TECHNICAL MEMORANDUM

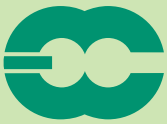
728

Comparison of NCEP Stage IV precipitation composites with ECMWF model

Philippe Lopez

Research Department

September 12, 2014



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

Series: ECMWF Technical Memoranda

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

Abstract

A systematic comparison of ECMWF precipitation short-range forecasts with NCEP Stage IV rainfall (NEXRAD) composites over the U.S.A. was performed over the period January 2002 to June 2014. Statistics show that the match between the model and NEXRAD observations has been regularly improving over the years, particularly in terms of mean biases, but also correlations and threat scores. Correlations exhibit a strong seasonal cycle with maximum values in winter and minimum values in summer, in all regions but the West Coast. Besides, precipitation tends to be slightly over-predicted all day long during the cold season. Conversely, the warm season is characterized by a strong underprediction in the first half of the night and a strong overestimation in the morning, except over the West Coast. A spin-down of precipitation during the first 12 hours of the forecast could be identified over the southern states in the summer. However, this spin-down does not change the sign of the precipitation mean biases between model and NEXRAD when comparing forecasts started at 0000 and 1200 UTC. Lastly, it was confirmed that the recent changes applied to the convective parametrization of the model (cycle 40r1) helped to reduce the long-standing large phase advance in the diurnal cycle of the model's summer precipitation over most regions.

1 Introduction

Over the past decades, several networks of ground-based radars providing precipitation estimates have been deployed over several continents, with typically between 160 and 200 radar sites each. The NEXRAD network has been operational over the U.S.A. since the mid-1990s (NEXRAD[1]; Fulton *et al.*1998). More recently, radar data from individual European countries have been merged on a continental-scale into the EUMETNET/OPERA network (Huuskonen *et al.* 2014). China has also installed its own large-scale network of weather radars (Bai 2013), while other less dense or less extended networks also exist in Australia and Japan, for instance.

In the U.S.A., NCEP has been producing 2D precipitation composites, the so-called NCEP Stage IV dataset, based on NEXRAD radar observations and rain gauge measurements, over the past 15 years (Lin and Mitchell 2005). A significant advantage of the NEXRAD network compared to its European equivalent (OPERA) lies in its greater homogeneity in terms of instrument characteristics (e.g. frequency, brand). This advantage made it possible to begin the assimilation of NCEP Stage IV 2D precipitation composites in ECMWF's operational 4D-Var system on 15 November 2011 (Lopez 2011). This latter study showed that the assimilation of these data could be beneficial to atmospheric analyses but also medium-range forecasts, and not only in terms of precipitation. The high spatial resolution and coverage of precipitation radar composites mean that they can also be very useful for validating precipitation forecasts obtained from numerical weather prediction (NWP) models. Their availability at high temporal frequency (hourly or shorter) also permits the assessment of the realism of the simulated diurnal cycle of precipitation, which has been one the major challenges in NWP so far, particularly in convective situations (Bechtold et al. 2014).

The present study therefore proposes a validation of ECMWF's operational short-range precipitation forecasts against NCEP Stage IV composites over the period 2002-2014. Statistics have been computed on both hourly and 6-hourly precipitation accumulations during the first day of the forecasts. It should be noted that the focus was deliberately put on the first 24 hours of the forecast to assess potential issues

[1] See list of acronyms in Appendix 1

of precipitation spin-up or spin-down and because the performance of the model at very short range has direct implications for the performance of 4D-Var precipitation data assimilation.

The NCEP Stage IV and ECMWF model datasets used in this study are described in section 2. Section 3 provides a brief reminder of the precipitation climatology over the U.S.A.. Results of the comparison of radar composites with ECMWF model short-range forecasts are presented in section 4. Section 5 summarizes the main findings of this study.

2 Datasets

2.1 Radar and gauge precipitation composites

The NCEP Stage IV precipitation observations used in this work are the result of the compositing of precipitation estimates from about 150 Doppler NEXt-generation RADars (NEXRAD) and about 5,500 hourly rain-gauge measurements over the conterminous USA (Baldwin and Mitchell 1996; Lin and Mitchell 2005). Technically speaking, NEXRAD corresponds to the so-called WSR-88D (Weather Surveillance Radar 1988 Doppler; Fulton *et al.* 1998). Each NCEP Stage IV precipitation analysis is initiated 35 min after the end of each hourly collection period and may be updated over a period of several hours with new data coming from the twelve USA regional centres. A first inflow of automatically generated precipitation data is available within a few hours after the accumulation time, while a second inflow of updated manually-quality-controlled data becomes available later (with a delay of up to 12 hours). In this work, the composites were obtained from the NCAR/UCAR/EOL archive (website: *http://data.eol.ucar.edu/cgi-bin/codiac/fgr_form/id=21.093*).

The original precipitation composites are available on a 4-km polar-stereographic grid, but for the purpose of comparison to the model, these were averaged onto the reduced Gaussian grid used at ECMWF (see section 2.2). The original NCEP Stage IV hourly precipitation accumulations were used to study the diurnal cycle of precipitation for recent years. However, the main focus will be laid on statistics of 6-hourly accumulations since hourly model outputs only became available in December 2011. Another justification is that six hours corresponds to the accumulation period used in the assimilation of ground-based precipitation data at ECMWF. In this respect, one should mention that after November 2011, NCEP Stage IV composites cannot be regarded as truly independent data for validating the model's short-range forecasts, since these observations started to be assimilated in ECMWF's 4D-Var system (Lopez 2011). In other words, from November 2011 onwards, the statistics presented here provide a validation of the forecast model and the precipitation data assimilation procedure altogether. For simplicity, the NCEP Stage IV observations will be referred to as "NEXRAD" in the rest of this document.

2.2 ECMWF forecasts

The model data consists of short-range precipitation forecasts produced by ECMWF's operational Integrated Forecasting System (IFS; e.g. Dee *et al.* 2011 and ECMWF 2013). Forecasts initiated at both 0000 UTC and 1200 UTC were used in order to identify the sensitivity of the statistics to the starting time of the forecast. Hourly forecast ranges between 1 and 24 hours were considered to study the diurnal cycle of precipitation since December 2011, while forecast ranges of 6, 12, 18 and 24 hours were retrieved to study 6-hourly accumulations over the 2002-2014 period. The horizontal resolution of the operational model was upgraded from 40 km to 25 km in February 2006 and from 25 km to 16 km in February 2010. All statistics presented here were computed at the appropriate model resolution. Besides, one should

note that the number of vertical levels used in the forecast computations was increased from 91 to 137 levels on 25 June 2013, although this should not directly affect the comparison performed here.

3 Climatological background

For reference, the seasonal mean precipitation distribution over the U.S.A. was computed over the period December 2001 to May 2014 from PRISM data (Di Luzio *et al.* 2008) and is depicted in Fig. 1. The original PRISM data consists of monthly precipitation 4-km gridded data generated from about 7,000 rain-gauges over mainland USA by the PRISM Climate Group (Oregon State University, http://www.prism.oregonstate.edu). Note that the PRISM data were averaged to the ECMWF model's most recent resolution (16 km) prior to plotting.

Figure 1 evidences the sharp contrast between the dry conditions which prevail over the Rocky Mountains throughout the year (seasonal total usually below 150 mm) and the wet regime that characterizes the Southern Great Plains and the East Coast in all seasons (seasonal total above 250 mm). From April to September, most of the precipitation over the two latter regions comes from intense convection. A north-south gradient is observed over the Rocky Mountains with semi-arid conditions in the south (often below 100 mm in each season) and locally much heavier precipitation associated with the highest orography in the north of the region. Over the Northern Great Plains, winter is usually on the dry side (below 150 mm) due to anticyclonic conditions, while spring and above all summer receive large amounts of rainfall (seasonal total between 150 and 400 mm), mainly from mesoscale convective systems. The extreme southeast of the U.S.A. (Florida) experiences relatively dry conditions in winter (around 150 mm) but becomes extremely wet during the summer (in excess of 500 mm), remaining rather wet during the autumn (around 300 mm), with the occasional passage of hurricanes. The West Coast north of 38°N is very wet in all seasons (totals over 400 mm), except in the summer (below 200 mm), while its southern part is usually much drier, except for relatively wet winters along the Californian coast.

4 Results

Monthly statistics of ECMWF precipitation forecasts against NEXRAD composites were computed in terms of mean differences, correlations and threat scores over the period January 2002 to June 2014. Statistics, which include land points only, will be shown for each sub-domain displayed in Fig. 2: West Coast, Rocky Mountains, Northern Great Plains (NGP), Southern Great Plains (SGP) and Eastern U.S.A. (East), as well as for the conterminous U.S.A. as a whole.

One should also stress the fact that the results presented here will focus on model forecasts initiated at 0000 UTC only, since statistics based on forecasts started at 1200 UTC led to very similar conclusions. The only difference between the two sets of forecasts will be addressed in section 4.5.

4.1 Time series

Figure 3 displays the time series of monthly normalized mean bias (*NMB*) and mean correlation between NEXRAD and ECMWF model for the entire U.S.A. mainland as well as for each sub-domain of Fig. 2 (see curve legend). *NMB* for a given month is defined as the mean NEXRAD−ECMWF bias divided by the mean of the two datasets (to ensure a symmetrical normalization). With this definition, *NMB*

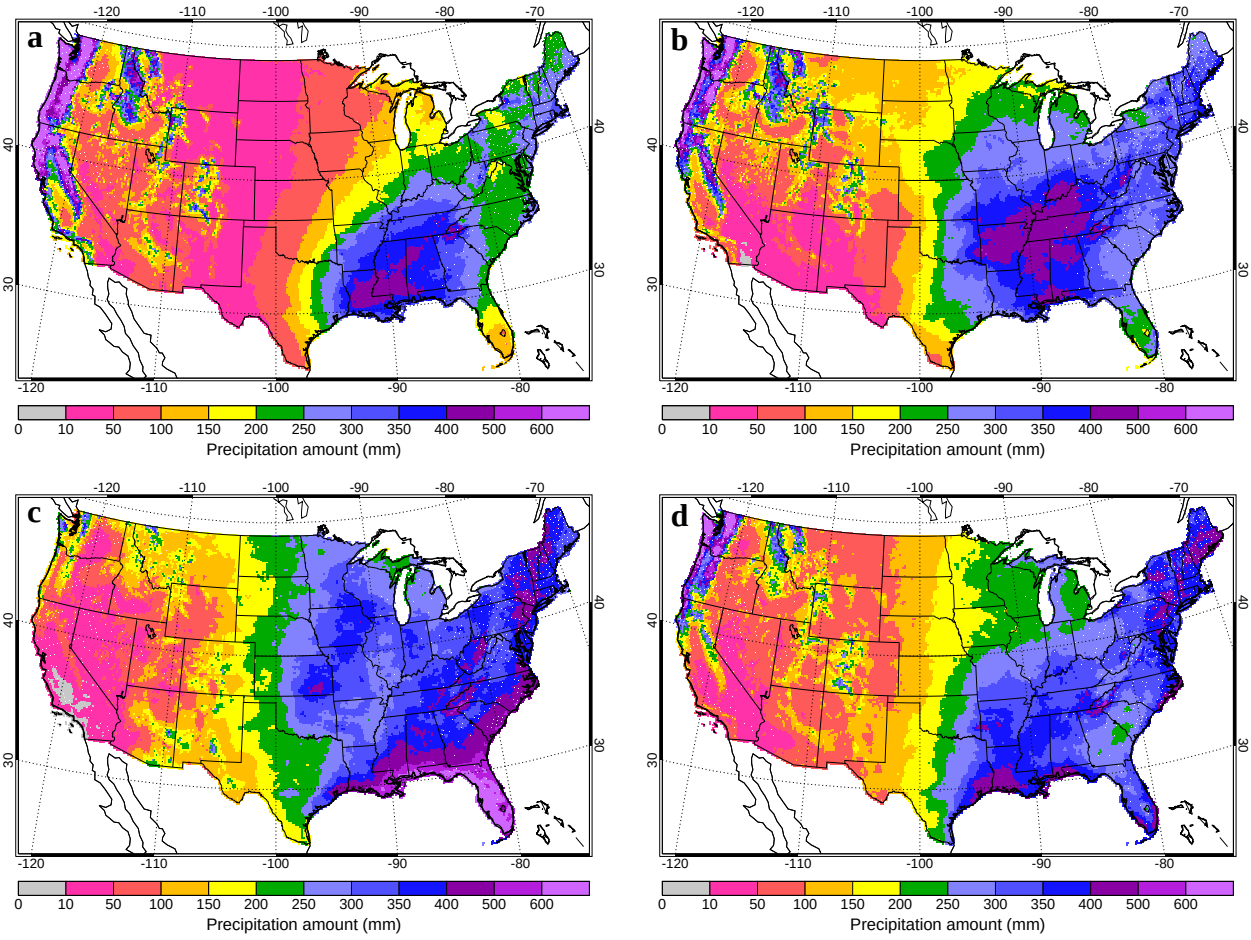


Figure 1: Seasonal mean precipitation amounts (in mm) over the U.S.A., as computed from PRISM gridded rain gauge data over the period December 2001-May 2014 for (a) winter, (b) spring, (c) summer and (d) autumn.

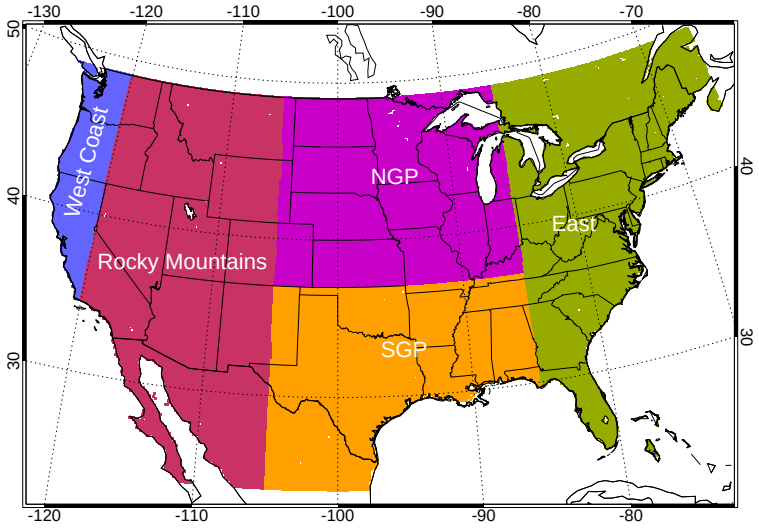


Figure 2: Geographical sub-domains for statistical computations: West Coast, Rocky Mountains, Northern Great Plains (NGP), Southern Great Plains (SGP) and Eastern U.S.A. (East).

varies between -2 and +2 and is unitless. A value of +1 (resp. -1) would be obtained when the mean precipitation in NEXRAD is three times as high (resp. as low) as in the model. The twelve-month running mean (thick line) is superimposed onto the monthly curve (thin line) to filter out seasonal fluctuations. Here, statistics are based on 6-hourly precipitation accumulations without any stratification according to the time of the day. Such stratification will be considered later in section 4.3.

4.1.1 Twelve-month running means

Focusing first on the 12-month running means (thick lines), Fig. 3 shows that for all sub-domains the mean *NMB* values (odd rows) have been steadily increasing towards zero throughout the whole period, which indicates that NEXRAD and ECMWF model precipitation have gradually converged over the years, on average. The predominance of negative values of *NMB* for all sub-domains suggests that the model systematically overestimates precipitation compared to NEXRAD. The strongest improvements in *NMB* are found over the Rocky Mountains (Fig. 3.i; from -0.8 to 0.0) and the NGP area (Fig. 3.c; from -0.5 to 0.05). The West Coast region (Fig. 3.k) exhibits the lowest rate of improvement in *NMB*. Panels on even rows in Fig. 3 exhibit a simultaneous regular improvement in the 12-month running mean correlation between NEXRAD and ECMWF forecasts, but at a somewhat more gentle pace than for *NMB*. It is very likely that all this amelioration can be explained by the successive beneficial upgrades made to the ECMWF forecasting system, in particular in physical parametrizations as well as in data assimilation. However, a contribution from a possible improvement in the quality of NEXRAD composites themselves should not be excluded, particularly in earlier years.

4.1.2 Monthly time series

Focusing now on the unsmoothed curves (thinner lines) on even rows in Fig. 3, one cannot fail to notice the well-defined seasonal cycle in the correlation, which oscillates between a minimum in mid-summer (between 0.20 and 0.35) and a maximum in mid-winter (slightly above 0.60) for all domains but the

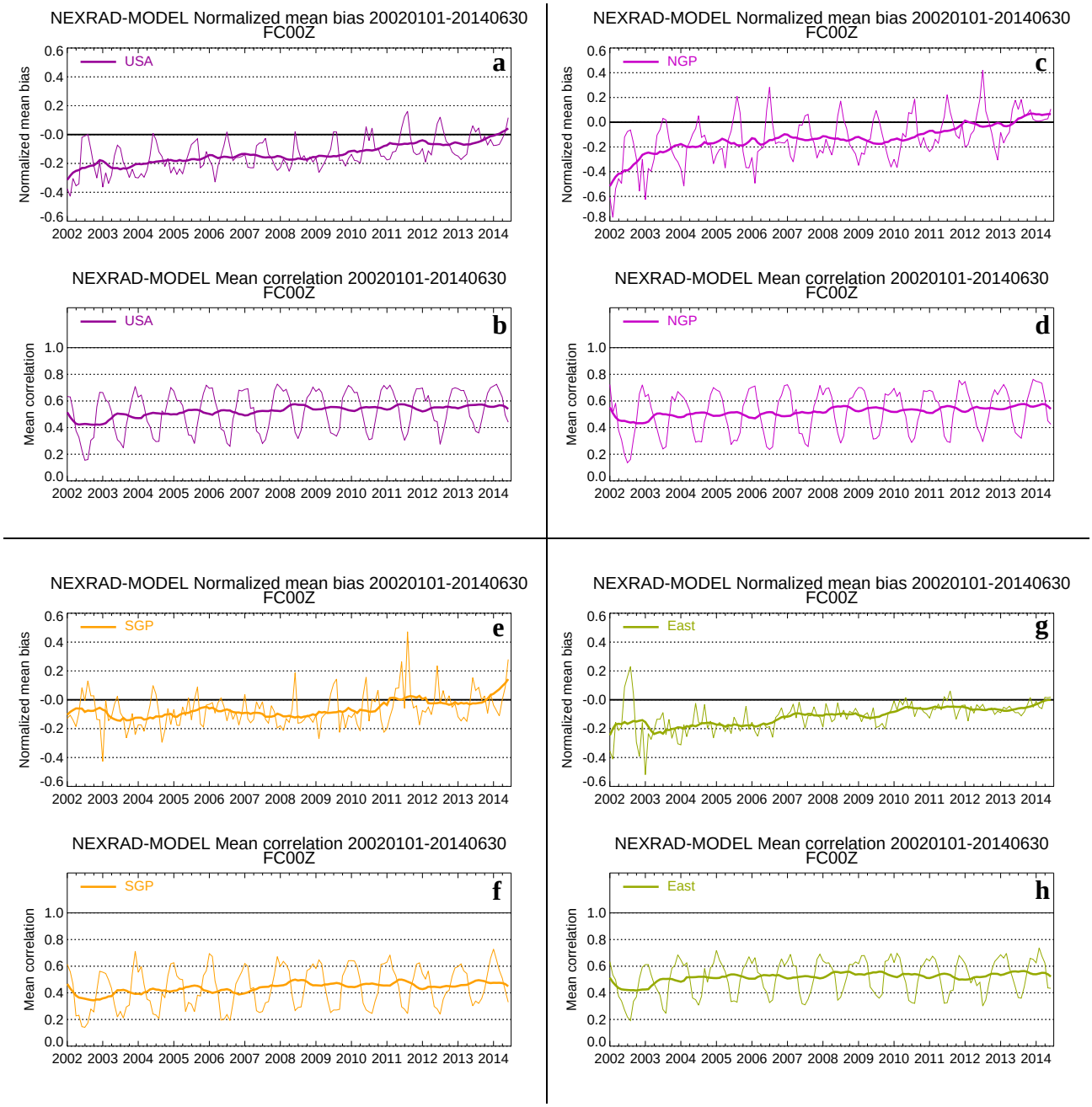


Figure 3: (Continued on next page.)

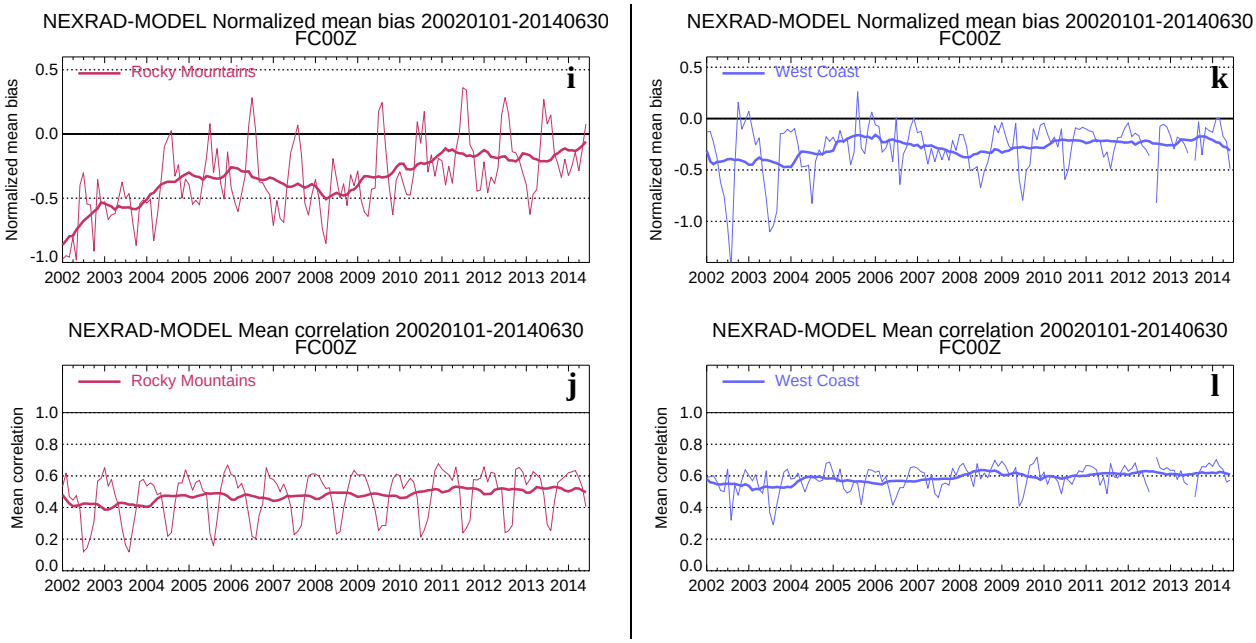


Figure 3: (Continued from previous page) Time series of monthly normalized mean bias (NMB; odd rows) and mean correlation (even rows) between NEXRAD composites and ECMWF 6-hourly forecasts (first forecast day) for the entire U.S.A. and for each of the sub-domains defined in Fig. 2 (see curve legend). The period for the statistics is January 2002 to June 2014. Statistics shown along the y-axis are all unitless. A positive value of NMB corresponds to an underestimation of precipitation in the model compared to NEXRAD. The twelve-month running mean (thick solid line) is superimposed onto the monthly curve (thin solid line).

West Coast. The difficulties of the model to simulate the precise location and intensity of convective precipitation events can explain the low correlations during the warm season. In contrast, the model is usually much better at forecasting large-scale stratiform precipitation systems, hence the higher correlation values in the wintertime. The predominance of non-convective precipitation throughout the year might also account for the steadily higher correlations (around 0.6) found over the West Coast (Fig. 3.l).

As far as NMB is concerned (odd rows in Fig. 3), there is less hint of a well-established seasonal cycle than for correlation, except for sub-domains NGP (panel c) and Rocky Mountains (panel i). Over these two regions, a clear sudden increase of NMB (towards an underestimation in the model after 2005) occurs in mid-summer and a decrease (strong overestimation in the model) in winter.

4.2 Threat scores

To complement previous statistics, Equitable Threat Score (ETS; see Appendix 2) values are briefly presented here. ETS provides a measure of how well ECMWF forecasts match NEXRAD observations for precipitation amounts above a specified threshold. ETS can theoretically range from $-\frac{1}{3}$ (very poor match) to 1.0 (perfect match), but a zero value would already suggest that the model has no skill. Figure 4.a displays the 12-year time series of ETS over the U.S.A. and for a minimum threshold of 3 mm day^{-1} (i.e. moderate precipitation). This figure clearly illustrates the regular improvement in model's skill from 0.3 to 0.4 in terms of the 12-month running mean (thick line). The monthly curve for ETS (thin line) exhibits the same seasonality as the correlation curve shown in Fig. 3.b. Similar conclusions could be drawn for other values of the minimum precipitation threshold ranging from 0 to 50 mm day^{-1} (not shown). For information, Fig. 4.b shows that the monthly sample size used in the ETS computations roughly increased from 600,000 to 4,500,000 points, as a result of the jumps in the operational model's horizontal resolution in early 2006 and 2010 (see section 2.2). It is worth noting that a study of other scores such as False Alarm Rate and Probability of Detection revealed the same kind of regular improvement as seen in ETS (not shown).

4.3 6-hourly precipitation statistics

It was also deemed interesting to compute mean monthly statistics stratified according to the four 6-hourly accumulation time slots of the first day of the forecast. This can be seen as a poor man's way of assessing the diurnal cycle of precipitation and is justified by the unavailability of hourly data from the model for most of the 12-year period. Note however that a better insight into the diurnal cycle using recent hourly model data will be given in section 4.4.

Figure 5 displays monthly NMB and correlation values averaged between January 2002 and June 2014 and for each sub-domain. The four curves in each panel correspond to accumulation periods 0-6, 6-12, 12-18 and 18-24 hours of the forecast (started at 0000 UTC), which in local time roughly corresponds to the first half of the night (blue), the second half of the night (green), the morning (red) and the afternoon (yellow), respectively.

First focusing on NMB on the odd-row panels of Fig. 5, one can see that overall the model slightly overestimates precipitation with respect to NEXRAD observations throughout the day during the cold season (from October to March). This excess in model precipitation is strongest over the Rocky Mountains ($NMB \approx 40\%$ on panel i) and weakest over the East, SGP and West Coast regions ($NMB \approx 10\%$; panels e, g and k, respectively). Conversely, during the warm season (April/May to September), the model tends to strongly underestimate rainfall amounts during the first half of the night (blue line) over all regions,

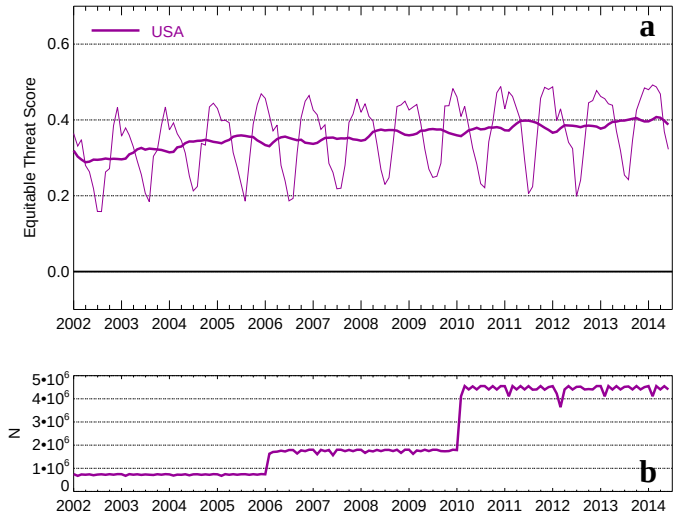


Figure 4: Monthly time series of (a) NEXRAD-ECMWF model Equitable Threat Score (ETS) and (b) corresponding sample population over the U.S.A. for the period January 2002-June 2014. ETS is shown for a minimum precipitation threshold set to 3 mm day^{-1}. The higher the ETS values are, the more skillful the model at predicting precipitation above the specified threshold. The twelve-month running mean (thick solid line) is superimposed onto the monthly curve (thin solid line).

except the West Coast (panel k). By mid-summer, this deficit peaks between 40% and 60% over NGP, SGP and the Rocky Mountains, but only 15% over the East region. On the other hand, morning precipitation (red line) is strongly overestimated by the model over most regions, with a peak excess of about -65% over the SGP and East regions and most of all around -100% over the Rocky Mountains in July and August. Much lower biases are found in the second half of the night (green line) and in the afternoon (yellow line). Throughout the year, the West Coast (panel k) exhibits the weakest diurnal cycle in *NMB*, with a persistent model overestimation peaking to around -70% by midday during the warm season.

Now considering correlations on even-row panels of Fig. 5, the pronounced seasonal cycle brings them from above 0.6 in the winter down to 0.2 in the summer for most sub-domains, except over the West Coast (panel l) where correlations remain between 0.5 and 0.6 all year round. These results are in agreement with the time series depicted on even-row panels of Fig. 3. A final noteworthy remark is that there appears to be very little dependency of mean correlations on the time of the day.

4.4 Diurnal cycle of precipitation

In addition to the previous statistics which were based on 6-hourly precipitation accumulations, original NEXRAD hourly accumulations were compared to equivalent quantities from ECMWF operational forecasts for the period December 2011 to June 2014 (prior to that hourly model outputs were not available). The main purpose was to more precisely assess the simulated diurnal cycle of precipitation against observations for each geographical sub-domains. The diurnal cycle was computed as a function of local solar time by taking into account the longitude of each grid point. Figure 6 displays the resulting curves obtained from NEXRAD (blue line) and ECMWF forecasts (red line) for each sub-domains and for the May-June period in 2012 (left), 2013 (middle) and 2014 (right). These three spring periods (with omnipresent convective activity) were selected to illustrate the benefits of the crucial change that was recently made to the diagnostic closure of the convective parametrization (Bechtold et al. 2014) and

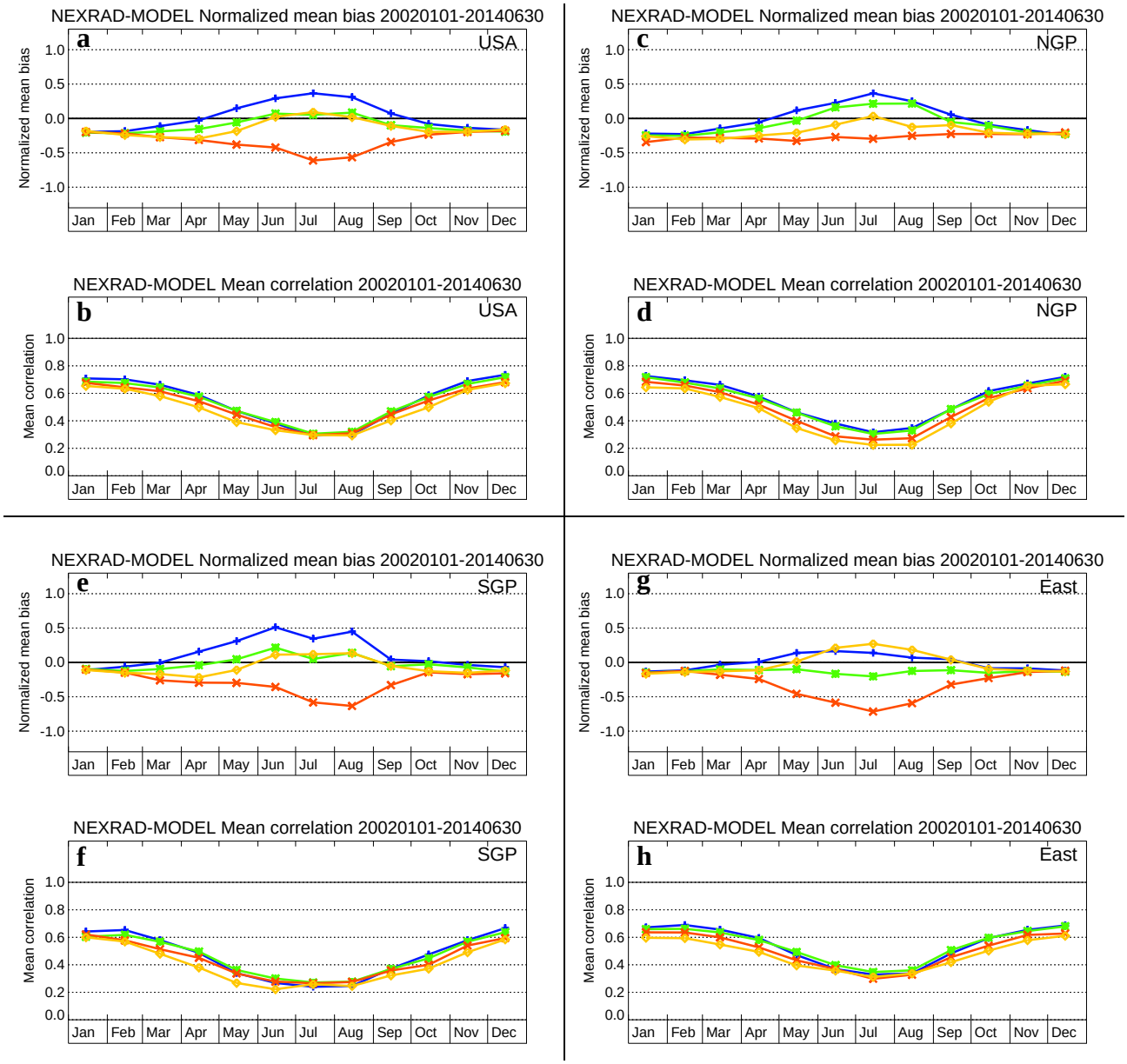


Figure 5: (Continued on next page.)

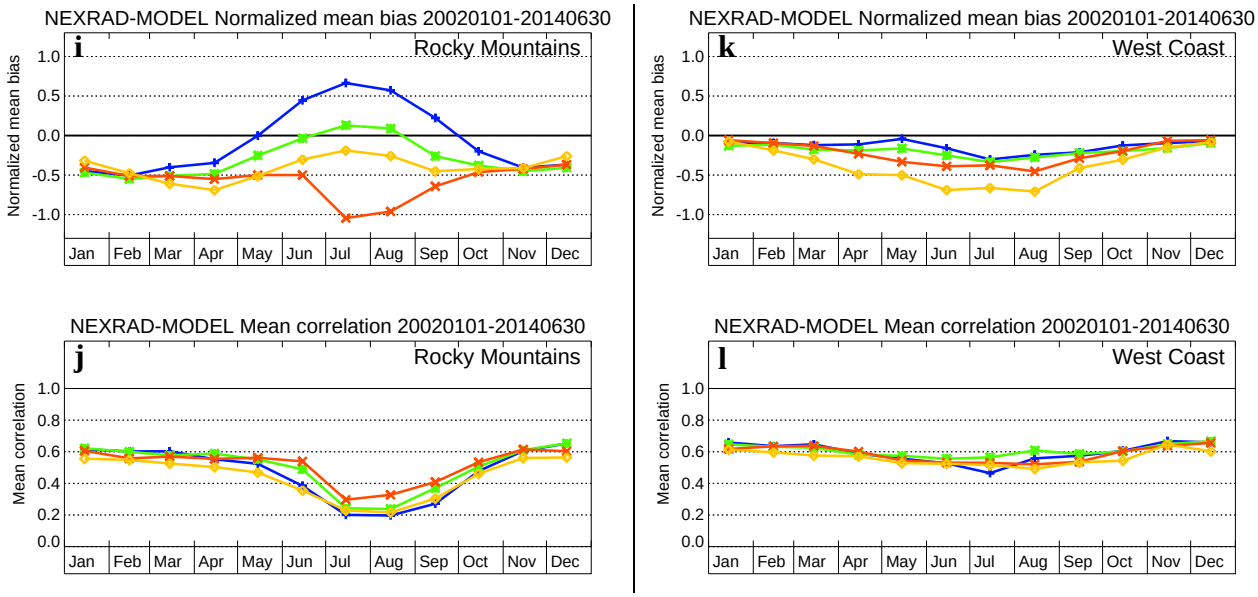


Figure 5: (Continued from previous page) Monthly normalized mean bias (NMB; odd rows) and mean correlation (even rows) between NEXRAD composites and ECMWF 6-hourly forecasts (first forecast day) for the entire U.S.A. and for each of the sub-domains defined in Fig. 2 (see label on each panel). Averages were calculated over the period January 2002 to June 2014 and for four 6-hourly accumulation periods in the day, which in local time correspond to the first half of the night (blue), the second half of the night (green), the morning (red) and the afternoon (yellow), respectively. Statistics shown along the y-axis are all unitless. A positive value of NMB corresponds to an underestimation of precipitation in the model compared to NEXRAD.

which became operational in November 2013 (ECMWF model cycle 40r1). It should be noted that the West Coast was discarded from the hourly statistics since some issues were identified in NEXRAD data over that region.

All left and middle panels in Fig. 6 show that the daily maximum of precipitation in 2012 and 2013 was simulated at around 14:00 local time by the model in all sub-domains, while this maximum was observed between 16:00 (East region; panel g) and 22:00 (NGP; panel c) according to NEXRAD data. This indicates that the model was converting convective available potential energy (CAPE) into precipitation far too early in the day. In 2014 (right panels), the revision of the convection scheme has led to a rather dramatic reduction of the phase shift between the model and the observations (down to a couple of hours). In particular, the rising portion of the red curve has been shifted to the right in 2014 compared to 2012 and 2013. This emphasizes that the model now simulates the intensification of precipitation later in the morning, which better agrees with the observations. Still, the model precipitation tends to wane too quickly during the afternoon.

4.5 Sensitivity of simulated precipitation amounts to forecast starting time

As mentioned earlier, statistical results turned out to be rather insensitive to forecast starting time (not shown). However, it was noticed that forecasts initiated at 1200 UTC (FC12) had a tendency to emphasize the amplitude of the summertime precipitation diurnal cycle, compared to forecasts starting at 0000 UTC (FC00). This is illustrated in Table 1 which compares 6-hourly summer precipitation accumulations from FC00 and FC12 averaged over the period 2002-2013 and over the United States. Mean NEXRAD precipitation amounts are also provided for reference and forecast lengths are given in parentheses.

Local times (approx.)	NEXRAD	FC00	FC12	FC12−FC00
00-06h	1.882	**1.756** (+12)	**1.733** (+24)	**−0.023**
06-12h	1.609	**2.765** (+18)	**3.213** (+6)	**+0.448**
12-18h	3.172	**3.030** (+24)	**3.187** (+12)	**+0.157**
18-00h	2.668	**1.928** (+6)	**1.588** (+18)	**−0.340**

Table 1: Comparison of 6-hourly summer precipitation accumulations averaged over the period 2002-2013 and over the U.S.A., and obtained from forecasts started at either 0000 (FC00) or 1200 UTC (FC12) and run for up to 24 hours. Forecast ranges (in hours) are indicated in parentheses. Mean NEXRAD precipitation is also reported, for information. Six-hourly precipitation amounts are in mm day^{-1}.

Table 1 confirms that the amplitude of the diurnal cycle of precipitation is larger in FC12 than in FC00, with less precipitation during the night (18-06h local time) and more during the day (06-18h local time). It is important to emphasize that other seasons than the summer do not exhibit this difference (not shown).

A more detailed examination of the summer statistics for each sub-domain revealed that the SGP and East regions were the two main contributors to the signal found in Table 1. To further analyze the larger 6-hourly precipitation amounts produced in the morning by FC12+6 (second line of Table 1), Fig. 7 gives an illustration of the mean differences in 6-hour rainfall accumulations between FC12+6 and FC00+18 during the summer 2012.

The largest excess in precipitation (above 2 and up to 7 mm day^{-1}) in FC12+6 compared to FC00+18 is attained along the Gulf of Mexico and extends along the Mississippi Valley. To understand the origin

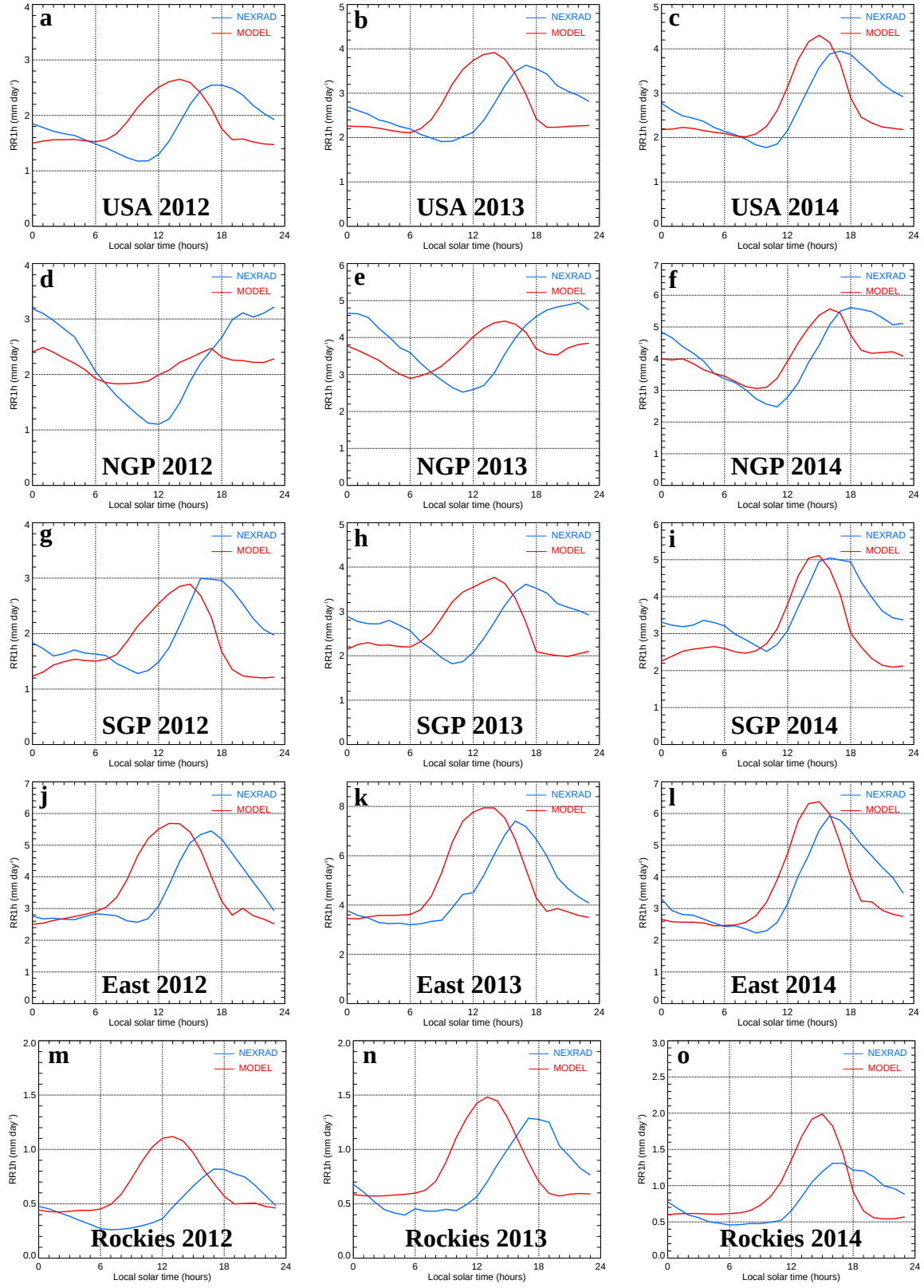


Figure 6: Diurnal cycle of mean hourly precipitation as a function of local solar time (hours) from NEXRAD composites (blue line) and ECMWF forecasts (red line; first day of forecast) for the entire U.S.A. and for sub-domains defined in Fig. 2 (see legend). The periods for the statistics are May-June 2012 (left), 2013 (middle) and 2014 (right). Mean precipitation along the y-axis is in mm day^{-1}.

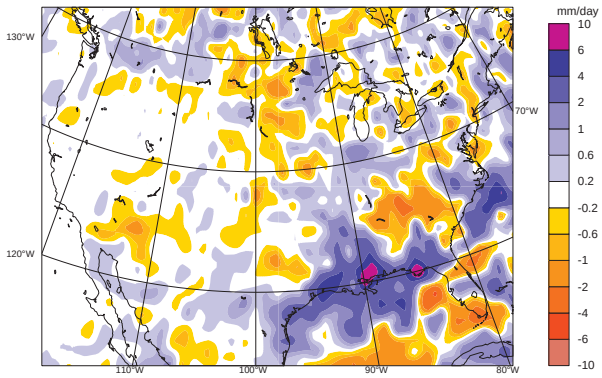


Figure 7: Mean differences in 6-hour precipitation accumulations between FC12+6 and FC00+18 in June-August 2012. Positive (resp. negative) differences (in mm day^{-1}) are displayed with bluish (resp. reddish) shading.

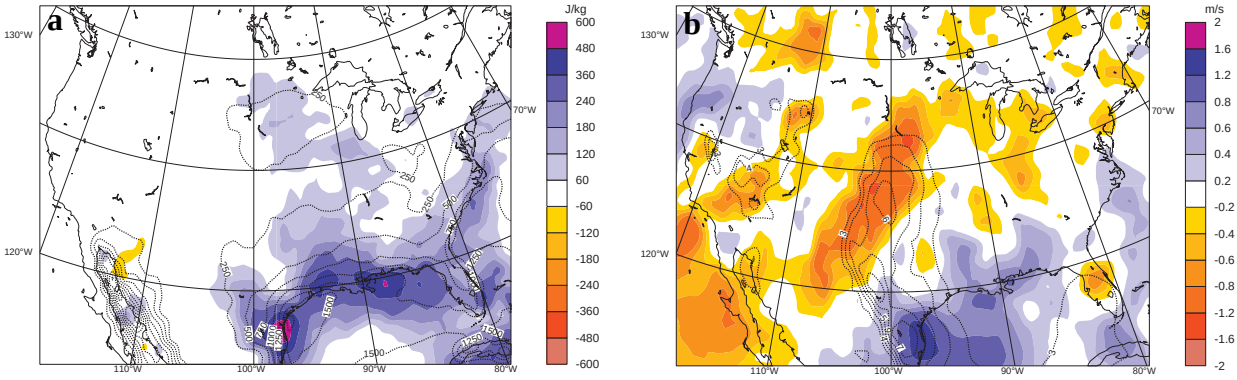


Figure 8: Mean differences in (a) CAPE and (b) meridional wind at model level 80 (i.e. at around 900 m above ground) between 1200 UTC analyses and corresponding 12-hour forecasts started at 0000 UTC in June-August 2012. CAPE is in J kg^{-1} and meridional wind is in m s^{-1}. Positive (resp. negative) differences are displayed with bluish (resp. reddish) shading. Mean CAPE values from the 0000 UTC forecasts are shown with dotted isolines (250 J kg^{-1}interval) in panel a. In panel b, mean meridional wind isotachs are plotted above +3 m s^{-1} only, to identify the low-level southerly jet just east of the Rocky Mountains.

of this excess, a comparison of mean meteorological conditions between FC00+12 and corresponding analyses at 1200 UTC (AN12) was carried out over the summer 2012. Figure 8.a indicates that the mean CAPE along the Gulf of Mexico was much higher (by up to 35% near the coast) in AN12 than in the corresponding FC00+12. At the same time, Fig. 8.b shows that in AN12 the mean southerly low-level jet (dotted isolines) around 100°W was stronger over the Gulf of Mexico but slower in its northern part. The combination of the resulting enhanced low-level convergence with the higher CAPE can explain why FC12+6 gave more rainfall than FC00+18 over the southeast of the United States. It should be stressed that the patterns seen in Fig. 8 turned out to be robust features of all summers throughout the studied period (not shown).

Furthermore, from the second and fourth lines of forecasts statistics in Table 1), one can deduce that for a given verification time, FC+6 produces more precipitation on average than FC+18, regardless of the forecast starting time. In contrast, the first and third lines of Table 1) show that FC+12 and FC+24 (valid at the same time) yield comparable mean amounts of precipitation. This suggests that some spin-down affects precipitation forecasts within their first 12 hours or so. However, this spin-down does not change the sign of the largest biases with respect to NEXRAD, as can be seen by comparing columns ”FC00” and ”FC12” with column ”NEXRAD” in Table 1). Finally, it is worth stressing that this spin-down happens during the summer season only.

5 Conclusions

A systematic comparison of ECMWF’s operational short-range forecasts with NCEP Stage IV (NEXRAD) precipitation composites was carried out over the period January 2002 to June 2014. Statistical results show that the match between the numerical forecasts and NEXRAD observations has kept improving over the years, as a result of the regular upgrades made to the various components of the IFS (e.g. physical parametrizations, dynamics representation and data assimilation). This improvement is particular obvious in terms of long-term mean biases and to a lesser extent on mean correlations as well. A clear seasonal cycle exists in the correlations between model and NEXRAD over most of the U.S.A., with higher values during winter and lower values in the summer, as expected given the universal challenge of forecasting convective activity at the right place and time. The only exception is the West Coast where correlations remain high throughout the year due to the prevalence of stratiform precipitation. A regular improvement of threat scores also occurred over the last decade.

It is also found that the model tends to systematically overestimate precipitation throughout the first day of forecast during the cold from October to March, with no marked diurnal cycle. From April to September, a strong underestimation occurs in the first half of the night while a strong overestimation takes place during the morning, both peaking in mid-summer. Systematic biases are less pronounced during the rest of the day. The West Coast is the only region not to be affected by the former large summer biases.

In other respects, forecasts started at 0000 UTC and 1200 UTC lead to monthly mean biases that have the same sign with respect to NEXRAD. However, some precipitation spin-down can be detected within the first 12 hours of the forecast, but only over the southern states and only during the summer. This spin-down seems to originate from differences in convective available potential energy and atmospheric circulation (low-level jet) over the latter region. Other seasons do not exhibit any significant precipitation spin-up or spin-down.

Using hourly forecast data from the past few years, it is clearly demonstrated that the long-standing phase advance in the diurnal cycle of convective precipitation was substantially reduced after the recent

changes made to the closure of the convection scheme (Bechtold *et al.* 2014).

It would certainly be beneficial to continue this evaluation of ECMWF forecasts against NEXRAD composites in the future as it provides a source of verification that nicely complements the more traditional verification against radiosondes, surface SYNOP observations and model analyses. Furthermore, it might also be relevant to extend this validation exercise beyond the first day of the forecast.

Acknowledgements

I am very grateful to NCEP for producing and NCAR/UCAR/EOL for providing the Stage IV precipitation composites used in this study. The PRISM Climate Group (Oregon State University, U.S.A.) is acknowledged for granting access to their monthly precipitation gridded data. I would also like to thank Anton Beljaars and Peter Bechtold (ECMWF) for their comments on the manuscript.

APPENDIX 1

List of acronyms used in the text (alphabetical order)

ECMWF	=	European Centre for Medium-range Weather Forecasts.
EOL	=	Earth Observing Laboratory (U.S.A.).
EUMETNET	=	EUropean METeorological services NETwork.
NCAR	=	National Center for Atmospheric Research (U.S.A.).
NCEP	=	National Centers for Environmental Prediction (U.S.A.).
NEXRAD	=	NEXt-generation RADars (U.S.A.).
OPERA	=	Operational Program for the Exchange of weather RAdar information (Europe).
PRISM	=	Parameter-elevation Regressions on Independent Slopes Model (U.S.A.)
UCAR	=	University Corporation for Atmospheric Research (U.S.A.).

APPENDIX 2

The Equitable Threat Score (ETS) is defined as

$$ETS = \frac{H - H_e}{H + M + F - H_e} \tag{1}$$

where H is the number of correct hits, M is the number of misses and F is the number of false alarms. H_e is the number of correct hits obtained by pure chance and is computed as

$$H_e = \frac{(H+F)(H+M)}{N} \tag{2}$$

where N is the sample size.

References

Bai, L. (2013). Regional Report on the current status of the exchange of weather radar data - RA II. WMO Commission for Basic Systems/Open Program Area Group on Integrated Observing Systems, Workshop on Radar Data Exchange, 24-26 April 2013, Exeter, United Kingdom, 13 pages, available at http://www.wmo.int/pages/prog/www/OSY/Meetings/ET-SBO_Workshop_Radar_Data_Ex/Doc_Plan.html.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J.-P., Beljaars, A., and Bormann, N. (2014). Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, 71:734–753.

Dee, D. P. and co authors (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, 137:553–597.

Di Luzio, M., Johnson, G. L., Daly, C., Eischeid, J. K., and Arnold, J. G. (2008). Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. *J. Appl. Meteor.*, 47:475–497.

ECMWF (2013). *Integrated forecast system documentation*. ECMWF, Shinfield Park, Reading, UK. Available online at http://www.ecmwf.int/research/ifsdocs/.

Fulton, R. A., Breidenbach, J. P., Seo, D. J., Miller, D. A., and O'Bannon, T. (1998). The WSR-88D rainfall algorithm. *Weather Forecast.*, 13:377–395.

Huuskonen, A., Saltikoff, E., and Holleman, I. (2014). The Operational Weather Radar Network in Europe. *Bull. Am. Meteorol. Soc.*, 95. doi:10.1175/BAMS-D-12-00216.1.

Lin, Y. and Mitchell, K. E. (2005). The NCEP Stage II/IV Hourly Precipitation Analyses: Development and Applications. In *Proceedings of the 19th AMS Conference on Hydrology, San Diego, CA (USA), 5–14 January 2005*.

Lopez, P. (2011). Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF. *Mon. Weather Rev.*, 139:2098–2116.