ECMWF – ARM Report Series

1. Assessing Physical Processes in the ECMWF Model Forecasts using the ARM SGP Observations

S. Cheinet, A. Beljaars, M. Köhler, J.-J. Morcrette
and P. Viterbo

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen terme

Series: ECMWF - ARM Report Series

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/publications/

Contact: library@ecmwf.int

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, England

ECMWF - ARM Report Series No. 1

Assessing Physical Processes in the ECMWF Model Forecasts using the ARM SGP Observations

S. Cheinet[1], A. Beljaars, M. Köhler, J.-J. Morcrette and P. Viterbo

July 2005

[1]Fellowship funded by the ARM program

ARM – Atmospheric Radiation Measurement

Abstract

In this study, we compare short-term weather forecasts of the ECMWF model (IFS) to observations at the Southern Great Plains (SGP) site of the Atmospheric Radiation Measurement (ARM) program in July 2003. By using different ARM instruments and complementary satellite and radar network data, a number of systematic deficiencies in the representation of cloud and mixing processes are characterized in the IFS.

The IFS correctly predicts the mean flow. However, it misrepresents the intensity of rainy deep convective systems, and their synoptic-scale propagation from the Rockies. It also underestimates the nocturnal Low-Level Jet (LLJ) and the diurnal temperature range. The night-time deficiencies are improved by decreasing the intensity of mixing in the stable boundary layer. The afternoon cold bias in fair-weather days is consistent with a lack of vertical mixing related to an underestimation of the frequency of shallow cumuli, which probably results from the systematic dry bias at low levels.

The lower level moisture in the SGP is remotely controlled through the competition between the LLJ advection and the convective processes that reduce the LLJ moisture along the path from the Gulf of Mexico. Sensitivity experiments performed in this study suggest that the model dry bias is caused by a misrepresentation of this mechanism. Implications of our results in terms of parameterization development and evaluation are discussed.

1 Introduction

The acquisition of global-scale measurements of the atmospheric state is a corner-stone of modern-day meteorology. It is performed on a routine basis, through different and complementary techniques. Worldwide meteorological stations form extended networks of routinely acquired data. Satellite remote sensing also provides measurements at the global scale. The formed data bases adequately provide initial conditions for the Numerical Weather Prediction (NWP) models (through the assimilation stage). Independently, they can also be used to reveal some systematic (and usually model-dependent) discrepancies between model predictions and observations. However, present global-scale observing systems usually monitor a relatively small number of simultaneous fields. Hence, as soon as the observed features involve feedbacks between a large number of physical processes, it becomes virtually impossible to attribute an identified model deficiency to a single model module. This is notably the case for cloud-mixing-radiation interactions, that are parameterized in atmospheric models. This explains the persistence of well-documented deficiencies in many NWP models (e.g. the underestimation of oceanic sub-Tropical stratocumulus).

The development of heavily-instrumented site measurements was a breakthrough, as it appeared to overcome this issue. By observing a number of simultaneous atmospheric variables, one can document the various aspects of the considered issue. The construction of site facilities dedicated to that purpose was pioneered by the US Department Of Energy (DOE), through the ARM (Atmospheric Radiation Measurement) program. The three ARM sites now deploy unique sets of in-situ, passive and active remote sensing instruments, with routine retrievals of thermodynamics, cloud / radiation and surface properties in the same atmospheric column. The existence of such heavily-instrumented site measurements has been exploited at the European Center for Medium-Range Weather Forecasts (ECMWF). Past and current studies have involved diagnostic comparisons between model and observations. Morcrette (2002) uses spring observations at the ARM Southern Great Plains (SGP) site to evaluate the IFS cloud and radiation parameterizations. Mace et al. (1998) rely on winter time cloud radar measurements to analyze the IFS cloud cover forecasts. Betts and Jakob (2002a,b) attempt to tackle some of the IFS issues with respect to the diurnal cycle of Tropical deep convection through comparisons with measurements at an Amazonian site.

The current study aims at documenting the parameterization of physical processes at mid-latitudes over continents, through diagnostic comparisons between model forecasts and site observations. Our interest here is

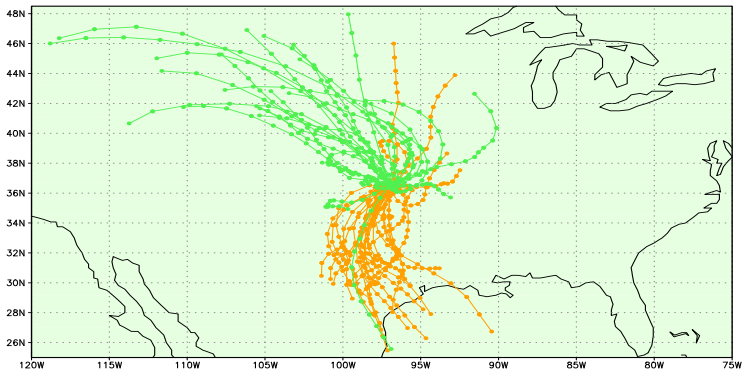

Figure 1: 36 hours backward trajectories ending at the model grid-point closest to the SGP Central Facility, at 12*LT for all days in July 2003, at model levels 54 (≈* 925*mb, orange) and 42 (≈* 600*mb, green). Dots are plotted every 3 hours.*

on the representation of the various types of mixing (stable boundary layer, dry convection, shallow and deep cumulus convection). It is known that the diurnal cycle of deep convection remains problematic in the IFS. Another documented issue is that the model tends to underestimate shallow convection over continents. The underlying question we try to answer is: using a comprehensive set of site observations, how far can we go in deducing parameterization deficiencies from forecast errors? Occasionally sensitivity experiments will be performed, in order to illustrate the proper impact of one closure or parameterization. However, at this stage, our study is not intended to improve the model predictions, since validating changes in the parameterization package of a NWP model requires global scale forecast experiments, at various time ranges.

The site selection is indeed a crucial point. In a search for focus, we have chosen to concentrate on the ARM SGP site in July 2003, because of the relevance of this choice for our purpose. Despite its possible deficiencies, one can use the IFS to infer the general structure of the flow reaching the SGP site. Fig 1 shows backward trajectories generated from the resolved flow at two model levels, starting each day in July 2003. It clearly emphasizes that the mean tropospheric flow is governed by mid-latitude Westerly winds passing over the Rockies. On the other hand, the boundary layer air mostly comes from the Gulf of Mexico, through the Low Level Jet (LLJ, e.g. Higgins et al., 1997). Given this particular flow structure, various types of convective activity affect the Southern Great Plains in summer: non-cloudy, shallow cumulus, and deep cumulonimbus. Not only is the SGP site the most heavily instrumented ARM site, but also the data can be enriched with other observations accessible over the US (from weather monitoring observational networks, satellites etc). This enables further comparisons between model forecasts and observations in the SGP surroundings. Besides, this guarantees that the model is initialized with a realistic atmospheric state, since some observations enter the initialization process through the assimilation stage. Occasionally, July months of other years and regional considerations will be emphasized, so that one can assess the climatological representativeness of the selected site and period.

The present study is organized as follows. Section 2 summarizes the IFS model and the ARM observations. The next two sections perform extensive comparisons between observations and forecasts. Surface properties and their diurnal cycle are considered in Section 3, atmospheric levels being analyzed in Section 4. Site observations give a snapshot of the atmospheric dynamics and physics at one place, but they are intrinsically local, so they are not sufficient to assess synoptic-scale processes. Satellite data or ground-based instruments networks have been used to perform such larger-scale analysis. This is done in Section 5, which discusses the impact of remote processes on the meteorology of the Southern Great Plains. Section 6 summarizes the findings of this study.

2 Model and observations set-up

The ARM SGP site consists of about twenty meteorological stations North of Oklahoma and South of Kansas. It stands in an area characterized by a relatively regular topography (East-West gradient), with the Rockies Mountains 700*km* to the West. The so-called Extended Facilities spread in a square centered on the Central Facility (C1, located 36.5N, 97.5W), within a 200*km* distance. The hereafter-called Local Time (LT) stands for UTC time minus six hours (only Fig 16 uses a longitude-dependant definition of local time). Unless precised, the data shown in this study were acquired at C1 (also indexed as extended facility E13 for some instruments). One station data can not be representative of fields that are marked by strong spatial heterogeneities (e.g. surface fluxes, rain rate). We then have averaged over the full network of available facilities. For these applications, a model domain has been defined as a counterpart of the ARM domain area, with boundaries being 95.4-99.7W in longitude and 34.8-38.8N in latitude.

All the observational data used here have been monitored by ARM quality checks (so-called *b1* quality level). Generally speaking, the time-resolution of the samplings is very high, so the data has been averaged over one hour (this time-averaging is performed so that, for example, hour indexed as 18 gathers data between hour 17 and 18). Profiling instruments (radar retrievals and radiosondes) use height as vertical coordinate. This has been converted into pressure by using the model surface pressure (which is always less than 2*mb* from the observation in July 2003). Radiosonde measurements are performed every 6 hours at C1 (i.e. 4 times a day). In July 2003, twelve radiosonde measurements were missing over the 124 expected, with no marked trend for one particular time of the day. Hence the available radiosondes allow an unbiased and reliable assessment of the atmospheric properties. Also, the horizontal displacement of sondes along their ascent has been neglected in this study. Unless specifically mentioned, the cloud cover observations shown hereafter are derived from the Actively Remote Sensed Cloud Location (ARSCL) algorithm (Clothiaux et al., 2000), which relies on a cloud radar (among other instruments). This retrieval yields the cloud boundary heights of a number of simultaneous cloudy layers. They are converted into cloud fraction through the above-mentioned one-hour averaging.

The simulations considered hereafter are based on Cycle 28R1 of the IFS, which was officially released in March 2004. The model is run with a time-step of 15*mins* and a horizontal resolution of approximately 80*km* in the area of interest (so-called T255 spatial resolution). There are 60 vertical levels, which follow $\eta-$coordinates (combination of pressure and height, see Appendix), with typically 10-15 levels in the boundary layer and 15-20 in the free troposphere. Model results shown in this study combine 24-hour forecasts over one month at the closest grid point to C1. Occasionally averages over the whole ARM SGP domain will be considered. Simulations always start at 12UTC, and results shown are 18-42 hours forecasts[1]. Note that the verification period starts at local midnight so that the morning and afternoon are produced by a same simulation. The choice of relatively short-term forecasts is justified by the relative vicinity of the initialization: model spin-up should have ceased, whereas the mean state should not be too far from observations. Unless defined otherwise, model outputs have been averaged over one hour. This procedure is consistent with the time-averaging used for the observations. There is however no direct correspondence with respect to the spatial sampling. It has been recognized that more elaborated comparison tools may also be valuable in quantifying the IFS forecast skill (e.g. Jakob et al., 2004).

3 Surface properties

The surface fluxes and thermo-dynamical properties are indeed strongly marked by the diurnal cycle. Whereas the day-to-day variability is of importance, systematic deficiencies were noted in the IFS with respect to the

[1] The first 30 hours of the month are thus omitted in this study.

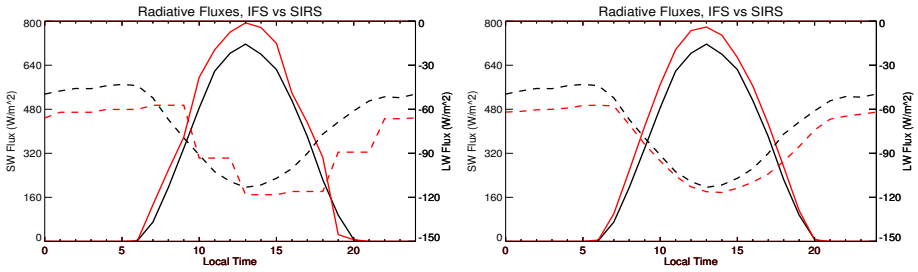

Figure 2: Monthly averaged diurnal cycle of the net Short-Wave and Long-Wave radiative fluxes at the surface, in the IFS (red) and as measured by the SIRS (black). Fluxes are positive downward. (b) is the same as (a) but with radiation scheme called every hour instead of every 3 hours.

radiative fluxes. Fig 2 shows the monthly average of the diurnal cycle of net radiative fluxes at the surface (positive downward). These quantities are measured by the Solar Infrared Radiation Stations (SIRS) at the extended facility 13. The model exhibits an important and almost systematic (i.e. every day during the month) excessive warming of the surface in the afternoon, already reported over the SGP site by Morcrette (2002), using Cycle 23R1 of the IFS. It is of course tempting to attribute this feature to the underestimation of low cloud cover, which will be shown later to occur in the afternoon. However, non-cloudy days still exhibit excessive incoming solar flux (by about $40\,Wm^{-2}$). As already noted by Morcrette (2002), the almost systematic underestimation of the precipitable water content (by $3\,kgm^{-2}$, see also below) is consistent with this result. It also does not preclude an underestimation of aerosols impact or an underestimation of the surface albedo.

The IFS also suffers from important discrepancies with respect to the Long-Wave net flux, which are again quasi systematic. The underestimation at night presumably relates to the lack of water vapor in the lower atmospheric levels (see below), which results in less downward emission. Furthermore, the phasing is lagged by three hours compared to the observations. The model behavior suggests that the surface does not respond rapidly enough in terms of Long-Wave emissions. The step-like aspect of the Long-Wave net flux relates to the numerical time-step of three hours used in IFS 28R1 for the full radiation module call. This Long-Wave flux deficiency of ECMWF forecasts is well-known and documented in other areas of the globe. As a response, the most recent version of the operational IFS calls the full radiation scheme every hour. Fig 2b shows exactly the same comparison as 2a, but with an hourly call to the radiation scheme as only difference with the reference numerical experiment. Very satisfactorily, the step-like behavior disappears, as well as most of the phase lag.

The surface latent and sensible heat fluxes are measured by the Energy Balance Bowen Ratio (EBBR) systems at the extended facilities. These fluxes vary greatly from place to place, according to previous atmospheric conditions (e.g. rain), and surface type (e.g. type of soil, land use, etc). In a general sense, the ARM SGP domain is known to have a strong West-East gradient in the Bowen ratio. In order to assess this aspect of the SGP site meteorology, we have performed a station-to-station analysis of the ratio between the monthly mean latent heat flux by the monthly mean sensible heat flux. Model data are here considered at the closest grid point to the available stations (extended facilities 2, 4, 7, 8, 9, 12, 13, 15, 18, 19, 20, 22, 26, 27). First, some stations appear to sustain ratios of less than 0.5 for the whole month (typical of dry surfaces), which is not reproduced by the IFS. Besides, whereas the ARM EBBR network does show an increase of the ratio Eastwards, it reveals a much greater variety than predicted by the IFS (Fig 3a). The IFS and the EBBR stations apparently agree in the Northern part of the SGP. In the IFS, the most Southern part of the SGP resembles the Northern, with a further Westwards displacement of the maximum gradient (the only outlier station in the model is E26, at 98.1W and 35.0N). This behavior is not confirmed by the EBBR stations, as the most South-East stations of the ARM domain account for ratios of less than 2.

Generally speaking, during summer over mid-latitude continents, the sensible heat flux progressively increases

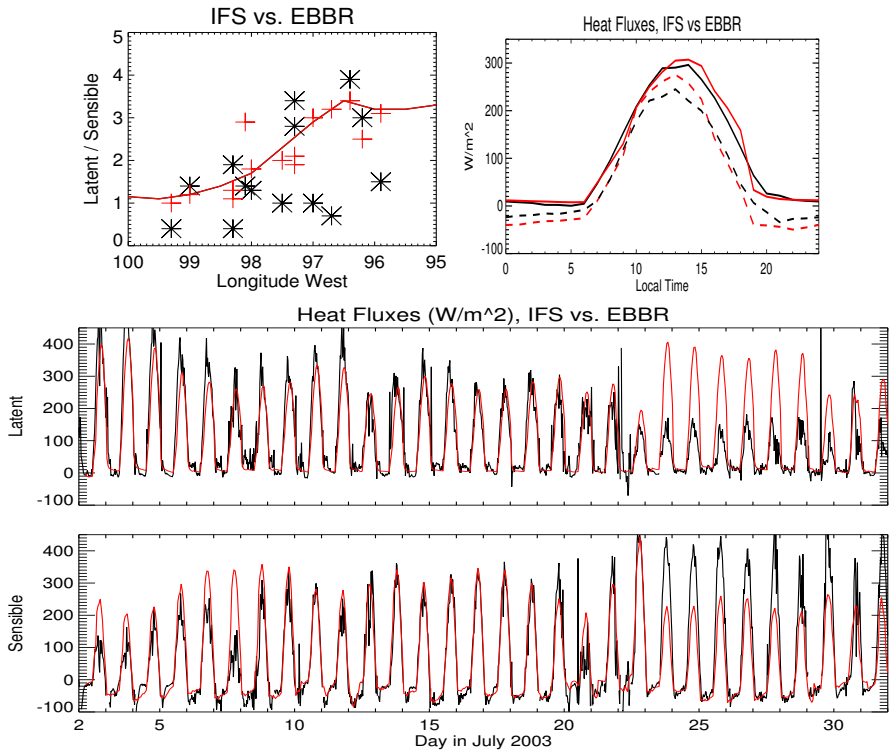

Figure 3: Comparison between IFS forecasts (red) and EBBR measurements (black). (a) Monthly averaged ratio between latent and sensible fluxes at the surface, for the ARM network stations (symbols) and for the model domain (line). (b) Monthly averaged diurnal cycle of latent (full lines) and sensible (dashed lines) fluxes, averaged over the ARM domain (IFS) and over ARM network stations (EBBR). (c) time series at C1.

whereas the latent heat flux decreases, as the surface dries. The day-to-day variations for E13 are shown in Fig 3c. Like other stations, E13 is marked by a well-defined drying during the month, which is captured by IFS until 22 July. After this day, the IFS strongly overestimates (underestimates) the latent (sensible) heat flux. The model may have forecast too much rain over E13, or the soil water cycle in the IFS may not coincide with the processes at work at E13 (e.g. local harvest). In all cases, E13 is clearly not representative of the other ARM stations with respect to the late July 2003 EBBR measurements. In order to assess systematic IFS biases in the diurnal cycle, it was thought more appropriate to use the early-July period, or equivalently the whole ARM EBBR network average, noting that the two yield comparable conclusions. The monthly averaged diurnal cycle of sensible and latent heat fluxes averaged over the ARM facilities is shown on Fig 3. Whereas the latent heat flux agrees well with the EBBR network, the sensible heat flux suffers from an underestimation at night (too negative), and an overestimation during the day. This latter feature is expected given the excessive net incoming radiative flux in the afternoon.

Fig 4 shows the monthly averaged diurnal cycle of the $2m$-temperature and specific humidity at E13. Exactly like for the radiative fluxes, given the possible impact of late July atypical behavior, it has been checked that the results for these days are consistent with the early-July days. The IFS underestimates the diurnal temperature range compared to SMOS measurements, at almost all days of the month. This includes two presumably distinct biases, one at day-time and one at night-time. Probably, the late afternoon cold bias reflects the surface sensible heat flux bias noted above. It may also involve (a combination of) the following reasons: (1) a misrepresentation of the vertical structure of the boundary layer (e.g. lapse rate in the super-adiabatic layer), and (2) a lack of incorporation of tropospheric air into the boundary layer (see below). Similarly, the night-time deficient cooling of the surface air may relate to the parameterization of the boundary layer under stable stratification.

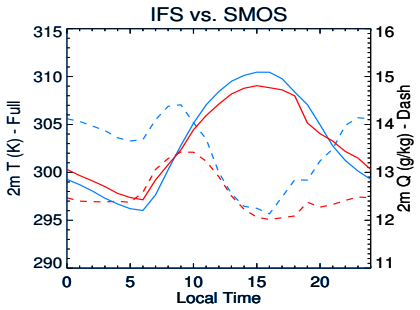

Figure 4: Monthly averaged diurnal cycle of 2m−specific humidity (dashed lines) and 2m−temperature (full lines), in the IFS (red) and as measured by the SMOS (blue).

Too efficient mixing with upper levels would yield an excessive warming from above.

Considering the surface specific humidity, in both the model and the observations, the moisture peaks in the morning, as surface evaporation accumulates in a relatively shallow boundary layer (Fig 4). It then drops as the boundary layer grows and incorporates drier air from above. However, the observed moisture increases in the evening, which is barely accounted for in the model. The lack of local evaporation is not likely to explain this feature, as the observed surface latent heat flux is rather small at night. The overly thick stable boundary layer (see above) is consistent with the excessive dry night-time surface layer. However, if the local vertical mixing were only at fault, there would be a compensating moisture excess elsewhere in the column, which is not the case. Hence the most likely cause for this underestimation is that the modelled nocturnal Low-Level Jet does not advect enough moisture over the SGP (Higgins et al., 1997). This issue will be further investigated later in this study.

The surface rain is routinely measured in ARM extended facilities by the Surface Meteorological Observation System (SMOS). Given the spatial heterogeneity of rain patterns, it was decided to average over all the available data (including facilities indexed 1, 3, 4, 5, 6, 7, 8, 11, 13, 15, 21, 27) for which the measurements time series was fairly complete. The resulting accumulated rain time series can be compared to the IFS forecasts for the model ARM domain (as defined above). We also used some mosaic analysis data based on rain radars and surface gauges (see description below), averaging it over the same domain. The three time series of accumulated rain are displayed on Fig 5. Qualitatively, one satisfactory feature is the excellent match between the ARM measurements, the mosaic data and the model forecasts with respect to the detection of significant rain events. Nevertheless, the IFS apparently tends to underestimate the rain rate over the ARM domain (this is notably true for the 10 July rain event).

It is difficult to go beyond qualitative arguments, because there is a large discrepancy between the two observing networks. Examination of the sampled ARM facilities locations showed that half of them are in the North West quadrant of the ARM domain. This may not by itself explain the underestimation of the first significant event by the SMOS network, because the rain that night spread throughout the domain. A direct comparison of each station measurements with the 4x4 *km* mosaic data exhibits large discrepancies, which can be positive or negative. Apart from possible calibration problems in the observing networks, our results suggest that the domain-averaged accumulated rain might not be quantitatively captured by the ARM network.

4 Atmospheric properties

We now compare IFS and observed properties in the atmosphere. Instruments used in this section include radiosondes and remote sensing techniques (passive and active).

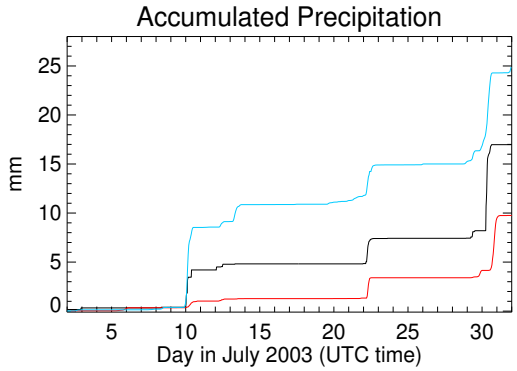

Figure 5: Comparison of accumulated precipitation rates in July 2003. IFS 18-42 hours forecasts (red) and NCEP Stage 4 rain analysis (blue) are averaged over the model-defined domain, SMOS observations (black) averaged over all the available ARM facilities (see text).

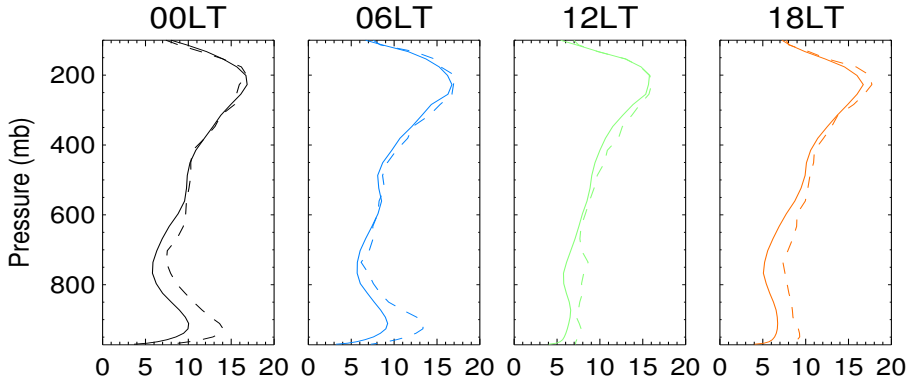

Figure 6: Monthly averaged wind speed profiles ($m.s^{-1}$) every six hours of the day. Dashed lines: ARM radiosonde measurements, full lines: IFS forecasts.

4.1 Wind

The radiosondes enable assessment of the performance of the IFS with respect to the mean wind. The comparison between forecasts and observations stresses the good performance of the forecasts for the mean wind in the troposphere (e.g., Fig 6 for wind speed). For example, at 350 *hPa* (in the upper troposphere), the mean observed wind speed is $11.9\,ms^{-1}$, and the model bias is $-0.8\,ms^{-1}$, i.e. less than 7%. Also, the model captures the variability at these levels, with a standard deviation (around monthly mean value) of $5.3\,ms^{-1}$ compared to $5.7\,ms^{-1}$ in the observations. As expected, the tropospheric flow, initially tied to the observations through the assimilation system, remains realistic in short term integrations. This agreement is also verified for the wind direction (not shown).

Somewhat surprisingly, the picture is much more contrasted in the boundary layer (Fig 6). First, the model bias ($-3.2\,ms^{-1}$, or -30%) is much stronger, and secondly, the model underestimates the time variability of the wind speed (the standard deviation is $3.2\,ms^{-1}$ in the model, $4.3\,ms^{-1}$ according to the sondes). This is also apparent on Fig 7a,b, which present hodographs of model and observed winds (sampling times when observations are available). Because the wind in the boundary layer is marked by a diurnal cycle, Fig 7a,b distinguish between values according to the time of the day. The day-to-day variability is underestimated in the model at all times. Fig 7c compares the monthly averaged winds. The model hour-by-hour data emphasizes the Northwards LLJ, with its progressive East-to-West tilting during the night (note the lack of tilting in the forecasts compared to the

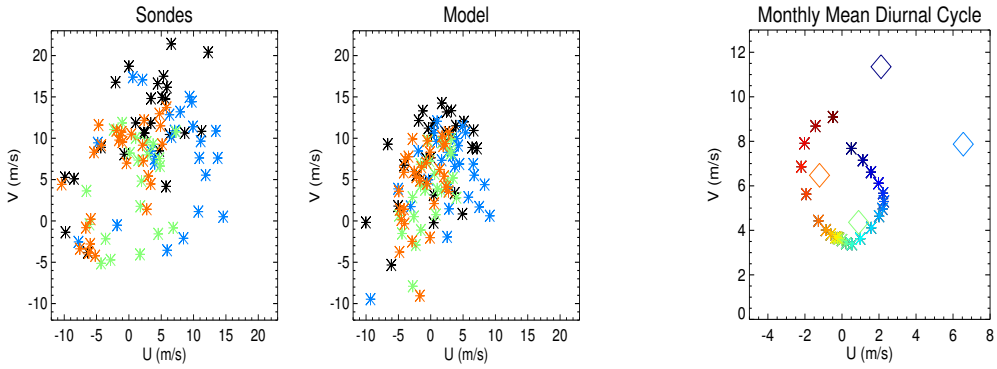

Figure 7: Wind hodographs at 935 hPa at times sampled by ARM. (a) radiosondes measurements, (b) IFS 18-42 hours forecasts, (c) monthly averaged values (triangle: model, diamond: observations). Colors give the Local Time: black=00, blue=06, green=12, orange=18. In (c), the stars give the model wind hodograph diurnal cycle based on a hour-per-hour averaging (colorcode consistent with a and b).

radiosondes average). This tilting is caused by an inertial oscillation initiated by the a-geostrophic component of the wind, which is a well-known phenomenon in the SGP (e.g. Garratt, 1992). Fig 7c shows that greater discrepancies are found for strong (North-Eastwards) winds at night and early morning. This corresponds to the LLJ peak (e.g. Higgins et al., 1997).

The general underestimation of the nocturnal LLJ at C1 seems consistent with the above-mentioned overestimation of the stable boundary layer thickness. This explanation needs caution, because there is a negative feedback at play: less shear implies less mixing which in turn allows more shear and then more mixing. The wind and turbulence profiles thus result from subtle balances between processes. In order to clarify this point, we have performed a sensitivity experiment, referred to as STBL (the simulation used up to now being CTRL). STBL uses a different stability function formulation. It results in much smaller (factor of 10) eddy-diffusivity mixing coefficients than CTRL in the case of very stable boundary layers. This formulation has been documented previously in the context of the IFS (Beljaars, 2001; Beljaars and Viterbo, 1998). Fig 8 shows exactly the same plots as 7b and 4, but for STBL. Comparing with CTRL, it is clear that STBL yields a much better match with respect to the minimum $2m-$temperature in the early morning. Although the wind bias still maintains in STBL, the monthly averaged diurnal cycle of the wind in the boundary layer is more realistic (stronger at night by more than $2m.s^{-1}$). Overall, the STBL better results at C1 clearly support the use of the alternative stability function in the IFS. However, this conclusion is hampered by the reported decrease of the IFS forecast performance at the global-scale (Beljaars and Viterbo, 1998).

4.2 Clouds and water

Fig 9 shows the time series of modelled and observed cloud fraction with height. A threshold of 2% has been imposed, for it is unlikely that such a small cloud cover in the model can be detected through a beam operating system. Actually the ARSCL algorithm produces only very few instances of cloud fraction below that threshold, whereas the model tends to maintain (probably spurious) cloud fractions of less than 0.1% in a number of occasions. Note also that the cloud radar detects rain particles, so rainy clouds base is retrieved at the surface (when rain reaches the surface). This is not the case in the model, since rain water is not considered part of the cloud. The overall impression is that the model satisfactorily predicts the deep clouds that reached C1 in July 2003 (with some errors in the precise timing, see also below). Most of these events are associated with rain (see also Fig 5), e.g. 10, 13, 21, 23, 30 July. Comparing a beam measurement with a wide grid-box

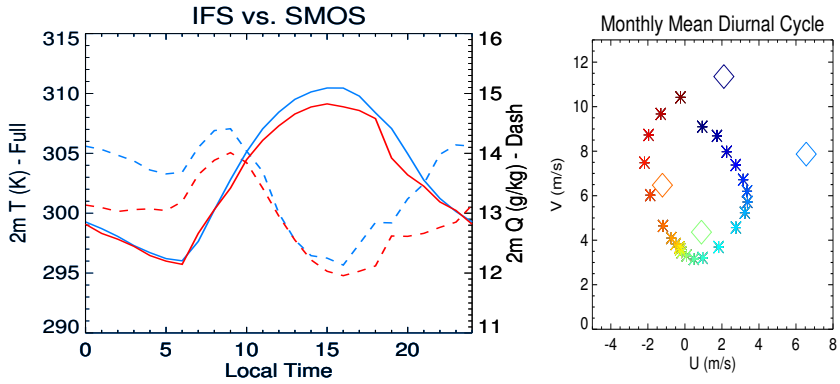

Figure 8: (a) same as 4, (b) same as 7b, but using the formulation of Beljaars and Viterbo (1998) in the IFS parameterization of the stable boundary layer.

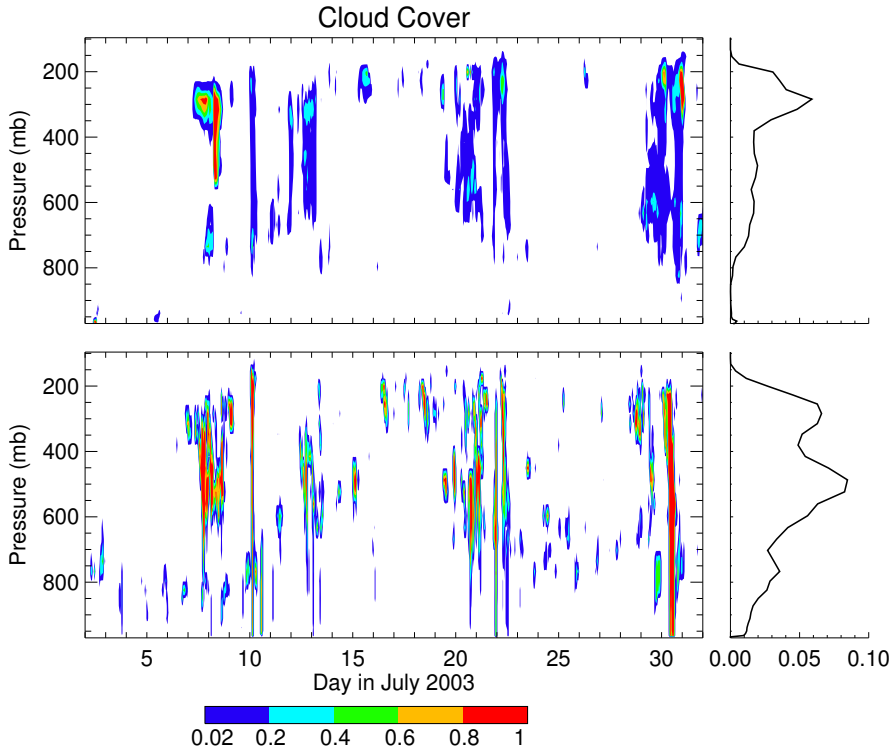

Figure 9: Cloud fraction with time and pressure, top: IFS forecasts, bottom: one-hour processing of ARSCL retrieval, left panels: time series, right panels: monthly averages.

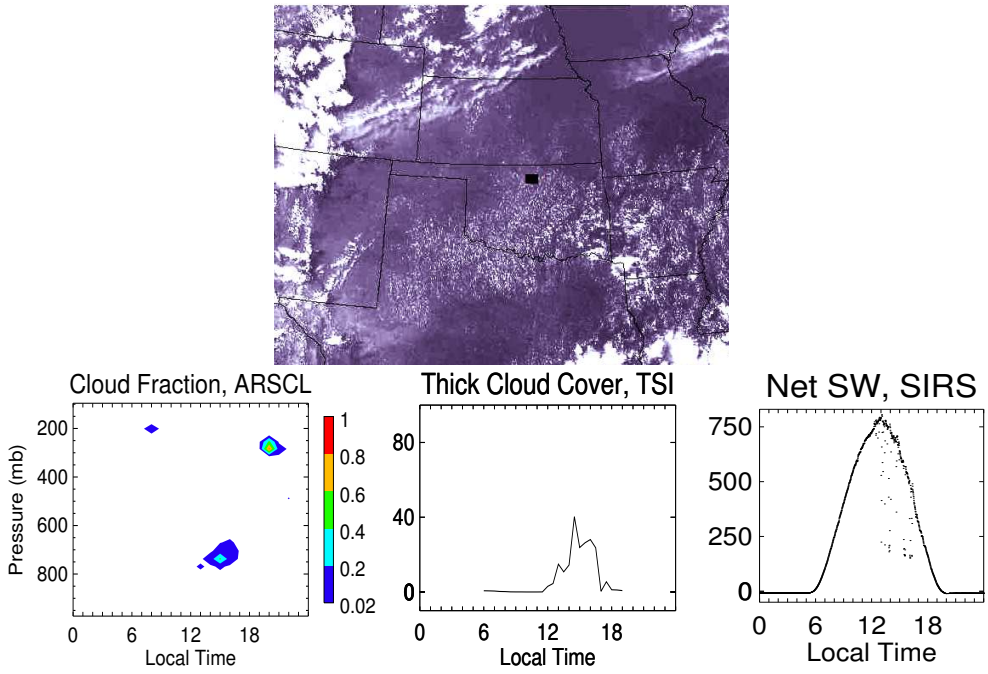

Figure 10: Measurements made on the 26 July 2003, (a) GOES visible channel at 14*LT (from NASA, the added black central square locates C1), (b) ARSCL retrieval of cloud fraction with time and pressure, (c) time series of cloud fraction (%) inferred from the TSI, and (d) one minute measurements of the net solar flux* (Wm^{-2}) *at the surface.*

average, it is not surprising that the ARSCL treatment exhibits more a zero-or-one cloud fraction that the model. The monthly average should be comparable, though, and this suggests that the model tends to underpredict low and middle cloud cover. In particular, the frequent occurrence of shallow clouds (with relatively low cloud fractions) around $800-700mb$ (at boundary layer top) seems to be missed by the model.

Insects in the day-time boundary layer may be detected by the cloud radar used in the ARSCL retrieval. Fig 10 compares the ARSCL detection with other instruments measurements for the 26 July 2003. According to the ARSCL, this day is typical of boundary layer cloudiness, with an afternoon cloud cover of $20-40\%$ at $750mb$, and no high clouds. The Langley Research Center (NASA) provides an archive of GOES-satellite pictures centered on the ARM sites[2]. The visible channel of GOES-Midwest (spatial resolution of $4*4km$) clearly shows some scattered clouds extending all over Oklahoma (Fig 10a). The InfraRed channel (not shown) shows no particular signal in that region, suggesting low cloud tops. The ARM Total Sky Imager (TSI) has also been used, restricting to the central quadrants (i.e. avoiding low zenith angles contributions, but still gaining from the solid angle detection of the instrument). The cloud cover has been inferred by one-hour averaging of the ratio between the number of opaque cloudy pixels and clear pixels. This simple retrieval is consistent with the ARSCL (Fig 10c). Finally, the net incoming Short-Wave flux (based on a 1-minute sampling) also shows a strong time variability, with sunny and shadow periods alternated, typical from small scattered clouds (Fig 10d). This analysis has been repeated for all the days with boundary layer clouds, and the agreement between all instruments remains as good. Hence both the ARSCL and GOES products show some skill in detecting shallow cumuli over the SGP, whereas a rather simple treatment of the TSI produces realistic matches with the ARSCL for these clouds.

Fig 11 shows the time series of the precipitable water content (PWC) in the IFS, inferred from the radiosondes measurements, and retrieved by a MicroWave Radiometer (MWR). The two observing systems and the model

[2]See the related Web page: http://www-angler.larc.nasa.gov/armsgp/g8sgp.html

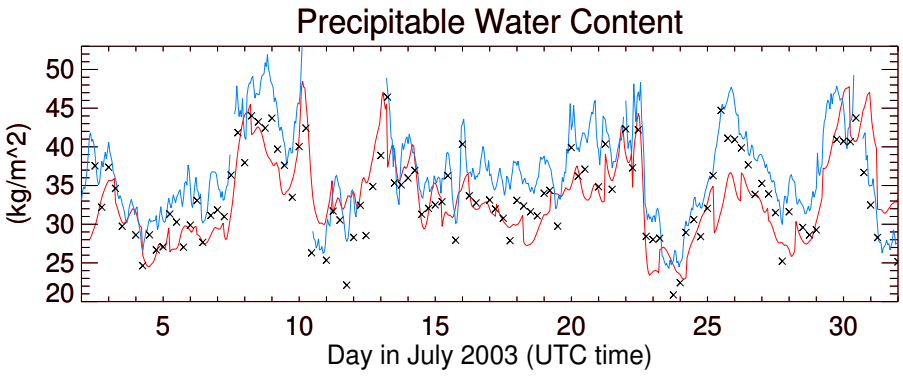

Figure 11: Precipitable Water Content in the IFS (red) and as measured by the MWR (blue). Missing data in the MWR time series correspond to rain events at C1 (no measurements possible).

show a very good agreement in terms of time variability during the month. However, the radiosondes PWC is always underestimated compared to the MWR, by about 8% (the monthly averages are 36.6 kgm^{-2} for the MWR, 33.7 kgm^{-2} for the sondes). This feature probably relates to the generally dry bias of radiosondes as reported by Turner et al. (2003). Whereas ARM radiosondes are not assimilated in the IFS, other radiosondes are, which might explain why the IFS shows a similar negative bias (monthly average of 33.0 kgm^{-2}). In order to improve the information provided by the radiosondes, each ARM radiosonde specific humidity profile has been scaled to match the MWR corresponding observation (following the recommendation of Turner et al., 2003). The applied artificial scaling neglects the dependence of the bias on pressure or temperature, and in theory may produce relative humidity exceeding 100%. The re-scaling procedure is still believed to be adequate to understand the vertical structure of the model bias, because the artificially supersaturated cases were found to occur only occasionally.

If the model dry bias were only caused by the radiosondes bias, it would keep a constant sign with height. According to Fig 12, the model bias is located at low and middle levels, whereas upper model levels tend to show an excess. This change of sign with height stresses the impact of meteorological processes along the numerical integration. The relative humidity underestimation at lower levels is consistent with the surface layer dry bias reported above. The lower levels are the main contributors to the PWC, which explains the underestimation of the latter in the IFS. Although the upper levels moist bias (in relative humidity) is expected to produce a cloud cover overestimation, this is actually not the case in the IFS (see above Fig 9).

4.3 Boundary layer properties

In order to assess the boundary layer characteristics in more details, we have empirically defined two weather regimes. The first regime consists of fair-weather days that combine (1) no middle-to-high clouds, and (2) dry convection or shallow moist convection. It includes eighteen days: 2-6, 11, 13-18, 23-27, and 31 July (see Fig 9). The other 12 days of July 2003 (excluding the 1st July, see above) show a more perturbed activity. The fair-weather regime is supposed to characterize situations with a strong vertical entrainment at the top of the boundary layer. Fig 13 shows the potential temperature and specific humidity profiles in the boundary layer and lower troposphere for the two regimes. Although averaging over days smooths inversions, the daytime growth of the well-mixed boundary layer is obvious in both regimes. Still, the disturbed regime is significantly more moist above 800mb, and colder during the day. These features can be interpreted as a cause (a moist atmosphere favors deep convection) or as a result (detrainment and evaporation of liquid water) of the disturbed weather. Satisfactorily, the IFS captures these qualitative differences between the two regimes.

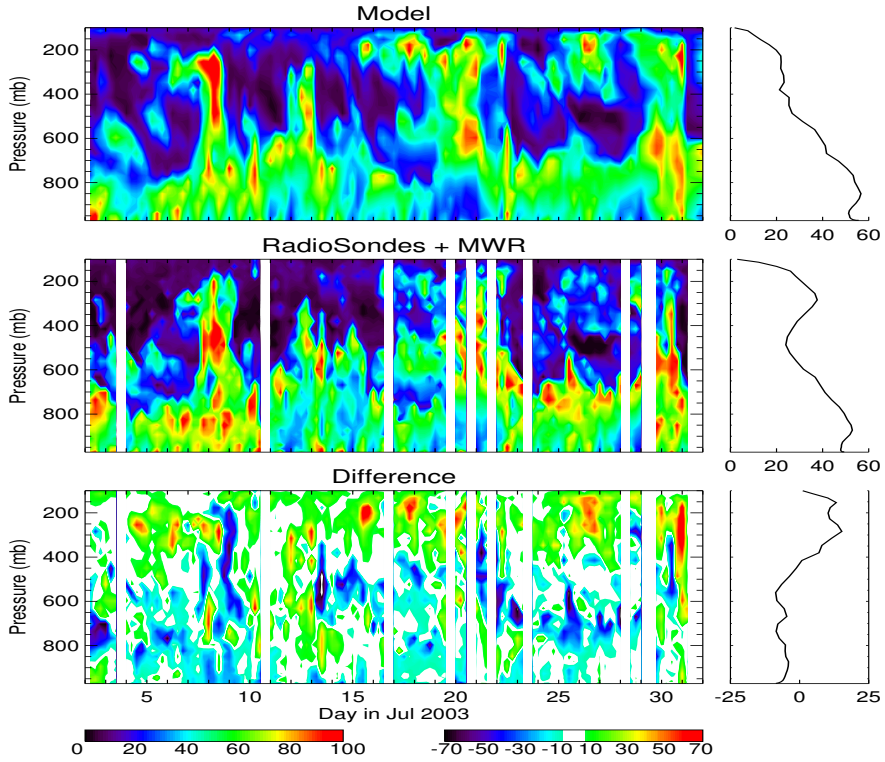

Figure 12: Time series of the relative humidity (%) at the SGP C1, according to (top) simulated by IFS, (middle) measured by the radiosondes (with a scaling to MWR, see text), and (bottom) difference. Right panels shows the monthly means. Colorcode for the top and middle (resp. bottom) figures is shown on the left (resp. right).

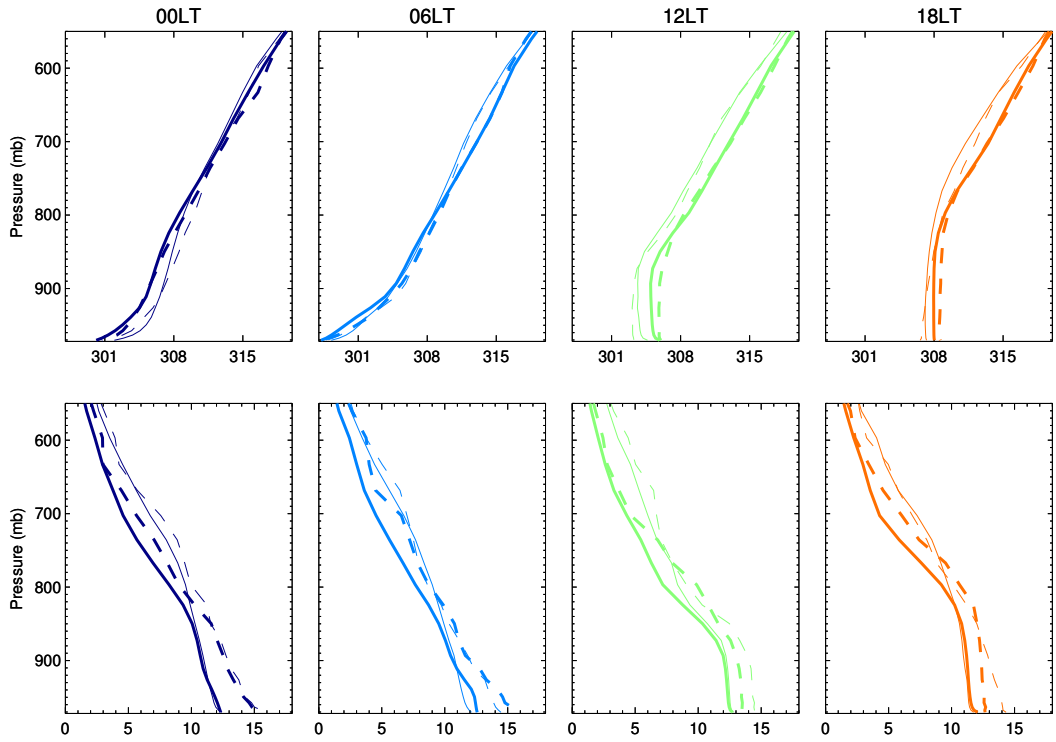

Figure 13: Monthly averaged vertical profiles of temperature (K) and specific humidity (gkg^{-1}) according to pressure and time of the day, in the IFS (full lines) and averaging over radiosondes (dashed lines). Thick lines: fair-weather days, thin lines: perturbed weather days (see text).

In terms of moisture, there is a dry bias of the IFS below $800mb$ in both regimes - although somewhat stronger in the fair-weather regime. It is maximum at low levels at night and in the morning. It spreads upward as the boundary layer grows, which leads to an apparent reduction of the bias at lower levels (see also the $2m-$specific humidity comparison). Hence, as far as moisture is concerned, the emerging picture does not necessarily incriminate a local overestimation of day-time boundary layer top entrainment. It can rather be interpreted as a lack of moistening at night (e.g. resulting from advection errors), that propagates upward later in the day.

The modelled monthly mean temperature (not shown) reveals a model cold bias of $-0.4K$ up to $600mb$. Fig 13 gives some insight on that diagnostic. In the fair-weather regime, the IFS day-time forecasts are clearly too cold in the boundary layer, consistently with an underestimation of the boundary layer height. However, the bias is opposite for disturbed days. Hence the lack of boundary layer top entrainment in fair-weather days is thought to explain the IFS monthly mean cold bias in the boundary layer. A day-by-day examination of the fair-weather regime supports the idea that days with dry convection and days with cumulus-topped boundary layer equally contribute to this deficiency. As already mentioned, the IFS underestimates the monthly mean day-time $2m-$temperature maximum. According to Fig 13, the model super-adiabatic layer is too pronounced, which tends to reduce the cold bias at the surface. Hence the two errors (lack of entrainment, overestimated lapse rate) partly compensate for the $2m$-temperature. This is typical of meteorological fields for which observational data heavily impact the model assimilation and evaluation.

In order to get an insight on the frequency of fair-weather cumuli at C1, Fig 14 shows the diurnal cycle of cloudy hours (defined as hours in which at least one cloud is detected), as retrieved by ARSCL. Four years are displayed (from 2000 to 2003), in order to illustrate the year-to-year variability of July months. The most striking and robust feature of Fig 14 is that all years show a high frequency of afternoon clouds extending from 1500 to $3000m$. The boundary layer growth can even be tracked for some years (e.g. 2002, 2003), through the rise of cloud base and top during the morning. Guichard et al. (2002) also report a frequent presence of boundary layer clouds over the SGP site in Spring 1998. Since the cloud frequency contains no information on the associated cloud fraction, a day-by-day analysis has been performed, which shows that at least seven among the days that form this frequency maximum are typical of a defined shallow cumulus activity (e.g. see Fig 9).

An unexpected feature of Fig 14 is that high clouds frequently reach C1 in the early morning (around $4am$). It varies from year to year, with an apparent shift to the afternoon in 2002. One may speculate that these high clouds are attached to a vertical structure (note the higher cloud frequency at $4km$, e.g. in 2000). Continental deep convection is the most likely process susceptible to cause a diurnal cycle of high cloud cover (through anvils), but it is expected to initiate in the afternoon / evening. Hence the early morning peak supports the idea that remotely initiated convective systems are responsible for the observed signature at C1. This will be investigated below.

Figure 15 compares the cloud frequency forecast by IFS and retrieved by ARSCL. It uses the model pressure levels to perform the vertical discretization, which explains the (minor) differences with Fig 14 with respect to the ARSCL retrieval. First, the night-time high cloud frequency is only qualitatively represented. Second, the IFS produces some clouds in the afternoon, but they are higher and deeper than the expected shallow cumuli (e.g. typically above 700 mb). This suggests that the IFS fundamentally misses the development of shallow clouds over the SGP site, and tends to replace it by locally-generated middle or deep convection. This conclusion is fully consistent with the findings above, and the conclusions of Mace et al. (1998), who compare measurements by the ARM cloud radar and IFS forecasts in winter 1997.

One may briefly recall some of the results discussed above with respect to the boundary layer. First, the IFS underestimates the nocturnal LLJ. The representation of the stable boundary layer at night (too thick) clearly contributes to this deficiency. Second, the model atmosphere is found to be generally too dry at lower levels,

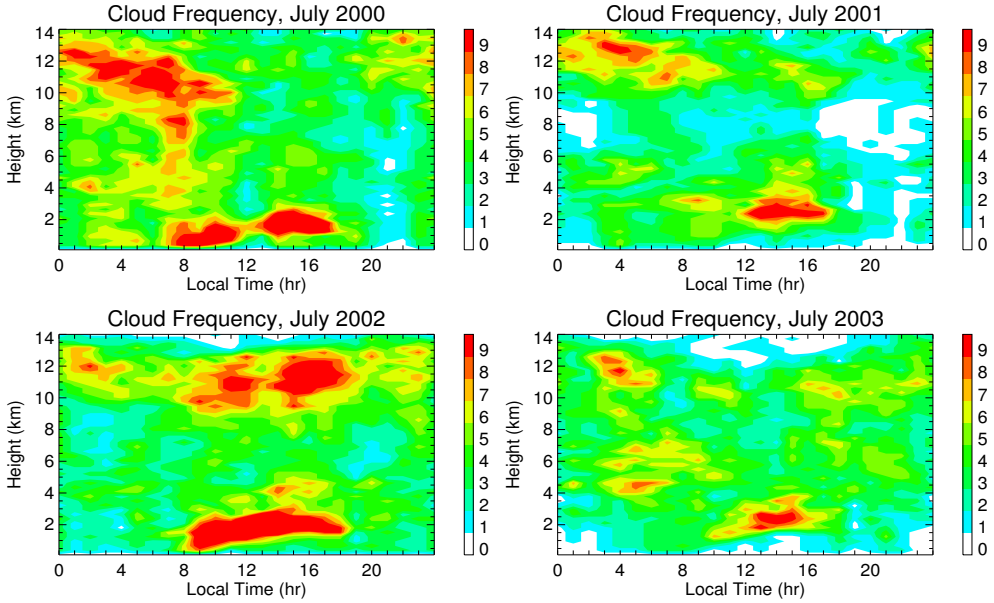

Figure 14: Number of cloudy hours with height and time of the day, according to radar-based retrieval (ARSCL product), for July months of 2000, 2001, 2002 and 2003.

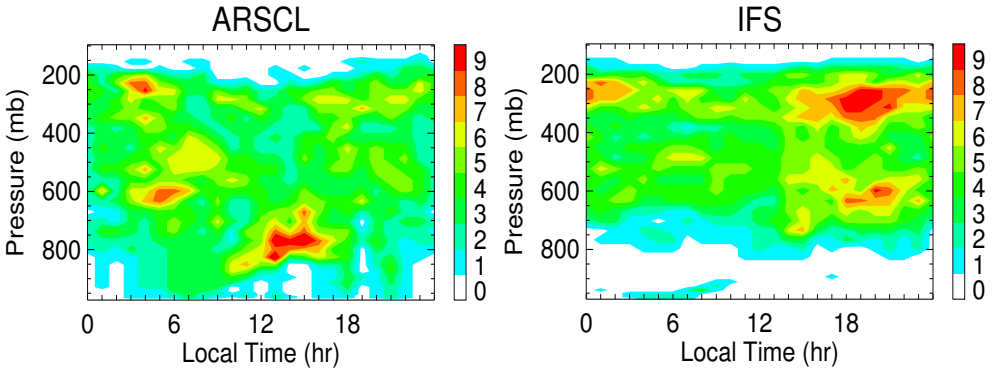

Figure 15: Number of cloudy hours with pressure and time of the day in July 2003, according to radar-based retrieval (ARSCL product, left) and to the IFS (right).

whereas the IFS misses the frequent development of fair-weather cumuli. It can of course be speculated that the formation of shallow cumuli is hindered by the boundary layer dry bias. In all cases, both the dry bias and the lack of shallow cumuli contribute to the overestimation of net incoming solar radiation (lack of absorption and reflexion). This excessive heat is transmitted to the lower atmospheric levels through the sensible heat flux, but these levels are still too cold in the IFS during fair-weather days, as the model locally lacks some vertical entrainment at the top of the fair-weather boundary layer. In clear days, the lack of entrainment directly points back to a deficiency in the IFS turbulence scheme (see also below). The lack of entrainment in shallow convection cases is more expected since the model misses the afternoon shallow cumuli.

5 Impact of remote processes

The IFS misrepresents the diurnal cycle of high cloud cover at the SGP site. As mentioned above, advected convective processes are susceptible to modulate this aspect of the observations. In relation, it was noted that the model tends to misrepresent the precise intensity and phasing of the rain events. As far as the boundary layer is concerned, two deficiencies emerge (LLJ too weak, lower levels too dry). They may be related, as the dry bias at lower levels may be attributed to the underestimated LLJ advection of moist air from the Gulf of Mexico. The present section intends to clarify the impact of these advection processes on the IFS behavior.

5.1 Propagating deep convection

Visual inspection of the GOES data (InfraRed channel, see above) suggests that the early morning high clouds relate to convective systems generated over the Rockies in the afternoon (e.g. left of Fig 10a), after being advected Eastwards along the Jet over the SGP. In order to validate this picture, the following analysis relies on the so-called D2 data produced in the framework of the International Satellite Cloud Climatology Project (ISCCP). The data has been collected and processed by the NASA Goddard Institute for Space Studies[3], and is available at the global-scale, with a spatial resolution of 280 km and a temporal resolution of 3 hours. Figure 16 shows the amplitude and the phase of the diurnal cycle of high cloud cover over North America. Data are not shown when the amplitude of the diurnal cycle is less than 4%. Note that such a diurnal cycle analysis is not informative on the cloud fraction daily average, as it focuses on contrasts between different times of the day.

ISCCP D2 data are not available after 2001, and July 2000 is here selected for its representativeness. Inspection of July months of 1998, 1999, 2000 and 2001 leads to the following conclusions: first, there is a small year-to-year variability in the diurnal cycle of the July high cloud cover. A late afternoon maximum is always found over Florida. The strong afternoon peak along California peninsula is also a systematic feature. The North and then Eastward development of the latter is always present, although its magnitude varies from year to year. The afternoon-to-night phase shift Eastwards is a robust characteristic of the four years. It strongly supports the idea that the cirrus in the Great Plains result from advection / propagation of deep convective towers (or associated anvils) along the Westerlies. This explains the large cloud frequency noted in the early morning in Fig 14. One may also speculate that the mid-level component results from the advection of detrained liquid water, whereas the high-level component results from the ice detrainment counterpart (faster advection at higher levels, i.e. reaching the SGP before). We have extended our analysis to the summertime Northern Hemisphere (not shown). The results suggest that a frequent advection of convective towers is generally found East (downstream) of mid-latitude mountains (Rockies, Alps, Himalaya).

For comparison, the IFS has been re-run for July 2000 with the same set-up as used above. Model results have

[3] see the related Web page: http://isccp.giss.nasa.gov/products/isccpDsets.html

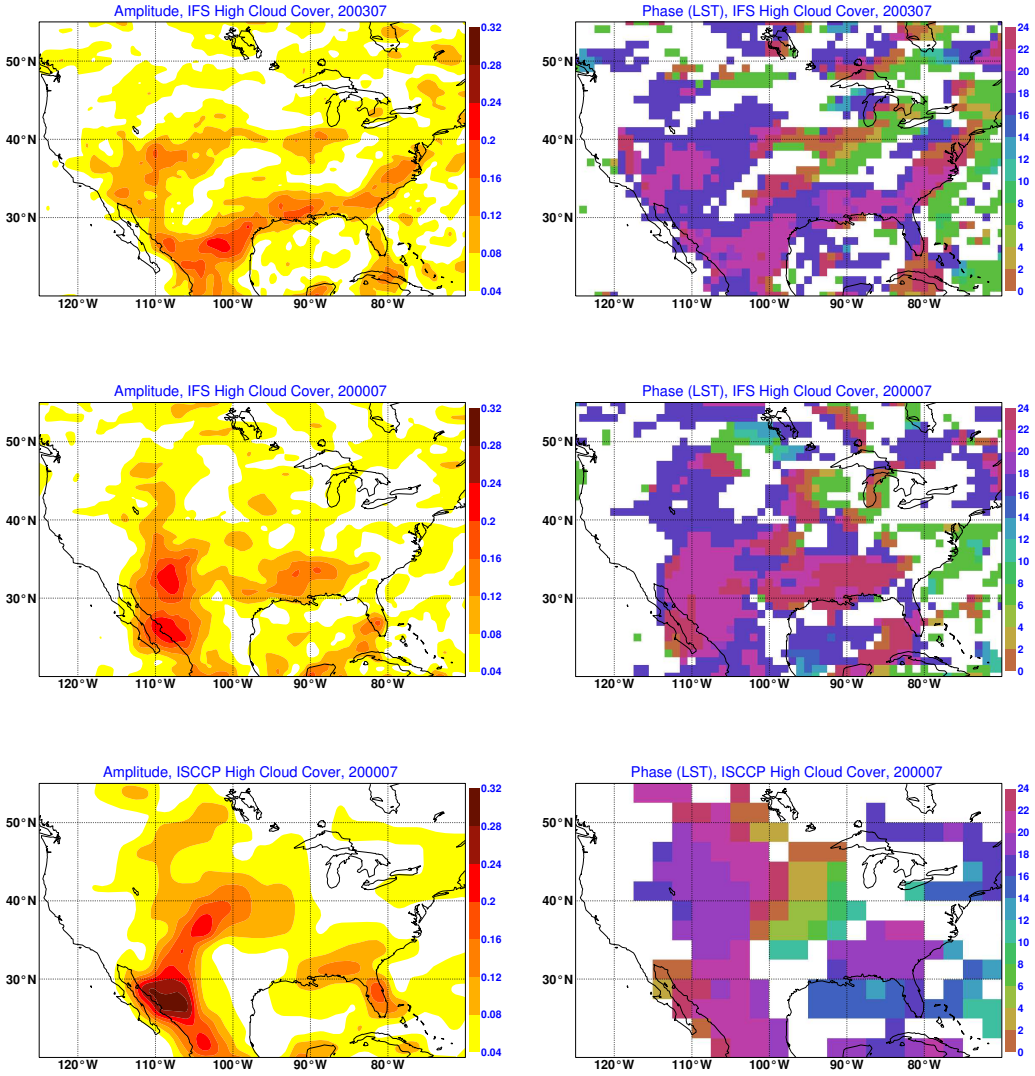

Figure 16: Amplitude (left) and phase (right) of the high cloud fraction in the IFS in July 2003 (top), in the IFS in July 2000 (middle), and according to ISCCP data (bottom). Note that the phase is expressed in longitude-dependant Local Time.

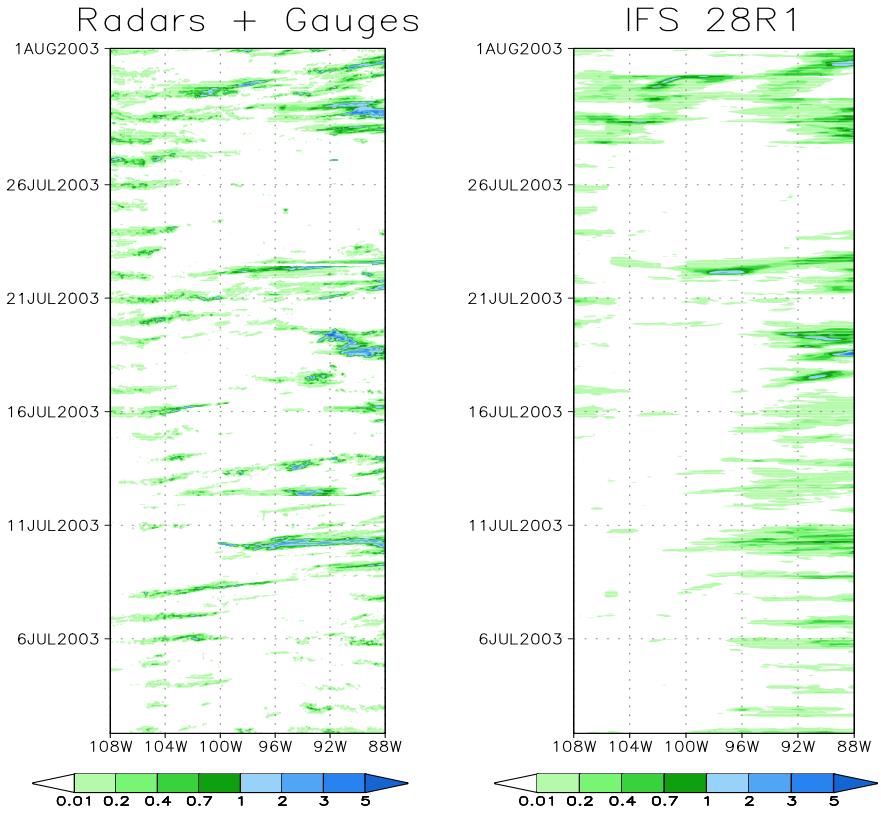

Figure 17: Hovmuller diagram of the precipitation rate (mm/hr) averaged over the latitudes from 35N to 40N in July 2003. NCEP Stage 4 mosaic observational data (left), and IFS forecasts (right).

been output on a degraded resolution of $1 * 1$ degree and every 3 hours. Extracted simulation hours are here $[12;36]$, so that the late afternoons and nights are produced by a same simulation. Considering July 2000, the model captures the Pacific Mexican coast maximum. It misrepresents the North Eastwards development in terms of amplitude, but the phase shift Eastwards in the Great Plains is captured. Now examining 2003, the amplitude maximum predicted by the IFS is shifted East of $105W$, which no year examined with ISCCP data supports. The IFS still predicts the afternoon generation of high clouds over the Rockies and their overnight Eastwards advection. Compared to July 2000, the IFS diurnal cycle in July 2003 is reduced over the SGP, in agreement with the ARSCL retrievals of Fig 14. Focusing on the IFS forecasts at ARM site in 2003, Fig 16 suggests that the high cloud frequency maximum in the afternoon (Fig 15) does not significantly impact the diurnal cycle of high cloud cover. This suggests that early morning advected high clouds are less frequent but with larger cloud fractions than afternoon high clouds.

In order to characterize the displacement of the convective systems themselves (rather than the high clouds they produce), it is appropriate to consider rain patterns. The mosaic data produced at NCEP by merging rain radars and gauges has been used[4](so-called Stage *IV*). The data has a spatial resolution of $4 * 4$km, and a temporal resolution of one hour. The spatial coverage includes the ARM SGP domain, at the exception of a small part at the extreme West of the domain. The temporal coverage also suffers from missing data (e.g. 12 and 22 July). Fig 17 shows the precipitation rate averaged over 35N to 40N, from 108W to 88W, for July (Hovmuller diagram). It emphasizes the very frequent rain events in the afternoon over the Rockies (around 105W), and their Eastwards propagation / advection (e.g. Wallace, 1975, Dai et al., 1999, among many others). These results are fully consistent with the cloud fraction diurnal cycle analysis above. They suggest that the local

transition from shallow to deep convection is not the most frequent situation for rain in the Great Plains. The statistical picture rather relates to orography-induced convective systems passing over the SGP. As shown on Fig 15, the ARM site-measurements often capture the mature stage of the propagation.

Fig 17 also shows the model counterpart, here again with extracted hours ($[8;32]$) chosen such that the afternoon and evening are produced by the same simulations. The IFS qualitatively captures some locally-generated rainy deep convective events over the Rockies. It however clearly underestimates their frequency. It also misses the Eastward advection/propagation of the rain patterns. This misrepresentation either suggests that the model is not able to produce propagating convective systems in the Great Plains, or that it produces convective systems that do not rain enough (as found at C1, see above). For example, the rain event of 29 July is triggered at 103W, in agreement with the observations. But it does not propagate in the IFS as fast as observed. The next simulation recovers from this deficiency thanks to the assimilation system, which creates an (irrealistic) discontinuity in the visual pattern of Fig 17. As mentioned above, the model reproduces the tropospheric winds, so this deficiency is likely to relate to the convection parameterization (or its feedbacks with other atmospheric processes). This marked underestimation of the convective activity leads to a lack of heating, which is expected to propagate as forecast errors over the North Atlantic. Interestingly, the IFS tends to overpredict the frequency of light rain afternoons at the East of the SGP domain, precisely where ratios between latent and sensible heat fluxes were found larger than observed.

In conclusion, the following picture emerges. The IFS tends to underpredict the frequency and intensity of afternoon raining convective systems over the Rockies. Whereas it captures the detrained cirrus advection Eastwards, it misrepresents the overnight propagation of the raining convective systems themselves over the Great Plains. On the contrary, it frequently generates some afternoon middle or deep convection over the SGP, with light rain. Although they may not have a large horizontal extent, the associated high clouds reduce the early morning vs. afternoon contrast in high cloud coverage. This explains the low amplitude of the diurnal cycle of high cloud cover over the SGP in the IFS.

5.2 Boundary layer moisture advection

As suggested above, the IFS may misrepresent the LLJ moisture advection along the Gulf of Mexico-to-SGP trajectory. A first hypothesis is that the trajectory over land may start with too dry an atmosphere over the Gulf of Mexico. A dry bias has been seen in the model climate (based on 12 month integrations) in the Gulf of Mexico (at least $-2\,kgm^{-2}$ in IFS Cycle 26R1, according to Jung and Tompkins, 2003). Further assessment of that latter bias is clearly required in order to understand the climatology of sub-Tropical and Central America as simulated by the IFS. A second hypothesis is that the IFS dry bias over the SGP is caused by a too slow / too diluted LLJ along the path. The implicit assumption in this view is that the moisture at C1 results from the competition between (1) the LLJ moisture advection in the boundary layer, and (2) other dilution processes that damp the LLJ moisture. This picture is supported by the analysis of Fig 8. The LLJ reinforcement in STBL suggests that (moist) air reaching the SGP from the Gulf of Mexico has less time to experience dilution processes on its way. This could explain why STBL better reproduces the observed increase of the $2m-$specific humidity at the nocturnal LLJ onset (e.g. the midnight value changes from $12.5\,gkg^{-1}$ in CTRL to more than $13\,gkg^{-1}$ in STBL). Without precluding an underestimation of the surface latent heat flux, convective processes were thought more likely to play a first-order role in the dilution, through precipitation at the surface or ventilation in the troposphere.

In order to test this Lagrangian view hypothesis, we hereafter focus on a domain from 26^oN to 40^oN and from 95^oW to 98^oW. The domain contains the SGP site and its Southwards (i.e. upwind) extension up to the Gulf

[4] See the related Web site: http://wwwt.emc.ncep.noaa.gov/mmb/ylin/pcpanl/

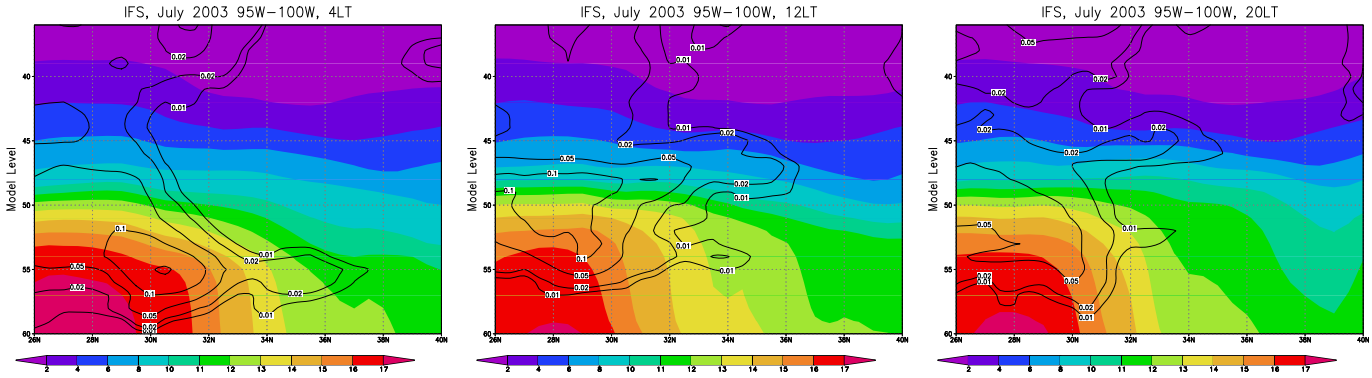

Figure 18: IFS forecasts with height and latitude, averaged over the fair-weather days of July 2003 and from 95W to 100W, at 4LT (left), 12LT (middle) and 20LT. Specific humidity (gkg^{-1}, shaded) and cloud cover (contoured).

of Mexico (Fig 1 and 7). It is basically a meridional cross-section, the zonal average is taken into account for representativeness. The IFS forecasts have been averaged over the fair-weather days defined above, noting that perturbed weather days show a similar sensitivity. Figure 18 displays the specific humidity and the cloud cover in the domain, at three times of the averaged forecast (hours 22, 30 and 38, i.e. 04LT, 12LT and 20LT). It reveals the vertical structure of the strong PWC gradient from the Gulf of Mexico (above $45\,kgm^{-2}$) to the SGP (less than $35\,kgm^{-2}$). It also emphasizes the strong convective activity over the Gulf of Mexico and the coast (see also Fig 10a, bottom right corner). The near-coastal deep convection is responsible for the peak in accumulated rain (Fig 19a). The land convection shows a well-marked diurnal cycle as the cumulus clouds form during the day.

In the sensitivity experiment labelled NOSH, the shallow convection parameterization has been shut down. This experiment is designed to assess the impact of this process (and its parameterization) on the moisture of the domain. As shown on Fig 19b, the shallow convection has a strong impact on the evaporation over oceans, even in our short-term forecasts. This result is consistent with the conclusions of Tiedtke et al. (1988). According to Fig 20, this is caused by the afternoon convection, which dries the lower boundary layer and moistens the lower free troposphere. The less water injection in NOSH leads to a deficit of $2\,kgm^{-2}$ in the PWC[5]over the Gulf of Mexico (Fig 19c). Whereas the latent heat flux is greatly reduced above the model level 45 (smaller convective towers), the rain does not change much, and actually slightly increases in NOSH. This result is consistent with the idea that the deep convection parameterization tends to replace the shallow convection scheme by generating some small rainy cumuli in NOSH. It also illustrates that shallow cumuli precede the development of deeper cumuli (e.g. Cheinet, 2004). The vertical anomaly of moisture comparably affects the in-land, noting that the PWC anomaly decreases Northwards and becomes negligible North of 34^oN. Hence the surface feedback (enhancement of evaporation) is weaker over land than over the ocean. Note that the moist anomaly in the lower boundary layer propagates with the LLJ during the night, whereas the upper boundary layer almost recovers from the dry bias by $4\,LT$.

In a second sensitivity experiment (LSCV), the parameterized cloud base mass flux is arbitrarily divided by two in all types of moist convection, but only over land. LSCV is intended to characterize the role of moist convective processes on the way from the ocean to the SGP, keeping an initial condition (over the ocean) relatively similar. This is actually the case, since the surface latent heat flux, rain and PWC remain similar to CTRL over the Gulf of Mexico. However, the forecasts are significantly changed over the coastal area, with a reduction of the rain and (thus) lesser surface evaporation (Fig 19a,b). Compared to NOSH, this illustrates

[5]As a crude indication, $1\,kgm^{-2}$ (or equivalently $1mm$ of rain) approximately corresponds to $1\,gkg^{-1}$ of specific humidity in the first $1000m$ of the atmosphere, and to a moisture flux of $\approx 30\,Wm^{-2}$ accumulated over 24 hours.

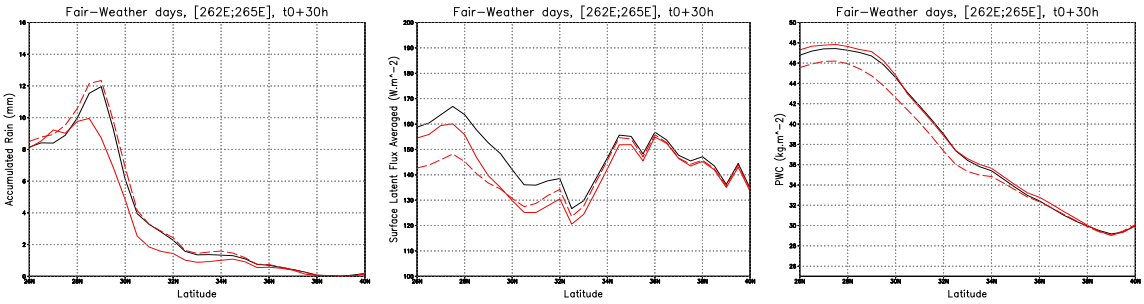

Figure 19: Same as Fig 18, but comparing CTRL, NOSH and LSCV. (a) Accumulated rain after 30 hours of model integration, (b) surface latent heat flux averaged over the 30 hours of integration, (c) PWC at the 30th hour of simulation. Line code is black: CTRL, full red: LSCV, dashed red: NOSH.

the role of deep convection parameterization in that area, i.e. to increase the water cycle (evaporation and precipitation) without significantly altering the PWC. This is a fundamental difference with the role of shallow convection parameterization noted above (leaving precipitation unchanged, enhancing the evaporation and the PWC). Here again, the absence of ventilation (i.e. of drying) of the lower boundary layer over the SGP yields a moist anomaly, that propagates Northwards with the LLJ. This anomaly is stronger than in NOSH, and yields a non negligible increase (more than $0.5\,kgm^{-2}$) of the PWC over the SGP.

These results clearly demonstrate the qualitative role of convective processes in diluting the LLJ moisture. However, the PWC increase in LSCV is still smaller than the dry bias noted over C1 (see above). Hence, if the LLJ wind were correct, the overactivity of the modeled moist convective processes over land would not be likely to explain by themselves the IFS bias in PWC over the ARM site. Comparably, the moistening in STBL restricts to the lower $200-300m$ of the atmosphere, and does not significantly impact the PWC at C1. One should keep in mind that the LLJ is still underestimated in STBL and LSCV, potentially overexposing the LLJ to convective processes. Hence, in terms of phenomenology, the competition between LLJ advection and convective processes remains as a likely mechanism of remote modulation of the boundary layer moisture at C1. In terms of the IFS performance, the model dry bias at lower levels could thus relate to a combination of a too slow LLJ and an overestimated coastal convective activity.

6 Summary and discussion

This study compares the IFS short-term forecasts (Cycle 28R1) with a posteriori observations. The focus is on mixing and cloud processes over one continental site (Northern America SGP) in summertime (July 2003). The ARM measurements prove to be highly suitable to evaluate many aspects of the physical package of the model. The quality of the data, together with its extensive temporal coverage, is clearly adequate for this study. Multiple observations of some features are also helpful in building confidence in the data (e.g. Fig 10). Indeed, in the case of spatially highly heterogeneous meteorological fields (rain, surface fluxes), it is appropriate to average over the ARM SGP stations network. Even with this procedure, it appears that our conclusions with respect to these fields should be complemented by extended data coverage. For example, the water cycle at the surface is a long-memory process, and should preferably be analyzed on a seasonal basis. Also, as soon as advective (remote) processes play a non negligible role, it is necessary to combine site measurements analysis with further considerations of the synoptic flow.

Our results document several deficiencies in the IFS forecasts with respect to the SGP summertime meteorology. First, the IFS underestimates the nocturnal LLJ. The representation of the stable boundary layer at

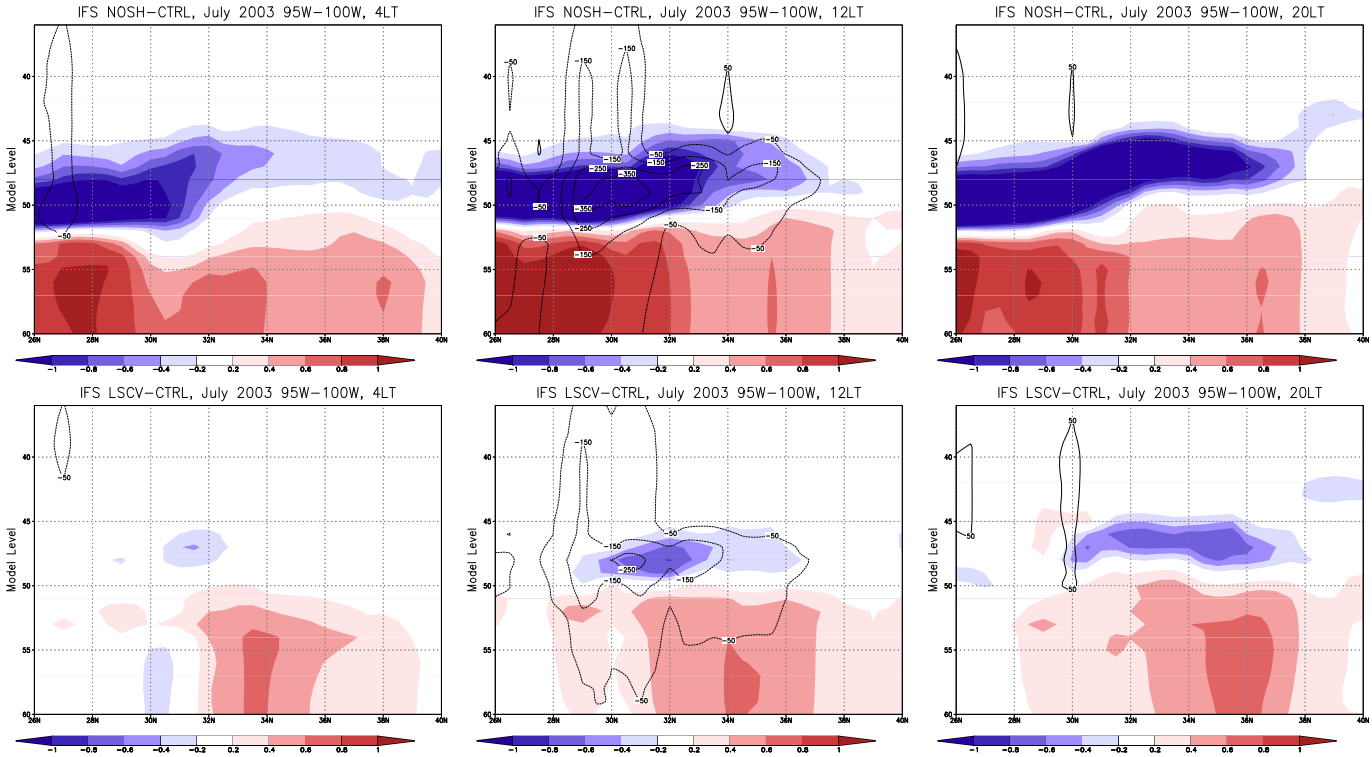

Figure 20: Same as Fig 18, but comparing CTRL, NOSH and LSCV. Differences between NOSH and CTRL in specific humidity (gkg^{-1}, shaded) and instantaneous latent heat flux (Wm^2, contoured).

night clearly contributes to this deficiency. More specifically, the stable boundary layer mixing is found to be overactive, leading to a warm bias and an underestimation of the wind at night. An alternative formulation previously documented cancels the warm bias and partly overcomes the wind underestimation. Second, the model atmosphere is found to be generally too dry. A dry bias from the assimilated radiosondes could play a role, but it can not explain the vertical structure of the model bias (concentrated at lower levels). This low level dry bias presumably explains why the IFS misses the very frequent development of shallow cumuli in fair-weather afternoons. Both the dry bias and the lack of shallow cumuli contribute to the overestimation of net incoming solar radiation. This excessive heat is transmitted to the lower atmospheric levels through the sensible heat flux. Nevertheless, these levels are still too cold in the IFS in fair-weather days, which suggests that the model lacks some vertical entrainment at the top of the fair weather boundary layer.

In clear days, the lack of entrainment points back to a deficiency in the representation of the dry convective boundary layer. A new turbulence scheme based on mixing of moist conserved variables (Tompkins et al., 2004) has been evaluated in our context, and it apparently does not improve the forecasts for the ARM SGP site. A lack of vertical mixing due to shallow convection cases is more likely, since the model misses the afternoon shallow cumuli - due to the dry bias (see above). In our context, this lower levels dry bias precludes a direct evaluation of the IFS representation of cumulus-topped boundary layer at C1. Generally speaking, the analysis of local processes is complicated by any uncertainty on the remote processes that pre-condition the local atmosphere. The Single-Column Modelling approach prescribes the advective forcings, so it would be appropriate to evaluate the parameterization of shallow convection in the IFS (e.g. Lenderink et al., 2005). Still, our study suggests that the first-order IFS deficiency with respect to the summertime shallow convection at C1 may not be the representation of the local processes.

The two emergent deficiencies (LLJ too weak, lower levels too dry) may be related, as the dry bias at lower

levels may be attributed to the underestimated LLJ advection of moist air from the Gulf of Mexico. The implicit physical view is that the boundary layer moisture at C1 results from the competition between the LLJ moisture advection and processes that damp the LLJ moisture along the path. The hypothesis considered in this study is that the moisture underestimation at C1 is caused by an underestimation of the LLJ combined with an overactive coastal convective activity. Sensitivity experiments performed in this study show that the boundary layer moisture at C1 increases with a more active LLJ, or with less active moist convective processes between the Gulf of Mexico and the SGP. Our results thus qualitatively support the above mechanism. Quantitative evaluation is difficult, because the IFS is not presently able to realistically reproduce the strength of the LLJ at C1. In the proposed mechanism, the local lack of cumuli results from a remote overestimation of cumuli. This paradoxical view emphasizes the complexity of parameterization evaluation and development in a NWP model.

Other model deficiencies may also play a role, but have not been tested in this study. For example, the trajectory over land may actually start with a too dry atmosphere over the Gulf of Mexico. The model may also underestimate the surface latent heat flux along the path to the SGP. In all cases, understanding the dry bias of the model over the SGP (and over the Gulf of Mexico) is the next step to refine the IFS representation of the SGP meteorology. In that perspective, our study suggests a possible methodology to calibrate parameterizations over the SGP, based on the analysis of their impact integrated along the air mass trajectory at the synoptic scale. More specifically, the role of physical processes on the continental boundary layer can be evaluated provided observations upstream (in the Gulf of Mexico) and downstream (in the SGP). This is possible because of the systematic and well-defined structure of the boundary layer flow, with a differential advection according to height (role of the Low Level Jet). As shown here, ARM measurements could account for the downwind component of such a calibration.

Finally, our study has also focused on the diurnal cycle of deep convection over the SGP. According to various observational data sets (satellite, reanalysis of rain radars and gauges, ARM cloud radar), the diurnal cycle of deep convection over the SGP is strongly modulated by the advection / propagation of convective storms generated over the Rockies. From the deep convection point of view, the ARM site is not statistically representative of locally-generated deep cumuli. Here again, ARM measurements document a specific stage of some processes essentially marked by meso- to synoptic-scale dynamics. Our results show that the IFS misrepresents rain events that are advected over the SGP, although it captures the advection of the generated high clouds. Addressing the underlying model issues appears to be a major scientific challenge, which involves not only the deep convection parameterization but also the interactions between parameterized and resolved scales in a GCM.

Acknowledgments

This research was sponsored by an ARM fellowship at ECMWF, under program element 11-7529. S. Cheinet sincerely wishes to acknowledge the USA Department Of Energy and more specifically the ARM program for this opportunity. He thanks the ECMWF staff for maintaining an exceptional working environment in all respects. The help by A. Tompkins (ECMWF) and C. Jakob (BMRC/ARM) in post-processing some data shown in this study, the prompt and efficient support from K. Costulis (NASA), R. Mc Cord (ORNL) and D. Cook (ANL) in analyzing the data, and the comments by Martin Miller on the manuscript, are very much appreciated. GOES satellite archive over the SGP, made available by the NASA, helped this study. The amplitude / phase treatment of ISCCP data was inspired by a discussion with C. Rocken (UCAR). Electronic correspondence with S. Cheinet can be addressed to sylvain.cheinet@m4x.org.

Appendix: vertical levels in the IFS

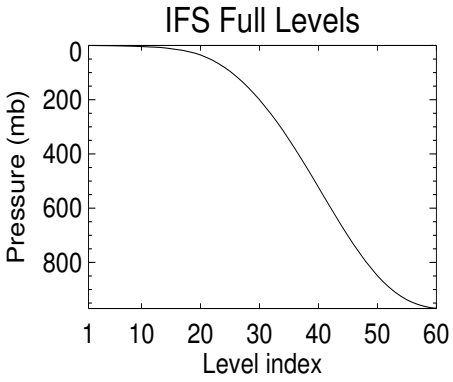

Figure 21: Equivalence between the IFS full levels index and pressure on the 2nd of July at 16*LT*.

The IFS pressure at half levels (index i, $i = [0,60]$) are determined following $P_i = a_i + b_i P_s$, where a (in hPa) and b are a-priori fixed coefficients. Pressure at full model levels results from an interpolation of half levels. The equivalence between levels index and pressure is displayed on fig 21 for one precise time of July 2003 at C1 (altitude 313m). The surface pressure varies by less than 1% at C1 in July 2003, so this equivalence remains virtually unchanged in our study.

References

Beljaars, A. (2001). Issues in boundary layer parametrization for large scale models. *Proc. ECMWF seminar on Key Issues in the Parametrization of Subgrid Physical Processes.* ECMWF, Reading (UK), 71–87.

Beljaars, A. and P. Viterbo (1998). The role of the boundary layer in a numerical weather prediction model. *In "Clear and Cloudy Boundary Layers", Kluwer Academic Publishers, Amsterdam*, 287–304.

Betts, A. and C. Jakob (2002a). Evaluation of the diurnal cycle of precipitation, surface thermodynamics and surface fluxes in the ecmwf model using lba data. *J. Geophys. Res. 107*, 8045–8053.

Betts, A. and C. Jakob (2002b). study of the diurnal cycle of convective precipitation over amazonia using a single column model. *J. Geophys. Res. 107*, 4732–4757.

Cheinet, S. (2004). A multiple mass-flux parameterization for the surface-generated convection. part 2: Cloudy cores. *J. Atmos. Sci. 61*, 1093–1113.

Clothiaux, E., T. Ackerman, G. Mace, K. Moran, R. Marchand, M. Miller, and B. Martner (2000). Objective determination of cloud heights and radar reflectivities using a combination of activ remote sensors at the arm cart sites. *J. Appl. Meteorol. 39*, 645–665.

Dai, A., F. Giorgi, and K. Trenberth (1999). Observed and model-simulated diurnal cycles of precipitation over the contiguous united states. *J. Geophys. Res. 104*, 6377–6402.

Garratt, J. R. (1992). *The Atmospheric Boundary Layer*. Cambridge University Press.

Guichard, F., D. Parsons, J. Dudhia, and J. Bresch (2002). Evaluating mesoscale model predictions of clouds and radiation with sgp arm data over a seasonal timescale. *Mon. Weather. Rev. 131*, 926–944.

Higgins, R. W., Y. Yao, E. S. Yarosh, J. E. Janowiak, and K. C. Mo (1997). Influence of the great plains low-level jet on summertime precipitation and moisture transport over the central united states. *J. Climate. 10*, 481–507.

Jakob, C., R. Pincus, C. Hannay, and K.-M. Xu (2004). The use of cloud radar observations for model evaluation: A probabilistic approach. *J. Geophys. Res. 109*, DO3203, 10.1029/2003JD003473.

Jung, T. and A. Tompkins (2003). Systematic errors in the ecmwf forecasting system. *ECMWF Technical Memorandum No 422*, 72.

Lenderink, G., P. Siebesma, S. Cheinet, S. Irons, C. Jones, P. Marquet, F. Muller, D. Olmeda, J. Calvo, E. Sanchez, and P. Soares (2005). The diurnal cycle of shallow convection over land: a single-column models intercomparison study. *Quart. J. Roy. Meteor. Soc. 130*, 3339–3364.

Mace, G., C. Jakob, and K. Moran (1998). Validation of hydrometeor occurence predicted by the ecmwf model using millimeter wave radar data. *Geophy. Res. Letters 25*, 1645–1648.

Morcrette, J.-J. (2002). Assessment of the ecmwf model cloudiness and surface radiation fields at the arm sgp site. *Mon. Weather. Rev. 130*, 257–277.

Tiedtke, M., W. Heckley, and J. Slingo (1988). Tropical forecasting at ECMWF: the influence of physical parameterization on the mean structure of forecasts and analyses. *Quart. J. Roy. Meteor. Soc. 114*, 639–664.

Tompkins, A., P. Bechtold, A. Beljaars, A. Benedetti, S. Cheinet, M. Janiskova, M. Kohler, P. Lopez, and J.-J. Morcrette (2004). Moist physical processes in the ifs: Progress and plans. *ECMWF Technical Memorandum No 452*, 91.

Turner, D. (2003). Dry bias and variability in vaisala rs80-h radiosondes: The arm experience. *J. Atmos. Oce. Techn. 20*, 117–132.

Wallace, J. M. (1975). Diurnal variations in precipitation and thunderstorm frequency over the conterminous united states. *Mon. Weather. Rev. 103*, 406–419.