The ECMWF re-analysis for the AMMA observational campaign

Anna Agustí-Panareda,[a]* Anton Beljaars,[a] Maike Ahlgrimm,[a] Gianpaolo Balsamo,[a] Olivier Bock,[b] Richard Forbes,[a] Anna Ghelli,[a] Françoise Guichard,[c] Martin Köhler,[a] Remi Meynadier,[b] and Jean-Jacques Morcrette[a]

[a]*ECMWF, Reading, UK*
[b] *LATMOS, CNRS–Univ. Paris VI, Paris, France*
[c] *CNRM, CNRS and Météo-France, Toulouse, France*
*Correspondence to: A. Agustí-Panareda, ECMWF, Shinfield Park, Reading, Berkshire, RG2 9AX, UK. E-mail: Anna.Agusti-Panareda@ecmwf.int

During the 2006 African Monsoon Multidisciplinary Analysis (AMMA) field experiment, an unprecedented number of soundings were performed in West Africa. However, due to technical problems many of these soundings did not reach the Global Telecommunication System and therefore they could not be included in the operational numerical weather prediction (NWP) analyses. This issue, together with the realization that there was a significant bias in the radiosonde humidity, led to the conclusion that a re-analysis effort was necessary. This re-analysis was performed at the European Centre for Medium-Range Weather Forecasts (ECMWF) spanning the wet monsoon season of 2006 from May–September. The key features of the ECMWF AMMA re-analysis are presented, including the use of a newer model version with improved physics, all the AMMA radiosonde data available from the AMMA database and a new radiosonde humidity bias-correction scheme. Data-impact experiments show that there is a benefit from these observations, but also highlight large model physics biases over the Sahel that cause a short-lived impact of the observations on the model forecast. The AMMA re-analysis is compared with independent observations to investigate the biases in the different parts of the physics. In the framework of the AMMA project, a hybrid dataset was developed to provide a best estimate of the different terms of the water cycle. This hybrid dataset is used to evaluate the improvement achieved from the use of extra AMMA observations and of a radiosonde humidity bias-correction scheme in the water cycle of the West African monsoon. Finally, future model developments that offer promising improvements in the water cycle are discussed.

Key Words: West African monsoon; water cycle; energy budget; radiosonde

Received 18 December 2009; Revised 7 May 2010; Accepted 26 May 2010; Published online in Wiley Online Library 31 August 2010

Citation: Agustí-Panareda A, Beljaars A, Ahlgrimm M, Balsamo G, Bock O, Forbes R, Ghelli A, Guichard F, Köhler M, Meynadier R, Morcrette J-J. 2010. The ECMWF re-analysis for the AMMA observational campaign. *Q. J. R. Meteorol. Soc.* **136**: 1457–1472. DOI:10.1002/qj.662

1. Introduction

Defining the state of the atmosphere as an initial condition for forecasts is an important aspect of numerical weather prediction (NWP). NWP systems use a forecast model to propagate the state of the atmosphere in time and continuously feed in observations to obtain so-called analyses. A well-designed analysis system obtains an

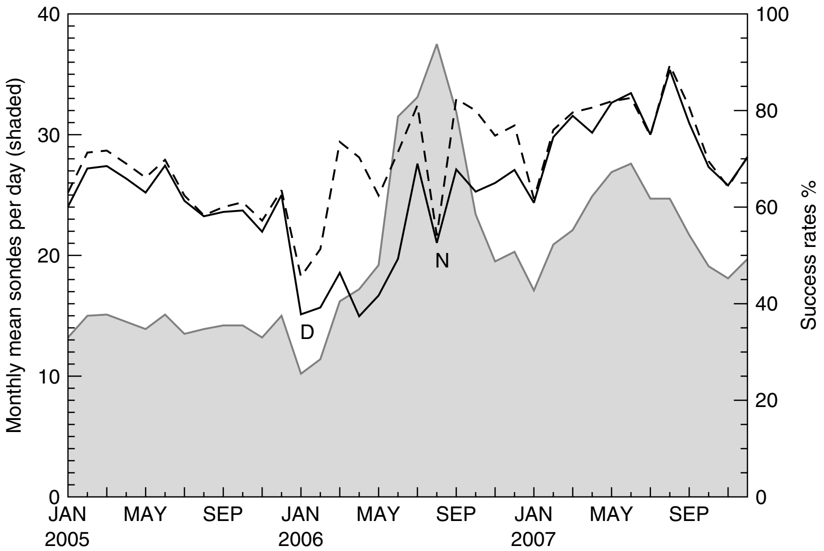

Figure 1. Shaded area: number of soundings (monthly mean sondes per day) acquired operationally by ECMWF from the AMMA network from January 2005–November 2007. Black solid line: percentage success rate of data reception for the 21 primary stations in the network. Black dashed line: percentage success rate excluding the four stations with no direct GTS link that used satellite and e-mail transmission. 'D' identifies the effects of a GTS failure at Dakar, while 'N' denotes lightning damage at Niamey, which interrupted transmission for several stations (from Parker, *et al.*, 2008). © American Meteorological Society. Reprinted with permission.

optimal estimate of the atmospheric state as a blend between a short-range forecast (information propagated from previous observations) and current observations. The observations are of different types, ranging from conventional observations such as radiosondes, pilot balloons and surface observations (SYNOPs) to aircraft observations and satellite data. Although satellite data are becoming increasingly important, they are still predominantly limited to cloud-free points, and over land no channels are used that have their peak sensitivity in the lower troposphere, due to uncertainty in the radiance contribution from the surface. Therefore radiosondes are still the dominant information source to define the thermodynamic and dynamic state of the atmosphere.

These atmospheric analyses, as a side product of NWP, give a consistent description of the atmosphere in time and space on a defined grid. NWP analyses aim to be consistent with all the available observations and therefore are very popular for research, climate monitoring and diagnostic studies.

As part of the African Monsoon Multidisciplinary Analysis (AMMA) project, a lot of effort was put into improving and enhancing the radiosonde network (Parker, *et al.*, 2008) with several special and intensive observing periods (SOP and IOP) in 2006 (Redelsperger, *et al.*, 2006). Special attention was paid to the telecommunication network (the global telecommunication system (GTS)) to ensure that all the observations would reach the operational NWP centres in real time. The European Centre for Medium-Range Weather Forecasts (ECMWF) was tasked with the monitoring of the AMMA radiosondes to ensure quality and to provide timely feedback on observation problems. Figure 1 shows that many more data were assimilated in the ECMWF operational NWP analysis in 2006 than in 2005. However, there were also intermittent failures of the data communication system, e.g. due to equipment failure, resulting in data loss. For example, the GTS link did not work for all stations, particularly some of the new and revived AMMA stations with high-frequency soundings. This is depicted in Figure 1 by the large deviation between the percentage success rate with and without the four stations with no direct GTS link in 2006. In 2007, both e-mail and satellite transmissions were working well for those stations (see Parker, *et al.*, 2008, for further details), leading to increased success in the reception of soundings.

It was therefore decided to create a special archive of AMMA radiosondes and to re-run the ECMWF data assimilation and forecasting system for the period of the AMMA field experiment in 2006. Because re-analysis is rather heavy on human and computer resources (at ECMWF a research data assimilation stream runs typically one day per day), it was decided to limit it to the period 1 May–30 September 2006. The advantages of the AMMA re-analysis are as follows:

- The maximum possible data coverage can be obtained, because data availability is not limited any more to real-time observations transmitted through the GTS.
- The observations are made at high vertical resolution, but only transmitted through the GTS at low resolution (so-called standard and characteristic levels). The re-analysis provides the opportunity to insert the full observed vertical resolution into the system.
- Since 2006, ECMWF has made substantial improvements to the system, in particular to the model formulation (e.g. the convection scheme (Bechtold, *et al.*, 2008), the land-surface hydrology, (Balsamo, *et al.*, 2009a) and the radiation code (Morcrette, *et al.*, 2008)). These model changes are particularly relevant for the Tropics.
- Moisture-budget studies are an important application for analysis data, but it turns out that a number of sonde stations show substantial biases in the moisture observations. Therefore it was decided to develop a bias-correction scheme (Agustí-Panareda, *et al.*, 2009a). This bias-correction scheme is used in the re-analysis.
- A substantial model change was implemented on 12 September 2006 with the introduction of a new cycle

(CY31R1, see section 2 for more details) that makes the operational analysis less consistent across this date.

The purpose of this article is to give a description of the ECMWF AMMA re-analysis, including the main improvements achieved by using an increased number of radiosonde data and a radiosonde bias-correction scheme, as well as documenting the errors in the energy and water budgets. A brief description of the data assimilation system is given in section 2. Section 3 presents the sounding data from West Africa used in the AMMA re-analysis. An evaluation of the re-analysis is provided in section 4 for different aspects of the model physics using the available observations and other elaborated products from AMMA. Finally, the biases found in the evaluation are discussed in section 5 and future model developments with expected impact on those biases are also presented. A summary of the main findings is given in section 6.

2. The data assimilation system

The ECMWF data assimilation and forecasting system is called the Integrated Forecasting System (IFS). It relies heavily on a forecast model to propagate the state of the atmosphere in time. For the re-analysis, $T511$ resolution is used (40 km resolution in grid-point space) with 91 vertical levels. The lowest model level is at about 10 m above the surface and the top of the model is at 0.01 hPa. The distribution of model levels is given in: http://www.ecmwf.int/products/data/technical/model_levels/model_def_91.html. The model is a state of the art spectral model with a comprehensive physics package to describe subgrid processes. Full model documentation is given in http://www.ecmwf.int/research/ifsdocs/CY31r1/index.html.

The AMMA re-analysis uses the CY32R3 version of the system, which was operational between 6 November 2007 and 3 June 2008. ECMWF improves the operational system on a regular basis, so the most recent version of the system was selected at the start of the re-analysis. The main model changes with respect to the model operational during the summer of 2006 are (i) introduction of ice supersaturation (Tompkins, *et al.*, 2007), (ii) new short-wave radiation scheme and introduction of McICA (Morcrette, *et al.*, 2008), (iii) new land-surface hydrology (Balsamo, *et al.*, 2009a) and (iv) convection entrainment closure based on relative humidity rather than moisture convergence (Bechtold, *et al.*, 2008). Furthermore, major changes were made to the configuration of the data assimilation system (three outer loops instead of two, with resolutions $T95$, $T159$ and $T255$) and a more comprehensive physics package was included for the linear version of the model. Also changes were made to the assimilation of satellite radiances. The full list of changes that were made since the AMMA campaign in 2006 can be found in http://www.ecmwf.int/products/data/technical/model_id/index.html.

The data assimilation system is a four-dimensional variational analysis (4D-Var) system working in 12 h time windows (Rabier, *et al.*, 2000; see Andersson and Thépaut, 2008 for a description). The basic principle is to perform a model integration over a 12 h interval, to evaluate the distance of the model trajectory to the observations (J_o) and to adjust the initial condition iteratively in such a way that a cost function is minimized (Figure 2). As well as the observation term J_o, the cost function also has a background term (J_b) representing the distance of the initial condition to the previous forecast or background, and it includes a balance constraint. The solution is a weighted mean of the observations and model background and the weights are given by the background-error covariance matrix and by an estimate of observation-error variances.

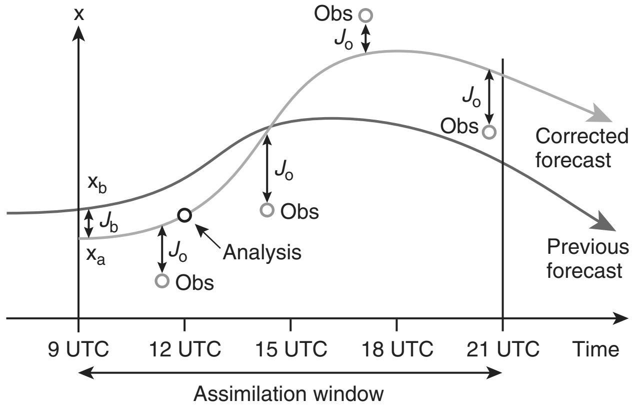

Figure 2. Illustration of 4D-Var data assimilation in 12 h windows (updated from Andersson and Thépaut, 2008).

It is worth noting that the distance to observations is predominantly evaluated in observation space, so for satellite observations a forward operator is used to convert model profiles into radiances, which allows direct comparison with the observed radiances. For many satellite channels, bias models (e.g. biases dependent on scan angle or air mass) are included in the forward model. A limited number of bias coefficients are included in the variational optimization to correct for biases. Radiosondes anchor the system, so they are a crucial component of the analysis system.

3. Sounding data from West Africa used in the AMMA re-analysis

The AMMA field experiment has provided the largest number of sounding data ever recorded in West Africa during the period of the wet monsoon in 2006, even more than during the GARP Atlantic Tropical Experiment (GATE) in 1974. All these data have been collected in the AMMA database. The AMMA re-analysis covers the period from 1 May–30 September 2006. This choice was a compromise between the wish to cover as much of the pre-monsoon and monsoon seasons as possible and the limitations on computer time. Before starting the re-analysis, it was necessary to complement the data that were not available from the ECMWF archive (as filled real time through the GTS) by data directly from the stations. For this purpose, the AMMA archiving group acquired all the available AMMA sondes at the highest possible resolution. After this, ECMWF retrieved the data from the AMMA archive and coded it into BUFR.†

The BUFR-coded data were also made available to Météo-France and NCEP where they were used in their analysis efforts. The AMMA database for this period includes the following:

†BUFR is a WMO standard format for coding observations that can be used in most data assimilation systems.

- 6063 high-resolution radiosondes/dropsondes collected from 21 stations, 3 research vessels and 2 research aircrafts. These have been thinned from about 2500 to approximately 300 vertical levels.
- Radiosondes launched from operational stations, research stations, vessels and research aircrafts via the GTS. The radiosonde data from West Africa typically contain 70–100 levels. These data are only used when there are no corresponding high-resolution data available.
- 101 dropsondes from research aircrafts obtained via GTS.
- 110 dropsondes from gondolas, also known as driftsondes. The development and deployment of the driftsonde system was a collaborative effort between the Earth Observing Laboratory (EOL/NCAR) and the French Space Agency (CNES) as part of the SOP3 period to investigate the development of tropical cyclogenesis downstream of West Africa. It is the first time these will be assimilated in an analysis experiment. Preliminary comparison with operational analysis showed good agreement.
- 7317 pilot balloons measuring wind profiles obtained via the GTS.

Station information is given in Table I, and the location of the AMMA stations as well as their total number of soundings used by the analysis is illustrated in Figure 3. Niamey (13.48°N, 2.17°E) is the station that has the highest number of soundings. The mobile Atmospheric Radiation Measurement (ARM) site was also deployed in Niamey (Niger) in 2006. Figure 4 displays a time series of the number of daily soundings from the AMMA database (solid line) and GTS (dotted line) spanning the AMMA re-analysis period. The period covers the special observing periods (SOP) dedicated to the monsoon pre-onset (SOP1) and the monsoon onset and peak (SOP2), as well as part of SOP3 which focused on the development of tropical cyclones downstream of West Africa over the Atlantic. During the SOPs there were intensive observing periods (IOPs) of 1–4 days focusing on specific events of the monsoon. IOPs are classified into patterns depending on the area or type of event covered. During some of the IOPs, intensive regional observations included the launching of eight radiosondes per day at six stations as shown in Table I. This happened during two periods within SOP1 and SOP2 from 20–29 June and 1–15 August. Figure 4 depicts these observing periods at Niamey.

It was clear from the beginning of the radiosonde monitoring at ECMWF that some of the soundings had large biases in the humidity. This was also confirmed with the help of independent Global Positioning System (GPS) total column water vapour (TCWV) estimates from six AMMA GPS sites (Bock, *et al.*, 2008). Thus, a bias-correction scheme was developed with the main purpose of applying it to the AMMA soundings used in the re-analysis. A full description of the bias-correction method, its testing and impact is described by Agustí-Panareda, *et al.* (2009a). After applying the bias correction, a better match with GPS data was obtained. The impact of the bias correction on the analysis turned out to be substantial, as shown by the TCWV increase at many places in West Africa, resulting in increased convective available potential energy

Table I. Radiosonde network during August 2006 monitored in AMMA.

Station name	WMO station ID	Lat (°N)	Lon (°E)	Altitude (m)	Frequency of AMMA soundings planned	High resolution data per day	pre-AMMA soundings per day
Sal	08594	16.73	−22.95	53	1	No	1
Tamanrasset	60680	22.80	5.43	1364	4	No	2
Agadez	61024	16.97	7.98	502	4 (8)	Yes	1
Niamey	61052	13.48	2.17	227	4 (8)	Yes	2
Tombouctou	61223	16.72	−3.00	264	2	Yes	0
Bamako/Senou	61291	12.53	−7.95	381	2	Yes	2
Nouadhibou	61415	20.93	−17.03	3	1	Yes	0
Nouakchott	61442	18.10	−15.95	3	1	Yes	0
Dakar/Yoff	61641	14.73	−17.50	24	2	Yes	2
Tambacounda	61687	13.77	−13.68	50	1	Yes	0
Conakry	61831	9.56	−13.61	48	1	Yes	0
Addis Ababa-Bole	63450	9.03	38.75	2354	1	No	1
Bangui	64650	4.40	18.52	366	2	Yes	0
N'Djamena	64700	12.13	15.03	295	2	Yes	0
Ngaoundere	64870	7.35	13.57	1104	1	Yes	0
Douala R.S	64910	4.02	9.70	15	2	Yes	2
Abuja	65125	9.25	7.00	344	4 (8)	Yes	0
Parakou	65330	9.35	2.62	393	4 (8)	Yes	0
Cotonou	65344	6.35	2.38	9	4 (8)	Yes	0
Tamale	65418	9.50	−0.85	173	4 (8)	Yes	0
Ouagadougou	65503	12.35	−1.52	306	2	Yes	1
Abidjan	65578	5.25	−3.93	8	2	Yes	0

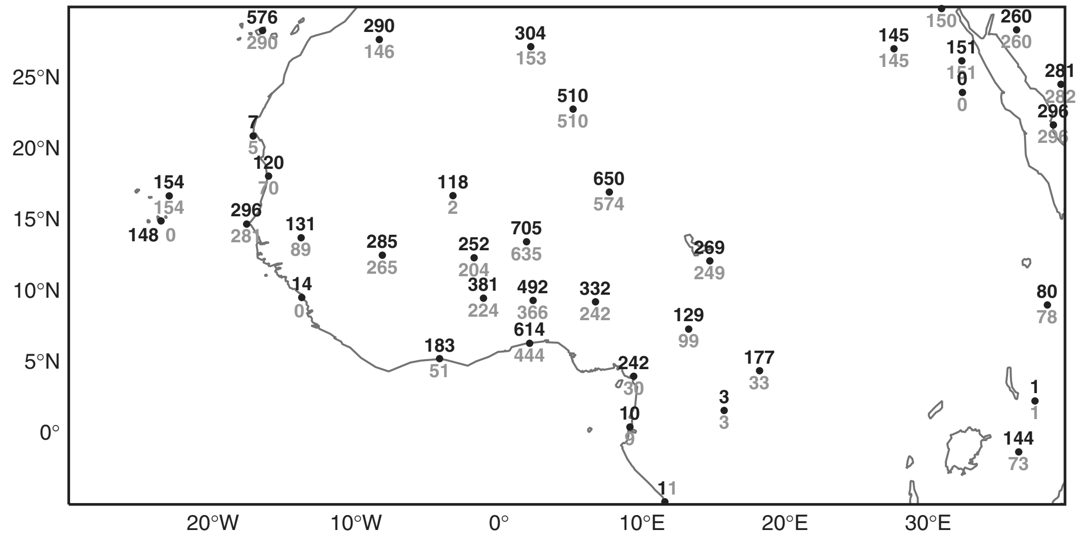

Figure 3. Total number of radiosonde soundings used in the AMMA re-analysis (black) and operational analysis (grey) from 1 May–30 September 2006.

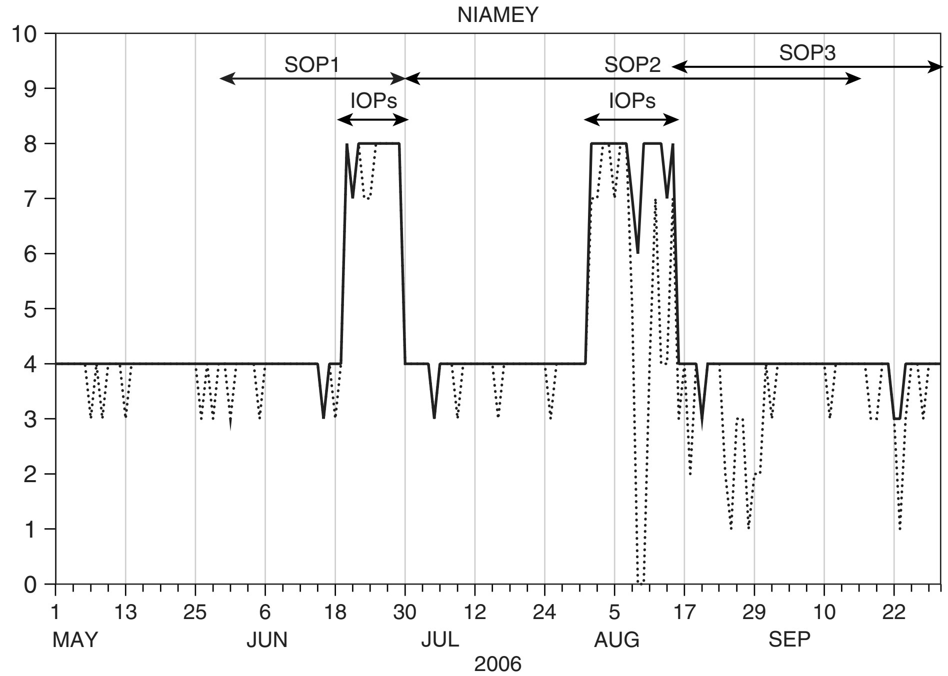

Figure 4. Number of soundings per day used from Niamey airport in the operational (dotted line) and AMMA (solid line) analysis from 1 May–30 September 2006. The different special observing periods (SOP) and intensive observing periods are highlighted with arrows.

(CAPE) and precipitation (see figures 12, 14 and 16 from Agustí-Panareda, *et al.* (2009a)).

4. Evaluation of physical processes in the AMMA re-analysis

Studies of the impact of the AMMA radiosonde data and bias-correction scheme on the AMMA re-analysis were performed by Agustí-Panareda, *et al.* (2009a, 2010a). These demonstrated the benefit of having both an enhanced radiosonde network and a correction for the radiosonde humidity bias on the wind, temperature and humidity analyses as well as the short-range precipitation forecast for the West African monsoon in August 2006. However, they also showed large systematic errors in precipitation, boundary-layer temperature, humidity and the monsoon flow over the Sahel. Previous work by Guichard (2009) and Tompkins, *et al.* (2005) suggests that these model biases can be largely attributed to radiation biases caused by aerosol and cloud biases, as well as deficiencies in the current land-surface parametrization.

In this section, an assessment of the model physical biases is performed for the AMMA re-analysis. It makes use of several independent datasets characterizing aerosols, surface radiation and heat fluxes at local scales, as well as cloud, precipitation and evapotranspiration at larger scales. The independent observations are from radiometers at the ARM-mobile facility, CALIPSO, CLOUDSAT, TRMM and AERONET, as well as from a hybrid dataset for the water cycle elaborated by Meynadier, *et al.* (2010a). This hybrid water budget allows inferences to be drawn on the vertically integrated moisture-flux convergence. The evaluation of the water cycle in the AMMA re-analysis is based on the

Table II. Description of the experiment configuration. The radiosonde network used is either the enhanced AMMA network with data collected from the AMMA database at high vertical resolution for most stations, or the pre-AMMA network by only using data from the GTS from stations reporting reliably in 2005 (see Table I and Figure 3 for further details).

Experiment name	Radiosonde network	Radiosonde humidity bias correction applied
AMMA	AMMA	Yes
PREAMMA	pre-AMMA	Yes
NOBIASCOR	AMMA	No

comparison of three analysis experiments that test the impact of the AMMA radiosonde data in the IFS (see Table II and section 3).

4.1. Cloud

The active lidar and radar instruments on board the CALIPSO and CloudSat satellites, which fly as part of the A-Train constellation (Stephens, *et al.*, 2002), provide an opportunity to evaluate the vertical profile of cloud and precipitation occurrence across West Africa. Due to the narrow footprint of the instruments (1–2 km) and the configuration of the orbit of the satellites with only one or two tracks across West Africa every day, the sampling is sparse in the east–west direction but very high-resolution in the north–south direction. Averaging all tracks between 10°W and 10°E for the whole of August provides a more robust statistical assessment of the meridional variation in the model cloud/precipitation across the region. Figure 5 shows the zonal cross-sections between the Equator and

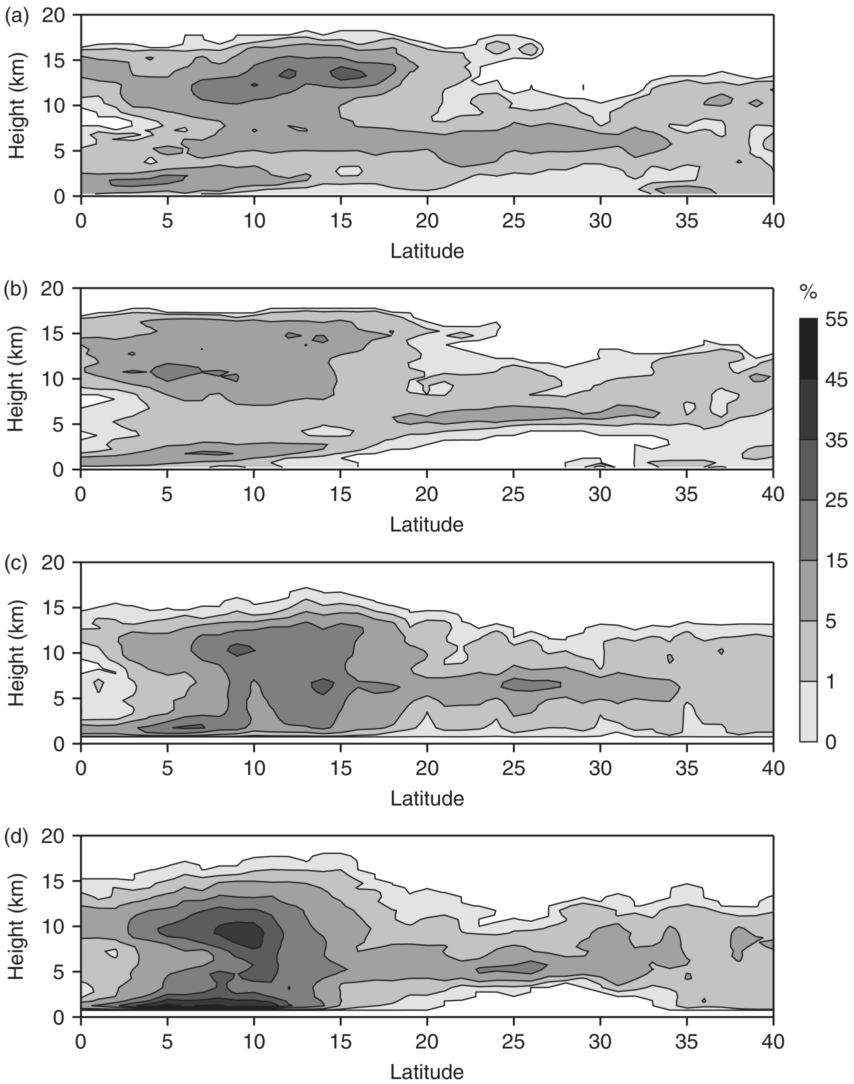

Figure 5. North–south cross-section of 10°W–10°E zonal average frequency of cloud/precipitation occurrence during August 2006 derived from (a) observed backscatter from the CALIPSO lidar, (b) modelled backscatter from the IFS model along the CALIPSO track, (c) observed radar reflectivity from CloudSat and (d) modelled radar reflectivity from the IFS model along the CloudSat track. All data have been binned into 1° latitude bins and 500 m vertical height bins. The lowest 1 km in (c) and (d) has been masked out due to contamination from the surface backscatter signal.

40°N for the CALIPSO observations of cloud occurrence derived from lidar backscatter, the CloudSat observations of cloud and precipitation occurrence derived from radar reflectivity and the model equivalents calculated from lidar and radar forward models.

Model data are extracted along the satellite's track from daily forecasts from the AMMA re-analysis at $T511$ resolution. Three-hourly outputs from the 12–36 h forecast range are stitched together to provide a series of vertical profiles of model data along the satellite track that are always within 1.5 h of the time of observations.

The CloudSat 94 GHz radar observes cloud and precipitation particles above a certain size threshold. Large cloud ice, snow and precipitation particles result in a large observed reflectivity, while ice clouds with small particles or warm clouds with small droplet size may be missed. Since the particle sizes range smoothly between cloud ice and precipitation, it is impossible to separate radar returns from cloud only from those of frozen precipitation. To facilitate a fair comparison, a radar forward model is applied to the IFS cloud and precipitation fields to simulate the reflectivity that the CloudSat radar would observe. Each model column is divided into 20 subcolumns and a maximum random cloud overlap is applied. Then the attenuated radar reflectivity is calculated from the 'in-cloud' values for cloud liquid, cloud ice and rain and snow precipitation in each subcolumn using the model cloud and precipitation subgrid fraction. The proportion of the grid box that is assumed to be cloud-covered is predicted directly by the model, which includes stratiform cloud cover detrained from convection (Tiedtke, 1993), but the subgrid precipitation fractions are diagnosed. The stratiform precipitation fraction is determined by the maximum cloud fraction in the column above, but the convective precipitation fraction is assumed to be a fixed value of 0.05 (i.e. 5% of the grid box is covered by convective cores). The sensitivity threshold of the CloudSat radar is approximately −30 dBZ, and so only model reflectivities exceeding this threshold are included in the model cloud/precipitation occurrence cross-section comparison (Figure 5(c) and (d)). The higher frequency of occurrence of hydrometeors seen in the model (Figure 5(d)) at low altitudes (0–2 km) between 0° and 12°N is at least partly due to an overprediction of drizzle in the model shallow convective clouds, which results in model-derived radar reflectivities above the sensitivity threshold of the radar. In reality, many of these low liquid-water clouds are not drizzling and thus remain below the sensitivity threshold of the CloudSat radar (Figure 5(c)).

The CALIPSO lidar is sensitive to small particles in the atmosphere and can detect optically thin features, such as aerosol layers and subvisible cirrus clouds. However, the lidar's signal is fully attenuated in clouds with optical depth exceeding approximately 3. In the case of deep convective clouds, for example, only the top of the cloud will be observed, while all clouds underneath the level of full signal attenuation are missed by the lidar. In contrast to CloudSat, CALIPSO does not observe precipitation. Typically, the clouds producing precipitation are so optically thick that the signal is fully attenuated within the cloud before reaching levels where precipitation is falling. Again, a lidar forward model is employed to provide a model cloud cover comparable to the observations. The forward model calculates a simulated backscatter profile for each subcolumn. In cases where the model clouds become fully attenuated, all levels below are excluded from the comparison. The cloud occurrences observed by CALIPSO and calculated from the model are shown in Figure 5(a) and (b).

When comparing the hydrometeor occurrence from the two observational sources (Figure 5(a) and (c)), it is apparent that the frequency of occurrence observed from CloudSat is higher, particularly in the areas with deep convection. Here, the lidar will miss considerable amounts of cloud due to signal attenuation. At the same time, CloudSat not only observes cloud cover but also precipitation, which increases the frequency of occurrence. On the other hand, CALIPSO's sensitivity to small particles results in detection of more cirrus clouds above 15 km and more low-level liquid-water cloud, both of which can be below the sensitivity of the CloudSat radar. In addition, CloudSat reflectivity observations in the lowest 1 km above ground are removed, due to contamination with the strong backscatter signal from the surface.

The size of the domain (10°W–10°E, Equator–40°N) means that some tracks south of 5°N and north of 30°N lie over the ocean. In particular, at the northern end of the cross-section the westernmost tracks observe the Atlantic off the Moroccan coast, while tracks in the east observe the Sahara. The very low clouds seen in the CALIPSO data (0–2 km altitude, 0–6°N and 32–40°N) correspond to marine and coastal boundary-layer clouds.

Both sets of observations place the area with deepest convection roughly between 6°N and 18°N. In contrast, the model's intertropical convergence zone (ITCZ) is displaced further to the south and confined between approximately 2°N and 12°N. The Sahel region, between 15°N and 20°N, is an area with intermittent deep convection, which occurs much less often in the IFS model.

Differences also exist to the north of the areas with deep convection (20°N to 35°N). CALIPSO detects clouds about twice as often above the Sahara as found in the model. However, this difference does not appear in the CloudSat observations and corresponding model hydrometeor occurrence. It is possible that CALIPSO mislabels some aerosol observations as clouds and thus overestimates cloud occurrence. However, it is also possible that the similarities between CloudSat and the model are due to a similar frequency of occurrence of precipitation, rather than similar cloud amounts.

4.2. *Aerosol*

Aerosols play a significant role in the radiative budget over West Africa (Milton, *et al.*, 2008). Mineral dust aerosols are particularly abundant over the Sahel region before the monsoon onset and over the convergence zone in the heat low region during the wet monsoon season (Bou Karam, *et al.*, 2008, 2009). The dust aerosols reduce the incoming solar radiation at the surface by scattering radiation back to space as well as by absorption of radiation (Haywood, *et al.*, 2003). This has a first-order direct effect on the surface energy balance and leads to a reduction of surface temperature (Mallet, *et al.*, 2009). This direct effect on the solar radiation is crucial in the region of the heat low, as surface temperature determines the location and intensity of the monsoon trough (Lavaysse, *et al.*, 2009).

The aerosol optical depth is a measure of the integrated effects of scattering and absorption by aerosols. The IFS

Table III. Monthly mean aerosol optical depth from the climatology used in the IFS and AERONET observations. The AERONET observations are used if available for at least 20 days over a given month. Values between parentheses correspond to 5–19 days of observations; N/A corresponds to fewer than 5 days. See main text for further details.

AERONET site	Lat (°N)	Lon (°E)	May Clim.	Ob.	Jun Clim.	Ob.	Jul Clim.	Ob.	Aug Clim.	Ob.	Sep Clim.	Ob.
Banizoumbou	13.3	02.4	0.29	0.75	0.27	0.95	0.22	0.67	0.19	0.73	0.17	0.61
Blida	36.3	02.5	0.24	0.29	0.28	0.32	0.30	0.22	0.14	0.27	0.28	N/A
Djougou	09.5	01.4	0.19	0.76	0.20	0.73	0.16	0.72	0.15	0.86	0.14	1.07
Niamey	13.3	02.1	0.27	N/A	0.26	N/A	0.20	N/A	0.17	0.33	0.16	0.52
Ouagadougou	12.1	−01.2	0.24	1.00	0.22	0.81	0.19	0.56	0.15	0.71	0.14	0.46
Tamanrasset	22.5	05.3	0.30	0.32	0.31	0.60	0.31	0.23	0.27	(0.31)	0.23	(0.38)
Agoufou	15.2	−01.3	0.29	0.78	0.27	0.86	0.24	0.71	0.19	0.61	0.17	0.49
Capo Verde	16.4	−22.6	0.20	0.26	0.22	0.66	0.25	0.62	0.21	(0.53)	0.16	N/A
IER Cinzana	13.2	−05.6	0.25	0.63	0.23	1.08	0.21	0.62	0.15	0.48	0.14	(0.68)
Dakar	14.2	−16.6	0.21	0.48	0.25	0.77	0.23	0.85	0.19	0.57	0.15	0.52
Ilorin	08.2	04.2	0.18	0.65	0.21	0.52	0.16	0.41	0.16	0.40	0.14	0.53
DMN Maine Soroa	13.1	12.0	0.25	0.93	0.27	1.03	0.20	0.63	0.20	0.55	0.20	0.54
Santa Cruz Tenerife	28.3	−16.1	0.17	0.06	0.16	0.04	0.19	0.14	0.18	0.10	0.14	0.16

uses a fixed seasonally varying climatology of aerosol optical depths based on Tegen, *et al.* (1997). In this section, the aerosol optical depth from the climatology used in the model is evaluated with independent observations from the AErosol RObotic NETwork (AERONET) photometers (Haywood, *et al.*, 2008) at several sites over West Africa (see Table III). The AERONET observations are given at 440 nm, whereas the climatological values are for the spectral interval 440–690 nm of the short-wave radiation scheme of the ECMWF model. Such a difference in spectral properties would make the climatological values lower by about 50% of the value they would have at 440 nm. However, the difference between the climatological and observed values shown in Table III is generally much larger than this expected difference (i.e. larger than 50% of the observed value). This is certainly the case for stations affected by the Saharan Air Layer (SAL) from May–September, e.g. Banizoumbou, Djougou, Ouagadougou, Agoufou, Capo Verde, Cinzana, Dakar and Maine Soroa. In those stations, the monthly mean observed aerosol optical depth reaches high values and is much larger than the climatology, sometimes by a factor of four. Tamanrasset is an exception because of its location in the Hoggar massif at 1362 m above sea level. On the other hand, coastal stations to the north and west of the Sahara (e.g. Blida and Tenerife) have low values of aerosol optical depth with no significant underestimation in the climatology because they are not within the SAL during the period of the re-analysis.

4.3. *Radiation and surface heat fluxes in the Sahel*

Measurements of radiative and surface heat fluxes were carried out in 2006 at different sites (Lebel, *et al.*, 2010). In particular, the ARM mobile facility was deployed in 2006 at Niamey airport (Niger), where it contributed to the AMMA field campaign (Miller and Slingo, 2007). In this section, surface heat flux and radiation measurements from the ARM site are used to evaluate the AMMA re-analysis and the ECMWF operational analysis. Except when otherwise mentioned, results presented below are overall consistent across three Sahelian sites, from Southern Niger (Niamey, 13°N) to Central (Gourma area around Agoufou, 15°N) and Nothern Sahel (Bamba, 17°N) where surface radiative and heat fluxes were also measured (Guichard, *et al.*, 2009; Ramier, *et al.*, 2009; Timouk, *et al.*, 2009).

Agustí-Panareda, *et al.* (2010a) showed that the impact of the enhanced radiosonde network on the ECMWF forecast is short-lived. Within one diurnal cycle, the forecast initialized from the AMMA analysis develops similar biases to the pre-AMMA experiment and operational model. In particular, the model's boundary layer is too warm and it quickly becomes too deep and well-mixed. The question to be addressed is how the fluxes at the surface relate to the boundary-layer biases.

The energy balance at the surface consists of

$$G = SW_{\mathrm{dn}} - SW_{\mathrm{up}} + LW_{\mathrm{dn}} - LW_{\mathrm{up}} - SH - LH,$$

where G refers to the energy storage in the surface (usually a small term), SW is the short-wave radiative flux, LW is the long-wave radiative flux, SH the surface sensible heat flux and LH the surface latent heat flux. Subscripts 'dn' and 'up' refer to downwards (i.e. incoming) and upwards (i.e. outgoing) radiative fluxes respectively.

Observations of incoming and outgoing SW radiation show brief periods of decreased values during the dry phase (January–April), which are lacking in the model (Figure 6). While some of these events may correspond to occasions with cloudy conditions, periods with heavy aerosol loading also contribute to the observed drop in SW radiation. The green bars at the bottom of Figure 6 show the 11 μm aerosol optical depth multiplied by a factor of 100, averaged for each day (Turner, 2008). In the model a constant climatological value of aerosol optical depth is used (see Table III). Thus, episodes with high aerosol loading are missing, resulting in an overestimation of net SW radiation absorbed by the ground (Figure 7). Similarly, several precipitation events are missing in the model during the pre-onset and monsoon phases (May–mid-September). SW radiation reaching the

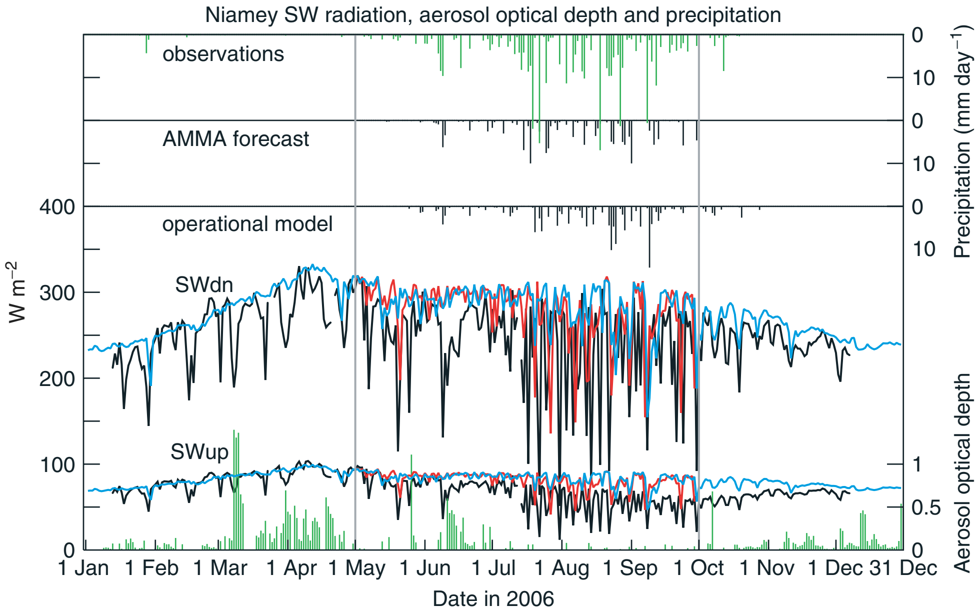

Figure 6. Surface radiation at the ARM mobile facility, Niamey. Lower panel: upward and downward short-wave radiation at the surface with observations in black, operational one-day forecast in blue and forecast initialized from AMMA analysis in red; the green bars indicate daily average aerosol optical depth, multiplied by a factor 100. Two upper panels: precipitation derived from satellite (FEWS RFEv2 dataset, courtesy of Climate Data Centre, NOAA) and the AMMA one-day forecast (mm day^{-1}). The thick grey vertical lines delimit the period of the AMMA re-analysis.

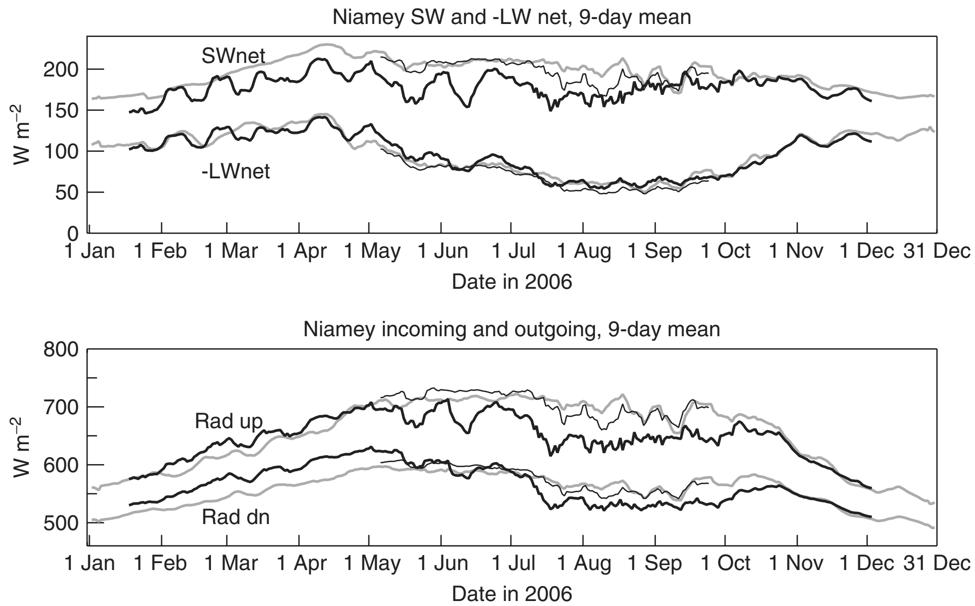

Figure 7. Surface radiation at the ARM mobile facility, Niamey. Nine-day running mean of net short-wave and net long-wave radiation at the surface (the thick black line denotes observations, the thick grey line the operational one-day forecast and the thin black solid line the forecast initiated from AMMA re-analysis).

surface is also overestimated during these events. This is true for the operational model, as well as the forecast initialized from the AMMA analysis.

Both outgoing and incoming LW radiation are underestimated in the model during the dry periods of the year (not shown), but are in reasonable agreement with the observations during the pre-onset and monsoon phases. However, compensation of the two LW fluxes leads to a realistic estimation of the net LW flux at the surface, while the net SW flux is overestimated (Figure 7). As a consequence, in the model the surface absorbs up to 50 W m^{-2} more solar radiation than observed. This value is consistent with results from Guichard (2009) (see her figures 20(c) and (d)) based on comparing the operational ECMWF model with observed incoming and outgoing radiation at the surface for a semi-arid Sahelian site (Agoufou at 15.2°N and 1.3°W). During the wet season, the net SW radiation from the AMMA re-analysis is up to 25 W m^{-2} closer to observations than that from the operational analysis, due to a decrease in incoming SW radiation associated with an increase in cloud (see section 4.1). The extra energy from the net radiation is released into the atmosphere through surface heat fluxes. The partitioning into latent and sensible heat flux is controlled by the available surface moisture.

In the ECMWF surface analysis, the soil moisture is adjusted to address biases in the 2 m temperature and

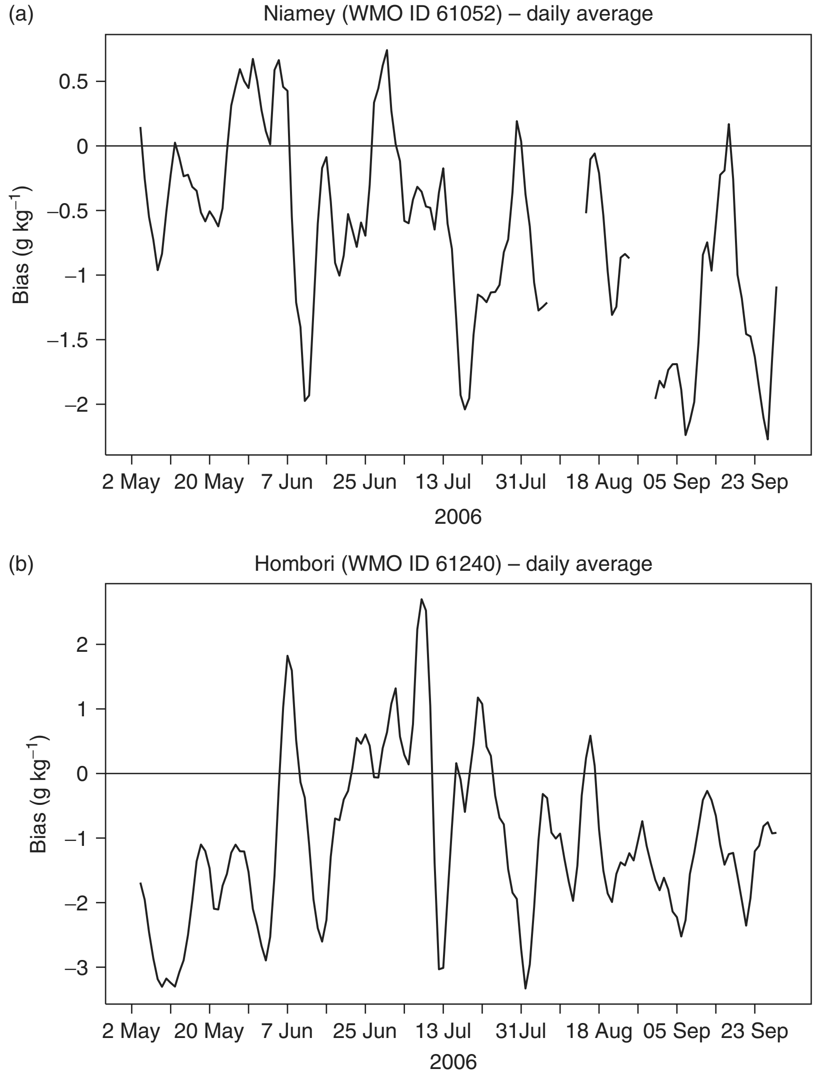

Figure 8. Time series of 2 m specific humidity bias (g kg^{-1}) of the one-day forecast relative to SYNOP observations for (a) Niamey (13.48°N, 2.17°E) and (b) Hombori (15.33°N, 1.68°W). The daily averages are calculated using the six-hourly observations (when available) and forecasts. A five-day running mean has been applied to the time series.

humidity (Mahfouf, *et al.*, 2000). Figure 8 shows the bias of 2 m specific humidity from the 1 day forecast with respect to SYNOP data at Niamey and Hombori (15.33°N,1.68°W), the SYNOP station closest to Agoufou. The mean bias from May–September is -1.06 g kg^{-1} in Hombori and -0.69 g kg^{-1} in Niamey. At both locations, the bias is mainly negative throughout the period, except for some instances when the intraseasonal variability of the observed value is not well represented in the forecast. In Hombori the bias can reach values of up to -3 g kg^{-1} and in Niamey up to -2 g kg^{-1}. These dry biases in the forecast are consistent at all the SYNOP stations within the region of the Sahel from 12°N–15°N and 10°W–10°E (not shown). The surface analysis increments performed in order to correct this 2 m humidity bias lead to an increase in soil moisture (Balsamo, *et al.*, 2009b). Consequently, the magnitude of latent heat flux in the model is very large–at times close to the magnitude of the sensible heat flux –before and after the rainy period from July–September. On the other hand, observations from all sites over the Sahel show that the latent heat flux is very small prior to the monsoon onset (Figure 9). This finding is also consistent with the study of Drusch and Viterbo (2006). In summary, it indicates that the model latent heat flux before the monsoon onset is too large over the Sahel.

During the wet monsoon phase, the surface heat fluxes are very much site-dependent, as shown by large differences in fluxes between neighbouring sites. This is the case for the mesoscale region around Niamey. For instance, surface heat fluxes from Niamey airport at 13.48°N, 2.17°E (see Figure 9) are more than twice as high as those measured at Wankama, located to its northeast at 13.64°N, 2.64°E (see figure 6 of Ramier, *et al.*, 2009). Therefore, it is not possible to use those to infer regional biases in the model. Further north, the Gourma area around Agoufou is also characterized by significant heterogeneities of surface heat fluxes at the mesoscale (Timouk, *et al.*, 2009). However, the relative simplicity of surface and soil types makes it possible to provide an estimate of a mesoscale sensible heat flux that can be more readily compared with the forecasts (grey shading in Figure 10). Beyond a possible influence of interannual variability, this comparison suggests an overestimation of the sensible heat flux during the monsoon pre-onset from May–June and particularly during the core of the monsoon over Northern Sahel from July–August, close to the heat low region. Upscaled and forecast sensible heat fluxes are much

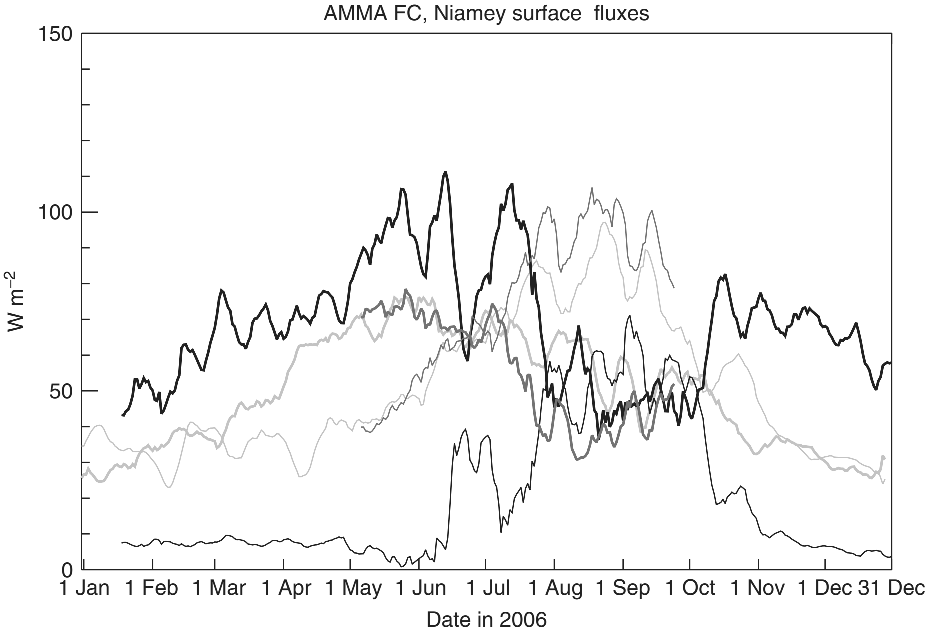

Figure 9. Time series of surface latent and sensible heat fluxes at the ARM mobile facility, Niamey. The sensible heat flux is shown by thick lines, the latent heat flux by thin lines. A nine-day running mean has been applied. Black corresponds to observations, light grey is the operational one-day forecast and dark grey is the forecast initiated from the AMMA re-analysis.

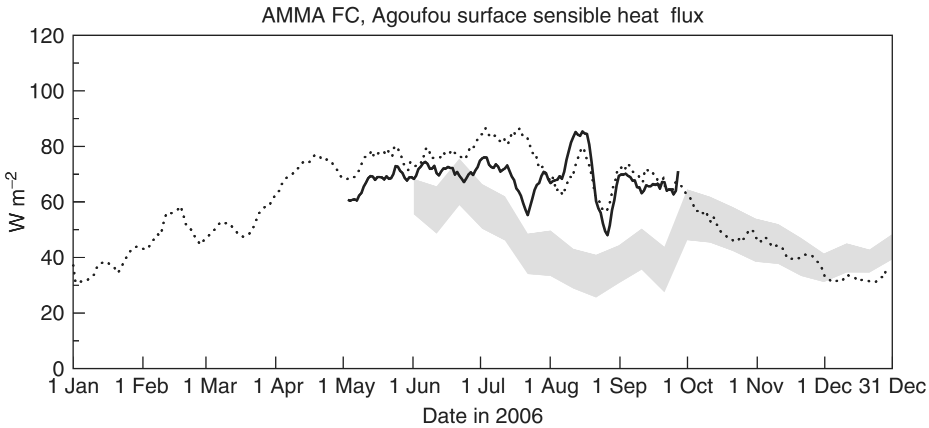

Figure 10. Simulated surface sensible heat flux (dotted line is the operational analysis and solid line is the AMMA re-analysis) smoothed with a nine-day running mean. The grey shading is the composite 10-day mean mesoscale upscaled surface sensible heat flux, estimated from individiual station data in the Gourma area around Agoufou (15.2°N and 1.3°W). The estimate was obtained from data from 2005–2007 and the thickness of the shading indicates the uncertainty due to surface heterogeneities.

closer after the monsoon, at the end of September, when there is less cloud and aerosol optical depth has decreased.

4.4. *The water cycle*

The atmospheric water budget for an atmospheric column is described by the following equation:

$$\frac{\partial W}{\partial t} = -\nabla_h \cdot \mathbf{Q} - (P - E). \tag{1}$$

Here W is the total column water vapour, t is time, P and E are precipitation and evapotranspiration rates respectively and $\mathbf{Q}$ is the vertically integrated horizontal moisture flux through the atmospheric column, given by

$$\mathbf{Q} = \left(\frac{1}{g} \int qu \, \mathrm{d}p, \ \frac{1}{g} \int qv \, \mathrm{d}p \right), \tag{2}$$

where g is gravity, q specific humidity, p pressure and u and v are the zonal and meridional wind components. The integration is over the entire atmospheric column.

Meynadier, *et al.* (2010a) presented a hybrid dataset for the atmospheric water budget in West Africa based on a combination of data that provide a best estimate for the different terms in Eq. (1). This dataset provides a powerful tool with which to assess the atmospheric water cycle of NWP models. Here, the same dataset is used to provide monthly mean values for the different terms of Eq. (1). Following Meynadier, *et al.* (2010a), precipitation is obtained from the TRMM satellite (product 3B42, see Huffman, *et al.*, 2007) and evapotranspiration is given by the offline ECMWF land-surface model (HTESSEL, see Balsamo, *et al.*, 2009a) forced by TRMM precipitation. Both precipitation and evapotranspiration were processed within the AMMA Land-surface Model Intercomparison Project (ALMIP, see Boone, *et al.*, 2009). The total column water-vapour tendency is

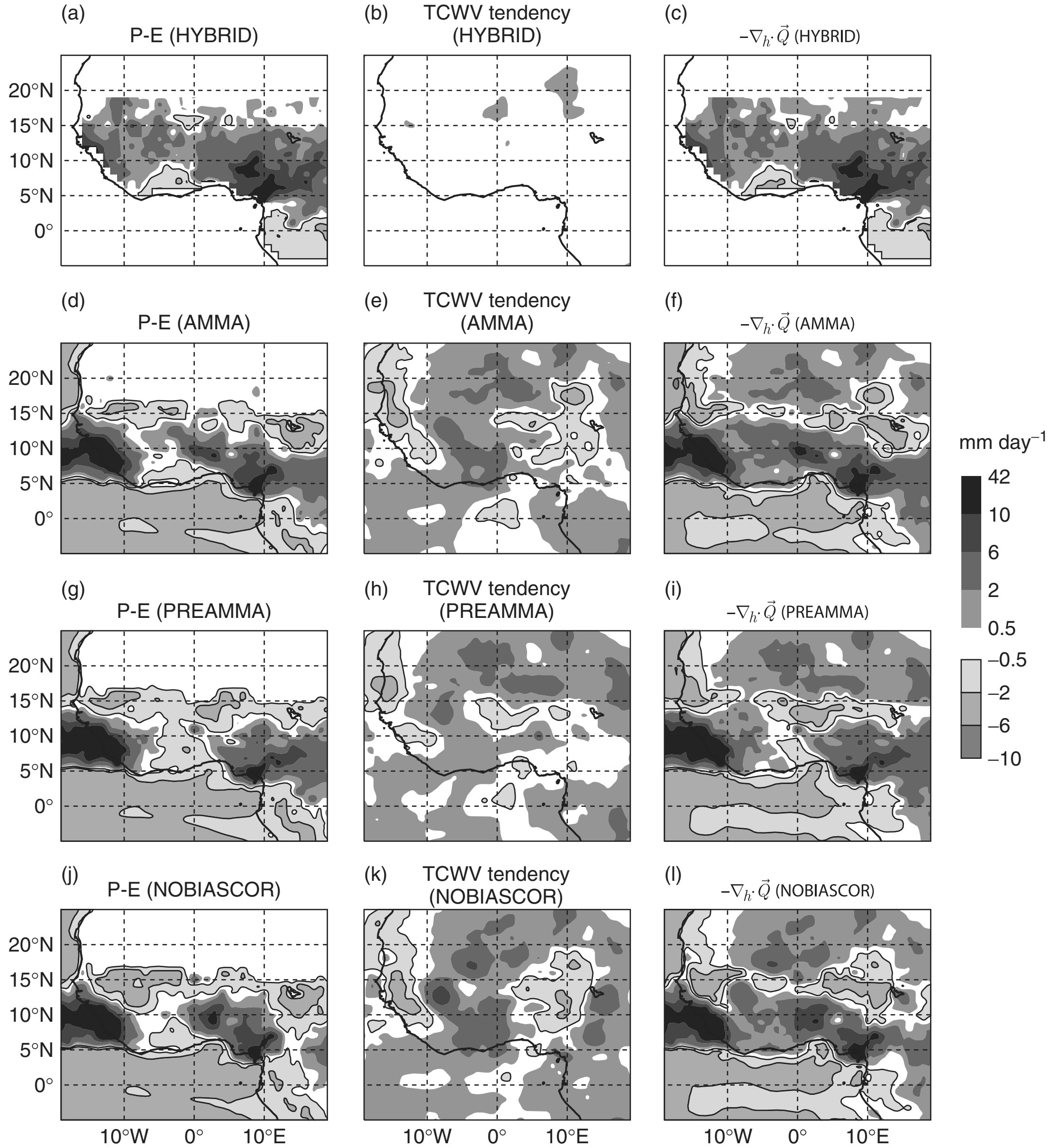

Figure 11. Monthly mean atmospheric water-budget terms for August 2006 from the hybrid dataset described in section 4.4 and the different analysis experiments (AMMA, PREAMMA and NOBIASCOR) described in Table II. See section 4.4 for a description of the different terms: $P - E$ in (a), (d), (g), (j); TCWV tendency in (b), (e), (h), (k); vertically integrated moisture-flux convergence in (c), (f), (i), (l).

obtained from the AMMA re-analysis. Finally, the residual of the three terms above provides the vertically integrated moisture-flux convergence. Note that the hybrid dataset covers the same domain used in ALMIP, and therefore there are no data available north of 19°N.

In the model forecast from the three analysis experiments (in Table II), the different terms of the water budget in Eq. (1) are obtained as follows. The first term is the change in total column water vapour ($\partial W/\partial t$) during the forecast period. This is balanced by the convergence of the vertically integrated moisture flux ($-\nabla_h \cdot \mathbf{Q}$) and the difference between evapotranspiration and precipitation ($-(P - E)$). All the terms from Eq. (1) are integrated over the same 12 h forecast period. The forecasts are initialized from 0000 UTC and 1200 UTC analyses. The monthly mean of each term is then computed.

A comparison of the three atmospheric water-budget components between the three experiments (AMMA, PREAMMA and NOBIASCOR: see Table II) and the reference hybrid dataset is presented in Figure 11. Overall, the pattern of the fields is similar between the different experiments and rather different from the reference. However, there are still interesting differences between the experiments, which emphasize the impact of the data in the analysis. The fact that the data impact is small compared with the large differences between model and reference highlights the large biases present in the model.

$P - E$ represents the net moisture sink/source that connects the atmospheric and terrestrial water reservoirs. Most of West Africa is a net moisture sink in August, except for a small area near the Guinea coast between 10°W and the Greenwich meridian. The first striking difference between the hybrid dataset and the AMMA re-analysis is around 15°N, where the sign is mainly positive in the hybrid dataset (Figure 11(a)), indicating a net sink of moisture, and negative in the AMMA re-analysis (Figure 11(d)) implying a moisture source. This problem is also present in the other experiments (Figure 11(g) and (j)). Moreover, the

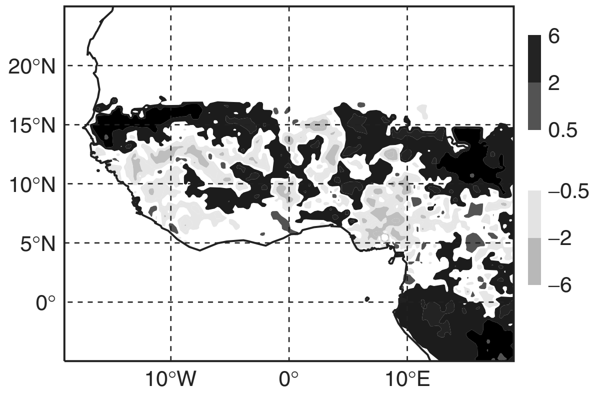

Figure 12. Top 1 m soil-moisture increments (mm day^{-1}) averaged for August 2006 from the AMMA re-analysis.

PREAMMA experiment also shows large differences, with negative instead of positive values of $P - E$ over the transect around the Greenwich meridian. This region is where AMMA revived and introduced new radiosonde stations. The radiosonde humidity bias-correction scheme also has an important impact over the West and Central African region (20°W–20°E) between 7°N and 15°N, where there were many Vaisala RS80-A sondes with large relative humidity biases (Agustí-Panareda, *et al.*, 2009a). In the NOBIASCOR experiment, this region has a net moisture source, whereas the AMMA experiment has a net moisture sink like the hybrid dataset. The fact that the $P - E$ from the forecast is negative (i.e. a moisture source) when it should be positive (i.e. a sink) is due to an underestimation of precipitation and overestimation of evapotranspiration. The underestimation of precipitation over Sahel in the IFS has been shown by Agustí-Panareda, *et al.* (2009a), Agustí-Panareda, *et al.* (2010a) and Meynadier, *et al.* (2010b). The overestimation of evapotranspiration is also discussed in Agustí-Panareda, *et al.* (2010b), and it is due to the large soil-moisture increments–equivalent to up to 6 mm day^{-1} in the top 1 m of soil–performed by the surface analysis (see Figure 12).

The TCWV tendency is also much smaller in the hybrid dataset (i.e. the AMMA re-analysis) than in the short-range forecast of the experiments (see Figure 11(b), (e), (h) and (k)). Monthly mean values of TCWV tendency in the hybrid dataset remain below 0.5 mm day^{-1} over West Africa. The three experiments all show large mean tendencies in the TCWV within the first 12 h of the forecast. This implies that there is a problem in the relationship between the other two terms in the water budget. Namely, $P - E$ and the vertically integrated moisture-flux convergence do not balance as they should.

Because the monthly mean TCWV tendency is very small in the hybrid dataset, the moisture-flux convergence that is obtained as a residual (Figure 11(c)) is very close to the $P - E$ term. That is to say, the moisture-flux convergence is balanced by the precipitation and evapotranspiration. Thus, in August most of West Africa has net moisture convergence, except for the region near the Guinea coast between 10°W and the Greenwich meridian. There are two regions where the difference between the forecast and the hybrid dataset is substantial. Near the coast the forecast from the AMMA re-analysis and the other experiments has too much convergence and over the Sahel (around 15°N) the moisture flux from the forecast is too divergent (Figure 11(f), (i) and (l)). Without the extra AMMA radiosondes and their humidity bias corrected, the divergence between 10°N and 15°N is further enhanced. East of 15°E and north of 15°N, the data from the few extra radiosonde stations increase the divergence instead of decreasing it. This has been attributed to the detrimental effect of very localized and large analysis increments that attempt to reduce the large model biases over those data-sparse regions of the Sahel (see Agustí-Panareda, *et al.*, 2010a, for further details).

From Figure 11 it is clear that the different components have a large latitudinal variability. Since most of the radiosonde observations are located between 10°W and 10°E, this region is chosen to compute the atmospheric water budget for three distinct latitude bands of 3.5° width across the steep north–south precipitation gradient. These are the Guinean [6°N–9.5°N], Soudanian [9.5°N–13°N] and Sahelian [13°N–16.5°N] bands. The monthly mean values over these three latitudinal bands are computed for each month of the AMMA re-analysis from May–September 2006. The resulting budget shown in Figure 13 indicates that the seasonal variability of the different terms is well represented in the forecast for all latitude bands. However, it is evident that considerable biases are present in the forecast throughout the wet monsoon season.

Near the Guinea coast (Figure 13(a)) the moisture-flux convergence is overestimated by approximately 1 mm day^{-1} from June–August. $P - E$ is mainly overestimated in June and underestimated in September by just under 1 mm day^{-1}. These are the two months with peak rainfall over the Guinean coast. The moisture-flux convergence is not balanced by the $P - E$ term, as in the hybrid dataset. Therefore, the mean monthly TCWV tendency is approximately 0.5 mm day^{-1} from May–August instead of being zero throughout the season as in the hybrid dataset.

Over the Soudanian band, like the coastal region, the moisture-flux convergence is balanced by the $P - E$ term in the hybrid dataset (Figure 13(b)). Nevertheless, there is a small underestimation in the moisture-flux convergence of approximately 0.5 mm day^{-1} and a large underestimation in the net moisture sink $P - E$ of 1–1.5 mm day^{-1} due to the combination of an underestimation in precipitation and an overestimation in evapotranspiration (not shown). Note that in June the forecast has a net moisture source instead of the very small net moisture sink present in the hybrid dataset. This is due to large soil-moisture increments in the surface analysis. As a result of the unbalanced relationship between the $P - E$ term and the moisture-flux convergence, there is an overestimation of the TCWV tendency in the forecast throughout the season.

The discrepancy between the forecast and the hybrid dataset becomes larger over the Sahel region compared with the Soudanian and coastal regions. Over the Sahel band (Figure 13(c)), there is a substantial underestimation of the moisture-flux convergence, particularly in July and August, of 1–2 mm day^{-1}, and an even larger underestimation of the net $P - E$ moisture sink (≥ 2 mm day^{-1}) throughout the season. This is again a combination of the lack of precipitation and the enhanced evapotranspiration due to large soil moisture analysis increments (e.g. Figure 12).

5. Discussion

The ECMWF AMMA re-analysis constitutes a valuable dataset for studying the West African monsoon, as it combines the most comprehensive observational dataset

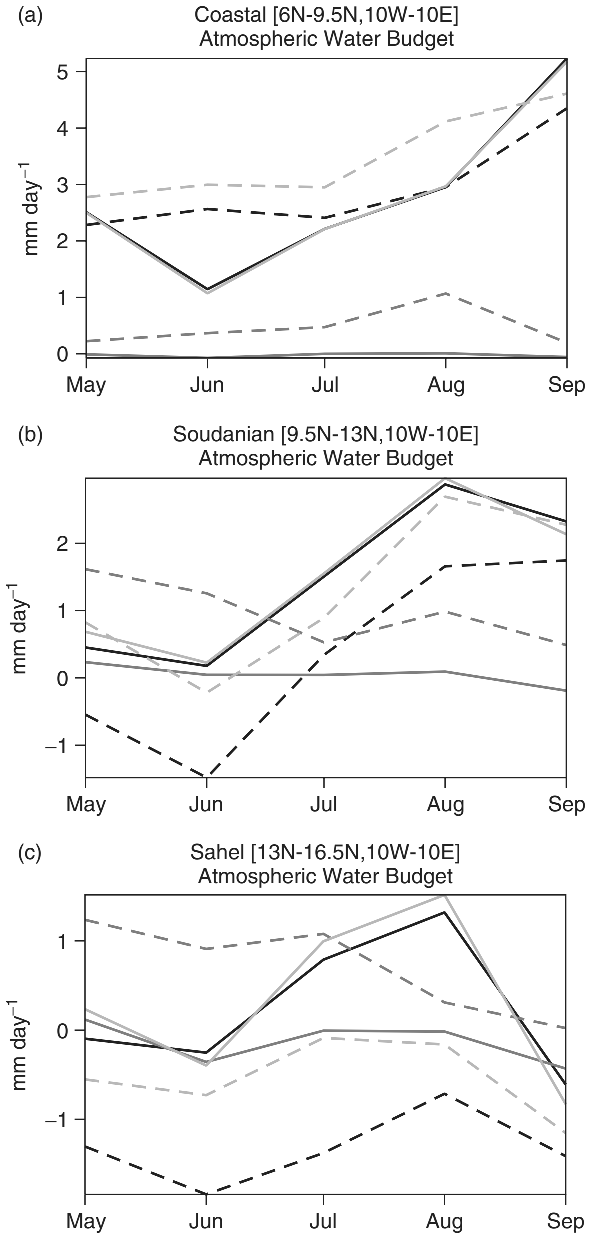

Figure 13. Time series of monthly averaged values in mm day^{-1} of atmospheric water-budget terms ($P - E$ is in black, TCWV tendency in dark grey and vertically integrated moisture-flux convergence in light grey) within three different latitude bands (panels a, b and c) from May–September 2006 for the AMMA re-analysis (dashed lines) and the hybrid dataset (solid lines).

of the region so far and a state-of-the-art NWP model and data assimilation. It supersedes the operational ECMWF analysis, as it includes a new model cycle with significant improvements in the physical parametrizations and data assimilation, a larger number of soundings at higher vertical resolution and the use of a bias-correction scheme specially tailored for the AMMA radiosonde humidity data. The re-analysis covers two crucial periods in the year 2006, i.e. the pre-monsoon and monsoon seasons. Therefore, it is suitable for studies focusing on intraseasonal variability and monsoon-onset issues. It can also provide direct benefit for a number of studies making use of analysis data. For example, the advantage of using the AMMA re-analysis has already been proven for mesoscale modelling as a result of the improvement in the initial and boundary conditions (Nicole Asencio, personal communication).

The use of AMMA and other independent satellite data has also enabled the identification of biases in the forecast model associated with physical processes and their components, namely cloud, aerosol, radiation and surface heat fluxes, as well as the water cycle. These play a key role in the West African monsoon and therefore influence the short-range forecast greatly, as shown by Agustí-Panareda, *et al.* (2010a).

The biases found in the different physical processes (see previous section) are all interrelated. The lack of aerosol and cloud over the Sahel is consistent with an overestimation of incoming SW radiation and sensible heat flux at the surface. Concurrently, the lack of convective cloud is associated with an underestimation of moisture-flux convergence. Indeed, the results from the evaluation of different analysis experiments show an overestimation of moisture-flux divergence over the Sahel, which is consistent with the well-known southward shift of the ITCZ in the ECMWF model. This is improved by the use of AMMA observations and the radiosonde humidity bias correction. However, the problem still remains on the whole, because it involves the heat low where not many observations are available in the re-analysis.

The heat low is a major driver of the meridional monsoon circulation and moisture flux during the wet monsoon season (Parker, *et al.*, 2005). The overestimation of the moisture-flux divergence over the Sahel is linked to the acceleration of the flow on the southern flank of the heat low in the forecast shown by Agustí-Panareda, *et al.* (2010a). All this evidence points towards the heat low circulation in the IFS being too strong. This conclusion is supported by an intermodel comparison of the monsoon circulation showing that the IFS has stronger wind speeds within the low-level inflow, stronger mid-tropospheric outflow and stronger vertical motion associated with the heat low than other operational NWP models (e.g. GFS from NCEP and Arpege from Météo-France; Olivier Bock and Remi Meynadier, personal communication). The intensification of the heat low is also consistent with a large underestimation of aerosol optical depth and a large net radiation bias over the region of the heat low (e.g. Agoufou in May, see Table III and section 4) in the model, which will contribute to the intensification of the heat low during the forecast.

Future model developments are expected to improve the modelling of the heat low and West African monsoon water cycle, including the following:

- Assimilation of satellite data (AMSU-B and MERIS) to constrain TCWV over land (see Bauer, 2009; Karbou, *et al.*, 2010a,b).
- Assimilation of soil moisture from ASCAT and SMOS to obtain better surface fluxes (see Drusch, *et al.*, 2008).
- Interaction of forecast aerosol from the GEMS/MACC project with radiation to reduce radiation biases.
- Improvement of soil texture dataset over deserts from the Food and Agriculture Organization (FAO) 2003 to the new Harmonised World Soil Database (HWSD) 2009, as well as revision of bare-ground evaporation to allow drier soil.
- Use of seasonally varying vegetation from climatology and eventually real-time observed Leaf Area Index (LAI).

6. Summary

An AMMA re-analysis has been performed from May–September 2006 for the AMMA field experiment. For information on the access of the AMMA re-analysis archive, see the appendix in Agustí-Panareda, *et al.* (2009b). The AMMA re-analysis makes use of all the sounding data from West Africa collected from the AMMA database, as well as a new humidity bias-correction scheme developed within the AMMA project and an improved model cycle with respect to the operational model in 2006. The combination of these new elements in the analyses has a beneficial impact on the analyses and forecasts, particularly for the water cycle. In this article the atmospheric water budget has been assessed using a hybrid dataset, which contains the best estimates of the different terms of the water budget. This is a powerful tool introduced by Meynadier, *et al.* (2010a), which provides a reference to investigate NWP model biases.

In summary, the evaluation shows that there is too much precipitation over the Guinea coast and too little over Sahel. The ECMWF AMMA re-analysis with enhanced radiosonde network and a radiosonde humidity bias-correction scheme presents improvements over the Soudanian region ($\sim 10°$N). However, the ECMWF model has too much divergence and subsidence over Sahel. This is consistent with the southern shift of the rain belt found by previous studies (Agustí-Panareda, *et al.*, 2009a, 2010a). Biases in the monsoon circulation, aerosol and sensible heat flux suggest that the heat low is primarily responsible for those biases, as it is largely unconstrained by observations in the analysis.

The plan to assimilate AMSU and MERIS data, which are sensitive to low-level moisture, over land in the operational IFS promises improvement for the water cycle in both the model analysis and forecast. Future model developments regarding the interaction of aerosol from the forecast with radiation and improved vegetation dynamics are also expected to have a significant impact, as well as the assimilation of soil moisture from satellites.

Acknowledgements

Based on a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, UK, USA and Africa. It has been beneficiary of a major financial contribution from the European Community's Sixth Framework Programme. Detailed information on scientific coordination and funding is available on the AMMA international web site 'http://www.amma-international.org'. The work presented in this manuscript was also partly supported by a fellowship from the ARM program at ECMWF (Grant No. DE-FG02-07ER64400).

The authors acknowledge the AMMA database group, in particular Laurence Fleury and Sophie Bouffiès-Cloché, for their help in retrieving the sounding data from the AMMA database and also for archiving the ECMWF AMMA re-analysis in the AMMA database. They also acknowledge Jean Blaise Ngamini for his important role in collecting the high-resolution data from the AMMA radiosonde stations, as well as the AMMA-TT1 radiosonde core group. The driftsonde data were obtained directly from NCAR/ATD and the authors are grateful to Junhong Wang, Kate Young and Dave Parsons for their help and support in the usage of this data.

Many thanks are given to Alfred Hofstadler and Ioannis Mallas for their help in allowing the data from stations that had problems with the GTS connections to be sent by e-mail to be assimilated by the model as a special request. Thanks are also given to Milan Dragovac for his help in converting the AMMA sounding data from the AMMA database into BUFR format; Antonio Garcia Mendez for his help in the monitoring of the GTS soundings that were used in the operational analysis; Carla Cardinali, Lars Isaksen, Mike Fisher, Drasko Vasiljevic and Jan Haseler for their advice and support on the data assimilation experiments; Fernando Ii and the metview team for their help with the graphics; Laurent Kergoat for providing the upscaled flux in Figure 10, Rob Hine and Anabel Bowen for their help in improving the presentation of the figures; and Jean-Philippe Lafore and Jan Polcher for their continuous support throughout this work. The comments from Erik Andersson and the anonymous reviewer, which helped to improve the manuscript, are also much appreciated.

References

Agustí-Panareda A, Vasiljevic D, Beljaars A, Bock O, Guichard F, Nuret M, Lafore JP, Mendez AG, Andersson E, Bechtold P, Fink A, Hersbach H, Ngamini JB, Parker D, Redelsperger JL, Tompkins A. 2009a. Radiosonde humidity bias correction over the West African region for the special AMMA reanalysis at ECMWF. *Q. J. R. Meteorol. Soc.* **135**: 595–617.

Agustí-Panareda A, Beljaars A, Ahlgrimm M, Balsamo G, Bock O, Forbes R, Ghelli A, Guichard F, Köhler M, Meynadier R, Morcrette JJ. 2009b. '*The ECMWF reanalysis for the AMMA observational campaign*', ECMWF Technical Memorandum, No. 612. ECMWF: Reading, UK.

Agustí-Panareda A, Beljaars A, Cardinali C, Genkova I, Thorncroft C. 2010a. Impact of assimilating amma soundings on ECMWF analyses and forecasts. *Weather and Forecasting* **25**: 1142–1160. (In Special Collection: West African Weather Prediction and Predictability).

Agustí-Panareda A, Balsamo G, Beljaars A. 2010b. Impact of improved soil moisture on the ECMWF precipitation forecast in West Africa. *Geophys. Res. Lett* In press.

Andersson E, Thépaut JN. 2008. ECMWF's 4D-Var data assimilation system – the genesis and ten years in operations. *ECMWF Newsletter* **115**: 8–12.

Balsamo G, Viterbo P, Beljaars A, van den Hurk B, Hirschi M, Betts AK, Scipal K. 2009a. A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *J. Hydrometeorol.* **10**: 623–643.

Balsamo G, Viterbo P, van den Hurk B, Hirschi M, Betts A, Scipal K. 2009b. A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *J. Hydrometeorol.* **10**: 623–643.

Bauer P. 2009. 4D-Var assimilation of MERIS total column water-vapour retrievals over land. *Q. J. R. Meteorol. Soc.* **135**: 1852–1862.

Bechtold P, Köhler M, Jung T, Leutbecher M, Rodwell M, Vitart F. 2008. Advances in simulating atmospheric variability with IFS cycle 32R3. *ECMWF Newsletter* **114**: 29–38.

Bock O, Bouin MN, Doerflinger E, Collard P, Masson F, Meynadier R, Nahmani S, Koité M, Balawan KGL, Didé F, Ouedraogo D, Pokperlaar S, Ngamini JB, Lafore J, Janicot S, Guichard F, Nuret M. 2008. The West African monsoon observed with ground-based GPS receivers during AMMA. *J. Geophys. Res.* **113**: D21105. DOI:10.1029/2008JD010327.

Boone A, de Rosnay P, Beljaars A, Balsamo G, Chopin F, Decharme B, Delire C, Ducharme A, Gascoin S, Grippa M, Guichard F, Gusev Y, Harris PP, Jarlan L, Kergoat L, Mougin E, Nasonova O, Norgaard A, d'Orgeval T, Ottlé C, Poccard-Leclercq I, Polcher J, Sandholt I, Saux-Picart S, Taylor C, Xue Y. 2009. The AMMA Land Surface Model Intercomparison Project (ALMIP). *Bull. Am. Meterol. Soc.* **90**: 1865–1880.

Bou Karam D, Flamant C, Knippertz P, Reitebuch O, Pelon J, Chong M, Dabas A. 2008. Dust emissions over the Sahel associated with the West

African monsoon intertropical discontinuity region: A representative case study. *Q. J. R. Meteorol. Soc.* **134**: 621–634.

Bou Karam D, Flamant C, Tulet P, Chaboureau JP, Dabas A, Todd M. 2009. Estimate of Sahelian dust emissions in the intertropical discontinuity region of the West African monsoon. *J. Geophys. Res.* **114**: D13106. DOI: 10.1029/2008JD011444.

Drusch M, Viterbo P. 2006. Assimilation of screen-level variables in ECMWF's integrated forecast system: A study on the impact on the forecast quality and analyzed soil moisture. *Mon. Weather Rev.* **135**: 300–314.

Drusch M, Scipal M, de Rosnay P, Balsamo G, Andersson E, Bougeault P, Viterbo P. 2008. '*Exploitation of satellite data in the surface analysis*', Technical Report 576, ECMWF Technical Memorandum. ECMWF: Reading, UK.

Guichard F. 2009. 'What can we learn from AMMA about physical processes and models?' In Seminar on Parametrization of Subgrid Physical Processes, 1–4 September 2008. ECMWF: Reading, UK.

Guichard F, Kergoat L, Mougin E, Timouk F, Baup F, Hiernaux P, Lavenu F. 2009. Surface thermodynamics and radiative budget in the Sahelian gourma: seasonal and diurnal cycles. *J. Hydrol.* **375**: 161–177. DOI:10.1016/j.jhydrol.2008.09.007.

Haywood J, Francis P, Osborne S, Glew M, Loeb N, Highwood E, Tanré D, Myhre G, Formenti P, Hirst E. 2003. Radiative properties and direct radiative effect of Saharan dust measured by the C-130 aircraft during shade. *J. Geophys. Res.* **108**: DOI:10.1029/2002JD002687.

Haywood J, Pelon J, Formenti P, Bharmal B, Brooks M, Capes G, Chazette P, Chou C, Christopher S, Coe H, Cuesta J, Drimian Y, Desboeufs K, Geed G, Harrison M, Heese B, Highwood E, Johnson B, Mallet M, Marticorena B, Marsham J, Milton S, Myhre G, Osborne S, Parker D, Rojot JL, Schulz M, Slingo A, Tanré D, Tulet P. 2008. Overview of the dust and biomass-burning experiment and African Monsoon Multidisciplinary Analysis Special Observing Period-0. *J. Geophys. Res.* **113**: DOI:10.1029/2008JD010077.

Huffman G, Adler R, Bolvin D, Gu G, Nelkin E, Bowman K, Stocker E, Wolff D. 2007. The TRMM multi-satellite precipitation analysis: Quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. *J. Hydrometeorol.* **8**: 33–55.

Karbou F, Gérard E, Rabier F. 2010a. Global 4D-Var assimilation and forecast experiments using AMSU observations over land:. Part I: Impact of various land surface emissivity parametrizations. *Weather and Forecasting* **25**: 5–19. DOI:10.1175/2009WAF2222243.1.

Karbou F, Rabier F, Lafore J, Redelsperger J. 2010b. Global 4D-Var assimilation and forecast experiments using AMSU observations over land. Part II: Impact of assimilating surface sensitive channels on the African monsoon during AMMA. *Weather and Forecasting* **25**: 20–36. DOI:10.1175/2009WAF2222244.1.

Lavaysse C, Flamant C, Janicot S, Parker D, Lafore J, Sultand B, Pelon J. 2009. Seasonal evolution of the West African heat low. *Clim. Dyn.* **33**: 313–330.

Lebel T, Parker D, Flamant C, Bourlès B, Marticorena B, Mougin E, Peugeot C, Diedhiou A, Haywood J, Ngamini J, Polcher J, Redelsperger JL, Thorncroft CD. 2010. The AMMA field campaigns: Multiscale and multidisciplinary observations in the West African region. *Q. J. R. Meteorol. Soc.* **136**(S1): 8–33. DOI: 10.1002/qj.486.

Mahfouf J, Viterbo P, Douville H, Beljaars A, Saarinen S. 2000. A revised land-surface analysis scheme in the integrated forecasting system. *ECMWF Newsletter*, Summer–Autumn.

Mallet M, Tulet P, D Ser C, Solmon F, Dubovik O, Pelon J, Pont V, Thouron O. 2009. Impact of dust aerosols on the radiative budget, surface heat fluxes, heating rate profiles and convective activity over West Africa during March 2006. *Atmos. Chem. Phys.* **9**: 7143–7160.

Meynadier R, Bock O, Guichard F, Boone A, Roucou P, Redelsperger JL. 2010a. Investigation of the West African monsoon water cycle. Part I: A hybrid water budget dataset. *J. Geophys. Res.* DOI:10.1029/2010JD013917, in press.

Meynadier R, Bock O, Gervois S, Guichard F, Redelsperger JL, Agustí-Panareda A, Beljaars A. 2010b. The West African monsoon water cycle. Part II: Assessment of NWP models. *J. Geophys. Res.* doi:10.129/2010JD013919, in press.

Miller M, Slingo A. 2007. The ARM mobile facility and its first international deployment: Measuring radiative flux divergence in West Africa. *Bull. Am. Meteorol. Soc.* **88**: 1229–1244. DOI:10.1175/BMAS-88-8-1229.

Milton S, Borrks M, Haywood J, Johnson B, Allan R, Slingo A, Grey W. 2008. Modeled and observed atmospheric radiation balance during the West African dry season: Role of mineral dust, biomass burning aerosol and surface albedo. *J. Geophys. Res.* **113**: 1–24, D00C02, DOI:10.1029/2007JD009741.

Morcrette JJ, Barker H, Cole J, Iacono M, Pincus R. 2008. Impact of a new radiation package, MCRAD, in the ECMWF integrated forecasting system. *Mon. Weather Rev.* **136**: 4773–4798.

Parker D, Burton R, Diongue-Niang A, Ellis R, Felton M, Taylor C, Thorncroft C, Bessemoulin P, Tompkins A. 2005. The diurnal cycle of the west African monsoon circulation. *Q. J. R. Meteorol. Soc.* **131**: 2839–2860. DOI:10.1256/qj.04.52.

Parker DJ, Fink A, Janicot S, Ngamini JB, Douglas M, Afiesimama E, Agustí-Panareda A, Beljaars A, Dide F, Diedhiou A, Lebel T, Polcher J, Redelsperger JL, Thorncroft C, Wilson GA. 2008. The AMMA radiosonde programme and its implications for the future of atmospheric monitoring over Africa. *Bull. Am. Meteorol. Soc.* **89**: 1015–1027.

Rabier F, Järvinen H, Klinker E, Mahfouf JF, Simmons A. 2000. The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.* **126**: 1143–1170.

Ramier D, Boulain N, Cappelaere B, Timouk F, Rabanit M, Lloyd C, Boubkraoui S, Métayer F, Descroix L, Wawrzyniak V. 2009. Towards an understanding of coupled physical and biological processes in the cultivated Sahel – 1. Energy and water. *J. Hydrol.* **375**: 204–216. DOI:10.1016/j.jhydrol.2008.12.002.

Redelsperger JL, Thorncroft CD, Diedhiou A, Lebel T, Parker DJ, Polcher J. 2006. African Monsoon Multidisciplinary Analysis. *Bull. Am. Meteorol. Soc.* **86**: 1739–1746.

Stephens G, Vane D, Boain R, Mace G, Sassen K, Wang Z, Illingworth A, O'Connor E, Rossow W, Durden S, Miller S, Austin R, Benedetti A, Mitrescu C, Team TCS. 2002. The Cloudsat mission and the A-train. *Bull. Am. Meteorol. Soc.* **83**: 1771–1790.

Tegen I, Hoorig P, Chin M, Fung I, Jacob D, Penner J. 1997. Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res.* **102**: 895–915.

Tiedtke M. 1993. Representation of clouds in large-scale models. *Mon. Weather Rev.* **121**: 3040–3061.

Timouk F, Kergoat L, Mougin E, Lloyd C, Ceschia E, Cohard JM, de Rosnay P, Hiernaux P, Demarez V, Taylor C. 2009. Response of surface energy balance to water regime and vegetation development in a Sahelian landscape. *J. Hydrol.* **375**: 178–189. DOI:10.1016/j.jhydrol.2009.04.022.

Tompkins A, Cardinali C, Morcrette JJ, Rodwell M. 2005. Influence of aerosol climatology on forecasts of the African Easterly Jet. *Geophys. Res. Lett.* **32**: 1–4. DOI:10.1029/2004GL022189.

Tompkins A, Gierens K, Rädel G. 2007. Ice supersaturation in the ECMWF integrated forecast system. *Q. J. R. Meteorol. Soc.* **133**: 53–63.

Turner D. 2008. Ground-based infrared retrievals of optical depth, effective radius, and composition of airborne mineral dust above the Sahel. *J. Geophys. Res.* **113**: D00E03.

