characteristic of the kriging estimator is unbiased and the estimation variance is minimized (Lebel and Amani, 1999). The kriging technique relies on a structure function of the rain field γ called the variogram. We take into account information from all the realizations to infer a single and robust climatological variogram (Lebel *et al.*, 1987). The estimation variance, corresponding to the sampling error, allows assessment of the estimation quality as function of the spatial structure of the rain field and the relative position of the rain gauge network versus the geographical support of interest A . The estimation variance requires the numerical evaluation of two integrals of the variogram function over the domain A (Journel and Huijbregts, 1978). Its expression is

$$\begin{aligned}
Var[R_{ref}(T, A) - R(T, A)] = & - \frac{1}{a^2} \int_A \int_A \gamma(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}' \\
& + \sum_{j=1}^{N_g} \lambda^j \left[\frac{1}{a} \int_A \gamma(\mathbf{x}, x_j) d\mathbf{x} \right] \\
& - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \lambda^i \lambda^j \gamma_{ij}
\end{aligned} \tag{2.4}$$

where $R(T, A)$ is the true area-averaged rainfall accumulation for the period T and area A , \mathbf{x} is the location vector and x_j denotes the location of the gauge j , a is the area of the domain A , and $\gamma(\mathbf{x}, \mathbf{x}')$ is the variogram function computed on the Euclidean interdistance between the two locations \mathbf{x} and \mathbf{x}' . The first term denotes the block-to-block computation of the variogram function over the domain A , with \mathbf{x} and \mathbf{x}' belonging to A ; the second term corresponds to the computation of the variogram function between each gauge and the domain A ; and the last term is the in-between gauge variogram computation. The estimation variance depends on the variogram characteristics, which are related to the rain field variance. The spatial structures and the variances of rain fields are linked to the integration time period considered (Lebel *et al.*, 1987; Berne *et al.*, 2004; Ali *et al.*, 2005a).

2.2.3 The satellite-based rain products

2.2.3.1 The satellite rain estimates

EPSAT-SG The Estimation of Precipitation by Satellites-Second Generation (EPSAT-SG) method has been developed at Laboratoire de Météorologie Dynamique in the framework of the AMMA research program (Bergès *et al.*, 2010). It consists in downscaling the Global Precipitation Climatology Project one-degree daily (GPCP-1DD) product to the Meteosat Second Generation (MSG) spatial and temporal resolutions (3 km and 15 min). It combines a rainfall potential intensity and a rainfall probability derived, because of a feed forward neural network, from the MSG multispectral channels and Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (PR) information. The product coverage extends from 5°S to 20°N and from 25°W to 25°E, and it is referred to as EPSAT.

GSMAP The Global Satellite Mapping of Precipitation (GSMAP) algorithm is inspired by a morphing technique (Joyce *et al.*, 2004) combining passive microwave data from the current fleet of LEO satellite microwave radiometers with cloud motion vectors and brightness temperature from geostationary infrared images. GSMAP covers the globe from 60°N to 60°S at the baseline resolution of 0.1°×0.1° and 1 h. A description of the methodology can be found in Ushio *et al.* (2009) and Aonashi and Liu (2000).

TMPA The TRMM multiplatform algorithm (TMPA) precipitation dataset is an operational product of the TRMM mission (Huffman *et al.*, 2007). It spans the latitudes from 50° N to 50° S at the resolution of 0.25° × 0.25° and 3 h. It combines precipitation estimates from multiple satellite passive microwave imagers [Special Sensor Microwave Imager (SSM/I), TRMM Microwave Imager (TMI), Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), and Advanced Microwave Sounding Unit-B (AMSU-B)] converted to precipitation estimates using the Goddard profiling algorithm (GPROF; Kummerow *et al.*, 2001), as well as microwave-calibrated infrared precipitation estimate data from the operational fleet of geostationary meteorological satellites. The Global Precipitation Climatology Project monthly rain gauge analysis produced by the Global Precipitation Climatological Center is used for a monthly rescaling (Huffman *et al.*, 1997).

2.2.3.2 The error model

The evaluation of the satellite products sampling errors when averaged at some scales from instantaneous estimates is a difficult task. Several approaches have been explored, all involving the modeling of the spatial covariance functions of the rain field or its autocorrelation function. Direct estimation using the integrated estimates of the satellite rain retrievals and the "true" rain fields (Bell *et al.*, 1990; North and Nakamoto, 1989) or derivation from the discrepancy between the gauges data and the true rain field (Ali *et al.*, 2005b) are among those. North and Nakamoto (1989) developed a spectral formalism to derive a general expression of the sampling errors, which can be interpreted in the Fourier space-time domain as the covariance rain field filtered by a function depending on the rain field sampling scheme. This expression reduces to the well-known expression of the uncertainty on the estimation of the mean of independent data in cases of simple sampling scheme,

$$S = \frac{\sigma}{\sqrt{N_{ind}}} \quad (2.5)$$

where σ is the standard deviation of the samples and N_{ind} is the number of independent data (e.g., Larsen and Marx, 2001). Bell *et al.* (1990) used a stochastic and parametric model of the covariance field with Global Atmospheric Program Atlantic Tropical Experiment (GATE) data to estimate the uncertainties expected at monthly scale and various spatial scales. Bell and Kundu (2003) further used this model to compare different sampling schemes of satellite and rain gauge measurements to optimize network design. The error model developed in Morrissey and Wang (1995) is close to a kriging technique and uses a variance reduction factor similar to Eq. (2.4) but without weights, which collapses to the Eq. (2.5) if the correlation among the samples is zero. The modeling of sampling errors in the case of combined satellite rainfall estimation products become even more complicated because the sampling scheme depends on which data sources are

combined but also on the way the algorithm itself combines the data. Huffman (1997) developed an error function at the monthly scale, following (North and Nakamoto, 1989), based on the computation of a variance of the observations and the evaluation of a number of independent samples. Recently, ensemble methods have been used to estimate satellite rainfall uncertainty (Bellerby and Sun, 2005; Teo and Grimes, 2007).

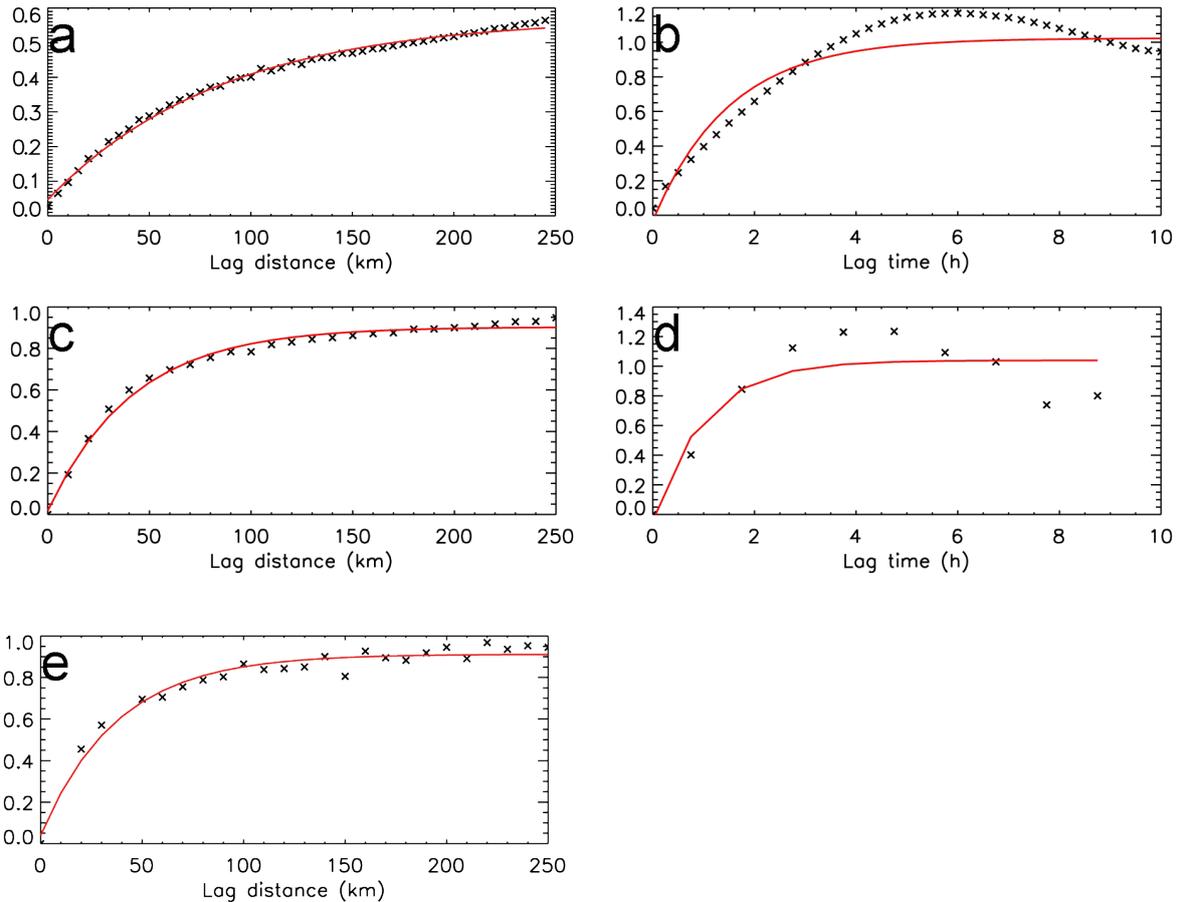


Figure 2.3 – (left) Spatial and (right) temporal variograms calculated over the $5^\circ \times 5^\circ$ window (9° - 14° N, 0° - 5° E) for (a),(b) EPSAT-SG ; (c),(d) GSMAP ; and (e) TMPA. The \times 's correspond to the calculated values of the variogram functions, and the lines correspond to the fitted exponential model.

Here, a simple error model is developed with a similar approach : it is assumed that the variance of the true precipitation field could be computed as the variance S^2 of the product data for a given area A and during an integration period T . The number of degrees of freedom is estimated independently of the native resolution of the products and depends on the level of autocorrelation of the product data. Over an area A and during a period T , it is defined as

$$N_{ind} = \frac{A T}{d^2 \tau} \quad (2.6)$$

where d and τ are the e-folding space and time distances, respectively, of a structural function computed using the same variogram function as the one used in section 2.2.2.2 assuming isotropy. The variogram functions of the satellite products is calculated over a $5^\circ \times 5^\circ$ region (9° - 14° N, 0° - 5° E) including both the Niamey and Ouémé sites and instantaneous estimations over the whole season are used. Sensitivity tests have shown very little dependence of the calculations to the selected window. An exponential model is fitted to the average variogram to derive the space and time e-folding distances shown in Fig. 2.3. The e-folding distance is 37 km for TMPA, 41 km for GSMAP, and 85 km for EPSAT (Figs. 2.3 a,c,e). The temporal variograms are computed for GSMAP and EPSAT only and yield e-folding times of 1 and 1.5 h, respectively (Figs. 2.3 b,d). It is not computed for TMPA because its 3-hourly estimates are considered to be independent.

2.2.4 Error computation results

The error models provide an estimation of errors at each time step of space-time-averaged satellite rainfall estimates. The time-averaged relative errors of the three satellite products along with the rain gauge are reported in Table 2.2 for each of the considered scales. The present error estimates are consistent with the sampling errors reported by Gebremichael and Krajewski (2004, 2005) at daily scale over areas of various sizes and using various sampling frequencies. They estimated the sampling errors of 3-hourly sampled rain fields for daily mean values around 140% for $32 \times 32 \text{ km}^2$ area and 30% for $256 \times 256 \text{ km}^2$ area. For daily average from the TMPA product (3-hourly sampling), the present results give 30% over the site of Niamey ($140 \times 180 \text{ km}^2$), 34% over Ouémé ($110 \times 140 \text{ km}^2$), and 94% over the Dakar site ($30 \times 40 \text{ km}^2$). Figure 2.4 shows the distribution of the errors. The dispersion of errors of the satellite products is comparable to the one of the rain gauge estimates but is larger for TMPA. The errors are not normally distributed. The relative errors decrease with increasing space-time-averaged rainfall estimate, following a power law similar to that reported in Huffman (1997). Because of its small extension and high density, the subset site of Dakar stands out with the smallest sampling errors for the rain gauges and the largest for the satellite estimates. The ratio of the mean error to the variance of the time series is shown in Table 2.3. It varies from 0% to 50% depending on the considered scale and product, and it further confirms the need to account for the errors in running the comparison as discussed in section 2.2.1.

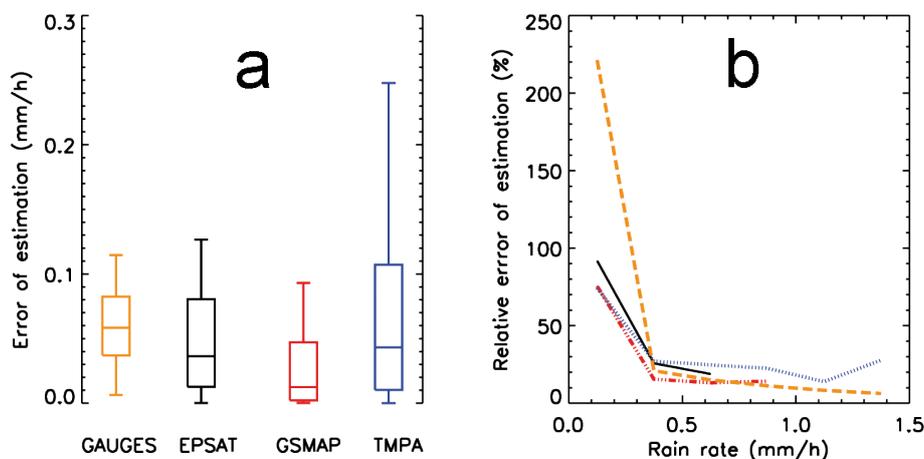


Figure 2.4 – (a) Box plots of the error of estimation on the gauge and satellite daily estimates over Niamey. Each box shows the 25th and 75th percentiles of the uncertainty distributions. The horizontal line shows the median of the distributions, and the whiskers extend out to largest and smallest values within 1.5 times the interquartile range. (b) Relative errors as a function of rain rate for the gauges (dashed line), EPSAT (solid line), GSMAP (dashed-dotted line), and TMPA (dotted line).

Scale	Site	Gauge	EPSAT	GSMAP	TMPA
Diurnal composite (postonset)	Niamey	4	23	21	31
	Ouémé	3	22	20	31
	Dakar	9	46	74	139
10 days (1-10 Aug 2006)	CILSS	27	17	13	20
10 days (10-21 Jun 2006)	CILSS	86	32	27	32
1 day ($>0.1 \text{ mm.h}^{-1}$)	Niamey	21	28	18	30
	Ouémé	17	34	20	34
	Dakar	8	50	61	94
3 hourly ($>0.1 \text{ mm.h}^{-1}$)	Niamey	13	54	26	38
	Ouémé	19	88	35	50
	Dakar	15	115	76	69

Tableau 2.2 – Mean percentages of errors for the seasonal diurnal, 10-day, 1-day, and 3-hourly time scales for the gauge and satellite product estimates.

Scale	Site	Gauge	EPSAT	GSMAP	TMPA
10-day (1-10 August 2006)	CILSS	34	11	5	12
10-day (10-21 June 2006)	CILSS	26	10	5	17
1-day	Niamey	4	18	4	13
	Ouémé	3	19	4	16
	Dakar	0.2	45	14	14
3-hourly	Niamey	0.6	25	4	12
	Ouémé	1	42	7	12
	Dakar	0.3	72	32	84

Tableau 2.3 – Ratio (%) of the mean error to the variance of the precipitation estimates. For Dakar, the median of the error distributions of the daily estimates is computed rather than the mean error because the latter does not represent the distribution of errors well as a result of the small sample size.

2.3 Methodology of comparison using error estimates

2.3.1 Approach

The comparison between the satellite products and the ground rainfall estimates are performed taking into account their respective estimated errors. The rain gauges data are considered as the reference, and the satellite estimates are expected to resemble this reference. Statistics are hence computed to characterize the expected linear relationship between them and the departures from this one-to-one relationship. First, a linear regression approach has been selected, and the results of this regression are compared to the perfect agreement $y = x$ line. Although linear regression between two samples of data assuming some errors on one of them is readily accessible, the problem is less trivial when assuming errors on both samples. Leng *et al.* (2007) reviewed the most useful classical techniques to overcome this issue (ordinary least squared, orthogonal regression, etc.) and their respective underlying validity assumptions. These are difficult to meet (e.g., Carroll and Ruppert, 1996). As shown in the previous section, here the errors from the satellite estimates are not Gaussian and are not directly comparable to the ground reference ones. The approach of Kelly (2007) has been chosen instead. It requires no assumption on the distribution of errors and is based on a maximum likelihood estimate technique, and a structural linear model (or equation) is used. Noting η as the dependent variable, it is related to ξ following

$$\eta_i = b + a\xi_i + E_i \quad (2.7)$$

where E_i is a random variable representing the intrinsic scatter of the regression relationship and a and b are the linear regression coefficients. The mean of E_i is assumed to be zero, and the

variance of E_i is assumed to be constant. The measurements data are usually accompanied by error estimates. Noting that x and y are the measurement data and that $E_{x,i}$ and $E_{y,i}$ are the associated errors on x_i and y_i , respectively, their relationship to ξ and η reads as follows :

$$\begin{aligned} x_i &= \xi_i + E_{x,i} \\ y_i &= \eta_i + E_{y,i} \end{aligned} \tag{2.8}$$

A Bayesian method is used to solve the regression using an elaborated implementation, which is described at length in Kelly (2007). In the present case, $E_{x,i}$ and $E_{y,i}$ are the errors presented in the previous section for the gauges and satellites, respectively.

2.3.2 Score indices

Among the various "scoring" indices available to quantify the degree of agreement between the satellite and rain gauge estimates Ebert (2007), a subset is selected that relies on the direct -or indirect through the fitting technique use of the individual errors (for another selection, see Ebert (2007)). Hence, the following indices are kept for the analysis :

- the mean coefficient of correlation R and
- the mean slope and mean intercept of the regression line a and b

The distributions of the coefficients a , b , and R , which are available from Kelly's technique, were analyzed and the significance of the mean values of the three coefficients were confirmed; they are thus used in the following. From these coefficients, two more indices are computed, the bias and the root-mean-square error (RMSE) of the regression :

- the bias of the regression is $\text{BIAS}_{reg} = b + (a - 1) \cdot \text{RainG}$, where RainG is the average value of the gauge estimates, and
- the root-mean-square of the regression is $\text{RMS}_{reg} = \sqrt{(1 - R^2)} \cdot \sqrt{S_y}$, where S_y is the variance of the satellite product.

Because the bias and the root-mean-square of the regression are functions of the coefficients a , b , and R , they indirectly account for the individual errors.

Similarly, the well-used probability of detection of rain (POD) and false-alarm rate (FAR) indices have been adapted to this error context for the analysis of the daily scale. Finally, a new index is computed. It consists in the frequency of error bars overlap (FEBO). The error bars are interpreted as a confidence interval, and FEBO represents the frequency of cases for which the satellite and the ground estimate error bars overlap. FEBO equals 1 if the two series are close enough given their errors and 0 if none of the values are close. FEBO is computed on the raw series as well as on the unbiased series. Both the regression and the FEBO computations are run systematically for the time-accumulated comparison over the whole Sahel as well as for the daily average comparisons over the three dense rain gauge network sites. All the regressions are computed only for the conditional rainfall estimates.

2.4 Results

The results in this section described the ability of the satellite products to mimic the rainfall variability derived from the gauges. The variability is investigated at various scales of accumulation. First, the seasonal cycle is analyzed using 10-day-averaged products, then the synoptic-scale variability is analyzed using daily means, and finally the diurnal cycle of rainfall is analyzed at both the seasonal scale using a composite and at the diurnal scale using 3-hourly accumulations.

2.4.1 Monsoon preonset and postonset and seasonal scales

In 2006, the monsoon onset occurred between 4 and 10 July (Janicot *et al.*, 2008). Therefore two 10-day periods were arbitrarily selected as representative of the conditions of the preonset (11-20 June) and postonset (1-10 August) to analyze the 10-day accumulation seasonal variability. The results of the comparison are presented in Table 2.4. For the second period, high correlation coefficients are found for the three satellite products with similar values for EPSAT and TMPA ($R \sim 0.95$). By contrast, correlations calculated without errors (not shown) were significantly lower but similar for GSMAP and TMPA. The biases are all negatives, with a bias of -0.12 mm.h^{-1} for GSMAP and equivalent lower values for EPSAT (-0.04 mm.h^{-1}) and TMPA (-0.03 mm.h^{-1}). Figure 2.5 shows the scatter diagram of the satellite versus gauge products for the second period. The error bars and the regression line are also plotted. A no-error regression line is also reported (dashed line). It shows that EPSAT and TMPA suffer from over-estimation of the large precipitation intensities. If the error bars had not been taken into account into the regression process, the result would have been the opposite. The RMS_{reg} is 0.02 mm.h^{-1} for EPSAT but 0.05 mm.h^{-1} for GSMAP and TMPA, meaning a higher precision in the fitted linear model for EPSAT. The FEBO index shows an agreement greater than 80% in both biased and unbiased cases for EPSAT. If GSMAP has the lowest linear performance scores, and similar unbiased agreements larger than 60% are found for TMPA and GSMAP. Statistics for the first period (pre- onset) yield similar conclusions but with smaller biases and RMS_{reg} values. At this 10-day scale, the satellite products capture the regional and seasonal variability of the gauges.

2.4.2 Synoptic scales

2.4.2.1 Daily time series

Both the satellite and gauge daily time series are characterized by a succession of days with and without rain. The POD and FAR allow the agreement between the two time series to be quantified, but the two indices have to be adapted to account for the errors. In fact, as seen in section 2.2.4, the error can be larger than the rain estimate itself for the small values. In such a case, the rainy day, because of the error, could also be considered as a non rainy day. Hence, POD and FAR must be computed using a variety of time series accounting for various numbers of non

(i) 11-20 June 2006	EPSAT	GSMAP	TMPA
Correlation	0.97	0.81	0.96
Slope	0.95	0.81	0.83
Intercept (mm/h)	-0.02	-0.01	-0.00
BIAS _{reg} (mm/h)	-0.02	-0.02	-0.01
RMS _{reg} (mm/h)	0.01	0.03	0.02
FEBO	0.77	0.58	0.75
FEBO-unbiased	0.77	0.65	0.75
(ii) 1-10 August 2006	EPSAT	GSMAP	TMPA
Correlation	0.98	0.82	0.93
Slope	1.15	0.80	1.29
Intercept (mm/h)	-0.08	-0.06	-0.11
BIAS _{reg} (mm/h)	-0.04	-0.12	-0.03
RMS _{reg} (mm/h)	0.02	0.05	0.05
FEBO	0.82	0.35	0.69
FEBO-unbiased	0.84	0.64	0.69

Tableau 2.4 – Statistics of the regressions between the three satellite products and the gauge estimates (i) for the 10-day period (11-20 Jun 2006) and (ii) for the 10-day period (1-10 Aug 2006). Sample size = 137.

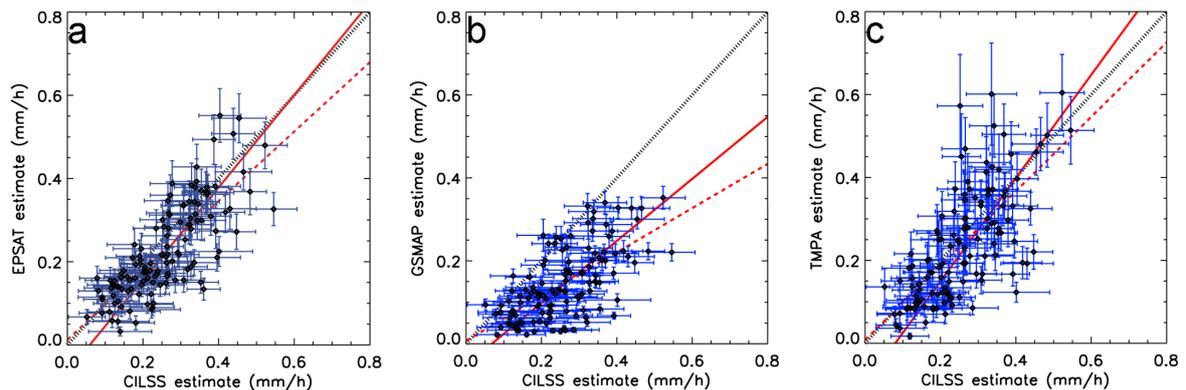


Figure 2.5 – Scatterograms of the 10-day rainfall (a) EPSAT, (b) GSMAP, and (c) TMPA estimates vs CILSS rain gauge estimates ($\text{mm}\cdot\text{h}^{-1}$) for the 1-10 Aug 2006 period. One-sigma error bars are over plotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 :1 line.

	Daily estimates			Filtered (3-5-day) daily estimates		
	EPSAT	GSMAP	TMPA	EPSAT	GSMAP	TMPA
Niamey						
Correlation	0.83	0.61	0.72	0.91	0.81	0.86
Slope	0.43	0.39	0.61	0.50	0.51	0.82
Intercept (mm/h)	0.10	0.05	0.09	0.07	0.02	0.03
BIAS _{reg} (mm/h)	-0.07	-0.14	-0.04	-0.04	-0.09	-0.02
RMS _{reg} (mm/h)	0.10	0.18	0.25	0.03	0.06	0.07
FAR	0.56	0.40	0.40	-	-	-
POD	0.92	0.83	0.79	-	-	-
FEBO	0.55	0.43	0.53	0.75	0.48	0.77
FEBO-unbiased	0.40	0.24	0.53	0.78	0.64	0.77
Sample size	67	63	57	67	63	57
Ouémé						
Correlation	0.92	0.64	0.90	0.92	0.71	0.95
Slope	0.66	0.38	0.75	0.68	0.44	1.24
Intercept (mm/h)	0.07	0.01	0.01	0.06	0.01	-0.03
BIAS _{reg} (mm/h)	-0.05	-0.21	-0.08	-0.07	-0.12	-0.02
RMS _{reg} (mm/h)	0.09	0.18	0.21	0.05	0.08	0.14
FAR	0.34	0.26	0.29	-	-	-
POD	0.93	0.81	0.90	-	-	-
FEBO	0.64	0.25	0.56	0.80	0.25	0.78
FEBO-unbiased	0.56	0.30	0.48	0.79	0.63	0.80
Sample size	89	81	81	89	81	81
Dakar						
Correlation	0.89	0.97	0.96	0.85	0.65	0.71
Slope	0.53	0.52	0.51	0.68	0.63	0.95
Intercept	0.15	0.04	0.06	0.13	0.13	0.11
BIAS _{reg} (mm/h)	0.03	-0.12	-0.12	0.07	0.05	0.10
RMS _{reg} (mm/h)	0.12	0.10	0.23	0.07	0.11	0.26
FAR	0.56	0.32	0.15	-	-	-
POD	0.76	0.59	0.32	-	-	-
FEBO	0.55	0.56	0.88	0.70	0.815	0.84
FEBO-unbiased	0.58	0.56	0.48	0.80	0.85	0.68
Sample size	40	27	25	40	27	25

Tableau 2.5 – Statistics of the regressions between the three satellite products and the gauge daily rain estimates and between the three satellite products and the gauge filtered (3-5 days) daily rain estimates over Niamey, Ouémé, and Dakar.

rainy days. Thus, a probability distribution function of POD and FAR values is considered ; in the following results, we give only the respective worst values : that is, the smallest POD and the largest FAR values.

The regression computation accounting for errors, described in section 2.3.1, is applied to the time series limited to the rainy days in both series (satellite and gauges). The scatterplots are shown in Fig. 2.6 for the three products and the three sites. The results of the statistical parameters described in section 2.3.2 along with FAR, POD, and FEBO indices are given in Table 2.5.

Figures 2.6 a-f, over Niamey and Ouémé, indicate a general underestimation of the large precipitation intensities. In Figs. 2.6 g-i, it is noticeable that, for the site of Dakar, the errors are small for the rain gauges because of the high density of the network, whereas the satellite products errors are large because of the small size of the Dakar site area. These large errors have an important impact on the characteristics of the regression accounting for the errors, which can be seen by the strong difference between the regression line (full line) with errors and the regression line (dashed line) without errors.

Considering the correlation coefficients, they are of the same order for EPSAT and TMPA and almost as high as for the 10-day accumulations. The correlation coefficients are lower for GSMAP, at least for Niamey and Ouémé. Similarly, considering the biases ($BIAS_{reg}$), EPSAT and TMPA provide the same order of negative bias (ranging from -0.04 to -0.08 mm.h⁻¹), much smaller than the biases for GSMAP (-0.14 and -0.21 mm.h⁻¹). The RMS_{reg} values obtained for EPSAT range between 0.09 and 0.12 mm.h⁻¹, the ones for GSMAP range between 0.10 and 0.18 mm.h⁻¹, and the ones for TMPA range between 0.21 and 0.25 mm.h⁻¹. EPSAT present the worst values of FAR (ranging from 34% to 56%) over the three sites but also the largest POD values, over 92% for Niamey and Ouémé. The POD values are slightly lower for the two other products. Finally, the FEBO values show lower skills than for the 10-day scale, ranging from 24% to 64% over Niamey and Ouémé. The FEBO unbiased values are always lower than their corresponding FEBO but have a similar range.

To summarize, the statistics are consistent over Niamey and Ouémé but lead to different conclusions over the site of Dakar. In addition, its smaller extension, the period considered for Dakar, is also shorter, which results in smaller sample sizes giving less strength to the statistics. In general, these results show that EPSAT and TMPA have similar skills at the daily scale, whereas regression characteristics of GSMAP are lower over Niamey and Ouémé.

2.4.2.2 3-5 Daily filtered time series

The ability of the satellite products to reproduce the well-recognized mode of synoptic variability of rainfall associated with the passing of AEW compared to the gauges is quantified by using simple spectral filtering of the 3-5-day band for each estimate. Figure 2.7 exemplifies the

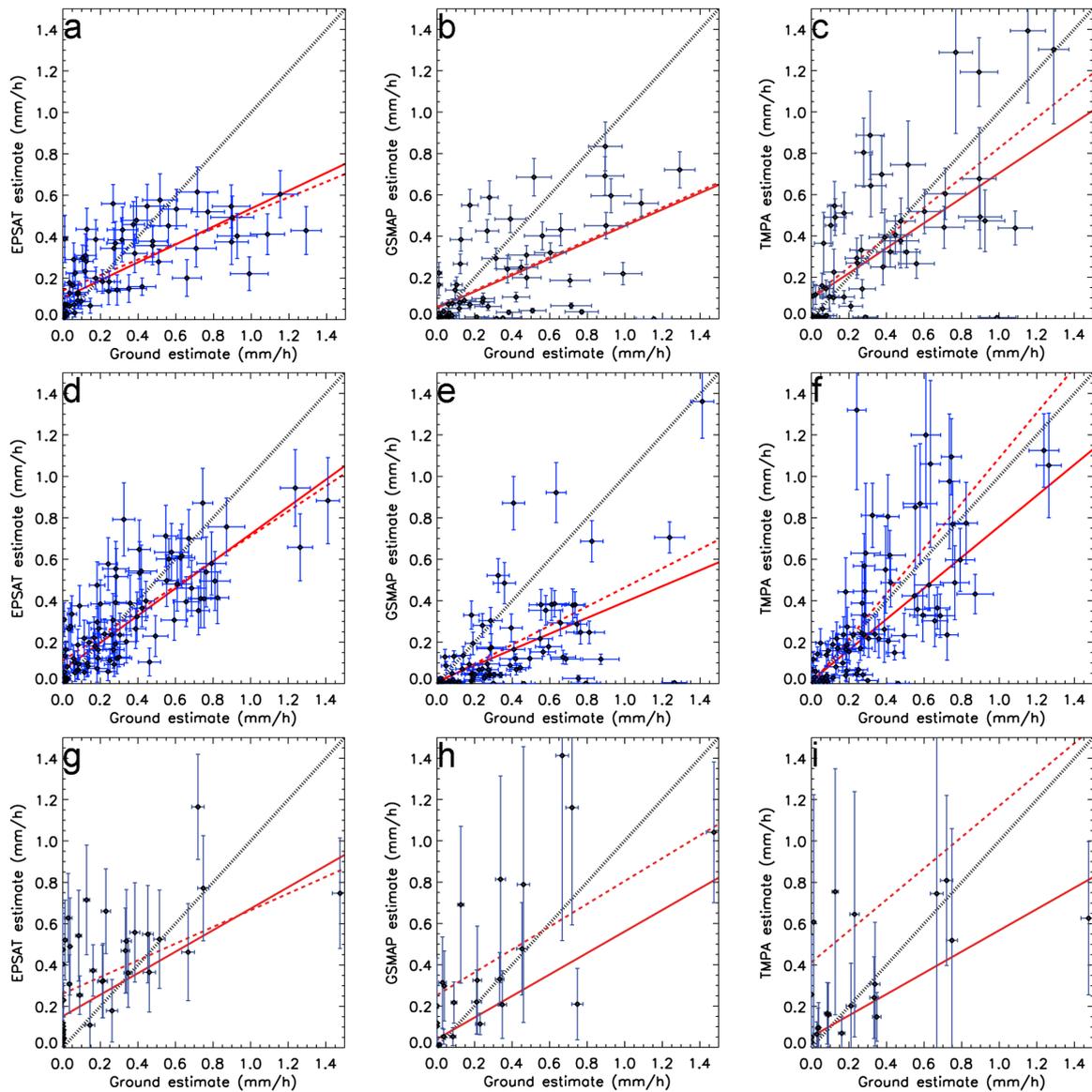


Figure 2.6 – Scattergrams of the daily rainfall ($\text{mm}\cdot\text{h}^{-1}$) of (left) EPSAT, (middle) GSMAP, and (right) TMPA estimates over the sites of (a)-(c) Niamey, (d)-(f) Ouémé, and (g)-(i) Dakar. One-sigma error bars are overplotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 : 1 line.

results for Ouémé and shows that indeed very good agreement is found at this scale. Both EPSAT and GSMAP tend to underestimate the high rain rates, whereas TMPA tends to overestimate them. The overall good agreement is confirmed by the statistics reported in Table 2.5 where high correlation, low bias, and RMS are reported. Over the Ouémé and Niamey sites, EPSAT and TMPA outperform GSMAP, which reveals a slightly worse overall fit to the gauges. At Dakar, the scores are still high, although less than for the previous sites, as expected from the smaller network under consideration (see section 2.2.4). Again, GSMAP exhibits less skill than the other products. Generally, the satellite estimates reproduce well the rainfall variability resulting from AEW, better than the unfiltered synoptic data.

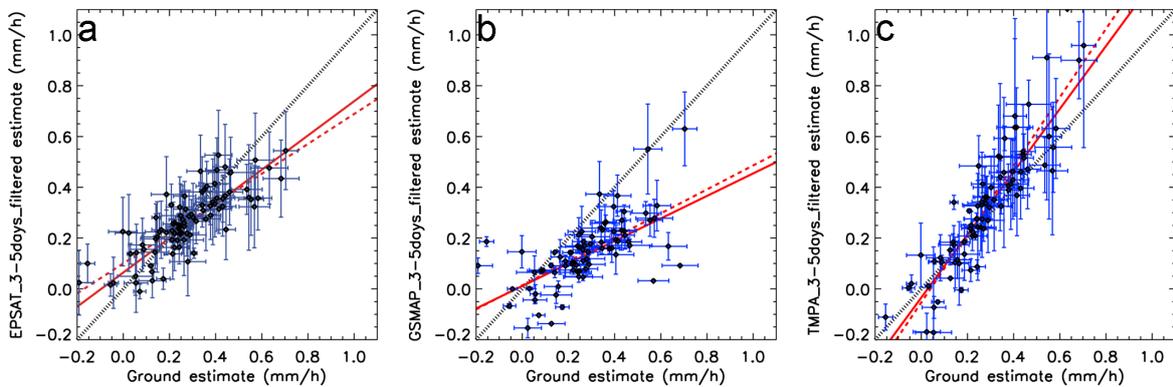


Figure 2.7 – Scattergrams of the 3-5-day filtered daily rainfall ($\text{mm}\cdot\text{h}^{-1}$) of (a) EPSAT, (b) GSMAP, and (c) TMPA estimates vs the 3-5-day filtered daily rainfall gauge estimates over the Ouémé site. One-sigma error bars are over plotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 :1 line.

2.4.3 Diurnal scale

An index is computed to quantify the influence of diurnal cycle of the insolation on rainfall. A simple Fourier analysis is performed on the 3-hourly time series, and all the scales less than or equal to one day are low-pass filtered. The filtered time series is then correlated with the raw series using the same method as earlier. The original errors at 3 h are used for both the original and the filtered series. The coefficient of determination R^2 provides the fraction of the common variance between the two series; subtracted from 1, it gives the fraction of variance associated with the diurnal scale. Results are summarized in Table 2.6. As expected, the gauges confirm the important role of the diurnal scale in the total variance of the rainfall field with value around 64% for Niamey after the onset. The satellite products also show such strong contributions of the diurnal scale. At Niamey, the satellite products all show a stronger contribution of the diurnal

cycle after the onset (Mathon et al. 2002). At Ouémé, ground measurements reveal a stronger contribution before the onset than after, and the satellite estimates do not reveal a consistent behavior. This is consistent with the findings of Depraetere *et al.* (2009) based on systems tracking at ground using gauges and radar. At Dakar, the gauge network indicates a 43% fraction, which TMPA reproduces well and is over-estimated in both EPSAT and GSMAP.

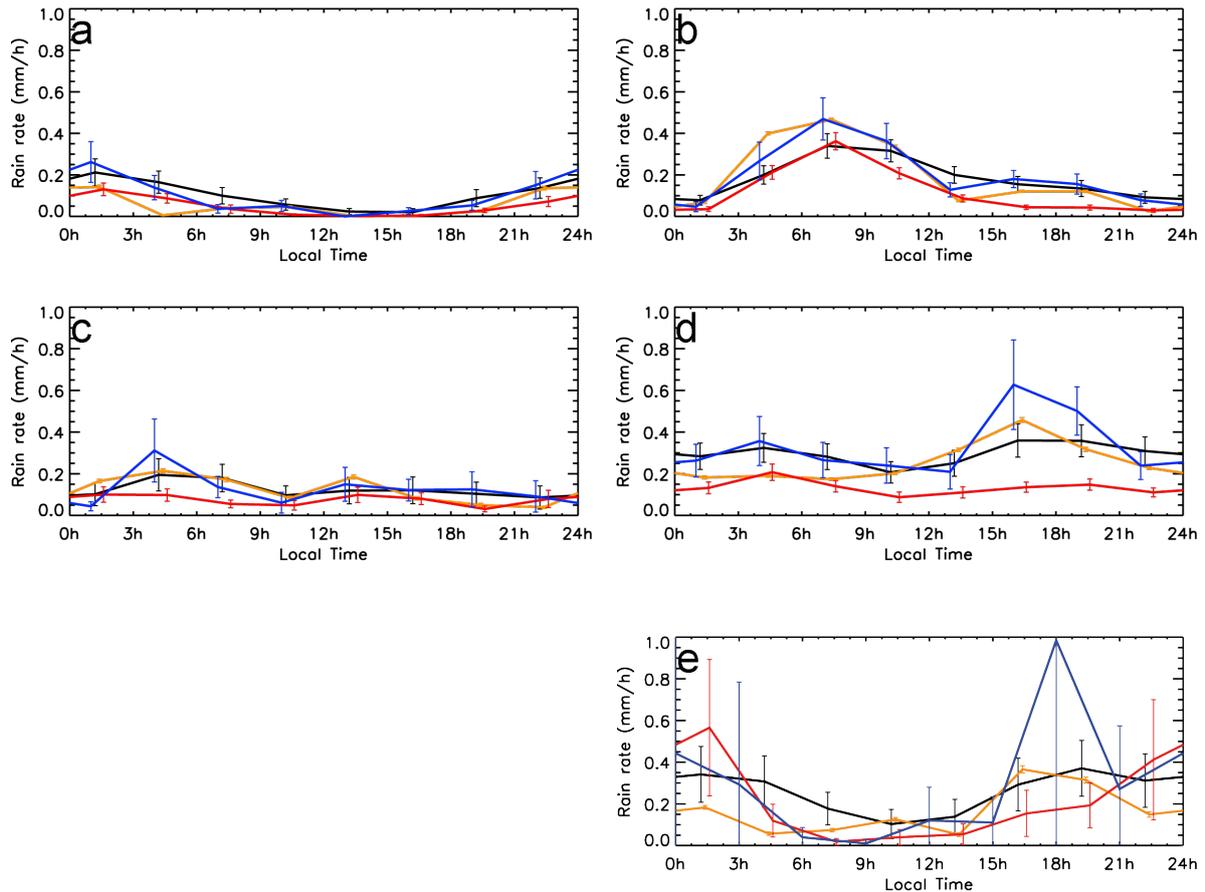


Figure 2.8 – Composite diurnal cycles of rain rates obtained from Gauges (orange line), EPSAT (black line), GSMAP (red line), and TMPA (blue line) for the (left) preonset and (right) postonset periods for (a),(b) Niamey; (c),(d) Ouémé, and (e) Dakar. Two-sigma error bars are plotted over the lines.

The composite or mean diurnal cycle is further computed and presented in Fig. 2.8. Before the onset, the mean cycle is very flat over both Sahelian regions; when the monsoon has come, Niamey exhibits a marked late night-early morning maximum and Ouémé shows a maximum, although smoother, occurring between 15 and 18 h. Over Dakar, the rain peaks between 15 and 21 h are presumably associated with local convection rather than traveling systems (e.g., Jenkins et al. 2010, unpublished manuscript). The satellite products rather exhibit a smooth maximum during the night, but TMPA also reproduces the late afternoon maximum. Note that, because

of the small size of the analyzed region, the satellite estimates here are prone to large errors that make the comparison less direct. These well-documented features of the WAM diurnal cycle are thought to be associated with a shift in the nature of the convective events with well-organized squall lines, initiated eastward and earlier in the day, contributing to this night time maximum for Niamey after the onset (Fink and Reiner 2003). Generally, the satellite products capture well this seasonal variability in the composite diurnal cycle (Fig. 2.8). For Niamey, the three products roughly behave similarly as indicated by the computation of the FEBO reported in Table 2.7. For Ouémé and Dakar, only TMPA captures the gauges features. Overall, TMPA has the higher FEBO statistics followed by EPSAT and GSMAP, which resemble each other.

The fraction of common variance between the original time series and the time series from which the mean cycle is repeatedly subtracted highlight the significance of the composite diurnal cycle (Nesbitt and Zipser, 2003). The stronger contribution of the composite cycle is found in the gauges for Dakar with a value of 34% (Table 2.8). The satellite estimates reproduce this, but for the EPSAT product the composite cycle hardly contributes to the variance. For Niamey, gauges indicate a 13% contribution of the mean cycle that is not seen in the EPSAT results. GSMAP and TMPA overestimate ($\sim 37\%$) their composite cycle influence. For Ouémé, the mean cycle only weakly contributes to the variance, a fact that satellite products agree upon only after the onset. Over the preceding period, both GSMAP and TMPA show a significant contribution of their respective composite cycles, which are very flat (Fig. 2.8c). The detailed analysis nevertheless reveals that the TMPA products best behave with respect to the gauges than other satellite products. Such good behavior of this product was previously noted over the Gulf of Mexico and was attributed to the final adjustment step that relies on the monthly gauges analysis (Sapiano and Arkin, 2009).

Site	Period	GAUGE	EPSAT	GSMAP	TMPA
Niamey	Post-onset	64	32	48	54
	Pre-onset	26*	31	27	44
Ouémé	Post-onset	36	32	16	67
	Pre-onset	61	30	56*	25
Dakar	1 Aug. 30 Sept.	43	37	80*	43

Tableau 2.6 – Percentage of variance explained by the subdaily variability of the gauge rain rate and of the three rainfall satellite products during the postonset and preonset periods for the sites of Niamey, Ouémé, and Dakar. The variance explained is calculated as $1-R^2$, where R is the correlation between the 1-day low-pass filtered 3-hourly time series and the 3-hourly time series. Low values ($<0.1 \text{ mm.h}^{-1}$) are filtered out, yielding to a varying population to perform the computations. * Not significant at the 99.9% level.

Site	Period		EPSAT	GSMAP	TMPA
Niamey	Post-onset	FEBO	0.50	0.25	0.50
		FEBO-unbiased	0.38	0.38	0.38
	Pre-onset	FEBO	0.50	0.62	0.62
		FEBO-unbiased	0.50	0.50	0.50
Ouémé	Post-onset	FEBO	0.38	0.12	0.50
		FEBO-unbiased	0.50	0.25	0.88
	Pre-onset	FEBO	1.00	0.38	0.88
		FEBO-unbiased	0.88	0.38	0.88
Dakar	1 Aug. 30 Sept.	FEBO	0.50	0.38	0.75
		FEBO-unbiased	0.62	0.38	0.62

Tableau 2.7 – FEBO and FEBO unbiased indices calculated with the mean diurnal cycles of EPSAT, GSMAP, and TMPA with 2-sigma errors.

Site	Period		EPSAT	GSMAP	TMPA
Niamey	Post-onset	13	0.8	38	36
	Pre-onset	34*	7	23	56
Ouémé	Post-onset	3	1	3	4.7
	Pre-onset	6	2	46*	31
Dakar	1 Aug. 30 Sept.	34	0.9	60	27

Tableau 2.8 – Percentage of variance explained by the mean conditional diurnal cycle of the Gauge rain rate and of the three rainfall satellite products, during the post-onset and pre-onset periods, for the sites of Niamey, Ouémé and Dakar. The variance explained is calculated as $1-R^2$, R being the correlation between the 3-hourly time series anomalies to the mean conditional diurnal cycle and the 3-hourly time series. * Not significant at the 99.9% level.

2.5 Summary and discussion

Ground data and satellite products of surface rainfall have been compared using a novel approach that accounts for the errors in the areal mean rainfall estimates during the AMMA 2006 campaign. The analysis focused on three well-instrumented sites covering a wide range of rain regimes. Simple sampling error models based on the spatial and temporal correlations of the rain field have been used both for the gauges (block kriging) and for the satellite estimates (variograms). This study completes previous efforts with climatic (e.g., Nesbitt *et al.*, 2004) and hydrologic (Hossain and Huffman, 2008) perspectives by focusing on the meteorological scales

relevant to the WAM. The findings of this study are as follows :

- at a $1^\circ \times 1^\circ$ scale, the block-kriging errors varies from 4% to 86% depending on the accumulation time and less than 15% over the denser network at Dakar ;
- the satellite errors are of the same order as the gauges errors, except for the smallest network, over which the satellite errors are greater ;
- at a 10-day scale, all the tested products show high skill in reproducing the ground network results, including the preonset-postonset variability ;
- at a daily scale, satellite products skill range from high to moderate ;
- at the AEW scale, all the tested products show high skill in reproducing the ground network variability ;
- the mean diurnal cycle and its variability in space and during the season are relatively well captured by some products, and others revealed difficulty to account for such scale.

Over the range of scales and rain regimes and types discussed here, the combined IR-MW satellite products are describing the rain variability similar to that of the ground measurements, with some being better than others at different scales. These products can then be used to investigate the physical functioning of the monsoon. For instance, Peyrillé *et al.* (2007) and Peyrillé and Lafore (2007) established an idealized conceptual model of the WAM to analyze the seasonal and diurnal cycle of the monsoonal rainfall that can be confronted to these satellite products. Using the National Centers for Environmental Prediction (NCEP) operational analysis, Bielli and Roca (2009) reported during the summer 2006 a lead-lag relationship between surface evaporation and rainfall in the WAM characterized by regional and temporal scale dependence (AEW). The use of the satellite products in such a water budget analysis would greatly enhance the findings of this model-based analysis. Yang and Smith (2006) provide an extensive, perhaps exhaustive, review on the physical processes associated to the diurnal variability of rainfall over the whole tropics. Elucidating the relative role of the various diurnal mechanisms for the whole West African monsoon region can now be attempted using the strength (coverage and sampling) of the validated new generation of combined satellite level 2 products.

The present methodology could easily be applied again if a better estimation of the error budget becomes available. The elaboration of our simple error sampling model toward a more complete error model (including the algorithm error term) is one venue for future research. An extension of the present effort to a larger number of years and satellite products making use of the ground data of the EOP of the AMMA program could strengthen the present results, especially for diurnal cycle prone to large interannual fluctuations (Mohr, 2004). Finally, the documentation of the individual convective system scale ($\Delta t < 1$ h ; $\Delta x \sim 10$ km) at which rain radar, gauges, and satellites provide different perspectives, which are difficult to compare (e.g., Roca *et al.*, 2010b), should be undertaken. The present error-based approach might be one way to improve on the situation, but such downscaling would require the challenging estimation of

an error budget at finer scale.

2.6 Conclusion du chapitre

Dans ce chapitre, trois produits d'estimation des précipitations par satellite ont été évalués à des échelles pertinentes en météorologie tropicale : l'échelle des pluies accumulées sur 10 jours, l'échelle des pluies accumulées sur 1 jour ainsi que celle du cycle diurne. À ces échelles de temps, il est nécessaire de prendre en compte les incertitudes des estimations des précipitations par satellite et des estimations issues des mesures de surface afin de comparer ces données entre elles.

Les incertitudes associées à l'échantillonnage des mesures ponctuelles des pluviomètres sont caractérisées par une technique de block-krigeage. Les incertitudes d'échantillonnage pour les produits satellite sont estimées grâce à un modèle d'erreur basé sur le calcul de variogrammes spatio-temporels. Une nouvelle méthode de comparaison prenant en compte ces incertitudes permet alors d'évaluer la qualité des produits d'estimation des pluies par satellite aux échelles sélectionnées en Afrique de l'Ouest grâce à quatre réseaux de pluviomètres. Il apparaît que :

- à l'échelle $1^\circ \times 1^\circ$, les erreurs de block-krigeage varient entre 4% et 86% selon la période d'accumulation ; ces erreurs sont inférieures à 15% pour le réseau de Dakar qui présente la plus forte densité de pluviomètres.
- les incertitudes sur les produits satellites sont équivalentes aux incertitudes liées au block-krigeage des données de surface, mis à part sur le réseau de Dakar pour lequel les incertitudes sur les produits satellites sont plus importantes.
- à l'échelle des pluies accumulées sur 10 jours, tous les produits testés reproduisent très bien les estimations de pluie issues du réseau de pluviomètres sur la bande Sahélienne, en particulier la variabilité pre-onset/post-onset caractéristique de la mousson en Afrique de l'Ouest.
- à l'échelle des pluies accumulées sur 1 jour, certains produits reproduisent convenablement les estimations de pluie issues des trois réseaux de pluviomètres de Niamey, Ouémé et Dakar.
- à l'échelle de variabilité des ondes d'Est africaines, tous les produits testés reproduisent très bien la variabilité des pluies mesurées par ces trois réseaux de pluviomètres.
- le cycle diurne moyen ainsi que sa variabilité spatiale et intra-saisonnière sont relativement bien reproduits par certains produits ; d'autres présentent des erreurs d'estimation plus importantes à cette échelle.

Cette analyse montre que les produits satellites combinés IR/MO ont atteint un degré de qualité suffisant pour être utilisés de manière quantitative aux échelles pertinentes en météorologie. Il apparaît également nécessaire d'associer une erreur à chaque estimation de pluie du fait de leurs incertitudes importantes aux échelles d'accumulation considérées.

L'élaboration d'un modèle d'erreur simple a permis d'évaluer la contribution à l'erreur totale liée à l'échantillonnage. Le chapitre 3 présentera une méthode d'évaluation de la contribution de chacune des composantes du bilan d'erreurs associées aux estimations des précipitations par satellite.

Étude du bilan d'erreurs des précipitations estimées par satellite

Comme nous l'avons vu dans le chapitre 1, les applications possibles des estimations quantitatives des précipitations par satellite sont variées. L'évaluation des produits satellites issus des développements les plus récents réalisée dans le chapitre 2 démontre leur capacité à être utilisés de manière quantitative à des échelles de temps inférieures à 10 jours. À ces échelles de temps, il est également apparu nécessaire d'utiliser les estimations de cumuls des précipitations avec leurs barres d'erreurs associées.

Divers efforts de caractérisation des erreurs ont été menés afin d'associer une erreur aux estimations de pluie à partir de modèles paramétriques et non paramétriques. Pourtant, peu de méthodes ont été développées afin d'estimer conjointement cumuls de pluie et erreurs associées. Le produit GPCP (Global Precipitation Climatology Project) d'estimation de cumuls mensuels à la résolution de 2.5° est ainsi un des rares produits à proposer une démarche intégrée vers l'estimation quantitative de cumuls des précipitations et des erreurs associées.

Une nouvelle méthode d'estimation des précipitations sur l'ensemble de la ceinture tropicale, appelée TAPEER (Tropical Amount of Precipitation with an Estimate of ERrors), a été développée et est introduite dans ce chapitre dans le but d'estimer des cumuls de pluie et leurs erreurs à l'échelle $1^\circ/1$ -jour, échelle clé pour les études météorologiques et hydrologiques. Cette approche est fondée sur une méthode de combinaison de données Infrarouge issues d'une flotte de satellites géostationnaires et d'estimations de taux de pluie issues de radiomètres Micro-ondes d'une constellation de satellites défilants.

Une investigation détaillée du bilan d'erreurs des estimations des précipitations par satellite est réalisée dans ce chapitre. Ce bilan permet d'identifier plusieurs sources d'erreurs : une erreur liée à l'échantillonnage des mesures satellitales, une erreur d'algorithmie liée à l'algorithme utilisé pour l'estimation des pluies à partir des mesures des différents systèmes d'observation utilisés, ainsi qu'une erreur liée à la calibration des instruments satellitaires.

Dans le chapitre 2, on a pu observer que le modèle d'erreur d'échantillonnage présenté permet d'évaluer le premier terme du bilan d'erreurs. Dans ce chapitre, une méthode de propagation d'erreurs permettra d'estimer les deux autres contributions à l'erreur totale. Les amplitudes de chaque source d'erreurs seront comparées ce qui permettra alors d'identifier les contributions les

plus importantes.

Le travail présenté ci-dessous a fait l'objet d'une publication, révisée pour le Quarterly Journal of the Royal Meteorological Society pour la Special Issue Megha-Tropiques :

Chambon, P., Jobard, I., Roca, R., and Viltard, N. (2011). **An investigation of the error budget of tropical rainfall accumulation derived from merged passive microwave and infrared satellite measurements.** Quart. J. Roy. Meteorol. Soc.

3.1 Introduction

Rainfall is central to the water and energy cycle of the Tropics and is characterized by an important space and time variability, making its measurement a challenge. The satellite perspective offers a unique opportunity to measure precipitation over open oceans and over land where surface measurement networks are scarce. The last decades of efforts accompanying the very successful TRMM mission has confirmed the importance of merging Passive Microwave (PM) instantaneous rain rates together with infrared imagery to build useful accumulated rainfall estimates in the Tropics (Arkin and Meisner, 1987; Adler *et al.*, 1993; Jobard and Desbois, 1994; Huffman *et al.*, 1997; Hsu *et al.*, 1997; Grimes *et al.*, 1999; Xu *et al.*, 1999; Huffman *et al.*, 2001; Kidd *et al.*, 2003; Joyce *et al.*, 2004; Turk and Miller, 2005; Huffman *et al.*, 2007; Kubota *et al.*, 2007). Indeed the suitability of the merged approach to the accumulated rainfall problem clearly stands out from the results of a number of validation studies (Negri *et al.*, 1995; Ebert *et al.*, 1996; Ebert and Manton, 1998; Xie *et al.*, 2007; Ebert *et al.*, 2007; Zeweldi and Gebremichael, 2009a; Sapiano and Arkin, 2009, see also International Precipitation Working Group website <http://www.isac.cnr.it/~ipwg/>). This is true in the Tropics, and for instance, well illustrated over West Africa, during the monsoon season, where recent validation exercises highlight significant higher skills of these types of products compared to other ones (IR only, microwave only) (Jobard *et al.*, 2011; Gosset *et al.*, 2011; Roca *et al.*, 2010a, hereafter R2010). At large enough space and time scales, sampling errors tend to be negligible (Gebremichael and Krajewski, 2004; Jobard *et al.*, 2011, R2010). On the other hand, at meteorologically relevant scales (<10 days, <1°), corresponding to a certain extent to some large basin hydrological scales, emerging applications like hydrology/crop model alimentation (Teo, 2006), meteorological quantitative precipitation forecasts evaluation (e.g. Bowler, 2006; Candille and Talagrand, 2008; Mittermaier, 2008; Ghelli and Santos, 2010), fine scale analysis of water cycle processes (Roca *et al.*, 2010b), or simply validation efforts based on ground measurements (e.g., R2010), all require an estimation of rainfall with an error associated to the accumulation. This illustrates the need for a repositioning of the theoretical framework from deterministic to probabilistic satellite rainfall estimations, where the accumulated amount can be thought of as the ensemble mean and the error estimate as the ensemble standard deviation (e.g. Gebremichael *et al.*, 2011). Assuming the various error

sources to be independent, the error budget of rain accumulation can be written as :

$$S_{Sampling}^2 + S_{Algorithm}^2 + S_{Calibration}^2 \quad (3.1)$$

where the sampling term refers purely to the error due to the discrete nature of observations (in space or time) when estimating the space/time accumulation, the algorithm term is associated to the intrinsic errors of the retrieval and the calibration term refers to satellite Level-1 data error. Note that separating the sampling and algorithm terms is usually considered as a difficult task since the precision and accuracy of retrievals depend upon the scale of interest ; hence the sampling term is sometimes considered as a combination of the two (Huffman, 1997). The difficulty of estimating each one of these terms for the various available algorithms (as well as for ground validation datasets) has often yielded to neglecting the investigation of a detailed error budget (Ebert *et al.*, 2007). The sampling term has nevertheless been discussed through the elaboration of various sampling error models (e.g. North and Nakamoto, 1989; Bell and Kundu, 1996; Steiner *et al.*, 2003; Gebremichael and Krajewski, 2005; Yan and Gebremichael, 2009, R2010). The last two terms are usually considered as small with respect to the first term and have rarely been explored in the context of rain accumulation (e.g. Bell *et al.*, 1990; Xu *et al.*, 1999).

The objectives of the paper are first to investigate the various contributors to the whole error budget of rainfall accumulations, within the Universally Adjusted GPI (UAGPI) framework, focusing on the 1-degree/1-day scale, and then to present a rainfall product developed for the Megha-Tropiques mission (Roca *et al.*, 2011). The UAGPI technique is indeed well suited to build rainfall accumulation estimates using passive microwave (PM) instantaneous rain rates and thermal infrared (TIR) data (Kummerow and Giglio, 1995; Xu *et al.*, 1999; Kidd *et al.*, 2003). Furthermore the sampling error term can be computed within the UAGPI framework with the use of a dedicated error model (R2010) and the computation of the remaining terms is rather straightforward by using the intermediate outputs of the UAGPI technique as discussed below. Hence such a framework is implemented and used to analyse the relative magnitude of each term in the error budget through a variety of sensitivity experiments.

Section 3.2 presents the datasets used in this study. Section 3.3 details the rainfall accumulation error budget ; the functioning and the present implementation of the UAGPI are then summarized together with the error propagation methodology. Section 3.4 is dedicated to presenting the results of the impact of propagating various errors in the UAGPI framework as well as the experiment set up to compare the magnitudes of errors. The performances of the rainfall product developed for the Megha-Tropiques mission are presented in Section 3.5. The consequences of this study for upcoming satellite mission algorithms development, for both passive microwave rain rates retrievals and merged products, are discussed in Section 3.7.

3.2 Data

3.2.1 Thermal infrared geostationary images

Two thermal infrared (TIR) datasets are used in the present study. Time series of TIR data from Meteosat-9 geostationary satellite are used to perform an investigation on the error budget of rainfall accumulations, during the June-to-September 2006 period. The SEVIRI imager aboard Meteosat-9 provides TIR datasets having one-sigma random errors of 0.013 K for standard scenes at 286 K (see report on Error budget Meteosat-9/SEVIRI - MetopA/IASI, GSICS Processing and Research Centre at Eumetsat). This high radiometric quality makes SEVIRI TIR datasets well suited to perform sensitivity analysis of error propagation and are therefore used in the present paper. A second dataset is also used to compute tropical wide rain estimates for illustration during the June-to-September 2009 period (Section 3.5); to this end a dataset from five geostationary satellites of the operational fleet was collected (Meteosat-9, Meteosat-7, MTSAT, GOES-11 and GOES-12) and a data quality process set up in the frame of the upcoming Megha-Tropiques mission (see Table 1 of Fiolleau and Roca (2011b) for details about the regions covered and the sampling frequencies of the different imagers).

3.2.2 Passive Microwave instantaneous rain estimates

BRAIN (Bayesian Rain Algorithm Including Neural network) is an algorithm originally developed for the TRMM Microwave Imager (TMI) and implemented for the Megha-Tropiques mission to provide Level-2 instantaneous rain estimates from Level-1 data of the MADRAS PM radiometer as well as DMSP-SSMIs and AQUA-AMSRE observing systems (Viltard *et al.*, 2006, 2011). BRAIN is a Bayesian-Monte-Carlo type of algorithm similar in concept to Gprof (Kummerow *et al.* 2001). Although some changes have been made, the general principle of BRAIN for TMI is thoroughly described in Viltard *et al.* (2006). The algorithm provides the most probable state of the atmosphere for each vector of measured brightness temperatures based on a distance-weighted average of all the atmospheric states stored in a retrieval database. A "state of the atmosphere" in this retrieval database is indeed made of a profile of hydrometeors (rain content, ice content, cloud liquid content and cloud ice content), a corresponding surface rain and the fraction of this surface rain due to convective precipitation and the associated vector of brightness temperatures. The main difference between the version presented in Viltard *et al.* (2006) and the present version lies in the process used to associate the brightness temperature vector and its corresponding atmospheric state. In the TMI-only version of 2006, a simple co-location of PR rain rates and TMI brightness temperatures prevented the use of a radiative transfer model and its related sources of errors. In the present multi-platform version, no such thing was possible since only TRMM offers the possibility for co-locating rain intensities and brightness temperatures. Hence, a more complex process was set up in order to compute a vector

of brightness temperatures for each atmospheric state of the retrieval database for each specific characteristic of SSMI, AMSR-E, TMI and MADRAS. As in the original co-located retrieval database, the rain profiles are derived from the 2A25 TRMM product (precipitation profiles retrieved from PR reflectivities). But this product does not describe thoroughly the vertical profile of neither cloud content nor ice content for obvious reasons of frequency and sensitivity threshold of the PR. An ancillary database of cloud model simulations (Goddard Cumulus Ensemble, Ferrer *et al.*, 1995) is used to match the PR rain profile to the most similar simulated rain profile. When a match is found, the PR rain profile is kept and is completed with the ice and cloud profile from the simulation. This completed profile is used as an input in a radiative transfer model to simulate the relevant brightness temperatures for each instrument's specific channels and geometry. This simulation is performed also for TMI, following the exact same procedure as for the other instruments but the simulated brightness temperatures are compared to the co-located observed ones. This comparison provides an accurate evaluation of the simulation error for TMI. This error is then extrapolated for AMSR-E, SSMI and MADRAS since it is required by the Bayesian approach to provide a variance error of both observation and simulation to compute the weighted average of the retrieval database individuals.

Validation studies of BRAIN showed that : (i) over Tropical ocean, BRAIN rain estimations have systematic errors ranging from +40% for small rain rates to -30% for large rain rates in comparison to TRMM Precipitation Radar estimates (Viltard *et al.*, 2006) ; (ii) over land in Western Africa, BRAIN rain estimations, between 2 mm h^{-1} and 10 mm h^{-1} , have systematic errors ranging from +100% and -40% (resp. +175% and -30%) over Ouémé in Benin (resp. over Niamey in Niger) in comparison to high-resolution rain gauges data (Kirstetter *et al.*, 2011).

Errors in the PM rain retrievals are characterized by a large variability depending on the algorithm, rain regimes and regions over the globe (see International Precipitation Working Group website). Presently the characterization of BRAIN errors over wider regions than Niamey and Ouémé in Western Africa is not achieved and the error characteristics needed for the present study are therefore idealized. In the following, BRAIN rain estimations will refer to several datasets respectively from the TRMM-TMI, DMSP-SSMIs or AQUA-AMSRE observing systems, for the summer 2006 monsoon over Western Africa and for the summer 2009 monsoon over the whole Tropical belt.

3.2.3 Validation datasets from ground-based rain measurements

The ground-based rain datasets are built from dense rain gauge networks covering two validation sites of about one square degree area, set up for the AMMA experiment (Lebel *et al.*, 2009), over Niamey in Niger and Ouémé in Benin, during the June-to-September 2006 period. A block kriging technique is applied to estimate areal daily accumulations together with variance of estimations giving an estimation of areal sampling errors. The other contributions of the error

budget of rain gauge-derived areal estimates can be neglected (e.g. R2010) in comparison to areal sampling errors that can be large depending on the density of the network. The relative sampling errors decrease with increasing rain accumulation intensity, following a power law. The seasonal areal average sampling errors are of 21% over the Niamey network and of 17% over the Ouémé network equipped with the highest density of gauges. The block kriging method as well as the error distributions of the validation datasets are fully described in R2010.

3.3 Methodology : TAPEER

The evaluation of a Tropical Amount of Precipitation with an Estimation of ERrors (TAPEER) can be achieved with various methodologies. In this paper, 1-degree/1-day rain accumulations are evaluated using the UAGPI framework merging TIR imagery and instantaneous rain estimates derived from the BRAIN passive microwave retrieval algorithm. The methodology to estimate the associated errors consists in modelling the sampling error term and estimating the remaining terms of the error budget with a forward error propagation method. Indeed, these remaining terms are related to errors on the PM instantaneous rain retrievals and calibration of Level-1 datasets. The latter errors can be large, especially on the PM rain retrievals, and possibly be magnified at the accumulated scale.

In the following, TAPEER-BRAIN estimates refer to the association of rain accumulation and error estimates.

3.3.1 Rain accumulation estimation

The Universally Adjusted GPI (UAGPI) framework relies on the inheritance of the GPI algorithm and the GARP Atlantic Tropical Experiment (GATE) (Arkin and Meisner, 1987); it also benefits from the increasing number of PM radiometers in orbit. The UAGPI adapts the 235-Kelvin rain/no-rain threshold and the 3 mm h^{-1} rain rate of the original GPI to current and local meteorological conditions with available PM rain estimations. For a given one-degree/one-day precipitation estimation, a training step, aiming at the adjustment of the rain/no-rain TIR temperature threshold and the conditional rain rate, is first performed. In a second stage, a generalization step is performed where the fraction of clouds colder than the threshold is used to compute the rain accumulation. The training process requires three sub-steps : (i) PM rain estimates are collocated with TIR pixels in a running space- and time- training domain, (ii) a temperature threshold $T_{threshold}$ is determined so that the fraction of TIR pixels colder than this threshold is equal to the fraction of rainy PM estimates in the domain, (iii) a rain rate intensity R_{cond} is set to the average of collected rainy PM estimates. The generalization process also requires three steps : (i) $T_{threshold}$ is used to segment the TIR images for the nominal one-degree/one-day domain, (ii) the pixels colder than the threshold are set as rainy areas with a

rain rate R_{cond} to produce intermediate fields at the resolution of TIR data, (iii) the intermediate fields are then accumulated to the one-degree/one-day resolution.

For the training step, a sufficient number of PM rain estimates collocated with TIR data is required to ensure robustness and representativeness of the $T_{threshold}$ and R_{cond} parameters. It is a trade-off between a large training domain with a large number of collocated data resulting in a robust training, and a small training domain fitting well to current and local meteorological conditions (e.g. Kidd *et al.*, 2003; Turk and Miller, 2005). An analysis performed with PM rain estimates from TMI/TRMM, AMSRE/AQUA and three SSMI/DMSP's over two African sites led to select a $5^\circ \times 5^\circ \times 5$ -day training domain, running 2 days both ahead and before the current day, and geographically centred on the current one square degree area. The impact of the size of the training domain on the accumulation and errors computations is further explored in another work (see chapter 4 of this manuscript).

3.3.2 Error estimation

The error budget is usually considered as a whole, precisely because it is difficult to separate the three sources of errors. A total error can be estimated by comparing satellite and ground rainfall measurements, accounting for the errors of the ground rainfall (e.g. Morrissey and Wang, 1995; Ali *et al.*, 2005b). Ensemble methods and neural network techniques have also been used to estimate satellite rainfall total error (Bellerby and Sun, 2005; Teo and Grimes, 2007; Bellerby, 2007). The separation of the various terms in the error budget (Eq. (1)) requires another approach. For instance, a parametric model using stochastic generation of rain fields, constrained by geostatistical properties of rainfall observed during the GATE experiment, led to the conclusion that sampling errors are the major contribution of the error budget at the monthly scale and over $512 \times 512 \text{ km}^2$ areas (Bell *et al.*, 1990). Under the same assumption, several parametric models (e.g. Gebremichael and Krajewski, 2005) or non-parametric models of error function (e.g. Gebremichael *et al.*, 2011) were developed based on radar datasets.

TAPEER-BRAIN estimates of the sampling error term are computed using variance calculations involving variogram computation (R2010). The key parameter for computing a variance of estimation is the number of independent samples $N_{independent}$. For a rain accumulation over an area A and a period T , $N_{independent}$ is evaluated using product-dependent space decorrelation distance d and time decorrelation duration τ . It is defined as follows :

$$N_{Independent} = \frac{A.T}{d^2.\tau} \quad (3.2)$$

The space and time decorrelation parameters are defined as the e-folding distances of exponential models, which are fitted on the variograms for space and time respectively. Details on the computation of variograms as well as some practical considerations on the sampling error model

concerning the variability of variograms in space and time are given in the appendix (section 3.7).

3.3.3 Error propagation method

Forward error propagation techniques have proved useful for the error budget analysis of non-linear algorithms/model using thresholding assumptions and for which building an adjoint model is not straightforward. These error propagation techniques are for instance applied in hydrology to study the response of hydrological catchment models to initial rainfall measurements errors from radar data or more recently from satellite data (e.g. Hossain *et al.*, 2004; Hong *et al.*, 2006; Moradkhani and Meskele, 2010).

In this paper, a forward error propagation method was developed to quantify the two terms of the rain accumulation error budget related to errors of the TIR and BRAIN datasets. The Level-1 TIR datasets are characterized by systematic and random errors. The UAGPI framework, described in Section 3.1, uses an adaptive threshold set during a training step, which makes use of local TIR data in a $5^\circ \times 5^\circ \times 5$ -day domain. Hence, systematic errors of the TIR datasets are taken into account in the process and do not affect cloud fraction calculations. The TAPEER-BRAIN framework thus prevent both rain and error estimations to be affected by systematic errors in Level-1 datasets. Therefore, only random errors in TIR brightness temperatures data are considered in the error propagation analysis. The Level-1 Passive Microwave brightness temperature datasets needed for BRAIN instantaneous rain estimations are subject to calibration errors (e.g. in the MADRAS specifications, calibration errors should be lower than 1 to 2 K depending on the channel). When Level-1 data are converted into rain rates by the BRAIN algorithm, these calibration errors propagate within the systematic and random errors of BRAIN estimates. Therefore, only systematic and random errors in BRAIN are considered in this paper. In order to assess the impact of a systematic bias, a rain rate dependent approach is followed. Indeed the comparison of BRAIN retrievals and TRMM PR-2A25 product revealed that BRAIN has a positive systematic error greater than 10% for rain rates smaller than 2 mm h^{-1} , but a negative systematic error of roughly 30% for rain rates greater than 10 mm h^{-1} (Viltard *et al.*, 2006). As a consequence, the systematic errors selected are applied on either low rain rates (0 mm h^{-1} to 2 mm h^{-1}), medium rain rates (2 mm h^{-1} to 10 mm h^{-1}) or high rain rates (greater than 10 mm h^{-1}).

For the various error scenarios considered below, bias distributions between reference TAPEER-BRAIN $1^\circ/1$ -day rain accumulations and $1^\circ/1$ -day rain accumulations obtained after the propagation of errors within TAPEER-BRAIN, are built and analysed for land-only cases over Western Africa during the summer 2006 monsoon.

3.4 Results of the error budget analysis

The results described in this section present the impact of idealized errors scenarios in BRAIN rain estimates (resp. in TIR brightness temperatures) in Section 3.4.1 (resp. in Section 3.4.2). Section 3.4.3 is then dedicated to the comparison of error sources magnitudes.

3.4.1 Propagating errors from PM-derived rain rates to accumulated-rain

In the case of random errors, noises propagation was performed, with uniform distributions ranging from $[-10\% ; +10\%]$ to $[-100\% ; +100\%]$. The statistical moments of the resulting relative bias distributions are reported in Table 3.1. All the distributions have a negative mean increasing with noise magnitude, and are also all negatively skewed. Excess of kurtosis decreases from 1115 to 10.5 and standard deviation increases from 1.1% to 17.6% with increasing noise magnitude. All distributions being leptokurtic with small mean values, random errors have very limited impact on TAPEER-BRAIN rain accumulations. Using a Gaussian PDF for error rather than a uniform PDF leads to similar results.

Systematic error propagations result in : (i) non-Gaussian distributions, (ii) distribution means smaller than the propagated error magnitude showing that systematic errors are attenuated from the instantaneous to the accumulated scale, (iii) the decrease of the distribution kurtosis with increasing error magnitude. Figure 3.1 shows the distributions for errors propagated on medium rain rates (2 mm h^{-1} to 10 mm h^{-1}), and Table 3.2 reports their statistical moments.

In this medium rain rate case, the -100% (resp. +100%) error in BRAIN rain estimates leads to a mean error of -73.6% (resp. +65.7%) on TAPEER-BRAIN estimates. The standard deviation of the distribution increases with systematic error magnitude. The negative systematic errors have more impact on the distributions; indeed the standard deviation of the distribution is +27.4% for the -100% bias in BRAIN, and is +12.7% for the +100% bias. Negative systematic errors induce negatively skewed distributions (except for the -80% error) and positive systematic errors induce positively skewed distributions. All the distributions are leptokurtic with an excess of kurtosis ranging from 1.0 to 1.5 that is smaller than in the case of random error propagations. The -100% systematic error in BRAIN medium rain estimates leads to the non-detection of a quite high number of one-degree/one-day estimates; this change can also be observed for the -80% error in BRAIN rain estimates (see Figure 3.1 (a) and (b)).

Systematic errors in BRAIN tend to show a more significant impact than random errors that seem to average out much more in TAPEER-BRAIN rain accumulations. These systematic errors also lead to smaller impacts when they are propagated on small rain rates and high rain rates (not shown). A -100% systematic error (resp. +100%) in BRAIN low rain rates induces a -16% (resp. +6%) error in TAPEER-BRAIN. A -100% (resp. +100%) systematic error in BRAIN high

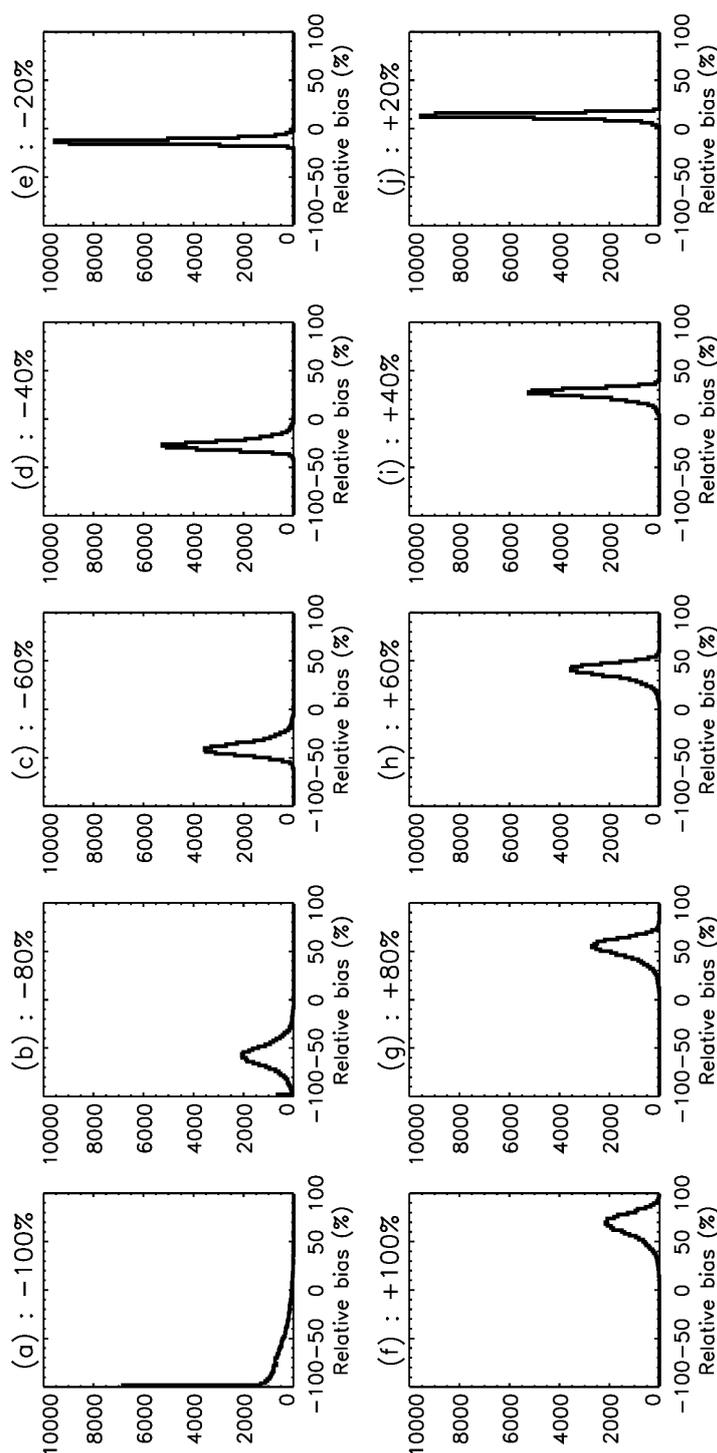


Figure 3.1 – Histograms of relative bias between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with systematic errors on BRAIN medium rain rates, and the reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the region $[20^{\circ}\text{W}-20^{\circ}\text{E}; 0^{\circ}\text{N}-30^{\circ}\text{N}]$, over land only, and for the June-to-September period of 2006. (a) corresponds to a systematic error on BRAIN medium rain rates of -100%, (b) of -80%, (c) of -60%, (d) of -40%, (e) of -20%, (f) of +100%, (g) of +80%, (h) of +60%, (i) of +40%, (j) of +20%.

Random error (%) on all rates	Number of of samples	Mean (%)	Standard deviation (%)	Skewness	Excess Kurtosis
[-10 % ; +10 %]	30054	-0.02	1.1	-17.9	1115
[-20 % ; +20 %]	30055	-0.03	1.5	-14.1	1133
[-30 % ; +30 %]	30054	-0.1	1.9	-15.4	757
[-40 % ; +40 %]	30055	-0.3	2.8	-15.3	483
[-50 % ; +50 %]	30054	-0.6	3.7	-13.4	286
[-60 % ; +60 %]	30054	-0.7	4.7	-10.7	189
[-70 % ; +70 %]	30055	-1.3	6.2	-8.2	104
[-80 % ; +80 %]	30056	-2.5	8.8	-6.4	58
[-90 % ; +90 %]	30055	-4.9	12.3	-4.4	26
[-100 % ; +100 %]	30055	-8.3	17.6	-2.9	11

Tableau 3.1 – Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with a random error on BRAIN rain rates, and the reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the area [20°W- 20°E ; 0°N-30°N], over land only, and for the June-to-September period of 2006.

Systematic error on medium rates	Number of of samples	Mean (%)	Standard deviation (%)	Skewness	Excess Kurtosis
-100 %	30056	-73.6	27.4	1.3	1.5
-80 %	30056	-58.7	14.4	-0.3	1.0
-60 %	30056	-39.5	7.6	0.8	1.2
-40 %	30056	-26.3	5.1	0.8	1.2
-20 %	30056	-13.2	2.6	0.8	1.2
+20 %	30056	13.2	2.6	-0.8	1.2
+40 %	30056	26.3	5.1	-0.8	1.2
+60 %	30056	39.5	7.6	-0.8	1.2
+ 80%	30056	52.6	10.2	-0.8	1.2
+100 %	30035	65.7	12.7	-0.8	1.2

Tableau 3.2 – Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with systematic errors on BRAIN medium rain rates, and reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the same area and period as Table 3.1

rain rates induces a -34% (resp. +24%) error in TAPEER-BRAIN. Therefore, systematic errors on medium rain rates have the most important impact on TAPEER-BRAIN rain accumulations. This is consistent with the percentage of population of medium rain rates in the datasets : small rain rates (0 mm h^{-1} to 2 mm h^{-1}) account for 37% of BRAIN rainy pixels, medium rain rates (2 mm h^{-1} to 10 mm h^{-1}) for 54% and large rain rates (greater than 10 mm h^{-1}) for 9%.

3.4.2 Propagating errors from TIR geostationary data to accumulated-rain

The scenarios selected for error propagation in TIR dataset consist in applying white noises with maximum magnitudes ranging from $[-1 \text{ K}; +1 \text{ K}]$ to $[-10 \text{ K}; +10 \text{ K}]$. Statistical moments of the resulting relative bias are reported in Table 3.3. The mean ranges from -0.02% to -0.8%, and the standard deviation reaches 6.2% for the largest noise magnitude. All distributions are leptokurtic with an excess of kurtosis always greater than 100, and are also negatively skewed.

White noise on TIR data	Number of of samples	Mean (%)	Standard deviation (%)	Skewness	Excess Kurtosis
$[-1 \text{ K}; +1 \text{ K}]$	30051	-0.02	2.8	-7.1	497.2
$[-2 \text{ K}; +2 \text{ K}]$	30047	-0.02	3.2	-2.8	311.2
$[-3 \text{ K}; +3 \text{ K}]$	30042	-0.02	3.5	-3.4	257.6
$[-4 \text{ K}; +4 \text{ K}]$	30046	-0.1	4.1	-5.8	222.0
$[-5 \text{ K}; +5 \text{ K}]$	30037	-0.1	4.2	-4.9	192.3
$[-6 \text{ K}; +6 \text{ K}]$	30032	-0.3	4.7	-5.8	183.0
$[-7 \text{ K}; +7 \text{ K}]$	30033	-0.4	5.2	-4.8	135.7
$[-8 \text{ K}; +8 \text{ K}]$	30038	-0.6	5.5	-6.1	131.5
$[-9 \text{ K}; +9 \text{ K}]$	30031	-0.7	5.8	-5.0	108.7
$[-10 \text{ K}; +10 \text{ K}]$	30034	-0.8	6.2	-5.4	102.8

Tableau 3.3 – Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with white noises on TIR brightness temperatures, and reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the same area and period as Table 3.1.

Thus TAPEER-BRAIN is more robust to errors in the TIR geostationary data than in BRAIN rain rates : for comparison, the impact of the largest random error applied, with a uniform distribution between $\pm 10 \text{ K}$, is similar to the impact of random error applied on rain rates with a magnitude of $\pm 60\%$ (see Table 3.1 and Table 3.3); the latter error magnitude being the most credible. The precision of TIR data calibrated with classical vicarious method is better than 2 K (Mueller *et al.*, 2004); this prevents from reaching such unrealistic option. Hence, in a typical case, TIR calibration error can easily be considered as negligible in the accumulated rainfall error

budget.

3.4.3 Comparing error sources

A first approach is set up to compare the magnitude of each error contribution with their respective impact on the quality of TAPEER-BRAIN precipitation estimations, over the two African sites of Niamey and Ouémé. It consists in comparing satellite rainfall and ground rainfall time series accounting for their respective uncertainties (R2010); it quantitatively estimates the relationship between BRAIN errors and both the precision and accuracy of TAPEER-BRAIN rain accumulations. Because all the different scenarios lead to a different total error, the comparison scores computed with ground-based rain for each case are then compared to reference scores computed with sampling errors only (i.e. assuming that the input data of TAPEER-BRAIN are free of errors).

In each scenario, a time series of total error is computed from the addition of the $S_{sampling}^2$ time series and $S_{calibration}^2$ or $S_{algorithm}^2$ time series extracted from the error distributions discussed in Section 4.1 and 4.2. This is achieved by : (i) assuming the algorithmic error term is independent from the rain accumulation amount, sub samples of these distributions are selected randomly, (ii) using a two-sided Kolmogorov-Smirnov test at the 99.9% confidence level to ensure that the sub samples reflect the characteristics of the original error distributions, (iii) adding $S_{calibration}^2$ or $S_{algorithm}^2$ to $S_{sampling}^2$ to build the time series of total errors. One series of errors for each of the two African sites was extracted with this technique, but prevented us from sampling the very extreme cases (e.g. -100% of systematic error on the medium rain rates). In this experiment, time series of $S_{sampling}^2$ were computed without taking into account the intra-seasonal variability of variograms for simplicity.

Figure 3.2 (a) and (b) show respectively the time series of TAPEER-BRAIN daily accumulations over Ouémé for the June-to-September 2006 period and the time series of sampling error. Figure 3.2 (c) to (h) show the time series of sampling error and total error for BRAIN's large rain rate systematic error scenarios. The total error differs from the sampling error mainly for the extreme magnitudes of systematic errors, which emphasizes the importance of performing a complete error budget analysis.

Figure 3.3 reports all comparison scores over Ouémé for the different scenarios discussed above, apart from the errors in TIR geostationary data found to be negligible. Row (a) shows the mean ratio of algorithmic error to sampling error. If this ratio never exceeds 30% in the case of random errors in BRAIN rain rates (Col. 1), it largely exceeds 100% for extreme cases of systematic errors on medium rain rates (Col. 3). Row (b) and Row (c) show that the precision of the regression increases with the error magnitude in BRAIN. Indeed, the agreement between two time series improves when errors increase (Kelly, 2007). However, the accuracy of the TAPEER-BRAIN estimates decreases when error magnitude increases. From Row (d) and Row (e), one

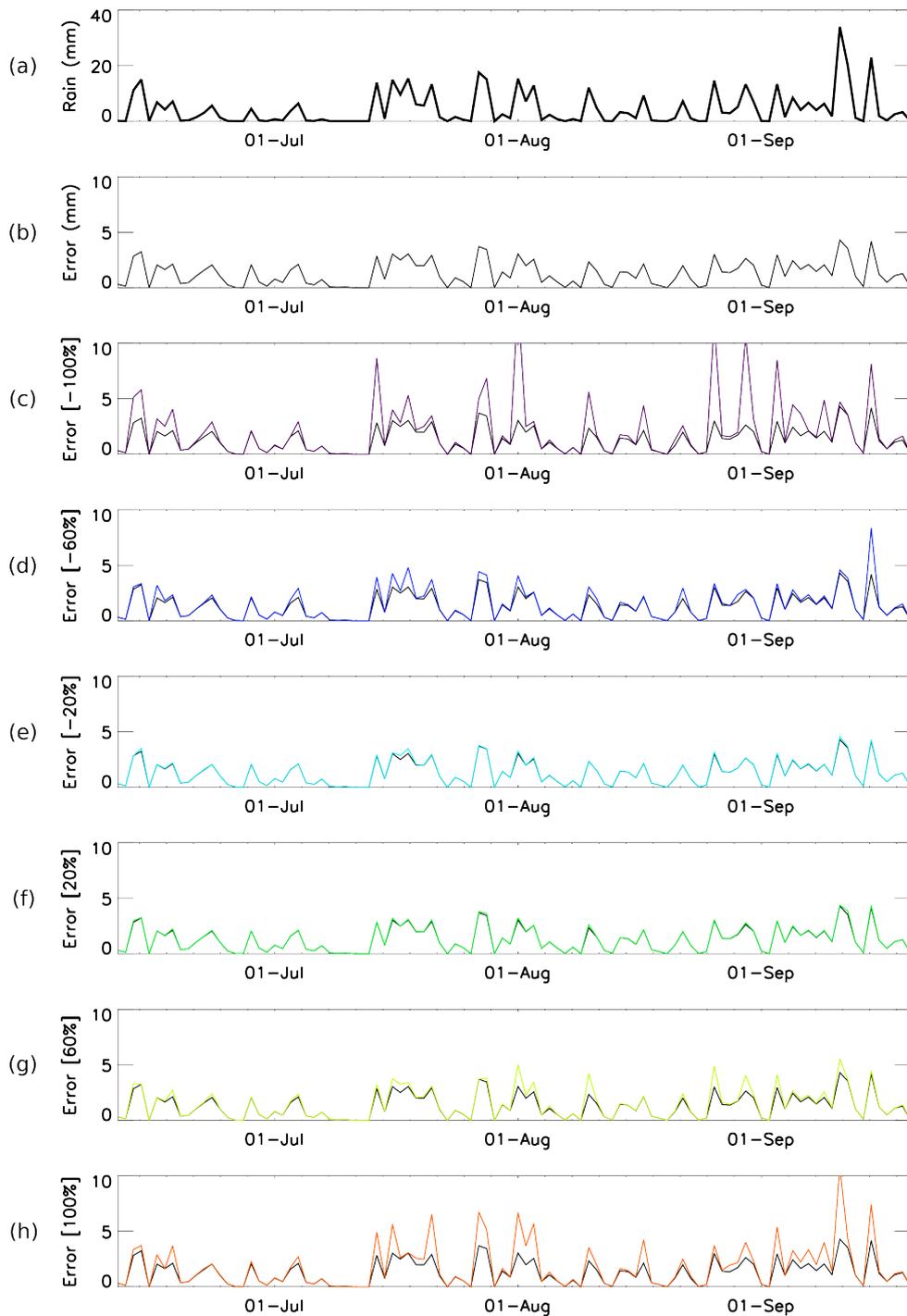


Figure 3.2 – Time series (a) of the daily rainfall (mm) of TAPEER-BRAIN estimates over the Ouémé site, from 10 June to 20 September 2006, and (b) of the corresponding daily sampling error (mm). (c-h) show the time series sampling errors (black lines) and total errors (coloured lines) where algorithmic errors are computed from the scenarios of systematic errors on BRAIN large rain rates. The black lines are almost covered by the coloured lines in the +20 % and -20% cases. (c) corresponds to a systematic error on BRAIN large rain rates of -100%, (d) of -60%, (e) of -20%, (f) of +20%, (g) of +60%, (h) of +100%.

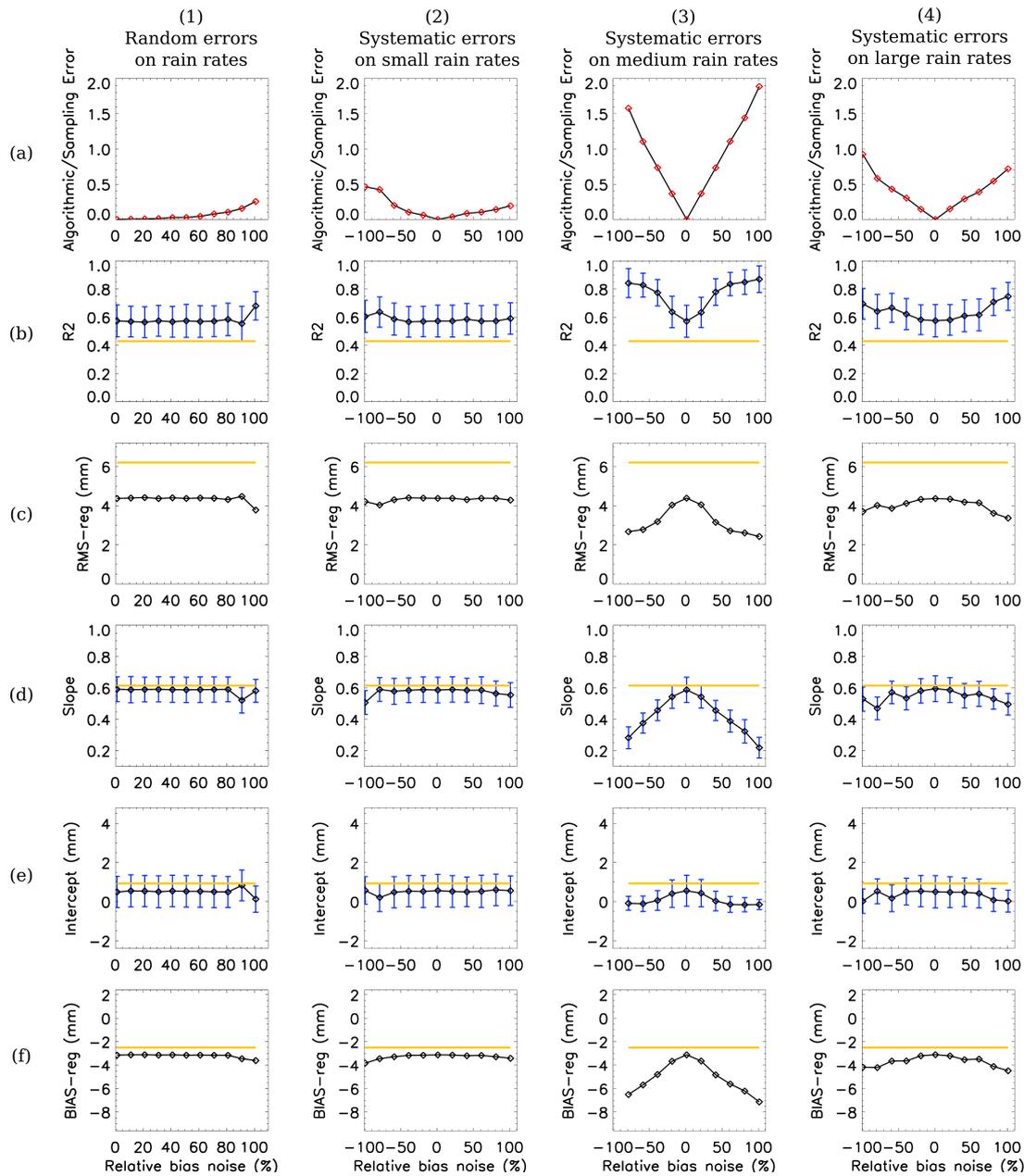


Figure 3.3 – Evolution of validation scores of the daily TAPEER-BRAIN rainfall, over Ouémé and for the June-to-September 2006 period, when algorithmic and sampling errors are considered in the validation process. (Col. 1) corresponds to scores where algorithmic errors are issued from random errors on BRAIN rain rates, (Col. 2) from systematic errors on BRAIN small rain rates, (Col. 3) from systematic errors on BRAIN medium rain rates, (Col. 4) from systematic errors on BRAIN large rain rates. (Row a) shows the TAPEER-BRAIN algorithmic to sampling errors ratio averaged over the season, as a function of magnitude of errors on BRAIN. (Row b) shows the coefficient of determination R^2 , (Row c) the RMS of the regression, (Row d) the slope of the regression, (Row e) the intercept and (Row f) the BIAS of the regression. Where scores are computed from a distribution of scores, the blue bars correspond to the standard deviation of the distribution. The scores computed without errors are over plotted in yellow.

can compute the bias of the regression reported in Row (f). Random errors and systematic errors on small and large rain rates induce a decrease in TAPEER-BRAIN's accuracy down to -4 mm (Row (f) - Col. 1, Col. 2 and Col. 4). Systematic errors in medium rain rates induce a decrease down to -7 mm in the worst cases (Row (f) - Col. 3). Comparison scores over Niamey (not shown) display similar behaviours as those over Ouémé : TAPEER-BRAIN's accuracy is almost constant for random and systematic errors in small and large rain rates, but decreases for errors in medium rain rates of roughly the same magnitude as over Ouémé.

In each of the idealized scenarios in BRAIN errors, biases of the regression give the corresponding level of accuracy at the accumulated scale ; this approach can lead to formulate elements of specifications on required BRAIN error characteristics. For instance, if user's constraint on TAPEER-BRAIN estimations over Ouémé is an overall accuracy better than 5 mm, compliant specifications in BRAIN error characteristics would be low systematic errors on small and high rain rates, and a systematic bias smaller than 40% on medium rain rates (see Figure 3.3, Row (f) - Col. 3).

3.5 Performance of TAPEER-BRAIN rain estimations

A preliminary TAPEER-BRAIN product is elaborated for summer 2006 over Western Africa and for summer 2009 over the entire Tropical belt using the whole series of geostationary TIR datasets and BRAIN instantaneous rain rates from multiple instruments (AMSRE/AQUA, SSMI/DMSP F13 & F14 and TMI/TRMM). A validation of this prototype is discussed below.

3.5.1 Validation over two African sites

Over the two well-instrumented sites in Niamey (Niger) and in Ouémé (Benin), the TAPEER-BRAIN skills are compared to those of TMPA (Huffman *et al.*, 2007). Figure 3.4 (a) and (b) show scatterograms of TAPEER-BRAIN and TMPA daily rainfall estimates over Niamey and Figure 3.4 (c) and (d) show scatterograms of TAPEER-BRAIN and TMPA over Ouémé. The comparison scores, reported in Table 3.4, all take into account the sampling errors in both the rain gauge and satellite datasets (see R2010 for their definitions).

Over Niamey, the bias of the regression ($BIAS_{reg}$) is smaller for TAPEER-BRAIN than for TMPA. The rms of the regression (RMS_{reg}) is also better for TAPEER-BRAIN. FAR is similar for TAPEER-BRAIN and for TMPA. POD is almost 0.8 for both algorithms.

Over Ouémé, $BIAS_{reg}$ is larger for TAPEER-BRAIN than for TMPA. One can also stress out that both biases are larger over Ouémé than over Niamey. RMS_{reg} and FAR are similar for both algorithms. POD is lower for TAPEER-BRAIN (0.74) than for TMPA (0.91). Although TMPA estimates include rain gauge calibration while TAPEER-BRAIN estimates do not, their statistics are similar for Niamey, but are not as good for Ouémé over this one-season time series.

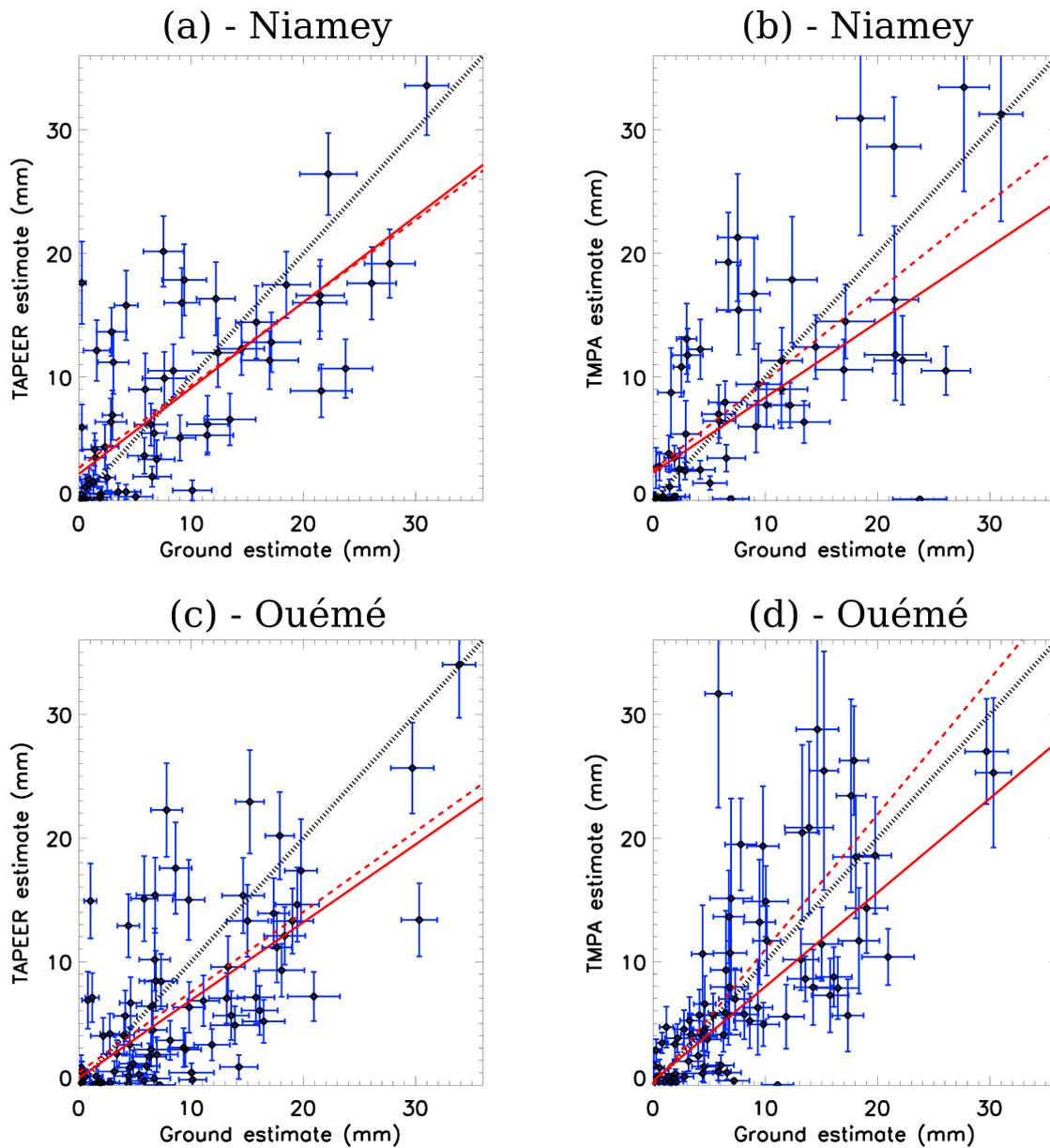


Figure 3.4 – Scattergrams of the daily rainfall estimates (mm) of TAPEER-BRAIN ((a) and (c)) and TMPA ((b) and (d)) over the two sites of Niamey ((a) and (b)), and Ouémé ((c) and (d)) for the June-to-September period of 2006. One-sigma error bars are over plotted in blue. The full red lines are the regression lines taking the error bars into account, and the dashed red lines are the regression lines without errors.

	Niamey		Ouémé	
	TAPEER	TMPA	TAPEER	TMPA
Sample size	60	56	79	80
Correlation (no error)	0.72	0.70	0.68	0.70
Correlation (with error)	0.77	0.72	0.76	0.91
Slope	0.69	0.61	0.63	0.76
Intercept	2.1	2.3	0.6	0.3
$BIAS_{reg}$ (mm)	-0.33	-0.9	-2.7	-1.8
RMS_{reg} (mm)	4.8	6.0	4.6	4.9
FAR	0.35	0.39	0.30	0.28
POD	0.77	0.78	0.74	0.91

Tableau 3.4 – Validation scores of the two satellite products TAPEER-BRAIN and TMPA daily rain estimates over the two sites of Niamey and Ouémé, for the June-to-September period of 2006. Sampling errors are taken into account for both satellite and gauge daily rain estimates in the validation process.

3.5.2 A first global comparison

Figure 3.5 (a) shows the TMPA rain accumulation map for the 30th of July 2009, and Figure 3.5 (b) and (c) show TAPEER-BRAIN rain accumulation and associated sampling error maps for the same day. A series of rainy events with varying magnitudes are located all along the Inter Tropical Convergence Zone. The day-to-day statistics of a classical comparison without errors (not shown) between global maps of TAPEER-BRAIN and TMPA demonstrate a stable behaviour along the June-to-September 2009 period, apart from the bias :

- The correlation between rainy one-degree/one-day accumulations (more than 8000 samples per day) ranges between 0.75 and 0.85 along the season.
- The POD ranges between 70% and 80% along the season for a rain/no-rain threshold of 1 mm.
- The FAR ranges between 10% and 20% along the season for a rain/no-rain threshold of 1 mm.
- The bias between rainy one-degree/one-day accumulations ranges between -10% and 10%. The bias is more frequently negative in June and July, and more frequently positive in August and September.

As a preliminary work, a qualitative preview of the total error magnitude for the TAPEER-BRAIN product was built with the same methodology as in Section 3.4.3. A best-case scenario of + 20% bias and a worst-case scenario of + 60% bias on medium rain rates were selected ; for both of them, small and high rain rates are assumed to have low biases. Figure 3.5 (d) and Figure

3.5 (e) are previews of the corresponding best and worst total error maps of TAPEER-BRAIN. The quantitative information of these maps should be interpreted with caution but it gives a qualitative first preview of the range in which the total error magnitude of the TAPEER-BRAIN product may vary depending on the error of BRAIN.

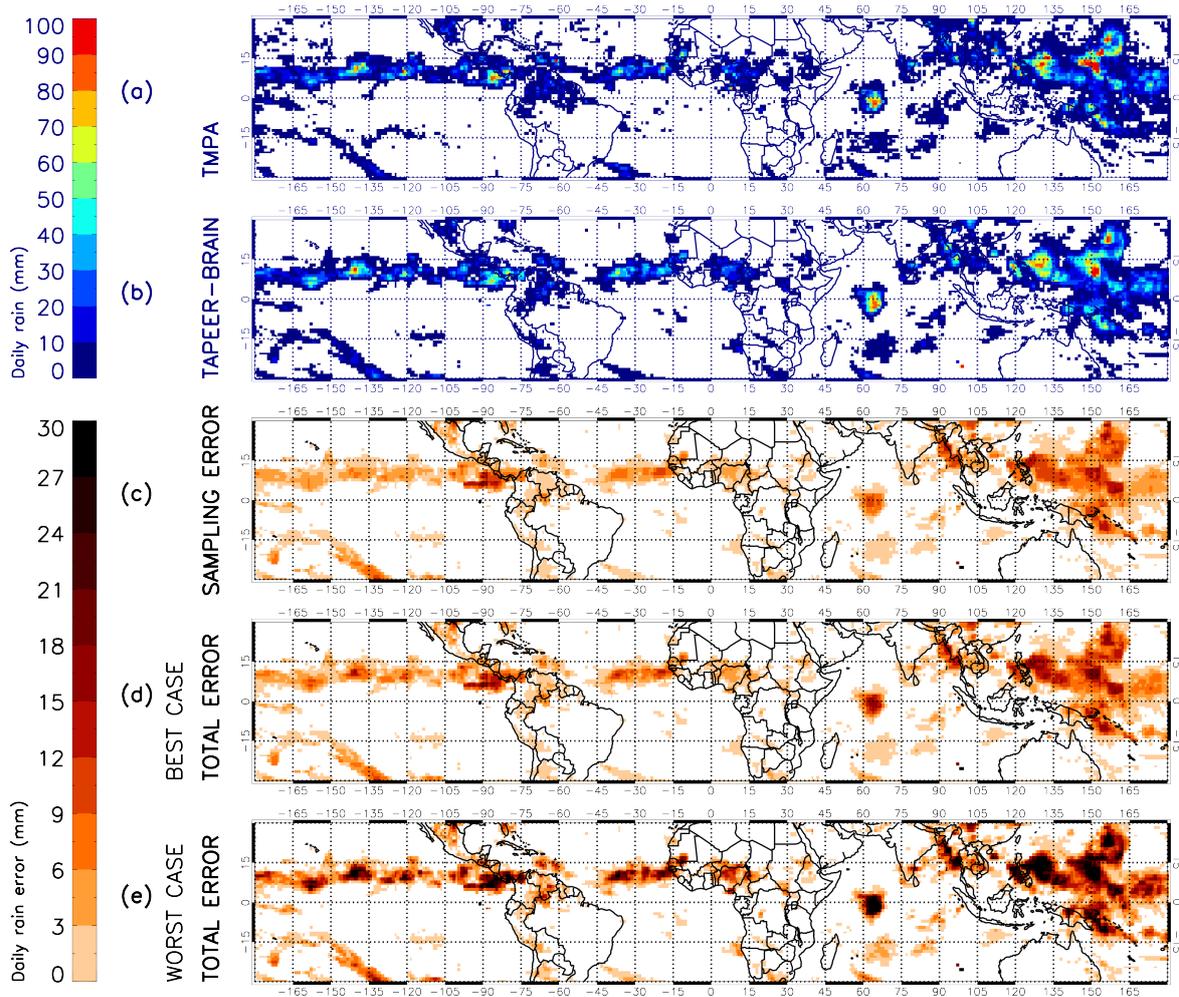


Figure 3.5 – Maps of 1-degree rain accumulations for TMPA (a) and TAPEER-BRAIN (b) for the 30th of July 2009 (accumulations smaller than 0.5 mm are set to zero). The corresponding maps of TAPEER-BRAIN sampling errors (c), as well as TAPEER-BRAIN total errors for a best case (d) and a worst case (e), are computed with the e-folding distance maps of the last 10-day period of July 2009.

3.6 Summary and perspectives

3.6.1 Summary of the error budget analysis

The magnitudes of errors related to algorithmic and calibration issues have been evaluated with a forward error propagation approach. The main findings are :

- Rain accumulations built within the TAPEER-BRAIN framework are robust to large random errors in TIR data. By construction, the framework is insensitive to systematic errors in TIR data. Calibration errors in the TIR imagery can thus be neglected in the total error budget.
- For a given precision and accuracy on TAPEER-BRAIN accumulations, it is possible to express a requirement on the error characteristics of the PM instantaneous rain retrievals.
- Random errors in PM instantaneous rain retrievals have very limited impact on rain-accumulations accuracy in comparison to systematic errors.
- The impact of a systematic error depends on the rain regime : medium rain rates, ranging from 2 mm h^{-1} to 10 mm h^{-1} , have the largest impact on rain accumulations. Therefore, improving the one-degree/one-day rain accumulations requires improving PM rain algorithms in this rain regime.

Based on this framework and the comparison between error sources, it is now possible to formulate a satellite rainfall product using the UAGPI technique for the accumulation and an error estimate composed of two terms : a PM algorithm error and a sampling term computed from the present model (the calibration term of the error budget being neglected). The sampling error term requires the characterization of d and τ parameters over the whole Tropical belt ; a $5^\circ \times 5^\circ$ and 10-day update of variograms was hence selected (see Appendix, section 3.7).

The satellite rainfall product is called TAPEER-BRAIN for Tropical Amount of Precipitation with Estimation of Errors, using BRAIN instantaneous rain estimates ; it will be one of the operational products for the Megha-Tropiques mission.

3.6.2 Perspectives

TAPEER-BRAIN estimates will be produced over the whole Tropical belt at the 1-degree/1-day resolution, using all available conical scanning Passive Microwave imagers. Because of both the low-inclination of Megha-Tropiques orbit and the large swath of the MADRAS instrument, one can expect the quality of the product to be enhanced in light of the encouraging performances of the TAPEER-BRAIN prototype presented in this paper. Indeed a recent study involving synthetic PM retrievals showed that TAPEER-BRAIN estimates could benefit from Megha-Tropiques and additional observing systems by reducing the training volume in the algorithm (see chapter 4 of this manuscript). Plans are thus made to use the PM observations of future satellite missions in the TAPEER-BRAIN framework, including the Global Precipitation Measurement

mission (GPM). Users of PM data from multiple platforms will benefit from the work of the GPM community with the help of the GPM-Core satellite. Its dual frequency radar is crucial to inter-calibrate and harmonize the observations of the upcoming constellation.

In this paper, the algorithm error term was computed assuming (i) idealized error scenarios in BRAIN estimates, (ii) no dependence of the algorithmic error term on the rain accumulation amount. The error propagation results are valid for the present implementation of the TAPEER-BRAIN product, but it would be worth to further assess the sensitivity of the results to the size of the selected training domain. Plans are made to apply the error propagation methodology using to the data collected during the Megha-Tropiques Calibration/Validation campaign over various rain regimes in Africa, India and Brazil. The expected enhanced characterization of BRAIN errors will help investigate to what degree the sampling error model gives the major contribution to the total error budget of TAPEER-BRAIN rain accumulations. If BRAIN errors cannot be neglected in the total error budget, a modelling of the algorithmic error term, as a function of rain accumulations, would then be needed to provide algorithmic error maps (Gebremichael *et al.*, 2011). Non-detection of low rain rates remains a critical issue in PM retrieval algorithms (e.g. Stephens and Kummerow, 2007). Currently, the rain/no rain mask in BRAIN leads to a loss of at least 20% of the rainfall volume over West Africa (Kirstetter *et al.*, 2011) and work is ongoing to improve the rain/no-rain mask of BRAIN. The impacts of non-detection (and of its improvement) of the lowest rain rates on TAPEER-BRAIN error budget will be explored in the future.

3.7 Appendix : Variogram computation

Space and time variograms are computed over the intermediate fields produced during the generalization process of the UAGPI technique to estimate the d and τ parameters. These intermediate fields (step (ii) of the generalization process detailed in Section 3.3.1) are subject to large random errors due to the very loose physical/statistical relationship between TIR data and rainfall. Random errors can contaminate the computation of variograms. They are thus computed over indicator fields (rain/no-rain) to reduce the contamination. Indeed, when data are possibly contaminated by large random errors, the nugget of the variogram is close to the sample variance : the samples appear to be randomly correlated despite the nature of the data, which might be strongly correlated. The generation of an indicator field (noted "IF") can be used to isolate part of the random errors due to the assignment of an intensity value to the samples (Journel, 1983; Carr *et al.*, 1985). The use of an indicator field is well adapted to this study because the UAGPI method provides a $1^\circ/1$ -day rainy cloud fraction. This cloud fraction is computed from the rain/no-rain samples of the intermediate fields and then scaled to a rain accumulation with a PM-derived rain intensity. The fields, with rain intensities, are by-products

of the computation process and are characterized by heterogeneous and discontinuous errors.

Over an area A and for a slot t_0 of an IF time series, a space variogram $\gamma(\Delta x, t_0)$ can be computed with Equation 3.3 where Δx is the lag space between a pair of two samples \vec{x}_0 and \vec{x}_1 , $n(\Delta x)$ is the number of pairs of samples distant of Δx , and σ_{SPACE}^2 the variance of the IF field over the area and at the slot t_0 where the space variogram is computed.

$$\gamma(\Delta x = \| \vec{x}_1 - \vec{x}_0 \|, t_0) = \frac{\frac{1}{n(\Delta x)} \sum_{i=0}^{n(\Delta x)} [IF(\vec{x}_1, t) - IF(\vec{x}_0, t)]^2}{\sigma_{SPACE}^2(t_0)} \quad (3.3)$$

A single space variogram at time t_0 does not give sufficiently reliable information to estimate the correlation pattern and hence the d parameters for a $1^\circ/1$ -day rain accumulation. One solution is to average several variograms over a period T to ensure robustness of the information (e.g. Lebel *et al.*, 1987; Grimes, 2009b). The choice of an area A and a period T is thus a trade off between robustness and representativeness of the correlation pattern that can be extracted from the mean variogram $\Gamma_{A,T}(\Delta x)$ (see Equation 3.4)

$$\Gamma_{A,T}(\Delta x) = \frac{1}{N_{t_0 \in T}} \sum_{t_0 \in T} \gamma(\Delta x, t_0) \quad (3.4)$$

The robustness of variograms depends on the number of samples used to compute them. In hydrology, this number can be limited where rain gauge networks are sparse, therefore climatological variograms using at least one season are often preferred (e.g. Lebel and Barbé, 1997). The variograms for the TAPEER-BRAIN estimates were found to be robust for 10-day periods and $5^\circ \times 5^\circ$ areas over the Tropical belt because a large number of pixels in the intermediate fields can be used if enough rain events occur. The d parameter is hence estimated with $\Gamma_{A,T}(\Delta x)$, which is the average of roughly 500 $\gamma(\Delta x, t_0)$.

Similarly, for a location \vec{x}_0 of an IF time series, a time variogram $\gamma(\vec{x}_0, \Delta t)$ can be computed with Equation 3.5 where Δt is the lag time between two samples t_0 and t_1 belonging to the period T , $m(\Delta t)$ is the number of pairs of samples distant of Δt and $\sigma_{TIME}^2(\vec{x}_0)$ the variance of the IF fields over the period and the location \vec{x}_0 where the time variogram is computed.

$$\gamma(\vec{x}_0, \Delta t = | t' - t |) = \frac{\frac{1}{m(\Delta t)} \sum_{i=0}^{m(\Delta t)} [IF(\vec{x}_0, t') - IF(\vec{x}_0, t)]^2}{\sigma_{TIME}^2(\vec{x}_0)} \quad (3.5)$$

In addition, a robust time variogram $\Gamma_{A,T}(\Delta t)$ needs to be estimated to provide reliable information of the time correlation pattern and hence the τ parameter. $\Gamma_{A,T}(\Delta t)$ is an area average of time variograms, following Equation 3.6.). For a 10-day period and a $5^\circ \times 5^\circ$ area, the τ parameter is hence estimated with the average of roughly 20 000 $\gamma(\vec{x}_0, \Delta t)$.

$$\Gamma_{A,T}(\Delta t) = \frac{1}{M_{\vec{x}_0 \in A}} \sum_{\vec{x}_0 \in A} \gamma(\vec{x}_0, \Delta t) \quad (3.6)$$

Figure 3.6 shows an example of 10-day space and time variograms computed over a $5^\circ \times 5^\circ$ area in Western Africa for the TAPEER-BRAIN rain estimates. These variograms are characterized by d and τ values of 63 km and 1.5h respectively.

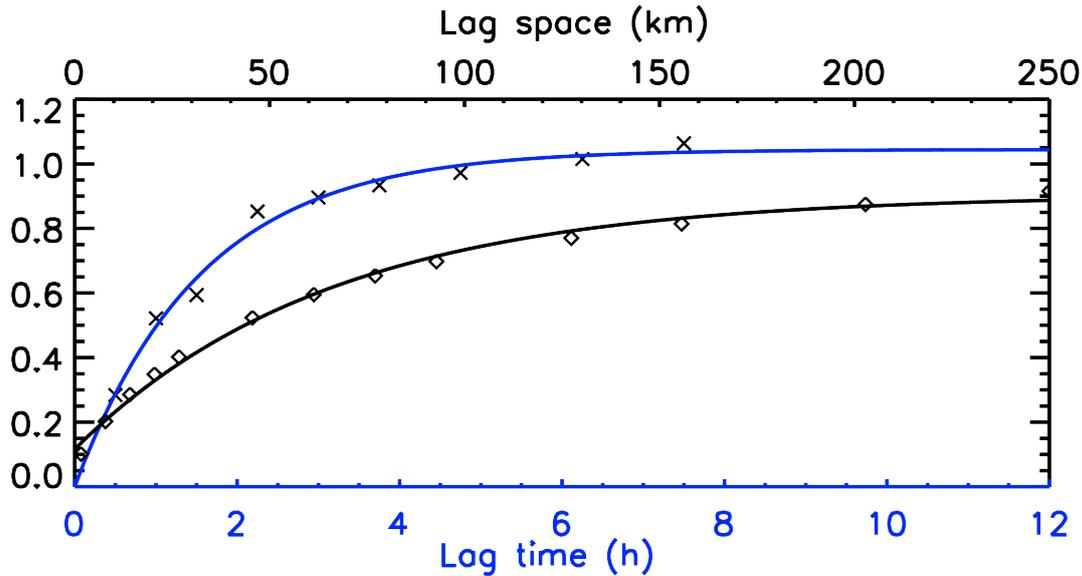


Figure 3.6 – Space (diamonds) and time (crosses) variograms of TAPEER-BRAIN daily rain accumulations over the $5^\circ \times 5^\circ$ region $[2^\circ\text{-}7^\circ\text{E}; 12^\circ\text{N}\text{-}17^\circ\text{N}]$, for the last 10-day period of August 2009. Exponential fits are plotted for the space variogram (black) and for the time variogram (blue), corresponding to d and τ of respectively 63 km and 1.55 h.

The representativeness of variograms is a key issue for a local estimation of d and τ parameters. Indeed, averaging variograms leads to the combination of correlation patterns of different rainfall events. The more variograms from different meteorological situations are combined, the less $\Gamma_{A,T}(\Delta x)$ and $\Gamma_{A,T}(\Delta t)$ are representative of a one-degree/one-day.

Figure 3.8 shows the space variability of d and τ over the whole Tropical belt for the first 10-day period of July 2009 and Figure 3.7 shows the 10-day variability of d and τ for a $5^\circ \times 5^\circ$ box over Western Africa. Variogram computation is highly CPU consuming, non-overlapping 10-day periods and $5^\circ \times 5^\circ$ boxes were thus selected although overlapping areas and periods would prevent discontinuities in the d and τ fields. Where 10-day rain accumulations are too small and lead to non-robust variograms, d and τ are set to maximum values and thus maximize the sampling error that will be provided. The space variability of d is depicted on Figure 3.8 (b) where d ranges

between 40 km and 70 km over homogenous areas in Western Africa and the Western Pacific ocean. Maxima of d greater than 100 km correspond to areas with 10-day rain accumulations greater than 240 mm (Figure 3.8 (a)), such as in the Bay of Bengal. One can see in Figure 3.8 (c) that τ demonstrates a similar variability with homogeneous areas where it ranges between 1 h and 3 h. Maxima of τ range between 3 h and 5 h with a few extremes values greater than 6 h in the Warm Pool. The time variability of d and τ is depicted on Figure 3.7 where d ranges between 20 km and 60 km, and τ ranges between 0.7 h to 1.5 h for the selected $5^\circ \times 5^\circ$ box. One can see the increase of both d and τ e-folding distances, following the monsoon activity. Indeed, the time series of rain accumulations over the same area shows the increase in the number and in the intensity of rainy events from pre-monsoon season in June to active monsoon in July and August. Variations of variograms in both space and time are essential for the sampling error because biased parameters can lead to an over- or under-estimation of a factor of two of the error.

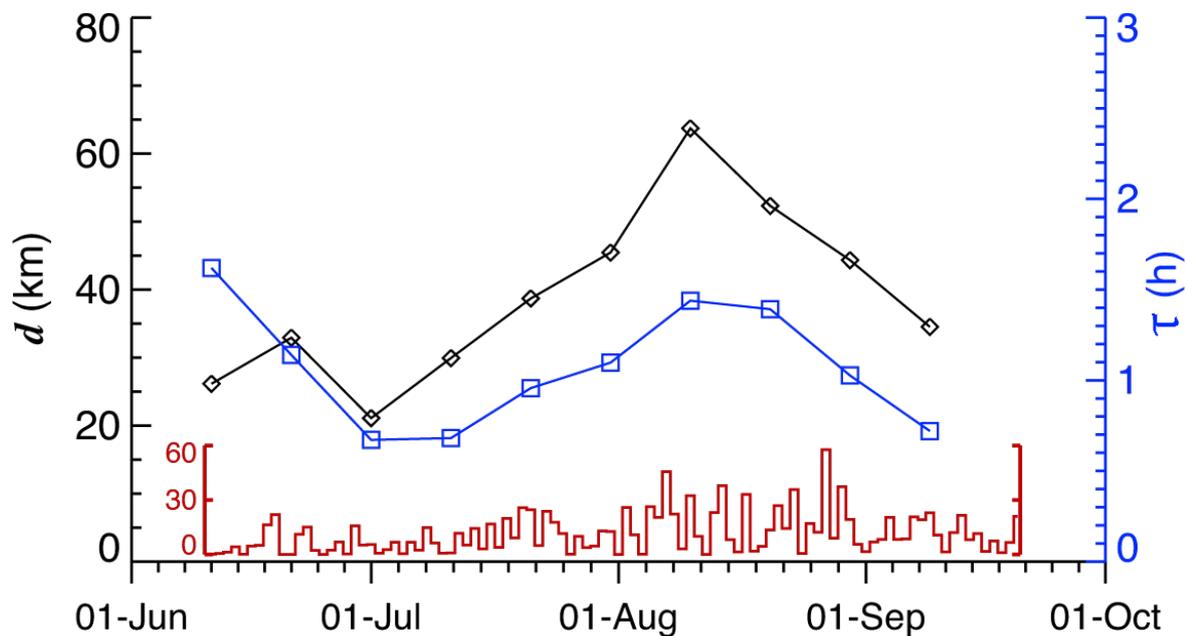


Figure 3.7 – Time series of TAPEER-BRAIN space (diamonds) and time (squares) e-folding distances, over the $5^\circ \times 5^\circ$ window $[7^\circ\text{-}12^\circ\text{E}; 12^\circ\text{N}\text{-}17^\circ\text{N}]$, for 10-day periods from 10 June 2006 to 20 September 2006. TAPEER-BRAIN daily rain accumulations (in mm) are over plotted in red for the period 10 June to 30 September 2006.

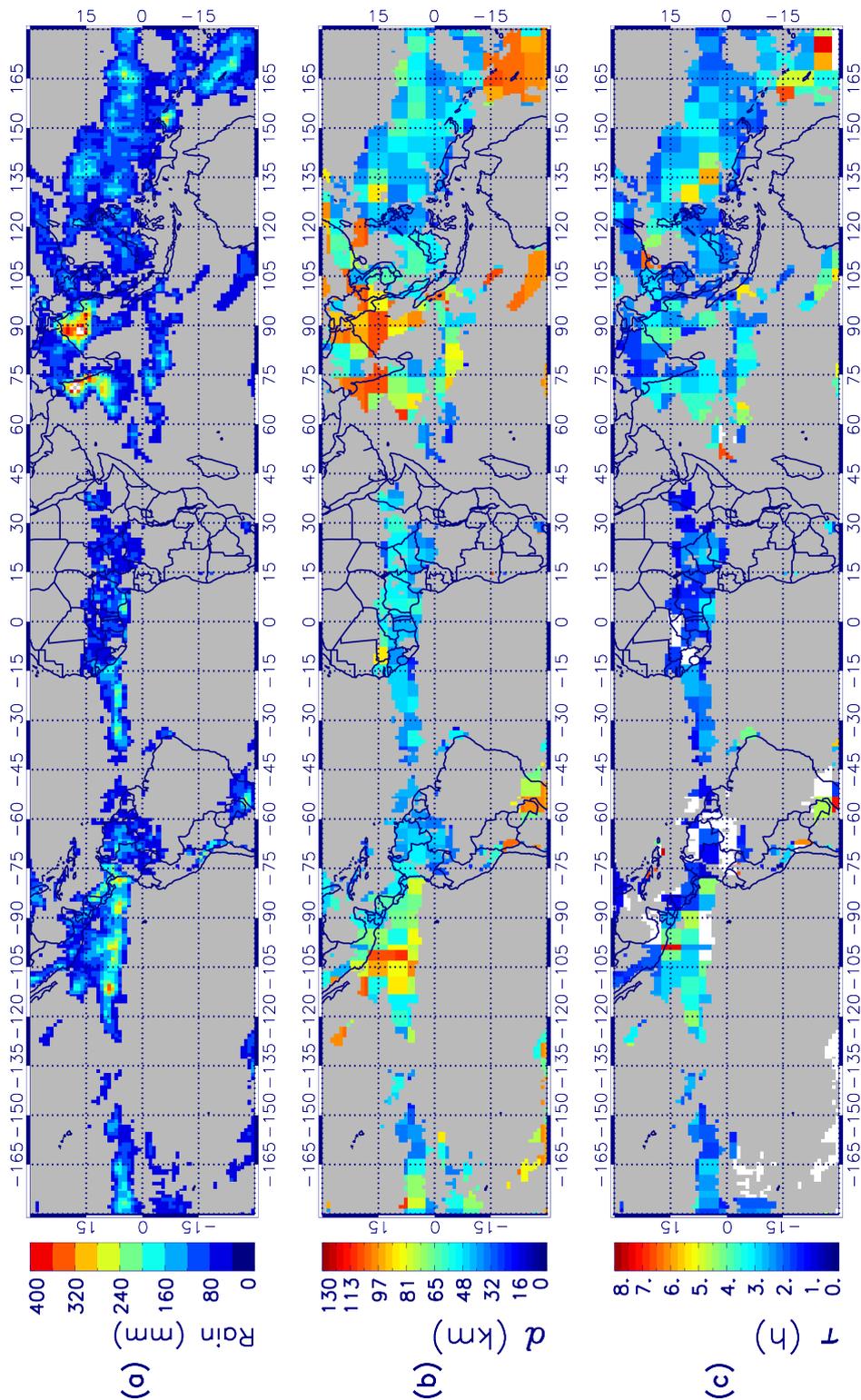


Figure 3.8 – TAPEER-BRAIN rain accumulations for the first 10-day period of July 2009 (a) and associated space (b) and time (c) e-folding distances. Areas in grey correspond to areas where 10-day rain accumulation is smaller than 40 mm. White areas correspond to areas where space or time e-folding distances are greater than the value of the upper limit shown on the colour scale.

3.8 Conclusion du chapitre

Un nouvel algorithme d'estimation quantitative des précipitations, appelée TAPEER, a été introduit dans le chapitre 3. Cette méthode consiste en l'estimation de cumuls de pluie à la résolution $1^\circ/1$ -jour et de leurs incertitudes associées.

Le bilan d'erreurs des estimations des précipitations par satellite peut se décomposer en trois termes : une première contribution liée à l'échantillonnage des observations spatiales, une deuxième contribution qui découle de l'algorithme utilisé pour les estimations de pluie, et une dernière contribution liée à la calibration des instruments.

Un algorithme de combinaison IR/MO permet d'estimer des cumuls de pluie. Les erreurs d'échantillonnage associées sont évaluées grâce au modèle d'erreur introduit au cours du chapitre 2. Une méthode de propagation d'erreurs, de l'échelle instantanée vers l'échelle accumulée $1^\circ/1$ -jour, permet alors d'estimer les deux autres contributions à l'erreur totale. Il est apparu que :

- les erreurs de calibration des imageurs IR à bord des satellites géostationnaires sont négligeables dans le bilan d'erreurs.
- la méthode TAPEER est sensible aux erreurs d'estimation des taux de pluie instantanés issus des mesures micro-ondes, en particulier les erreurs systématiques sur les taux de pluie d'intensité comprise entre 2 mm.h^{-1} et 10 mm.h^{-1} .
- pour une erreur maximale tolérée sur un cumul de pluie TAPEER à la résolution de $1^\circ/1$ -jour, il est possible de caractériser l'erreur maximale acceptable pour les taux de pluie instantanés.

Dans le cadre de la mission Megha-Tropiques, le développement de la méthode TAPEER a conduit à la mise en place d'un nouveau produit d'estimation quantitative des précipitations pour les Tropiques appelé TAPEER-BRAIN. Ce produit fournira des cumuls de pluie et erreurs associées à la résolution $1^\circ/1$ -jour sur l'ensemble de la ceinture tropicale. L'annexe B est consacrée au produit TAPEER-BRAIN dont les bases théoriques ainsi que sa réalisation technique sont explicitées.

L'algorithme d'inversion BRAIN (voir annexe A) utilisé dans cette étude pour l'estimation de taux de pluie instantanés a été validé sur l'océan et en Afrique de l'Ouest pour le système d'observation TMI/TRMM. Les caractéristiques d'erreurs de BRAIN mises en évidence indiquent que le terme d'algorithmie serait négligeable par rapport au terme d'échantillonnage pour l'estimation de cumuls de pluie. Afin de s'en assurer, la campagne de validation de la mission Megha-Tropiques permettra de caractériser les erreurs d'estimations des taux de pluie BRAIN sur différents régimes de pluie en Afrique, en Inde et au Brésil. Les résultats de cette validation permettront alors de réévaluer la contribution réelle de l'erreur algorithmique à l'erreur totale des estimations TAPEER-BRAIN.

La validation des estimations TAPEER-BRAIN réalisée dans ce chapitre montre que la qualité

des estimations produites est équivalente à celle des cumuls dérivés du produit TRMM-3B42 (les résultats d'une évaluation de TAPEER-BRAIN aux échelles d'accumulation sur 10 jours et à l'échelle de la variabilité des ondes d'Est africaines sont résumés dans l'annexe B).

Les cumuls de pluie et erreurs associées TAPEER-BRAIN présentés dans ce chapitre ont été réalisés à partir des produits de Niveau-2 BRAIN-TMI/TRMM, BRAIN-AMSRE/AQUA et BRAIN-SSMI/DMSF. Pendant la phase d'exploitation de la mission Megha-Tropiques, les estimations TAPEER-BRAIN seront réalisées à partir du produit de Niveau-2 BRAIN-MADRAS ainsi que des produits de Niveau-2 BRAIN issus d'une constellation de satellites défilants. Le volume de données pris en compte dans l'apprentissage de la méthode TAPEER dépend des produits de Niveau-2 utilisés. Afin d'anticiper l'impact du produit BRAIN-MADRAS et d'autres produits de Niveau-2 sur les estimations de pluies cumulées, le chapitre suivant est dédié à la sensibilité de la méthode TAPEER aux caractéristiques de la constellation de satellites disponibles.

Simulations de la constellation GPM dans les Tropiques

4.1 Introduction

Le domaine des micro-ondes est le domaine de fréquence privilégié pour l'observation des pluies (voir annexe A). En presque quarante ans, l'évolution de la densité des observations satellitaires dans ce domaine de fréquence permet d'envisager la mesure quantitative des précipitations depuis l'espace, à des résolutions spatio-temporelles de plus en plus fines.

Sonder l'atmosphère dans le domaine des micro-ondes à partir d'un satellite géostationnaire serait la technologie spatiale qui conduirait à la densité d'observations la plus élevée. Cependant, la taille de l'antenne requise est actuellement problématique si l'on veut effectuer des mesures aux basses fréquences micro-ondes sensibles à l'émission/absorption des gouttes d'eau (voir annexe A), à une résolution spatiale équivalente à celle des sondeurs à bord de satellites défilants en orbite basse. Les instruments étudiés pour équiper les futurs satellites géostationnaires sont des sondeurs dans le domaine des longueurs d'ondes sub-millimétriques/millimétriques sensibles au contenu en glace des nuages; la taille de l'antenne nécessaire pour de tels instruments est de plusieurs mètres (exemple du projet Geostationary Observatory for Microwave Atmospheric Sounding (GOMAS) qui permettrait un échantillonnage à la résolution de 10 km / 15 minutes dans le domaine des fréquences 50 à 430 GHz, Bizzarri *et al.*, 2007).

La formation de constellations de satellites en orbite basse est une solution alternative qui permet un échantillonnage accru tout en laissant la possibilité de sonder l'atmosphère dans les basses fréquences micro-ondes. Plusieurs projets tel que FLORAD (FLOWer constellation of millimeter-wave RADiometers and RADio occultation receivers) (Marzano and Cimini, 2010) et Iridium Next (Iridium, 2011) sont basés sur ce principe de constellation. FLORAD est une constellation de satellites sur orbite elliptique portant à leur bord des radiomètres dans le domaine des longueurs d'ondes sub-millimétriques pour la mesure des précipitations. Iridium Next est une constellation de satellites de télécommunication qui embarqueraient chacun des instruments dédiés à l'observation de la Terre (imageurs multi-spectraux, radiomètres à bande large pour l'étude du bilan radiatif Terrestre).

La future constellation GPM (Global Precipitation Measurement, Hou *et al.*, 2011) est dédiée

à la mesure des précipitations par radiométrie micro-onde. Cette constellation est constituée d'un satellite "mère" américano-japonais, le GPM-Core équipé d'un radar bi-fréquence, d'un radiomètre à micro-ondes passives GMI (GPM Microwave Imager) et d'un ensemble de satellites également équipés de radiomètres (figure 4.1). L'objectif de la constellation GPM est de fournir des observations micro-ondes sur l'ensemble du globe, au moins toutes les 3 heures.



Figure 4.1 – Vue d'artiste de la future constellation de satellites, pourvus de radiomètres à micro-ondes passives, issue de la coopération internationale du programme GPM (Source : <http://pmm.nasa.gov/>)

L'échantillonnage temporel d'une constellation de satellites peut se caractériser par deux grandeurs qui ne dépendent que de la latitude du lieu observé : le nombre moyen d'observations effectuées en un lieu donné dans une journée ainsi que la distribution des observations pendant cette journée. Les deux grandeurs sont importantes pour la mesure des précipitations cumulées. Un nombre d'observations élevé est une condition nécessaire pour que l'erreur d'échantillonnage associée à un cumul soit petite (Iida *et al.*, 2005), mais ce n'est pas une condition suffisante. En effet, le temps d'auto-corrélation caractéristique des pluies intégrées à l'échelle de la centaine de kilomètres carrés (taille d'un pixel micro-onde) est de l'ordre de 30 minutes à 1 heure (voir Chap 3). Deux observations successives en un lieu donné sont non corrélées et donc "utiles" à l'estimation de pluies cumulées si l'intervalle de temps qui les sépare est supérieur au temps

d'auto-corrélation de 30 minutes à 1 heure.

Ce temps d'auto-corrélation varie géographiquement et temporellement, il est donc intéressant de définir plusieurs nombres moyens de passages par jour pour une constellation de satellites : (i) N_{TOT} , le nombre moyen de passages par jour prenant en compte toutes les observations et équivalant à un nombre de passages "utiles" si l'auto-corrélation des pluies est nulle, (ii) $N_{0,5h}$, le nombre moyen de passages par jour prenant en compte toutes les observations séparées d'au moins 30 minutes, (iii) et N_{1h} , un nombre moyen de passages par jour prenant en compte toutes les observations séparées d'au moins 1 heure.

Un système d'observation, noté SO, de type SSMI/DMSP (radiomètre / satellite) apporte des observations une fois par jour dans les Tropiques. La figure 4.2 montre que l'échantillonnage d'une constellation de deux SO de type SSMI/DMSP peut être caractérisé par deux passages utiles par jour ou bien un seul si les mesures sont trop rapprochées dans le temps.

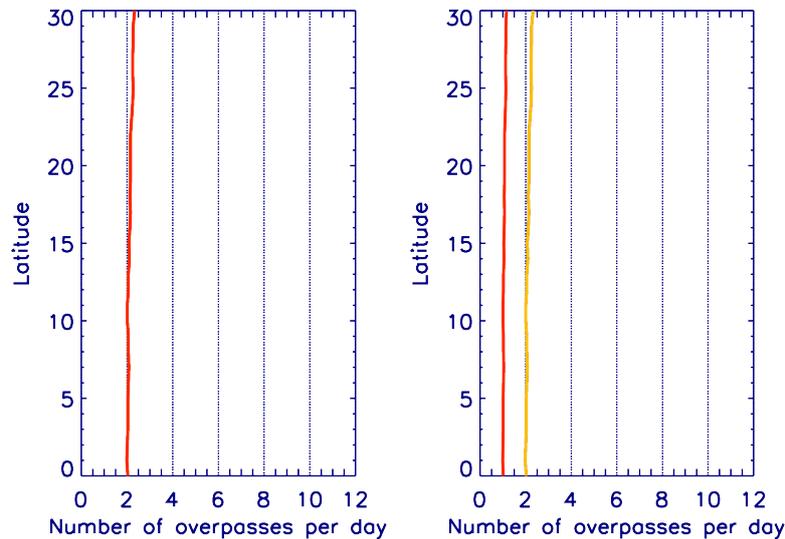


Figure 4.2 – Nombre moyen de passages par jour (N_{TOT} , courbe noire), pour une durée d'auto-corrélation de la pluie de 30 minutes ($N_{0,5h}$, courbe jaune) et de 1 heure (N_{1h} , courbe rouge), correspondant à deux configurations de SO pour les latitudes de 0° à 30° N : 2 SSMI/DMSP à 00H et 17H LTC (à gauche), 2 SSMI/DMSP à 00H et 00H30 LTC (à droite). Sur la figure de gauche, les 3 courbes (noire, jaune et rouge) sont superposées. Sur la figure de droite, la courbe jaune est superposée à la courbe noire.

Chaque ajout d'un nouveau système d'observation amène à une augmentation de N_{TOT} , le nombre total de passages par jour d'une constellation. La figure 4.3 montre que chaque SO de la constellation considérée contribue plus ou moins à l'augmentation de N_{TOT} dans la bande de latitude $[0^\circ \text{ N}; 30^\circ \text{ N}]$. En particulier, la mission Megha-Tropiques apporte la contribution la plus

importante car N_{TOT} augmente de 3 à 5 passages par jour en fonction de la latitude du lieu observé.

Lorsqu'on considère une constellation de SO, l'ensemble des intervalles de temps entre deux passages consécutifs pour un lieu donné forme une distribution. Cette distribution présente des caractéristiques qui varient en fonction de la configuration de la constellation à savoir le nombre de SO et leurs propriétés d'observation.

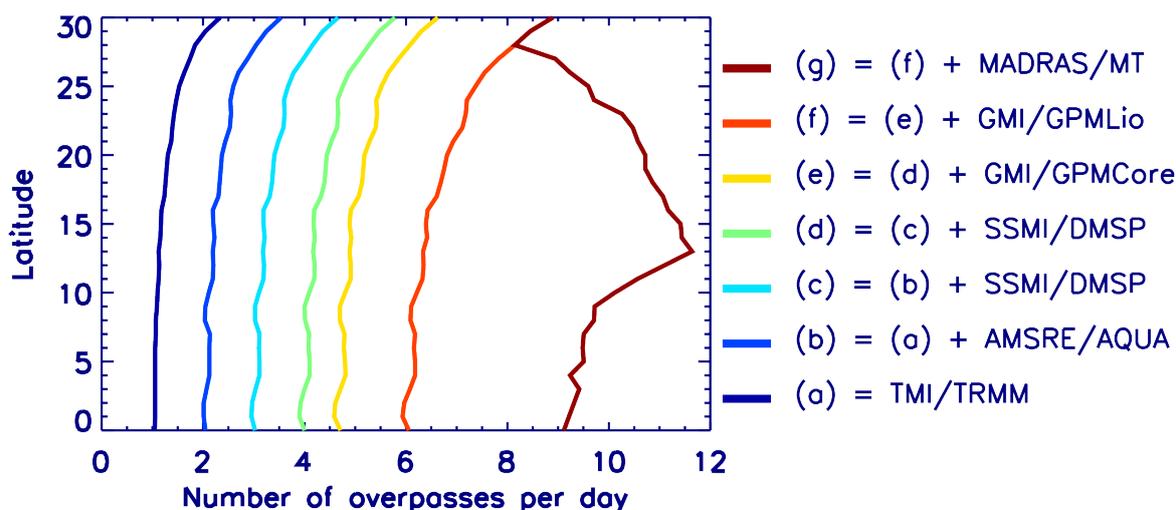


Figure 4.3 – Nombre moyen de passages par jour N_{TOT} pour 7 configurations de constellation de SO pour les latitudes de 0° à 30° . L'évolution de N_{TOT} en fonction de la latitude est symétrique par rapport à l'équateur, seul l'échantillonnage pour l'hémisphère Nord est représenté. Les SO considérés sont : TMI/TRMM, AMSRE/AQUA, SSMI/DMSP, GMI/GPM-Core, GMI/GPM-Lio, MADRAS/Megha-Tropiques). La configuration (a) comporte 1 SO, la configuration (g) comporte 7 SO.

La figure 4.4 montre que pour une constellation formée de 4 SO dont 3 sont en orbite héliosynchrone avec des heures locales de passage à l'équateur espacées de plus d'une heure : (i) 50% des intervalles de temps entre deux passages successifs sont inférieurs à 4 heures à toutes les latitudes, (ii) N_{TOT} , $N_{0,5h}$ et N_{1h} sont quasiment égaux et compris entre 4 et 6 passages par jour selon la latitude du lieu observé.

La figure 4.5 montre que pour une constellation formée des 4 mêmes SO et de MADRAS/Megha-Tropiques : (i) 50% des intervalles de temps entre deux passages successifs sont inférieurs à 1,5 heures à toutes les latitudes, (ii) pour un temps d'auto-corrélation caractéristique des pluies de 0,5 h, une observation par jour n'est pas utile pour l'estimation des précipitations cumulées, (iii) pour un temps d'auto-corrélation caractéristique des pluies de 1 heure, ce sont deux observations par jour qui ne sont pas utiles.

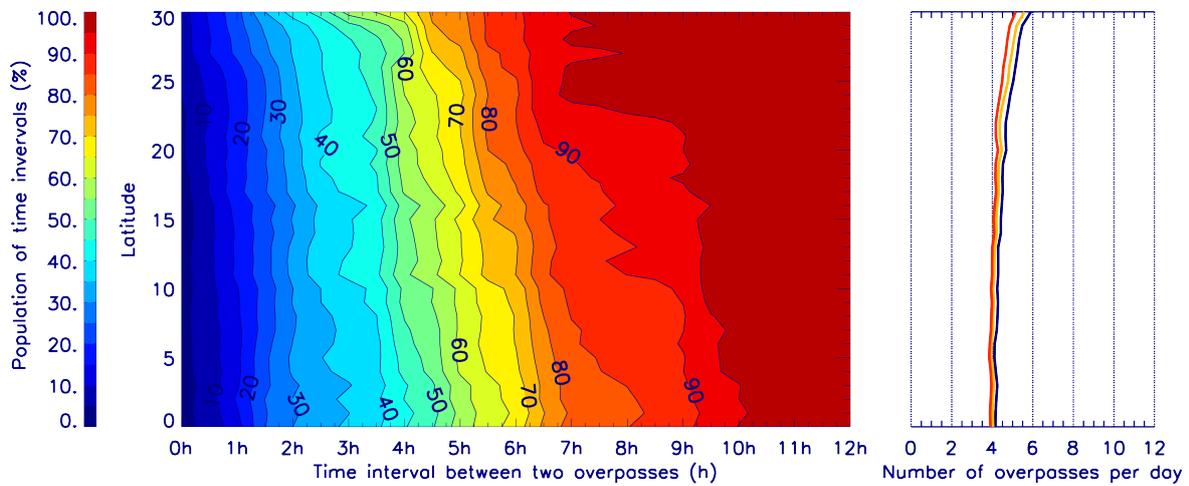


Figure 4.4 – Distribution cumulée des intervalles de temps entre deux passages successifs en fonction de la latitude du lieu observé, pour une constellation formée de 4 SO : 2 SSMI/DMSF à 00H et 17H LTC, AMSRE/AQUA à 13h30 LTC et TMI/TRMM. N_{TOT} , $N_{0,5h}$ et N_{1h} sont tracés en fonction de la latitude respectivement en noir, jaune et rouge.

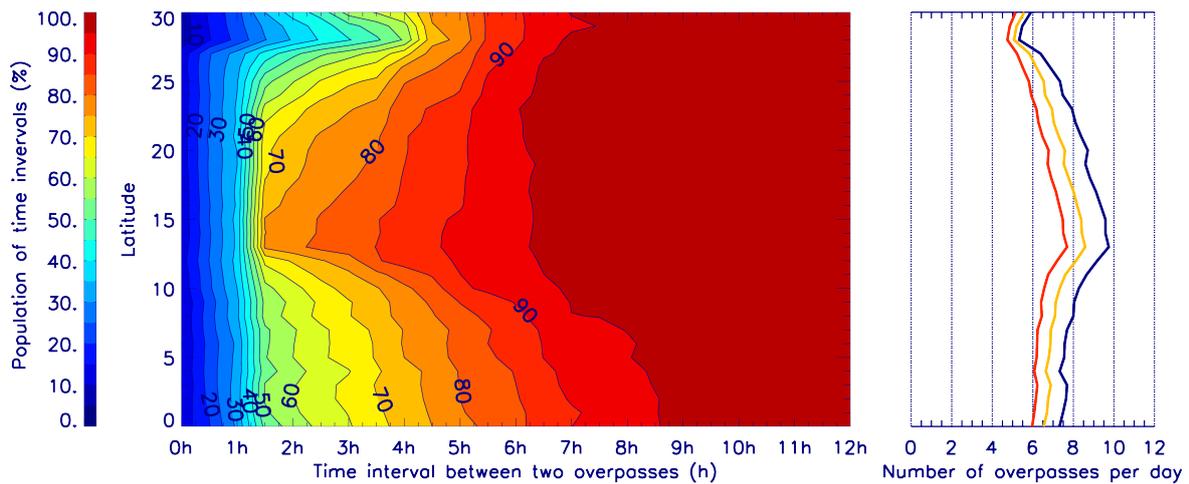


Figure 4.5 – Idem que pour la figure 4.4, pour une constellation de 5 SO dont MADRAS/Megha-Tropiques et les 4 SO de la figure 4.4

Le mode d'observation de MADRAS/Megha-Tropiques a été conçu pour étudier le bilan d'énergie des systèmes convectifs de mousson (Roca *et al.*, 2011). Dans la bande de latitude $[10^\circ \text{ N}; 20^\circ \text{ N}]$, les passages de MADRAS/Megha-Tropiques sont rapprochés dans le temps (101 minutes au plus court) ce qui permet d'échantillonner des systèmes convectifs plusieurs fois pendant leurs cycles de vie (Fioleau and Roca, 2011a). Au sein d'une constellation, la présence de MADRAS/Megha-Tropiques ajoute des observations qui sont séparées d'une durée proche des temps d'auto-corrélation caractéristiques de la pluie.

La simulation de différentes configurations de constellation permet à la fois de mieux appréhender mais également d'anticiper l'apport de systèmes d'observation variés pour les estimations TAPEER-BRAIN. Le nombre de plateformes équipées de radiomètres à micro-ondes passives varie en effet d'année en année, avec les nouvelles missions d'observations de la Terre programmées ainsi que les aléas des lancements ou des problèmes techniques rencontrés.

L'étude présentée dans ce chapitre vise à montrer quels sont les paramètres clés de la configuration d'une constellation de satellites pour l'estimation des pluies cumulées par combinaison IR/MO. Ainsi, on s'intéresse à quantifier l'impact des caractéristiques d'une constellation idéalisée que sont le nombre de SO présents dans la constellation et leurs orbites respectives (orbites tropicales ou orbites héliosynchrones caractérisées par l'heure locale de passage à l'équateur). Dans la suite de ce manuscrit, on appellera "produit de Niveau-2" les estimations de taux de pluie instantanés réalisées à partir des mesures de températures de brillance micro-ondes d'un SO donné.

Une première approche possible consiste à utiliser des produits de Niveau-2 issus des observations de radiomètres à bord de satellites existants. On peut alors évaluer différentes configurations de constellation pour l'estimation de cumuls des précipitations (exemple avec le produit CMORPH, Joyce and Xie, 2011). La limitation principale de cette méthode est que l'on ne peut tester qu'une constellation existante.

Il est aussi possible d'utiliser une base de données, telles que des données radar, afin de simuler des produits de Niveau-2. Ces produits simulés sont dans ce cas des échantillons de la base de données, sélectionnés selon le mode d'échantillonnage d'une constellation (exemple avec des données radar, Xu *et al.*, 1999). On peut ainsi évaluer l'impact de constellations futures. Cette méthode pose cependant un problème de représentativité des résultats, car on n'étudie que le régime pluviométrique dans la zone d'acquisition de la base de données utilisée.

L'approche développée dans ce chapitre consiste à simuler des produits de Niveau-2, non pas à partir de données de surface mais à partir de données satellites à haute résolution spatio-temporelle. Cette démarche permet d'évaluer l'impact de constellations idéalisées sur de grandes zones géographiques. La méthode de génération des produits de Niveau-2 simulés est détaillée dans la section 4.2, les résultats de l'impact des caractéristiques d'une constellation sur les pluies cumulées TAPEER sont ensuite présentés dans les sections 4.3 et 4.4.

4.2 Méthodologie

4.2.1 Simulation de produits de Niveau-2

Le produit d'estimation des précipitations GSMaP est utilisé afin de simuler des produits de Niveau-2. GSMaP est fourni toutes les heures à haute résolution spatiale ($0,1^\circ$) (Aonashi *et al.*, 2009; Ushio *et al.*, 2009).

La première étape consiste à calculer les positions géographiques des pixels d'un SO donné à l'aide du logiciel IXION (Capderou, 2005, 2009). La figure 4.6 montre un exemple des fauchées de MADRAS/Megha-Tropiques ; une fauchée comporte 224 positions de pixels (pour le canal à 89 GHz) de 6 km de rayon.

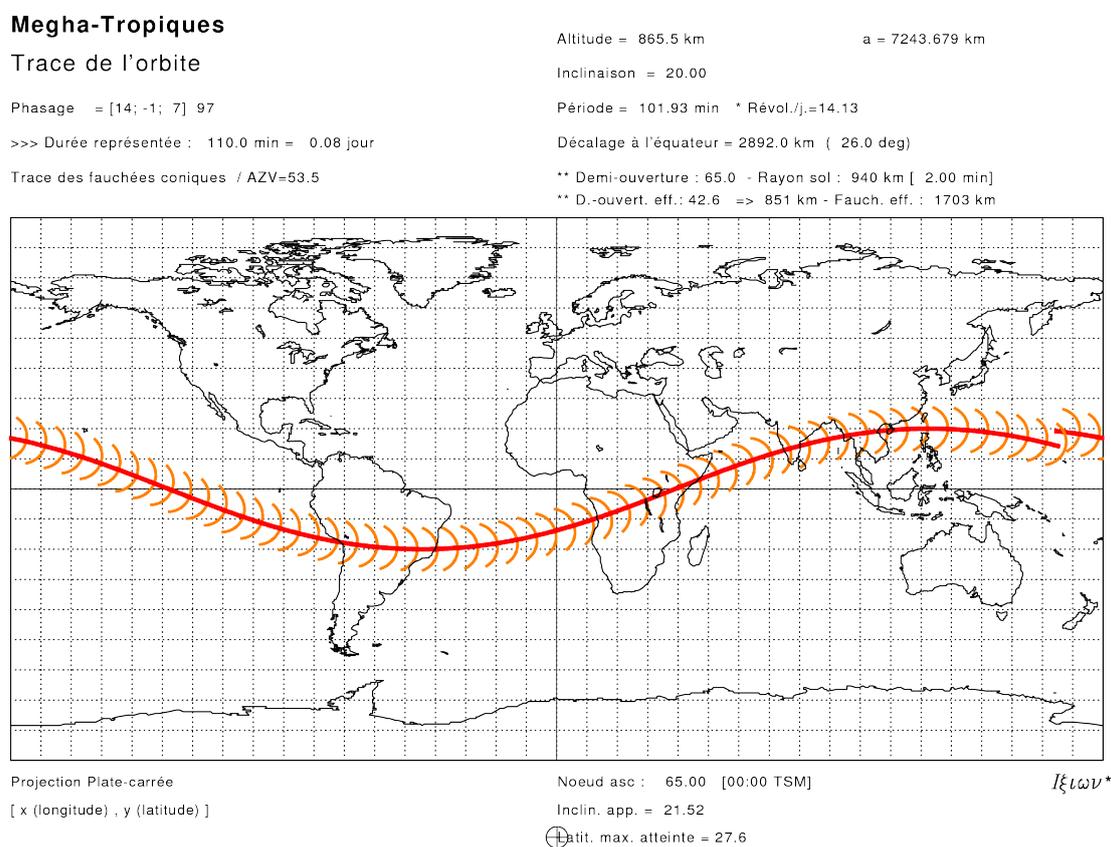


Figure 4.6 – Projection au sol de l'orbite de Megha-Tropiques (en rouge) pour 110 minutes d'observations. Les fauchées du radiomètre MADRAS sont représentées en orange ; seulement une fauchée toutes les 2 minutes est représentée au lieu d'une fauchée toutes les 2,5 secondes. Source : IXION (Capderou, 2005)

La seconde étape de la simulation consiste à attribuer un taux de pluie aux positions des

pixels. Chaque pixel est colocalisé dans l'espace avec les données GSMaP à l'aide de la méthode Point Spread Function (Gif *et al.*, 2011). Pour pouvoir attribuer un taux de pluie à chaque position de pixel à partir des champs de pluie GSMaP horaire, la colocalisation temporelle est au maximum de trente minutes.

La figure 4.7 montre un exemple de champ de pluie GSMaP, ainsi que le résultat de la procédure de création d'un produit de Niveau-2 simulé. La position géographique des fauchées générées pour ce cas permet de simuler des estimations de taux de pluie instantanés pour deux évènements pluvieux dans la baie du Bengale et en mer de Chine.

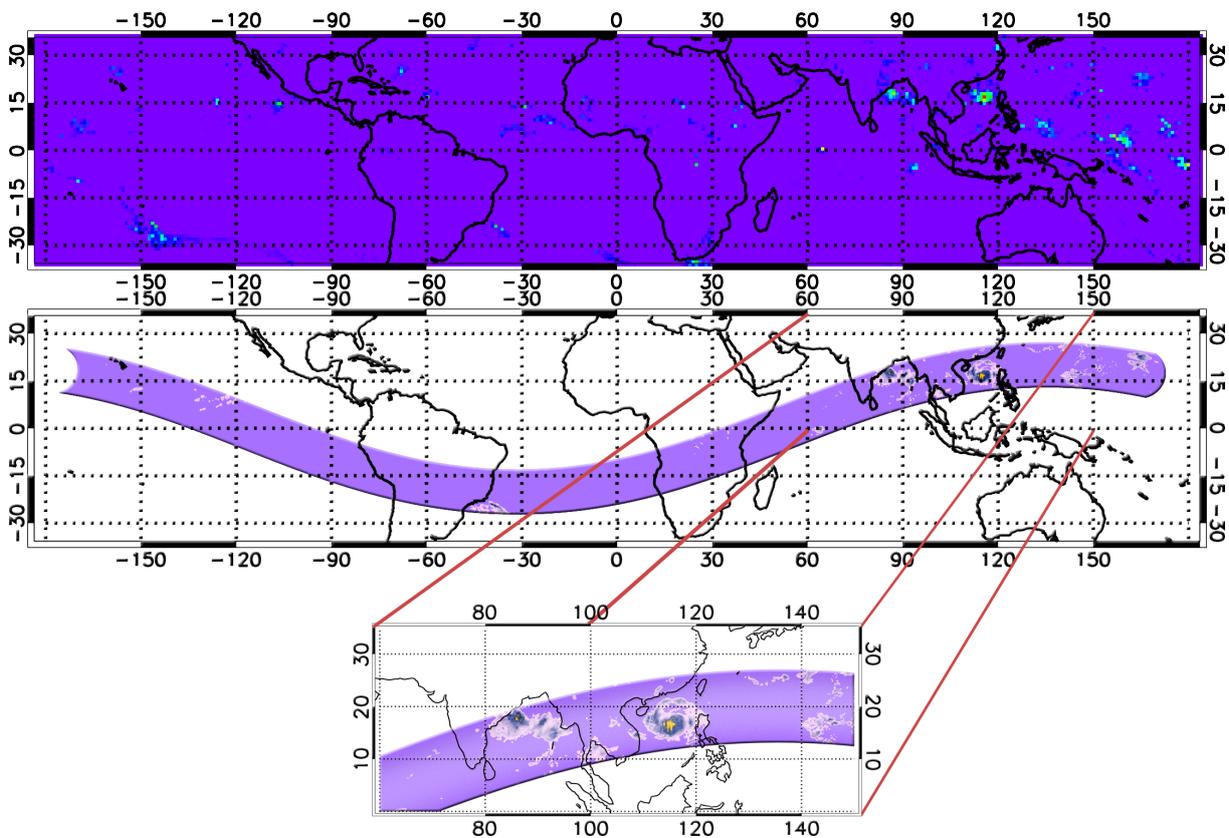


Figure 4.7 – Figure du haut : champ GSMaP du 1^{er} août 2006 à 15H00-UTC. Figure du bas : illustration de la simulation d'un produit de Niveau-2 pour MADRAS/Megha-Tropiques pendant 110 minutes d'observations.

Un produit de Niveau-2 est caractérisé par sa résolution spatiale qui diffère selon le système d'observation considéré (voir annexe A). La variété des résolutions de produits de Niveau-2 induit des erreurs dans l'estimation de cumuls de pluie. Dans ce chapitre, on étudie l'influence des caractéristiques d'une constellation de satellites, en s'affranchissant des problèmes de résolution. Chaque pixel des produits de Niveau-2 simulés présente donc une résolution commune (rayon de 6 km).

La fauchée de MADRAS/Megha-Tropiques a une largeur effective de l'ordre de 1700 km. Les fauchées respectives de SSMI/DMSP, AMSRE/AQUA et SSMIS/DMSP sont respectivement de 1500, 1500 et 1700 km. Dans cette étude, on fait donc l'hypothèse que chaque SO simulé, sur orbite tropicale ou sur orbite héliosynchrone, présente la même fauchée que MADRAS/Megha-Tropiques. L'acquisition des 224 observations d'une fauchée s'effectue alors toutes les 2,5 secondes comme pour Megha-Tropiques. On simule ainsi des produits de Niveau-2 pour sept SO différents sur la période de juin à septembre 2006. Cinq d'entre eux sont sur des orbites héliosynchrones dont les heures locales de passage à l'équateur au noeud ascendant sont respectivement 00H, 03H, 06H, 09H, 13H30. Les deux derniers produits de Niveau-2 simulent chacun le SO MADRAS/Megha-Tropiques ; ces deux SO sont placés en quadrature de phase l'un par rapport à l'autre.

La méthode TAPEER (voir Annexe B) permet d'estimer des cumuls de pluie et leurs erreurs associées, à la résolution $1^\circ/1$ -jour, pour chacune des constellations idéalisées choisies, à partir des produits de Niveau-2 simulés correspondant. Ces estimations de cumuls et d'erreurs seront appelées estimations TAPEER-GSMaP dans la suite de ce chapitre.

Le produit GSMaP original est pris comme référence pour les comparaisons des estimations TAPEER-GSMaP afin de mesurer l'impact de la configuration des constellations simulées. En effet, l'étude menée dans le chapitre 3 a montré que les 2 termes prépondérants du bilan d'erreurs de la méthode TAPEER sont : (i) l'erreur d'échantillonnage, (ii) l'erreur algorithmique liée aux biais systématiques des produits de Niveau-2. Dans le cas des estimations TAPEER-GSMaP, il est possible d'estimer le premier terme d'échantillonnage, pour chaque cas de configuration de constellation, en recalculant $T_{threshold}$ et R_{COND} et les variogrammes spatio-temporels (voir Annexe B). Les produits de Niveau-2 simulés étant des sous-échantillons des champs de pluie GSMaP, ils ne présentent par construction aucun biais par rapport aux champs de référence. Choisir le produit GSMaP original comme la référence permet donc d'annuler le terme algorithmique du bilan d'erreurs.

Des variogrammes spatiaux et temporels sont aussi calculés pour le produit GSMaP afin de construire les estimations de référence (cumuls et erreurs) à la résolution de $1^\circ/1$ -jour. Les scores de comparaison $BIAS_{reg}$, RMS_{reg} , POD et FAR prenant en compte les erreurs de chaque estimation ont été introduits au chapitre 2. Ils permettent d'évaluer les résultats des simulations dans leur intégralité (cumuls et erreurs). Afin de faciliter les comparaisons de simulations, un score intégré (noté F) est introduit. Il s'agit de la somme de chacun des scores normalisé par sa variance (Eq. 4.1).

Les poids w_{Bias} , w_{RMS} , w_{FAR} valent $+1$ et w_{POD} vaut -1 . Ainsi plus le score intégré F est faible, meilleures sont les estimations TAPEER-GSMaP (en fixant tous les poids à ± 1 , on ne privilégie aucun score dans l'évaluation de TAPEER mais il serait possible de pondérer différemment les scores afin de privilégier une caractéristique d'erreurs particulière des estimations

TAPEER).

$$F = 1 + w_{Bias} \cdot \left| \frac{BIAS_{Reg}}{\langle R \rangle_{Ref}} \right| + w_{RMS} \cdot \frac{RMS_{Reg}}{\langle R \rangle_{Ref}} + w_{POD} \cdot POD + w_{FAR} \cdot FAR \quad (4.1)$$

4.2.2 Degrés de liberté de la méthode TAPEER : configuration de la constellation de satellites et volume spatio-temporel d'apprentissage

La méthode TAPEER comprend une phase d'apprentissage effectuée à partir de l'imagerie IR géostationnaire et des taux de pluie issus des observations d'une constellation de satellites défilants. Pendant cette phase décrite dans l'annexe B, deux paramètres $T_{threshold}$ et R_{COND} sont calculés à partir des données disponibles dans un certain volume d'apprentissage.

Différentes tailles de zone géographique et durées de période d'apprentissage ont été utilisées et évaluées dans la littérature pour des méthodes similaires à la méthode TAPEER : des périodes d'apprentissage mensuelles (Huffman *et al.*, 1997, 2001), des périodes plus courtes de quelques jours (Kidd *et al.*, 2003) jusqu'à des apprentissages par évènement pluvieux (Kidd, 1999) ou par passage d'un SO micro-ondes et pour des zones géographiques à l'échelle régionale (Turk and Miller, 2005).

Ces études ont montré que pour une échelle d'accumulation des précipitations donnée, les meilleures valeurs estimées des paramètres $T_{threshold}$ et R_{COND} sont obtenues lorsque la taille du volume d'apprentissage est proche de l'échelle d'accumulation. La phase d'apprentissage permet de déterminer une relation locale IR/taux de pluie (voir annexe B). Certaines conditions météorologiques peuvent influencer sur la relation IR/taux de pluie. Ces phénomènes météorologiques sont caractérisés par des temps et distances typiques. Dans un espace spatio-temporel de taille inférieure ou égale à ces temps et distances caractéristiques, on peut considérer que la relation IR/taux de pluie est à peu près stable. Toutefois, plus ces temps et distances sont courts, moins le volume d'apprentissage ad hoc est susceptible de contenir un évènement pluvieux observé par micro-ondes. Le nombre de pixels pluvieux disponible pour la phase d'apprentissage dépend alors de la configuration de la constellation de satellites prise en compte. Avec peu d'échantillons, la relation IR/taux de pluie est difficile à estimer car le calcul des deux paramètres $T_{threshold}$ et R_{COND} nécessite un minimum d'occurrence de pluie (Kidd, 1999; Todd *et al.*, 2001).

Ainsi, une étude d'impact sur l'estimation des pluies cumulées d'une constellation de satellites défilants comporte trois degrés de liberté. Le premier est la configuration de la constellation, le deuxième est la taille de la zone géographique d'apprentissage, le troisième est la durée de la période d'apprentissage.

La figure 4.8 illustre l'impact de ces trois degrés de liberté par un exemple d'estimations des précipitations sur 10 jours issues de la méthode TAPEER. En particulier, on s'intéresse à la zone géographique [86° E - 88° E; 18° N - 20° N] sur la côte Indienne au nord de la baie du

Bengale. Cette zone est composée de quatre $1^\circ \times 1^\circ$ présentant des cumuls de pluie importants. Le maximum de cette zone n'est pas reproduit par toutes les simulations. À partir des quatre estimations de pluie de chaque simulation sur cette zone géographique, reportées dans le tableau 4.1, on peut calculer un score FEBO (Frequency of Error Bars Overlap, voir chapitre 2 pour la définition) :

- le score FEBO vaut 100% pour les simulations réalisées avec un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour. Ce volume permet de reproduire les QPE de référence sur cette zone
- le score FEBO est inférieur à 75% pour les simulations réalisées avec un volume d'apprentissage de $5^\circ \times 5^\circ / 5$ -jours.

	Cumul (mm)		Barres d'erreurs (mm)		Recouvrement barres d'erreurs	
Référence GSMaP	515,6	659,8	$\pm 43,0$	$\pm 87,2$	-	-
	487,6	502,4	$\pm 40,3$	$\pm 62,9$	-	-
Un SO - $1^\circ \times 1^\circ / 1$-jour	414,6	563,9	$\pm 68,9$	$\pm 99,2$	✓	✓
	623,2	556,1	$\pm 113,8$	$\pm 89,6$	✓	✓
Un SO - $5^\circ \times 5^\circ / 5$-jours	393,2	393,3	$\pm 95,9$	$\pm 97,9$	✓	✗
	458,7	458,2	$\pm 96,0$	$\pm 99,0$	✓	✓
Trois SO - $1^\circ \times 1^\circ / 1$-jour	551,1	602,3	$\pm 55,1$	$\pm 94,2$	✓	✓
	524,2	514,2	$\pm 55,9$	$\pm 82,7$	✓	✓
Trois SO - $5^\circ \times 5^\circ / 5$-jours	380,2	382,4	$\pm 45,2$	$\pm 55,4$	✗	✗
	440,1	434,4	$\pm 45,4$	$\pm 56,1$	✓	✓

Tableau 4.1 – Estimations de cumuls de pluie sur 10 jours pendant la dernière décade d'août 2006, pour la zone géographique [86° E - 88° E; 18° N - 20° N] (voir figure 4.8) composée de quatre zones de $1^\circ \times 1^\circ$, par la référence GSMaP et par les estimations TAPEER-GSMaP issues de 4 combinaisons différentes de constellations et de volume d'apprentissage : (i) un seul SO et un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (ii) un seul SO et un volume d'apprentissage de $5^\circ \times 5^\circ / 5$ -jours, (iii) 3 SO et un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (iv) 3 SO et un volume d'apprentissage de $5^\circ \times 5^\circ / 5$ -jours. Le symbole ✓ (resp. ✗) indique qu'il y a recouvrement (resp. il n'y a pas recouvrement) entre les intervalles de confiance à 68% des estimations TAPEER-GSMaP et de la référence.

Une comparaison systématique des résultats de chaque simulation avec la référence, sur la région étendue à la boîte "mousson indienne", montre que dans ce cas, les biais et rms les plus forts sont associés aux simulations dont le volume d'apprentissage est de $1^\circ \times 1^\circ / 1$ -jour (voir tableau 4.2). On constate aussi une amélioration de la détection et des biais et rms de régression

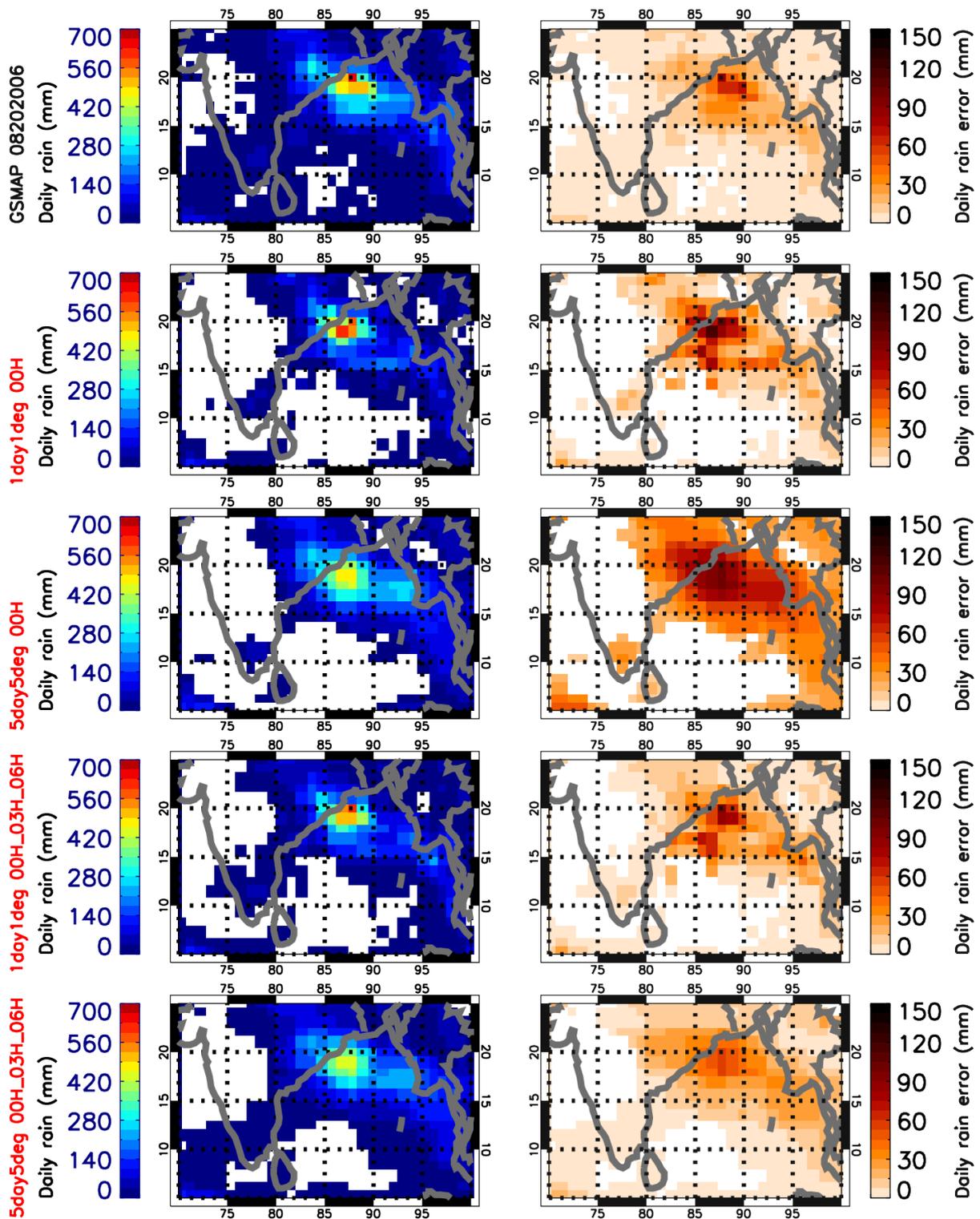


Figure 4.8 – Cartes d’estimations de cumuls de pluie pour la dernière décade du mois août 2006 (colonne de gauche) et d’erreurs d’échantillonnage associées (colonne de droite). La première ligne correspond à la référence GSMaP. Les quatre lignes suivantes correspondent à des estimations TAPEER-GSMaP pour : (i) 1 SO et un volume d’apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (ii) 1 SO et un volume de $5^\circ \times 5^\circ / 5$ -jours, (iii) 3 SO et un volume de $1^\circ \times 1^\circ / 1$ -jour, (iv) 3 SO et un volume de $5^\circ \times 5^\circ / 5$ -jours.

pour les simulations avec une constellation de 3 SO par rapport à un seul SO ce qui révèle une sensibilité de la méthode TAPEER à l'augmentation de N_{TOT} . Pour ce cas, N_{TOT} , $N_{0,5h}$ et N_{1h} sont égaux car les SO pour ces simulations sont sur des orbites héliosynchrones dont les heures locales de passage à l'équateur sont respectivement 00H, 03H et 06H. Les intervalles de temps entre 2 passages successifs pour un lieu donné sont donc d'au moins 3h.

	$BIAS_{reg}$ (mm)	RMS_{reg} (mm)	POD	FAR
Un SO - $1^\circ \times 1^\circ$/1-jour	-25,6	3,6	0,58	0,00
Un SO - $5^\circ \times 5^\circ$/5-jours	2,4	0,03	0,56	0,08
Trois SO - $1^\circ \times 1^\circ$/1-jour	-18,0	0,1	0,68	0,01
Trois SO - $5^\circ \times 5^\circ$/5-jours	-0,2	0,07	0,79	0,08

Tableau 4.2 – Scores de comparaison entre la référence GSMaP et les estimations TAPEER-GSMaP de cumuls sur 10 jours pendant la dernière décade d'août 2006, pour la zone géographique $[70^\circ \text{ E}; 100^\circ \text{ E} - 5^\circ \text{ N}; 25^\circ \text{ N}]$ (figure 4.8). Les simulations TAPEER-GSMaP évaluées correspondent à quatre combinaisons différentes de constellations et de volumes d'apprentissage : (i) un seul SO et un volume d'apprentissage de $1^\circ \times 1^\circ$ /1-jour, (ii) un seul SO et un volume d'apprentissage de $5^\circ \times 5^\circ$ /5-jours, (iii) 3 SO et un volume d'apprentissage de $1^\circ \times 1^\circ$ /1-jour, (iv) 3 SO et un volume d'apprentissage de $5^\circ \times 5^\circ$ /5-jours.

Grâce aux degrés de liberté liés à la phase d'apprentissage de la méthode TAPEER, il est possible d'illustrer la nature de l'erreur évaluée grâce au modèle introduit dans le chapitre 2. Cette erreur est une combinaison d'erreurs qui présente deux composantes : une erreur d'échantillonnage liée à la nature discrète dans l'espace et dans le temps des échantillons qui composent le cumul de pluie et une erreur d'algorithmie liée à la méthode de combinaison de données IR/taux de pluie (voir Annexe B). Les deux composantes de cette combinaison d'erreurs sont indissociables (Huffman, 1997); il est cependant possible de faire varier l'une ou l'autre des contributions. La figure 4.9 montre que 60% des estimations considérées présentent des erreurs relatives qui sont inférieures à 100% dans le cas d'un volume d'apprentissage de $5^\circ \times 5^\circ$ /5-jours et qui sont inférieures à 70% dans le cas du petit volume d'apprentissage de $1^\circ \times 1^\circ$ /1-jour. La diminution de l'erreur modélisée avec le volume d'apprentissage est donc une illustration de la combinaison d'erreurs : l'erreur algorithmique associée à la combinaison de données IR/taux de pluie est plus petite lorsque le volume d'apprentissage est de $1^\circ \times 1^\circ$ /1-jour. Dans un volume de $1^\circ \times 1^\circ$ /1-jour, la relation IR/taux de pluie serait à peu près stable alors que dans un volume de $3^\circ \times 3^\circ$ /3-jours ou $5^\circ \times 5^\circ$ /5-jours, on serait amené à mélanger des situations météorologiques diverses. Dans ces cas, l'erreur d'algorithmie de combinaison de données IR/taux de pluie serait plus forte.

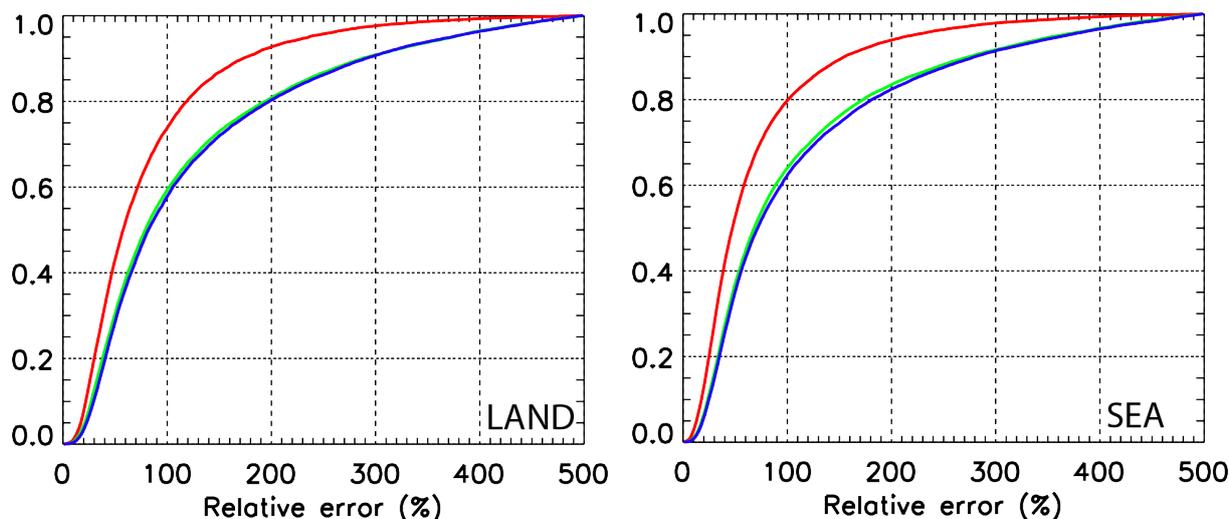


Figure 4.9 – Distributions cumulées des erreurs TAPEER-GSMaP pour la période de juin à septembre 2006, dans la région "mousson indienne" (terre à gauche, mer à droite) pour trois simulations différentes avec les volumes d'apprentissage : $1^\circ \times 1^\circ / 1$ -jour (rouge), $3^\circ \times 3^\circ / 3$ -jours (vert), $5^\circ \times 5^\circ / 5$ -jours (bleu). Chacune des simulations est réalisée à partir des trois mêmes SO sur des orbites héliosynchrones respectivement à 00H, 03H et 06H.

4.3 Étude de sensibilité des estimations TAPEER-GSMaP pour des constellations idéalisées

Différentes constellations idéalisées sont étudiées : (i) les constellations formées d'un seul SO dont le satellite est héliosynchrone, (ii) les constellations formées de plusieurs satellites héliosynchrones, (iii) les constellations formées de plusieurs satellites héliosynchrones et d'un satellite sur une orbite tropicale. Bien que le calcul des variogrammes spatio-temporels pour chaque simulation demande des ressources informatiques importantes, chacune des simulations a été réalisée sur deux régions distinctes : une boîte "mousson Indienne" [$70^\circ \text{ E} - 100^\circ \text{ E}$; $5^\circ \text{ N} - 25^\circ \text{ N}$] (voir figure 4.8) et une boîte "mousson Africaine" [$12^\circ \text{ W} - 12^\circ \text{ E}$; $2^\circ \text{ N} - 17^\circ \text{ N}$]. Les résultats obtenus étant similaires sur les deux régions, ceux-ci sont décrits uniquement pour la boîte "mousson Indienne".

4.3.1 Sensibilité à l'heure locale de passage d'un satellite héliosynchrone

Les résultats des deux simulations présentées dans cette section illustrent l'impact de l'heure locale de passage à l'équateur d'un satellite héliosynchrone sur les estimations TAPEER-GSMaP. Les deux simulations, notées SO/00H et SO/06H, correspondent respectivement à des simulations

avec une seul SO, dont l'heure locale de passage à l'équateur au noeud ascendant est 00H ou 06H.

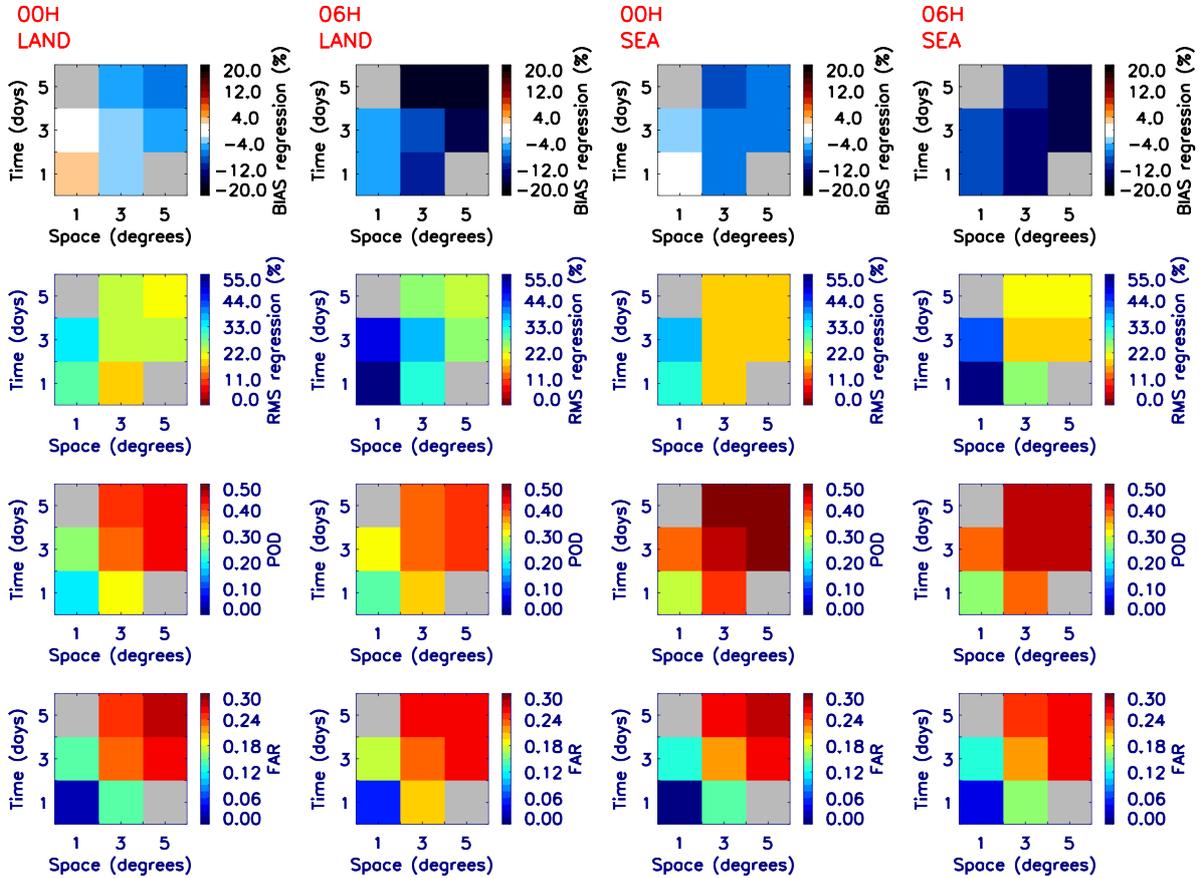


Figure 4.10 – Scores de comparaison des estimations TAPEER-GSMaP pour les simulations SO/00H (1^{ère} et 3^{ième} colonnes) et SO/06H (2^{ième} et 4^{ième} colonnes). Les scores sont calculés pour la boîte "mousson Indienne", pour la terre (1^{ère} et 2^{ième} colonnes) et pour la mer (3^{ième} et 4^{ième} colonnes). Les scores $BIAS_{reg}$, RMS_{reg} , POD et FAR sont reportés respectivement sur les 1^{ère}, 2^{ième}, 3^{ième} et 4^{ième} lignes. Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure.

La figure 4.10 montre comment chacun des scores $BIAS_{reg}$, RMS_{reg} , POD et FAR évolue selon la simulation pour les divers volumes d'apprentissage et la figure 4.11 montre la réponse du score intégré F. Le $BIAS_{reg}$ vaut quelques pourcents sur terre pour tous les volumes d'apprentissage quand le SO/00H est utilisé, alors qu'il décroît jusqu'à -20% avec le SO/06H. On voit aussi que $BIAS_{reg}$ diminue avec la taille du volume d'apprentissage. Sur terre, la RMS_{reg} varie entre 20% et 30% pour le SO/00H tandis que ce score varie entre 30% et 50% pour le SO/06H. La RMS_{reg} augmente lorsqu'on diminue la taille du volume d'apprentissage. Le contraste entre les

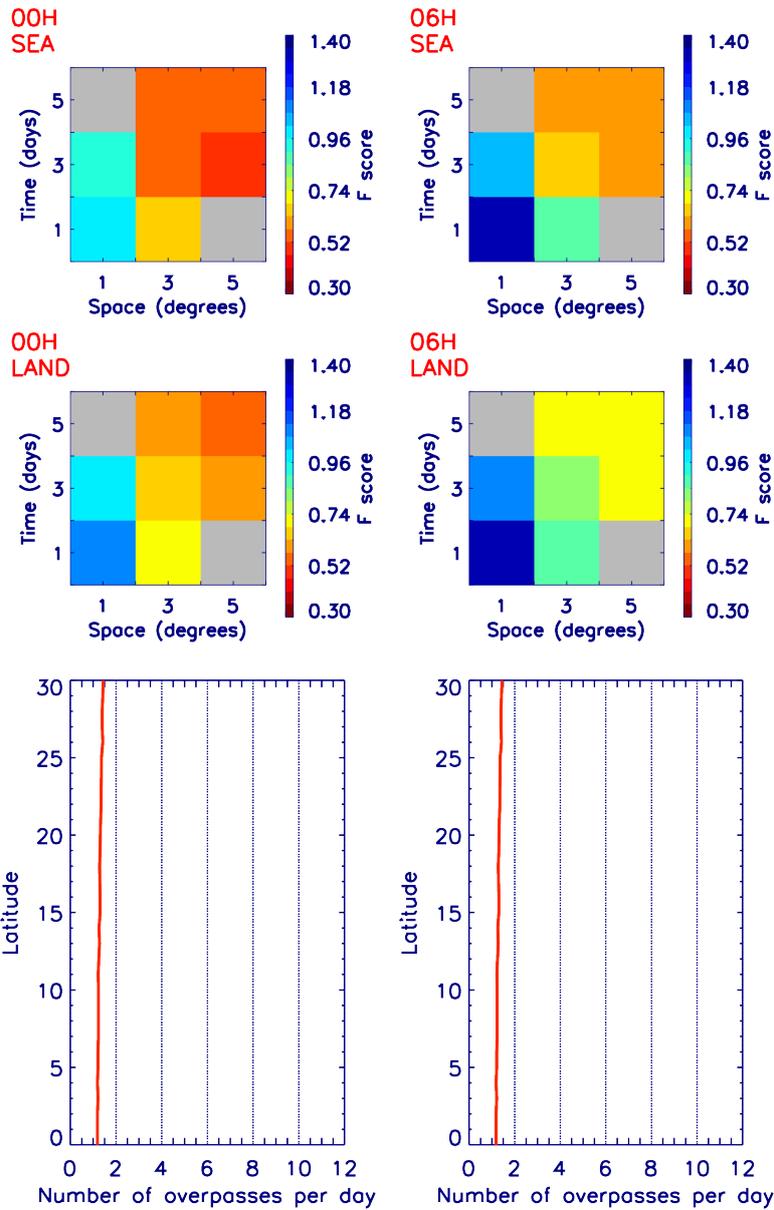


Figure 4.11 – Score intégré F calculé à partir des scores de comparaison des estimations TAPEER-GSMaP pour les simulations SO/00H (1^{ère} colonne) et pour les simulations SO/06H (2^{ème} colonne). Les scores sont calculés pour la boîte "mousson Indienne", pour la mer (1^{ère} ligne) et pour la terre (2^{ème} ligne). Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure. Le nombre moyen de passages par jour en fonction de la latitude du lieu observé pour chacune des 2 configurations est indiqué (3^{ème} ligne). N_{TOT} , $N_{0.5h}$ et N_{1h} sont tracés en noir, jaune et rouge. Les trois courbes (noire, jaune et rouge) sont superposées pour ces deux cas.

deux simulations SO/00H et le SO/06H est similaire sur mer.

En ce qui concerne les scores de détection des précipitations de TAPEER-GSMaP, la sensibilité des scores POD et FAR est équivalente pour les 2 SO, sur terre et mer. Le volume d'apprentissage est le degré de liberté influençant le plus les scores de détection : le score POD chute à moins de 30% pour le volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, tandis que le score FAR monte à presque 30% pour le volume de $5^\circ \times 5^\circ / 5$ -jours.

Sur la figure 4.11, on peut voir que le score F traduit les variations des quatre scores BIAS_{reg} , RMS_{reg} , POD et FAR pour les deux simulations SO/00H et SO/06H. Dans chacun des cas, F décroît quand on augmente la taille du volume d'apprentissage (par exemple dans le cas de SO/06H sur terre : F vaut 0,7 pour le volume de $5^\circ \times 5^\circ / 5$ -jours et 1,4 pour le volume de $1^\circ \times 1^\circ / 1$ -jour).

Sur terre, F varie entre 0,5 et 1,1 pour le SO/00H alors que F varie entre 0,7 et 1,4 pour celui de 06H. Sur mer, F varie entre 0,4 et 1,0 pour le SO/00H et entre 0,6 et 1,4 pour le SO/06H. Ainsi, dans notre région d'étude, c'est le SO/00H qui donne les meilleurs scores.

Le cycle diurne des précipitations est très variable d'une région à l'autre des Tropiques (Yang and Smith, 2006). Il est en général plus marqué sur terre que sur mer bien que la côte Est de la baie du Bengale présente une forte dépendance des précipitations aux effets diurnes (Islam and Ueyda, 2006). La variabilité du score F selon l'heure locale de passage du satellite est en effet à la fois observée sur terre et sur mer. Ces simulations montrent que dans le cas où TAPEER est alimenté par un seul instrument micro-ondes, les estimations des précipitations cumulées sont très variables selon la fraction du cycle diurne observée. D'autre part, ces résultats suggèrent que la taille du volume d'apprentissage influe au moins autant que la sélection des heures de passage du système d'observation.

4.3.2 Sensibilité au nombre de satellites héliosynchrones dans une constellation

L'objectif de la deuxième série de simulations est de quantifier l'impact du nombre de SO dans la constellation. Deux configurations sont testées avec 3 et 5 SO, dont les satellites sont tous placés sur orbite héliosynchrone, avec des heures locales de passage au noeud ascendant respectivement à 00H, 03H, 06H, 09H et 13H30.

La figure 4.12 reporte le score F pour ces deux configurations de constellation. On constate d'une part que la configuration avec 3 SO amène à des valeurs de F bien meilleures que celles calculées pour les configurations pourvues d'un seul SO. D'autre part, on voit que le score F est plus faible pour tous les volumes d'apprentissage lorsqu'on passe de 3 à 5 SO. Cette diminution est toutefois plus marquée pour les petits volumes : pour le volume de $1^\circ \times 1^\circ / 1$ -jour, F est de l'ordre de 0,75 pour la configuration de 3 SO et de l'ordre de 0,6 pour celle avec 5 SO.

Dans le cas des simulations avec un seul SO, la diminution de la taille du volume d'ap-

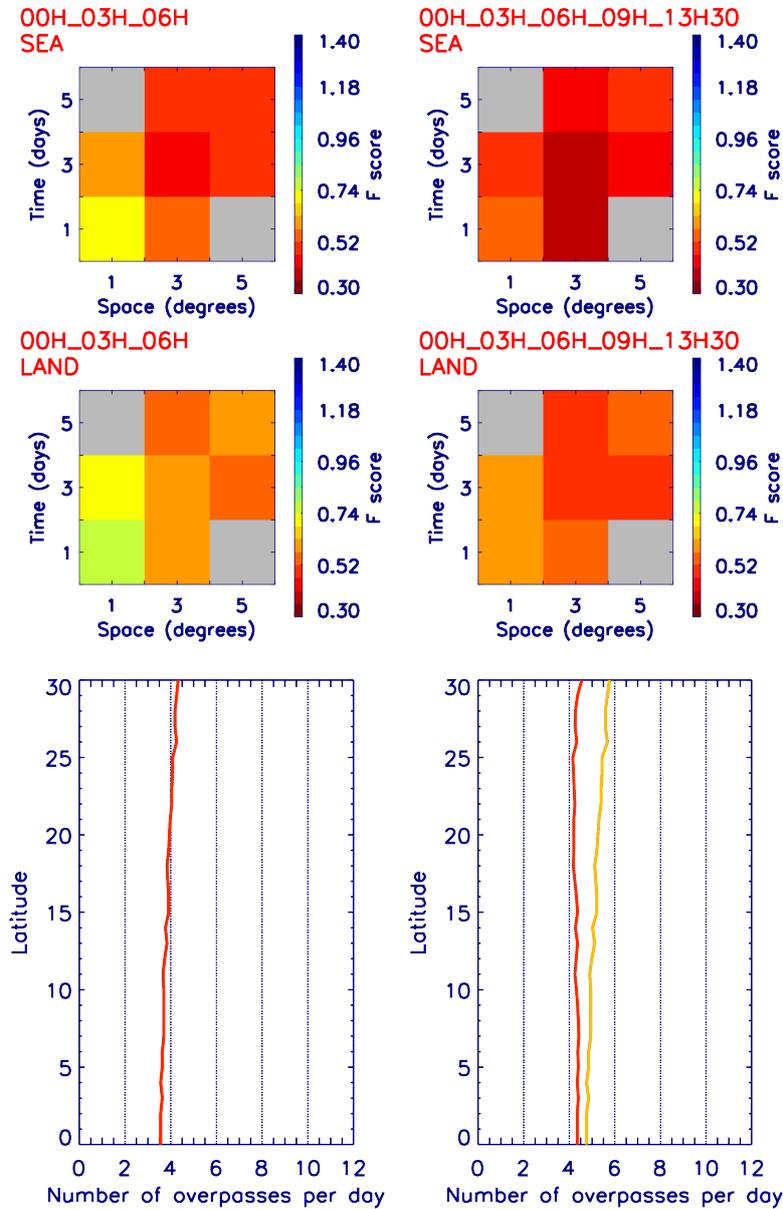


Figure 4.12 – Idem que pour la figure 4.11. Les 2 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/03H et SO/06H (1^{ère} colonne) ; 5 SO dont SO/00H, SO/03H, SO/06H, SO/09H et SO/13H30 (2^{ème} colonne). Les trois courbes (noire, jaune et rouge) sont superposées pour les simulations avec 3 SO. La courbe jaune est superposée avec la courbe noire dans le cas des simulations avec 5 SO.

prentissage dégrade les estimations TAPEER-GSMaP ; il y a de moins en moins d'échantillons disponibles pour la phase d'apprentissage ce qui a pour conséquence une augmentation de F . Dans le cas d'une constellation avec plusieurs SO, le nombre d'échantillons est plus conséquent : le minimum de F correspond alors non pas au volume de $5^\circ \times 5^\circ / 5$ -jours, mais à un volume de $3^\circ \times 3^\circ / 3$ -jours, dans le cas de l'utilisation de 5 satellites.

On observe aussi un phénomène de saturation lorsqu'on passe à 5 SO. Pour le volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, F diminue d'abord de 1,4 à 0,7 puis diminue seulement à 0,6 quand on passe de 1 à 3 puis 5 SO. Cette saturation apparaît aussi pour le volume de $5^\circ \times 5^\circ / 5$ -jours dès la configuration avec 3 SO.

Le calcul des nombres moyens de passages par jour montre que $N_{TOT} = N_{0,5h} = N_{1h}$ est proche de 4 dans le cas des simulations avec 3 SO. En revanche pour les simulations avec 5 SO, le nombre de passages utiles N_{1h} vaut également à peu près 4. Les observations apportées par 2 SO supplémentaires dans la constellation seraient donc corrélées avec les observations déjà disponibles grâce à 3 SO, ce qui pourrait expliquer la saturation observée.

Les résultats de ces simulations amènent aux conclusions suivantes :

- l'utilisation de plusieurs SO améliore les estimations de pluie de la méthode TAPEER
- si le nombre de SO disponible est important, il est possible de diminuer la taille du volume d'apprentissage de la méthode TAPEER, afin d'améliorer les estimations.
- l'apport d'un nouveau SO au sein d'une constellation est faible si le nombre de SO utilisé est déjà conséquent. Les estimations TAPEER sont moins facilement améliorables quand on augmente le nombre de SO dans la constellation.

4.3.3 Sensibilité à la présence d'un satellite placé sur une orbite tropicale dans une constellation

Le nombre total de passages par jour apporté par un SO sur orbite tropicale tel que MADRAS/ Megha-Tropiques est équivalent au nombre de passages apportés par plus de deux SO sur orbite héliosynchrone pour les latitudes $[0^\circ \text{ N} ; 30^\circ \text{ N}]$. En introduction, il a été précisé que la présence de MADRAS/Megha-Tropiques dans une constellation ajoute des observations séparées d'intervalles de temps de durées proches des temps d'auto-corrélation caractéristiques de la pluie (voir figure 4.5). Il est donc important de simuler l'impact d'un SO sur orbite tropicale afin d'évaluer sa contribution réelle à l'estimation des pluies cumulées.

Quatre simulations sont réalisées : la première est la simulation d'une constellation avec 3 SO dont les satellites sont sur orbite héliosynchrone (SO/00H, SO/03H et SO/06H). La deuxième est une simulation avec 2 SO (SO/00H et SO/06H) et MADRAS/Megha-Tropiques. La troisième est une simulation avec 1 SO (SO/00H) et MADRAS/Megha-Tropiques. Enfin, la quatrième est aussi une simulation avec 2 SO (SO/00H et SO/06H) et MADRAS/Megha-Tropiques. Le MT de cette dernière simulation est en quadrature de phase avec le MT de la deuxième et la troisième

simulation.

	Cumul (mm)		Barres d'erreurs (mm)		Recouvrement barres d'erreurs	
	Référence GSMP	238,3	242,7	±37,5	±33,5	-
	297,8	177,8	±30,7	±45,7	-	-
Nb - Orbite - Vol. d'apprentissage						
3 Hélios - 1°×1°/1-jour	-	-	-	-	×	×
	-	-	-	-	×	×
3 Hélios - 5°×5°/5-jours	173,3	136,0	±2,2	±17,6	×	×
	158,0	134,1	±5,9	±24,4	×	✓
2 Hélios et MT - 1°×1°/1-jour	228,2	224,4	±22,6	±43,3	✓	✓
	420,6	205,6	±12,4	±35,9	×	✓
2 Hélios et MT - 5°×5°/5-jours	164,2	131,2	±6,9	±25,4	×	×
	154,9	131,6	±2,9	±19,8	×	✓

Tableau 4.3 – Idem que pour le tableau 4.1, pour la journée du 2 juillet 2006 (voir figure 4.13). Les quatre simulations effectuées utilisent les configurations suivantes : 3 SO sur orbite héliosynchrone et un volume d'apprentissage de 1°×1°/1-jour, 3 SO sur orbite héliosynchrone et un volume d'apprentissage de 5°×5°/5-jours, 3 SO dont 2 sur orbite héliosynchrone et MADRAS/Megha-Tropiques et un volume d'apprentissage de 1°×1°/1-jour, 3 SO dont 2 sur orbite héliosynchrone et MADRAS/Megha-Tropiques et un volume d'apprentissage de 5°×5°/5-jours.

La figure 4.13 illustre l'impact du SO MADRAS/Megha-Tropiques sur un exemple d'estimations de cumuls de pluie journaliers. La journée choisie de juillet 2006 présente une zone de cumuls importants sur la côte Indienne au nord de la baie du Bengale. Cette zone n'est pas du tout détectée pour la simulation avec 3 SO sur orbite héliosynchrone et un volume d'apprentissage de 1°×1°/1-jour. Avec la même constellation, un volume de 5°×5°/5-jours permet de détecter cette zone pluvieuse. En revanche dans la configuration avec 3 satellites dont Megha-Tropiques, la zone de cumuls importants est détectée, même avec le volume de 1°×1°/1-jour.

En particulier, pour la zone [86° E - 88° E; 18° N - 20° N] composée de quatre zones de 1°×1°, toutes les simulations ne reproduisent pas les maxima de la référence. À partir des quatre estimations de pluie de chaque simulation, reportées dans le tableau 4.3, on peut calculer un score FEBO. Ce score vaut 25% pour trois des simulations hormis celle avec MADRAS/Megha-Tropiques et un volume de 1°×1°/1-jour. Cette zone de cumuls maximum se trouve à une des latitudes les plus observées par le SO MADRAS/Megha-Tropiques (voir figure 4.5). Ces résultats

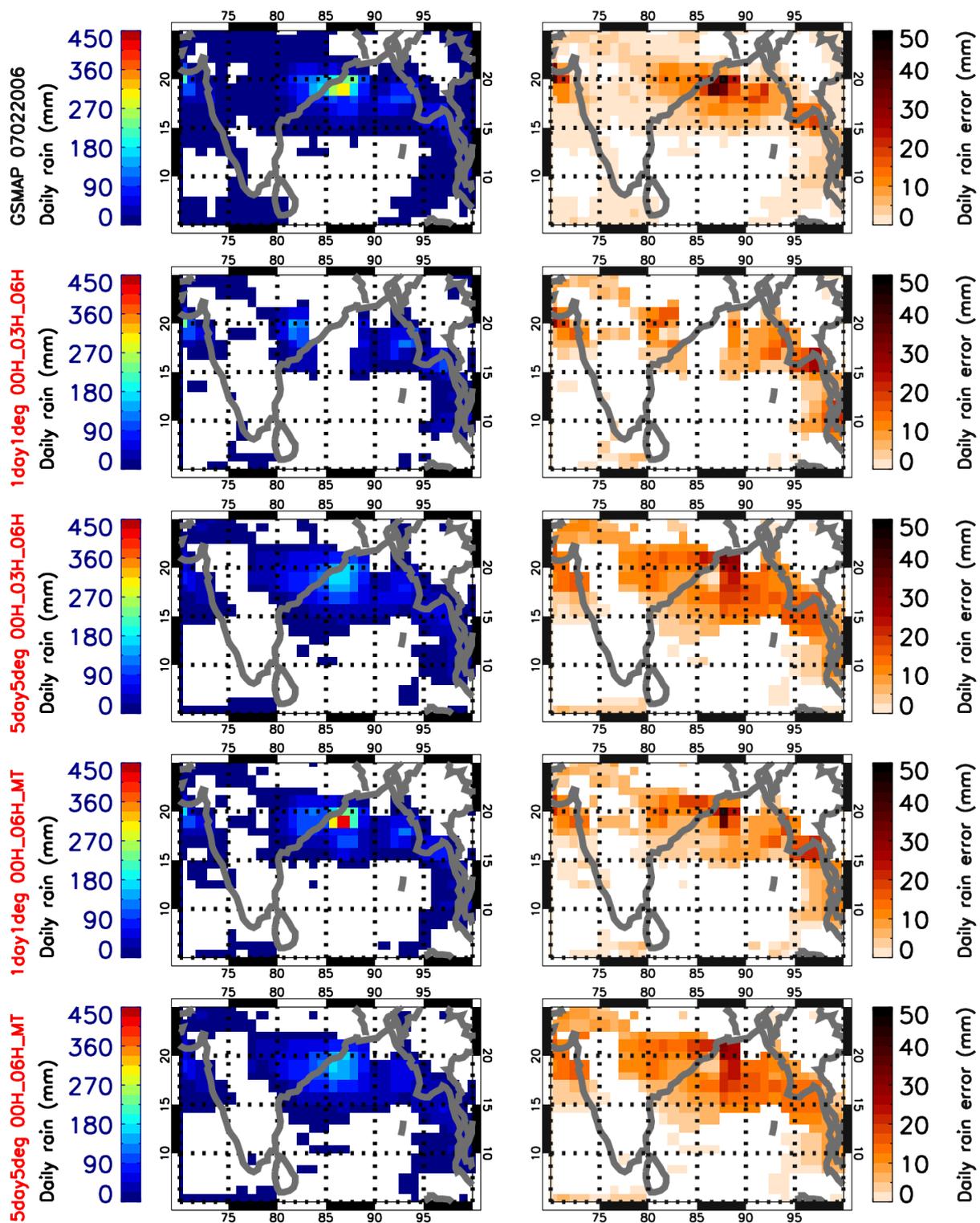


Figure 4.13 – Cartes d'estimations de cumuls de pluie pour le 2 juillet 2006 (colonne de gauche) et d'erreurs d'échantillonnage associées (colonne de droite). La première ligne correspond à la référence GSMaP. Les quatre lignes suivantes correspondent à des estimations TAPEER-GSMaP pour : (i) SO/00H, SO/03H et SO/06H et un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (ii) SO/00H, SO/03H et SO/06H et $5^\circ \times 5^\circ / 5$ -jours, (iii) SO/00H, SO/06H et MADRAS/MT et $1^\circ \times 1^\circ / 1$ -jour, (iv) SO/00H, SO/06H et MADRAS/MT et $5^\circ \times 5^\circ / 5$ -jours.

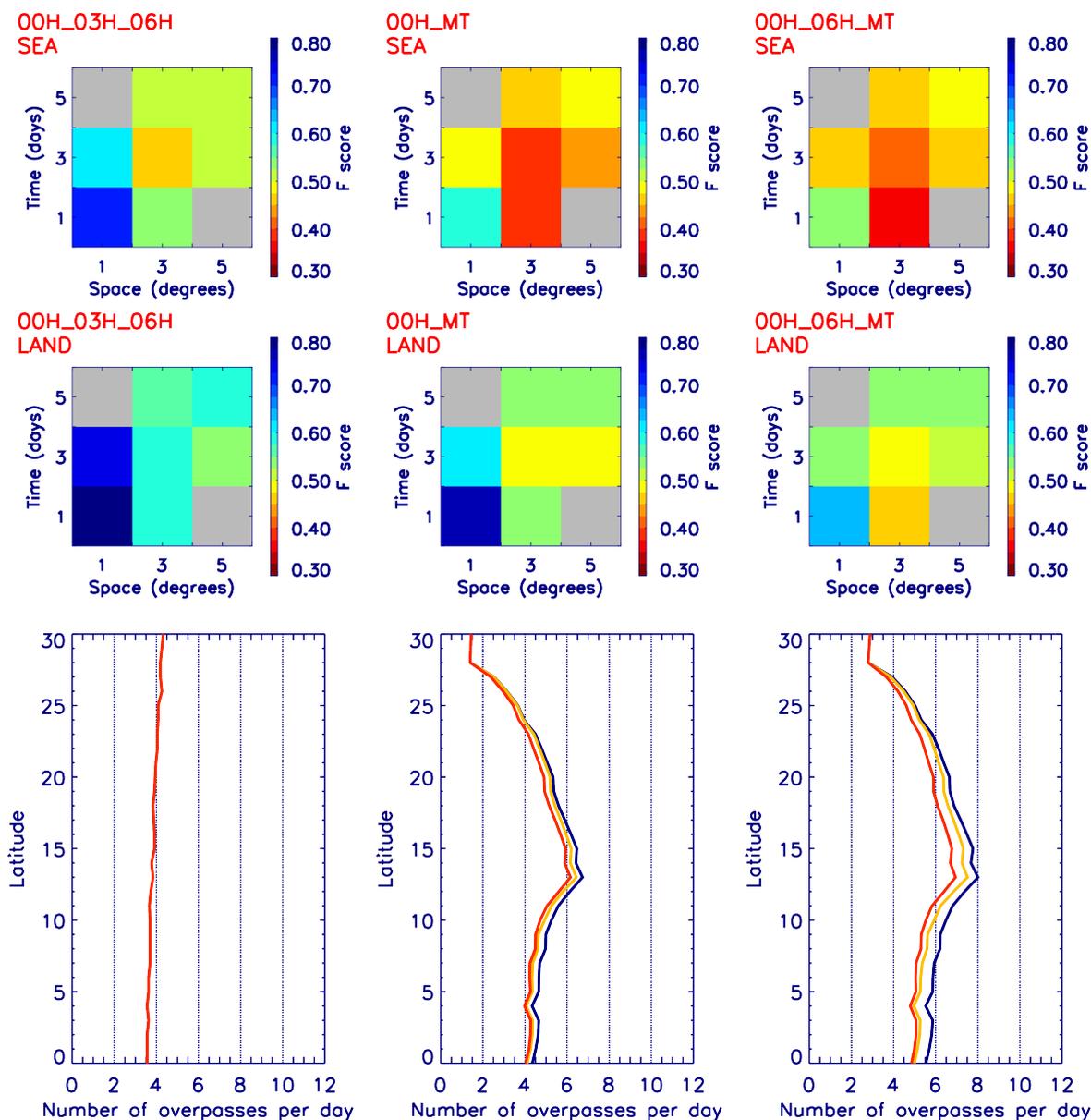


Figure 4.14 – Idem que pour la figure 4.11. Les 3 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/03H et SO/06H (1^{ère} colonne); 2 SO dont SO/00H et MADRAS/Megha-Tropiques (2^{ième} colonne); 3 SO dont SO/00H, SO/06H et MADRAS/Megha-Tropiques (3^{ième} colonne). Les trois courbes (noire, jaune et rouge) sont superposées pour les simulations avec 3 SO sur orbite héliosynchrone.

suggèrent que le mode d'échantillonnage particulier d'une orbite tropicale améliore les estimations de pluie TAPEER.

Les statistiques de comparaison calculées sur toute la saison 2006 confirment le résultat (figure 4.14). On peut voir que le score F est amélioré avec MT au sein de la constellation pour tous les volumes d'apprentissage. Son orbite inclinée lui confère la possibilité de remplacer deux satellites héliosynchrones tout en améliorant tout de même les estimations initiales.

Le calcul de N_{TOT} , $N_{0,5h}$ et N_{1h} montre qu'au maximum une unique observation par jour n'apporte pas d'information pour l'estimation des cumuls, dans le cas d'un temps d'auto-corrélation de la pluie de 1h. Pour les deux simulations avec 3 SO en orbite héliosynchrone et celle avec 3 SO dont MT, N_{1h} augmente au minimum de 4 à 5 au niveau de l'équateur et de 4 à 7 à 15° N ce qui explique les améliorations de F observées.

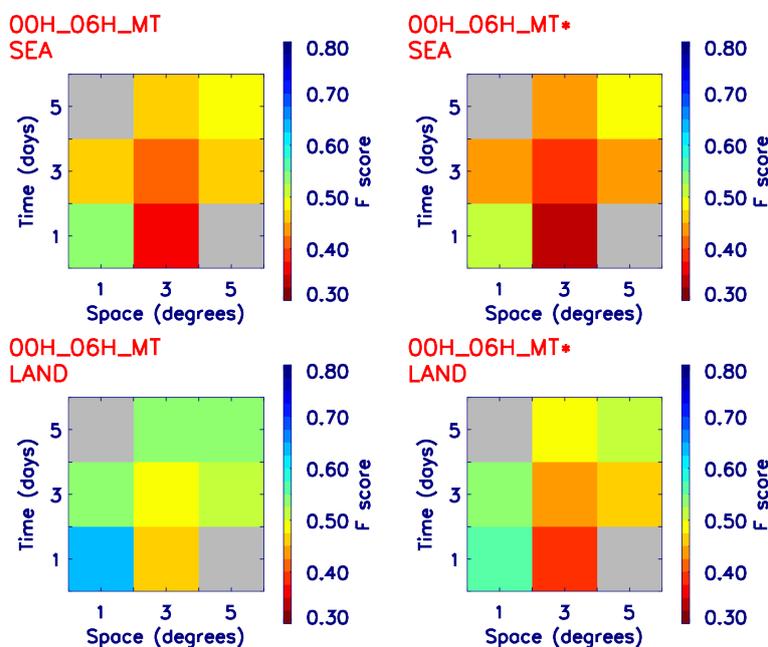


Figure 4.15 – Idem que pour la figure 4.11. Les 2 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/06H et MADRAS/MT (1^{ère} colonne) ; 3 SO dont SO/00H, SO/06H et MADRAS/MT* (2^{ème} colonne). Dans ces simulations, MT et MT* sont en quadrature de phase. Les trois courbes (noire, jaune et rouge) ne sont pas représentées car elles sont identiques à celles de la figure 4.14

Comme dans la Section 4.3.2, on constate ici aussi que le volume d'apprentissage optimum est un volume intermédiaire entre $5^\circ \times 5^\circ / 5$ -jours et $1^\circ \times 1^\circ / 1$ -jour : en présence de MADRAS/Megha-Tropiques, le score F diminue jusqu'à 0,3 sur terre et sur mer pour un volume de $3^\circ \times 3^\circ / 1$ -jour.

Ce minimum observé suggère que la relation IR/taux de pluie est à peu près stable dans une zone géographique plus grande que $1^\circ \times 1^\circ$. Les échelles caractéristiques des perturbations

tropicales en météorologie synoptique sont de l'ordre de 1000 km et de 1 jour (Beucher, 2010). On ne mélangerait donc pas de situations météorologiques différentes dans la phase d'apprentissage de TAPEER en étendant la zone géographique d'apprentissage à $3^\circ \times 3^\circ$.

On notera aussi que sur mer, les scores F sont très proches pour les simulations avec deux MADRAS/Megha-Tropiques en quadrature de phase; ils diffèrent en revanche plus sur terre (figure 4.15). Ceci montre une légère sensibilité des estimations TAPEER-GSMaP au phasage de Megha-Tropiques. Dans les deux cas, les configurations de constellation avec MT présentent de meilleurs scores que les estimations TAPEER-GSMaP réalisées avec 3 satellites en orbite héliosynchrone.

On peut donc conclure qu'un SO tel que MADRAS/Megha-Tropiques, est un élément contribuant fortement à l'augmentation du nombre d'observations utiles (N_{TOT} , $N_{0,5h}$ et N_{1h}) au sein d'une constellation, dans le contexte d'estimations de cumuls $1^\circ/1$ -jour et du calcul des barres d'erreurs associées. La méthode TAPEER peut pleinement bénéficier du mode d'échantillonnage d'un SO sur orbite tropicale, même pour les latitudes où deux observations consécutives sont séparées d'une durée de l'ordre du temps d'auto-corrélation caractéristique de la pluie (figure 4.13).

4.4 Simulation de l'apport de Megha-Tropiques : constellation réaliste

D'après les résultats préliminaires de validation du produit TAPEER-BRAIN (voir annexe B), sur les sites de Niamey et Ouémé en Afrique de l'Ouest, pour la saison 2006, les scores obtenus sont comparables ou meilleurs que ceux obtenus avec le produit TRMM-3B42 V6 (Huffman *et al.*, 2007). Ces estimations TAPEER-BRAIN sont réalisées avec les mesures issues d'une constellation de quatre SO : TMI/TRMM, AMSRE/AQUA et 2 SSMI/DMSF. Étant donné le phénomène de saturation mis en évidence à partir de trois à cinq SO, il est intéressant d'estimer l'impact de la présence de Megha-Tropiques en addition à cette constellation.

Deux simulations de constellation sont réalisées : la première constituée de 4 SO dont SO/00H, SO/03H, SO/06H et TMI/TRMM, la seconde constituée de 5 SO dont SO/00H, SO/03H, SO/06H, TMI/TRMM et MADRAS/Megha-Tropiques. Cette analyse est menée sur une région couvrant les 2 sites de validation de Niamey et Ouémé [12° W - 12° E; 2° N - 17° N].

La figure 4.16 montre les valeurs du score intégré F pour les 2 configurations. On peut voir que la présence de Megha-Tropiques améliore les estimations TAPEER-GSMaP, principalement pour les petits volumes d'apprentissage comme attendu. Par exemple, F diminue de 0,75 à 0,45 en ajoutant MADRAS/Megha-Tropiques dans la simulation pour le volume d'apprentissage de $1^\circ \times 1^\circ/1$ -jour.

Les 2 simulations réalisées sont des configurations de constellation avec 4 et 5 SO, le phéno-

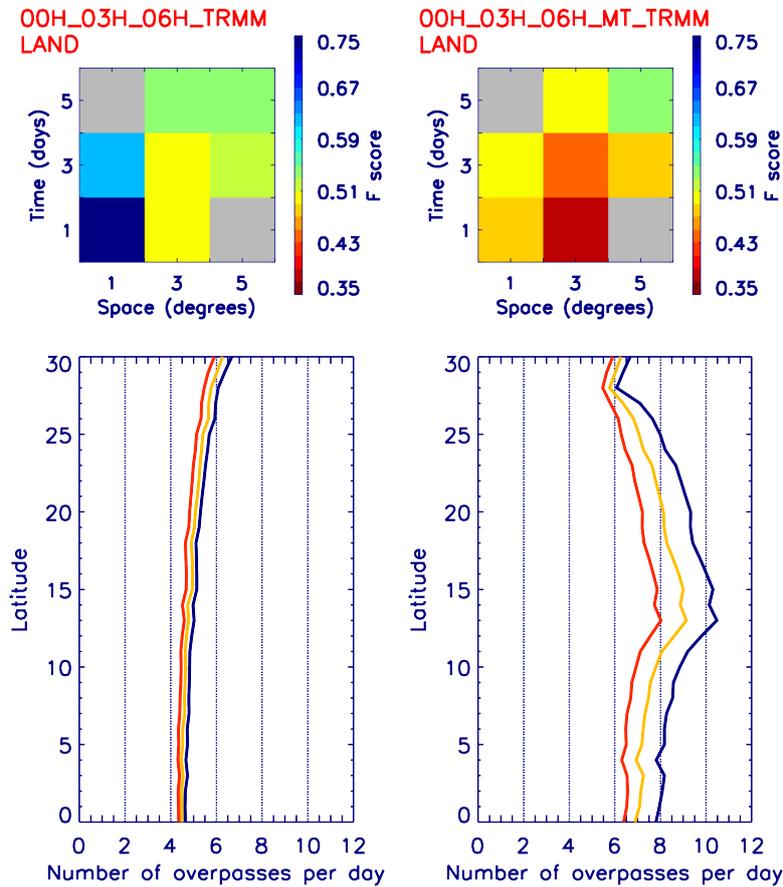


Figure 4.16 – Score intégré F calculé à partir des scores de comparaison des estimations TAPEER-GSMaP pour 2 configurations de constellation : 4 SO dont SO/00H, SO/03H, SO/06H et TMI/TRMM ; 5 SO dont SO/00H, SO/03H, SO/06H, TMI/TRMM et MADRAS/Megha-Tropiques. Les scores sont calculés pour la boîte "mousson Africaine" sur la terre. Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure. Le nombre moyen de passages par jour en fonction de la latitude du lieu observé pour chacune des 2 configurations est indiqué (3^{ième} ligne). N_{TOT} , $N_{0.5h}$ et N_{1h} sont tracés en noir, jaune et rouge.

mène de saturation discuté précédemment s’observe donc aussi dans cette expérience. En effet, le calcul de N_{TOT} , $N_{0,5h}$ et N_{1h} montre que pour la simulation avec 5 SO, deux observations par jour en moyenne ne sont pas utiles à l’estimation de pluies cumulées. Les améliorations d’une configuration à l’autre restent cependant quantifiables car, même en considérant un temps caractéristique d’auto-corrélation de la pluie de 1h, N_{1h} augmente de 4,5 à 6,5 au niveau de l’équateur et de 5 à 8 à 15°N.

Pour cette constellation de 5 satellites avec Megha-Tropiques, les estimations TAPEER-GSMaP présentent les meilleurs scores (F égal à 0,4) lorsque le volume d’apprentissage $3^\circ \times 3^\circ / 1$ -jour est sélectionné, comme dans la Section 4.3.3.

4.5 Conclusion

Les simulations réalisées dans ce chapitre ont permis dans un premier temps l’analyse des effets de trois caractéristiques d’une constellation de systèmes d’observation satellitaires pourvus de radiomètres à micro-ondes passives : l’heure locale de passage des satellites héliosynchrones, le nombre de satellites héliosynchrones dans une constellation, la présence d’un satellite placé sur une orbite tropicale dans une constellation.

Les résultats sont les suivants :

- la sélection du satellite en orbite héliosynchrone selon son heure locale de passage à l’équateur a un impact important sur l’estimation des précipitations cumulées TAPEER dans le cas où l’on utilise un seul SO. Cette sensibilité est a priori reliée au cycle diurne de la pluie.
- il est possible d’améliorer les estimations des précipitations cumulées TAPEER en utilisant un nombre important de SO. Ces améliorations sont alors quantifiables en diminuant la taille du volume d’apprentissage.
- un effet de saturation au nombre de SO a été identifié. Cette saturation peut être interprétée grâce au décompte d’un nombre moyen de passages par jour utile à l’estimation des pluies cumulées ($N_{0,5h}$, N_{1h}), prenant en compte les temps d’auto-corrélation caractéristiques de la pluie intégrée à l’échelle de la dizaine de kilomètres (taille d’un pixel micro-onde).

Par conséquent, on préconise l’utilisation d’un volume d’apprentissage de taille inférieure à $5^\circ \times 5^\circ / 5$ -jours en présence de Megha-Tropiques. L’analyse ne prend pour l’instant pas en compte l’inhomogénéité de résolution spatiale des différents radiomètres existants, or cette inhomogénéité induit une variabilité dans les produits de Niveau-2, L’impact de cette variabilité sur les estimations des pluies cumulées est minimisé seulement pour un large volume d’apprentissage pour lequel beaucoup d’échantillons de taux de pluie sont moyennés. Pendant la campagne de calibration/validation de Megha-Tropiques, des données pluviométriques seront recueillies à partir de réseaux de pluviomètres en Afrique, en Inde et au Brésil. Avec les données réelles de MADRAS, il sera possible d’optimiser le volume d’apprentissage de la méthode TAPEER et confirmer les

résultats présentés.

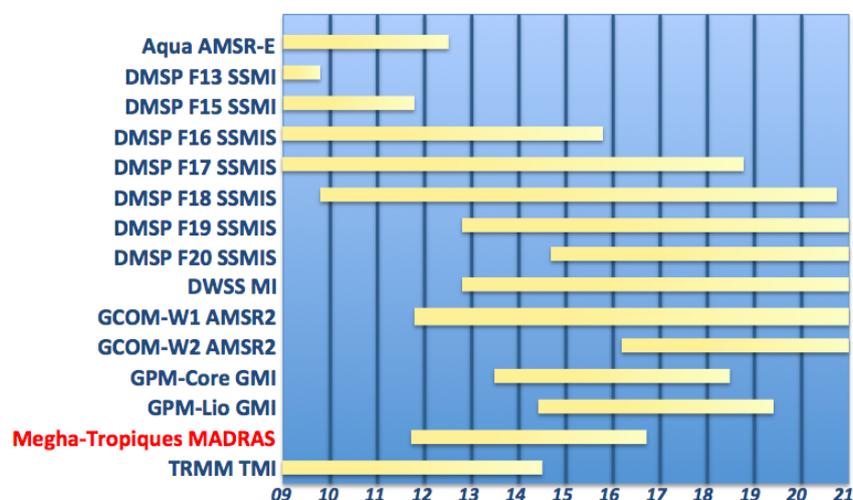


Figure 4.17 – Durées de vie des satellites portant des radiomètres à micro-ondes passives à scan conique (D’après Hou *et al.*, 2011; Huffman *et al.*, 2010).

De plus, l’effet de saturation identifié montre qu’il est possible de sélectionner les systèmes d’observation de la constellation à utiliser pour le produit TAPEER-BRAIN. On peut voir sur la figure 4.17 la durée de vie des satellites de la constellation GPM pourvus de radiomètres à micro-ondes passives depuis 2009 et projetées jusqu’à 2021. En particulier, d’après ces prévisions, le nombre de radiomètres qui opéreront au même moment s’élève à dix au maximum (en 2015 : DMSP F16, F17, F18, F19, F20, DWSS, GCOM-W1, GPM-Core, Megha-Tropiques et TRMM ; en 2016 : DMSP F17, F18, F19, F20, DWSS, GCOM-W1, GCOM-W2, GPM-Core, GPM-Lio, Megha-Tropiques). Un élément à prendre en compte dans la sélection parmi ces systèmes d’observation pourra être la résolution spatiale des autres radiomètres que MADRAS : en choisissant ceux dont les estimations BRAIN auraient une échelle d’intégration spatiale proche de celle du produit de Niveau-2 BRAIN-MADRAS, la variabilité liée à l’utilisation d’instruments multiples sera limitée. Cette optimisation sera menée dans le cadre de la phase d’exploitation de la mission Megha-Tropiques.

Conclusion, discussion et perspectives

5.1 Résumé des travaux présentés

L'intensité, l'occurrence ainsi que la répartition géographique des pluies sont trois caractéristiques des précipitations qui ont un impact important sur les populations dans les Tropiques. Le réchauffement global est susceptible de modifier ces trois éléments, dont les évolutions ont des conséquences sur le cycle de l'eau et de l'énergie à l'échelle de la planète toute entière. Dans ce contexte, il est donc important de développer des outils permettant d'estimer quantitativement les précipitations, à la fois sur les surfaces continentales et sur les surfaces océaniques.

L'estimation quantitative des précipitations (QPE) nécessite une densité d'observations élevée, accessible uniquement à partir d'observations spatiales pour l'ensemble des Tropiques. En particulier, l'évolution des moyens satellitaires à disposition conduit à une systématisation des mesures dans le domaine des micro-ondes.

Dans le cadre de la mission Megha-Tropiques et de la future constellation GPM (Global Precipitation Measurement), les travaux réalisés dans cette thèse ont permis de définir de nouveaux outils pour l'estimation quantitative des précipitations par satellite, sur l'intégralité de la ceinture tropicale.

Depuis plus de trente ans, différentes méthodes de QPE satellite ont été développées, principalement à partir de mesures dans les domaines des micro-ondes et de l'infrarouge. Grâce aux efforts menés, les techniques algorithmiques ont évolué du GOES Precipitation Index vers des méthodes multispectrales (IR/MO), assimilant les mesures de constellation de satellites afin d'estimer des cumuls de pluie à des échelles de plus en plus fines. La caractérisation des erreurs associées aux cumuls s'appuie principalement sur des comparaisons, ou validations, à partir de QPE issues de données de surface (pluviomètres ou radar) ; ces méthodes permettent d'évaluer les biais et incertitudes d'un ensemble de QPE. Des techniques de modélisation paramétrique et non paramétrique des erreurs ont également été développées afin d'adjoindre une erreur à une QPE individuelle. Ces techniques ont cependant rarement été associées à une méthode d'estimations des précipitations afin d'évaluer cumuls et erreurs de façon conjointe.

Il existe un large spectre de produits satellite, de qualité inégale. Dans le chapitre 2, une analyse des produits issus des développements les plus récents montre que la qualité des QPE satellite a atteint un degré suffisant pour que ces produits soient utilisés de manière quantitative

à des échelles de temps pertinentes en météorologie :

- à l'échelle des pluies cumulées sur 10 jours, tous les produits évalués reproduisent très bien la variabilité des pluies mesurées grâce au réseau de pluviomètres sur la bande Sahélienne, y compris la variabilité pre-onset/post-onset caractéristique de la mousson en Afrique de l'Ouest.
- à l'échelle des pluies cumulées sur 1 jour, certains produits reproduisent convenablement la variabilité des pluies mesurées par les réseaux de pluviomètres à Niamey (Niger), Ouémé (Bénin) et à Dakar (Sénégal).
- à l'échelle de variabilité des ondes d'Est africaines, tous les produits testés reproduisent très bien la variabilité des pluies mesurées par ces trois réseaux sols.
- à l'échelle du cycle diurne, certains produits reproduisent assez bien la variabilité des pluies observées au sol, y compris la variabilité pre-onset/post-onset de la mousson en Afrique de l'Ouest.

Il est également apparu qu'à ces échelles de temps, il est nécessaire de manipuler les estimations de cumul des précipitations avec leurs barres d'erreurs.

Les comparaisons des QPE satellite avec des QPE issues de mesures de surface (pluviomètres ou radars), considérées comme référence, est une méthode classique employée pour valider les produits satellite. Les résultats présentés dans cette thèse montrent que ces méthodes se révèlent être limitées lorsque les incertitudes de chaque QPE sont importantes. Dans ce cas, une technique mathématiques dédiée est nécessaire afin d'effectuer les comparaisons.

Pour cela, une nouvelle méthode de validation a alors été développée au cours du chapitre 2 ; celle-ci nécessite la connaissance des incertitudes de chaque QPE pour effectuer une comparaison. Une série de scores de validation est alors proposée, prenant tous en compte les incertitudes de chaque QPE comparée.

Les incertitudes calculées pour un produit satellite sont basées sur un nouveau modèle permettant d'évaluer la composante liée à l'échantillonnage du bilan d'erreurs. Un cumul de pluie estimé par satellite est constitué d'échantillons ; le modèle d'erreurs introduit dans cette thèse nécessite le calcul de la variance de ces échantillons ainsi que celui de N_{ind} , le nombre d'échantillons réellement indépendants qui constitue le cumul de pluie . Une méthode fondée sur le calcul de fonctions variogrammes permet de caractériser N_{ind} .

Il apparaît que les erreurs d'échantillonnage des produits satellites sont du même ordre de grandeur que les erreurs d'échantillonnage associées à la spatialisation par block-krigeage des données ponctuelles des réseaux de pluviomètres de Niamey et Ouémé.

La méthode TAPEER (Tropical Amount of Precipitation with an Estimate of ERrors), introduite dans cette thèse, sert à la compréhension de la mesure des pluies grâce à la mise en oeuvre d'une approche théorique. Cette approche vise à étudier le bilan des erreurs associées aux QPE satellite. Ce bilan comporte trois composantes : une composante d'échantillonnage liée à

la nature discrète dans l'espace et dans le temps des mesures satellitales, une composante d'algorithme liée à l'inversion des mesures de températures de brillance MO en taux de pluie ainsi qu'une composante liée à la calibration des divers systèmes d'observation utilisés.

Le modèle d'erreur introduit dans le chapitre 2 permet d'évaluer la contribution à l'erreur totale liée à l'échantillonnage. La propagation des erreurs de calibration des instruments et d'algorithme, de l'échelle instantanée vers l'échelle des pluies cumulées, permet d'évaluer les deux autres contributions du bilan d'erreurs. L'analyse de ce bilan a permis d'arriver aux conclusions suivantes :

- les erreurs de calibration IR ont très peu d'impact sur les QPE à la résolution de $1^\circ/1$ -jour : cette source d'erreurs peut être négligée.
- les erreurs sur les taux de pluie micro-ondes peuvent fortement impacter l'estimation de cumuls de pluie.
- les erreurs systématiques sur les taux de pluie d'intensité moyenne sont les plus critiques pour l'estimation des pluies cumulées.
- il est possible de caractériser les erreurs d'estimation à l'échelle des pluies instantanées nécessaire pour obtenir une erreur donnée à l'échelle des pluies cumulées.

D'après les validations du produit de Niveau-2 BRAIN-TMI effectuées en Afrique de l'Ouest, la composante algorithmique du bilan d'erreurs est négligeable par rapport à la composante d'échantillonnage. Pendant la phase d'exploitation de la mission Megha-Tropiques, les estimations BRAIN seront confrontées à des QPE de référence (pluviomètres et radar) en Afrique, en Inde et au Brésil. En particulier, les produits de Niveau-2 BRAIN/MADRAS, BRAIN/TMI, BRAIN/SSMI seront comparés afin de réduire les hétérogénéités de résolution liées aux différences de surfaces d'intégration spatiale des pluies. La caractérisation systématique des biais et des incertitudes associées aux taux de pluie BRAIN permettra de réévaluer la contribution de l'erreur algorithmique à l'erreur totale évaluée par la méthode TAPEER.

La simulation de la constellation GPM qui a été réalisée, permet de définir les caractéristiques de la constellation de satellites défilants dont il est nécessaire de faire usage, en supplément de Megha-Tropiques, afin d'améliorer les QPE à la résolution $1^\circ/1$ -jour. La sensibilité de la méthode TAPEER à plusieurs degrés de liberté a alors été évaluée : la configuration de la constellation de satellites utilisée, la zone géographique ainsi que la durée de la période d'apprentissage de l'algorithme de combinaison IR/taux de pluie.

Cette étude a permis de démontrer l'apport de Megha-Tropiques à l'estimation des pluies cumulées. Son mode d'échantillonnage tropical amène une contribution équivalente à celle d'au moins deux satellites héliosynchrones. Les simulations réalisées ont aussi révélé une sensibilité à l'heure locale des satellites héliosynchrones sélectionnés. Les durées d'auto-corrélation caractéristiques de pluies intégrées sur des surfaces de ~ 100 km² sont de l'ordre de 30 minutes à 1 heure. Les simulations réalisées ont aussi mis en évidence que des observations micro-ondes cor-

réelées dans le temps sont inutiles aux QPE $1^\circ/1$ -jour. Cette analyse permet donc de construire la constellation de systèmes d'observation en adéquation avec l'objectif scientifique.

Grâce à l'investigation menée, il devient alors possible de formuler un nouveau produit d'estimation des précipitations, appelé TAPEER-BRAIN, permettant de fournir des QPE (cumuls et erreurs) à la résolution de $1^\circ/1$ -jour sur l'ensemble de la ceinture tropicale. Ce produit utilisera les estimations de taux de pluie BRAIN du système d'observation MADRAS/Megha-Tropiques et d'une constellation d'autres satellites. Le produit TAPEER-BRAIN pourra bénéficier des observations de cette constellation grâce à un volume d'apprentissage de $3^\circ \times 3^\circ/1$ -jour.

5.2 Discussion

Dans un contexte de réchauffement global, des études ont été menées sur l'évolution de l'intensité, de l'occurrence et de la répartition géographique des pluies grâce aux observations satellitaires. Ces études ne prennent pas en compte les erreurs associées à ces observations. Pourtant, une première étude réalisée sur l'été 2009 révèle l'importance de l'utilisation de la barre d'erreur dans l'analyse de la distribution des pluies, en particulier pour les plus forts cumuls.

La caractérisation des erreurs de chaque estimation de pluie $1^\circ \times 1^\circ/1$ -jour sur l'ensemble des Tropiques permet en effet de définir un intervalle de confiance pour chacun d'entre eux. Il est possible de construire un ensemble de réalisations du champ de pluie en perturbant chaque cumul dans son intervalle de confiance. L'analyse de l'intensité, de l'occurrence et de la répartition géographique des pluies peut alors être effectuée sur cet ensemble de réalisations et prendre ainsi en compte les erreurs d'estimation des précipitations.

Une technique de *bootstrap* simple est choisie pour afin de générer différentes réalisations des cumuls de pluie TAPEER-BRAIN. On fait l'hypothèse que l'intensité de chaque cumul présente une probabilité uniforme d'occurrence dans son intervalle de confiance. Trois ensembles de 500 réalisations de champs TAPEER-BRAIN sont calculés avec un degré de confiance de 95% sur la saison 2009 et toute la ceinture intertropicale. Le premier ensemble prend uniquement en compte les erreurs d'estimations liées à l'échantillonnage. Le deuxième ensemble prend en compte les erreurs d'échantillonnage et la composante d'erreurs algorithmiques dans le cas idéalisé, appelé "Best Case" au chapitre 3, pour lequel les taux de pluie BRAIN présentent des biais systématiques de 20% pour le régime d'intensité de pluie de 2 mm.h^{-1} à 10 mm.h^{-1} . Pour le troisième ensemble de réalisations, appelée "Worst Case" au chapitre 3, un biais systématique de 60% sur les taux de pluie moyen est sélectionné.

La figure 5.1 montre la répartition géographique des 1000 cumuls de pluie $1^\circ \times 1^\circ/1$ -jour les plus forts sur la saison 2009, exprimée en densité de probabilité. La première carte correspond à une analyse de la distribution des pluies TAPEER-BRAIN sans prendre en compte les erreurs

des estimations. Les trois cartes suivantes correspondent à une analyse effectuée à partir des trois ensembles de 500 réalisations.

On constate que les quatre cartes présentent de nombreuses zones communes de forte densité de probabilité d'occurrence en cumuls de pluie intenses, principalement dans la baie du Bengale et dans l'océan Pacifique Ouest où de nombreux typhons ont contribué à d'importants cumuls de pluie journaliers pendant la saison 2009. La majeure partie des zones géographiques présentant une fraction des 1000 cumuls de pluie les plus forts sont des zones océaniques ou des zones côtières. Deux zones continentales au sud du Brésil et en Afrique de l'Ouest présentent également une forte probabilité d'occurrence de cumuls de pluie intenses : ces régions sont en effet identifiées dans la littérature comme des zones géographiques caractérisées par des événements orageux intenses (Zipser, 2006).

À l'échelle régionale, l'analyse prenant en compte les erreurs d'estimations permet d'une part de confirmer la présence de certains maxima et de révéler d'autre part des disparités régionales non représentées dans l'analyse sans erreurs.

La figure 5.2 montre que le nord de la baie du Bengale présente la densité de probabilité d'occurrence de cumuls de pluie intenses la plus forte de la région. La densité observée pour cette zone géographique est statistiquement robuste aux trois analyses avec erreurs. En revanche, la prise en compte des erreurs montre que l'Inde présente une densité de probabilité d'occurrence de cumuls de pluie intenses non nulle ; la zone géographique de forte probabilité au niveau de la côte ouest de l'Inde s'étend également vers le sud. La densité de probabilité au niveau de la zone à l'Est de l'Indonésie est renforcée dans l'analyse avec erreur. Dans le cas de l'analyse prenant en compte les erreurs algorithmiques, on observe une probabilité équivalente sur cette zone à celle de la côte ouest de l'Inde.

La figure 5.3 montre que la zone de l'océan Pacifique [120°W - 100°W ; 0°N - 20°N] à l'ouest du Mexique présente également une zone de forte probabilité d'occurrence de cumuls de pluie intenses. La densité de probabilité observée pour cette zone géographique est statistiquement robuste aux trois analyses avec erreurs. En revanche, la prise en compte des erreurs révèle une seconde zone de forte probabilité sur la côte colombienne. La figure 5.4 montre en effet que dans le cas d'une analyse sans erreurs, les intensités des cumuls de pluie dans la zone de l'océan Pacifique [120°W - 100°W ; 0°N - 20°N] sont plus importantes que celles des cumuls sur la côte colombienne. Les deux séries représentées sur cette figure présentent pourtant un score de comparaison FEBO de 84% (Frequency of Error Bars Overlap, voir Chapitre 2 pour la définition). Ainsi la prise en compte des incertitudes de chacun de ces cumuls révèle que ces deux zones géographiques sont caractérisées par une densité de probabilité d'occurrence de cumuls de pluie intenses équivalente.

Dans le cas des réalisations avec une erreur totale (échantillonnage et algorithmique), la côte colombienne présente l'une des probabilités d'occurrence de cumuls de pluie intenses les plus importantes de tout l'océan Pacifique Est.

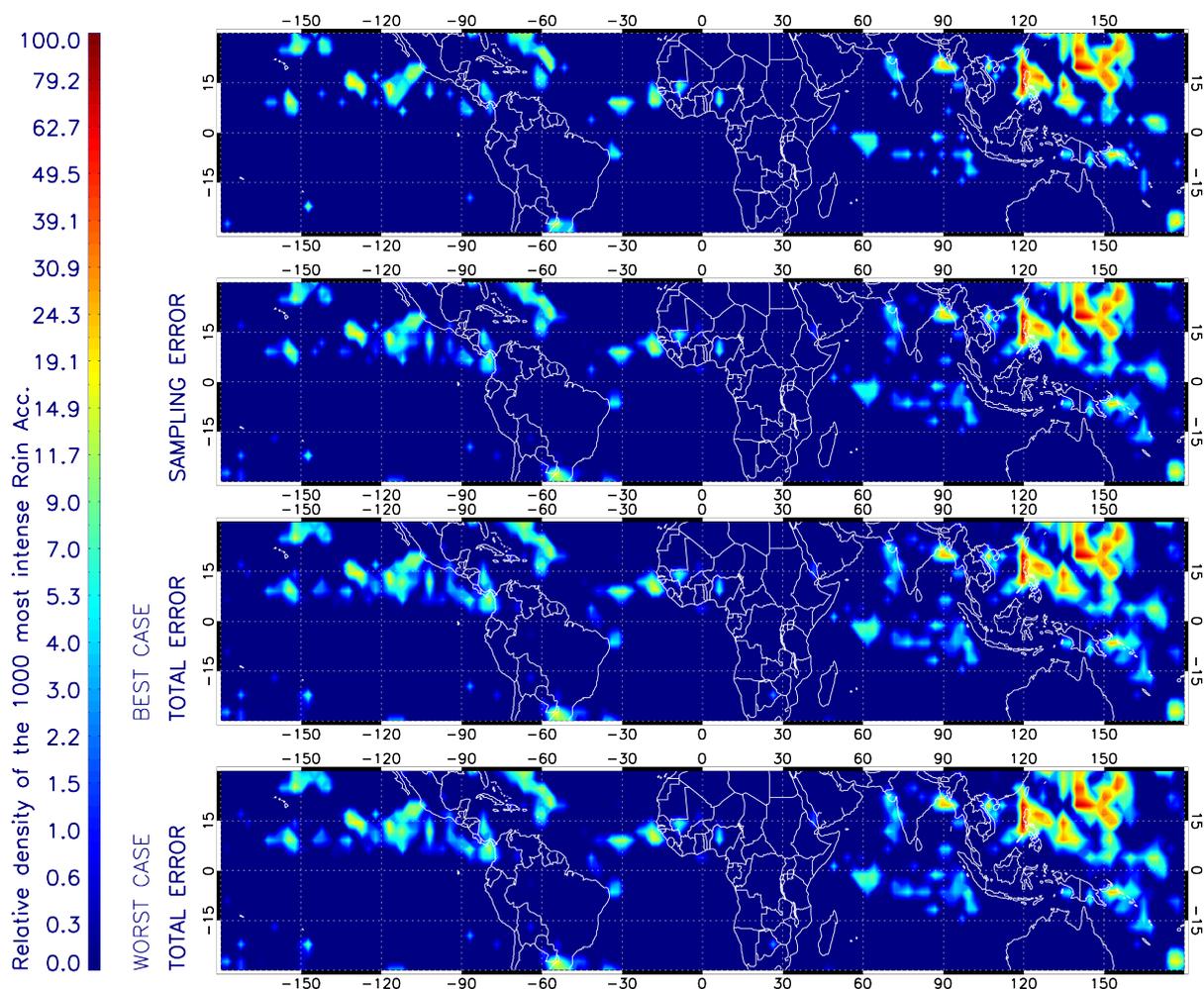


Figure 5.1 – Densités géographiques des 1000 cumuls de pluie $1^\circ \times 1^\circ / 1$ -jour les plus forts sur la saison 2009, normalisées par rapport au maximum d’occurrence de ces 1000 cumuls sur l’ensemble des Tropiques sur une grille de $2.5^\circ \times 2.5^\circ$. La première carte correspond à une densité calculée uniquement à partir des cumuls de pluie TAPEER-BRAIN. La deuxième carte correspond à une densité calculée à partir de 500 réalisations de champs TAPEER-BRAIN prenant en compte les erreurs d’échantillonnage avec un degré de confiance de 95%. La troisième carte (resp. quatrième carte) correspond à une densité calculée à partir de 500 réalisations de champs TAPEER-BRAIN prenant en compte les erreurs d’échantillonnage et les erreurs algorithmiques issues d’un biais systématique de BRAIN de 20% (resp. 60%) sur les taux de pluie 2 mm.h^{-1} et 10 mm.h^{-1} .

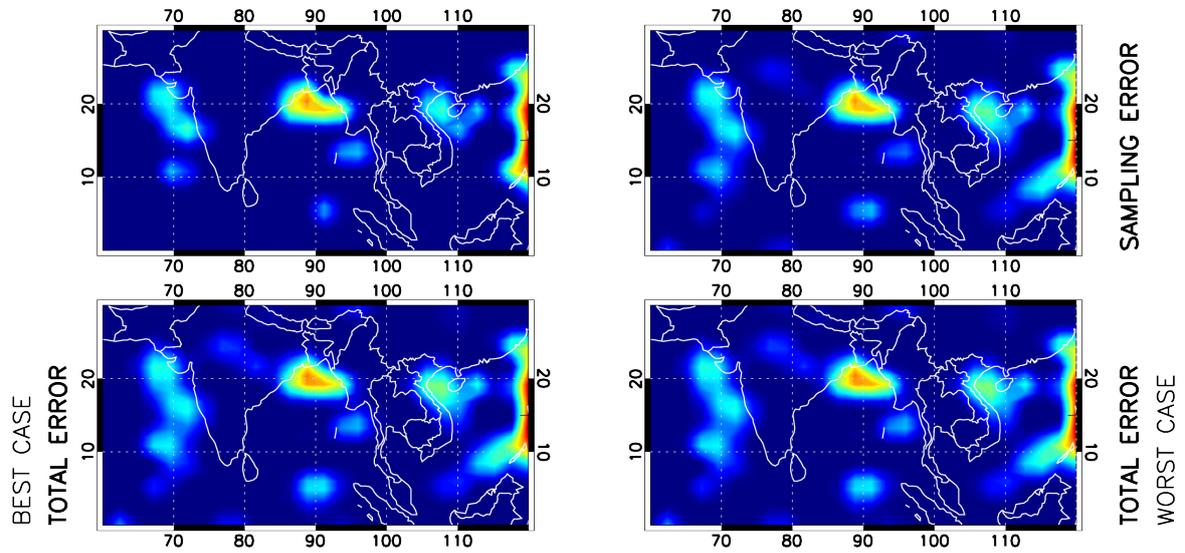


Figure 5.2 – Extraction des quatre cartes de la figure 5.1 sur la zone géographique $[60^{\circ}\text{E} - 120^{\circ}\text{E}; 0^{\circ}\text{N} - 30^{\circ}\text{N}]$

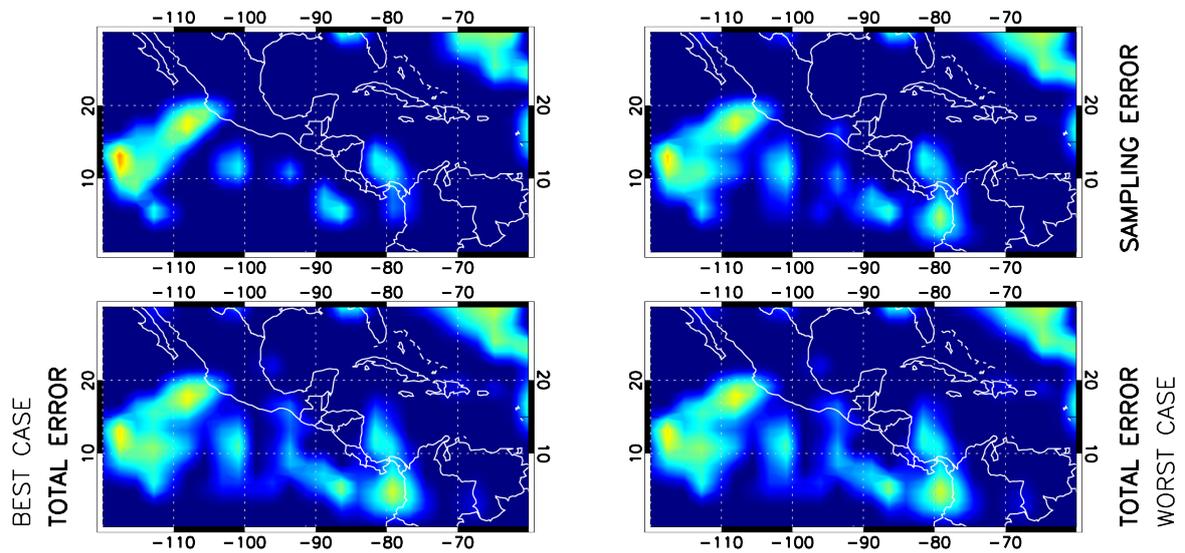


Figure 5.3 – Extraction des quatre cartes de la figure 5.1 sur la zone géographique $[120^{\circ}\text{W} - 60^{\circ}\text{W}; 0^{\circ}\text{N} - 30^{\circ}\text{N}]$

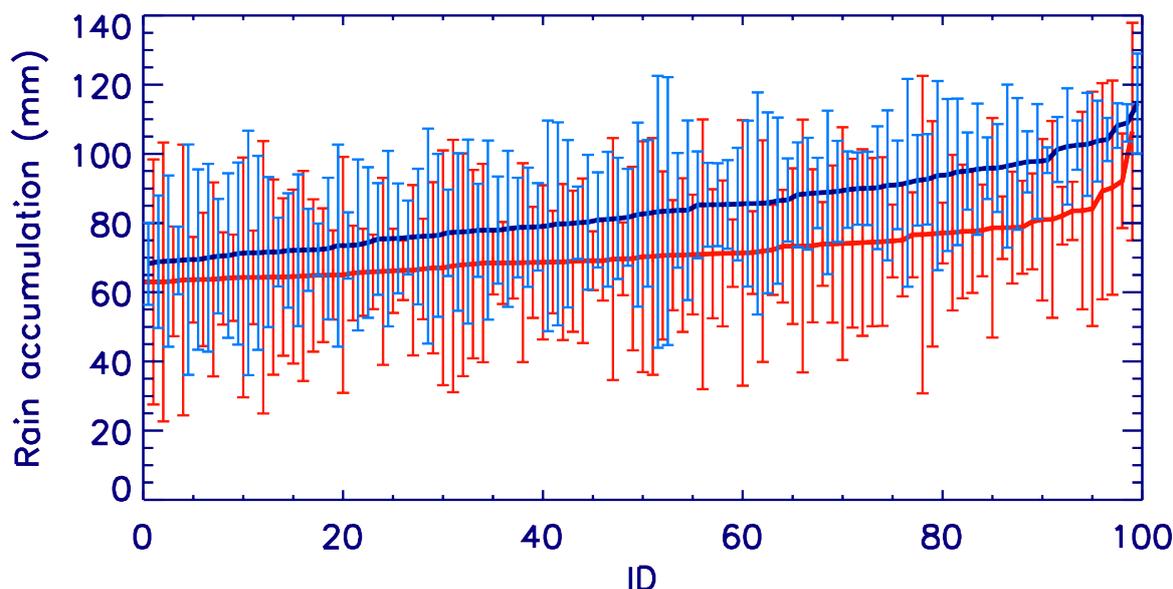


Figure 5.4 – Intensités et erreurs associées (degré de confiance de 95%) des 100 cumuls de pluie les plus forts pour deux zones géographiques : $[120^{\circ}\text{W} - 110^{\circ}\text{W} ; 10^{\circ}\text{N} - 20^{\circ}\text{N}]$ en bleu et $[85^{\circ}\text{W} - 75^{\circ}\text{W} ; 00^{\circ}\text{N} - 10^{\circ}\text{N}]$ en rouge

Dans la littérature, plusieurs études utilisent les observations micro-ondes issues de satellites défilants telles que TRMM afin de caractériser la répartition géographique des orages les plus intenses sur la planète (Mohr *et al.*, 1999; Zipser, 2006; Liu, 2011). Dans ces analyses, les températures de brillance acquises à diverses fréquences servent directement d'indicateurs pour l'intensité des orages. L'utilisation des températures de brillances est préférée à l'utilisation d'estimations de taux de pluie afin que les résultats ne soient pas contaminés par les erreurs liées aux inversions.

Au regard de cette étude préliminaire, la caractérisation systématique des erreurs d'estimations associées aux cumuls de pluie proposée par la méthode TAPEER apparaît comme une solution alternative pour l'étude de l'intensité, de l'occurrence et de la répartition géographique des pluies sur l'ensemble des Tropiques.

5.3 Perspectives

Une première perspective amenée par ce travail concerne le raffinement de notre étude de propagation des erreurs. En effet, dans le chapitre 3, on a fait l'hypothèse que les distributions d'erreurs algorithmiques sont indépendantes de l'intensité du cumul. Cette hypothèse pourra être remise en question si, pendant la campagne de validation de la mission Megha-Tropiques en Afrique, en Inde et au Brésil, les biais et incertitudes associés aux estimations de taux de

pluie BRAIN se révèlent importants. D'autre part, les erreurs de détection des cumuls de pluie ne sont pour l'instant pas prises en compte dans le bilan d'erreurs. Néanmoins, la méthode de propagation d'erreurs mise en place dans cette thèse s'avère très générale. En effet, la détection de cumuls est directement associée à la détection de la pluie à l'échelle instantanée par les produits de Niveau-2. Une approche possible serait la propagation d'estimations BRAIN avec des seuils de détection de pluie différents. Cette approche permettrait d'évaluer la sensibilité de la détection de cumuls par rapport à la détection de la pluie à l'échelle instantanée. On pourrait alors définir une erreur de détection à associer aux estimations des précipitations TAPEER-BRAIN. L'amplitude des erreurs algorithmiques sur les cumuls de pluie dépend de l'amplitude des erreurs sur les taux de pluie instantanés ; les caractéristiques des erreurs de détection sur les cumuls de pluie dépendraient également des erreurs de détection des estimations BRAIN.

Une deuxième perspective est liée aux simulations réalisées au cours du chapitre 3. L'hypothèse simplificatrice, concernant la résolution des produits de Niveau-2, a conduit à utiliser des Niveau-2 simulés dont chaque fauchée comporte le même nombre de pixels. Pourtant, les produits BRAIN-SSMI et BRAIN-AMSRE sont respectivement composés de 64 et 243 pixels au sein d'une fauchée. Ainsi, AMSRE/AQUA apporte le nombre d'échantillons de pluie instantanée le plus important pour les estimations de pluie cumulée TAPEER-BRAIN produites au cours de cette thèse. Au moment de la rédaction de ce manuscrit, AMSRE vient de subir une défaillance technique. Le satellite Megha-Tropiques a été lancée le 12 Octobre 2011 avec succès depuis la base de Sriharikota au Nord de Chennai, sur la baie du Bengale. Pendant la phase d'exploitation de la mission Megha-Tropiques, il sera donc nécessaire d'évaluer le nombre de systèmes d'observation requis en supplément de MADRAS/Megha-Tropiques, en prenant en compte la résolution réelle des produits de Niveau-2.

Une troisième perspective concerne l'utilisation de TAPEER-BRAIN pour les applications en hydrométéorologie. Comme discuté dans la section précédente, la prise en compte des barres d'erreurs est importante pour l'étude de la distribution des pluies. L'analyse des phénomènes de mousson pourrait donc bénéficier de l'approche intégrée aux QPE introduite dans cette thèse. Les modèles de prévision des crues nécessitent également des QPE dont les erreurs sont caractérisées. L'étude du bilan d'erreurs réalisée pourrait contribuer à l'amélioration des prévisions hydrologiques dans les Tropiques, sur des bassins versants peu instrumentés. Selon le modèle hydrologique utilisé, il pourrait être nécessaire de fournir des estimations TAPEER-BRAIN à une meilleure résolution, de $0.25^\circ/6$ heures par exemple. Dans ce but, il serait alors nécessaire de réévaluer la contribution des erreurs d'estimations de BRAIN au bilan d'erreurs à l'échelle des pluies cumulées. Certaines contributions telles que les erreurs aléatoires de BRAIN ne seraient plus nécessairement négligeables. D'autre part, la durée de la période d'apprentissage est supérieure à un jour. Résoudre le cycle diurne de la pluie, sur des périodes d'accumulation de

6 heures, nécessiterait un volume d'apprentissage adapté, éventuellement par tranche d'heures locales.

Une autre perspective amenée par ce travail serait la génération d'une série longue d'estimations TAPEER-BRAIN pour l'étude des cumuls de pluie intenses sur la ceinture intertropicale. Il serait alors nécessaire d'étendre la production de TAPEER à une période antérieure à la mission Megha-Tropiques et à la constellation GPM. Dans le chapitre 4, les simulations de produits de Niveau-2 ont montré une sensibilité des estimations TAPEER-BRAIN à la sélection des heures locales des satellites sur orbite héliosynchrone utilisés, en particulier lorsque la méthode TAPEER est alimentée par peu d'observations micro-ondes. La construction de cette série longue nécessiterait donc un soin particulier dans la sélection des systèmes d'observations utilisés, ainsi que la caractérisation de l'évolution des biais au sein de cette série. Une archive d'estimations des précipitations TAPEER-BRAIN pourrait débuter en 1997, année de lancement de la mission TRMM, afin d'utiliser au minimum deux systèmes d'observations micro-ondes. Il faut néanmoins attendre juillet 1998 pour que l'océan Indien soit observé depuis une orbite géostationnaire par Meteosat-5. La période de janvier 1999 à nos jours apparaît donc être une possibilité pour constituer une série longue TAPEER-BRAIN afin d'analyser les tendances récentes d'évolution du cycle de l'eau dans les Tropiques.

Radiométrie micro-ondes pour la télédétection spatiale des précipitations

A.1 Rappels théoriques sur le transfert radiatif dans le domaine des micro-ondes

A.1.1 Interaction entre le rayonnement et l'atmosphère

Pour la mesure de la pluie ou du contenu en eau liquide des nuages, on s'intéresse au domaine des micro-ondes (ou hyperfréquences) qui correspond aux fréquences comprises entre 1 et 300 GHz, c'est-à-dire aux longueurs d'onde entre 30 cm et 1 mm.

L'atmosphère gazeuse est responsable de l'absorption résonante à certaines fréquences : la vapeur d'eau absorbe à 22 GHz et à 183 GHz et l'oxygène absorbe entre 50 et 70 GHz et à 118 GHz comme on peut le voir sur la figure A.1 qui représente la variabilité du coefficient de transmission à travers l'atmosphère en ciel clair dans ce domaine de fréquences. Le choix des fréquences utilisées en radiométrie passive pour la détection des précipitations, se fait dans les "fenêtres atmosphériques". Ainsi, les canaux des radiomètres micro-ondes (MO), tel MADRAS qui sont dédiés à la mesure de la pluie, sont aux fréquences 18,7 GHz, 36,5 GHz et 89 GHz et présentent une transmission élevée en ciel clair. À ces canaux s'ajoute un canal à 23,8 GHz pour la mesure de la vapeur d'eau atmosphérique.

L'interaction du rayonnement avec les nuages et les précipitations (pluie et glace), se traduit par deux phénomènes : l'absorption (non résonante) et la diffusion par les hydrométéores (gouttelettes, gouttes d'eau et cristaux de glace) qui sont décrites par la théorie de Mie que l'on peut résumer par les considérations suivantes. Les termes de l'absorption et de la diffusion s'expriment en fonction du paramètre χ qui est le rapport de la taille D de la goutte à la longueur d'onde λ du rayonnement : $\chi=D/\lambda$. Le terme d'absorption varie en χ^3 , tandis que le terme de diffusion varie en χ^6 et ces termes dépendent de la distribution des tailles des gouttes. Le diamètre des gouttelettes dans les nuages est de l'ordre de 1 à 100 μm , tandis que celui des gouttes d'eau ou cristaux de glace précipitants est de l'ordre de 100 μm à quelques mm, certains cristaux pouvant atteindre plusieurs centimètres.

La diffusion est donc négligeable devant l'absorption pour les gouttelettes des nuages ($\chi \ll 1$);

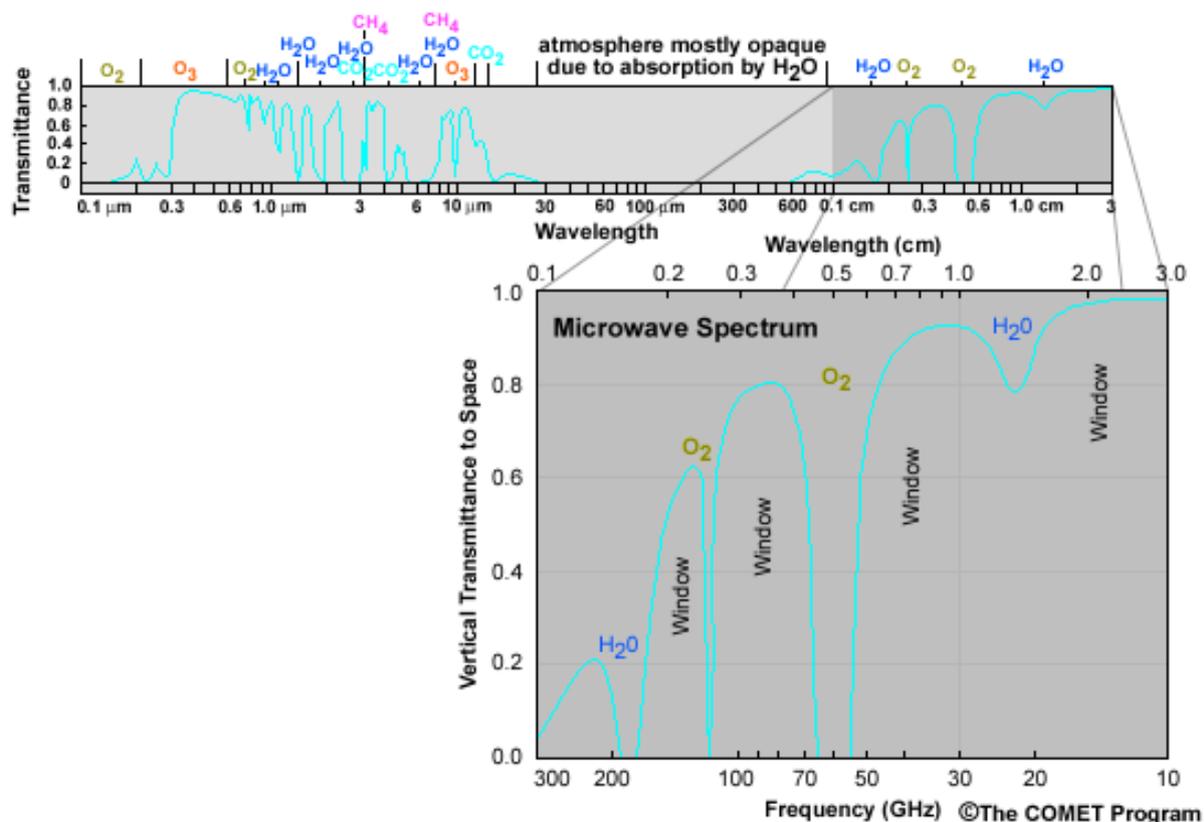


Figure A.1 – Coefficient de transmission à travers l’atmosphère terrestre en ciel clair en fonction de la fréquence du rayonnement. Source : COMET <http://med-ted.ucar.edu/>, University Corporation for Atmospheric Research

par contre, la diffusion est plus importante pour les gouttes les plus grosses qui précipitent, avec $\chi \geq 1$. Par ailleurs, les termes de l’absorption et de la diffusion dépendent de la constante diélectrique, donc de la phase de l’eau (liquide ou solide). Compte tenu des valeurs de la constante diélectrique de l’eau et de la glace, les coefficients d’absorption et les coefficients de diffusion varient avec le taux de pluie de façon différente pour l’eau et pour la glace, comme cela est représenté sur la figure A.2 pour les trois fréquences 19, 37 et 86 GHz. On voit sur la figure A.2a que l’absorption par les particules de glace est tout à fait négligeable devant l’absorption par les gouttes d’eau quelle que soit la fréquence ; ainsi, les cirrus (nuages constitués de petits cristaux de glace) sont transparents aux MO. Sur la figure A.2b, on voit que l’importance relative de la diffusion par l’eau ou par la glace dépend de la fréquence ; quand la fréquence est élevée, c’est la diffusion par la glace qui est prépondérante. Par exemple, dans les systèmes convectifs tropicaux, les fortes pluies sont produites par des cumulo-nimbus développés jusqu’à des altitudes élevées, contenant, sur une épaisseur de plusieurs kilomètres, de gros cristaux de glace qui sont responsables d’une forte diffusion du rayonnement MO aux fréquences élevées ; le signal MO

mesuré à 86 GHz est alors très marqué et est relié aux pluies intenses associées à la présence de ces cristaux de glace. En résumé, en présence des précipitations, l'absorption par les gouttes d'eau est prépondérante aux basses fréquences (19 et 36,5 GHz), tandis que c'est la diffusion par la glace qui est le phénomène le plus important aux fréquences élevées (86 GHz). Les relations entre la pluie et les radiances MO mesurées à bord de satellite sont présentées dans la section suivante.

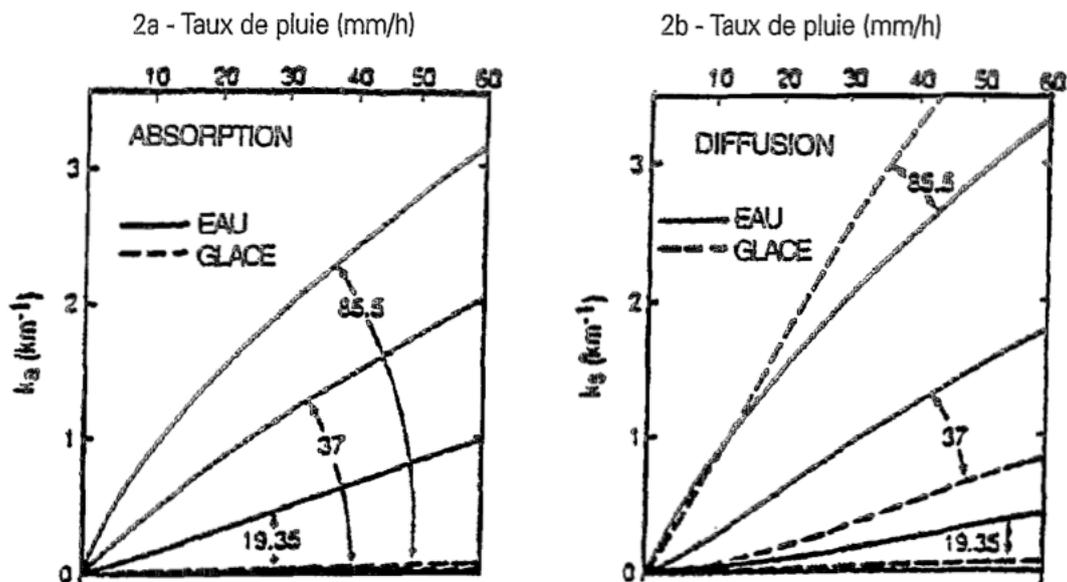


Figure A.2 – Coefficient d'absorption k_a (2a) et coefficient de diffusion k_s (2b) en km⁻¹, en fonction du taux de pluie, pour l'eau liquide (en trait plein) et pour la glace (en traits pointillés), aux trois fréquences 18,7 GHz, 37 GHz et 86 GHz. (Jobard, 1996)

A.1.2 Relation entre les données satellite Micro-ondes et la pluie

L'intensité du rayonnement MO, reçue par un radiomètre placé à bord d'un satellite, est constituée de l'intensité de l'ensemble des rayonnements MO émis, transmis et diffusés par la surface et par l'atmosphère nuageuse. La radiance ou l'intensité d'un rayonnement émis par un corps noir, à la température physique T et à la fréquence ν , est donnée par la fonction de Planck :

$$B(\nu, T) = \frac{2h\nu^5}{c^3} \cdot \left[\frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1} \right] \quad (W.m^{-2}.Sr^{-1}.\mu m^{-1}) \quad (A.1)$$

où h est la constante de Planck, k la constante de Boltzman et c la vitesse de la lumière. Dans le domaine des fréquences MO (1 à 300 GHz) et pour les températures de l'atmosphère (10

à 300 K), le produit $h\nu$ est toujours très inférieur à kT ; en utilisant un développement limité du premier ordre, la radiance s'écrit :

$$B(\nu, T) = \frac{2\nu^4}{c^3} \cdot kT \quad (W.m^{-2}.Sr^{-1}.\mu m^{-1}) \quad (A.2)$$

Ceci constitue l'approximation connue sous le nom de Rayleigh-Jeans. Elle permet de considérer que la radiance est proportionnelle à la température T du corps noir. La surface de la Terre n'est pas un corps noir, elle a une émissivité ε qui dépend de la fréquence d'observation ; un radiomètre micro-ondes mesure alors une température de brillance $T_B = \varepsilon.T$ qui est la température équivalente d'un corps noir qui émettrait la même énergie que la surface.

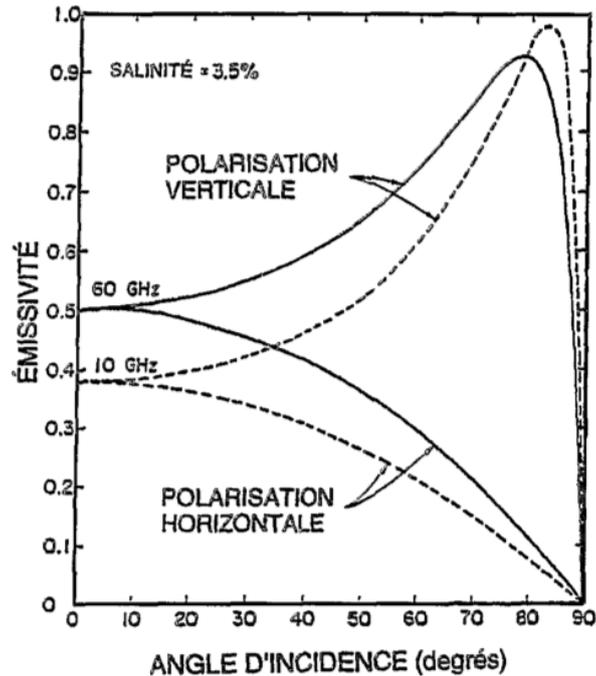


Figure A.3 – Émissivité de l'océan pour les polarisations horizontale et verticale en fonction de l'angle d'incidence à 10 GHz et à 60 GHz. (Jobard, 1996)

En ciel clair, le radiomètre voit la surface de la Terre, à travers l'atmosphère sans nuage qui est transparente aux MO, et mesure une température de brillance qui dépend de la température T de la surface et de son émissivité ε . L'émissivité dépend du type de surface, mer ou continent :

- pour la mer, l'émissivité varie de 0,3 à 0,7 environ, en fonction de la fréquence, de la polarisation et de l'angle de visée. Ceci est illustré par la figure A.3 qui représente l'émissivité pour les polarisations horizontale et verticale en fonction de l'angle d'incidence à 10 GHz et à 60 GHz. Les canaux de MADRAS dédiés à la mesure de la pluie, sont polarisés horizontalement et verticalement aux fréquences 18,7, 36,5 et 89 GHz et le radiomètre vise la

surface avec un angle de 53° . L'émissivité de la mer dépend également de l'état de surface (rugosité liée au vent, salinité et température) ;

- pour les continents, l'émissivité est forte ; elle varie de 0,9 à 1 principalement en fonction de l'humidité du sol.

Ainsi, le radiomètre, lorsqu'il vise la mer, enregistre des températures de brillance froides (typiquement : $T_B = \varepsilon.T = 0,5 \times 300 = 150$ K), tandis que lorsqu'il vise la terre, il enregistre des températures de brillance élevées (typiquement : $T_B = \varepsilon.T = 0,9 \times 300 = 270$ K).

Lorsqu'il y a des nuages et des précipitations, comme il a été dit dans la section précédente, il y a absorption et diffusion par les hydrométéores. Ainsi, le radiomètre ne reçoit plus seulement le signal venant de la surface mais celui provenant du nuage ou des précipitations.

Quand il y a absorption par les gouttes d'eau, celles-ci émettent à la température de la couche de l'atmosphère où elles se trouvent et le radiomètre mesure alors une température de brillance qui est plus élevée que la température de brillance de la surface en arrière-plan. Ainsi, l'augmentation de la température de brillance observée est un signal de l'absorption par les gouttes d'eau qui est une mesure du volume d'eau ; le signal est donc directement relié à la pluie. Toutefois, ce signal n'est mesurable que lorsqu'il y a un contraste assez grand avec la température de brillance de la surface, c'est-à-dire seulement lorsque la surface visée est la mer.

Quand il y a diffusion par les hydrométéores, principalement les cristaux de glace, le rayonnement est diffusé dans toutes les directions. La couche supérieure de glace agit comme un écran qui masque les couches du niveau inférieur et elle n'émet vers le satellite qu'une partie du rayonnement ; les températures de brillance mesurées sont donc basses. Ainsi, la diminution de la température de brillance observée, aussi bien au-dessus de la mer que sur la terre, est un signal de la diffusion par les grosses particules de glace au sommet du nuage qui sont généralement associées au processus de la formation des gouttes de pluie. C'est donc une mesure moins directe de la pluie que le signal d'absorption, mais il peut être mesuré sur les continents aussi bien que sur la mer. Toutefois il est difficile de relier quantitativement le signal de diffusion au taux de pluie, car le coefficient de diffusion dépend d'un grand nombre de paramètres tels que la distribution des particules de glace, la forme des cristaux de glace et leur orientation, la densité de la glace et l'épaisseur de la couche de glace.

Comme l'absorption est prépondérante aux basses fréquences (18,7 et 36,5 GHz) et que la diffusion est le phénomène le plus important aux fréquences élevées (89 GHz), on peut simplifier en disant qu'on observe pour les canaux à 89 GHz, une température de brillance qui décroît quand le taux de pluie augmente à cause de la diffusion par les cristaux de glace, tandis que pour les canaux à 18,7 et 36,5 GHz, on observe une température de brillance qui augmente avec le taux de pluie à cause de l'absorption par les gouttes d'eau. À 18,7 GHz, la température de brillance atteint la saturation pour les forts taux de pluie, tandis qu'à 36,5 GHz, l'effet de la diffusion par les très grosses particules de glace associées aux taux de pluie élevés se fait néanmoins sentir et

on observe une diminution de la température de brillance pour les forts taux de pluie.

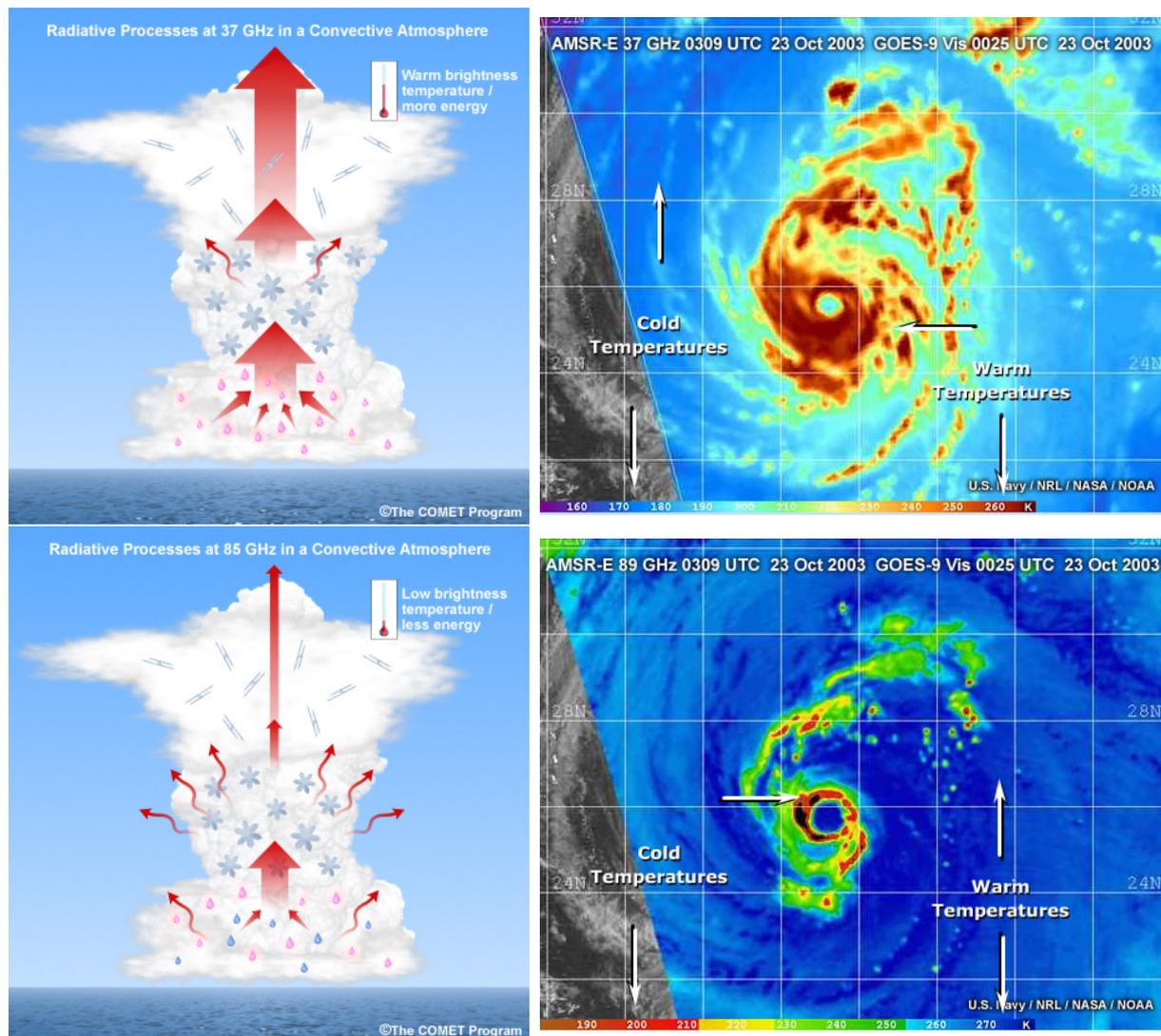


Figure A.4 – Processus d’absorption/émission à 37 GHz -V (en haut) et de diffusion à 85 GHz -V (en bas), illustrés par un scan d’AMSRE du Typhon Parma en mer des Philippines en Octobre 2003, superposé avec une image dans le domaine Visible acquise par le satellite GOES-9. Pour le canal 37 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus chaudes. Pour le canal 89 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus froides. Source : COMET <http://meted.ucar.edu/>

Pour les surfaces océaniques : la figure A.4 montre un exemple des T_B obtenues pour un typhon en mer des Philippines. À 37 GHz, la surface de la mer a une T_B basse à 170 K et l’émission des gouttes d’eau au sein du typhon entraîne des T_B plus chaudes entre 240 et 270 K ; à 89 GHz, la surface de la mer a une T_B élevée à 270 K et la diffusion par les cristaux de glace

au sein du typhon entraîne des T_B plus froides entre 190 et 210 K.

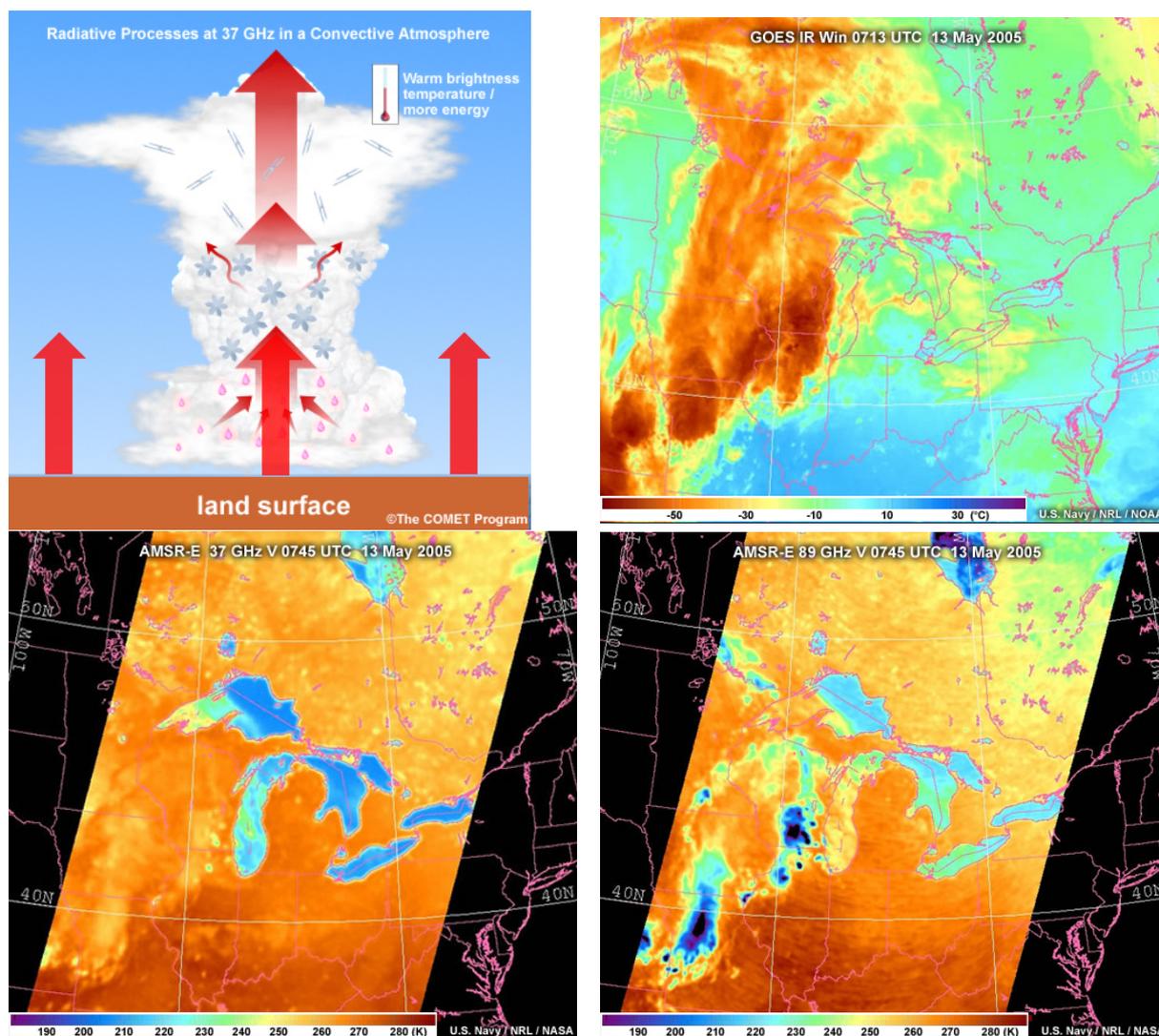


Figure A.5 – Processus d’absorption/émission à 37 GHz -V (en bas à gauche) et de diffusion à 85 GHz -V (en bas à droite), illustrés par un scan d’AMSRE d’un front dans le Midwest américain en Mai 2005. Pour les 2 canaux 37 GHz -V et 85 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus chaudes. L’image dans le domaine IR acquise par le satellite GOES montre la position géographique du front (en haut à droite). Source : COMET <http://meted.ucar.edu/>

Pour les surfaces continentales : la figure A.5 montre un exemple des T_B obtenues pour un front dans le Midwest américain. L’image du canal IR thermique d’un des satellites géostationnaires GOES montre la localisation du front. À 37 GHz, la surface continentale et la zone a priori pluvieuse présentent une T_B chaude au dessus de 250 K. Ceci illustre le fait qu’à basse fréquence

sur continent, l'émissivité du sol empêche d'avoir un contraste suffisant entre zone pluvieuse et ciel clair pour que le signal soit exploitable. À 87 GHz, la surface continentale a une T_B élevée à 270 K et la diffusion par les cristaux de glace au sein du front entraîne des T_B plus froides entre 180 et 210 K.

A.2 Exemple de l'algorithme BRAIN pour l'estimation des pluies instantannées

L'algorithme BRAIN (Viltard *et al.*, 2006, 2011) est un algorithme d'estimation des précipitations instantannées utilisant la radiométrie micro-ondes. Cet algorithme, à l'origine dérivé de l'algorithme GProf (Kummerow *et al.*, 2001), s'appuie sur une base de données observées conjointement par le radiomètre TMI et le radar PR du satellite TRMM afin d'estimer des taux de pluie de surface dans la zone d'observation de radiomètres micro-ondes à bord de satellites défilants.

Les profils de réflectivité radar dérivés du système d'observation PR-TRMM permettent d'obtenir une estimation du profil de pluie et de la pluie de surface sous certaines hypothèses sur la distribution des gouttes au sein des scènes nuageuses observées (Drop Size Distribution ou DSD : tailles, concentrations et phases des hydrométéores). La réflectivité radar Z et le profil de pluie R sont tous les deux reliés à la DSD par les équations suivantes :

$$Z = \int_D N(D).D^6.dD \quad (\text{A.3})$$

$$R = \int_D N(D).D^3.w(D).dD \quad (\text{A.4})$$

où D est le diamètre des gouttes, $N(D)$ est le nombre de gouttes par unité de volume pour un diamètre donné, et $w(D)$ la vitesse de chute des gouttes pour un diamètre donné. Une estimation de R connaissant Z peut être faite grâce à un modèle paramétrique en loi de puissance, en faisant des hypothèses sur les DSD, distinctes dans les parties stratiformes et convectives des nuages. Les estimations de profil de pluie et taux de pluie de surface utilisées dans l'algorithme BRAIN sont fournies par le produit TRMM-2A25 (Iguchi *et al.*, 2000).

Les mesures des deux systèmes d'observation PR-TRMM et TMI-TRMM peuvent être localisées dans l'espace et dans le temps. Grâce à plus de dix années d'observations du satellite TRMM, il est possible de constituer une base de données, dite base de données d'inversion, dans laquelle chaque profil de pluie R et taux de pluie de surface $R_{surface}$ sont associés à un vecteur de 9 T_B (10,65 GHz - V et - H, 19,35 GHz - V et - H, 21,3 GHz - V, 37 GHz - V et H, 85,5 GHz - V et - H). La figure A.6 montre la distribution des différents taux de pluie $R_{surface}$ de cette base de données ; on peut voir notamment que les taux de pluie faibles sont les plus nombreux.

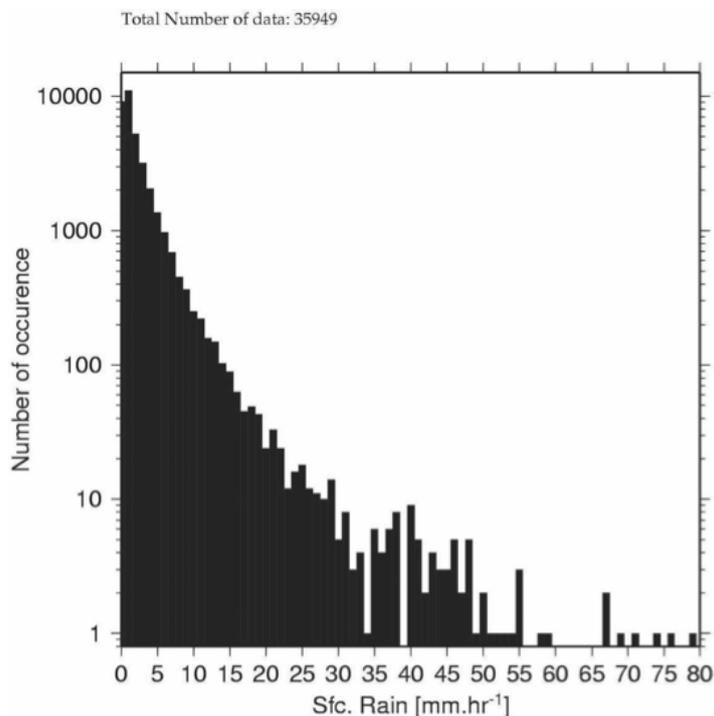


Figure A.6 – Distribution des taux de pluie dans la base de données d'inversions de BRAIN. (Viltard *et al.*, 2006)

L'algorithme BRAIN procède en deux temps pour estimer un taux de pluie de surface $R_{surface}$ à partir d'un vecteur de T_B observées, noté $[T_{B-OBS}]$. Un masque pluie/non-pluie est d'abord appliqué (Kacimi et al., 2011) puis, si $[T_{B-OBS}]$ est déterminé comme associé à un $R_{surface}$ non nul, BRAIN s'appuie alors sur une méthode d'inversion Bayésienne afin de déterminer R et $R_{surface}$. Au dessus d'une surface océanique, les T_B à toutes les fréquences sont utilisées ; au dessus d'une surface continentale, seules les T_B dans les hautes fréquences sont utilisables pour les raisons abordées précédemment.

Si le vecteur $[T_{B-OBS}]$ a été mesuré aux mêmes fréquences que celles du radiomètre TMI, il est possible de chercher les vecteurs les plus proches du $[T_{B-OBS}]$, noté $[T_{B-DATABASE}]$, dans la base de données d'inversion. Une fois ces vecteurs de $[T_{B-DATABASE}]$ déterminés, une estimation du taux de pluie de surface $R_{surface}$ à $[T_{B-OBS}]$ est la moyenne pondérée des taux de pluie de surface associés aux $[T_{B-DATABASE}]$ dans la base de données d'inversion (figure A.7).

Suivant le théorème de Bayes, si $Pr(R)$ désigne la probabilité d'observer un profil de pluie R , et $Pr([T_B]/R)$ la probabilité de mesurer un vecteur de températures de brillance $[T_B]$ pour ce profil, alors, la probabilité d'observer un profil de pluie R connaissant le vecteur de températures $[T_B]$, est donnée par :

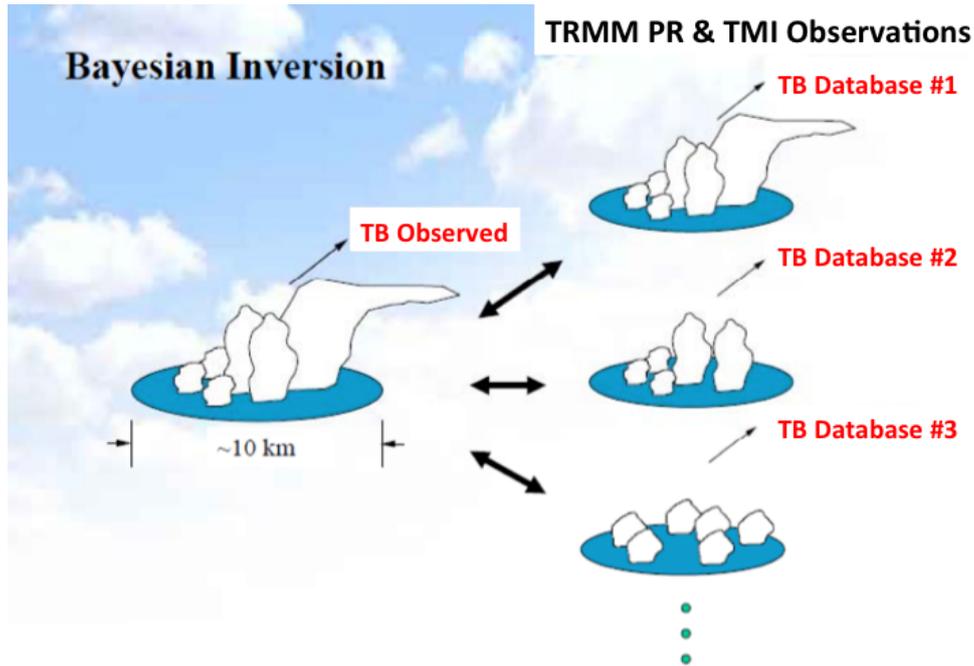


Figure A.7 – Base de données d’observations et estimations Bayésiennes. Source : IPWG, <http://www.isac.cnr.it/~ipwg/>

$$Pr(R/T_B) = Pr(R) \times Pr(T_B/R) \quad (\text{A.5})$$

En supposant que les éléments de la base de données d’inversion sont représentatifs de l’ensemble des cas pouvant être observés sur la région étudiée, la probabilité $Pr(R)$ est donnée par le nombre relatif d’occurrence d’un profil de pluie R dans cette base.

Dans le cas d’un radiomètre ne présentant pas exactement les mêmes fréquences que le radiomètre TMI, la procédure décrite ci-dessus ne peut être directement appliquée. Il est nécessaire de simuler des températures de brillance $[T_{B-SIMULATED}]$ aux fréquences de TMI à partir des températures de brillances observées $[T_{B-OBS}]$. L’algorithme BRAIN a été adapté au radiomètre MADRAS ainsi qu’à d’autres radiomètres de la future constellation GPM (Viltard *et al.*, 2011). Pour chaque radiomètre, une base de données d’inversion de profil de pluie R et $[T_{B-SIMULATED}]$ associées est générée grâce à un modèle de transfert radiatif. Afin de simuler un vecteur $[T_{B-SIMULATED}]$ à partir d’un des vecteurs $[T_{B-DATABASE}]$, le modèle de transfert radiatif nécessite l’utilisation d’un profil d’atmosphère. Un tel profil est sélectionné dans une base de données de simulations de nuages (Goddard Cumulus Ensemble, Tao *et al.*, 2003); le profil d’atmosphère choisi correspond à celui dont le profil de pluie est le plus proche de R .

Les estimations de BRAIN ont été validées sur océan et surface continentale. Un processus de validation de BRAIN est notamment possible en utilisant sa propre base de données d’inversion.

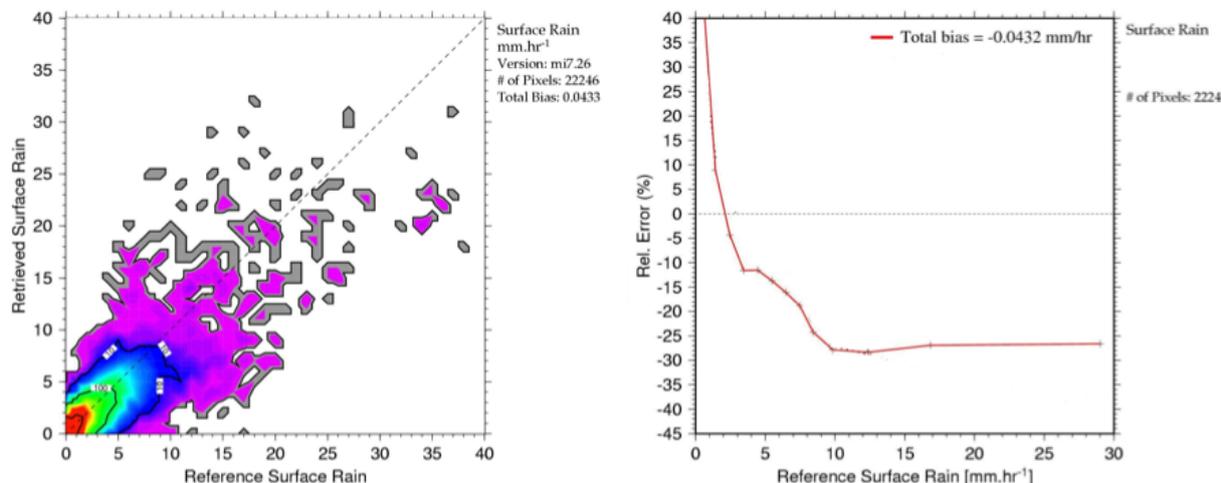


Figure A.8 – Comparaison des estimations de précipitations BRAIN (noté "Retrieved Surface Rain") avec les estimations de précipitations issues du PR (noté "Reference Surface Rain") sur l'océan (à gauche). Erreurs d'estimations relatives de BRAIN en fonction des taux de pluie. (Viltard *et al.*, 2006)

Si l'on scinde cette base en deux, la première demi base peut servir de $[T_{B-OBS}]$ et la seconde demi base sert alors à l'algorithme BRAIN. En comparant les taux de pluie estimés par BRAIN pour les $[T_{B-OBS}]$ avec les taux de pluie de la première demi base associées aux $[T_{B-OBS}]$, on obtient une caractérisation de l'erreur algorithmique de BRAIN. La figure A.8 montre le résultat de cette procédure, pour les estimations BRAIN sur l'océan. L'erreur algorithmique varie de +40% pour les taux de pluie faibles à -30% pour les taux de pluie les plus forts.

Pour les surfaces continentales, BRAIN a été comparé à des données de pluviomètres à haute résolution temporelle sur deux sites en Afrique de l'Ouest : pour les taux de pluie d'amplitudes de 2 mm.h^{-1} à 10 mm.h^{-1} , l'erreur obtenue est de +100% à -40% sur le bassin de l'Ouémé (Bénin) et de +175% à -30% à Niamey (Niger) (Kirstetter *et al.*, 2011).

A.3 Résolution spatiale des radiomètres micro-ondes

L'estimation de taux de pluie à partir de températures de brillance micro-ondes multi-canaux entraîne plusieurs incertitudes dont l'une est connue sous le nom de "beam filling". Le beam filling comporte deux aspects. Le premier est le remplissage partiel du champ de vue d'un pixel micro-ondes par la pluie. Le second est l'effet de non linéarité causé par la variabilité spatiale de la pluie au sein de la zone pluvieuse dans le champ de vue du pixel. La figure A.9 montre par exemple qu'une T_B mesurée de 180 K par le canal à 10GHz peut être associée à un taux de pluie moyen intégré allant de 8 mm.h^{-1} à 30 mm.h^{-1} .

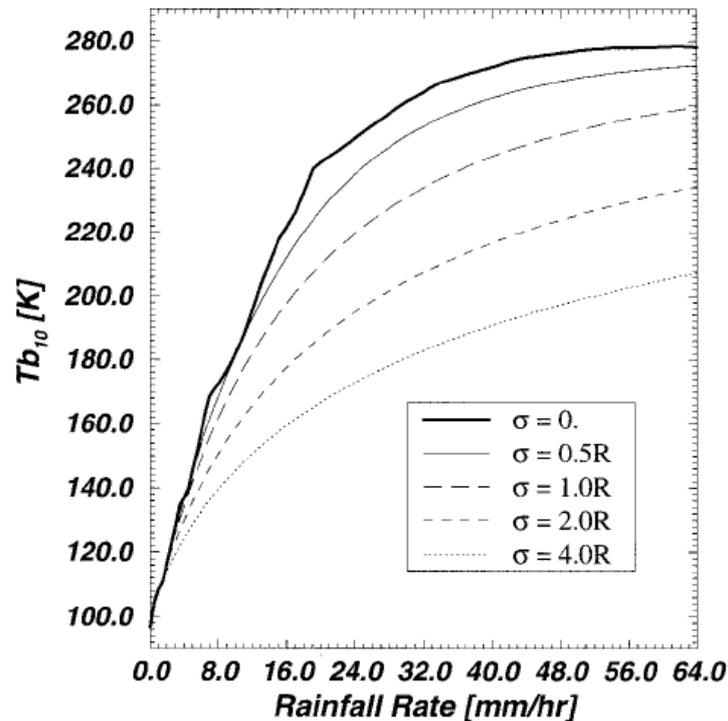


Figure A.9 – Relation T_B à 10 GHz, taux de pluie pour différents degrés d'inhomogénéités de champs de pluie à moyenne spatiale constante. (Kummerow, 1998)

L'utilisation de plusieurs canaux micro-ondes permet de lever une partie des indéterminations liées au beam filling dans la relation T_B /taux de pluie. En revanche, la taille des antennes utilisées pour un radiomètre à micro-ondes passives amène à une résolution spatiale différente pour chaque canal d'acquisition ; les tailles des pixels sont décroissantes des basses vers les hautes fréquences. Ainsi pour un pixel micro-ondes donné, l'estimation de taux de pluie multi-spectrale est issue d'intégrations spatiales de différents champs de vue. La résolution spatiale d'une estimation de taux de pluie dérivé de T_B micro-ondes n'est pas déterminée précisément ; cette résolution est intermédiaire à celle des canaux du système d'observation utilisé.

La figure A.10 montre les tailles relatives des pixels pour plusieurs radiomètres : on peut voir la taille décroissante des pixels des basses vers les hautes fréquences, tandis que pour un même canal les pixels sont de tailles variables d'un radiomètre à l'autre. En effet, chaque radiomètre répond à des spécifications techniques différentes : par exemple, dans les spécifications de MADRAS, les pixels du canal à 89 GHz sont jointifs d'une fauchée à l'autre. Les pixels des plus basses fréquences sont plus gros et se chevauchent. Le radiomètre GMI à bord du satellite GPM-Core présentera, lui, la résolution la plus fine des radiomètres à balayage conique.

Ainsi, selon l'instrument considéré, les estimations de pluie réalisées n'ont pas la même échelle d'intégration spatiale. Ces résolutions variées rendent l'utilisation des estimations BRAIN multi-

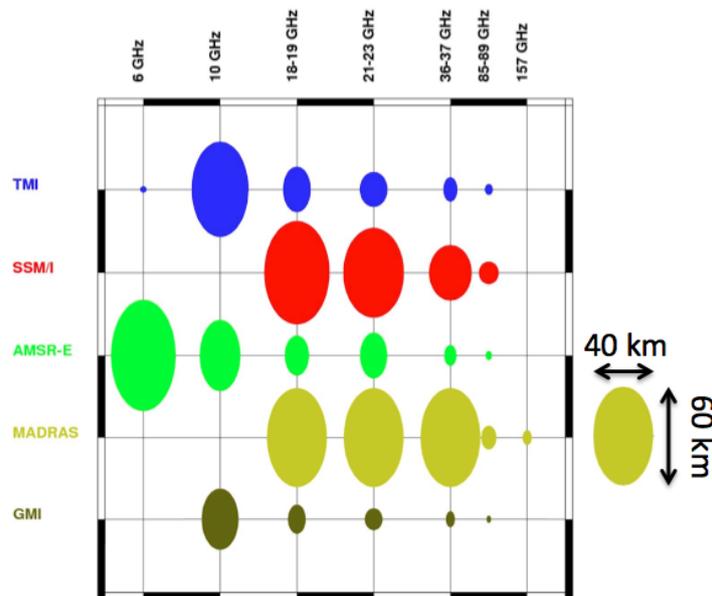


Figure A.10 – Représentation de la taille relative des pixels de différents radiomètres micro-ondes par canal. Source : N. Viltard

plateforme complexe au sein de TAPEER. La figure A.11 montre les restitutions de BRAIN-TMI, BRAIN-SSM/I et BRAIN-AMSR-E le long du cycle de vie du cyclone Gaëlle : entre le scan de TMI effectué à 17h55 UTC le 4 février et le scan de SSM/I effectué à 02h34 UTC le 5 février, on observe une diminution des taux de pluie estimés. L'observation suivante d'un scan de AMSRE à 10h06 UTC le 5 février montre des estimations de taux de pluie plus cohérentes avec celles issues du scan de TMI qu'avec les estimations issues du scan de SSM/I. Il est donc important de comprendre si ces différences sont dues à une diminution d'intensité pluviométrique au sein du cyclone lui-même ou à des questions de résolutions inter-radiomètres.

Dans l'algorithme BRAIN, les estimations de pluie du radar de précipitation sont colocalisées avec les mesures radiométriques de TMI afin de constituer la base de données d'inversion. Dans ce but, les estimations de pluie du radar de précipitation sont moyennées à la résolution du canal à 37 GHz du radiomètre ; on peut donc faire l'hypothèse que les pluies estimées par BRAIN correspondent à des pluies intégrées spatialement à l'échelle du canal 37 GHz de chaque radiomètre. Cette hypothèse est appliquée dans la méthode TAPEER pour la colocalisation des données IR et BRAIN : les T_B IR dont les pixels sont situés dans une zone de "rayon équivalent" au canal 37 GHz de chaque radiomètre sont moyennées. Le raffinement qui consisterait à prendre en compte la forme ovale des champs de vue effectifs des pixels micro-ondes a peu d'impact sur la méthode TAPEER étant donné la résolution variable des pixels des différents canaux utilisés dans l'algorithme BRAIN.

La colocalisation d'un ensemble d'estimations de taux de pluie BRAIN et de T_B IR amène à

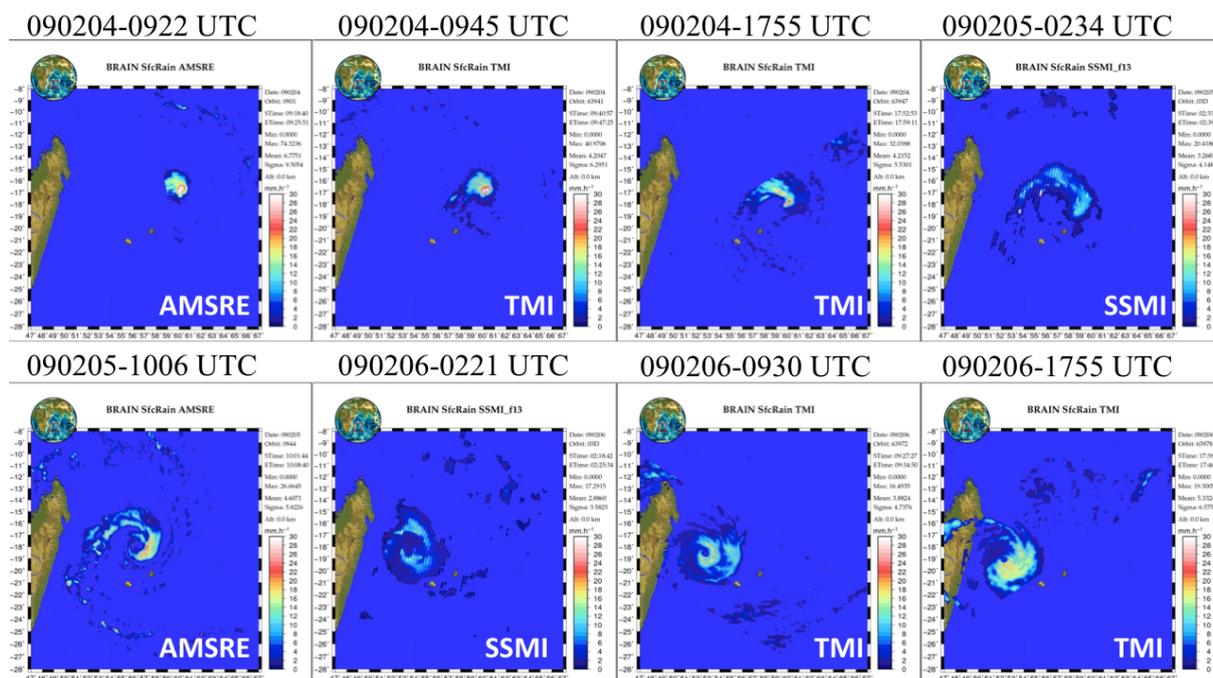


Figure A.11 – Exemples de restitutions de taux de pluie par BRAIN grâce aux instruments TMI, SSM/I et AMSRE pendant le cycle de vie du cyclone Gaëlle à son passage sur l'île de la Réunion du 4 au 6 février 2009. Source : N. Viltard

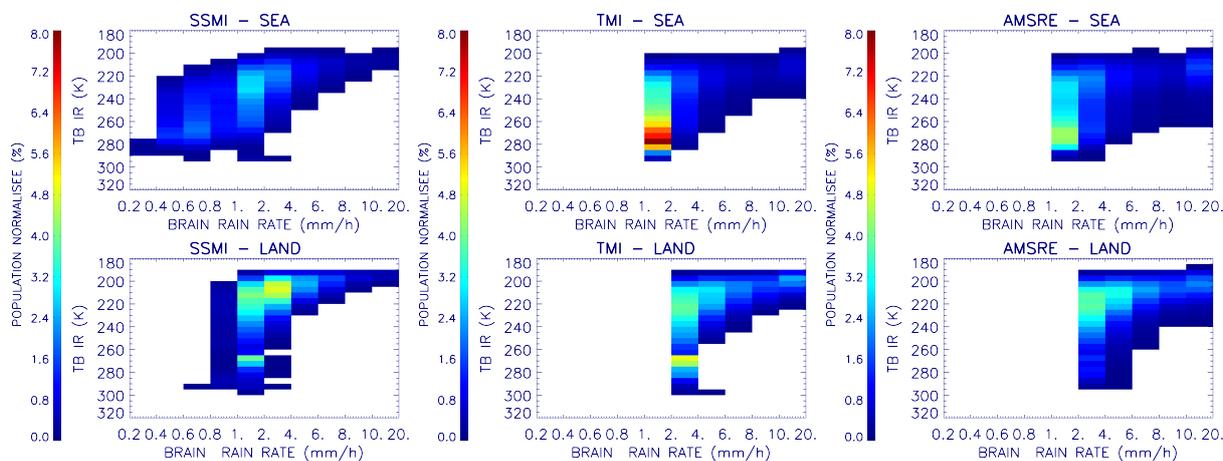


Figure A.12 – Histogrammes 2D des estimations de pluie BRAIN pour les radiomètres SSMI (colonne de gauche) TMI (colonne du milieu) et AMSRE (colonne de droite) géo localisées avec les données IR du satellite MSG-2 pour la période juin à septembre 2009, pour la zone $[40^{\circ} \text{W} - 40^{\circ} \text{E}; 30^{\circ} \text{S} - 30^{\circ} \text{N}]$, sur continent et sur océan

pouvoir comparer les estimations BRAIN issues de TMI, AMSRE et SSMI en fonction des T_B IR. La figure A.12 montre que les estimations BRAIN issues de TMI et AMSRE exhibent deux distributions équivalentes, sur la terre et sur l'océan : sur la terre on observe un maximum de population entre 2 et 4 mm.h⁻¹ pour des températures IR entre 210 et 230 K, et sur l'océan un maximum de population entre 1 et 2 mm.h⁻¹ pour des températures IR entre 260 et 280 K. Dans le cas de SSM/I, on peut voir que l'histogramme est différent avec un seuil de détection beaucoup plus bas, ce qui est cohérent avec l'intégration spatiale sur des surfaces plus grandes. Le maximum de population sur l'océan entre 1 et 2 mm.h⁻¹ correspond à des températures entre 220 et 240 K, températures beaucoup plus froides que pour BRAIN-TMI et BRAIN-AMSRE.

L'harmonisation des estimations de pluie micro-ondes multi systèmes d'observation, en termes de détection pluie/non-pluie et d'intensité de pluie, est une question qu'il est difficile de traiter même si des techniques ont été proposées (exemple dans TMPA, Huffman *et al.*, 2007). L'impact de la multi-résolution des radiomètres sur les estimations de cumuls est par extension une question difficile à résoudre. L'un des objectifs de la mission GPM est d'unifier, grâce au radar du GPM-Core, les observations de la constellation de radiomètres micro-ondes en vol, au niveau des températures de brillance micro-ondes et des restitutions de pluie. Les développements qui découleront de cette mission permettront de mieux appréhender les différences entre radiomètres dans les méthodes telles que TAPEER.

The TAPEER-BRAIN product

TAPEER-BRAIN is a combined Microwave-Infrared accumulated precipitation estimation product implemented as a Megha-Tropiques Level 4 product. TAPEER stands for **T**ropical **A**mount of **P**recipitation with an **E**stimate of **ER**rors. It provides precipitation estimations and associated errors at the one-degree/one-day accumulated scale. TAPEER-BRAIN is based on the BRAIN algorithm (Viltard *et al.*, 2006, MT Level 2 MW-derived instantaneous rain estimation product) and the TAPEER algorithm. Section B.1 describes the TAPEER algorithm, Section B.2 presents its implementation for the TAPEER-BRAIN Level 4 product and Section B.3 shows preliminary validation results over West Africa at several meteorologically relevant scales.

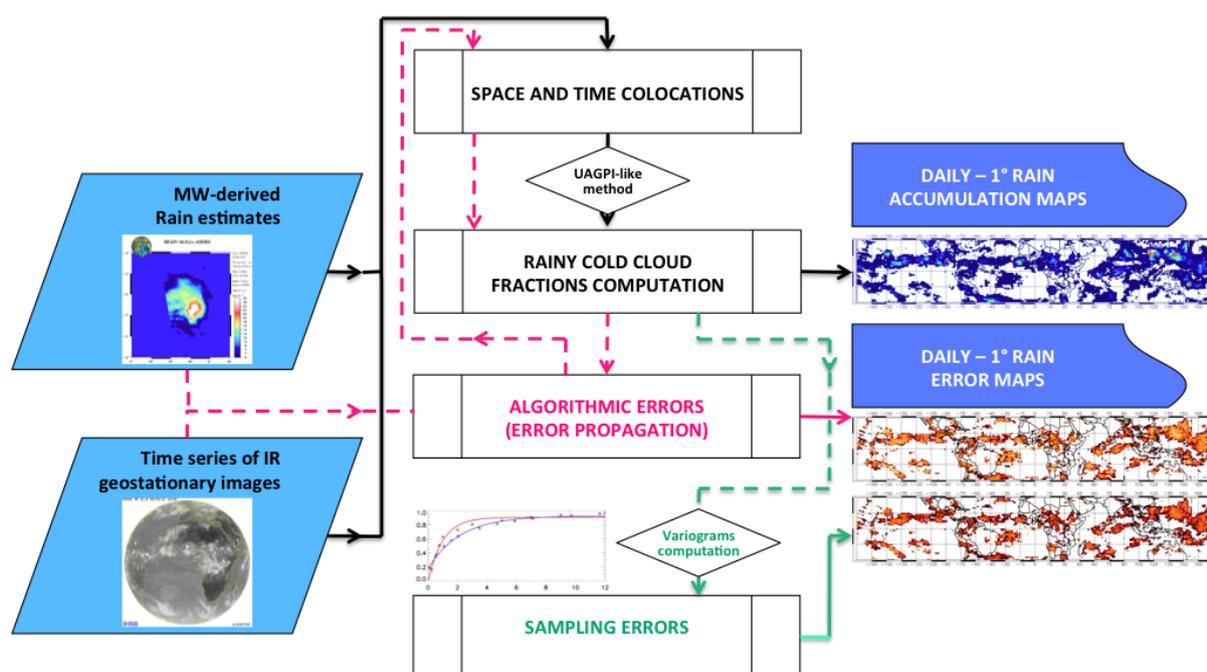


Figure B.1 – General diagram

B.1 TAPEER : Algorithm Theoretical Basis Document

The TAPEER algorithm relies on two separate algorithms to provide both the Tropical Amount of Precipitation and the Estimate of ERrors. Figure B.1 summarizes the functioning of the algorithm with black arrows standing for the Tropical Amount of Precipitation computations, green arrows and pink arrows standing for the Estimate of ERrors computations.

B.1.1 Tropical Amount of Precipitation

The Tropical Amount of Precipitation is estimated with a combination of MW-derived rain estimations and time series of geostationary Thermal Infrared data (TIR). Figure B.2 illustrates the need of MW and TIR data combination to estimate rainfall at the accumulated scale : a rainy event occurred on the 8th of September 2006 over Niamey and led to one of the most important daily rain accumulation (18.5 mm \pm 2.1 mm) in the rain gauge records for the 2006 monsoon season. This rainy event has not been observed by any MW imager of a constellation of four low earth-orbiting satellites, but was monitored by the multispectral imager SEVIRI onboard the geostationary satellite Meteosat Second Generation (top line). IR monitoring provides information only statistically related to rainfall accumulated on long periods and averaged over large areas. Thus MW-derived estimations are jointly used with IR data to provide an Amount of Precipitation. The method of combination of TIR data and MW-derived rain rates selected for TAPEER relies on the Universally Adjusted GOES Precipitation Index (UAGPI) technique (Xu *et al.*, 1999) which aims to estimate rain accumulation through cold cloud fraction calculations.

The UAGPI relates accumulated rainfall, in a given area A over a period T , $\langle R \rangle_{A,T}$ to the rainy cloud fraction F_C in this space-time-volume, working on the assumption that for large enough areas and accumulation times, rainfall rates in this volume tend towards a local mean value R_{COND} that may be determined through the use of available local microwave satellite precipitation data. For instance, Figure B.2 (bottom line) illustrates the available local microwave satellite precipitation data for the one-degree area of Niamey on the 8th of September 2006, in a volume of $5^\circ \times 5^\circ \times 5$ -day surrounding the $1^\circ/1$ -day volume. With an estimate of the rainy cloud fraction F_C and the local mean rain rate value R_{COND} , the following equation is used to compute the Tropical Amount of Precipitation :

$$\langle R \rangle_{A,T} = F_C \times R_{COND} \times T \quad (\text{B.1})$$

The rainy cloud fraction F_C , for a period T (24 h) and a given area A ($1^\circ \times 1^\circ$), is estimated with TIR data under the following assumption : TIR samples colder (resp. warmer) than a given threshold $T_{threshold}$ are assumed to be rainy (resp. non-rainy). From the GATE experiment in Tropical Atlantic, it was shown that a 235 K threshold was a good compromise when estimating monthly and $2.5^\circ \times 2.5^\circ$ rainy cloud fractions (Arkin, 1979).

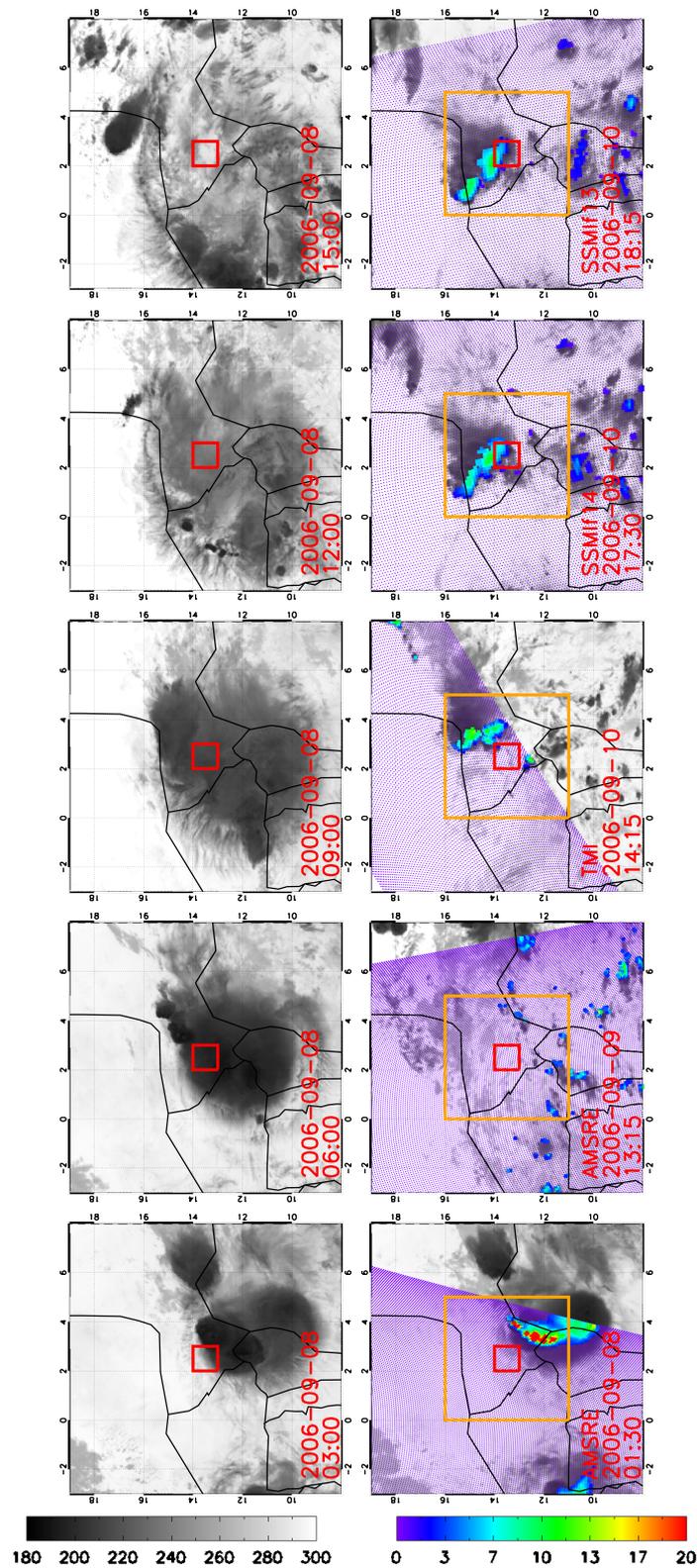


Figure B.2 – Time series (first column) of Infrared Brightness temperatures (in Kelvin, grey colorbar) images for the 8th of September 2006 over the region of Niamey, Niger (the red and orange squares correspond to the 1° and 5° areas close to the rain gauge network of Niamey). Time series (second column) of Infrared Brightness temperatures images during the period ± 2 -days off the 8th of September 2006, selected when the BRAIN algorithm, from a constellation of passive MW imagers, provides rain estimations (in $\text{mm}\cdot\text{h}^{-1}$, rainbow colorbar)

Numerous studies have demonstrated that $T_{threshold}$ can be adapted to local meteorological conditions for a more accurate estimation of the rainy cloud fraction, at scales shorter than a month and for areas smaller than $2.5^\circ \times 2.5^\circ$ (Adler *et al.*, 1993; Xu *et al.*, 1999; Huffman *et al.*, 2001; Todd *et al.*, 2001; Kidd *et al.*, 2003).

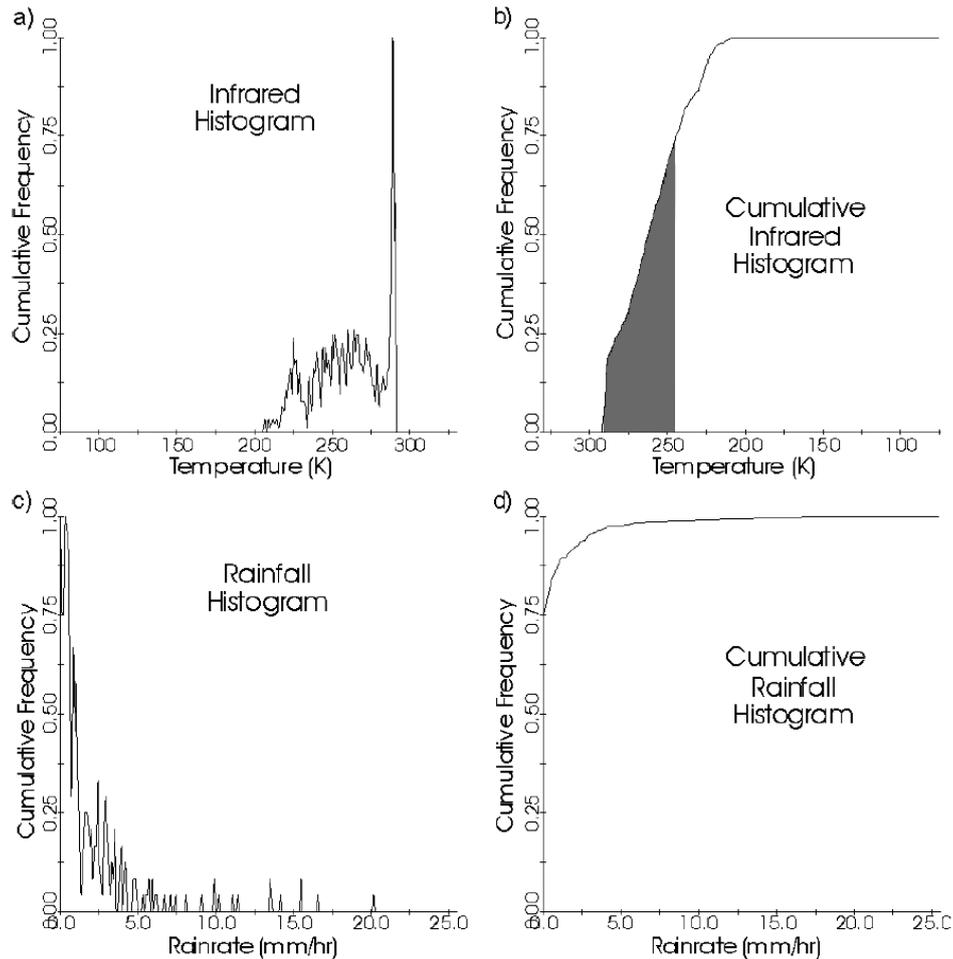


Figure B.3 – Distributions of Infrared pixels (top left) and MW-derived rain rates (bottom left). Cumulated distributions of IR pixels (top right) and MW-derived rain rates (bottom right) for a $5^\circ \times 5^\circ \times 5$ -days volume (Kidd *et al.*, 2003)

One method to adapt $T_{threshold}$ to local meteorological conditions is to work with local distributions of MW-derived rain estimations, collocated with TIR data, and perform a histogram matching. Figure B.3 shows an example of the procedure : local distributions of Infrared and Rainfall data are collected in a $5^\circ \times 5^\circ \times 5$ -day volume, surrounding a $1^\circ/1$ -day volume in which F_C needs to be estimated. These distributions are cumulated to derive cumulated distribution functions. $T_{threshold}$ is then set so that the fraction of collocated TIR data warmer than this threshold is equal to the fraction of non-rainy collocated Rain Rates. In the example of Figure

B.3, 75% of the MW-derived rain samples are non-rainy samples : $T_{threshold}$ is then set to approximately 240 K. The second parameter R_{COND} is set to the mean rainy value of this distribution for keeping the rain volume estimated in the MW-derived rain rate local distribution.

The cloud top temperature/rainfall relationship is likely to vary over a wide range of scales, from inter-annual to sub-daily scales depending on the physical properties of the environment of rainy events like the tropical wave dynamics (Machado *et al.*, 1993), the relative humidity available in the boundary layer (Roca *et al.*, 2005), soil moisture (Taylor *et al.*, 2010), local solar time (Laing *et al.*, 2008).

This simple histogram matching method converges toward values of $T_{threshold}$ and R_{COND} depending on the volume of data used for their computation. This volume of data, called training volume in the following, has two characteristics : the size of the geographical domain and the duration of the period considered for the training of a $1^\circ/1$ -day rain estimation.

The two parameters $T_{threshold}$ and R_{COND} are computed for all the $1^\circ \times 1^\circ \times 1$ -day volumes of the whole Tropics and for a given period. The local volumes of collocated data are chosen to overlap each other with a moving window of $5^\circ \times 5^\circ \times 5$ -day to avoid large discontinuities in the $(T_{threshold}, R_{COND})$ fields. For each day of the period, maps of parameters $(T_{threshold}, R_{COND})$ are then created. Figure B.4 shows an example of two maps of $T_{threshold}$ and R_{COND} used to compute the rain accumulation map of Figure B.7. One can see that in the Tropical Atlantic ocean, the two parameters are close to 235 K and 3 mm.h^{-1} as found during the GATE experiment, but exhibit a large variability over the whole Tropical belt. In particular, large regions like the Sahelian band or the Northern part of Bay of Bengal present very cold $T_{threshold}$ under 200 K to exclude raining from non-raining clouds. This variability demonstrates the geographical dependence of the cloud top temperature/rainfall relationship and confirms the need of their local computation.

The size of the geographical domain and the duration of the period of the training volume are two parameters of the TAPEER algorithm. They are closely related to a third degree of freedom : the configuration of the constellation of Passive MW Observing Systems. In chapter 4 of this manuscript, we show a sensitivity study of these three degrees of freedom on the Tropical Amount of Precipitation as well as on the Estimate of ERrors. Depending on the number of Passive MW Observing Systems available, the training volume can be selected in order to compute $1^\circ/1$ -day best estimates.

In order to compute the rainy cloud fraction F_C for each $1^\circ \times 1^\circ \times 1$ -day volume, time series of TIR geostationary images are segmented : T_B warmer than the local $T_{threshold}$ are set to "0" (non rainy) samples and T_B colder than the local $T_{threshold}$, are set to "1" (rainy) samples. These Rain/No-Rain maps are then aggregated and averaged to compute local F_C . $\langle R \rangle_{A,T}$ is then derived using Equation B.1. Another technical solution which was chosen in the present implementation of the TAPEER algorithm, is to assign 0 mm.h^{-1} to T_B warmer than the

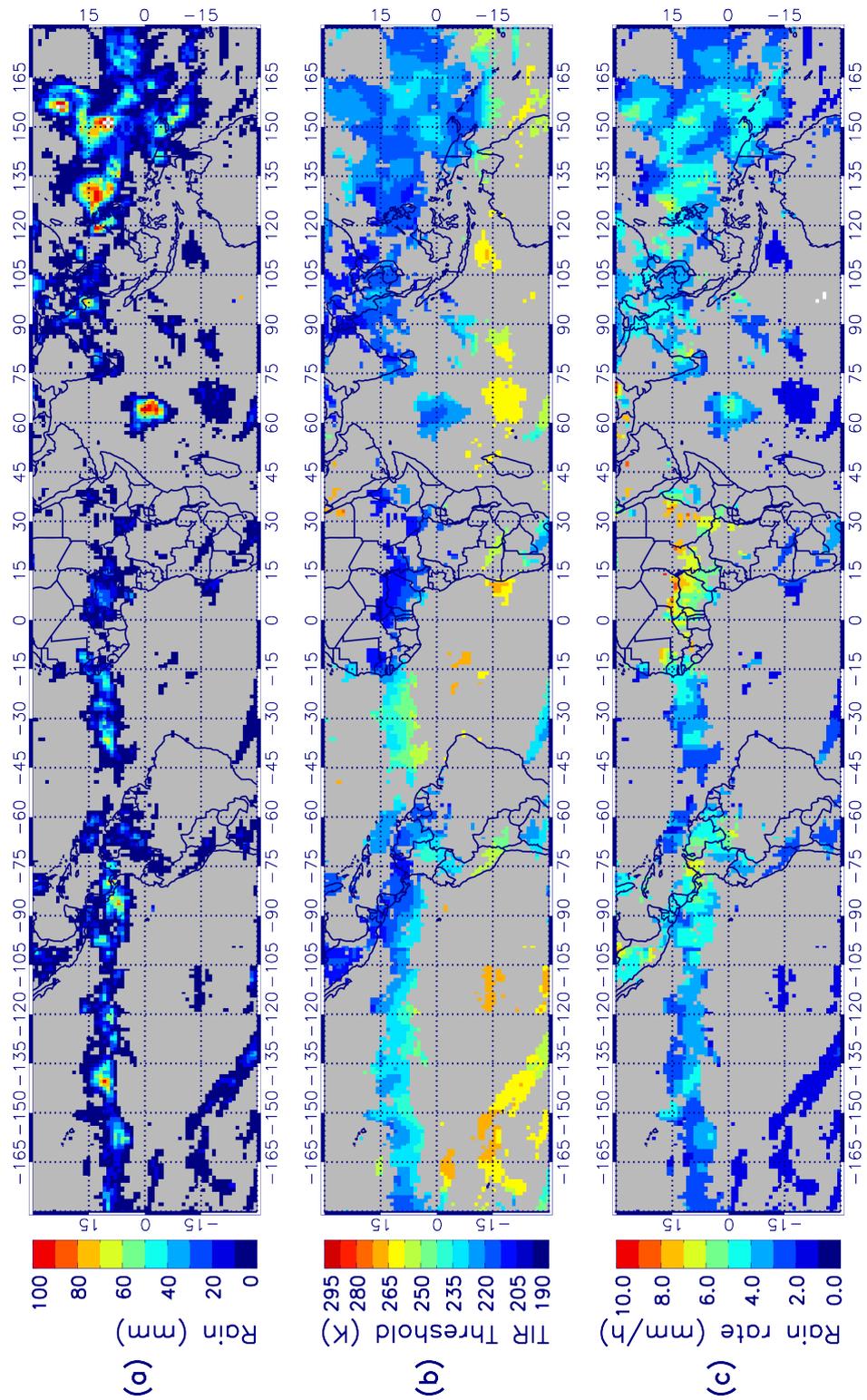


Figure B.4 – $T_{threshold}$ (middle) and R_{COND} (right) used to compute a TAPEER $1^\circ/1$ -day precipitation map (left)

local $T_{threshold}$, and R_{COND} mm.h⁻¹ to T_B colder than the local $T_{threshold}$. The rain rates maps are then averaged at the 1°×1°×1-day. In this case, the rain rates maps may be seen as instantaneous rain samples at the high resolution of geostationary TIR data. These rain samples are contaminated by very large random errors due the very loose physical/statistical relationship on which their computation is based. These random errors average out when the rain rates maps are aggregated at the accumulated scale.

B.1.2 Estimate of ERrors

The error budget of satellite rainfall estimations is composed of a sum of three terms, under the assumption that error sources are uncorrelated. These different terms are added to compute the Estimate of ERrors in the TAPEER algorithm :

$$S^2 = S_{Sampling}^2 + S_{Algorithm}^2 + S_{Calibration}^2 \quad (B.2)$$

Both $S_{Algorithm}^2$ and $S_{Calibration}^2$ terms are the consequence of the errors on the two data sources used in TAPEER : MW-derived rain estimations and TIR brightness temperature data. Errors on these two data sources can be large, especially for the MW-derived rain estimations, and possibly be magnified in the area-integrated rain accumulations. To account for these error sources in the error computed in the TAPEER algorithm, a forward error propagation technique was developed (Chambon *et al.*, 2011). This technique leads to an estimation of $S_{Algorithm}^2$ and $S_{Calibration}^2$ for realistic magnitude of errors on MW-derived rainfall estimation algorithms and on TIR brightness temperature data.

The $S_{Sampling}^2$ term represents the error related to the discrete nature of measurements in space and time, used to estimate an area-integrated rain accumulation. An error model involving variograms computation is used to estimate this first term (Roca *et al.*, 2010a). This model consists in the computation of the variance of estimation of a mean, given σ^2 the variance of the samples and $N_{Independent}$ the number of independent samples composing the mean rainfall accumulation :

$$S_{Sampling}^2 = \frac{\sigma^2}{N_{Independent}} \quad (B.3)$$

For this model, a local estimate of $N_{Independent}$ is required. To this end, local space and time variograms, $\gamma(\Delta x, t_0)$ and $\gamma(\vec{x}_0, \Delta t)$, are computed to extract $N_{Independent}$ from the samples constituting $\langle R \rangle_{A,T}$. These samples are the "0" (non rainy) samples and "1" (rainy) samples or the 0 mm.h⁻¹ samples and R_{COND} mm.h⁻¹ samples discussed above. In order to avoid the contamination of variograms by the large random errors in the rain rate maps, variograms are

computed over the rain/no-rain fields or Indicator fields (noted IF fields below) (Barancourt *et al.*, 1992)).

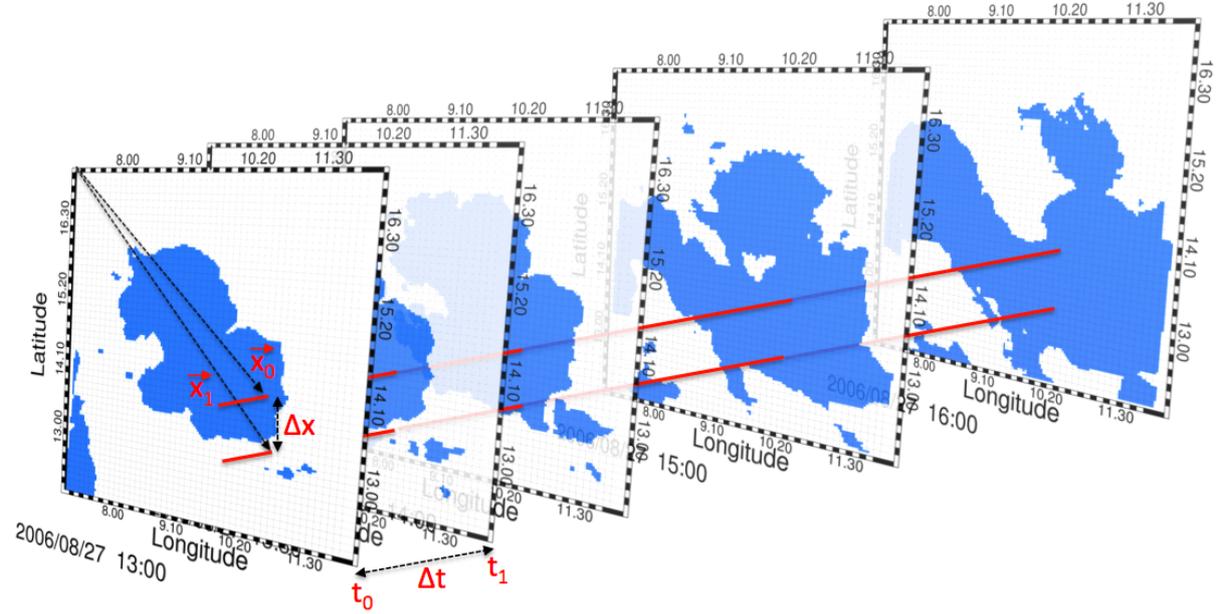


Figure B.5 – Time series of TAPEER-BRAIN Indicator fields for the region $[7^\circ \text{ E} - 12^\circ \text{ E}; 12^\circ \text{ N} - 17^\circ \text{ N}]$. The two red lines, positioned at \vec{x}_0 \vec{x}_1 show the time series over which two time variograms $\gamma(\vec{x}_0, \Delta t)$ and $\gamma(\vec{x}_1, \Delta t)$ can be computed.

For a slot t_0 of the IF time series, the space variogram $\gamma(\Delta x, t_0)$ can be expressed with the Equation B.4 where Δx is the lag space between two samples, $n(\Delta x)$ is the number of pairs of samples distant of Δx , and σ_{SPACE}^2 the variance of the IF field over the area and at the slot t_0 where the space variogram is computed.

For a location \vec{x}_0 of the IF fields, the time variogram $\gamma(\vec{x}_0, \Delta t)$ can be expressed with the Equation B.5 where Δt is the lag time between two samples, $m(\Delta t)$ is the number of pairs of samples distant of Δt and $\sigma_{TIME}^2(\vec{x}_0)$ the variance of the IF fields over the period and the location \vec{x}_0 where the time variogram is computed.

$$\gamma(\Delta x = \|\vec{x}_1 - \vec{x}_0\|, t_0) = \frac{\frac{1}{n(\Delta x)} \sum_{i=0}^{n(\Delta x)} [IF(\vec{x}_1, t) - IF(\vec{x}_0, t)]^2}{\sigma_{SPACE}^2(t_0)} \quad (\text{B.4})$$

$$\gamma(\vec{x}_0, \Delta t = |t' - t|) = \frac{\frac{1}{m(\Delta t)} \sum_{i=0}^{m(\Delta t)} [IF(\vec{x}_0, t') - IF(\vec{x}_0, t)]^2}{\sigma_{TIME}^2(\vec{x}_0)} \quad (\text{B.5})$$

Figure B.5 shows a part of the time series of IF fields used to compute Figure 6 of Chambon *et al.* (2011). Over each image of $5^\circ \times 5^\circ$, a space variogram can be computed : $\gamma(\Delta x, t_0)$ for the slot t_0 , $\gamma(\Delta x, t_1)$ for the slot t_1 , etc. For each pixel, a time series over a given period can be

extracted (red lines of Figure B.5); a time variogram can be computed : $\gamma(\vec{x}_0, \Delta t)$ for the pixel at position \vec{x}_0 , $\gamma(\vec{x}_1, \Delta t)$ for the pixel at position \vec{x}_1 .

Deriving an estimate of $N_{Independent}$ in a $1^\circ \times 1^\circ \times 1$ -day volume requires the averaging of variograms over the volume. Indeed, where a single variogram reflects the autocorrelation of only a fraction of the samples composing the mean rainfall accumulation, averaged space and time variograms $\Gamma_{A,T}(\Delta x)$ and $\Gamma_{A,T}(\Delta t)$ reflect the ensemble autocorrelation :

$$\Gamma_{A,T}(\Delta x) = \frac{1}{N_{t_0 \in T}} \sum_{t_0 \in T} \gamma(\Delta x, t_0) \quad (\text{B.6})$$

$$\Gamma_{A,T}(\Delta t) = \frac{1}{M_{\vec{x}_0 \in A}} \sum_{\vec{x}_0 \in A} \gamma(\vec{x}_0, \Delta t) \quad (\text{B.7})$$

Figure 3 of Roca *et al.* (2010a) shows $\Gamma_{A,T}(\Delta x)$ and $\Gamma_{A,T}(\Delta t)$ for three different satellite products and the modeling of them by an exponential function. In the TAPEER algorithm, $\Gamma_{A,T}(\Delta x)$ and $\Gamma_{A,T}(\Delta t)$ are also modeled by exponential functions and the derived space and time e-folding distances d and τ are used to compute $N_{Independent}$ using the following equation :

$$N_{Independent} = \frac{A}{d^2} \cdot \frac{T}{\tau} \quad (\text{B.8})$$

In terms of practical considerations, using a domain larger than $1^\circ \times 1^\circ \times 1$ -day to average variogram functions $\Gamma_{A,T}(\Delta x)$ and $\Gamma_{A,T}(\Delta t)$ leads to more robust variograms. A compromise between averaging variograms for robustness and detecting the variability of d and τ parameters was found to be an averaging volume of $5^\circ \times 5^\circ \times 10$ -days. Variogram calculations are time consuming, so the following assumptions are used to limit computation cost :

- A regular space grid is assumed in $5^\circ \times 5^\circ$ domains of geostationary TIR data, variations in inter-pixel distance are only taken into account from a $5^\circ \times 5^\circ$ domain to another $5^\circ \times 5^\circ$.
- The smaller time resolution of geostationary TIR data is 30 minutes.
- An exponential model is assumed for all variograms.
- $5^\circ \times 5^\circ \times 10$ -days domains do not overlap each other, d and τ are considered to be constant inside such a volume.

B.2 TAPEER implementation for Megha-Tropiques and GPM missions : the TAPEER-BRAIN product

B.2.1 Observing systems

B.2.1.1 IR Fleet

Time series of geostationary images in the form of 2D images are used for the calculation of the $T_{threshold}$ and R_{COND} parameters and the computation of the rain rates maps. The

full space-time resolution datasets acquired by the thermal IR channel of a fleet of geostationary satellites is used to process the gridded rainfall accumulation over the whole tropical belt ($\pm 30^\circ\text{N}$ of latitudes) (Table B.1). All geostationary satellites are processed independently over regions covering the Sub-Satellite Point $\pm 50^\circ\text{E}$ and then merged to provide the global maps.

Data	Satellite	Sub-Satellite Point	Space resolution	Time resolution
T_B IR - $10.8\mu\text{m}$ (K)	Meteosat-9	0°	3 km	15 min
T_B IR - $11.5\mu\text{m}$ (K)	Meteosat-7	57.5°E	5 km at nadir	30 min
T_B IR - $10.7\mu\text{m}$ (K)	GOES-11	135°W	5 km at nadir	30 min
T_B IR - $10.7\mu\text{m}$ (K)	GOES-12	75°W	5 km at nadir	30 min
T_B IR - $10.8\mu\text{m}$ (K)	MTSAT-1R	140°E	5 km at nadir	30 min

Tableau B.1 – Thermal Infrared Brightness Temperatures derived from geosynchronous Infrared Imagers.

B.2.1.2 Microwave Constellation

The Megha-Tropiques Level-2 MW-derived rain product (BRAIN algorithm) includes rain retrievals from MW data acquired by the MADRAS imager as well as a constellation of Passive MW observing systems. The list provided in Table B.2 is likely to change depending on the availability of imagers and other considerations detailed in chapter 4.

Data	Satellite	Radiometer	Space resolution
BRAIN (mm.h^{-1})	Megha-Tropiques	MADRAS	TBD
BRAIN (mm.h^{-1})	TRMM	TMI	~ 12 km
BRAIN (mm.h^{-1})	AQUA	AMSR-E	~ 9 km
BRAIN (mm.h^{-1})	DMSP-F13	SSM-I	~ 20 km
BRAIN (mm.h^{-1})	DMSP-F15	SSM-I	~ 20 km

Tableau B.2 – Instantaneous precipitation products derived from the microwave brightness temperatures acquired by various observing systems.

B.2.1.3 Space and Time GEO/LEO Collocation files

A collocation between the MT Level 2 rain product and geostationary IR data is performed using the following specifications :

- A maximum time interval between two observations of half the standard time resolution of each geostationary IR datasets is allowed (e.g 7.5 minutes over the region using Meteosat-9 data, 15 minutes over the region using Meteosat-7)
- Data of the MT Level 2 rain products are assumed to cover circular areas of diameters specified in Table B.2

The colocation technique involves distance calculation between IR pixels and Level-2 rain products pixels and a correspondence file between the MT Level 2 rain products and geostationary data is produced.

B.2.1.4 General Pre-processing

A general pre-processing of the geostationary images must be carried out upstream and can be useful for all users of these images. It permits to make or check the geometrical images correction, the radiometric correction (calibration) and to check the quality of the images (brightness temperature margin errors). The quality check procedure adopted to flag images with a fraction of Good Line Rate is detailed in Szantai *et al.* (2011). A Good Line Rate threshold is set to avoid the use of poor quality images based on the statistics presented in Table B.3.

Satellite	Number of tested images	Fraction of GLR of 100% over [30°S - 30°N]	GLR Threshold
Meteosat-9	11261	99.83	TBD
Meteosat-7	5589	99.71	TBD
GOES-11	1085	93.94	TBD
GOES-11 Northern sector	4091	91.38	TBD
GOES-12	1049	98.04	TBD
GOES-12 Northern sector	4720	84.12	TBD
GOES-12 Southern sector	3502	80.89	TBD
MTSAT-1R	-	-	TBD

Tableau B.3 – Good Line Rate (GLR) statistics of geostationary IR data for the 2009 June-to-September season

B.2.2 Algorithm processing

Figure B.6 shows the general functioning of the TAPEER-BRAIN product in which two parameters are computed with a co-located training dataset (Database preparation) and then applied to the all GEO dataset (Full resolution rain field generation). These two parameters are : a IR Brightness Temperature Threshold defining rainy areas and a Conditional Rain Rate

defining the intensity of rainfall. The threshold is calculated so that the occurrence of IR pixels colder than this threshold corresponds to the occurrence of rainy pixels in the training dataset; the Conditional Rain Rate is the mean rain rate observed in the training dataset. A space-time running window surrounding a nominal day and a nominal one-degree square defines the volume of co-located data for which the two parameters are calculated. The error calculated for a one-degree daily accumulation is a variance of estimation taking into account the over sampling in space and time among the samples of the full resolution estimations. This over sampling is estimated through space and time variogram computation and modeling (Variograms computation).

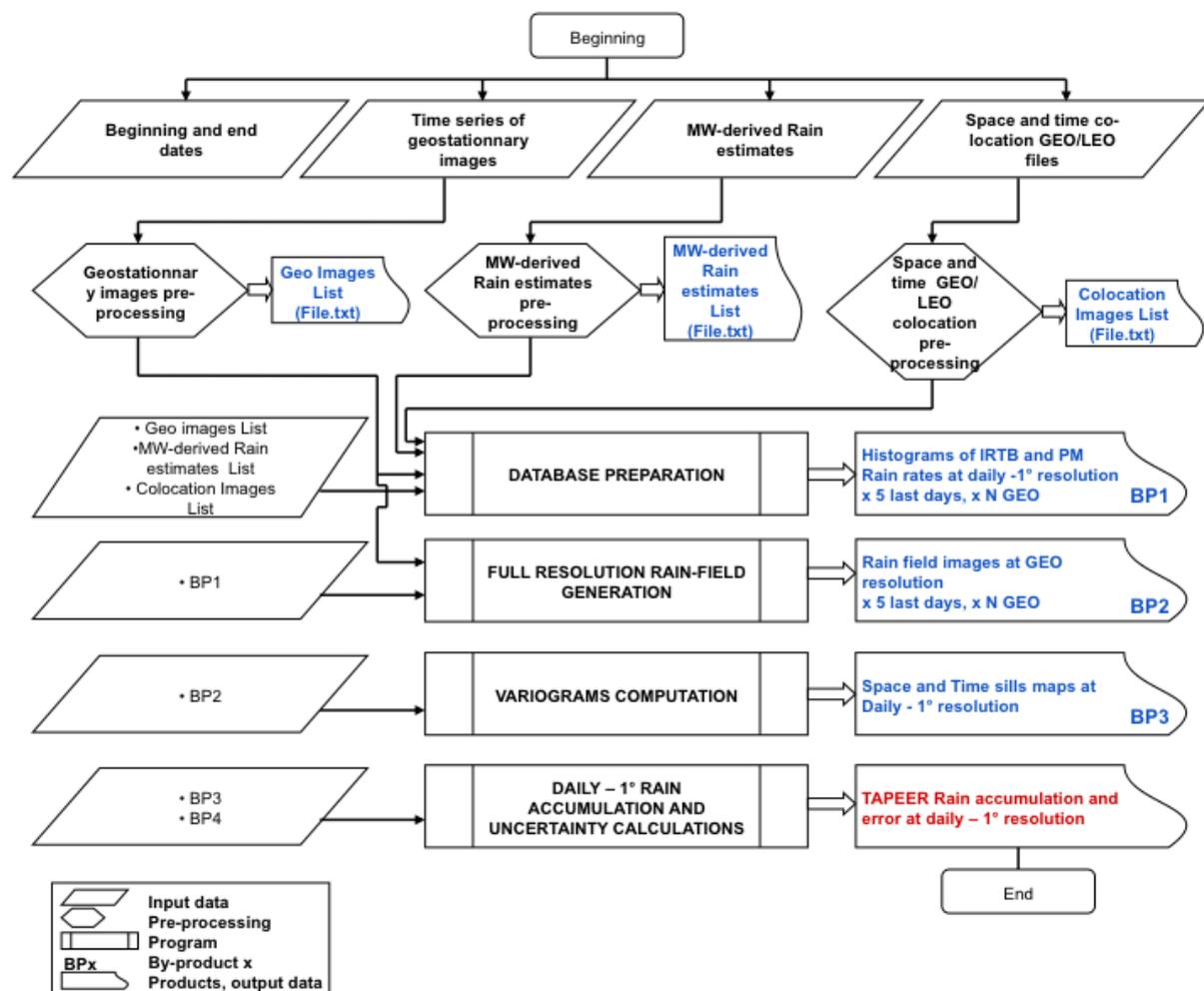


Figure B.6 – TAPEER flow chart

A preliminary product was built for the June-to-September 2006 period over Western Africa and for the June-to-September 2009 period over the whole Tropical belt, using BRAIN-TMI/TRMM, BRAIN-AMSRE/AQUA, BRAIN-SSMI/DMSP-F13 and BRAIN-SSMI/DMSP-F15 rain products.

Figure B.7 shows an example of two maps of rain accumulations and sampling errors derived by TAPEER-BRAIN for the 30th of July 2009. Figure B.8 shows relative error distributions obtained for the 2009 TAPEER-BRAIN preliminary product. One can see that sampling errors can be very large for rain accumulations lower than 2 mm; the errors decrease and range between 10% and 30 % for rain accumulations greater than 20 mm.

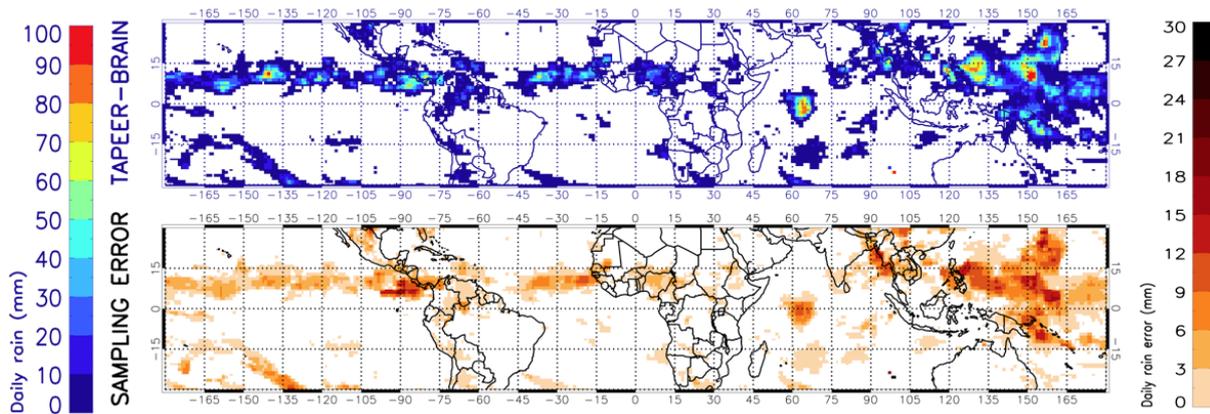


Figure B.7 – TAPEER-BRAIN accumulated rainfall (top) and sampling error (bottom) for the 30th of July 2009

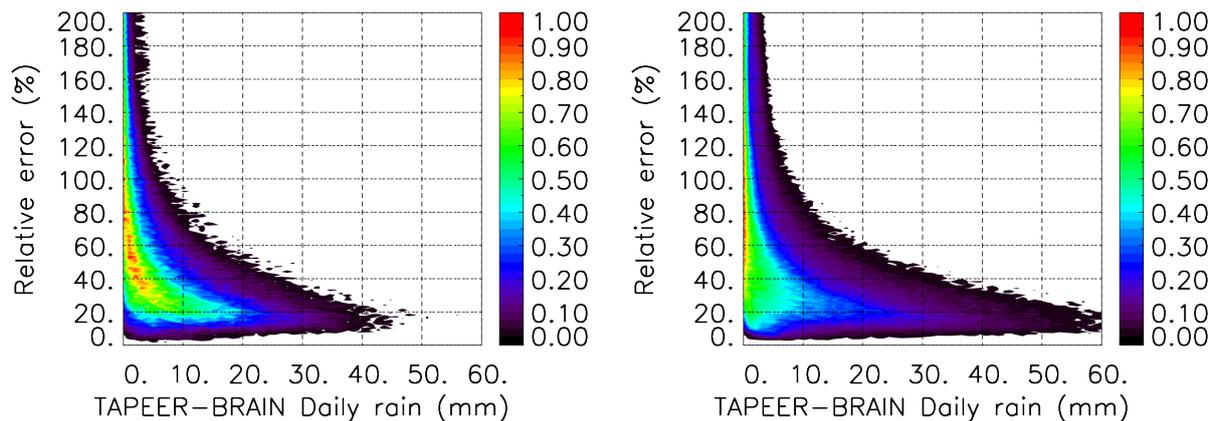


Figure B.8 – Relative sampling errors (%) as a function of TAPEER-BRAIN rain accumulations for the June-to-September 2009 period, over land (left) and over ocean (right).

B.3 Preliminary validation of the TAPEER-BRAIN product

B.3.1 Comparison with TRMM-3B42 over the whole Tropical belt

A comparison was performed between the preliminary TAPEER-BRAIN product and the TRMM-3B42 product aggregated at the $1^\circ/1$ -day scale. Day-to-day validation scores (without errors) are computed between maps of the two products over the whole tropical belt along the June-to-September 2009 period.

Over ocean (figure B.10), the various scores demonstrate a stable behavior (e.g. the correlation coefficient ranges between 0.75 and 0.85 along the season) except the bias which is more frequently negative in June and July but more frequently positive in August and September.

Over land (figure B.10), as only high frequency channels can be used due to surface emissivity, rain rate estimations exhibit larger errors than over ocean and by extent retrieval algorithms exhibit stronger disagreements. A larger disagreement between the TAPEER-BRAIN product and the TRMM-3B42 product is thus expected over land than over the ocean. Indeed, the comparison performed between TAPEER-BRAIN and TRMM-3B42 demonstrates results less stable than over ocean (e.g. the correlation coefficient ranges between 0.6 and 0.8 along the season).

B.3.2 Comparison with surface rainfall products over West Africa at Meteorologically relevant scales

The comparison of TAPEER-BRAIN, TRMM-3B42 V6, TRMM-3B42RT and GSMAP MVK was performed at the meteorologically relevant scales selected in chapter 2 of this manuscript. The analysis is performed using error estimates.

The results of the comparisons for the 10-day scale are reported in table B.4 and table B.5; for the daily scale they are reported in table B.6 and table B.7 and for the 3-5 day filtered daily scale in table B.8 and table B.9.

Figure B.11 and figure B.12 shows the scatterograms for the 10-day scale; figure B.13 and figure B.14 shows the scatterograms for the daily scale; figure B.15 and figure B.16 shows the scatterograms for the 3-5 day filtered daily scale.

The preliminary validation of the TAPEER-BRAIN product shows general good performances at the three meteorologically relevant selected scales. The comparison scores of TRMM-3B42 V6 and TRMM-3B42RT exhibit differences ranging from low to high. These differences are related to the rain gauge calibration performed improving the TRMM-3B42 V6 rain estimations.

The TAPEER-BRAIN product does not include rain gauge data, but even so it presents similar statistics than the TRMM-3B42 V6 over most of the presented cases.

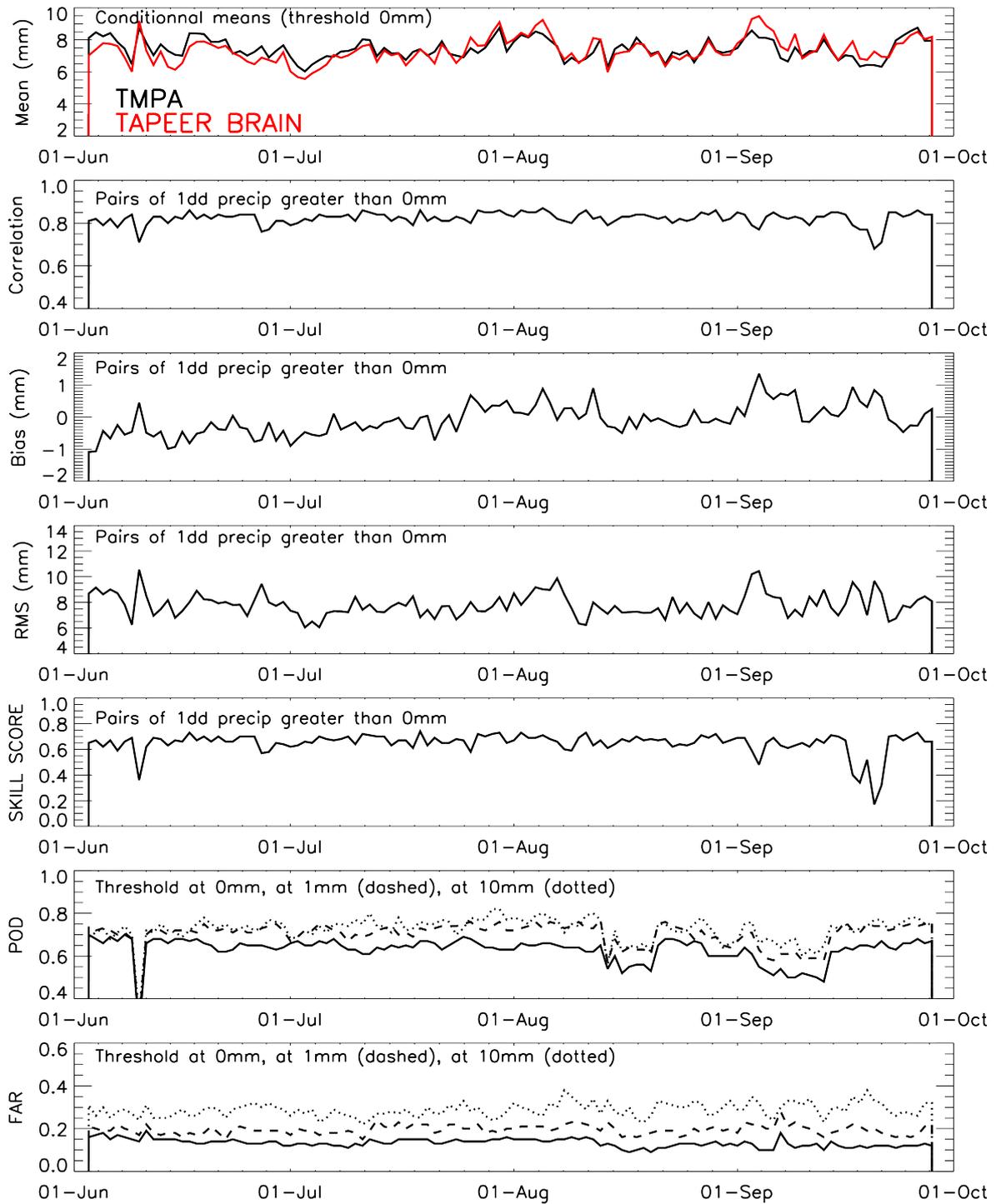


Figure B.9 – Comparison of TAPEER-BRAIN daily estimates and TMPA daily estimates over oceans of the whole Tropical belt for the June-to-September 2009 period

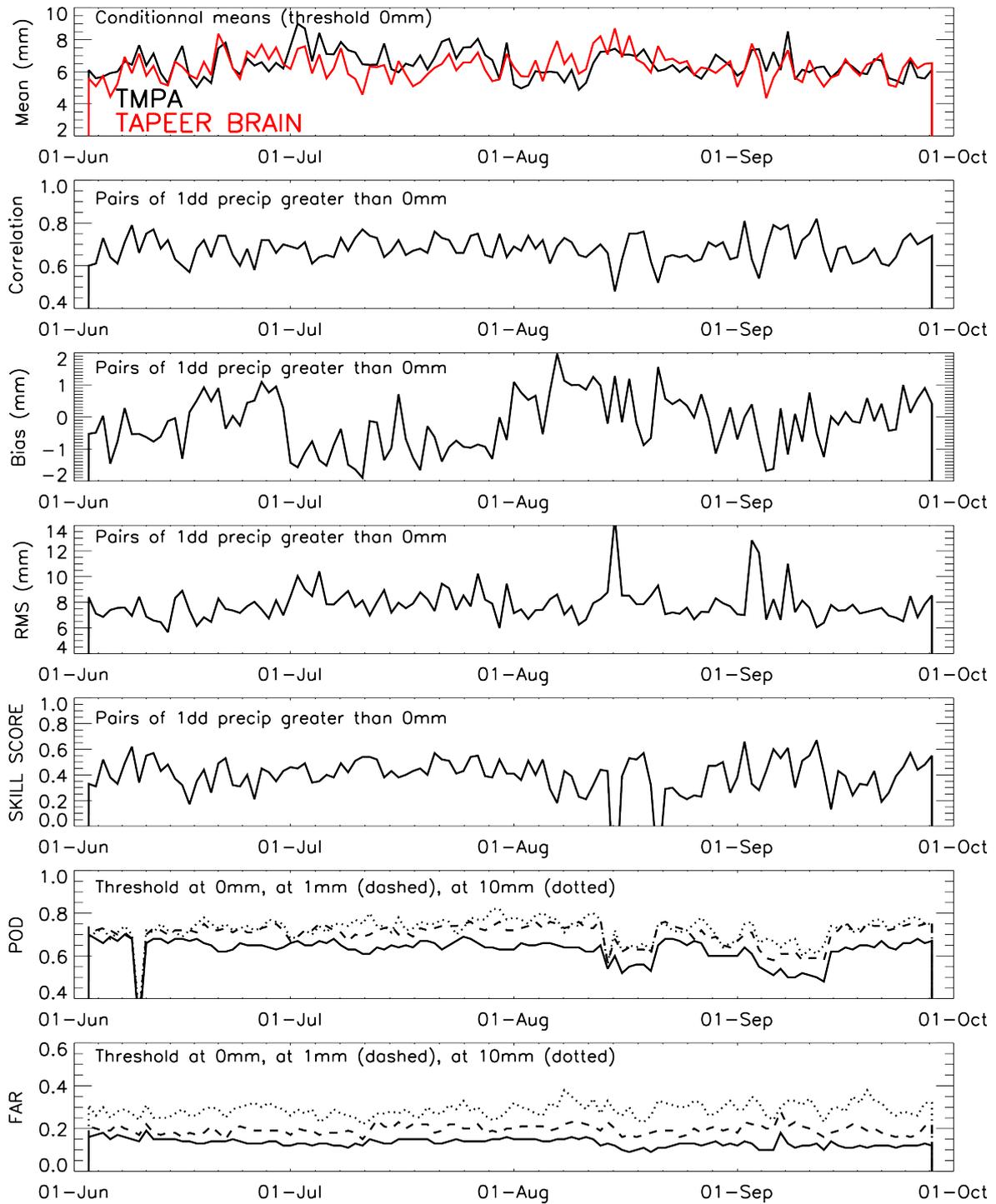


Figure B.10 – Comparison of TAPEER-BRAIN daily estimates and TMPA daily estimates over continents of the whole Tropical belt for the June-to-September 2009 period

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	F
GSMAP	125	0.80	0.79	-1.67	-5.34	8.33	1.05
TMPA-RT	122	0.99	1.83	-11.36	3.73	1.96	0.60
TMPA	125	0.96	0.80	-0.29	-3.75	3.34	0.67
TAPEER- BRAIN	120	0.72	1.13	0.73	3.18	14.51	1.32

Tableau B.4 – Statistical scores of comparison between 10-day rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 10-day rainfall gauge estimates over the Sahelian band for the second 10-period of June 2006

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	F
GSMAP	128	0.80	0.76	-13.13	-27.8	12.07	0.64
TMPA-RT	128	0.89	1.64	-35.16	4.27	20.91	0.40
TMPA	128	0.93	1.24	-23.71	-8.74	11.83	0.33
TAPEER- BRAIN	128	0.71	1.08	-1.70	3.3	23.5	0.43

Tableau B.5 – Statistical scores of comparison between 10-day rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 10-day rainfall gauge estimates over the Sahelian band for the first 10-period of August 2006

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	FAR	POD	F
GSMAP	63	0.61	0.39	1.31	-3.44	4.24	0.40	0.83	1.55
TMPA-RT	57	0.67	0.56	1.85	-1.76	6.93	0.33	0.74	1.62
TMPA	57	0.72	0.62	2.20	-0.92	5.85	0.40	0.78	1.44
TAPEER- BRAIN	60	0.77	0.69	2.15	-0.33	4.78	0.35	0.76	1.20

Tableau B.6 – Statistical scores of comparison between daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and daily rainfall gauge estimates over Niamey

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	FAR	POD	F
GSMAP	81	0.64	0.38	0.33	-5.01	4.38	0.26	0.80	1.54
TMPA-RT	82	0.87	0.72	0.70	-1.17	5.32	0.28	0.89	1.20
TMPA	81	0.89	0.74	0.34	-1.86	5.07	0.29	0.89	1.18
TAPEER- BRAIN	79	0.76	0.63	0.58	-2.67	4.56	0.29	0.73	1.38

Tableau B.7 – Statistical scores of comparison between daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and daily rainfall gauge estimates over Ouémé

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	F
GSMAP	63	0.81	0.51	0.41	-2.22	1.37	0.78
TMPA-RT	57	0.70	0.48	2.06	-0.80	2.74	0.81
TMPA	57	0.86	0.82	0.65	-0.34	1.75	0.51
TAPEER- BRAIN	60	0.83	0.90	0.30	-0.24	2.23	0.55

Tableau B.8 – Statistical scores of comparison between 3-5 day filtered daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 3-5 day filtered daily rainfall gauge estimates over Niamey

	Nb	R	Slope	Intercept (mm)	BIAS-reg (mm)	RMS-reg (mm)	F
GSMAP	81	0.71	0.44	0.24	-3.46	2.18	1.01
TMPA-RT	82	0.93	1.27	-1.05	0.80	2.30	0.56
TMPA	81	0.94	1.24	-0.85	0.75	2.06	0.52
TAPEER- BRAIN	79	0.89	0.85	-0.60	-1.57	1.88	0.63

Tableau B.9 – Statistical scores of comparison between 3-5 day filtered daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 3-5 day filtered daily rainfall gauge estimates over Ouémé

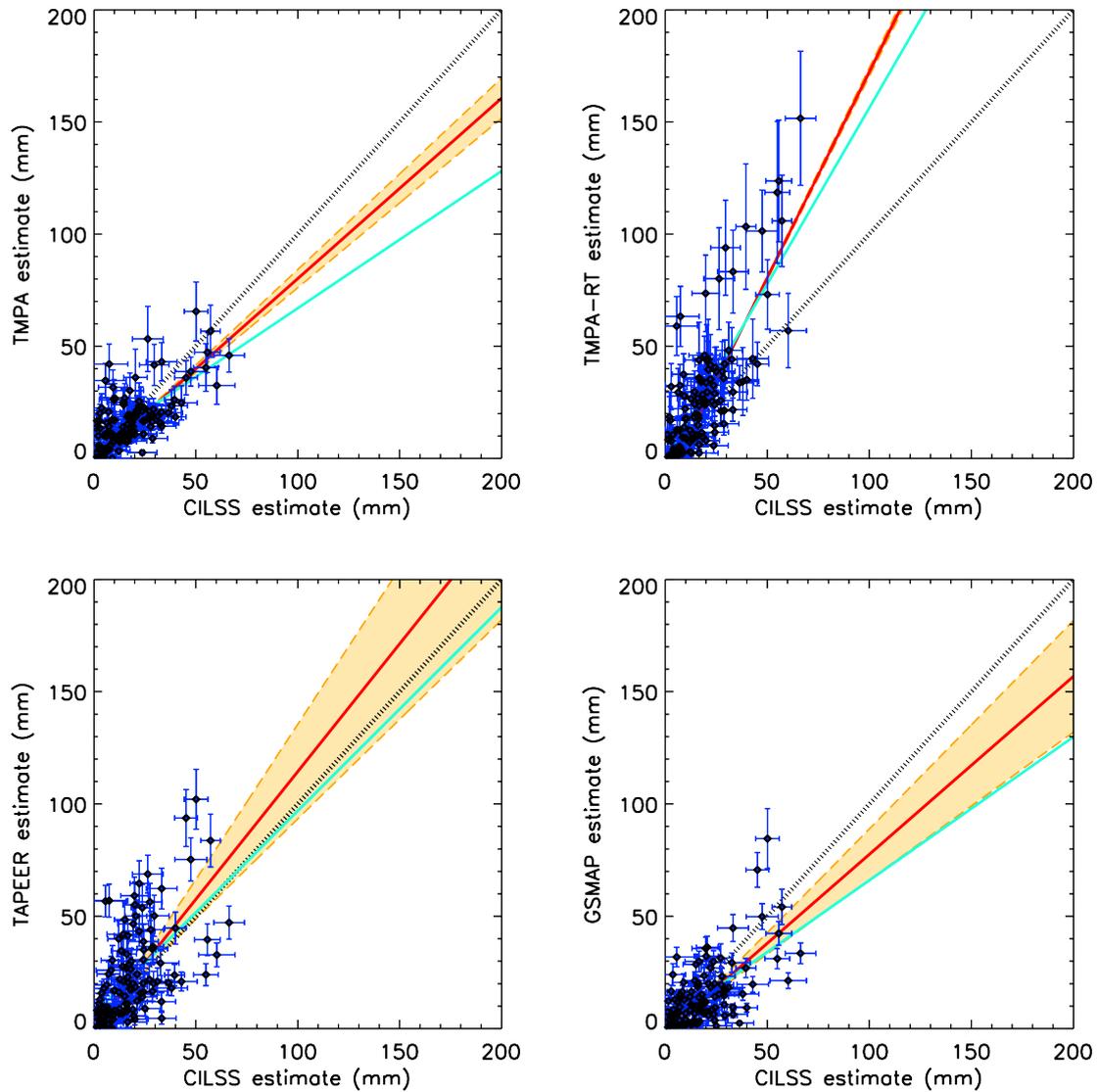


Figure B.11 – Scattergrams of the 10-day rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the 10-day rainfall gauge estimates over the Sahelian band for the second 10-day period of June 2006. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 : 1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.

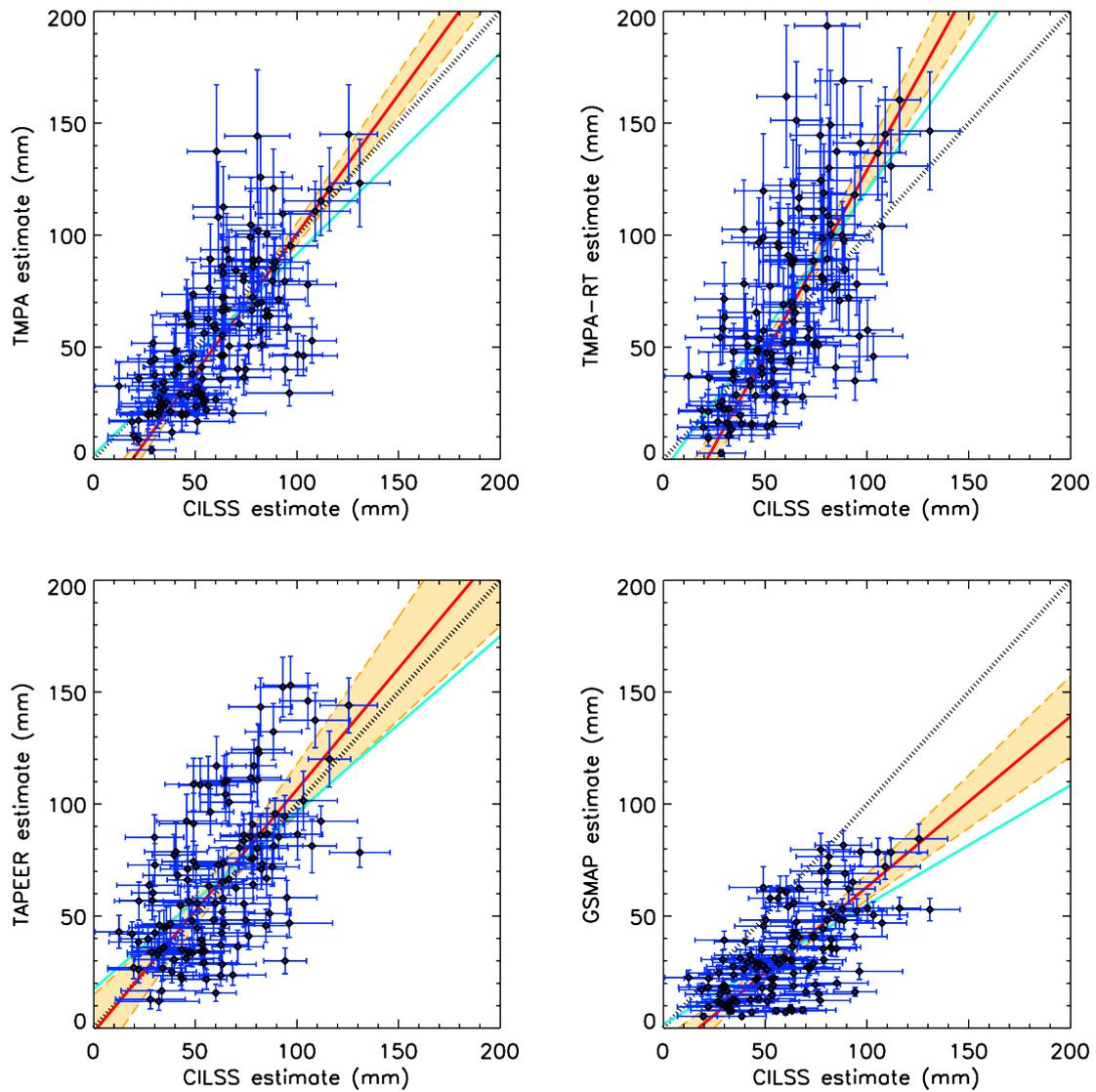


Figure B.12 – Same as figure B.11 for the first 10-period of August 2006.

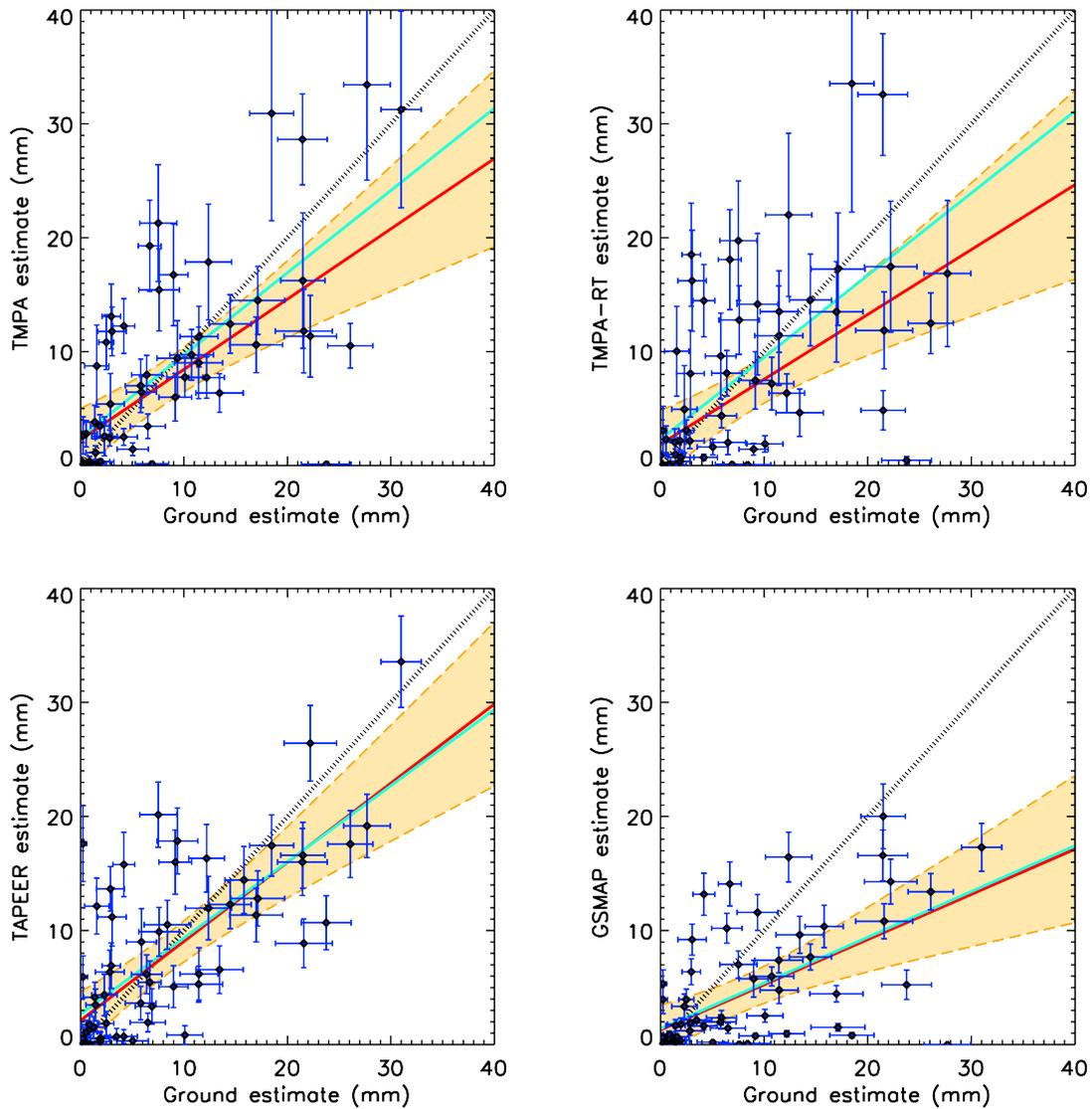


Figure B.13 – Scatterograms of the daily rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the daily rainfall gauge estimates over the Niamey site. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 :1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.

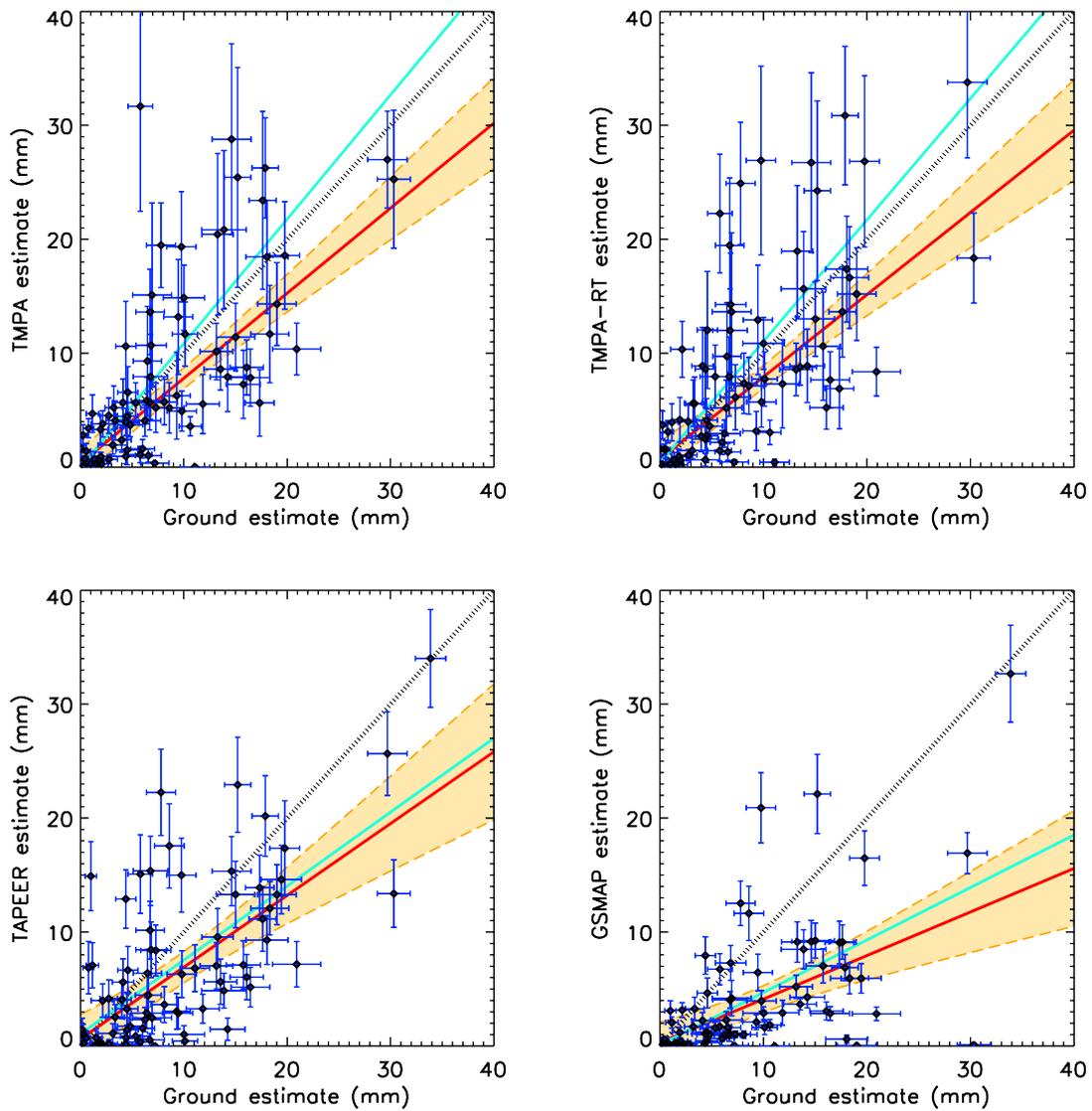


Figure B.14 – Same as figure B.13 for the Ouémé site.

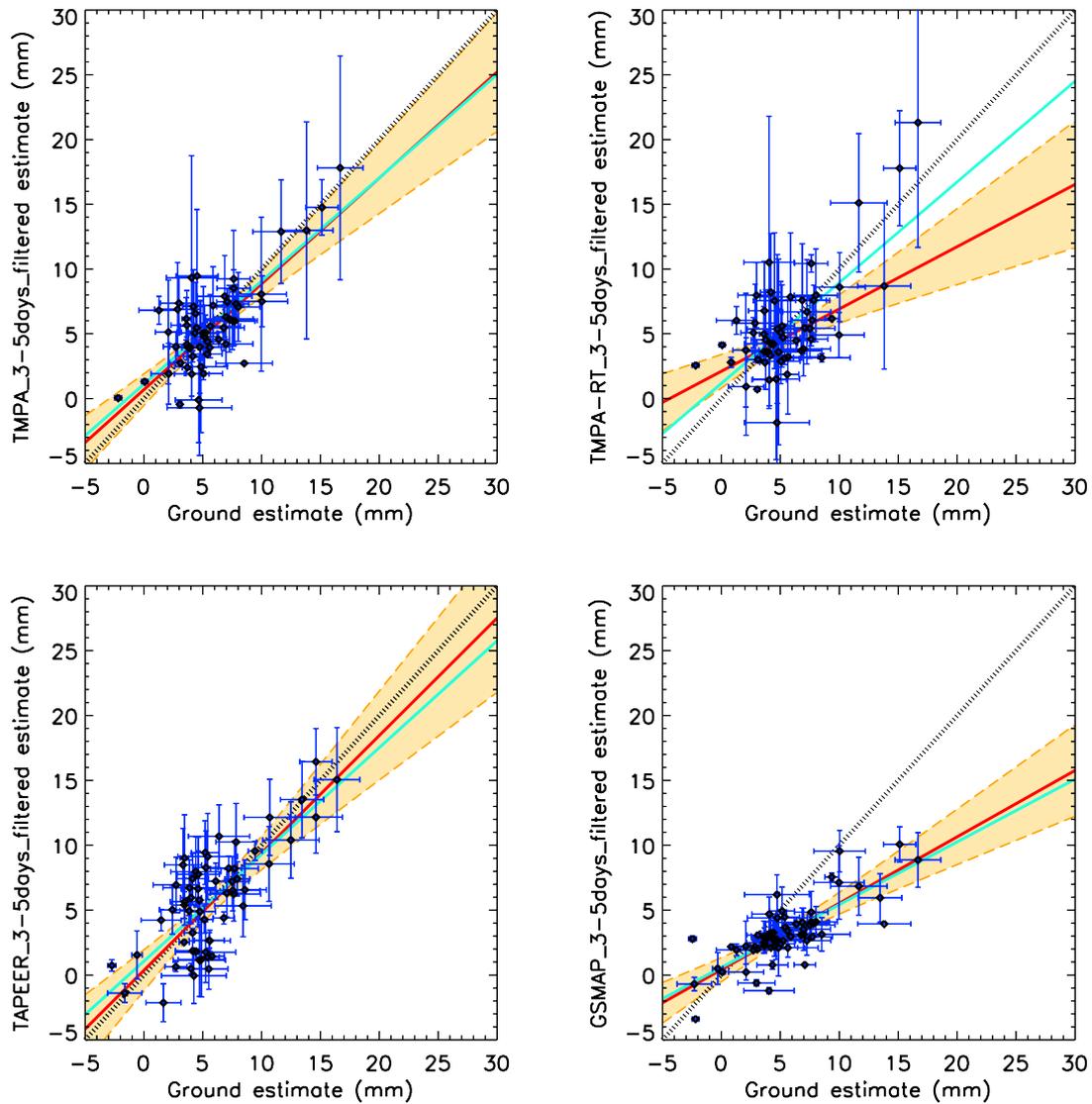


Figure B.15 – Scattergrams of the 3-5 day filtered daily rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the daily rainfall gauge estimates over the Niamey site. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 :1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.

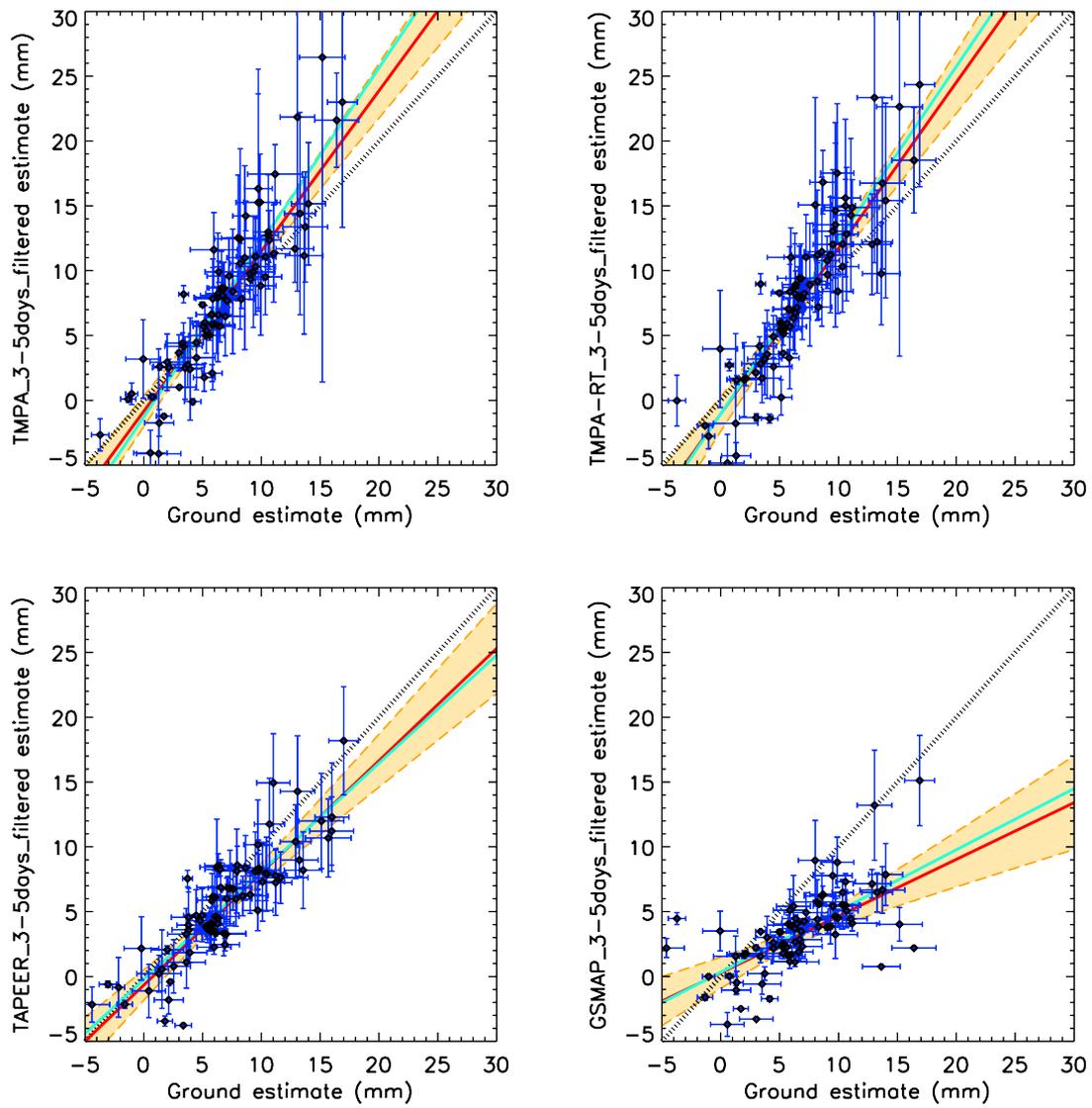


Figure B.16 – Same as figure B.15 for the Ouémé site.

Bibliographie

- Adler, R., Negri, A., Keehn, P., and Hakkarinen, I. (1993). Estimation of monthly rainfall over japan and surrounding waters from a combination of low-orbit microwave and geosynchronous ir data. *J. Appl. Meteorol.*, **32**(2), 335–356. (Cité en pages 56 et 140.)
- Adler, R., Kidd, C., Petty, G., Morrissey, M., and Goodman, H. (2001). Intercomparison of global precipitation products : The third precipitation intercomparison project (pip-3). *Bull. Amer. Meteorol. Soc.*, **82**(7), 1377–1396. (Cité en page 14.)
- Ali, A. and Lebel, T. (2008). The sahelian standardized rainfall index revisited. *Int. J. Climatol.* (Cité en page 30.)
- Ali, A., Lebel, T., and Amani, A. (2005a). Rainfall estimation in the sahel. part i : Error function. *J. Appl. Meteorol.*, **44**(11), 1691–1706. (Cité en pages 32, 34, 35, 36 et 183.)
- Ali, A., Amani, A., Dieddhiou, A., and Lebel, T. (2005b). Rainfall estimation in the sahel. part ii : Evaluation of rain gauge networks in the cilss countries and objective intercomparison of rainfall products. *J. Appl. Meteorol.*, **44**(11), 1707–1722. (Cité en pages 16, 31, 32, 37 et 61.)
- Allan, R., Soden, B., John, V., Ingram, W., and Good, P. (2010). Current changes in tropical precipitation. *Environmental Research Letters*, **5**. (Cité en pages 3, 4, 5, 13 et 181.)
- Allen, M. and Ingram, W. (2002). Constraints on future changes in climate and the hydrological cycle. *Nature*, **419**, 224–232. (Cité en page 3.)
- Andersson, A., Fennig, K., Klepp, C., Bakan, S., Garssl, H., and Schulz, J. (2010). The hamburg ocean atmosphere parameters and fluxes form satellite data - hoaps-3. *Earth Syst. Sci. Data*, **2**, 215–234. (Cité en pages 2 et 181.)
- Aonashi, K. and Liu, G. (2000). Passive microwave precipitation retrievals using tmi during the baiu period of 1998. part i : Algorithm description and validation. *J. Appl. Meteorol.*, **39**, 2024–2037. (Cité en page 36.)
- Aonashi, K., Awaka, J., Hirose, M., Kozu, T., Kubota, T., Liu, G., Shige, S., Kida, S., Seto, S., Takahashi, N., and Takayabu, Y. (2009). Gsmap passive, microwave precipitation retrieval algorithm : Algorithm description and validation. *J. Meteor. Soc. Japan*, **87A**, 119–136. (Cité en pages 12 et 89.)
- Arkin, P. (1979). The relationship between fractional coverage of high cloud and rainfall accumulations during gate over the b-scale array. *Mon. Weather Rev.*, **107**, 1382–1387. (Cité en pages 7 et 138.)

- Arkin, P. and Meisner, B. (1987). The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Mon. Weather Rev.*, **115**(1), 51–74. (Cité en pages 7, 56 et 60.)
- Arkin, P., Smith, T., Sapiano, M., and Janowiak, J. (2010). The observed sensitivity of the global hydrological cycle to changes in surface temperature. *Environmental Research Letters*, **5**. (Cité en page 13.)
- Barancourt, C., Creutin, J., and Rivoirard, J. (1992). A method for delineating and estimating rainfall fields. *Water Resour. Res.*, **28**(4), 1133–1144. (Cité en page 144.)
- Barbé, L. L., Lebel, T., and Tapsoba, D. (2002). Rainfall variability in west africa during the years 1950–90. *J. Climate*, **15**, 187–202. (Cité en page 30.)
- Barrett, E., Dodge, J., Goodman, M., Janowiak, J., Kidd, C., and Smith, E. (1994). The first wetnet precipitation inter-comparison project. *Remote Sensing Reviews*, **11**, 49–60. (Cité en page 14.)
- Bell, T. and Kundu, P. (1996). A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic model. *J. Climate*, **9**(6), 1251–1268. (Cité en pages 21 et 57.)
- Bell, T. and Kundu, P. (2000). Dependence of satellite sampling error on monthly averaged rain rates :comparison of simple models and recent studies. *J. Climate*, **13**(2), 449–462. (Cité en page 21.)
- Bell, T. and Kundu, P. (2003). Comparing satellite rainfall estimates with rain gauge data : Optimal strategies suggested by a spectral model. *J. Geophys. Res.*, **108**. (Cité en pages 13, 16, 21, 37 et 182.)
- Bell, T., Abdullah, A., Martin, R., and North, G. (1990). Sampling errors for satellite-derived tropical rainfall : Monte carlo study using a space-time stochastic model. *J. Geophys. Res.*, **95**, 2195–2205. (Cité en pages 21, 37, 57 et 61.)
- Bell, T., Kundu, P., and Kummerow, C. (2001). Sampling errors of ssm/i and trmm rainfall averages : Comparison with error estimates from surface data and a simple model. *J. Appl. Meteorol.*, **40**, 938–954. (Cité en page 13.)
- Bellerby, T. (2007). Satellite rainfall uncertainty estimation using an artificial neural network. *J. Hydrometeor.*, **8**(6), 1397–1412. (Cité en page 61.)
- Bellerby, T. and Sun, J. (2005). Probabilistic and ensemble representations of the uncertainty in an ir/microwave satellite precipitation product. *J. Hydrometeor.*, **6**, 1032–1044. (Cité en pages 22, 38 et 61.)

- Bergès, J.-C., Chopin, F., Jobard, I., and Roca, R. (2010). Epsat-sg : A satellite method for precipitation estimation ; its concept and implementation for amma experiment. *Ann. Geophys.*, **28**, 289–308. (Cité en pages 31 et 36.)
- Berne, A., Delrieu, G., Creutin, J.-D., and Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.*, **299**(3-4), 166–179. (Cité en page 36.)
- Beucher, F. (2010). *Météorologie Tropicale : Météorologie Tropicale : des alizés au cyclone*, volume 1. Météo France. (Cité en pages 3 et 106.)
- Bielli, S. and Roca, R. (2009). Scale decomposition of atmospheric water budget over west africa during the monsoon 2006 from ncep/gfs analyses. *Climate Dyn.* (Cité en page 52.)
- Bitew, M. and Gebremichael, M. (2011). Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. *Water Resour. Res.*, **47**(W06526). (Cité en page 26.)
- Bizzarri, B., Gasiewski, A., and Staelin, D. (2007). *Observing Rain by Millimetre-Submillimetre wave sounding from geostationary orbit*, volume Measuring Precipitation From Space : EUR-AINSAT and the Future of *Advances in Global Change Research*, pages 675–692. Springer. (Cité en page 83.)
- Bowler, N. (2006). Explicitly accounting for observation error in categorical verification of forecasts. *Mon. Weather Rev.*, **134**(6), 1600–1606. (Cité en pages 14 et 56.)
- Candille, G. and Talagrand, O. (2008). Impact of observational error on the validation of ensemble prediction systems. *Quart. J. Roy. Meteorol. Soc.*, **134**, 959–971. (Cité en pages 14 et 56.)
- Capderou, M. (2005). *Satellites : Orbits and missions*, volume 564pp. Springer. (Cité en pages 8, 89, 181 et 186.)
- Capderou, M. (2009). Sampling : Comparison with other meteorological satellites. Meghatropiques technical memorandum, LMD. (Cité en pages 8, 89 et 181.)
- Carr, J., Bailey, R., and Deng, E. (1985). Use of indicator variograms for an enhanced spatial analysis. *Mathematical Geology*, **17**(8), 797–811. (Cité en page 75.)
- Carroll, R. and Ruppert, D. (1996). The use and misuse of orthogonal regression estimation in linear errors-in-variables models. *Amer. Stat.*, **50**, 1–6. (Cité en page 41.)
- Chambon, P., Jobard, I., Roca, R., and Viltard, N. (2011). An investigation of the error budget of tropical rainfall accumulation derived from merged passive microwave and infrared satellite measurements. *Quart. J. Roy. Meteorol. Soc.* (Cité en pages 143 et 144.)

- Ciach, G. (2003). Local random errors in tipping-bucket rain gauge measurements. *J. Atmos. Ocean. Technol.*, **20**(5), 752–759. (Cité en pages 20, 32 et 182.)
- Ciach, G. and Krajewski, W. (1999a). On the estimation of radar rainfall error variance. *Adv. Water. Resour.*, **22**(6), 585–595. (Cité en pages 16, 19 et 32.)
- Ciach, G. and Krajewski, W. (1999b). Radar rain gauge comparisons under observational uncertainties. *J. Appl. Meteorol.*, **38**(1519-1525). (Cité en page 16.)
- Conway, D., Persechino, A., Ardoin-Bardin, S., Hamandawana, H., Dieulin, C., and Mahé, G. (2009). Rainfall and water resources variability in sub-saharan africa during the twentieth century. *J. Hydrometeorol.*, **10**, 41–59. (Cité en page 30.)
- Depraetere, C., Gosset, M., Ploix, S., and Laurent, H. (2009). The organization and kinematics of tropical rainfall systems ground tracked at mesoscale with gages : First results from the campaigns 1999–2006 on the upper ouémé valley (benin). *J. Hydrol.*, **375**, 143–160. (Cité en page 49.)
- Desbois, M., Kayiranga, T., Gnamien, B., Guessous, S., and Picon, L. (1988). Characterization of some elements of the sahelian climate and their interannual variations for july 1983, 1984 and 1985 from the analysis of meteosat isccp data. *J. Climate*, **1**(9), 867–904. (Cité en page 30.)
- Ebert, E. (2007). *Methods for Verifying Satellite Precipitation Estimates*, volume 28 of *Advances in Global Change Research*, pages 345–356. Springer. (Cité en pages 32 et 42.)
- Ebert, E. and Manton, M. (1998). Performance of satellite rainfall estimation algorithms during toga coare. *J. Atmos. Sci.*, **55**(9), 1537–1557. (Cité en pages 31 et 56.)
- Ebert, E., Manton, M., Arkin, P., Allam, R., Hoplin, C., and Gruber, A. (1996). Results from the gpcp algorithm intercomparison programme. *Bull. Amer. Meteorol. Soc.*, **77**(12), 2875–2887. (Cité en pages 14, 31 et 56.)
- Ebert, E., Janowiak, J., and Kidd, C. (2007). Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteorol. Soc.*, **88**(1). (Cité en pages 25, 31, 56 et 57.)
- Ferrier, B., Tao, W., and Simpson, J. (1995). A double-moment multiple-phase four-class bulk ice scheme. part ii : Simulations of convective storms in different large-scale environments and comparisons with other bulk parameterizations. *J. Atmos. Sci.*, **52**, 1001–1033. (Cité en page 59.)
- Fioleau, T. and Roca, R. (2011a). Composites of tropical convective systems life cycle using merged geostationary and low earth orbit satellites observations. *Quart. J. Roy. Meteorol. Soc.* (Cité en page 88.)

- Fiolleau, T. and Roca, R. (2011b). Toocan. *IEEE Trans. Geosci. Remote Sens.* (Cité en page 58.)
- Gebremichael, M. and Krajewski, W. (2004). Characterization of the temporal sampling error in space-time-averaged rainfall estimates from satellites. *J. Geophys. Res.*, **109**. (Cité en pages 14, 21, 22, 23, 39, 56 et 183.)
- Gebremichael, M. and Krajewski, W. (2005). Modeling distribution of temporal sampling errors in area-time-averaged rainfall estimates. *Atm. Research*, **73**(3-4), 243–259. (Cité en pages 22, 32, 39, 57 et 61.)
- Gebremichael, M., Krajewski, W., Morrissey, M., Langerud, D., Huffman, G., and Adler, R. (2003). Error uncertainty analysis of gpcp monthly rainfall products : A data-based simulation study. *J. Appl. Meteorol.*, **42**, 1837–1848. (Cité en page 19.)
- Gebremichael, M., Liao, G.-Y., and Yan, J. (2011). Nonparametric error model for a high resolution satellite rainfall product. *Water Resour. Res.*, **47**(W07504). (Cité en pages 23, 24, 56, 61, 75 et 183.)
- Ghelli, A. and Santos, C. (2010). Accounting for observational uncertainty when verifying precipitation forecasts. In *3rd WMO International Conference on Quantitative Precipitation Estimation and Quantitative Precipitation Forecasting and Hydrology, 18-22 October 2010, Nanjing China*, pages 200–204. World Meteorological Organization. (Cité en pages 14 et 56.)
- Giannini, A., Biasutti, M., and Verstraete, M. (2008a). A climate model-based review of drought in the sahel : Desertification, the re-greening and climate change. *Global Planet. Change*, **64**, 119–128. (Cité en page 30.)
- Giannini, A., Biasutti, M., Held, I., and Sobel, A. (2008b). A global perspective on african climate. *Climate Change*, **90**, 359–383. (Cité en page 30.)
- Gif, N., Chomette, O., and Raberanto, P. (2011). Co-location algorithms : Geophysical data projection using pixel point spread function. Megha-tropiques technical memorandum, LMD. (Cité en page 90.)
- Gosset, M., Viarre, J., and et al. (2011). Evaluation of several rainfall products used for hydrological applications over west africa, using two high resolution gages network. *Quart. J. Roy. Meteorol. Soc.* (Cité en page 56.)
- Grimes, D. (2009a). *An Ensemble Approach to Uncertainty Estimation for Satellite-Based Rainfall Estimates*, volume 63 of *Water Science and Technology Library*, pages 145–162. Springer. (Cité en pages 19 et 22.)

- Grimes, D. (2009b). *Geostatistical Tools for Validation of Satellite and NWP Model Rainfall Estimates*, volume 63 of *Water Science and Technology Library*, pages 117–143. Springer. (Cité en page 76.)
- Grimes, D., Pardo-Iguzquiza, E., and Bonifacio, R. (1999). Optimal areal rainfall estimation using raingauges and satellite data. *J. Hydrol.*, **222**(1-4), 93–108. (Cité en pages 22, 32 et 56.)
- Gu, G. and Adler, R. (2004). Seasonal evolution and variability associated with the west african monsoon system. *J. Climate*, **17**, 3364–3377. (Cité en page 30.)
- Gu, G., Adler, R., Huffman, G., and Curtis, S. (2003). Summer synoptic-scale waves over west africa observed by trmm. *Geophys. Res. Lett.*, **30**(14). (Cité en page 30.)
- Gu, G., Adler, R., Huffman, G., and Curtis, S. (2007). Tropical rainfall variability on interannual-to-interdecadal and longer time scales derived from the gpcp monthly product. *J. Climate*, **20**(15), 4033–4046. (Cité en page 4.)
- Habib, E., Krajewski, W., and Kruger, A. (2001). Sampling errors of tipping-bucket raingauge measurements. *J. Hydrol. Eng.*, **6**, 159–166. (Cité en page 32.)
- Hastenrath, S. (1991). *Climate Dynamics of the Tropics*. Number 488. Springer. (Cité en page 30.)
- Held, I. and Soden, B. (2006). Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699. (Cité en page 3.)
- Herman, A., Kumar, V., Arkin, P., and Kousky, J. (1997). Objectively determined 10-day african rainfall estimates created for famine early warning systems. *Int. J. Remote Sens.*, **18**, 2147–2159. (Cité en page 31.)
- Hong, Y., Hsu, K.-L., Sorooshian, S., and Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteorol.*, **43**(12), 1834–1853. (Cité en pages 12, 21 et 182.)
- Hong, Y., Hsu, K.-L., Moradkhani, H., and Sorooshian, S. (2006). Uncertainty quantification of satellite precipitation estimation and monte carlo assessment of the error propagation into hydrologic response. *Water Resour. Res.*, **42**. (Cité en pages 5, 14, 21, 62 et 182.)
- Hossain, F. and Huffman, G. (2008). Investigating error metrics for satellite rainfall data at hydrologically relevant scales. *J. Hydrometeorol.*, **9**(3), 563–575. (Cité en page 51.)
- Hossain, F., Anagnostou, E., and Dinku, T. (2004). Sensitivity analyses of satellite rainfall retrieval and sampling error on flood prediction uncertainties. *IEEE Trans. Geosci. remote Sens.*, **42**, 130–139. (Cité en page 62.)

- Hou, A., Kakar, R., Neeck, S., Ardeshir, A., Azarbarzin, A., Kummerow, C., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2011). The global precipitation measurement (gpm) mission. *Bull. Amer. Meteorol. Soc.* (Cité en pages 12, 83, 109 et 187.)
- Hsu, K.-L., Gao, X., Sorooshian, S., and Gupta, H. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.*, **36**(9), 1176–1190. (Cité en pages 15, 31 et 56.)
- Huffman, G. (1997). Estimates of root-mean-square random error for finite samples of estimated precipitation. *J. Appl. Meteorol.*, **36**(9), 1191–1201. (Cité en pages 25, 32, 38, 39, 57 et 95.)
- Huffman, G., Adler, R., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U. (1997). The global precipitation climatology project (gpcp) combined precipitation dataset. *Bull. Amer. Meteorol. Soc.*, **78**(1), 5–20. (Cité en pages 25, 37, 56 et 92.)
- Huffman, G., Adler, R., Morrissey, M., Bolvin, D., Curtis, S., Joyce, R., McGavock, B., and Susskind, J. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**(1), 36–50. (Cité en pages 31, 56, 92 et 140.)
- Huffman, G., Adler, R., Bolvin, D., Gu, G., E.J., N., Bowman, K., Hong, Y., Stocker, E., and Wolff, D. (2007). The trmm multisatellite precipitation analysis (tampa) : Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**(1), 38–55. (Cité en pages 12, 14, 15, 31, 37, 56, 70, 106 et 135.)
- Huffman, G., Adler, R., Bolvin, D., and Gu, G. (2009). Improving the global precipitation record : Gpcp version 2.1. *Geophys. Res. Lett.*, **36**(L17808). (Cité en page 25.)
- Huffman, G., Bolvin, D., Braithwaite, B., Hsu, K., Joyce, R., Xie, P., and Yoo, S.-H. (2010). Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg). Algorithm Theoretical Basis Document (ATBD) Version 1.0, NASA/GSFC. (Cité en pages 109 et 187.)
- Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., and Okamoto, K. (2000). Rain-profiling algorithm for the trmm precipitation radar. *J. Appl. Meteorol.*, **39**(12), 2038–2052. (Cité en page 128.)
- Iida, Y., Okamoto, K., and Ushio, T. (2005). Sampling simulation for estimating the sampling error of space-time average rain rate for trmm and gpm mission. *IEEE Trans. Geosci. Remote Sens.*, pages 87–90. (Cité en pages 21 et 84.)
- IPCC (2007). Fourth assessment report : Climate change. (Cité en page 4.)
- Iridium (2011). Iridium next - global climate observation. Technical report. (Cité en page 83.)

- Islam, M. and Ueyda, H. (2006). Trmm observed vertical structure and diurnal variation of precipitation in south asia. In *Proceedings in the IEEE International Geoscience and Remote Sensing Symposium (IGARSS06), August, Denver, USA*, pages 1–4. (Cité en page 99.)
- Janicot, S., Thorncroft, C., Ali, A., Ascencio, N., Berry, G., Bock, O., Bourles, B., Caniaux, G., Chauvin, F., Deme, A., Kergoat, L., Lafore, J.-P., Lavaysse, C., Lebel, T., Marticorena, B., Mounier, F., Nedelec, P., Redelsperger, J.-L., Ravegnani, F., Reeves, C., Roca, R., de Rosnay, P., Schlager, H., Sultan, B., Tomasini, M., Ulanovsky, A., and forecasters team, A. (2008). Large-scale overview of the summer monsoon over west africa during the amma field experiment in 2006. *Ann. Geophys.*, **26**, 2569–2595. (Cité en page 43.)
- Jenkins, G. and Gaye, A. (2010). Increasing research opportunities in the atmospheric sciences for underrepresented groups through international field experiences in senegal. *Bull. Amer. Meteorol. Soc.*, **91**(7), 845–852. (Cité en page 35.)
- Jobard, I. (1996). Estimation des pluies par satellite : l’apport des données satellite micro-onde. (Cité en pages 123, 124 et 188.)
- Jobard, I. and Desbois, M. (1994). Satellite estimation of the tropical precipitation using the meteosat and ssm/i data. *Atm. Research*, **34**(1-4), 285–298. (Cité en pages 32 et 56.)
- Jobard, I., Chopin, F., Bergès, J., and Roca, R. (2011). An intercomparison of 10-day precipitation satellite products during west african monsoon. *Int. J. Remote Sens.*, **32**(9), 2353–2376. (Cité en pages 14, 15, 25, 31, 56 et 182.)
- Journel, A. (1983). Nonparametric estimation of spatial distributions. *Mathematical Geology*, **15**(3), 445–468. (Cité en page 75.)
- Journel, A. and Huijbregts, C. (1978). *Mining Geostatistics*. Number 600. Academic Press. (Cité en pages 17 et 35.)
- Joyce, R. and Xie, P. (2011). Kalman filter based cmorph. *J. Hydrometeor.* (Cité en pages 12, 26 et 88.)
- Joyce, R., Janowiak, J., Arkin, P., and Xie, P. (2004). Cmorph : A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**(3), 487–503. (Cité en pages 15, 31, 36 et 56.)
- Kawanishi, T., Sezai, T., Ito, Y., Imaoka, K., Takeshima, T., Ishido, Y., Shibata, A., Miura, M., Inahata, H., and Spencer, R. (2003). The advanced microwave scanning radiometer for the earth observing system (amsr-e), nasda’ s contribution to the eos for global energy and water cycle studies. *IEEE Trans. Geosci. Remote Sens.*, **41**, 184 – 194. (Cité en page 7.)

- Kelly, B. (2007). Some aspects of measurement error in linear regression of astronomical data. *Astrophys. J.*, **665**, 1489–1506. (Cit  en pages 17, 32, 41, 42, 67 et 182.)
- Kidd, C. (1999). Results of an infrared/passive microwave rainfall estimation technique. In *Proc. Remote Sensing Society, Cardiff, Wales, United Kingdom, Remote Sensing Society*, pages 685–689. (Cit  en page 92.)
- Kidd, C. and Huffman, G. (2011). Review global precipitation measurement. *Meteorol. Appl.*, **18**, 334–353. (Cit  en pages 7 et 12.)
- Kidd, C., Kniveton, D., Todd, M., and Bellerby, T. (2003). Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeor.*, **4**(6), 1088–1104. (Cit  en pages 56, 57, 61, 92, 140 et 189.)
- Kidd, C., Levizzani, V., Turk, F., and Ferraro, R. (2009). Satellite precipitation measurements for water resource monitoring. *J. Amer. Wat. Res. Assoc.*, **45**(3), 567–579. (Cit  en page 12.)
- Kidder, S. and Haar, T. V. (1995). *Satellite meteorology : an introduction*. Academic Press. (Cit  en page 12.)
- Kiladis, G., Thorncroft, C., and Hall, N. (2006). Three-dimensional structure and dynamics of african easterly waves. part i : Observations. *J. Atmos. Sci.*, **63**, 2212–2130. (Cit  en page 30.)
- Kirstetter, P.-E. (2008). *Estimation quantitative des pr cipitations par radar m t orologique : inf rence de la structure verticale des pluies, mod lisation des erreurs radar-pluviom tres*. 276pp, Universit  Joseph Fourier. (Cit  en page 12.)
- Kirstetter, P.-E., Viltard, N., and Gosset, M. (2011). Toward an error model for brain algorithm for instantaneous rainfall estimation in west africa. *Quart. J. Roy. Meteorol. Soc.* (Cit  en pages 26, 59, 75 et 131.)
- Krajewski, W., Ciach, G., McCollum, J., and Bacotiu, C. (2000). Initial validation of the global precipitation climatology project monthly rainfall over the united states. *J. Appl. Meteorol.*, **39**(7), 1071–1086. (Cit  en page 19.)
- Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y., Nakagawa, K., Iwanami, K., Ushio, T., Kachi, M., and Okamoto, K. (2007). Global precipitation map using satelliteborne microwave radiometers by the gsmep project : Production and validation. *IEEE Trans. Geosci. Remote Sens.*, **45**(2259-2275). (Cit  en page 56.)
- Kummerow, C. (1998). Beamfilling errors in passive microwave rainfall retrievals. *J. Appl. Meteorol.*, **37**, 356–370. (Cit  en pages 132 et 188.)

- Kummerow, C. and Giglio, L. (1995). A method for combining passive microwave and infrared rainfall observations. *J. Atmos. Ocean. Technol.*, **12**(1), 33–45. (Cité en page 57.)
- Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simpson, J. (1998). The tropical rainfall measuring mission (trmm) sensor package. *J. Atmos. Ocean. Technol.*, **15**, 809–817. (Cité en page 7.)
- Kummerow, C., Hong, Y., Olson, W., Yang, S., Adler, R., McCollum, J., Ferraro, R., Petty, G., Shin, D., and Wilheit, T. (2001). The evolution of the goddard profiling algorithm (gprof) for rainfall estimation from passive microwave sensors. *J. Appl. Meteorol.*, **40**(1801-1820). (Cité en pages 37 et 128.)
- Laing, A., Carbone, R., Levizzani, V., and Tuttle, J. (2008). The propagation and diurnal cycles of deep convection in northern tropical africa. *Quart. J. Roy. Meteorol. Soc.*, **134**(630), 93–109. (Cité en page 141.)
- Lamptey, B. L. (2008). Comparison of gridded multisatellite rainfall estimates with gridded gauge rainfall over west africa. *J. Appl. Meteorol. and Climatol.*, **47**(1), 185–205. (Cité en page 31.)
- Larsen, R. and Marx, M. (2001). *An Introduction to Mathematical Statistics and Its Applications : 3rd ed.* Number 790. Prentice Hall. (Cité en page 37.)
- Lau, K.-M. and Wu, H.-T. (2007). Detecting trends in tropical rainfall characteristics, 1979-2003. *Int. J. Climatol.*, **27**, 979–988. (Cité en pages 4 et 13.)
- Laughlin, C.-R. (1981). On the effect of temporal sampling on the observation of mean rainfall. In D. Atlas and O.-W. Thiele, editors, *Precipitation Measurements From Space : Workshop Report*, pages D59 – D66. Goddard Space Flight Cent., Greenbelt, Md. (Cité en page 21.)
- Laurent, H., Jobard, I., and Toma, A. (1998). Validation of satellite and ground-based estimates of precipitation over the sahel. *Atm. Research*, **47-48**, 651–670. (Cité en page 31.)
- Lebel, T. and Amani, A. (1999). Rainfall estimation in the sahel : What is the ground truth? *J. Appl. Meteorol.*, **38**(5). (Cité en pages 33 et 35.)
- Lebel, T. and Barbé, L. L. (1997). Rainfall monitoring during hapex-sahel. 2. point and areal estimation at the event and seasonal scales. *J. Hydrol.*, **188-189**, 97–122. (Cité en page 76.)
- Lebel, T., Bastin, G., Obled, C., and Creutin, J. (1987). On the accuracy of areal rainfall estimation : A case study. *Water Resour. Res.*, **23**(11), 2123–2134. (Cité en pages 32, 35, 36 et 76.)

- Lebel, T., Cappelaere, C., Galle, S., Hanan, N., Kergoat, L., Levis, S., Vieux, B., Descroix, L., Gosset, M., and Mougin, E. (2009). Amma-catch studies in the sahelian region of west-africa : an overview. *J. Hydrol.* (Cit  en pages 31, 33 et 59.)
- Lebel, T., Parker, D., Flamant, C., Bourles, B., Marticorena, B., Mougin, E., Peugeot, C., Diedhiou, A., Haywood, J., Ngamini, J., Polcher, J., Redelsperger, J.-L., and Thorncroft, C. (2010). The amma field campaigns : multiscale and multidisciplinary observations in the west african region. *Quart. J. Roy. Meteorol. Soc.*, **136**(S1), 8–23. (Cit  en page 31.)
- Leng, L., Zhang, T., Kleinman, L., and Zhu, W. (2007). Ordinary least square regression, orthogonal regression, geometric mean regression and their application in aerosol science. *J. Phys.*, **78**(012084). (Cit  en page 41.)
- Levizzani, V., Bauer, P., and Turk, F. (2007). *Measuring Precipitation From Space : EURAINSAT and the Future*, volume 28 of *Advances in Global Change Research*. Springer. (Cit  en pages 12 et 31.)
- Liu, C. (2011). Rainfall contributions from precipitation systems with different sizes, convective intensities, and durations over the tropics and subtropics. *J. Hydrometeorol.*, **12**, 394–412. (Cit  en page 118.)
- Loeb, N., Wielicki, B., Doelling, D., Smith, G., Keyes, D., Kato, S., Manalo-Smith, N., and Wong, T. (2009). Toward optimal closure of the earth’s top-of-atmosphere radiation budget. *J. Climate*, **22**, 748–766. (Cit  en pages 2 et 181.)
- Machado, L., Duvel, J., and Desbois, M. (1993). Diurnal variations and modulation by easterly waves of the size distributions of convective cloud clusters over west africa and the atlantic ocean. *Mon. Weather Rev.*, **121**, 37–49. (Cit  en pages 30 et 141.)
- Marzano, F. and Cimini, D. (2010). Flower elliptical-orbit constellation exploiting millimetre-wave radiometry and radio occultation for meteo-climatological applications. *Adv. Geosci.*, **25**, 167–177. (Cit  en page 83.)
- Mittermaier, M. (2008). Introducing uncertainty of radar-rainfall estimates to the verification of mesoscale model precipitation forecasts. *Natural Hazards and Earth System Sciences*, **8**(445–460). (Cit  en page 56.)
- Mohr, K. (2004). Interannual, monthly, and regional variability in the wet season diurnal cycle of precipitation in sub-saharan africa. *J. Climate*, **17**, 2441–2453. (Cit  en page 52.)
- Mohr, K., Famiglietti, J., and Zipser, E. (1999). The contribution to tropical rainfall with respect to convective system type, size, and intensity estimated from the 85-ghz ice-scattering signature. *J. Appl. Meteorol.*, **38**, 596–606. (Cit  en page 118.)

- Moradkhani, H. and Meskele, T. (2010). *Probabilistic Assessment of the Satellite Rainfall Retrieval Error Translation to Hydrologic Response.*, pages 229–242. Springer. (Cité en pages 14 et 62.)
- Morrissey, M. and Greene, J. (1993). Comparison of two satellite-based rainfall algorithms using pacific atoll raingage data. *J. Appl. Meteorol.*, **32**(2), 411–425. (Cité en page 16.)
- Morrissey, M. and Wang, Y. (1995). Verifying satellite microwave rainfall estimates over the open ocean. *J. Appl. Meteorol.*, **34**(4), 794–804. (Cité en pages 16, 35, 37 et 61.)
- Morrissey, M., Maliekal, J., Greene, J., and Wang, J. (1995). The uncertainty of simple spatial averages using rain gauge networks. *Water Resour. Res.*, **31**(8), 2011–2017. (Cité en page 32.)
- Mounier, F., Kiladis, G., and Janicot, S. (2007). Analysis of the dominant mode of convectively coupled kelvin waves in the west african monsoon. *J. Climate*, **20**, 1487–1503. (Cité en page 30.)
- Mueller, J., Hanson, C., Govaerts, Y., Heinemann, T., and Konig, M. (2004). Initial results from the validation of the meteosat-8 seviri calibration. In *Proc. of The 2004 EUMETSAT Meteorological Satellite Conference. Prague. Czech Republic : 31 May – 4 June 2004.* (Cité en page 66.)
- Negri, A., Nelkin, E., Adler, R., Huffman, G., and Kummerow, C. (1995). Evaluation of passive microwave precipitation algorithms in wintertime midlatitude situations. *J. Atmos. Ocean. Technol.*, **12**, 20–32. (Cité en pages 31 et 56.)
- Nesbitt, S. and Zipser, E. (2003). The diurnal cycle of rainfall and convective intensity according to three years of trmm measurements. *J. Climate*, **16**, 1456–1475. (Cité en page 50.)
- Nesbitt, S., Zipser, E., and Kummerow, C. (2004). An examination of version-5 rainfall estimates from the trmm microwave imager, precipitation radar, and rain gauges on global, regional, and storm scales. *J. Appl. Meteorol.*, **43**, 1016–1036. (Cité en page 51.)
- New, M., Hulme, M., and Jones, P. (2000). Representing twentieth-century space-time climate variability. part ii : development of 1901–1996 monthly grids of terrestrial surface climate. *J. Climate*, **13**(2217–2238). (Cité en page 5.)
- New, M., Todd, M., Hulme, M., and Jones, P. (2001). Precipitation measurements and trends in the twentieth century. *Int. J. Climatol.*, **21**, 1899–1922. (Cité en pages 5 et 6.)
- Nicholson, S., Klotter, D., Some, B., McCollum, J., Nelkin, E., Berte, Y., Diallo, B., Gaye, I., Kpabeba, G., Ndiaye, O., Noukpozoukou, J., Tanu, M., Thiam, A., Toure, A., and Traore, A. (2003a). Validation of trmm and other rainfall estimates with a high-density gauge dataset

- for west africa. part i : Validation of gpcc rainfall product and pre-trmm satellite and blended products. *J. Appl. Meteorol.*, **42**(10), 1337–1354. (Cité en page 31.)
- Nicholson, S., Klotter, D., Some, B., McCollum, J., Nelkin, E., Berte, Y., Diallo, B., Gaye, I., Kpabeba, G., Ndiaye, O., Noukpozoukou, J., Tanu, M., Thiam, A., Toure, A., and Traore, A. (2003b). Validation of trmm and other rainfall estimates with a high-density gauge dataset for west africa. part ii : Validation of trmm rainfall products. *J. Appl. Meteorol.*, **42**(10), 1355–1368. (Cité en page 31.)
- Nijssen, B. and Lettenmaier, D. (2004). Effect of precipitation sampling error on simulated hydrological fluxes and states : Anticipating the global precipitation measurement satellites. *J. Geophys. Res.*, **109**(D02103). (Cité en page 14.)
- North, G. and Nakamoto, S. (1989). Formalism for comparing rain estimation designs. *J. Atmos. Ocean. Technol.*, **6**(6), 985–992. (Cité en pages 19, 37, 38, 57 et 182.)
- Pall, P., Allen, M., and Stone, D. (2007). Testing the clausius clapeyron constraint on changes in extreme precipitation under co2 warming. *Climate Dynamics*, **28**, 351–363. (Cité en page 3.)
- Pardo-Iguzquiza, E., Grimes, D., and Teo, C.-K. (2006). Assessing the uncertainty associated with intermittent rainfall fields. *Water Resour. Res.*, **42**. (Cité en page 32.)
- Parent du Châtelet, J. (2003). Aramis, le réseau français de radars pour la surveillance des précipitations. *La Météorologie*, **40**, 44–52. (Cité en page 6.)
- Peterson, T. and Vose, R. (1997). An overview of the global historical climatology network temperature database. *Bull. Amer. Meteorol. Soc.*, **78**, 2837–2849. (Cité en page 5.)
- Peyrillé, P. and Lafore, J.-P. (2007). An idealized two-dimensional framework to study the west african monsoon. part ii : Large-scale advection and the diurnal cycle. *J. Atmos. Sci.*, **64**, 2783–2803. (Cité en page 52.)
- Peyrillé, P., Lafore, J.-P., and Redelsperger, J.-L. (2007). An idealized two-dimensional framework to study the west african monsoon. part i : Validation and key controlling factors. *J. Atmos. Sci.*, **64**, 2765–2782. (Cité en page 52.)
- QPE/QPF (2010). Proceeding of the 3rd wmo international conference on quantitative precipitation estimation and quantitative precipitation forecasting and hydrology, 18-22 october 2010. World Meteorological Organization, Nanjing University of Information Science and Technology. (Cité en page 12.)
- Redelsperger, J.-L., Diongue, A., Dieddhiou, A., Ceron, J.-P., Diop, M., Gueremy, J.-F., and Lafore, J.-P. (2002). Multi-scale description of a sahelian synoptic weather system representative

- of the west african monsoon. *Quart. J. Roy. Meteorol. Soc.*, **128**(582), 1229–1257. (Cité en page 30.)
- Redelsperger, J.-L., Thorncroft, C., Dieddhiou, A., Lebel, T., Parker, D., and Polcher, J. (2006). African monsoon multidisciplinary analysis : An international research project and field campaign. *Bull. Amer. Meteorol. Soc.*, **87**, 1739–1746. (Cité en page 30.)
- Roca, R., Lafore, J.-P., Piriou, C., and Redelsperger, J.-L. (2005). Extratropical dry-air intrusions into the west african monsoon midtroposphere : An important factor for the convective activity over the sahel. *J. Atmos. Sci.*, **62**(2), 390–407. (Cité en pages 30 et 141.)
- Roca, R., Chambon, P., Jobard, I., Kirstetter, P.-E., Gosset, M., and Bergès, J.-C. (2010a). Comparing satellite and surface rainfall products over west africa at meteorologically relevant scales during the amma campaign using error estimates. *J. Appl. Meteorol. and Climatol.*, **49**(4), 715–731. (Cité en pages 56, 143 et 145.)
- Roca, R., Bergès, J.-C., Brogniez, H., Capderou, M., Chambon, P., Chomette, O., Cloché, S., Fiolleau, T., Jobard, I., Lémond, J., Ly, M., Picon, L., Raberanto, P., Szantai, A., and Viollier, M. (2010b). On the water and energy cycles in the tropics. *Comptes Rendus Geoscience*, **342**(4-5), 390–402. (Cité en pages 52 et 56.)
- Roca, R., Srinivasan, J., and coauthors (2011). The megha-tropiques mission. *Quart. J. Roy. Meteorol. Soc.* (Cité en pages 11 et 88.)
- Rudolf, B., Becker, A., Schneider, U., Meyer-Christoffer, A., and Ziese, M. (2010). The new gpcc full data reanalysis version 5” providing high-quality gridded monthly precipitation data for the global land-surface is public available since december 2010. Gpcc status report, Global Precipitation Climatology Centre, gpcc.dwd.de. (Cité en pages 5 et 25.)
- Sapiano, M. and Arkin, P. (2009). An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeor.*, **10**(1), 149–166. (Cité en pages 31, 50 et 56.)
- Smith, E., Lamm, J., Adler, R., Alishouse, J., Aonashi, K., Barrett, E., Bauer, P., Berg, W., Chang, A., Ferraro, R., Ferriday, J., Goodman, S., Grody, N., Kidd, C., Kniveton, D., Kummerow, C., Liu, G., Marzano, F., Mugnai, A., Olson, W., Petty, G., Shibata, A., Spencer, R., Wentz, F., Wilheit, T., and Zipser, E. (1998). Results of wetnet pip-2 project. *J. Atmos. Sci.*, **55**(9), 1483–1536. (Cité en page 14.)
- Steiner, M. (1996). Uncertainty of estimates of monthly areal rainfall for temporally sparse remote observations. *Water Resour. Res.*, **32**, 373–388. (Cité en page 21.)

- Steiner, M., Bell, T., Zhang, Y., and Wood, E. (2003). Comparison of two methods for estimating the sampling-related uncertainty of satellite rainfall averages based on a large radar dataset. *J. Climate*, **16**, 3759–3778. (Cité en pages 21 et 57.)
- Stephens, G. and Kummerow, C. (2007). The remote sensing of clouds and precipitation from space : A review. *J. Atmos. Sci.*, **64**(3742-3765). (Cité en page 75.)
- Stephens, G., L'Ecuyer, T., Forbes, R., Gettleman, A., Golaz, J.-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P., and Haynes, J. (2010). Dreary state of precipitation in global models. *J. Geophys. Res.*, **115**(D24211), 1–13. (Cité en page 5.)
- Sultan, B. and Janicot, S. (2003). The west african monsoon dynamics. part ii : The "preonset" and "onset" of the summer monsoon. *J. Climate*, **16**, 3407–3427. (Cité en page 30.)
- Sultan, B., Janicot, S., and Dieddhiou, A. (2003). The west african monsoon dynamics. part i : Documentation of intraseasonal variability. *J. Climate*, **16**, 3389–3406. (Cité en page 30.)
- Szantai, A., Cloché, S., Six, B., and Roca, R. (2011). Quality control of geostationary data for the megha-tropiques mission. Megha-tropiques technical memorandum, LMD. (Cité en page 147.)
- Tao, W.-K., Simpson, J., Baker, D., Braun, S., Chou, M.-D., Ferrier, B., Johnson, D., Khain, A., Lang, S., Lynn, B., Shie, C.-L., Starr, D., Sui, C.-H., Wang, Y., and Wetzell, P. (2003). Microphysics, radiation and surface processes in the goddard cumulus ensemble (gce) model. *Meteorology and Atmospheric Physics*, **82**, 97–137. 10.1007/s00703-001-0594-7. (Cité en page 130.)
- Taylor, C., Harris, P., and Parker, D. (2010). Impact of soil moisture on the development of a sahelian mesoscale convective system : A case-study from the amma special observing period. *Quart. J. Roy. Meteorol. Soc.*, **136**(s1), 456–470. (Cité en page 141.)
- Teo, C.-K. (2006). *Application of satellite-based rainfall estimates to crop yield forecasting in Africa*. 242p, University of Reading. (Cité en pages 5 et 56.)
- Teo, C.-K. and Grimes, D. (2007). Stochastic modelling of rainfall from satellite data. *J. Hydrol.*, **346**(1-2), 33–50. (Cité en pages 22, 32, 38, 61 et 182.)
- Thorne, V., Coakley, P., Grimes, D., and Dugdale, G. (2001). Comparison of tamsat and cpc rainfall estimates with rainfall, for southern africa. *Int. J. Remote Sens.*, **22**(10), 1951–1974. (Cité en page 22.)
- Todd, M., Kidd, C., Kniveton, D., and Bellerby, T. (2001). A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall. *J. Atmos. Ocean. Technol.*, **18**(5), 742–755. (Cité en pages 92 et 140.)

- Turk, F. and Miller, S. (2005). Toward improved characterization of remotely sensed precipitation regimes with modis/amsr-e blended data techniques. *IEEE Trans. Geosci. remote Sens.*, **43**(5), 1059–1069. (Cité en pages 56, 61 et 92.)
- Ushio, T., Kubota, T., Shige, S., Okamoto, K., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., Kachi, M., Oki, R., Morimoto, T., and Kawasaki, Z. (2009). A kalman filter approach to the global satellite mapping of precipitation (gsmap) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Japan*, **87A**, 137–151. (Cité en pages 12, 31, 36 et 89.)
- Viltard, N., Burlaud, C., and Kummerow, C. (2006). Rain retrieval from tmi brightness temperature measurements using a trmm pr-based database. *J. Appl. Meteorol. and Climatol.*, **45**, 455–466. (Cité en pages 26, 32, 58, 59, 62, 128, 129, 131, 137 et 188.)
- Viltard, N., Kirstetter, P.-E., and Kacimi, S. (2011). The brain rainfall retrieval algorithm. *Quart. J. Roy. Meteorol. Soc.* (Cité en pages 26, 58, 128 et 130.)
- Wang, J., Hong, Y., Li, L., Gourley, J., Khan, S.-I., Yilmaz, K.-K., Adler, R., Policelli, F., Habib, S., Irwn, D., Limaye, A., Korme, T., and Okello, L. (2011). The coupled routing and excess storage (crest) distributed hydrological model. *Hydrological Sciences Journal*, **56**(1), 84–98. (Cité en page 14.)
- Wentz, F., Ricciardulli, L., Hilburn, K., and Mears, C. (2007). How much more rain will global warming bring? *Science*, **317**(5835), 233–235. (Cité en page 3.)
- WMO (2003). The role of satellites in wmo programmes in the 2010s. World meteorological organisation technical document, WMO Space Programme Satellite Reports. (Cité en page 18.)
- Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., and Liu, C. (2007). A gauge-based analysis of daily precipitation over east asia. *J. Hydrometeor.*, **8**(3), 607–626. (Cité en pages 31 et 56.)
- Xie, P., Joyce, R., Yoo, S.-H., and Yarosh, Y. (2010). A high-resolution gauge-satellite merged global precipitation analysis and its applications for model verifications. In *3rd WMO International Conference on Quantitative Precipitation Estimation and Quantitative Precipitation Forecasting and Hydrology, 18-22 October 2010, Nanjing China*, pages 88–91. World Meteorological Organization. (Cité en page 25.)
- Xu, L., Gao, X., Sorooshian, S., Arkin, P., and Imam, B. (1999). A microwave infrared threshold technique to improve the goes precipitation index. *J. Appl. Meteorol.*, **38**(5), 569–579. (Cité en pages 56, 57, 88, 138 et 140.)

- Yan, J. and Gebremichael, M. (2009). Estimating actual rainfall from satellite rainfall products. *Atm. Research*, **92**(4), 481–488. (Cité en page 57.)
- Yang, S. and Smith, E. (2006). Mechanisms for diurnal variability of global tropical rainfall observed from trmm. *J. Climate*, **19**(20), 5190–5226. (Cité en pages 52 et 99.)
- Zeweldi, D. and Gebremichael, M. (2009a). Evaluation of cmorph precipitation products at fine space–time scales. *J. Hydrometeor.*, **10**(1), 300–307. (Cité en page 56.)
- Zeweldi, D. and Gebremichael, M. (2009b). Sub-daily scale validation of satellite-based high-resolution rainfall products. *Atm. Research*, **92**(4), 427–433. (Cité en page 31.)
- Zipser, E. (2006). Where are the most intense thunderstorms on earth. *Bull. Amer. Meteorol. Soc.*, pages 1057–1071. (Cité en pages 115 et 118.)

Acronymes

AEW : African Easterly Wave
AMMA : African Monsoon Multidisciplinary Analyses
AMMA - CATCH : AMMA Couplage de l'Atmosphère Tropicale et du Cycle Hydrologique
AMMA - EOP : AMMA Enhanced Observing Period
AMSRE : Advanced Microwave Scanning Radiometer-Earth
BRAIN : Bayesian Rain Algorithm Including Neural network
CILSS : Comité Inter-Etats de Lutte contre la Sécheresse dans le Sahel
CMORPH : CPC MORPHing technique
CPC : Climate Prediction Center
CRU : Climate Research Unit
DMSP : Defense Meteorological Satellite Program
EPSAT : Estimation des Précipitations par SATellite
FAR : False Alarm Rate
FEBO : Frequency of Error Bars Overlap
FLORAD : FLOWer constellation of millimeter-wave RADimeters and RADio occultation receivers
GARP : Global Atmospheric Research Program
GATE : GARP Atlantic Tropical Experiment
GHCN : Global Historical Climatology Network
GMI : GPM Microwave Imager
GOMAS : Geostationary Observatory for Atmospheric Sounding
GOES : Geostationary Operational Environmental Satellites
GPCP : Global Precipitation Climatology Project
GPCC : Global Precipitation Climatology Center
GPI : GOES Precipitation Index
GPM : Global Precipitation Mission
GPROF : Goddard PROFiling algorithm
GTS : Global Telecommunication System
GSICS : Global Space-based Inter-Calibration System
GSMaP : Global Satellite Mapping of Precipitation
ICARE : Interactions Clouds Aerosols Radiation Etc
IASI : Infrared Atmospheric Sounding Interferometer
IDL : Interactive Data Language
IPWG : International Precipitation Working Group
IR : InfraRouge
ITCZ : Intertropical Convergence Zone
JMA : Japan Meteorological Agency
MADRAS : Microwave Analysis and Detection of Rain and Atmospheric Structure
MO : Micro-Ondes
MW : Microwaves

NCEP : National Centers for Environmental Prediction

NOAA : National Oceanic and Atmospheric Administration

PERSIANN : Precipitation Estimation from Remote Sensing Information using Artificial Neural Network

PM : Passive Microwave

POD : Probability Of Detection

PDF : Probability Distribution Function

PR : Precipitation Radar

QPE : Quantitative Precipitation Estimation

RFE : Rain Fall Estimation

SATMOS : Service d'Archivage et de Traitement Météorologique des Observations Satellitaires

SEVIRI : Spinning Enhanced Visible and Infrared Image

SO : Système d'Observation

SSMI : Special Sensor Microwave Imager

TAMSAT : Tropical Applications of Meteorology using SATellite data

TAPEER : Tropical Amount of Precipitation with an Estimation of ERrors

TIR : Thermal Infrared

TMI : TRMM Microwave Imager

TMPA : TRMM Multisatellite Precipitation Analysis

TRMM : Tropical Rainfall Measuring Mission

UAGPI : Universally Adjusted GOES Precipitation Index

WAM : West African Monsoon

Table des figures

1.1	Figure du haut : moyennes zonales du rayonnement solaire absorbé (trait plein), et du rayonnement émis vers l'espace (trait pointillé) par le système Terre-Océan-Atmosphère. Ces moyennes sont réalisées à partir des données de flux sortant au sommet de l'atmosphère CERES-EBAF pour la période 2000 à 2010 (Loeb <i>et al.</i> , 2009). La zone orange met en évidence la bande de latitude excédentaire en énergie. Figure du bas : moyennes zonales des précipitations océaniques (trait plein) et de l'évaporation océanique (trait pointillé). Ces moyennes sont réalisées à partir des produits Évaporation et Précipitations HOAPS-3 pour la période 1987 à 2005 (Andersson <i>et al.</i> , 2010). Les zones bleues (resp. beiges) mettent en évidence les bandes de latitude où le bilan [Évaporation - Précipitation] est négatif (resp. positif)	2
1.2	(a) Anomalie des précipitations tropicales en mm/jour (moyenne glissante de 3 mois) pour la période 1979-2010, (b) réponse linéaire des précipitations tropicales à l'ENSO et aux éruptions volcaniques, (c) anomalie des précipitations tropicales une fois isolée la variabilité des pluies liée au phénomène ENSO et aux éruptions volcaniques (Source : C. Kummerow)	4
1.3	Sensibilité linéaire de la fréquence des intensités journalières des précipitations à un changement de température de l'océan pour : des estimations par satellite (noté SSM/I), une moyenne d'ensemble de résultats de simulations de modèles, et une expérience contrôlée par la relation de Clausius-Clapeyron. (Allan <i>et al.</i> , 2010)	5
1.4	Images acquises dans le domaine IR thermique à 06h00-UTC le 9 août 2011 depuis 5 satellites en orbite géostationnaire. De gauche à droite, première ligne : GOES-W (NOAA), GOES-E (NOAA), Meteosat-9 (EUMETSAT), et deuxième ligne : Meteosat-7 (EUMETSAT), MTSAT-1 (JMA). (Source : goes.noaa.gov)	6
1.5	Figure du haut : photographie de l'instrument. Figure du bas : Trace au sol de l'orbite du satellite en noir et fauchées de l'instrument en orange pour une période de 6 h consécutives (calculées grâce au logiciel d'orbitographie IXION (Capderou, 2005, 2009). Instrument : SSMI, Satellite : DMSP-F17.	8
1.6	Figure du haut : vue d'artiste du satellite. Figure du bas : idem que figure 1.5. Instrument : TMI, Satellite : TRMM	9
1.7	Figure du haut : vue d'artiste du satellite. Figure du bas : idem que figure 1.5. Instrument : MADRAS, Satellite : Megha-Tropiques	10
1.8	Évolution des moyens d'observation des précipitations par radiométrie micro-ondes depuis 1990 (Source : d'après http://pmm.nasa.gov/)	11
1.9	Illustration des notions de biais et d'incertitudes d'un système de mesure pour un ensemble de QPE, par analogie avec une cible. La pluie réelle pour chacune des mesures est représentée au centre de la cible (point rouge). Les points noirs représentent des QPE associées. La figure de gauche illustre un système de mesure caractérisé par une incertitude faible mais un biais important. La figure de droite illustre un système de mesure dont le biais est faible mais l'incertitude importante.	13

1.10	Cumuls des précipitations pour la troisième décennie d'août 2005 sur l'Afrique de l'Ouest fournis par 10 produits satellites (Jobard <i>et al.</i> , 2011). La zone encadrée en blanc correspond à la zone du réseau AGRHYMET pour laquelle la comparaison statistique a été effectuée.	15
1.11	Estimation de $\frac{\sigma_{samp}}{R}$ l'amplitude relative de l'écart type observée lors de la comparaison de deux QPE à cause des incertitudes de chaque QPE. L'une des QPE est obtenue à partir du passage d'un SO satellitaire pendant la période d'accumulation et l'autre à partir d'un pluviomètre dans la zone géographique considérée. Les comparaisons sont effectuées pour des durées d'accumulations de 0,005 h à 50 h et pour des surfaces d'intégration spatiales de rayon a variant de 2 à 200 km. (Bell and Kundu, 2003)	16
1.12	Écart relatif entre une corrélation classique r entre deux séries d'estimations de pluie (x_i) et (y_i) et une corrélation ρ prenant en compte les erreurs associées à chaque estimation, en fonction R_x et R_y . S_x est la moyenne des incertitudes associées à chaque estimation de pluie x_i . Idem pour S_y . R_x est le rapport entre S_x^2 et la variance des estimations (x_i). Idem pour R_y . (Kelly, 2007)	17
1.13	Illustration d'une comparaison d'un ensemble de QPE dont chaque estimation est caractérisée par une incertitude importante. La QPE de référence est représentée au centre de la cible (point rouge) avec son incertitude (cercle rouge). Les QPE à valider par rapport à cette référence sont représentées par les points noirs, avec leurs propres incertitudes (cercles noirs).	17
1.14	Schéma illustrant la différence par nature des mesures pluviométriques intégrant la pluie dans le temps et les estimations par satellite intégrant la pluie dans l'espace (North and Nakamoto, 1989)	19
1.15	Photographie d'un dispositif expérimental (à gauche) d'une série de 15 pluviomètres très proches les uns des autres afin d'estimer le terme S^2 pour une QPE obtenue à partir de mesures de pluviomètres. Nuage de points (à droite) montrant la dispersion des mesures des 15 pluviomètres pour un temps d'intégration de 5 minutes, en fonction de l'intensité du cumul moyen mesuré. (Ciach, 2003)	20
1.16	Distributions empiriques d'erreurs d'échantillonnage relatives (ligne du haut) de QPE issues de l'algorithme PERSIANN CCS (Hong <i>et al.</i> , 2004), obtenues par comparaison avec des QPE estimées par radar, pour différentes échelles d'accumulations spatio-temporelles. Distributions modélisées d'erreurs d'échantillonnage relatives (ligne du bas) par optimisation des paramètres a , b , c et d de l'équation 1.6 afin de reproduire les distributions empiriques. Trois classes de distributions sont construites pour des taux de pluie de 0,1 mm/h (colonne de gauche), 2,0 mm/h (colonne du milieu), 4 mm/h (colonne de droite) intégrés sur différentes échelles d'accumulations (d'après : Hong <i>et al.</i> , 2006).	21
1.17	Illustration d'un ensemble de 8 réalisations du champ de pluie TAMSIM au dessus de la Gambie pour le 7 août 2001. CCD : Cold Cloud Duration ; RFE : TAMSAT Rainfall Estimation (estimation déterministe) ; Mean : moyenne de l'ensemble des réalisations. (Teo and Grimes, 2007)	22

1.18	Distributions d'erreurs d'échantillonnage relatives (en %), calculées à partir d'une méthode empirique utilisant des données radars, en fonction des intervalles de temps entre deux mesures allant de 1 heure à 12 heure. Ces distributions sont représentés pour des pluies cumulées sur des périodes de 1mois, 5 jours et 1 jour (colonnes) et pour des zones d'extension géographique de 512 x 512 km à 32 x 32 km (lignes), sur le bassin du Mississipi pour les mois de juin à août (à gauche) et Décembre à Février (à droite). (Gebremichael and Krajewski, 2004)	23
1.19	Validation d'un modèle d'erreur non paramétrique pour l'algorithme CMORPH fournissant une estimation de $S^2_{echantillonnage} + S^2_{algorithmie} + S^2_{calibration}$. Les distributions B1 à B25 correspondent aux probabilités d'intensités de la pluie estimée par radar dans le cas où l'estimation des précipitations de CMORPH est nulle : $P(\text{Pluie Ref} / \text{Pluie CMORPH} = 0)$. Chaque distribution de B1 à B25, correspond à des zones de validation différentes de $2,25^\circ \times 2,25^\circ$ dans le sud des États-Unis, le résultat du modèle (courbe noire) est superposé aux distributions empiriques (histogrammes rouges) pour validation du modèle. Les valeurs des probabilités $P(\text{Pluie Ref} = 0 / \text{Pluie CMORPH} = 0)$ pour le modèle d'erreur (en noir) et les valeurs des probabilités empiriques (en rouge) sont aussi indiquées pour chaque zone géographique. (Gebremichael <i>et al.</i> , 2011)	24
2.1	CILSS area rain gauge network distribution (Ali <i>et al.</i> , 2005a)	34
2.2	Rain gauge distributions of the three dense networks of (a) Niamey, (b) Ouémé, and (c) Dakar. The squares represent the area selected for the kriging interpolation. Black lines are the borders of Niger, Benin and Senegal.	34
2.3	(left) Spatial and (right) temporal variograms calculated over the $5^\circ \times 5^\circ$ window ($9^\circ\text{-}14^\circ\text{N}$, $0^\circ\text{-}5^\circ\text{E}$) for (a),(b) EPSAT-SG; (c),(d) GSMAP; and (e) TMPA. The \times 's correspond to the calculated values of the variogram functions, and the lines correspond to the fitted exponential model.	38
2.4	(a) Box plots of the error of estimation on the gauge and satellite daily estimates over Niamey. Each box shows the 25th and 75th percentiles of the uncertainty distributions. The horizontal line shows the median of the distributions, and the whiskers extend out to largest and smallest values within 1.5 times the interquartile range. (b) Relative errors as a function of rain rate for the gauges (dashed line), EPSAT (solid line), GSMAP (dashed-dotted line), and TMPA (dotted line).	40
2.5	Scatterograms of the 10-day rainfall (a) EPSAT, (b) GSMAP, and (c) TMPA estimates vs CILSS rain gauge estimates (mm.h^{-1}) for the 1-10 Aug 2006 period. One-sigma error bars are over plotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 :1 line. . .	44
2.6	Scatterograms of the daily rainfall (mm.h^{-1}) of (left) EPSAT, (middle) GSMAP, and (right) TMPA estimates over the sites of (a)-(c) Niamey, (d)-(f) Ouémé, and (g)-(i) Dakar. One-sigma error bars are overplotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 :1 line.	47

- 2.7 Scatterograms of the 3-5-day filtered daily rainfall (mm.h^{-1}) of (a) EPSAT, (b) GSMAP, and (c) TMPA estimates vs the 3-5-day filtered daily rainfall gauge estimates over the Ouémé site. One-sigma error bars are over plotted. The thin solid line is the regression line taking the error bars into account, and the dashed line is without the errors. The thick solid line is the 1 :1 line. 48
- 2.8 Composite diurnal cycles of rain rates obtained from Gauges (orange line), EPSAT (black line), GSMAP (red line), and TMPA (blue line) for the (left) preonset and (right) postonset periods for (a),(b) Niamey ; (c),(d) Ouémé , and (e) Dakar. Two-sigma error bars are plotted over the lines. 49
- 3.1 Histograms of relative bias between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with systematic errors on BRAIN medium rain rates, and the reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the region $[20^{\circ}\text{W}- 20^{\circ}\text{E}; 0^{\circ}\text{N}-30^{\circ}\text{N}]$, over land only, and for the June-to-September period of 2006. (a) corresponds to a systematic error on BRAIN medium rain rates of -100%, (b) of -80%, (c) of -60%, (d) of -40%, (e) of -20%, (f) of +100%, (g) of +80%, (h) of +60%, (i) of +40%, (j) of +20%. 64
- 3.2 Time series (a) of the daily rainfall (mm) of TAPEER-BRAIN estimates over the Ouémé site, from 10 June to 20 September 2006, and (b) of the corresponding daily sampling error (mm). (c-h) show the time series sampling errors (black lines) and total errors (coloured lines) where algorithmic errors are computed from the scenarios of systematic errors on BRAIN large rain rates. The black lines are almost covered by the coloured lines in the +20 % and -20% cases. (c) corresponds to a systematic error on BRAIN large rain rates of -100%, (d) of -60%, (e) of -20%, (f) of +20%, (g) of +60%, (h) of +100%. 68
- 3.3 Evolution of validation scores of the daily TAPEER-BRAIN rainfall, over Ouémé and for the June-to-September 2006 period, when algorithmic and sampling errors are considered in the validation process. (Col. 1) corresponds to scores where algorithmic errors are issued from random errors on BRAIN rain rates, (Col. 2) from systematic errors on BRAIN small rain rates, (Col. 3) from systematic errors on BRAIN medium rain rates, (Col. 4) from systematic errors on BRAIN large rain rates. (Row a) shows the TAPEER-BRAIN algorithmic to sampling errors ratio averaged over the season, as a function of magnitude of errors on BRAIN. (Row b) shows the coefficient of determination R^2 , (Row c) the RMS of the regression, (Row d) the slope of the regression, (Row e) the intercept and (Row f) the BIAS of the regression. Where scores are computed from a distribution of scores, the blue bars correspond to the standard deviation of the distribution. The scores computed without errors are over plotted in yellow. 69
- 3.4 Scatterograms of the daily rainfall estimates (mm) of TAPEER-BRAIN ((a) and (c)) and TMPA ((b) and (d)) over the two sites of Niamey ((a) and (b)), and Ouémé ((c) and (d)) for the June-to-September period of 2006. One-sigma error bars are over plotted in blue. The full red lines are the regression lines taking the error bars into account, and the dashed red lines are the regression lines without errors. 71

3.5	Maps of 1-degree rain accumulations for TMPA (a) and TAPEER-BRAIN (b) for the 30th of July 2009 (accumulations smaller than 0.5 mm are set to zero). The corresponding maps of TAPEER-BRAIN sampling errors (c), as well as TAPEER-BRAIN total errors for a best case (d) and a worst case (e), are computed with the e-folding distance maps of the last 10-day period of July 2009.	73
3.6	Space (diamonds) and time (crosses) variograms of TAPEER-BRAIN daily rain accumulations over the $5^\circ \times 5^\circ$ region [$2^\circ\text{-}7^\circ\text{E}$; $12^\circ\text{N-}17^\circ\text{N}$], for the last 10-day period of August 2009. Exponential fits are plotted for the space variogram (black) and for the time variogram (blue), corresponding to d and τ of respectively 63 km and 1.55 h.	77
3.7	Time series of TAPEER-BRAIN space (diamonds) and time (squares) e-folding distances, over the $5^\circ \times 5^\circ$ window [$7^\circ\text{-}12^\circ\text{E}$; $12^\circ\text{N-}17^\circ\text{N}$], for 10-day periods from 10 June 2006 to 20 September 2006. TAPEER-BRAIN daily rain accumulations (in mm) are over plotted in red for the period 10 June to 30 September 2006.	78
3.8	TAPEER-BRAIN rain accumulations for the first 10-day period of July 2009 (a) and associated space (b) and time (c) e-folding distances. Areas in grey correspond to areas where 10-day rain accumulation is smaller than 40 mm. White areas correspond to areas where space or time e-folding distances are greater than the value of the upper limit shown on the colour scale.	79
4.1	Vue d'artiste de la future constellation de satellites, pourvus de radiomètres à micro-ondes passives, issue de la coopération internationale du programme GPM (Source : http://pmm.nasa.gov/)	84
4.2	Nombre moyen de passages par jour (N_{TOT} , courbe noire), pour une durée d'auto-corrélation de la pluie de 30 minutes ($N_{0,5h}$, courbe jaune) et de 1 heure (N_{1h} , courbe rouge), correspondant à deux configurations de SO pour les latitudes de 0° à 30° N : 2 SSMI/DMSP à 00H et 17H LTC (à gauche), 2 SSMI/DMSP à 00H et 00H30 LTC (à droite). Sur la figure de gauche, les 3 courbes (noire, jaune et rouge) sont superposées. Sur la figure de droite, la courbe jaune est superposée à la courbe noire.	85
4.3	Nombre moyen de passages par jour N_{TOT} pour 7 configurations de constellation de SO pour les latitudes de 0° à 30° . L'évolution de N_{TOT} en fonction de la latitude est symétrique par rapport à l'équateur, seul l'échantillonnage pour l'hémisphère Nord est représenté. Les SO considérés sont : TMI/TRMM, AMSRE/AQUA, SSMI/DMSP, GMI/GPM-Core, GMI/GPM-Lio, MADRAS/Megha-Tropiques). La configuration (a) comporte 1 SO, la configuration (g) comporte 7 SO.	86
4.4	Distribution cumulée des intervalles de temps entre deux passages successifs en fonction de la latitude du lieu observé, pour une constellation formée de 4 SO : 2 SSMI/DMSP à 00H et 17H LTC, AMSRE/AQUA à 13h30 LTC et TMI/TRMM. N_{TOT} , $N_{0,5h}$ et N_{1h} sont tracés en fonction de la latitude respectivement en noir, jaune et rouge.	87
4.5	Idem que pour la figure 4.4, pour une constellation de 5 SO dont MADRAS/Megha-Tropiques et les 4 SO de la figure 4.4	87

4.6	Projection au sol de l'orbite de Megha-Tropiques (en rouge) pour 110 minutes d'observations. Les fauchées du radiomètre MADRAS sont représentées en orange ; seulement une fauchée toutes les 2 minutes est représentée au lieu d'une fauchée toutes les 2,5 secondes. Source : IXION (Capderou, 2005)	89
4.7	Figure du haut : champ GSMaP du 1 ^{er} août 2006 à 15H00-UTC. Figure du bas : illustration de la simulation d'un produit de Niveau-2 pour MADRAS/Megha-Tropiques pendant 110 minutes d'observations.	90
4.8	Cartes d'estimations de cumuls de pluie pour la dernière décade du mois août 2006 (colonne de gauche) et d'erreurs d'échantillonnage associées (colonne de droite). La première ligne correspond à la référence GSMaP. Les quatre lignes suivantes correspondent à des estimations TAPEER-GSMaP pour : (i) 1 SO et un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (ii) 1 SO et un volume de $5^\circ \times 5^\circ / 5$ -jours, (iii) 3 SO et un volume de $1^\circ \times 1^\circ / 1$ -jour, (iv) 3 SO et un volume de $5^\circ \times 5^\circ / 5$ -jours.	94
4.9	Distributions cumulées des erreurs TAPEER-GSMaP pour la période de juin à septembre 2006, dans la région "mousson indienne" (terre à gauche, mer à droite) pour trois simulations différentes avec les volumes d'apprentissage : $1^\circ \times 1^\circ / 1$ -jour (rouge), $3^\circ \times 3^\circ / 3$ -jours (vert), $5^\circ \times 5^\circ / 5$ -jours (bleu). Chacune des simulations est réalisée à partir des trois mêmes SO sur des orbites héliosynchrones respectivement à 00H, 03H et 06H.	96
4.10	Scores de comparaison des estimations TAPEER-GSMaP pour les simulations SO/00H (1 ^{ère} et 3 ^{ème} colonnes) et SO/06H (2 ^{ème} et 4 ^{ème} colonnes). Les scores sont calculés pour la boîte "mousson Indienne", pour la terre (1 ^{ère} et 2 ^{ème} colonnes) et pour la mer (3 ^{ème} et 4 ^{ème} colonnes). Les scores $BIAS_{reg}$, RMS_{reg} , POD et FAR sont reportés respectivement sur les 1 ^{ère} , 2 ^{ème} , 3 ^{ème} et 4 ^{ème} lignes. Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure.	97
4.11	Score intégré F calculé à partir des scores de comparaison des estimations TAPEER-GSMaP pour les simulations SO/00H (1 ^{ère} colonne) et pour les simulations SO/06H (2 ^{ème} colonne). Les scores sont calculés pour la boîte "mousson Indienne", pour la mer (1 ^{ère} ligne) et pour la terre (2 ^{ème} ligne). Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure. Le nombre moyen de passages par jour en fonction de la latitude du lieu observé pour chacune des 2 configurations est indiqué (3 ^{ème} ligne). N_{TOT} , $N_{0,5h}$ et N_{1h} sont tracés en noir, jaune et rouge. Les trois courbes (noire, jaune et rouge) sont superposées pour ces deux cas.	98
4.12	Idem que pour la figure 4.11. Les 2 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/03H et SO/06H (1 ^{ère} colonne) ; 5 SO dont SO/00H, SO/03H, SO/06H, SO/09H et SO/13H30 (2 ^{ème} colonne). Les trois courbes (noire, jaune et rouge) sont superposées pour les simulations avec 3 SO. La courbe jaune est superposée avec la courbe noire dans le cas des simulations avec 5 SO.	100

4.13	Cartes d'estimations de cumuls de pluie pour le 2 juillet 2006 (colonne de gauche) et d'erreurs d'échantillonnage associées (colonne de droite). La première ligne correspond à la référence GSMaP. Les quatre lignes suivantes correspondent à des estimations TAPEER-GSMaP pour : (i) SO/00H, SO/03H et SO/06H et un volume d'apprentissage de $1^\circ \times 1^\circ / 1$ -jour, (ii) SO/00H, SO/03H et SO/06H et $5^\circ \times 5^\circ / 5$ -jours, (iii) SO/00H, SO/06H et MADRAS/MT et $1^\circ \times 1^\circ / 1$ -jour, (iv) SO/00H, SO/06H et MADRAS/MT et $5^\circ \times 5^\circ / 5$ -jours.	103
4.14	Idem que pour la figure 4.11. Les 3 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/03H et SO/06H (1 ^{ère} colonne) ; 2 SO dont SO/00H et MADRAS/Megha-Tropiques (2 ^{ième} colonne) ; 3 SO dont SO/00H, SO/06H et MADRAS/Megha-Tropiques (3 ^{ième} colonne). Les trois courbes (noire, jaune et rouge) sont superposées pour les simulations avec 3 SO sur orbite héliosynchrone.	104
4.15	Idem que pour la figure 4.11. Les 2 configurations de constellations simulées sont : 3 SO dont SO/00H, SO/06H et MADRAS/MT (1 ^{ère} colonne) ; 3 SO dont SO/00H, SO/06H et MADRAS/MT* (2 ^{ième} colonne). Dans ces simulations, MT et MT* sont en quadrature de phase. Les trois courbes (noire, jaune et rouge) ne sont pas représentées car elles sont identiques à celles de la figure 4.14	105
4.16	Score intégré F calculé à partir des scores de comparaison des estimations TAPEER-GSMaP pour 2 configurations de constellation : 4 SO dont SO/00H, SO/03H, SO/06H et TMI/TRMM ; 5 SO dont SO/00H, SO/03H, SO/06H, TMI/TRMM et MADRAS/Megha-Tropiques. Les scores sont calculés pour la boîte "mousson Africaine" sur la terre. Les différents volumes d'apprentissage sont indiqués sur l'axe des abscisses (espace) et l'axe des ordonnées (temps) de chaque figure. Le nombre moyen de passages par jour en fonction de la latitude du lieu observé pour chacune des 2 configurations est indiqué (3 ^{ième} ligne). N_{TOT} , $N_{0,5h}$ et N_{1h} sont tracés en noir, jaune et rouge.	107
4.17	Durées de vie des satellites portant des radiomètres à micro-ondes passives à scan conique (D'après Hou <i>et al.</i> , 2011; Huffman <i>et al.</i> , 2010).	109
5.1	Densités géographiques des 1000 cumuls de pluie $1^\circ \times 1^\circ / 1$ -jour les plus forts sur la saison 2009, normalisées par rapport au maximum d'occurrence de ces 1000 cumuls sur l'ensemble des Tropiques sur une grille de $2.5^\circ \times 2.5^\circ$. La première carte correspond à une densité calculée uniquement à partir des cumuls de pluie TAPEER-BRAIN. La deuxième carte correspond à une densité calculée à partir de 500 réalisations de champs TAPEER-BRAIN prenant en compte les erreurs d'échantillonnage avec un degré de confiance de 95%. La troisième carte (resp. quatrième carte) correspond à une densité calculée à partir de 500 réalisations de champs TAPEER-BRAIN prenant en compte les erreurs d'échantillonnage et les erreurs algorithmiques issues d'un biais systématique de BRAIN de 20% (resp. 60%) sur les taux de pluie 2 mm.h^{-1} et 10 mm.h^{-1}	116
5.2	Extraction des quatre cartes de la figure 5.1 sur la zone géographique $[60^\circ \text{E} - 120^\circ \text{E}; 0^\circ \text{N} - 30^\circ \text{N}]$	117
5.3	Extraction des quatre cartes de la figure 5.1 sur la zone géographique $[120^\circ \text{W} - 60^\circ \text{W}; 0^\circ \text{N} - 30^\circ \text{N}]$	117

5.4	Intensités et erreurs associées (degré de confiance de 95%) des 100 cumuls de pluie les plus forts pour deux zones géographiques : [120°W - 110°W ; 10°N - 20°N] en bleu et [85°W - 75°W ; 00°N - 10°N] en rouge	118
A.1	Coefficient de transmission à travers l'atmosphère terrestre en ciel clair en fonction de la fréquence du rayonnement. Source : COMET http://meted.ucar.edu/ , University Corporation for Atmospheric Research	122
A.2	Coefficient d'absorption k_a (2a) et coefficient de diffusion k_s (2b) en km^{-1} , en fonction du taux de pluie, pour l'eau liquide (en trait plein) et pour la glace (en traits pointillés), aux trois fréquences 18,7 GHz, 37 GHz et 86 GHz. (Jobard, 1996)	123
A.3	Émissivité de l'océan pour les polarisations horizontale et verticale en fonction de l'angle d'incidence à 10 GHz et à 60 GHz. (Jobard, 1996)	124
A.4	Processus d'absorption/émission à 37 GHz -V (en haut) et de diffusion à 85 GHz -V (en bas), illustrés par un scan d'AMSRE du Typhon Parma en mer des Philippines en Octobre 2003, superposé avec une image dans le domaine Visible acquise par le satellite GOES-9. Pour le canal 37 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus chaudes. Pour le canal 89 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus froides. Source : COMET http://meted.ucar.edu/	126
A.5	Processus d'absorption/émission à 37 GHz -V (en bas à gauche) et de diffusion à 85 GHz -V (en bas à droite), illustrés par un scan d'AMSRE d'un front dans le Midwest américain en Mai 2005. Pour les 2 canaux 37 GHz -V et 85 GHz -V, les couleurs chaudes correspondent aux températures de brillance micro-ondes les plus chaudes. L'image dans le domaine IR acquise par le satellite GOES montre la position géographique du front (en haut à droite). Source : COMET http://meted.ucar.edu/	127
A.6	Distribution des taux de pluie dans la base de données d'inversions de BRAIN. (Viltard <i>et al.</i> , 2006)	129
A.7	Base de données d'observations et estimations Bayésiennes. Source : IPWG, http://www.isac.cnr.it/~ipwg/	130
A.8	Comparaison des estimations de précipitations BRAIN (noté "Retrieved Surface Rain") avec les estimations de précipitations issues du PR (noté "Reference Surface Rain") sur l'océan (à gauche). Erreurs d'estimations relatives de BRAIN en fonction des taux de pluie. (Viltard <i>et al.</i> , 2006)	131
A.9	Relation T_B à 10 GHz, taux de pluie pour différents degrés d'inhomogénéités de champs de pluie à moyenne spatiale constante. (Kummerow, 1998)	132
A.10	Représentation de la taille relative des pixels de différents radiomètres micro-ondes par canal. Source : N. Viltard	133
A.11	Exemples de restitutions de taux de pluie par BRAIN grâce aux instruments TMI, SSM/I et AMSRE pendant le cycle de vie du cyclone Gaëlle à son passage sur l'île de la Réunion du 4 au 6 février 2009. Source : N. Viltard	134

A.12	Histogrammes 2D des estimations de pluie BRAIN pour les radiomètres SSMI (colonne de gauche) TMI (colonne du milieu) et AMSRE (colonne de droite) géo localisées avec les données IR du satellite MSG-2 pour la période juin à septembre 2009, pour la zone $[40^\circ \text{W} - 40^\circ \text{E}; 30^\circ \text{S} - 30^\circ \text{N}]$, sur continent et sur océan	134
B.1	General diagram	137
B.2	Time series (first column) of Infrared Brightness temperatures (in Kelvin, grey colorbar) images for the 8th of September 2006 over the region of Niamey, Niger (the red and orange squares correspond to the 1° and 5° areas close to the rain gauge network of Niamey). Time series (second column) of Infrared Brightness temperatures images during the period ± 2 -days off the 8th of September 2006, selected when the BRAIN algorithm, from a constellation of passive MW imagers, provides rain estimations (in mm.h^{-1} , rainbow colorbar)	139
B.3	Distributions of Infrared pixels (top left) and MW-derived rain rates (bottom left). Cumulated distributions of IR pixels (top right) and MW-derived rain rates (bottom right) for a $5^\circ \times 5^\circ \times 5$ -days volume (Kidd <i>et al.</i> , 2003)	140
B.4	$T_{\text{threshold}}$ (middle) and R_{COND} (right) used to compute a TAPEER $1^\circ/1$ -day precipitation map (left)	142
B.5	Time series of TAPEER-BRAIN Indicator fields for the region $[7^\circ \text{E} - 12^\circ \text{E}; 12^\circ \text{N} - 17^\circ \text{N}]$. , The two red lines, positioned at \vec{x}_0 \vec{x}_1 show the time series over which two time variograms $\gamma(\vec{x}_0, \Delta t)$ and $\gamma(\vec{x}_1, \Delta t)$ can be computed.	144
B.6	TAPEER flow chart	148
B.7	TAPEER-BRAIN accumulated rainfall (top) and sampling error (bottom) for the 30th of July 2009	149
B.8	Relative sampling errors (%) as a function of TAPEER-BRAIN rain accumulations for the June-to-September 2009 period, over land (left) and over ocean (right).	149
B.9	Comparison of TAPEER-BRAIN daily estimates and TMPA daily estimates over oceans of the whole Tropical belt for the June-to-September 2009 period	151
B.10	Comparison of TAPEER-BRAIN daily estimates and TMPA daily estimates over continents of the whole Tropical belt for the June-to-September 2009 period	152
B.11	Scatterograms of the 10-day rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the 10-day rainfall gauge estimates over the Sahelian band for the second 10-day period of June 2006. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 :1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.	155
B.12	Same as figure B.11 for the first 10-period of August 2006.	156

B.13 Scatterograms of the daily rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the daily rainfall gauge estimates over the Niamey site. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 :1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.	157
B.14 Same as figure B.13 for the Ouémé site.	158
B.15 Scatterograms of the 3-5 day filtered daily rainfall (mm) of TMPA (top left), TMPA-RT (top right), GSMAP (bottom right) and TAPEER-BRAIN (bottom left) estimates vs the daily rainfall gauge estimates over the Niamey site. One-sigma error bars are over plotted. The red solid line is the regression line taking the error bars into account, and the green solid line is without the errors. The black dashed line is the 1 :1 line. Confidence bands at 99% for the regressions line taking the error bars into account are over plotted in orange.	159
B.16 Same as figure B.15 for the Ouémé site.	160

Liste des tableaux

2.1	Characteristics of the rain gauge networks.	33
2.2	Mean percentages of errors for the seasonal diurnal, 10-day, 1-day, and 3-hourly time scales for the gauge and satellite product estimates.	40
2.3	Ratio (%) of the mean error to the variance of the precipitation estimates. For Dakar, the median of the error distributions of the daily estimates is computed rather than the mean error because the latter does not represent the distribution of errors well as a result of the small sample size.	41
2.4	Statistics of the regressions between the three satellite products and the gauge estimates (i) for the 10-day period (11-20 Jun 2006) and (ii) for the 10-day period (1-10 Aug 2006). Sample size = 137.	44
2.5	Statistics of the regressions between the three satellite products and the gauge daily rain estimates and between the three satellite products and the gauge filtered (3-5 days) daily rain estimates over Niamey, Ouémé, and Dakar.	45
2.6	Percentage of variance explained by the subdaily variability of the gauge rain rate and of the three rainfall satellite products during the postonset and preonset periods for the sites of Niamey, Ouémé, and Dakar. The variance explained is calculated as $1-R^2$, where R is the correlation between the 1-day low-pass filtered 3-hourly time series and the 3-hourly time series. Low values ($<0.1 \text{ mm.h}^{-1}$) are filtered out, yielding to a varying population to perform the computations. * Not significant at the 99.9% level.	50
2.7	FEBO and FEBO unbiased indices calculated with the mean diurnal cycles of EPSAT, GSMAP, and TMPA with 2-sigma errors.	51
2.8	Percentage of variance explained by the mean conditional diurnal cycle of the Gauge rain rate and of the three rainfall satellite products, during the post-onset and pre-onset periods, for the sites of Niamey, Ouémé and Dakar. The variance explained is calculated as $1-R^2$, R being the correlation between the 3-hourly time series anomalies to the mean conditional diurnal cycle and the 3-hourly time series. * Not significant at the 99.9% level.	51
3.1	Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with a random error on BRAIN rain rates, and the reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the area [20°W- 20°E; 0°N-30°N], over land only, and for the June-to-September period of 2006.	65
3.2	Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with systematic errors on BRAIN medium rain rates, and reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the same area and period as Table 3.1	65
3.3	Statistical moments and number of samples of relative bias distributions between TAPEER-BRAIN 1-degree/1-day rain accumulations, generated with white noises on TIR brightness temperatures, and reference TAPEER-BRAIN 1-degree/1-day rain accumulations, for the same area and period as Table 3.1.	66

3.4	Validation scores of the two satellite products TAPEER-BRAIN and TMPA daily rain estimates over the two sites of Niamey and Ouémé, for the June-to-September period of 2006. Sampling errors are taken into account for both satellite and gauge daily rain estimates in the validation process.	72
4.1	Estimations de cumuls de pluie sur 10 jours pendant la dernière décade d'août 2006, pour la zone géographique [86° E - 88° E ; 18° N - 20° N] (voir figure 4.8) composée de quatre zones de 1°×1°, par la référence GSMaP et par les estimations TAPEER-GSMaP issues de 4 combinaisons différentes de constellations et de volume d'apprentissage : (i) un seul SO et un volume d'apprentissage de 1°×1°/1-jour, (ii) un seul SO et un volume d'apprentissage de 5°×5°/5-jours, (iii) 3 SO et un volume d'apprentissage de 1°×1°/1-jour, (iv) 3 SO et un volume d'apprentissage de 5°×5°/5-jours. Le symbole ✓ (resp. ✗) indique qu'il y a recouvrement (resp. il n'y a pas recouvrement) entre les intervalles de confiance à 68% des estimations TAPEER-GSMaP et de la référence.	93
4.2	Scores de comparaison entre la référence GSMaP et les estimations TAPEER-GSMaP de cumuls sur 10 jours pendant la dernière décade d'août 2006, pour la zone géographique [70° E ; 100° E - 5° N ; 25° N] (figure 4.8). Les simulations TAPEER-GSMaP évaluées correspondent à quatre combinaisons différentes de constellations et de volumes d'apprentissage : (i) un seul SO et un volume d'apprentissage de 1°×1°/1-jour, (ii) un seul SO et un volume d'apprentissage de 5°×5°/5-jours, (iii) 3 SO et un volume d'apprentissage de 1°×1°/1-jour, (iv) 3 SO et un volume d'apprentissage de 5°×5°/5-jours.	95
4.3	Idem que pour le tableau 4.1, pour la journée du 2 juillet 2006 (voir figure 4.13). Les quatre simulations effectuées utilisent les configurations suivantes : 3 SO sur orbite héliosynchrone et un volume d'apprentissage de 1°×1°/1-jour, 3 SO sur orbite héliosynchrone et un volume d'apprentissage de 5°×5°/5-jours, 3 SO dont 2 sur orbite héliosynchrone et MADRAS/Megha-Tropiques et un volume d'apprentissage de 1°×1°/1-jour, 3 SO dont 2 sur orbite héliosynchrone et MADRAS/Megha-Tropiques et un volume d'apprentissage de 5°×5°/5-jours.	102
B.1	Thermal Infrared Brightness Temperatures derived from geosynchronous Infrared Imagers.	146
B.2	Instantaneous precipitation products derived from the microwave brightness temperatures acquired by various observing systems.	146
B.3	Good Line Rate (GLR) statistics of geostationary IR data for the 2009 June-to-September season	147
B.4	Statistical scores of comparison between 10-day rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 10-day rainfall gauge estimates over the Sahelian band for the second 10-period of June 2006	153
B.5	Statistical scores of comparison between 10-day rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 10-day rainfall gauge estimates over the Sahelian band for the first 10-period of August 2006	153
B.6	Statistical scores of comparison between daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and daily rainfall gauge estimates over Niamey	153

B.7	Statistical scores of comparison between daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and daily rainfall gauge estimates over Ouémé	154
B.8	Statistical scores of comparison between 3-5 day filtered daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 3-5 day filtered daily rainfall gauge estimates over Niamey	154
B.9	Statistical scores of comparison between 3-5 day filtered daily rainfall of GSMAP, TMPA-RT, TMPA, TAPEER-BRAIN and 3-5 day filtered daily rainfall gauge estimates over Ouémé	154