





ÉVALUATION DES PRÉVISIONS DE PLUIE DU PROTOTYPE AROME OUTRE-MER

Rapport de stage IENM2 Projet de Modélisation

Coralie VOIRIN, Marie PORTÉ, Alexis CASTELEIN Encadrants : Philippe CHAMBON, Ghislain FAURE Date du stage : du 6 janvier 2014 au 14 février 2014 Centre National de Recherches Météorologiques (Groupe GMAP) Météo-France, Toulouse



IENM 2012-2015

Remerciements

Nos premiers et plus sincères remerciements vont à nos encadrants, Philippe et Ghislain, pour l'énorme travail qu'ils ont fourni pour préparer et accompagner le nôtre. C'était une expérience vraiment formidable. Pour votre disponibilité sans faille ni condition, pour vos encouragements constants, pour votre bonne humeur, pour votre aide et vos conseils avisés, mais aussi pour nous avoir laissé nous débrouiller tous seuls lorsque nous pouvions en apprendre quelque chose, merci beaucoup. Nous ressortons grandis de cette expérience, et sans vous ça n'aurait peut-être pas été aussi évident. Et parce que, vous aussi, vous en avez sans doute eu besoin durant ces cinq semaines de stage, à notre tour de vous souhaiter : « Bon courage! ».

Nous remercions également Michaël Zamo pour son aide en matière de score, et Alexandre Mary pour ses conseils en python. Enfin, merci à toute l'équipe du GMAP de nous avoir si bien accueillis ainsi qu'aux habitants de la Tchétchénie.

Résumé

La prévision des précipitations est un enjeu majeur dans les régions tropicales soumises fréquemment à de violents phénomènes météorologiques. L'amélioration continue des modèles de prévision numérique du temps y est primordiale et doit être validée par des scores objectifs. Pour cela, des observations utilisées en tant que référence sont nécessaires. Le domaine des Tropiques étant principalement océanique, il est important d'utiliser une alternative aux intruments de mesure traditionnels répartis de manière très éparse sur les terres émergées, comme les pluviomètres et les radars. Dans ce contexte, les produits satellitaires d'estimation de pluie sont une alternative intéressante pour la vérification des performances des modèles de prévision sur l'ensemble de leur domaine.

Un jalon important pour la prévision numérique du temps sur l'Outre-Mer à Météo-France est de déployer le modèle AROME, à 2,5 km de résolution, sur les zones où ALA-DIN est actuellement opérationnel, d'ici 2016.

Le projet consiste à étudier l'apport de ce modèle de méso-échelle pour la prévision des précipitations par rapport à ALADIN, en utilisant plusieurs produits d'estimation des précipitations par satellite comme référence.

Pour cela, une technique de comparaison avec tolérance spatiale a été mise en œuvre. Elle consiste à utiliser le « Fractions Skill Score », et permet de confronter les observations et les prévisions sur un voisinage et pour différents seuils d'intensité de pluie. L'originalité de la démarche utilisée consiste à prendre en compte les erreurs de la référence, en calculant les scores sur un ensemble d'observations perturbées.

Avant tout, la robustesse du score est validée via différents tests de sensibilité portant sur le choix du produit satellitaire, sur les hypothèses faites sur la distribution des erreurs du produit satellitaire et sur le nombre de perturbations. Ces tests ont rendu légitime d'utiliser par la suite ce score avec des observations perturbées 1000 fois suivant une loi normale à intervalle de confiance de $\pm \sigma$ où σ représente l'écart-type d'erreur.

Des comparaisons entre le futur modèle opérationnel AROME (initialisé avec IFS) et ALADIN puis entre AROME et ALADIN à conditions initiales égales (celles d'ALADIN) ont ensuite pu être réalisées sur des cumuls de pluie 6h grâce à la mise en œuvre de ce score. Les résultats montrent la capacité d'AROME à prévoir des cumuls supérieurs à 50 mm/6h, et ce, dès les six premières heures d'échéance. Pour les seuils de pluie inférieurs, AROME obtient de meilleurs scores à mesure que les échéances augmentent.

Afin d'évaluer l'impact des conditions initiales sur le modèle AROME, une nouvelle comparaison est réalisée entre AROME initialisé avec IFS et AROME initialisé avec ALA-DIN en utilisant GSMaP comme référence. Les résultats ne permettent pas de privilégier un choix d'initialisation puisque aucun score ne ressort de manière nette. L'utilisation de TMPA pour cette même comparaison montre une forte sensibilité au produit satellitaire.

Cette étude participe finalement à la justification de la mise en place du prototype AROME Outre-Mer en remplacement d'ALADIN.

Table des matières

Ren	iercien	nents	2
Rés	umé		4
Intr	oductio	on	8
1.	Donné	ées de référence et de prévision	10
	1.1.	Produits satellitaires 1	10
		1.1.1. TMPA	10
		1.1.2. GSMaP	13
		1.1.3. Erreurs associées	14
	1.2.	Modèles de prévision numérique	15
		1.2.1. AROME	15
		1.2.2. ALADIN	16
	1.3.	Situations météorologiques exploitées	Ι7
2.	Métho	odes de vérification des prévisions	22
	2.1.	Score point à point	22
	2.2.	Vérification avec tolérance	22
		2.2.1. Fractions Skill Score	23
		2.2.2. Exemple simplifié	24
	2.3.	Prise en compte des erreurs de la référence	26
3.	Mise e	en place d'un score et tests de sensibilité	30
	3.1.	Mise en place d'un score	30
	3.2.	Sensibilité au nombre de perturbations	31
	3.3.	Sensibilité à la prise en compte des erreurs	32
	3.4.	Sensibilité au produit satellitaire de référence	33
4.	Mise e	en œuvre du score	36
	4.1.	Comparaison AROME_IFS et ALADIN	36
	4.2.	Comparaison AROME_ALA et ALADIN	37
	4.3.	Comparaison des deux AROME	38
Con	clusion	n et perspectives	12
Réfé	erences	;	14
Acro	onymes	s	16

Introduction

L'estimation des précipitations au sol est un enjeu majeur dans le domaine hydro-météorologique, en particulier dans les régions tropicales où les tempêtes et cyclones causent d'importants dégâts humains et matériels. De plus, les précipitations font partie des phénomènes les plus difficiles à prévoir en raison de leurs grandes variabilités spatiale et temporelle, et de la complexité à modéliser la convection. Afin d'améliorer les modèles dans ce domaine, il est nécessaire de documenter leurs points forts et faibles, notamment par comparaison à une référence basée sur des observations.

Cette étude s'intéresse aux prévisions de précipitations dans les Tropiques, et en particulier dans les Petites Antilles. Pour cette zone principalement océanique, il est difficile d'obtenir des observations de référence car les instruments de mesure traditionnels des précipitations, tels que les pluviomètres ou les radars, ne peuvent être déployés que très localement sur les terres émergées. Les produits d'estimation de pluies satellitaires sont alors une référence alternative à étudier. En effet, ces produits, développés depuis une trentaine d'années, arrivent maintenant à un degré de maturité intéressant pour la vérification de modèles de prévision numérique du temps.

À terme, il est prévu de déployer le modèle AROME sur l'Outre-Mer où ALADIN est actuellement opérationnel. Du fait de sa résolution plus fine et de sa convection explicite, ce modèle est reconnu pour mieux représenter les fortes précipitations ayant lieu notamment dans le sud-est de la France métropolitaine. Il pourrait donc permettre une meilleure prévision des phénomènes pluvieux intenses, fréquents dans les régions tropicales.

L'objectif de cette étude est d'évaluer l'apport de ce modèle de méso-échelle pour la prévision des précipitations par rapport à ALADIN, en utilisant plusieurs produits d'estimation des précipitations par satellite comme référence, tout en prenant en compte leurs incertitudes.

Pour cela, il est possible d'utiliser une technique classique de vérification des pluies prévues, qui compare point à point l'observation et la prévision. Cependant cette méthode peut être insuffisante, c'est pourquoi la technique utilisée dans cette étude correspond à ce qu'on appelle le « Fractions Skill Score ». Il s'agit d'un score dit avec tolérance spatiale qui considère une notion de voisinage autour des observations (Ebert [2008]). De plus, pour prendre en compte les incertitudes des produits satellitaires, des perturbations sont appliquées à l'observation de référence.

Après avoir décrit les différents types de données et scores de vérification des prévisions, un score avec tolérance spatiale est mis en place et sa robustesse est vérifiée via différents tests de sensibilité. Les modèles ALADIN et AROME sont ensuite comparés grâce à ce score sur trois situations à enjeu autour de l'arc antillais de fin juillet à fin août 2012. Enfin, deux versions d'AROME aux conditions initiales différentes sont comparées.

1. Données de référence et de prévision

1.1. Produits satellitaires

L'estimation des précipitations sur l'ensemble du globe, à l'échelle des pluies instantanées et des pluies cumulées, est possible à travers plusieurs familles de produits alliant des mesures satellitaires dans trois domaines de fréquence : les micro-ondes, l'infra-rouge et le visible.

Dans cette étude, on s'intéressera uniquement aux produits combinant infra-rouges et micro-ondes, car des campagnes de validation ont montré que ces produits donnaient les meilleures estimations à l'échelle des pluies cumulées (série des *Precipitation Inter-comparison Project* et des *Algorithm Inter-comparison Programm* (Barrett [1994]; Ebert [1996]; Smith [1998]; Adler [2001])). Parmi ces produits, nous utiliserons TMPA et GS-MaP_MVK, actuellement considérés comme faisant partie des meilleurs produits opérationnels.

Une étape d'inversion des températures de brillance issues des images infra-rouges et micro-ondes est nécessaire pour obtenir, de manière indirecte, des données de précipitations. Différentes méthodologies existent pour réaliser cette étape (Kidd [2011]). Une première consiste à utiliser les données satellites infra-rouges. En effet, les sommets des nuages les plus froids apparaissent en blanc sur les images infra-rouges (exemple sur la figure 1) et représentent potentiellement des nuages à forte extension verticale. Dans cette méthode, ces derniers sont supposés être liés à des précipitations fortes (Arkin [1987]). Une seconde méthode, plus précise, se base sur l'émission et la diffusion des gouttelettes d'eau et des cristaux de glace dans les micro-ondes.

Cette dernière approche ne permet pas d'obtenir des cumuls de pluie sans subir une importante erreur d'échantillonnage; en effet, même avec une constellation de satellites défilants pourvus de radiomètres micro-ondes (comme l'exemple de la figure 3), l'échantillonnage reste faible par rapport à la variabilité spatio-temporelle de la pluie (figure 2). Or, les images infra-rouges n'ont pas ce défaut, car elles sont fournies par les satellites géostationnaires à intervalle de temps régulier (par exemple, toutes les 15 minutes pour Meteosat Seconde Génération) et sur tout le globe. Chacun des deux produits décrits ici utilise donc ces images pour compléter les intervalles dépourvus de données micro-ondes, chacun à sa propre manière (figure 2).

1.1.1. TMPA

TMPA est l'abréviation de *TRMM Multisatellite Precipitation Analysis*. TRMM, signifiant *Tropical Rainfall Measuring Mission*, est un satellite américano-japonais lancé en 1997 et dédié à l'estimation des précipitations tropicales grâce à son radiomètre microondes TMI et son radar de précipitations. La description détaillée de TMPA est présentée dans l'article Huffman [2007]. TMPA est en fait constitué de deux produits différents : un produit disponible 9h après le temps réel et un produit de recherche de meilleure qualité mais disponible seulement 10 à 15 jours après la fin du mois. Ce produit fournit des estimations de précipitations à la résolution de 0,25° avec une fréquence temporelle de 3h.

Deux algorithmes sont utilisés actuellement pour l'inversion des radiances micro-ondes observées en taux de précipitations : l'algorithme GPROF pour les instruments TMI,



FIGURE 1 – Image infra-rouge du satellite GOES13 le 22 août 2012 à 03h00UTC pour le canal 10,7 micromètres. Source : SATMOS http ://www.satmos.meteo.fr/

AMSR2 et SSMI/S à bord des satellites TRMM, GCOM-W1 et DMSPs ainsi que l'algorithme de Weng pour les instruments AMSU-B et MHS à bord des satellites MetOp et des satellites de la NOAA. Les estimations de pluie grâce à TMPA sont produites en quatre étapes :

- estimation des précipitations à partir des images micro-ondes;
- estimation des précipitations, à partir des images infra-rouges, interprétées grâce aux mesures micro-ondes passées;
- mise en commun de ces estimations;
- ajout des mesures au sol (seulement fait avec le produit recherche).

Pour la première étape, les données acquises avec les radiomètres micro-ondes passifs des satellites défilants sont converties en estimations de précipitations sur une grille spatiale de $0.25 \times 0.25^{\circ}$ et sur une échelle de temps de \pm 90 minutes toutes les 3h. Ces inversions sont réalisées à partir de précédentes estimations sur des cas similaires grâce au radar de précipitations à bord de TRMM. S'il y a eu plusieurs passages de satellites défilants pour le même point de grille 0.25° , la moyenne des taux de pluie est considérée.

La figure 2 montre un exemple de la couverture géographique obtenue en regroupant les estimations de pluie réalisées à partir des observations d'une constellation de satellites défilants, pour une période de trois heures. Pour les zones non-couvertes en trois heures,



FIGURE 2 – Image de la constellation des satellites défilants ayant à leur bord un radiomètre micro-ondes. Source NASA http://pmm.nasa.gov/



FIGURE 3 – Estimation de précipitations micro-ondes combinées sur une période de 3h centrée sur 00h00UTC le 25/05/04 en mm/h. Les régions noires indiquent les zones non observées par un instrument micro-ondes sur une période de 3h. Source : Huffman [2007]

une estimation de pluie est réalisée à partir d'observations infra-rouges des satellites géo-

stationnaires. Ces estimations se basent sur une relation statistique établie localement à partir d'observations infra-rouges géolocalisées et micro-ondes. Afin de construire cette relation statistique locale, les données colocalisées sont collectées sur une période d'un mois.

La dernière étape n'entre que dans le cadre du produit recherche mais pas dans celui du temps réel, pour des raisons de disponibilité des données.

1.1.2. GSMaP

Le second produit satellitaire est appelé GSMaP_MVK (abrégé GSMaP par la suite), et correspond au produit avec vecteur de mouvement et filtre de Kalman (Motion Vector with Kalman filter - MVK) développé par le groupe japonais GSMaP (Global Satellite Mapping of Precipitations). Une description plus complète de ce produit se trouve dans l'article Ushio [2009].

Tout comme TMPA, ce produit se base tout d'abord sur les images micro-ondes issues des satellites défilants pour évaluer les taux de pluie instantanés, grâce à un algorithme spécifique décrit dans les articles Joyce [2004] et Aonashi [2000]. De plus, il tient compte des températures de brillance issues des images infra-rouges pour améliorer cette estimation, comme TMPA. Sa spécificité tient au filtre de Kalman appliqué lors de cette fusion, qui permet d'estimer les précipitations en dehors des zones couvertes par les satellites défilants. Cela sert dans ce cas à propager les zones pluvieuses identifiées lors du dernier passage d'un satellite défilant.



FIGURE 4 – Illustration de la méthode utilisée par GSMaP : Propagation vers l'avant, puis rétro-propagation, et moyenne pour obtenir l'estimation finale. Source : Ushio [2009]

La figure 4 illustre le principe suivant : la propagation repose sur la comparaison de deux images infra-rouges successives. Si les éléments de l'image à t=1 sont décalés par rapport à leur position en t=0 en latitude et longitude, le champ des vecteurs d'advection

est déduit de ce décalage. Ce champ est ensuite appliqué aux zones pluvieuses identifiées lors du passage du satellite défilant.

De plus, ces zones sont rétro-propagées spatialement avec le même vecteur d'advection. L'estimation optimale des intensités pluvieuses est obtenue par moyenne pondérée entre les propagations avant et arrière. Cela permet de propager les zones pluvieuses de manière réaliste, en se servant uniquement des images infra-rouges.

L'inconvénient majeur de ce produit est la difficulté à décrire l'initiation des précipitations entre les passages du radiomètre micro-ondes. En effet, l'algorithme actuel propage et applique le filtre de Kalman uniquement sur les pixels de pluie repérés sur la dernière image micro-ondes. Ainsi, une nouvelle zone pluvieuse n'apparaît qu'après le passage du radiomètre micro-ondes. Cela explique une partie des erreurs d'estimation de ce produit.

En conclusion, ces deux produits d'estimation des précipitations par satellite présentent encore des incertitudes non négligeables, qu'il convient d'étudier. Dans la suite de cette étude, on s'intéresse aux pluies cumulées à l'échelle $0.5^{\circ}/6h$ car cette résolution permet de projeter les produits TMPA (0.25° de résolution native) et GSMaP (0.1° de résolution native) sur la même grille géographique.

1.1.3. Erreurs associées

Comme indiqué précédemment, et malgré leur nette amélioration depuis les premières générations de produits, il reste important de caractériser les erreurs des produits satellitaires. Par exemple, un groupe de travail international sur les précipitations (IPWG) s'attache à documenter quotidiennement les erreurs de plusieurs produits à partir de divers réseaux d'observations aux Etats-Unis, en Australie, en Europe, au Japon, en Amérique du Sud et en Afrique du sud¹.

Comme tout système de mesure, si l'on suppose les différentes sources d'erreur non corrélées entre elles, il est possible d'écrire un bilan d'erreur comme proposé ci-dessous, avec un terme lié à la calibration de l'instrument, un autre lié à l'algorithme d'inversion utilisé, et enfin un terme lié à l'échantillonnage :

$E^2 = E^2 calibration + E^2 algorithme + E^2 echantillonnage$

La distribution de ces erreurs est difficile à caractériser du fait de leurs variabilités géographique et saisonnière. Néanmoins, leur écart-type peut être modélisé. Un modèle d'erreur pour le terme d'échantillonnage a par exemple été proposé par Roca [2010] et est utilisé dans cette étude pour les deux produits satellitaires TMPA et GSMaP. Des études ont également été menées sur les termes d'erreurs algorithmiques et de calibration dont Chambon [2013] mais ne sont pas disponibles pour TMPA et GSMaP.

Ainsi, la figure 5 montre la distribution des erreurs d'échantillonnage associées à chaque point de mesure du cumul sur 6h à résolution 0.5°, pour chacun des produits et sur toute la période étudiée (fin juillet à fin août 2012). La plage de couleurs montre le logarithme du nombre de points de mesure présentant l'erreur relative associée à chaque cumul de précipitations. On remarque que les erreurs sont nettement plus grandes pour les plus faibles accumulations de pluie, sachant que les erreurs supérieures à 120% ont volontairement été tronquées. En revanche, pour les cumuls de plus de 20mm, les erreurs

^{1.} http://www.isac.cnr.it/~ipwg/



FIGURE 5 – Erreurs relatives en fonction du cumul de pluie sur 6h pour GSMaP (à gauche) et pour TMPA (à droite)

diminuent fortement, ceci étant plus marqué pour le produit GSMaP. De plus, les erreurs d'échantillonnage issues de TMPA comprises entre 50 et 100% sont plus fréquentes aux plus faibles cumuls que pour GSMaP, et ce, malgré une plus grande dispersion de leur distribution. Par exemple, pour un cumul estimé de 20mm, les erreurs GSMaP varient de 10 à 60% tandis que celles de TMPA balayent une gamme de 5 à 100% environ.

Ceci s'explique par le fait qu'on a ici caractérisé une erreur d'échantillonnage et que le produit GSMaP possède une meilleure résolution native que TMPA.

En résumé, les produits d'estimations satellitaires de pluies sont entachées d'erreurs qu'il est important de prendre en compte lors de l'évaluation d'un modèle de prévision numérique du temps.

1.2. Modèles de prévision numérique

Météo-France dispose de quatre modèles ALADIN pour les besoins de la prévision sur l'Outre-Mer. Il est prévu de les remplacer d'ici 2016 par des modèles AROME. Ces modèles sont décrits ci-après et leurs domaines respectifs sont indiqués sur la figure 6.

1.2.1. AROME

Après six ans de développement, AROME est déclaré opérationnel fin 2008 sur la métropole. Ses principales caractéristiques sont son noyau dynamique non-hydrostatique, sa physique humide spécifique à la méso-échelle, et le schéma variationnel d'assimilation associé (3D-Var). Sa résolution est de 2.5km, la convection est alors partiellement résolue (convection profonde). La convection sous-maille est paramétrée. Le modèle est entièrement décrit dans l'article Seity [2011].

Les bénéfices d'AROME ont été prouvés par des scores objectifs et par l'utilisation en conditions réelles sur la métropole. Il est meilleur que ses prédécesseurs pour le brouillard et les précipitations, grâce à sa physique plus réaliste et à son assimilation de données radar. En particulier, AROME améliore les prévisions des précipitations convectives méditerranéennes intenses et les structures précipitantes forcées par l'orographie sur la métropole (Bouttier [2007]). Les précipitations étant un enjeu majeur de la météorologie



FIGURE 6 – Illustration des domaines ALADIN Outre-Mer (domaines entourés en noir) et des futurs domaines AROME (domaines entourés en rouge). Les points verts représentent la climatologie des cyclones tropicaux et les territoires français sont en rouge. Source : GMAP

tropicale, ce modèle mérite d'être étudié pour ces zones.

C'est pourquoi des travaux sont en cours pour remplacer ALADIN, actuellement opérationnel sur les zones Outre-Mer, par ce modèle. Dans le cadre des expériences AROME sur l'Outre-Mer, l'assimilation de données n'est pas encore utilisée; les prévisions sont faites à partir d'une analyse fournie soit par le modèle du centre européen IFS (Integrated Forecast System) soit par ALADIN. On appellera par la suite AROME_IFS la version d'AROME qui prend ses conditions initiales et aux limites du modèle IFS, et AROME_ALA celle qui les prend du modèle ALADIN.

1.2.2. ALADIN

Les ALADIN Outre-Mer sont opérationnels depuis 2010 sur les zones Antilles-Guyane, Polynésie Française et Nouvelle-Calédonie et depuis 2006 sur l'océan Indien. Ils sont utilisés à la résolution de 8 kilomètres avec 70 niveaux verticaux.

Le temps sensible sur les territoires de l'Outre-Mer est complexe à prévoir, en raison du relief (supérieur à 3000m sur la Réunion par exemple), du régime des alizés, du manque d'observations et de la présence de phénomènes violents, qui réduisent également la prévisibilité. Le relief, conjugué à des conditions cycloniques, peut provoquer d'importants cumuls de pluies dans le cadre d'épisodes pluvieux de fortes intensités sur des périodes allant de 12 heures à 15 jours. Enfin, la particularité des domaines Outre-Mer par rapport à la métropole est leur caractère essentiellement maritime (souvent plus de 80%) visible sur la figure 6.

Les caractéristiques d'ALADIN sont les suivantes : il possède son propre système d'assimilation de données (3D-Var) et utilise comme modèle coupleur ARPEGE pour ses conditions aux limites à la Réunion et IFS pour les autres domaines. Une description plus détaillée d'ALADIN-Réunion est faite dans Faure [2008].

La résolution d'ALADIN lui permet de simuler de façon assez réaliste la structure des cyclones (gradient, intensité) par rapport aux modèles globaux. Le principal défaut d'ALADIN provient de ses prévisions des précipitations, en particulier en ce qui concerne leur localisation et leur intensité (sous-estimées en cas d'advection de précipitations sur le relief lors d'alizés humides, et surestimées en cas de précipitations convectives sous le vent), même si on note un apport par rapport aux modèles de résolution plus faible.

1.3.Situations météorologiques exploitées

L'étude se focalise sur trois situations météorologiques, de fin juillet à fin août 2012, sur une zone centrée autour des Petites Antilles (11°N à 22°N et 52°W à 67°W) : une dépression tropicale fin juillet et deux tempêtes tropicales début et fin août (nommées respectivement Ernesto et Isaac). Les images infra-rouges issues du satellite GOES13 illustrent chacun des cas étudiés (figure 7). Le sommet des nuages les plus hauts, donc plus froids, est caractérisé par de plus faibles valeurs de température de brillance.



FIGURE 7 – Images infra-rouges issues du satellite GOES13 avec le canal 10,7 micromètres pour les dates du 30 juillet 2012 à 09h45 UTC (a), du 03 août 2012 à 23h45 UTC (b) et du 22 août 2012 à 12h15 UTC (c)

L'une des situations étudiées est la tempête tropicale Isaac qui était la neuvième dépression tropicale de la saison (figure 7c). Elle s'est formée le 21 août 2012 à 06h UTC et a atteint le stade de tempête tropicale le 21 août à 18h UTC. Elle a suivi une trajectoire vers l'ouest puis s'est dirigée vers le nord-ouest à partir du 24 en matinée, et a atteint la côte sud d'Haïti le 25 août vers 3h locales. On peut estimer que sur certaines régions très localisées (sud et sud-ouest de Port au Prince), les cumuls ont atteint environ 200mm de pluie en 24 heures (Alizés [2012]).

Comme expliqué auparavant, le domaine des Petites Antilles étant essentiellement maritime, la répartition très éparse des mesures de pluie au sol nous amène à prendre en considération les estimations de pluie par satellite pour pouvoir étudier les situations et les comparer aux prévisions des modèles. Comme décrit précédemment, les produits satellitaires TMPA et GSMaP ainsi que les prévisions des modèles AROME et ALADIN ont été choisis pour cette étude.

Afin d'avoir un aperçu de la distribution des quantités de pluie sur toute la période et pour tous les points du domaine, des histogrammes des cumuls de pluie sur 6h sont étudiés. Ceci est réalisé pour les produits satellitaires d'une part, et pour une échéance donnée des prévisions des deux modèles d'autre part.



FIGURE 8 – Distributions des cumuls de pluies issus des deux produits satellitaires GSMaP et TMPA à l'échelle 0,5°/6h (à gauche) et distribution des cumuls de pluie sur 6h prévus à l'échéance +18h pour les deux modèles AROME_IFS et ALADIN (à droite) pour les trois situations étudiées

La figure 8 illustre le fait que les estimations GSMaP et TMPA présentent de plus nombreuses occurrences de pluies supérieures à 60mm en 6h que celles des modèles. Les deux produits satellitaires présentent une distribution très similaire pour les cumuls inférieurs à 60mm en 6h, mais divergent pour les plus forts cumuls. Il faut néanmoins se rappeler que chacun des produits comporte ses propres erreurs d'estimation qui ne peuvent être illustrées de façon simple sur ces figures.

Concernant les modèles, les histogrammes rassemblent toutes les occurrences des deux réseaux (00h et 12h UTC) disponibles en partie sur treize jours, donnant 24 prévisions. La figure 8 concerne l'échéance +18h qui répertorie ainsi les cumuls de pluie de +12 à +18h. La principale information qui ressort de ce graphique est la différence entre les deux modèles pour des cumuls supérieurs à 50mm/6h. En effet, AROME prévoit de tels cumuls de manière plus fréquente qu'ALADIN, et peut simuler des cumuls dépassant les maxima d'ALADIN. AROME se rapproche ainsi de la tendance des estimations des produits satellitaires. Cette constatation se généralise quelles que soient les échéances entre +6h et +24h (non montré ici). Enfin, ALADIN a tendance à produire des faibles cumuls plus souvent qu'AROME.

Pour tous les produits, les occurrences de cumuls de pluie les plus faibles sont évidemment les plus nombreuses, puisque les épisodes de forte pluie sont très localisés par rapport au domaine étudié.

Sur la figure 9, la situation de la tempête Isaac, plus précisément la période du 22 août 2012 de 12h à 18h, est prise comme exemple pour comparer les précipitations vues par les différents produits étudiés sur le domaine entier (vue en 2D) : GSMaP, TMPA et leurs écarts-type d'erreur associés respectifs, AROME et ALADIN à l'échéance +18h.

Au regard des figures 9a) et 9c), tandis que GSMaP indique un cumul maximal en 6h d'environ 200mm \pm 80mm (pour un intervalle de confiance à 68%) localisé vers 15°N et 59°W, le maximum des mesures TMPA est aux alentours de 100mm \pm 35mm (pour le même intervalle de confiance). Les deux produits présentent donc des erreurs importantes pour ce maximum; en revanche les intervalles de confiance se recoupent entre eux. On constate donc ici qu'associer une erreur à ces produits permet d'expliquer en partie leurs différences d'estimation.

Les figures 9e) et 9f) concernant les modèles montrent qu'ALADIN prévoit des maxima de pluie plus faibles répartis sur une zone plus étendue du domaine tandis qu'AROME prévoit plus de zones sèches et des pluies plus fortes qu'ALADIN et situe le phénomène pluvieux plus à l'est que les produits satellitaires. Ainsi, la comparaison des quatre produits sur cet exemple montre les deux principaux défauts des modèles : l'accumulation de pluie mal prévue en intensité et en localisation.



FIGURE 9 – Estimations des précipitations sur 6h issues du produit GSMaP (a) et du produit TMPA (c). Les écarts-types d'erreur de ces produits sont sur la figure (b) pour GSMaP et la figure (d) pour TMPA. Les figures (e) et (f) représentent les prévisions des cumuls de précipitations sur 6h à l'échéance +18h issues du modèle AROME_IFS et du modèle ALADIN respectivement. Toutes ces figures portent sur la journée du 22 août 2012 entre 12h00 et 18h00 UTC.

2. Méthodes de vérification des prévisions

2.1. Score point à point

La méthode la plus courante et la plus simple d'évaluation d'une prévision consiste à la comparer avec l'observation par point de grille. On suppose pour cela que l'on dispose d'un champ observé dont la grille correspond à celle de la prévision, en utilisant une interpolation, par exemple. On considère ainsi, pour chacun des points de grille, un couple (valeur observée, valeur prévue) que l'on traite indépendamment des autres points.

Les traitements que l'on applique à ces couples de valeurs sont variés, mais relèvent en général d'opérations statistiques standards. On calculera ainsi, par exemple, le biais moyen, c'est-à-dire la différence moyenne entre valeurs observée et prévue. De même, on étudiera la corrélation entre ces valeurs, et les moyennes quadratiques pour chaque couple (champ modèle, champ d'observation).

Il est également possible de tracer des diagrammes de dispersion, représentant chaque couple de valeurs par un point sur un graphique d'axes la valeur prévue et la valeur observée. Ce type de diagramme permet d'évaluer visuellement une éventuelle relation linéaire entre modèle et observation, ce qui marquerait une bonne prévision.

S'il s'agit d'une évaluation relativement simple à mettre en œuvre, elle présente l'inconvénient de ne pas tenir compte de ce que l'on pourrait qualifier de prévision « presque juste ». Tout décalage spatial sera sanctionné par une double peine : le résultat sera détérioré par l'absence de pluie prévue à la position observée, et par la présence de pluie prévue à une position sans pluie observée. Et ce, alors qu'une évaluation humaine aurait été très clémente avec cette même prévision, dans la mesure où la précision attendue pour un modèle n'est pas aussi fine que ne l'est la grille.

2.2. Vérification avec tolérance

Comme vu précédemment, les comparaisons point à point ont des limites, et pénalisent fortement un modèle lorsque celui-ci prévoit des précipitations décalées dans l'espace, même d'un seul point de grille. Par conséquent, il est difficile de quantifier l'apport d'un modèle de méso-échelle avec cette méthodologie, car ce type de modèle peut éventuellement introduire plus d'erreurs point à point que les modèles de plus grande échelle, surtout sur sa grille native. C'est dans l'optique d'introduire une nouvelle notion de score, en relation avec le voisinage, qu'ont été développées plusieurs méthodes de vérification avec tolérance. Ces techniques étudient la correspondance entre l'observation et la prévision dans les environs de l'observation (on regarde les proximités en espace, en temps et en intensité). Plusieurs méthodes de vérification avec tolérance existent dans la littérature (Ebert [2008]) et se différencient sur les concepts de proximité, sur la définition d'une prévision utile (modèle de décision) et sur le choix du score.

À Météo-France, ce type de score est utilisé de manière opérationnelle pour comparer les prévisions de précipitations des modèles ARPEGE et AROME sur la métropole. Ces travaux sont décrits dans l'article Amodei [2009].

Dans le cadre de ce projet, la méthode de vérification avec tolérance utilisée est appelée Fractions Skill Score (FSS). Dans un premier temps, le FSS est décrit puis son fonctionnement est explicité sur un exemple simplifié.

2.2.1. Fractions Skill Score

Le but du FSS est de comparer les observations à la prévision sur un même voisinage. Cette technique diffère de la comparaison entre une observation d'un point de référence et la prévision à un voisinage alentour. La notion de tolérance spatiale pour le FSS peut se définir par la figure 10 :



Domaine d'observation Domaine de prévision

FIGURE 10 – Définition du voisinage d'observation et du voisinage de la prévision pour la vérification de la performance de la prévision. Source : Ebert [2008]

La taille des deux domaines est identique.

Le FSS est décrit dans l'article Roberts [2008]. Cette méthode compare directement les fractions de la couverture des événements dans le voisinage entourant les observations et les prévisions. Le modèle de décision pour ce score est le suivant : une prévision est utile si la fréquence des événements prévus est similaire à la fréquence des événements observés. Plutôt que de déduire les événements du voisinage des fréquences, elles sont utilisées directement pour définir le Fractions Brier Score (FBS) comme suit :

$$FBS = \frac{1}{N} \times \sum_{i=1}^{N} (Po(i) - Pf(i))^2$$

N étant le nombre de points de grille dans la zone de tolérance, Po la fraction des points de grille avec de la pluie observée au-dessus d'un seuil, et Pf la fraction des points de grille avec de la pluie prévue au-dessus de ce même seuil. A partir du FBS, on peut déduire le FSS :

$$FSS = 1 - \frac{FBS}{FBS_{ref}}$$

avec

$$FBS_{ref} = \frac{1}{N} \times \sum_{i=1}^{N} (Po(i) \times Po(i) + Pf(i) \times Pf(i))$$

 FBS_{ref} est la pire prévision possible, c'est-à-dire qu'il n'y a aucune pluie prévue dans le voisinage des pluies observées. On l'obtient en mettant à 0 le terme croisé du FSS.

Le FSS prend des valeurs comprises entre 0 et 1 avec 0 pour une correspondance nulle et 1 pour une correspondance parfaite. Les meilleurs scores pour le FSS sont obtenus lorsque les tolérances spatiales sont importantes, car la prévision et l'observation sont comparées sur un plus grand domaine, et lorsque les seuils d'intensité sont modérés, car ce sont les seuils où les prévisions ont le moins de biais.

Comme ces techniques différencient les seuils d'intensité de pluies et les tolérances spatiales, on peut déterminer à partir de quels seuils et de quelles tolérances spatiales on arrive à quantifier, par exemple, l'impact d'une amélioration apportée au modèle de prévision. De plus, ce score permet de déterminer à quelles intensités et avec quelles tolérances spatiales la prévision est la plus fiable.

2.2.2. Exemple simplifié

Afin d'illustrer le fonctionnement du Fractions Skill Score, voici un exemple idéalisé de présence de deux zones pluvieuses sur un domaine de 5x5 points de grille, illustré sur la figure 11.

1	1	0	0	0
1	1	0	0	0
0	0	0	0	0
0	25	25	0	0
0	25	0	0	0

0	0	1	1	0
0	0	1	1	0
0	0	0	0	0
0	10	10	0	0
0	10	10	0	0

FIGURE 11 – Tableau d'observation à gauche et tableau de prévision à droite

La première zone pluvieuse présente des cumuls de 1mm et n'a pas la même position entre son état observé et la prévision. La seconde zone pluvieuse observée présente des cumuls de 25mm de pluie, tandis que la prévision n'est que de 10mm, mais sur une plus grande étendue spatiale.

Dans l'optique de simplifier l'exemple, seuls cinq seuils d'intensité (1mm, 5mm, 10mm, 20mm, 50mm) et cinq tolérances spatiales ont été considérés. La première tolérance compare les observations et les prévisions point à point et la dernière tolérance compare les observations et prévisions sur l'ensemble du domaine.

L'intégralité des scores est résumée sur la figure 12 (la tolérance 0 est la plus petite, la tolérance 4 correspond à l'ensemble du domaine). On retrouve ici les meilleurs scores du FSS lorsque les tolérances spatiales sont importantes car la prévision et les observations sont comparées sur un plus grand domaine.

La valeur 0.4 du FSS de la tolérance 0 et du seuil 1mm peut être expliquée ainsi :

	1 mm	5 mm	10 mm	20 mm	50 mm
Tolérance 0	0,4	0,857	0,857	0	-999
Tolérance 1	0,552	0,96	0,96	0	-999
Tolérance 2	0,807	0,958	0,958	0	-999
Tolérance 3	0,962	0,969	0,969	0	-999
Tolérance 4	0,991	0,96	0,96	0	-999

FIGURE 12 – Scores du FSS pour l'exemple

Pf = [0, 0, 1, 1, 0]
$0. \ 0. \ 1. \ 1. \ 0.$
0. 0. 0. 0. 0.
0. 1. 1. 0. 0.
0. 1. 1. 0. 0.]

FIGURE 13 – Les tableaux Po et Pf pour le seuil 1mm et la tolérance 0

N=25 (nombre de points de grille);

Po correspond au tableau des observations et Pf correspond au tableau des prévisions (les valeurs de Po et de Pf pour le seuil 1mm et la tolérance 0 sont explicitées sur la figure 13). La méthode de remplissage est la même pour les deux tableaux. La tolérance spatiale considérée implique qu'il s'agit d'un score point à point, c'est pourquoi la valeur 1 apparaît lorsque le seuil est dépassé sur la case correspondante.

Les différentes étapes de calcul sont détaillées ici :

$$FBS = \frac{1}{N} \times \sum_{i=1}^{N} (Po(i) - Pf(i))^2 = \frac{9}{25} = 0.36$$

FBS_{ref} = $\frac{1}{N} \times \sum_{i=1}^{N} (Po(i) \times Po(i) + Pf(i) \times Pf(i)) = \frac{15}{25} = 0.6$

d'où

$$FSS = 1 - \frac{FBS}{FBS_{ref}} = 1 - \frac{0.36}{0.6} = 0.4$$

Ce score n'est pas très élevé. Ceci est dû en majeure partie au fait que les valeurs de 1mm n'ont pas été prévues au bon endroit.

Pour la tolérance 4 avec un seuil de 1mm, le score est plus élevé.

N=1, Po = [0.28], Pf = [0.32]

La valeur 0.28 correspond au rapport du nombre 7 (nombre de valeurs d'observation dépassant 1mm sur le domaine) par le nombre 25 (nombre total de valeurs). De même, la valeur 0.32 est le rapport de 8 valeurs de prévisions dépassant 1mm sur le domaine par 25 valeurs au total.

$$\begin{split} FBS &= 0.0016\\ FBS_{ref} &= 0.1808\\ et\ FSS &= 0.991 \end{split}$$

Dans ce cas, peu importe la correspondance de l'emplacement des observations et des prévisions, seul compte le nombre d'occurrences dépassant les 1mm.

Pour le seuil 20mm, les scores pour l'ensemble des tolérances sont nuls. Sur cet exemple, aucun point de grille ne présente de prévision supérieure à 20mm, ce qui implique que Pf est nul pour toutes les tolérances. Par conséquent, FBS = FBSref, et donc FSS = 0.

Enfin, les valeurs -999 pour le seuil de 50mm signalent l'absence totale de point de grille à plus de 50mm. Les tableaux Po et Pf sont alors nuls, FBS et FBS_{ref} également, ce qui provoque une division par zéro; l'algorithme remplace alors ce calcul par la valeur -999 par défaut.

2.3. Prise en compte des erreurs de la référence

L'utilisation du Fractions Skill Score décrit précédemment induit la comparaison du modèle avec une référence supposée exacte. Or, les données d'observation issues des produits satellitaires présentés plus haut sont entachées d'erreurs qu'il est possible d'estimer. La prise en compte de ces incertitudes est donc nécessaire pour améliorer la confiance dans les scores calculés.

Les erreurs d'échantillonage estimées sont comparables à des écart-types de lois de probabilité. Il s'agit donc de faire une hypothèse sur la nature de cette loi, puis de perturber chaque point d'observation grâce à cette loi, de moyenne la valeur de l'observation et d'écart-type l'erreur estimée en ce point.

Reproduire ces perturbations un grand nombre de fois permet d'obtenir une fonction densité de probabilité plutôt qu'un score unique.

Afin d'illustrer le fonctionnement du FSS perturbé, voici un exemple idéalisé comportant deux zones pluvieuses. On a toujours une prévision (figure 15) et une observation (figure 14), mais on introduit également, pour l'exemple, deux autres observations qui sont perturbées (figures 16 et 17). Ces deux dernières présentent un décalage en intensité par rapport à l'observation car on ajoute une incertitude à la mesure. Celle-ci ne dépasse pas 100% de la mesure originale effectuée. L'incertitude n'est ajoutée que lorsque l'observation est non nulle (pour les points de grille où l'observation est nulle, l'observation perturbée sera nulle aussi). Dans cet exemple, les incertitudes ont été choisies volontairement conséquentes pour obtenir des écarts importants, mais elles restent réalistes : la majorité des incertitudes est comprise entre 40 et 100% de la valeur mesurée, en accord

1	1	0	0	0
1	1	0	0	0
0	0	0	0	0
0	11	11	0	0
0	11	0	0	0

FIGURE 14 – Tableau d'observation

2	1.6	0	0	0
0.6	0.9	0	0	0
0	0	0	0	0
0	<mark>15.2</mark>	9.1	0	0
0	6.2	0	0	0

FIGURE 16 – Tableau d'observation perturbé 1

0	0	1	1	0		
0	0	1	1	0		
0	0	0	0	0		
0	10	10	0	0		
0	10	10	0	0		
FIGURE 15 – Tableau de prévision						

1	0.2	0	0	0
1.9	1	0	0	0
0	0	0	0	0
0	5.3	16.8	0	0
0	11.1	0	0	0

FIGURE 17 – Tableau d'observation perturbé 2

avec les valeurs d'erreurs effectives sur les deux produits satellitaires (section 1.1.3.). Les tableaux d'entrée sont à nouveau composés d'une grille 5x5.

Pour simplifier cet exemple, seuls 3 seuils d'intensité (1mm, 10mm, 20mm) et 5 tolérances spatiales ont été considérés. La première tolérance spatiale permet de comparer les observations et les prévisions point à point, et la dernière tolérance spatiale de les comparer sur l'ensemble du domaine.

Le tracé des différents résultats du Fractions Skill Score est réalisé pour le premier seuil et la première tolérance sur la figure 18. On retrouve les trois valeurs du FSS pour les trois observations (une non perturbée et deux perturbées) : 0.4, 0.462 et 0.429. Les valeurs du FSS varient selon les observations et montrent donc que les incertitudes de la référence impactent le score final. Ce résultat se généralise pour l'ensemble des tolérances et seuils.

Un autre tracé des différents résultats du FSS avec un seuil de 10mm et la quatrième tolérance est présent sur la figure 19.

Sur cet exemple, la valeur du score peut passer presque du simple au double. Cette sensibilité accrue est liée au nombre d'échantillons dans cet exemple idéalisé, mais montre tout de même l'importance de la prise en compte de cette incertitude.

En résumé, il existe des incertitudes relatives à la mesure des précipitations par satellite qui sont prises en compte par l'introduction des perturbations. Ces dernières permettent d'avoir des valeurs des scores plus réalistes car on considère que les observations sont



FIGURE 18 – Fonction densité de probabilité du FSS pour un seuil de 1mm et la tolérance spatiale 0

entachées d'erreurs.



FIGURE 19 – Fonction densité de probabilité du FSS pour un seuil de 10mm et la tolérance spatiale 3

3. Mise en place d'un score et tests de sensibilité

3.1. Mise en place d'un score

L'objectif de cette étude est l'intercomparaison des modèles AROME et ALADIN pour les prévisions de précipitations dans les régions tropicales. C'est pourquoi tous les résultats présentés par la suite correspondent à une différence de FSS : celui d'ALADIN est soustrait à celui d'AROME, de sorte que la convention choisie fasse correspondre un résultat positif à un meilleur score pour AROME.

Cette différence de scores n'a de sens que si la référence utilisée pour les calculer est la même quel que soit le modèle. C'est pourquoi, pour chacun des états observés perturbés, les deux FSS puis leur différence sont évalués, avant de générer une nouvelle perturbation et d'effectuer un nouveau calcul de score, pour obtenir, comme expliqué précédemment, une densité de probabilité de différences de FSS plutôt qu'une seule valeur.

Dans la suite de ce rapport, on s'intéressera donc à des distributions de différences de FSS entre deux ensembles de prévisions, la distribution provenant du fait que l'on traite à la fois plusieurs réseaux et plusieurs perturbations sur la référence, donnant à chaque réalisation une nouvelle différence de FSS.

Avant d'utiliser la méthode de vérification avec tolérance et prise en compte des erreurs de la référence décrite précédemment, il convient d'étudier sa stabilité quand on fait varier certains de ses constituants. En effet, un score très dépendant de paramètres intrinsèques à la méthode ne pourrait pas être considéré comme très robuste. On parle alors de test de sensibilité à chacun des paramètres.

	Nombre de perturbations	Loi de probabilité des erreurs	Intervalle de confiance considéré pour ces erreurs	Produit satellitaire de référence	
Référence	1000	Hypothèse d'une gaussienne de variance l'erreur d'échantillonnage proposée pour chaque produit	±1σ (représentant 68% de la distribution)	GSMaP	
Test 1)	100	référence	référence	référence	
Test 2)	référence	Hypothèse d'une distribution uniforme	référence	référence	
	référence	référence	±1,96σ (représentant 95% de la distribution)	référence	
Test 3)	Test 3) référence référence réf		référence	ТМРА	

FIGURE 20 – Tests de sensibilité

Dans cette étude de sensibilité, les trois paramètres étudiés sont les suivants :

- le nombre de perturbations effectuées pour obtenir la fonction densité de probabilité;
- la prise en compte des erreurs, c'est-à-dire l'hypothèse sur la loi de probabilité suivie par les erreurs, et l'intervalle de confiance choisi borné dans lequel les valeurs aléatoires sont tirées;
- le produit satellitaire utilisé comme référence.

Les expériences réalisées pour cette étude de sensibilité sont décrites dans le tableau de la figure 20.

On rappelle que les données utilisées sont constituées de vingt-quatre prévisions de cumuls de pluie sur 6h effectuées sur les deux réseaux principaux durant l'été 2012; seule l'échéance +18h est utilisée (c'est-à-dire la quantité de pluie sur 6h de +12h à +18h après le réseau).

3.2. Sensibilité au nombre de perturbations

En statistiques, on considère qu'un nombre important d'expériences aléatoires répétées dans les mêmes conditions est essentiel pour parvenir à une conclusion valable. Cependant, en pratique, les temps de calcul imposent de limiter ce nombre de répétitions.

Dans notre cas, le nombre de points à partir desquels les fonctions densité de probabilité sont tracées dépend du nombre de perturbations des observations. C'est la sensibilité du score à ce nombre que l'on teste dans cette partie.



FIGURE 21 – Fonctions densité de probabilité de la différence de FSS pour un seuil de 10mm et une tolérance spatiale de 50km pour 1000 perturbations (à gauche) et pour 100 perturbations (à droite)

La figure 21 représente les fonctions densité de probabilité avec 1000 perturbations et 100 perturbations, pour un seuil d'intensité de 10mm et une tolérance spatiale de 50km. Les différents maxima des distributions que l'on observe correspondent aux différents réseaux que l'on traite dans cette comparaison, tandis que la dispersion autour de ces maxima est engendrée par la prise en compte des erreurs dans le processus de calcul.

Les maxima sont moins élevés sur le graphique de droite car il y a seulement 2400 points représentés, contre 24000 sur le graphique de gauche. Cependant, on constate que l'allure de la fonction densité pour 100 perturbations se rapproche de cette même fonction pour 1000 perturbations. La fonction densité de probabilité avec 1000 perturbations est seulement plus lisse.

L'analyse des autres tolérances spatiales et seuils montre un comportement similaire. On peut en conclure que les différences entre 100 et 1000 perturbations sont assez faibles. La valeur de 1000 perturbations a été utilisée dans la suite de cette étude, mais une valeur de 100 semble suffisante pour une utilisation ultérieure.

3.3. Sensibilité à la prise en compte des erreurs

Le second test de sensibilité consiste à évaluer l'influence de la loi de distribution des erreurs de la référence. Les distributions réelles des erreurs des produits satellitaires étant très difficiles à caractériser, il a été nécessaire de faire des hypothèses quant aux lois suivies par ces erreurs. D'une part, on étudie la sensibilité à la loi de la distribution et d'autre part, on évalue l'influence de l'intervalle de confiance choisi pour borner la distribution des erreurs, uniquement pour la loi normale.

L'observation perturbée est issue d'un tirage aléatoire d'après une loi de probabilité choisie et centrée sur l'observation non perturbée. L'erreur estimée du produit satellitaire (section 1.1.3.) est prise comme écart-type de la loi. Ainsi, chaque cumul $0.5^{\circ}/6h$ est perturbé. Deux lois de distribution sont testées : gaussienne et uniforme.

Le premier test de sensibilité consiste à évaluer l'influence de l'utilisation de la loi uniforme bornée (ou porte) par rapport à l'expérience de référence réalisée avec la loi normale, bornée également à $\pm 1\sigma$. La loi uniforme correspond à une estimation pessimiste des erreurs, dans la mesure où les erreurs fortes sont beaucoup plus fréquentes qu'avec la loi gaussienne.



FIGURE 22 – Fonctions densité de probabilité de la différence de FSS pour un seuil de 10mm et une tolérance spatiale de 00km pour une loi normale (à gauche) et pour une loi uniforme (à droite)

Sur la figure 22 sont montrés les résultats obtenus pour les fonctions densité de probabilité dans le cas de la référence et dans le cas de l'expérience avec la loi uniforme pour un seuil de 10mm et une tolérance spatiale nulle. On remarque que l'allure des densités est similaire d'un cas à l'autre. En effet, les plus grands nombres d'occurrences, visibles sous forme de maxima sur les figures, sont associés aux mêmes valeurs de différences de scores. Ces dernières sont réparties dans les deux cas entre -0.4 et 0.45 environ. On note également que les fractions de cas avec une différence de score positive (respectivement de 59.3% et 59.1%) restent proches, ce qui n'impacte donc pas une conclusion sur la qualité des prévisions d'un modèle par rapport à un autre par exemple. Ce résultat se généralise aux autres échelles spatiales et aux autres seuils de cumuls de précipitations.

Le calcul du score est donc robuste au changement de la loi de probabilité. Or, on aurait attendu une différence plus importante, compte tenu de l'impact statistique que pourrait avoir ce changement de loi. Cette faible sensibilité peut s'expliquer par le calcul d'une différence de score au lieu d'un score absolu, induisant une compensation des différences sur les scores absolus.

Le second test de sensibilité consiste à évaluer l'influence de l'intervalle de confiance choisi pour borner la distribution des erreurs. Dans l'expérience de référence, celle-ci est une loi normale bornée à $\pm 1\sigma$ soit 68%. Dans l'expérience de sensibilité, on utilise un intervalle de confiance à 95%, la loi normale est donc tronquée à $\pm 1,96\sigma$ pour les tirages aléatoires.

La figure 23 montre les fonctions densité de probabilité obtenues pour le seuil 50mm et la tolérance spatiale de 250km. On remarque de nouveau que leur allure est similaire. Les pourcentages ne sont que peu modifiés par le changement de loi ; dans cet exemple, ils sont même égaux (99,1%). Ce résultat se généralise aussi aux autres échelles et intensités de pluie. Le score choisi est donc stable à ce paramètre, ce qui rend son utilisation légitime.



FIGURE 23 – Fonctions densité de probabilité de la différence de FSS pour un seuil de 50mm et une tolérance spatiale de 250km pour $\pm 1\sigma$ (à gauche) et pour $\pm 1.96\sigma$ (à droite) avec une loi normale

3.4. Sensibilité au produit satellitaire de référence

Les deux produits satellitaires utilisés pour ces travaux sont GSMaP et TMPA. S'ils ne sont pas identiques, on peut tout de même se demander si la prise en compte de leurs erreurs suffit à limiter leurs différences dans le calcul du score. Le test de sensibilité sur la référence permet d'observer l'impact du changement de produit satellitaire sur les scores.

La partie analyse des erreurs n'a pas permis d'écarter un produit satellitaire qui serait de moindre qualité, c'est pourquoi l'un comme l'autre peut servir de référence. En effet, ont été comparés dans cette partie uniquement des écarts-types d'estimation de la moyenne liés à l'échantillonnage, les biais de ces produits ne sont pas pris en compte car difficiles à caractériser sur un domaine océanique.

Pour cette étude, le produit de référence choisi est GSMaP et on compare les résultats de vérification de modèle avec ceux que l'on obtient si on utilise le produit TMPA comme référence.



FIGURE 24 – Fonctions densité de probabilité de la différence de FSS pour un seuil de 20mm et une tolérance spatiale de 250km pour le produit satellitaire GSMaP (à gauche) et pour le produit satellitaire TMPA (à droite)

Sur la figure 24, les fonctions densité de probabilité sont représentées pour un seuil de 20mm et une tolérance spatiale de 250km. Il y a un décalage des différences de scores vers des valeurs plus élevées lorsque l'on passe de GSMaP vers TMPA (pour les seuils de 20 et 50mm).

Il existe donc des seuils et tolérances spatiales pour lesquels les résultats sont sensibles au choix du produit satellitaire. Cependant, ces cas représentent moins de 15% des cas étudiés. Un cas plus représentatif de la majorité est sur la figure 25.



FIGURE 25 – Fonctions densité de probabilité de la différence de FSS pour un seuil de 5mm et une tolérance spatiale de 200km pour le produit satellitaire GSMaP (à gauche) et pour le produit satellitaire TMPA (à droite)

Pour ce seuil de 5mm et cette tolérance spatiale de 200km, aucun décalage n'est à noter. On retrouve les mêmes maxima représentant un nombre important d'occurrences

sur les deux fonctions densité de probabilité de TMPA et de GSMaP, ce qui montre que les allures de ces deux densités sont assez semblables.

En résumé, les fonctions densité de probabilité sont en grande majorité similaires mais peuvent présenter, dans certains cas, des différences notables pour certains seuils d'intensité de pluie et certaines tolérances spatiales. Ces éléments appellent d'une part à prendre en compte un bilan d'erreur plus complet pour les produits satellitaires dans la méthodologie de vérification mise en œuvre, et d'autre part, à utiliser plusieurs produits de référence afin de confirmer la robustesse des résultats.

4. Mise en œuvre du score

Comme précédemment décrit, la vérification avec tolérance et prise en compte de l'erreur de la référence est suffisamment robuste pour être utilisée pour l'évaluation comparative des modèles ALADIN et AROME sur l'Outre-Mer.

Pour rappel, les fonctions densité de probabilité sont des différences de scores. En plus des densités de probabilité par seuil d'intensité et par tolérance spatiale, est calculée ici la fréquence de l'événement (dans le cas où on compare AROME moins ALADIN) « le score d'AROME est supérieur à celui d'ALADIN ». Sa significativité est ensuite évaluée grâce à un test du Chi2. Cela permet d'associer à chaque fréquence une interprétation : soit la fréquence des différences favorables à AROME est significativement supérieure à 50%, soit elle est significativement inférieure à 50%. Dans le cas où le résultat est trop proche de 50%, ou qu'il y a trop peu d'échantillons pour que la différence soit significative d'après le test de Chi2, on ne peut conclure qu'un modèle est meilleur que l'autre.

Les résultats qui suivent sont présentés sous forme de tableaux contenant les pourcentages décrits dans le paragraphe précédent, en fonction des seuils d'intensité des précipitations et des tolérances spatiales. Les cases sont également colorées en fonction de la significativité de ces pourcentages avec un degré de confiance de 95% : si l'on compare AROME moins ALADIN, la couleur jaune signifie que le score d'AROME est le plus souvent supérieur à celui d'ALADIN, la couleur bleue signifie l'inverse, et la couleur grise signifie qu'aucune conclusion n'est possible quant à la supériorité de l'un des modèles sur l'autre.

Dans cette partie, chaque version d'AROME (AROME_ALA et AROME_IFS) sera comparée avec le modèle ALADIN, puis AROME_ALA et AROME_IFS seront comparés entre eux.

4.1. Comparaison AROME IFS et ALADIN

Sur la figure 26 se trouvent les tableaux des fréquences de l'évènement « le score d'AROME est supérieur à celui d'ALADIN », pour les échéances +6h (en haut à gauche), +12h (en haut à droite), +18h (en bas à gauche) et +24h (en bas à droite).

En se basant sur la couleur bleue (qui, pour rappel, signifie ici que le score d'ALADIN est souvent supérieur à celui d'AROME_IFS), on remarque que les prévisions ARO-ME_IFS sont meilleures que les prévisions ALADIN au fur et à mesure des échéances. Pour les seuils d'intensité supérieurs à 50mm, les prévisions AROME_IFS obtiennent toujours des meilleurs scores que les prévisions ALADIN. De plus, ces pourcentages sont très importants et augmentent avec les échéances.

Pour les cumuls de moins de 50mm, ALADIN est meilleur en début de prévision (les six premières heures). Cependant, dès la deuxième échéance étudiée, les cumuls de pluie supérieurs à 20mm sont mieux prévus par AROME_IFS, et ceux qui sont supérieurs à 1mm et 10mm le deviennent également pour certaines tolérances spatiales.

En ce qui concerne le seuil 5mm, AROME_IFS est meilleur qu'ALADIN seulement pour les grandes tolérances spatiales et à partir de l'échéance +18h. On retrouve cette conclusion également pour le seuil 1mm où ALADIN obtient des meilleurs scores qu'ARO-ME_IFS pour les échéances +18h et +24h à des faibles tolérances spatiales.

	•		(c)							(d)			•
	1 mm	5 mm	10 mm	20 mm	50 mm	100 mm		1 mm	5 mm	10 mm	20 mm	50 mm	100 mm
0 km	38,4	38,2	59,1	99,1	100,0	nan	0 km	45,4	40,5	65,3	77,2	100,0	nan
50 km	50,2	46,8	59,3	92,3	100,0	nan	50 km	50,0	45,2	59,2	79,2	100,0	nan
100 km	55,3	51,0	60,7	89,8	100,0	nan	100 km	53,8	45,3	60,1	85,4	100,0	nan
150 km	55,7	58,1	60,7	79,2	100,0	100,0	150 km	56,4	43,3	61,1	86,8	100,0	100,0
200 km	56,8	59,0	60,6	79,1	99,9	100,0	200 km	62,9	45,6	67,2	<mark>85,3</mark>	100,0	100,0
250 km	63,3	60,5	59,5	78,9	99,0	100,0	250 km	69,5	50,9	74,2	<mark>85,3</mark>	100,0	100,0
			(a)							(b)			
	1 mm	5 mm	10 mm	20 mm	50 mm	100 mm		1 mm	5 mm	10 mm	20 mm	50 mm	100 mm
0 km	32,6	4,3	19,0	48,8	57,1	nan	0 km	<mark>52,9</mark>	25,6	44,6	74,5	70,5	nan
50 km	34,7	4,6	17,1	46,9	70,3	100,0	50 km	61,2	26,1	49,4	72,1	83,0	100,0
100 km	35,8	4,8	16,9	38,4	74,1	100,0	100 km	64,4	28,4	51,2	70,2	85,4	100,0
150 km	36,1	5,2	16,9	40,5	78,3	100,0	150 km	68,5	31,1	49,6	67,7	85,4	100,0
200 km	36,0	6,8	15,8	38,6	82,9	100,0	200 km	74,1	33,3	48,9	62,5	87,2	100,0
250 km	35,4	9,0	11,4	34,6	68,2	100,0	250 km	78,1	33,5	45,5	63,3	87,2	100,0

FIGURE 26 – Tableaux des pourcentages de AROME_IFS supérieur à ALADIN pour les 24 réseaux pour les échéances +06h (a), +12h (b), +18h (c) et +24h (d)

Pour les petits seuils, ALADIN obtient des meilleurs scores qu'AROME_IFS aux petites tolérances spatiales pour la majorité des échéances temporelles et il faut attendre l'échéance +18h pour qu'AROME_IFS commence à devenir meilleur qu'ALADIN aux grandes tolérances spatiales. Une explication possible à cette évolution en fonction des échéances est liée à l'initialisation d'AROME. En effet, pendant les premières heures de prévision AROME, le modèle doit s'adapter à des conditions initiales provenant d'IFS, et donc d'une physique différente; ce n'est donc qu'après une période d'ajustement que les prévisions d'AROME deviendraient supérieures à celle d'ALADIN, qui possède son système d'assimilation permettant de démarrer d'un état initial équilibré.

4.2. Comparaison AROME ALA et ALADIN

La seconde étape consiste à comparer les modèles AROME_ALA et ALADIN à conditions initiales égales. Cette comparaison permet de savoir si les prévisions AROME_IFS se révèlent être meilleures que celles d'ALADIN uniquement grâce aux analyses IFS qui seraient de meilleures qualités que les analyses d'ALADIN. Cela revient aussi à comparer la qualité des prévisions des deux modèles avec la même initialisation.

Les conclusions de cette comparaison sont similaires à celles de la comparaison AROME-_IFS versus ALADIN : (1) AROME a toujours un meilleur score qu'ALADIN pour les cumuls de pluie supérieurs à 50mm et ce, pour les quatre échéances ; (2) les seuils inférieurs à 50mm sont favorables à ALADIN pour les six premières heures de prévision ; (3) plus l'échéance devient grande, plus la fonction densité de probabilité tend à se décaler vers une différence de score positive, montrant ainsi qu'AROME est meilleur qu'ALADIN. Un exemple est sélectionné pour illustrer cette comparaison AROME_ALA versus ALADIN : sur la figure 27 est représenté un exemple de fonctions densité de probabilité pour le seuil de 20mm et la tolérance spatiale de 50km pour les quatre échéances (+06h, +12h, +18h et +24h). On constate que le pourcentage de prévisions en faveur d'AROME_ALA passe de 46% pour l'échéance de +6h, à 76.3% puis 79.3% et enfin 96.8% pour les cumuls de pluie de l'échéance +24h.



FIGURE 27 – Fonctions densité de probabilité pour AROME_ALA moins ALADIN pour un seuil de 20mm et une tolérance spatiale de 50km pour l'échéance +06h (a), l'échéance +12h (b), l'échéance +18h (c) et l'échéance +24h (d). La couleur jaune signifie qu'AROME_IFS est meilleur qu'ALADIN, le bleu signifie l'inverse et la couleur grise indique une différence non significative entre les deux modèles

4.3. Comparaison des deux AROME

Les résultats montrés précédemment illustrent le fait qu'AROME_ALA et AROME_ IFS proposent des prévisions de précipitations meilleures que celles d'ALADIN dès six heures de prévision, et à la résolution spatiale de 0.5° étudiée dans ce rapport. Une des perspectives entreprises pour l'amélioration future d'AROME Outre-Mer est l'étude des performances du modèle avec des conditions initiales différentes.

La troisième et dernière étape de cette étude consiste donc à comparer directement A-ROME_ALA avec AROME_IFS. Les figures 28 et 29 montrent les deux tableaux des fré-

	1 mm	5 mm	10 mm	20 mm	50 mm	100 mm
0 km	59,7	48,8	55,7	57,5	52,9	nan
50 km	57,5	50,4	52,9	62,1	53,4	0,0
100 km	58,1	50,0	47,7	65,9	54,2	0,0
150 km	54,4	49,7	46,2	70,9	67,8	0,0
200 km	52,2	50,0	48,1	70,1	76,8	7,9
250 km	50,7	49,4	52,3	75,6	80,5	75 <i>,</i> 8

FIGURE 28 – Tableau des pourcentages d'AROME_IFS supérieurs à AROME_ALA pour les 24 réseaux à l'échéance +18h avec le produit GSMaP

	1 mm	5 mm	10 mm	20 mm	50 mm	100 mm
0 km	47,5	37,9	50,8	64,5	83,4	0,0
50 km	48,5	42,8	52,6	52,4	89,2	0,0
100 km	48,8	37,5	52,8	59,3	79,7	0,0
150 km	40,7	35,7	51,9	64,2	78,9	0,0
200 km	38,6	34,4	51,5	67,4	78,0	0,0
250 km	40,7	32,1	43,4	71,6	75,7	0,0



quences de l'évènement « le score d'AROME_IFS est supérieur à celui d'AROME_ALA » pour l'échéance +18h avec GSMaP et TMPA respectivement comme produit satellitaire de référence pour le calcul du score du modèle. La couleur jaune indique ici que la version d'AROME initialisée avec IFS a un meilleur score que celle initialisée avec ALADIN.

Les pourcentages d'AROME_IFS supérieurs à AROME_ALA pour l'échéance +18h et le produit satellitaire GSMaP sont dans le tableau de la figure 28. Pour les seuils de 1mm, 20mm et 50mm, AROME_IFS est meilleur qu'AROME_ALA avec parfois des pourcentages importants (80.5% pour un seuil de 50mm et une tolérance spatiale de 250km). Pour le seuil de 100mm, AROME_ALA obtient de très bons scores (100% pour une tolérance spatiale de moins de 150km) ce qui signifie qu'AROME_IFS ne prévoit pas de cumul de pluie supérieur à 100mm avec une bonne précision. En ce qui concerne les seuils 5mm et 10mm, aucune version d'AROME n'est globalement meilleure que l'autre. On peut conclure qu'AROME_ALA et AROME_IFS n'obtiennent pas les mêmes scores, particulièrement pour des seuils supérieurs à 20mm, avec le produit satellitaire GSMaP.

Dans le cadre des tests de sensibilité présentés précédemment, les résultats des comparaisons AROME versus ALADIN ont toutes été réalisées avec GSMaP et TMPA en référence, mais seuls les résultats avec GSMaP ont été montrés car les résultats étaient peu sensibles à ce choix, sauf quelques cas de seuils et de tolérances spatiales, comme discuté dans la section 3.4.. Un autre test de sensibilité aux produits satellitaires a été fait ici avec les deux versions d'AROME.

On remarque que quel que soit le produit satellitaire utilisé, AROME_IFS présente de meilleurs scores que l'autre version pour des cumuls de pluie supérieurs à 20 et 50mm, et ce quelle que soit la tolérance spatiale. C'est en revanche AROME_ALA qui obtient les meilleurs scores pour les cumuls supérieurs à 100mm. En effet, AROME_ALA a 100% de scores supérieurs à ceux d'AROME_IFS et ceci, pour la grande majorité des tolérances spatiales (on rappelle que 100% de scores favorables à AROME_ALA correspond au 0% de scores favorables à AROME_IFS dans les tableaux des figures 28 et 29). Cependant il faut relativiser ce résultat devant la petite taille des échantillons et le fait que l'on ne considère que l'échéance +18h.

Pour des seuils de pluie de 1 et 5mm, l'utilisation du produit satellitaire TMPA est favorable aux prévisions AROME_ALA tandis que l'emploi de GSMaP rend les scores d'AROME_IFS meilleurs pour le seuil 1mm, et ne permet pas de conclure quant à la performance des modèles pour le seuil 5mm.

Ainsi, dans le cas des comparaisons AROME_IFS versus AROME_ALA, les résultats obtenus sont nettement sensibles au choix du produit satellitaire de référence.

En résumé, lorsqu'on compare chacune des versions d'AROME avec ALADIN (cf. parties précédentes), les conclusions sont semblables en utilisant l'un ou l'autre des produits satellitaires, mais lorsqu'on compare les deux versions d'AROME, des changements de performance des modèles sont à noter, notamment aux seuils d'intensité les plus faibles. Ceci pourrait s'expliquer par le fait que les différences entre les prévisions AROME et ALADIN sont plus importantes que celles entre AROME_IFS et AROME_ALA; les différences de score sont donc potentiellement plus simples à mettre en évidence.

La sensibilité décrite dans ce paragraphe appelle à nouveau à prendre en compte un bilan d'erreur plus détaillé pour les produits satellitaires, ou bien à étudier des cumuls de pluie à une échelle plus agrégée, du type $1^{\circ}/1$ jour permettant une meilleure convergence entre les produits.

Conclusion et perspectives

Cette étude propose une première évaluation des prévisions de pluie produites par AROME Outre-Mer, en comparaison avec celles du modèle ALADIN actuellement opérationnel sur le domaine des Petites Antilles. Les références choisies pour cette étude sont des produits d'estimation de pluie par satellite afin de pouvoir évaluer les prévisions de précipitations sur tout le domaine des modèles et pas seulement sur les terres émergées.

Une méthode de comparaison avec tolérance spatiale a été mise en œuvre, ce qui permet de confronter les observations et les prévisions sur un voisinage et pour différents seuils d'intensité de pluie. L'originalité de la démarche utilisée consiste à prendre en compte les erreurs de la référence, en calculant les scores sur un ensemble d'observations perturbées.

Différents tests de sensibilité aux paramètres de la méthode ont été effectués : un test portant sur le choix du produit satellitaire, un deuxième concernant les hypothèses faites sur la distribution des erreurs du produit satellitaire et un dernier test sur le nombre de perturbations. Concernant ces deux derniers, les scores se sont révélés être robustes pour l'ensemble des tolérances spatiales et des seuils d'intensité, ce qui a rendu légitime de fixer par la suite des valeurs de référence pour ces paramètres. En ce qui concerne le choix du produit satellitaire, les scores se sont également révélés être stables sauf pour environ 15% des cas testés, principalement pour les pluies faibles et très intenses.

Cette nouvelle méthodologie de comparaison a pu être appliquée à des cumuls de pluie 6h issus des prévisions AROME et ALADIN, pour quatre échéances entre +6h et +24h. Les prévisions ALADIN utilisées sont celles produites en opérationnel; celles d'AROME sont celles du futur modèle opérationnel, initialisé avec IFS.

Une première étape consistait à comparer AROME, initialisé avec IFS, avec ALADIN. Elle a permis de faire ressortir la capacité d'AROME_IFS à prévoir des cumuls supérieurs à 50 mm/6h, et ce, dès les six premières heures d'échéance, par rapport à ALADIN dont les cumuls de pluie sont systématiquement plus faibles. Pour les seuils de pluie inférieurs, AROME_IFS obtient de meilleurs scores à mesure que les échéances augmentent. Ce résultat est lié, en partie, au fait que le modèle est utilisé avec des conditions initiales issues d'un autre modèle. Une mise en équilibre s'opère alors dès les premières échéances. Une seconde étape, comparant AROME initialisé cette fois-ci avec ALADIN à ALADIN, a permis d'arriver aux mêmes conclusions. Au final, la mise en place du prototype AROME Outre-Mer en remplacement d'ALADIN peut être justifiée par ces résultats.

Une comparaison a également été menée entre les prévisions d'AROME initialisé avec IFS et AROME initialisé avec ALADIN afin de tester les limites de la méthodologie en comparant des prévisions plus proches. Une plus grande sensibilité est apparue au choix du produit satellitaire pour les faibles et forts cumuls de pluie.

Cela montre les limites dans la formulation actuelle du modèle d'erreur des produits satellitaires, notamment en matière de détection de pluie (sensibilité plus forte au choix du produit pour le seuil 1mm) et en matière de biais pour les cumuls de pluie forts (sensibilité plus forte pour les cumuls de pluie supérieurs à 100 mm).

La méthode sera utile pour poursuivre le développement du démonstrateur AROME Outre-Mer, y compris sur d'autres domaines géographiques. De plus, la méthode étant générale, elle pourra être appliquée pour d'autres évaluations, par exemple concernant ARPEGE. En effet, l'utilisation du schéma de convection PCMT modifie de manière importante la distribution des pluies produites par ARPEGE et il serait intéressant de quantifier les différences en matière de pluies tropicales entre la chaîne ARPEGE opérationnelle et une chaîne utilisant PCMT.

Références

- Adler, R. et al., « Intercomparison of global precipitation products : The third precipitation intercomparison project (pip-3) », American Meteorological Society, Vol. 82, pp. 1377-1396, 2001
- Alizés, Journal de la DIRAG, n°64, septembre 2012
- Amodei, M. and Stein, J., « Deterministic and fuzzy verification methods for a hierarchy of numerical models », Met. Apps, Vol. 16, pp. 191-203, 2009
- Aonashi, K. et Liu, G., « Passive microwave precipitation retrievals using tmi during the baiu period of 1998. part i : Algorithm description and validation », J. Appl. Meteorol., Vol. 39, pp. 2024-2037, 2000
- Arkin, P. et Meisner, B., « The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-1984 », *Monthly Weather Review*, Vol. 115, pp. 51-74, 1987
- Barrett, E. et al., « The first wetnet precipitation inter-comparison project », Remote Sensing Reviews, Vol. 11, pp. 49-60, 1994
- Bouttier, F., « Arome, avenir de la prévision régionale », La Météorologie , n°58, août 2007
- Chambon, P. et al., « An investigation of the error budget of tropical rainfall accumulation derived from merged passive microwave and infrared satellite measurements », Q.J.R. Meteorol. Soc. , Vol. 139, pp. 879-893, 2013
- Chambon, P., « Contribution à l'estimation des précipitations tropicales : préparation aux misssions Megha-Tropiques et Global Precipitation Measurement », Thèse de doctorat de l'Université Paris-Est, 2011
- Ebert, E., « Fuzzy verification of high-resolution gridded forecasts : a review and proposed framework », *Meteorological Applications*, Vol. 15, pp. 51-64, 2008
- Ebert, E. et al, « Results from the gpcp algorithm intercomparison programme », Bulletin of the American Meteorological Society, Vol. 77, pp. 2875-2887, 1996
- Faure, G., « Un nouveau modèle de prévision à Météo-France : Aladin-Réunion », *La Météorologie*, n°60, février 2008
- Hou, A. Y. et al., « The Global Precipitation Measurement (GPM) Mission », Bulletin of the American Meteorological Society, 2013

- Huffman, G. et al., « The TRMM Multisatellite Precipitation Analysis (TMPA) : Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales », Journal of Hydrometeorology, Vol. 8, pp. 38-55, 2007
- Joyce, R. et al., « Cmorph : A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution », Journal of Hydrometeorology, Vol. 5, pp. 487-503, 2011
- Kidd, C. et Levizzani, V., « Status of satellite precipitation retrievals », Hydrology and Earth System Sciences, Vol. 15, pp. 1109-1116, 2004
- Roberts, N.M. et Lean, H.W., « Scale-selective verification of rainfall accumulations from highresolution forecasts of convective events », *Monthly Weather Review*, Vol. 136, pp. 78-97, 2008
- Roca, R. et al., « 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., Vol. 49, pp. 715-731, 2010
- Seity, Y. et al., « The AROME-France Convective-Scale Operational Model », Monthly Weather Review, Vol. 139, pp. 976-991, 2011
- Smith, E. et al., « Results of wetnet pip-2 project », J. Atmos. Sci., Vol. 55, pp. 1483-1536, 1998
- Ushio, T. et al., « A Kalman Filter Approach to the Global Satellite Mapping of Precipitation(GSMaP) from Combined Passive Microwave and Infrared Radiometric Data », *Journal* of the Meteorological Society of Japan, Vol. 87A, pp. 137-151, 2009

Acronymes

ALADIN : Aire Limitée, Adaptation dynamique, Développement InterNational **AMSR** : Advanced Microwave Scanning Radiometer AMSU-B : Advanced Microwave Sounding Unit-B **AROME** : Application de la Recherche à l'Opérationnel à Méso-Échelle **ARPEGE** : Action de Recherche Petite Échelle Grande Échelle **DMSP** : Defense Meteorological Satellite Program **FBS** : Fractions Brier Score **FSS** : Fractions Skill Score GCOM-W : Global Change Observation Mission - Water **GMAP** : Groupe de Modélisation et d'Assimilation pour la Prévision **GOES** : Geostationary Operational Environmental Satellites $\mathbf{GPROF}: \mathbf{Goddard} \ \mathbf{PROFiling} \ \mathbf{algorithm}$ **GSMaP** MVK : Global Satellite Mapping of Precipitation Motion Vector with Kalman filter **IDL** : Interactive Data Language **IFS** : Integrated Forecast System **IPWG** : International Precipitation Working Group **MetOP** : Meteorological Operational Polar **MHS** : Microwave Humidity Sounder **NASA** : National Aeronautics and Space Administration **NOAA** : National Oceanic and Atmospheric Administration **PCMT** : Prognostic Condensates Microphysics and Transport **SATMOS** : Service d'Archivage et de Traitement Météorologique des Observations Satellitaires **SSMI/S** : Special Sensor Microwave Imager Sounder **TMI** : TRMM Microwave Imager **TMPA** : TRMM Multisatellite Precipitation Analysis

TRMM : Tropical Rainfall Measuring Mission