

Le système d'information AMMA

Laurence Fleury⁽¹⁾, Nicole Asencio⁽²⁾, Jean-Luc Boichard⁽¹⁾,
Guillaume Brissebrat⁽¹⁾, Sophie Cloché⁽³⁾, Laurence Eymard⁽⁴⁾,
Florence Favot⁽²⁾, Laurence Mastrorillo⁽¹⁾, Oumarou Moulaye⁽⁵⁾,
Karim Ramage⁽³⁾ et Odile Roussot⁽²⁾

(1) SEDOO, OMP UMS 831, Toulouse, France

(2) GAME, URA CNRS/Météo-France, Toulouse, France

(3) ESPRI, CNRS/IPSL, Palaiseau, France

(4) LOCEAN, CNRS/IPSL, Paris, France

(5) CRA, AGRHYMET, Niamey, Niger

Contact : Laurence.Fleury@obs-mip.fr

Résumé

Le système d'information AMMA vise à faciliter la communication des données et des résultats scientifiques entre les différentes équipes et disciplines du programme, et à diffuser ces informations à l'ensemble de la communauté scientifique. Il est constitué de plusieurs bases de données – concernant les mesures des instruments sur le terrain, les produits satellitaires, les résultats de modèles, les articles scientifiques et les images utiles aux applications temps réel – et d'interfaces Internet permettant de consulter, sélectionner et télécharger ces données. Initialement développé par deux centres de données français, le système d'information est intégralement dupliqué et opérationnel au Niger.

Abstract

The AMMA information system

The AMMA information system aims at facilitating data and scientific results exchange between the AMMA research teams and domains. It also aims at distributing the information at the complete scientific community. It includes several databases associated with display, select and download Webtools. The databases manage field data, satellite products, model outputs, publications and real time quick looks. All the applications have been developed by two French data services and are fully duplicated and operated in Niger.

Le système d'information AMMA constitue un outil transverse aux différents thèmes du projet AMMA. Il rassemble les informations utiles ou produites par chacun d'entre eux et les structure de façon à faciliter les échanges entre les communautés de recherche. Ses composantes – sites d'affichage d'images produites régulièrement, bases de données scientifiques numériques ou qualitatives et base bibliographique – et les interfaces permettant d'y accéder ont été développées en utilisant exclusivement des logiciels et langages gratuits par les équipes ESPRI (Ensemble de services pour la recherche à l'IPSL, Institut Pierre-Simon Laplace) à Palaiseau et SEDOO (Service de données de l'OMP, Observatoire Midi-Pyrénées) à Toulouse. La totalité des données et des fonctionnalités est également opérationnelle au Centre régional AGRHYMET (CRA) de Niamey au Niger. L'ensemble du système est plus précisément décrit dans Fleury et al. (2011).

Sites d'images en temps réel

Un site Internet (<http://aoc.amma-international.org/>) a été mis en place en 2006 pour rassembler les nombreuses informations utiles aux décisions opérationnelles des périodes d'observation spéciales, notamment la planification et le suivi des vols des avions

instrumentés. Cet outil propose l'affichage d'images issues de modèles de prévision atmosphérique globaux ou régionaux, de produits satellitaires et d'observations *in situ* opérationnelles ou de recherche. Il permet également la consultation de documents, tels les rapports instrumentaux ou les comptes rendus des réunions biquotidiennes du centre de décision basé alors à Niamey. Conformément à l'objectif du site, la navigation est prioritairement basée sur la date : une page regroupe toutes les images d'un même produit à une date donnée, ces images étant mises à jour selon une fréquence prédéfinie qui va du quart d'heure (images satellitaires) au jour (mesures de terrain). L'alimentation du site se fait à l'aide de procédures automatiques ou directement par les producteurs des informations. Aujourd'hui encore, ce site fournit un moyen aisé de parcourir les différentes situations rencontrées pendant la campagne et, par exemple, de choisir des cas d'études.

Un site analogue (<http://fenoc.sedoo.fr/>) a été développé pour la campagne de l'été 2011 du projet FENNEC portant sur le système climatique du Sahara, notamment la dynamique de la dépression saharienne qui joue un rôle central dans la mousson d'Afrique de l'Ouest. Une autre application de ces sites d'images est de suivre l'évolution de la saison de mousson. Ainsi le site <http://isv.sedoo.fr/> permet-il à un groupe de scientifiques français et africains de partager et diffuser des

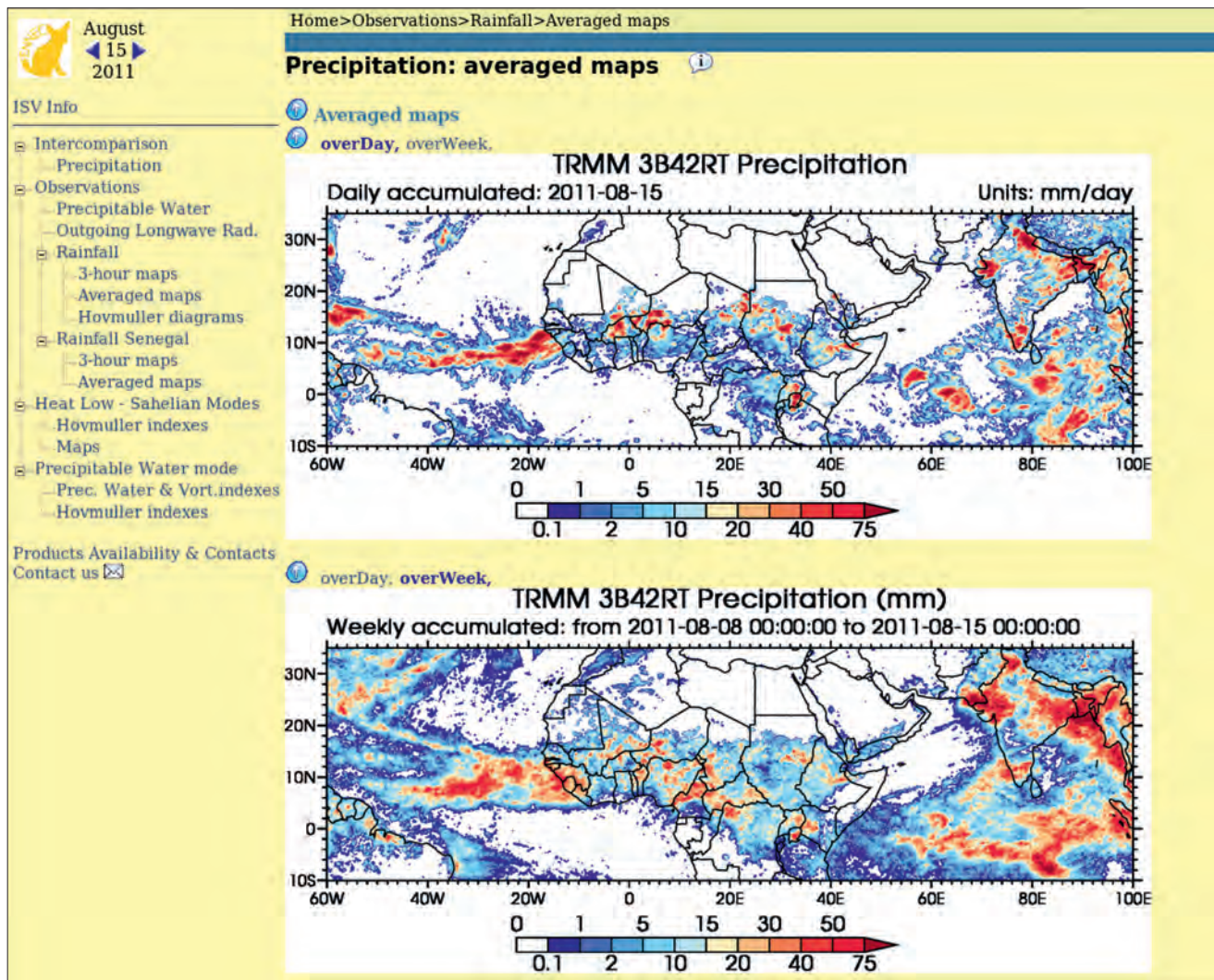


Figure 1 - Site de suivi de l'activité de la mousson au cours de la saison : page d'affichage de cumuls de précipitations issus d'observations satellitaires.

indicateurs caractéristiques de l'activité météorologique (voir figure 1). Initié en 2011, ce site sera opérationnel de mai à octobre en 2012 et 2013.

Bases de données scientifiques

La base de données AMMA a pour but de rassembler les jeux de données produits par les scientifiques du projet, mais aussi celles qu'ils utilisent pour conduire leurs études. Ces jeux de données émanent donc de toutes les thématiques du projet (agronomie, hydrologie, météorologie, océanographie, santé, sciences humaines...). Ils sont issus d'observations et d'enquêtes sur le terrain, de réseaux de mesures opérationnels, de mesures satellitaires, de simulations numériques ou de travaux croisant plusieurs sources d'information. À l'heure actuelle, les données AMMA représentent un volume de plus de 20 téraoctets composé de :

- plus de 200 jeux de données de terrain, correspondant aux instruments du programme AMMA pendant ses différentes périodes d'observation, mais aussi aux mesures effectuées par des réseaux opérationnels météorologiques et hydrologiques actuels ou historiques, des services d'observation et des campagnes intensives antérieures ;
- 1 350 réponses à un questionnaire de socio-économie portant sur l'adaptation au changement climatique à l'échelle de l'unité de production agricole ;
- environ 60 produits satellitaires opérationnels couvrant la période de 1980 à nos jours et plusieurs produits recherche élaborés dans le cadre de AMMA comme par exemple EPSAT-SG (Estimation des pluies par satellite-seconde génération, Bergès et al., 2010) ;
- 10 ensembles de sorties de modèles opérationnels météorologiques et océaniques, ainsi que des résultats de simulations effectuées dans le cadre du programme AMMA, dont une réanalyse dédiée du Centre européen pour les prévisions météorologiques à moyen terme et les résultats de l'exercice de

comparaison des modèles de surface (AMMA Land surface Model Inter-comparison Program).

D'origines très diverses, ces jeux de données ont tous été documentés dans

Les métadonnées

Les métadonnées sont l'ensemble des informations permettant de décrire les jeux de données. Ainsi une donnée géographique est caractérisée par sa localisation, sa date, le nom du paramètre, l'instrument de mesure, le numéro du capteur, le nom du responsable, etc. L'ensemble des informations nécessaires à la description des données géophysiques fait l'objet de normes (ISO 19115) et de recommandations au niveau européen (directive INSPIRE). De plus, l'utilisation de thésaurus ou listes hiérarchiques de mots permet de référencer les données avec un vocabulaire standard. Cette taxonomie porte notamment sur les disciplines, les paramètres et les instruments.

une même base de métadonnées (voir l'encadré « Les métadonnées ») et classés dans un catalogue selon un thésaurus reconnu (Global Change Master Directory, <http://gcmd.nasa.gov/>). Mais plusieurs bases de données ont été mises en place selon la « dimension » des données. Ainsi, les séries temporelles d'un paramètre en un point ou le long de la trajectoire d'une plateforme mobile (avion, ballon, bateau...) sont considérées comme monodimensionnelles et ont été intégrées dans une base de données relationnelle spécifique (voir l'encadré « Avantages d'une base de données relationnelle »). Les données bi- ou tridimensionnelles (cartes, produits satellitaires, résultats de simulations) ont toutes été projetées sur des grilles latitude/longitude carrées et converties dans un même format NetCDF, (Network Common Data Form, <http://www.unidata.ucar.edu/software/netcdf/>). Les fichiers résultants ont été nommés selon une nomenclature prédéfinie et placés dans une archive structurée. Enfin, une base relationnelle spécifique a été développée pour gérer les réponses numériques et textuelles à un questionnaire de socio-économie.

Ces différentes bases de données sont accessibles depuis une même interface Internet facile d'utilisation décrite dans Fleury et al. (2010). Elle permet aux

utilisateurs de commander les données à partir de différents critères : zone géographique, période, paramètre, etc. (voir figure 2). La requête est envoyée à toutes les bases de données et l'utilisateur peut donc à partir d'une seule

Avantages d'une base de données relationnelle

Un modèle de données relationnel répartit les données dans différentes tables liées par des relations algébriques. Cette structure permet de ne stocker qu'une fois une même information, ce qui économise de la place et évite les erreurs de saisie. Par exemple, dans la table contenant les paramètres, les différents enregistrements de type température (température de l'air en surface, température océanique, etc.) seront tous liés à un seul enregistrement de la table des unités : le degré Celsius. D'autre part, cette structuration permet de formuler des requêtes très complexes portant sur la satisfaction de différents critères simultanément. Dans la base de données AMMA, il est par exemple possible de rechercher toutes les données où, dans une zone de l'espace et un intervalle de temps, un paramètre excède un certain seuil.

opération obtenir des fichiers de données mono-, bi- ou tridimensionnelles. Le système de gestion et distribution de données est schématisé en figure 3 : en France (<http://database.amma-international.org/>) les bases sont réparties entre les services ESPRI et SEDOO ; au Niger l'ensemble des bases et des fonctionnalités est géré au CRA (<http://amma.agrhymet.ne/>). Les deux systèmes sont synchronisés hebdomadairement quant au contenu des bases et chacun est parfaitement autonome quant aux fonctions de distribution de données et d'administration.

Dans la mise en place de ces outils, une attention particulière a été prêtée à l'échange de données entre les chercheurs de différentes disciplines ou utilisant des outils différents. Ainsi, l'existence d'une base relationnelle de métadonnées unique respectant les standards internationaux et d'une interface permettant d'interroger l'ensemble des bases de données permet à tous d'accéder à la totalité des jeux de données. D'autre part, la projection sur des grilles régulières des produits satellitaires ou de modélisation et leur conversion en NetCDF ont permis de mettre à la disposition des utilisateurs non experts des produits bien plus simples à manipuler que les produits natifs, ces derniers étant néanmoins disponibles pour les utilisateurs avertis.

Figure 2 - Interface de requête des données AMMA : onglet de sélection selon les critères géographiques et temporels.

The screenshot displays the AMMA data query interface. At the top, there is a section titled "Types de données" with three checked options: "Données In Situ", "Données Satellite", and "Données Modèle". Below this, a navigation bar includes tabs for "Zones et période", "Jeux et Plateformes", "Satellites", "Modèles", "Paramètres", "Options", "Valider la requête", and "Suivi des requêtes In Situ". The "Zones et période" section is active, showing a "Période" (Period) selection with "Début" (Start) set to 01/07/2006 and "Fin" (End) set to 15/09/2006. Below the period selection, there is a "Zones prédéfinies" (Predefined zones) section with a tree view showing "Sub-regional windows", "Meso-scale Sites" (with sub-items: Gourma, Niamey, Oueme), and "Super sites". A "Ré-initialiser" (Reset) button is located at the bottom of this section. The central part of the interface features a "Carte" (Map) showing a satellite-style map of West Africa, with a "Global" button below it. To the right of the map is a "Zoom" section with input fields for latitude and longitude. The top coordinates are Lat: 25° 54' 0" and Lon: -24° 42' 1". A zoomed-in area is highlighted with a dashed box, showing coordinates Lat: -3° 11' 57" and Lon: 26° 0' 0". "Unzoom" and "Apply" buttons are at the bottom of the zoom section.

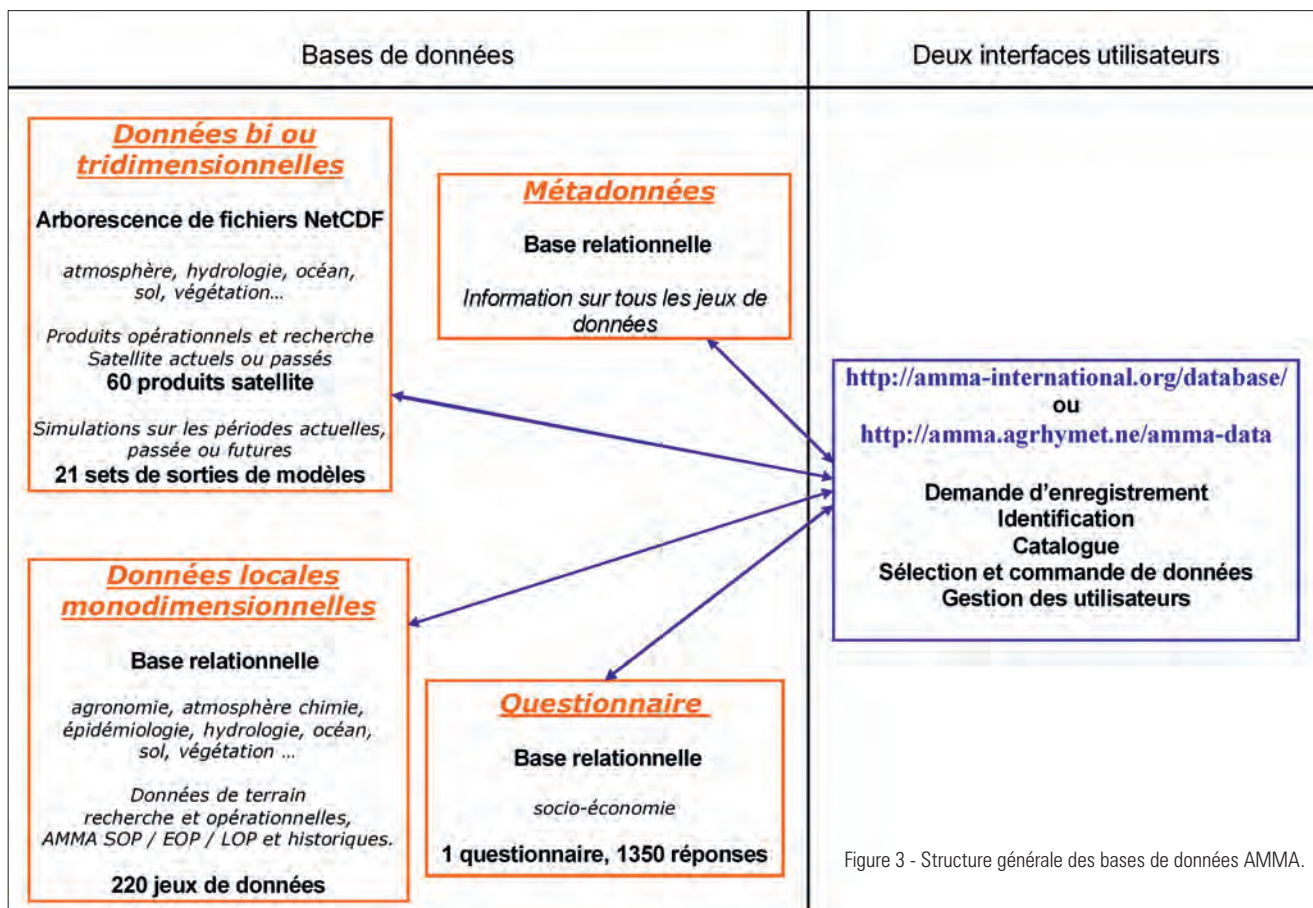


Figure 3 - Structure générale des bases de données AMMA.

Concernant les données de terrain, l'insertion dans une base relationnelle permet d'homogénéiser les jeux de données (nom des paramètres, unités...) et de proposer deux formats standardisés en réponse aux requêtes : un format de type CSV compatible avec les éditeurs de textes ou les tableurs et le format NetCDF compatible avec de nombreux logiciels graphiques.

Afin de favoriser l'échange de données, des règles de bon usage assurent le respect des droits de propriété intellectuelle des personnes, projets et organismes à l'origine des données. Tout utilisateur souhaitant accéder aux données doit signer en ligne ces règles. Il s'engage notamment à ne pas redistribuer ni faire un quelconque usage commercial des données et à proposer aux personnes ayant produit les données de collaborer à ses travaux scientifiques. Dans tous les cas, l'utilisateur s'engage à remercier le fournisseur de données dans ses publications. Le système de données AMMA gère donc également un annuaire des utilisateurs et garde une trace de tous les accès aux données. Aujourd'hui, l'annuaire compte environ 700 utilisateurs enregistrés, dont 24 % sont africains, 35 % français et 20 % européens non français. Les interfaces d'accès traitent une centaine de demandes de données par mois.

Base bibliographique

Pendant la première phase du programme AMMA, le portail collaboratif du programme européen permettait de mettre à la disposition de ses participants les publications dès leur soumission. Vecteur de communication et de collaboration dans la communauté, cette procédure a permis de rassembler un grand nombre de publications. Afin d'étendre cette dynamique de partage des publications à l'ensemble de la communauté AMMA et de préserver cette collection d'articles, un outil collaboratif basé sur le logiciel libre WIKINDX a été mis en place (<http://biblio.amma-international.org/> et <http://amma.agrhymet.ne/ammaBiblio/>). Il regroupe l'ensemble de la production scientifique du projet en termes d'articles, de thèses, de rapports techniques, de posters et de communications, mais aussi un grand nombre d'articles de référence concernant la mousson africaine. Le catalogue et les résumés d'articles sont accessibles à tous ; les textes intégraux sont soit accessibles à tous, soit réservés aux utilisateurs enregistrés, selon la politique éditoriale des journaux.

Actuellement la base bibliographique compte environ 1 200 références et constitue la plus grande collection de

documents scientifiques sur la mousson africaine. Cet outil public contribue au rayonnement du programme AMMA.

Conclusion

Aujourd'hui, le système d'information AMMA représente un héritage concret du programme tant en termes de contenu que de fonctionnalités. De nouveaux projets l'ont déjà choisi pour gérer leur propre production et contribuent à l'enrichir : phase 2 du programme AMMA aux niveaux national et international, projet ESCAPE (Environmental and Social Changes in Africa: Past, present and future) financé par l'Agence nationale de la recherche (ANR), projet FENNEC soutenu entre autres par l'ANR et le NERC (Natural Environment Research Council, Grande-Bretagne), projet WAMME2 (West African Monsoon Modeling and Evaluation)... Des rapprochements avec d'autres projets en cours ou à venir et l'intégration de données historiques se poursuivent. Déjà reconnu comme un outil efficace d'accès aux données et aux informations scientifiques, le système d'information AMMA a vocation à se développer encore et à devenir une référence sur l'Afrique de l'Ouest pour l'ensemble de la communauté scientifique.