Use of re-analyses and re-forecasts at NCEP

Huug van den Dool and Suranjana Saha

NCEP, Washington DC

1. Introduction

The NCEP Climate Forecast System Reanalysis (CFSR) was completed for the 31-year period from 1979 to 2009, in January 2010 (Saha et al 2010). Earlier NCEP Reanalyses (R1; Kalnay et al 1996; Kistler et al 2001) and R2 (Kanamitsu et al 2002), along with their respective real time climate data assimilation systems (CDAS) updates, have been widely used. However, both R1 and R2 are becoming technologically old. The CFSR was primarily motivated by the need to provide initial conditions for retrospective seasonal predictions (also called Reforecasts or hindcasts) that are required for the operational implementation of an upgraded CFS seasonal prediction model (version 2), at NCEP in January 2011. The current operational CFS (version 1) was implemented in 2004 and consists of diverse components, such as R2 for the atmospheric initial conditions, a global atmospheric model (GFS) of vintage 2003, a near global MOM3 ocean model, near global ocean initial condition produced by a stand alone GODAS, and a set of hindcasts over the period 1981-2003 (Saha et al 2006). The new CFSv2 has more consistency between the historical initial conditions and the reforecasts, as well as with the real time coupled data assimilation and prediction system.

As described in some detail in Saha et al(2010), atmosphere-ocean-land-seaice initial conditions were generated for CFSv2, by creating a high resolution coupled reanalysis from 1979-present. The new Climate Forecast System Reanalysis (CFSR) consists of an atmospheric model (GFS, 2010) at T382L64 (~38 Km) , a MOM4 ocean model at 0.25 degree meridional resolution near the equator and 0.5 degree global, a 3-layer sea-ice model and the 4-level Noah land model. The 30-minute coupling to the ocean (during the generation of 9-hr guess fields) is perhaps the only true novelty, compared to the many Reanalyses made internationally in Europe, the US and Japan. There are, however, noteworthy features and/or improvements over the earlier reanalysis efforts at NCEP, namely:

a) assimilation of satellite radiances over the entire period, b) coupling to a sea-ice model, c) increasing GHG gas concentrations, d) high resolution, e) cleaned up observational input, which is an ongoing international effort, f) real time monitoring of the output by a team of scientists from the Climate Prediction Center at NCEP, and g) much improved data assimilation and forecast models. The Reanalysis (CFSR) is run at T382L64 even though the hindcasts (1982-2009) are made at T126L64. The hindcasts have been completed for the period 1982-2009, and both the CFS Reanalysis and Reforecasts (CFSRR) are ongoing for 2010. A paper on the CFS version 2 is forthcoming (Saha et al, 2011). The reader can view a comprehensive atlas of maps at http://cfs.ncep.noaa.gov/cfsr , or send questions to cfs@noaa.gov or download the data from sites: at
NCDC (http://nomads.ncdc.noaa.gov/data.php?name=access#cfsr) or
NCAR http://dss.ucar.edu/pub/cfsr.html .

2. The SST – Precipitation correlation

One of the reasons to conduct a coupled Reanalysis was to address criticism that R1 and R2 have an erroneous relationship between SST (which are prescribed in R1 and R2 and all other Reanalyses) and model generated precipitation (P). The error was especially clear in the tropical west Pacific. As shown in Fig.1 R1 and R2 have a simultaneous positive correlation, suggesting that an SST anomaly in the westPac would set off a P response of the same sign, this is of course also a problem in AMIP runs. Observations suggest otherwise (the black curve), namely SST leads P, and the simultaneous correlation is near zero. It appears that CFSR is reproducing the observed relationship quite well. The improved coupled initial states may have contributed to the vastly improved forecast skill of certain tropical fields, see section 6.

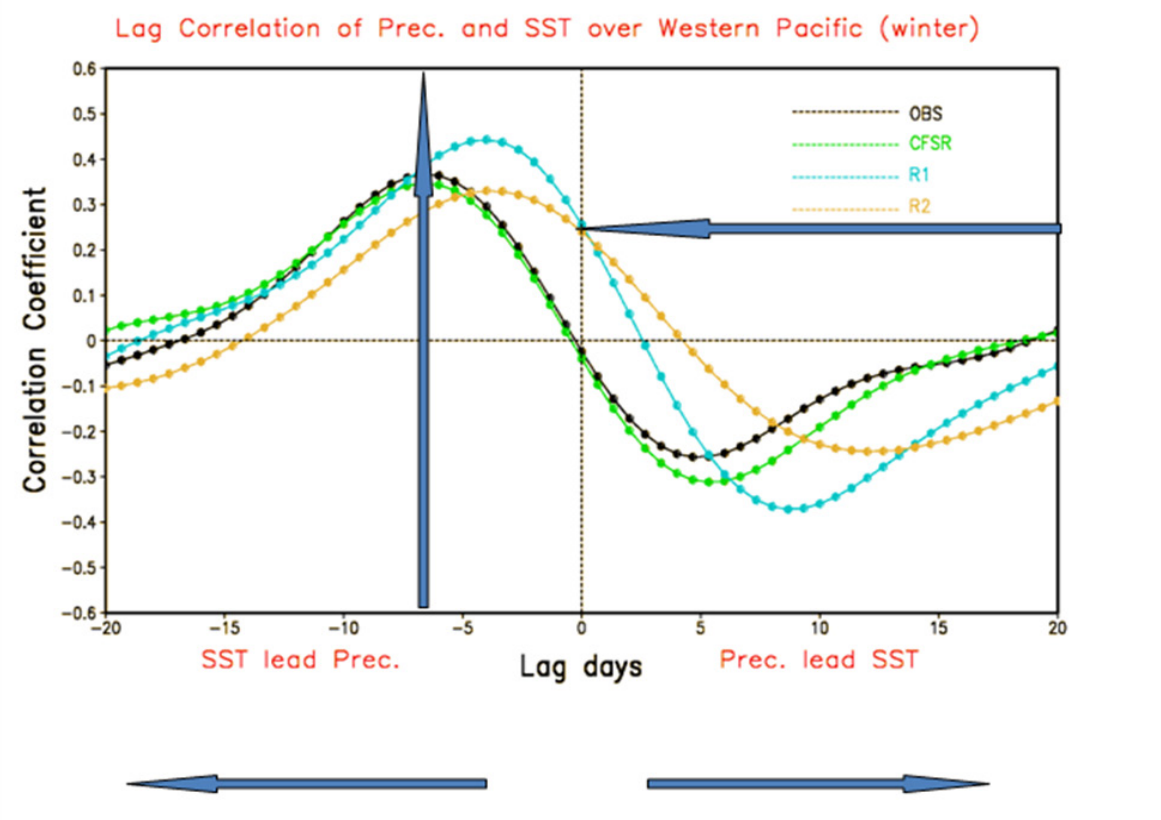

Figure 1: Temporal lag correlation coefficient between precipitation and SST in the tropical Western Pacific (averaged over 10S-10N, 130E-150E) in R1 (red), R2 (brown), CFSR (green) and observation (black). GPCP daily precipitation and Reynolds ¼º daily SST are used as observational data. Negative (positive) lag in days on the x-axis indicates the SST leads (lags) the precipitation. Data for the boreal winter (Nov-Apr) over the period 1979-2008 are band-pass filtered for 20-100 days after removing the climatological mean.

3. Accuracy of the initial states

Fig. 2 shows a time series of the spatially and annually aggregated RMS difference between the 6 hour forecast (guess field) and the verifying surface pressure observations (wherever they are). In order to drive the point home we aggregated the RMS for SH-ocean (previously the most difficult) and NH-land (where we used to have, relatively speaking, a lot of data). Over the SH oceans, the RMS has been coming down steadily from about 2 hPa in 1979 to around 1.1 hPa in 2009, i.e. the analyses are getting better and better all the time. In contrast, the RMS over land in the NH starts already at about 1.1 hPa in 1979, stayed virtually unchanged through 1996, then dropped to 0.90-0.95 hPa in May 1997 when the Metar data became abundant. Curiously, no such single outstanding reason for improvement exists in the SH, where many new data sources helped reduce RMS on, what appears to be, a continuous basis. As had already been noted in real time since 2000 by tracking day 5 forecast verification scores, the SH has practically caught up with the NH in being observed and predicted nearly equally accurately. A final point relates to climate monitoring: the CFSR is obviously not homogeneous over the SH oceans.

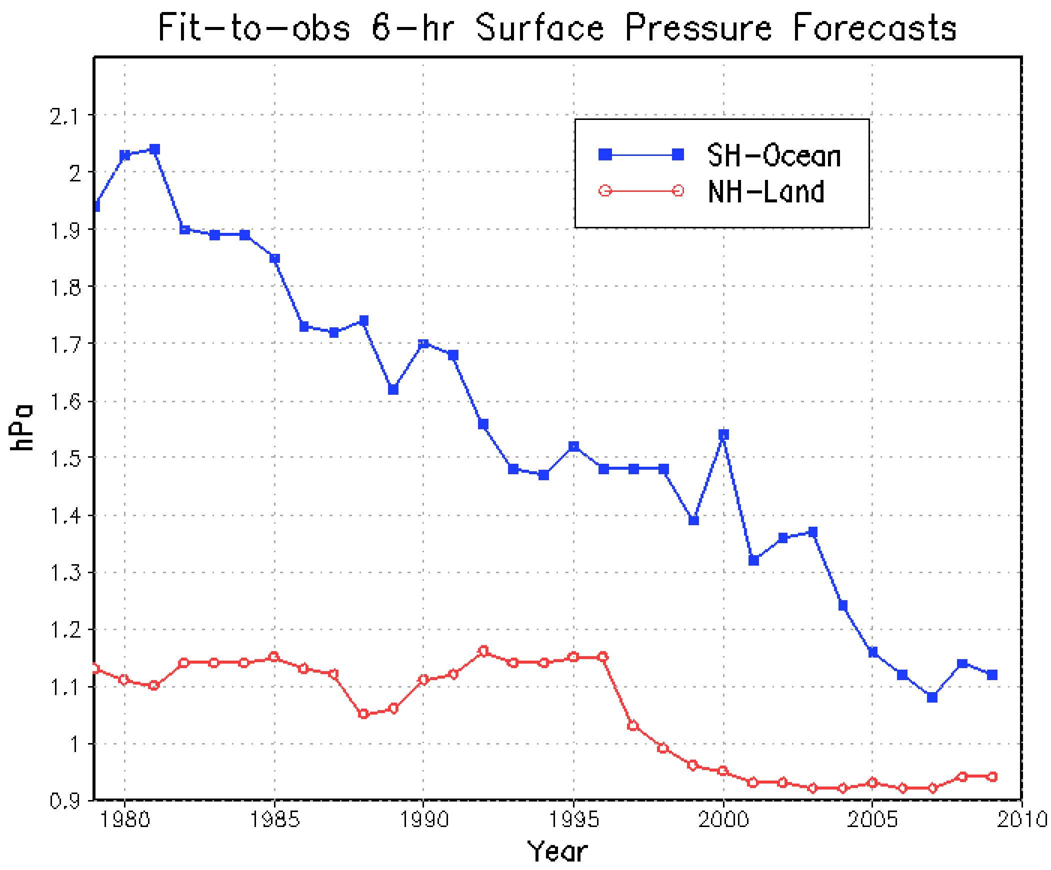

Figure 2: The fit of 6 hour forecasts of instantaneous surface pressure against irregularly distributed observations. Shown are annually compiled fit-to-obs data 1979-2009 in units of hPa. In blue SH ocean, in red NH land.

4. Climate Monitoring

Beyond the use of CFSR for initial conditions of forecasts is its use for climate monitoring. NCEP itself has a strong program at CPC for real time short-term climate monitoring (ENSO and such), but R1 has also been scrutinized (and criticized) for global change information. Fig.3 shows a time series of the global mean 2 meter surface air temperature (T2m) over land produced by three data sets, R1, CFSR and GHCN-CAMS (Fan and Van den Dool 2008). The latter is based on T2m observations only; GHCN is a well known source, and the CAMS addition refers to all the GTS data received at NCEP by which we augmented GHCN enormously, especially for recent years. Both R1 and CFSR do NOT assimilate T2m at all, so it is interesting to see how similar these three curves are; obviously T2m is not a stand-alone variable, but it relates to the winds and pressure that ARE assimilated in R1 and CFSR. No offset was applied to these three lines, so we do find R1 (GHCN-CAMS) to be the coldest (warmest) of the three, and CFSR in between. This may be because R1 was running on a CO2 value that was not only kept constant from 1948-present but was quite old (~330ppm) to begin with for the post 1979 period. In comparison to R1 (trend +0.66K per 31 years) , CFSR has a more realistic upward trend, rather close to GHCN-CAMS's 1 K per 31 years, so one might surmise that the upward trend can be recovered for 2/3^{rds} by data assimilation alone (but not assimilating T2m directly) but to get the other 1/3rd one needs increasing GHG concentration. See also Simmons et al (2010) for ECMWF's experience in this regard. Over the same years, 1979-present, the global mean SST (the other 70% of the planet) went up by only 0.3K and here R1 and CFSR agree closely. The ocean component of CFSR is relaxed to daily 1/4th degree SST analysis, so its trend is pretty much a given and not left to modeling.

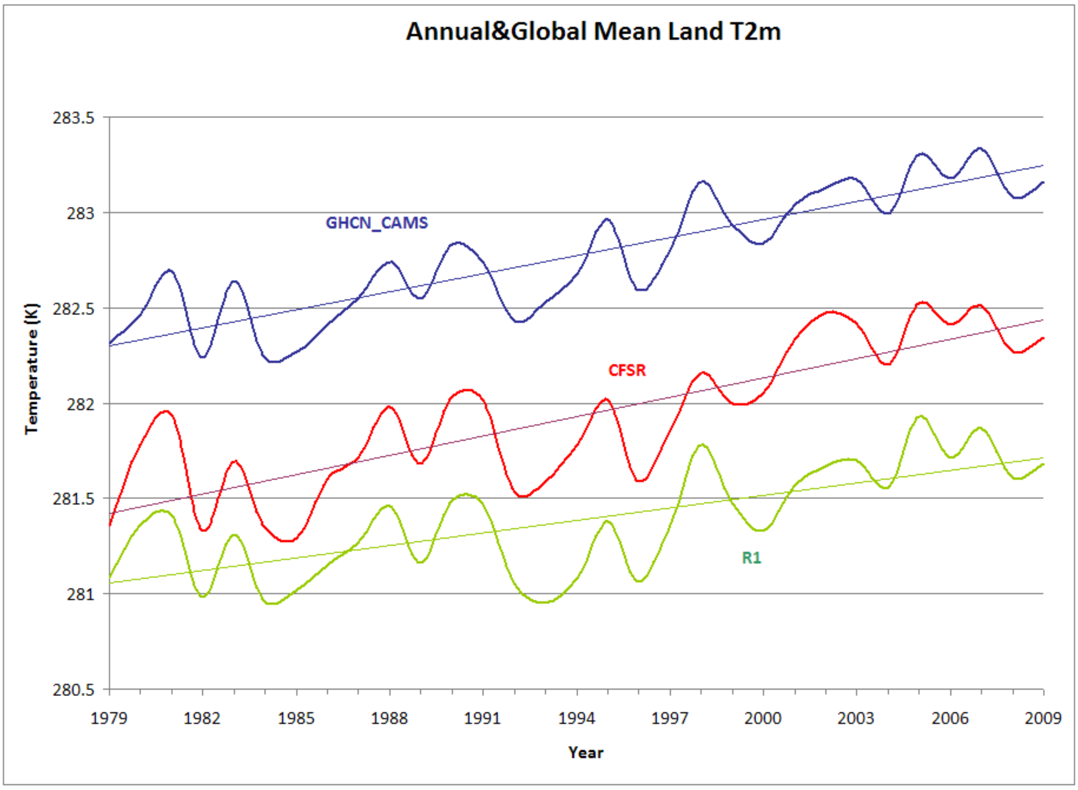

Figure 3: The annual global mean 2-meter temperature over land in R1 (green), CFSR (red) and GHCN_CAMS (blue) over the period 1979-2009. Unit is Kelvin. Thin lines are least square linear fits of the three time series against time. The linear trends are 0.66, 1.02 and 0.94K per 31 years for R1, CFSR and GHCN_CAMS respectively. (Keep in mind that straight lines may not be perfectly portraying climate change trends).

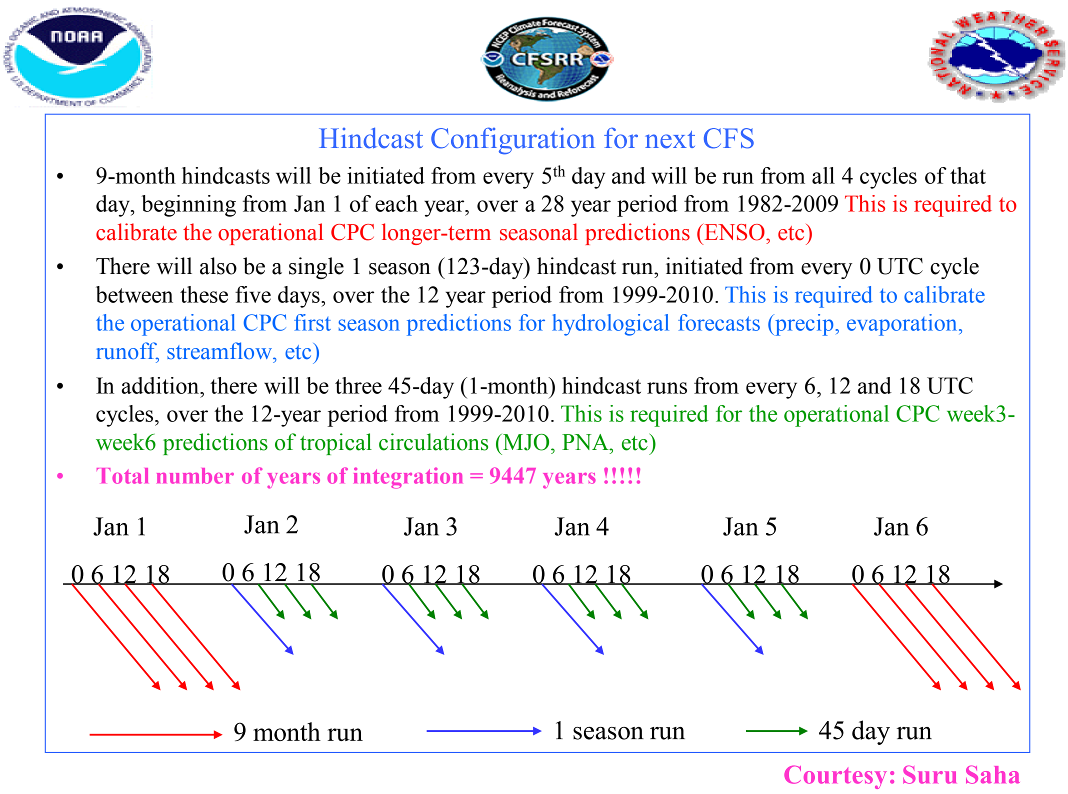

Figure 4: The lay-out of the hindcasts. Nine-month runs have been made every 5 days (all 4 cycles), 90 days runs every day from 0Z, and 45 day runs every day all four cycles.

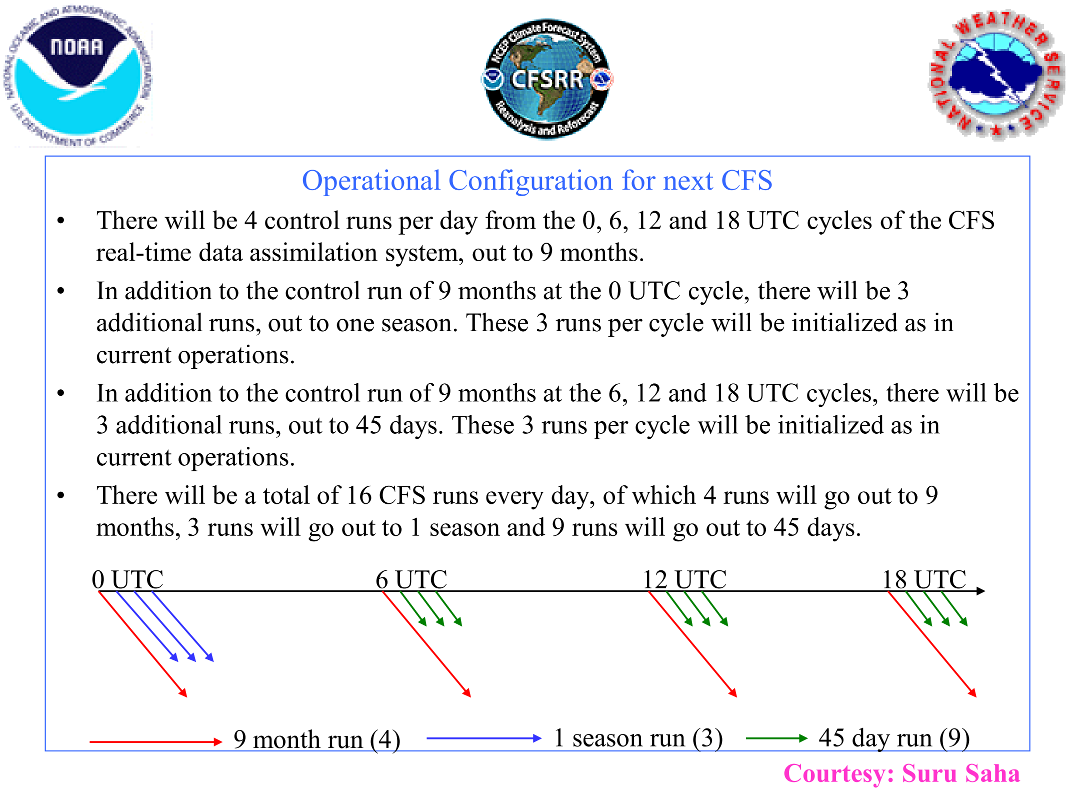

Figure 5: The lay-out of the CFSv2 forecasts in real time. Nine-month runs will be made every day (all 4 cycles), 90 days runs every day from all four cycles, and (a total of sixteen) 45 day runs every day four from all four cycles.

5. The lay-out of the hindcasts and real time forecasts

As an information item for those who want to use the hindcast data we add here figures 4 and 5 that detail the lay-out of the hindcasts over 1982-2010, and the lay-out of the daily real time forecasts as we go forward. We believe the figures + text speak for themselves. An important point is that the real time forecasts will be run on a schedule without delay. Previously, the CFS and forerunners (SFM, b9x) were run days, if not weeks, after real time because a) this was thought to be OK for seasonal prediction, and b) some input data sets were late in getting assembled. All of this has changed, and the CFSv2 forecasts target all prediction ranges, from week 1, week 2-6, MJO and seasonal to interannual. Not only has skill improved, but the models are run early enough to be relevant.

6. An example of improvement in forecast skill

Although the CFSv2 has been developed nominally as a seasonal prediction system, there will be applications to much shorter forecasts leads also. Given that the forecasts will be run in real time without delay, CFSv2 forecasts will have utility for the next few days, week1 – week 6, MJO, then seasonal out to almost a year. Below we highlight skill in hindcasts in the range of days 1-30.

Fig. 6 shows the skill (as per anomaly correlation (AC)) of CFS in predicting the MJO, as expressed by the Wheeler and Hendon (WH) index (two EOFs of combined zonal wind&OLR). The period is 1982-2008. On the left is CFSv2, on the right CFSv1. On top a period in Nov/Dec, on the bottom a period in Feb/Mar. For each initial day, say March, 1 we have 28 forecasts, one in each year. The thin blue lines are the anomaly correlation for each year (based on about 35 forecasts per year), the red line is the aggregate for all years. It is quite clear that CFSv2 has much better skill than CFSv1. In fact this

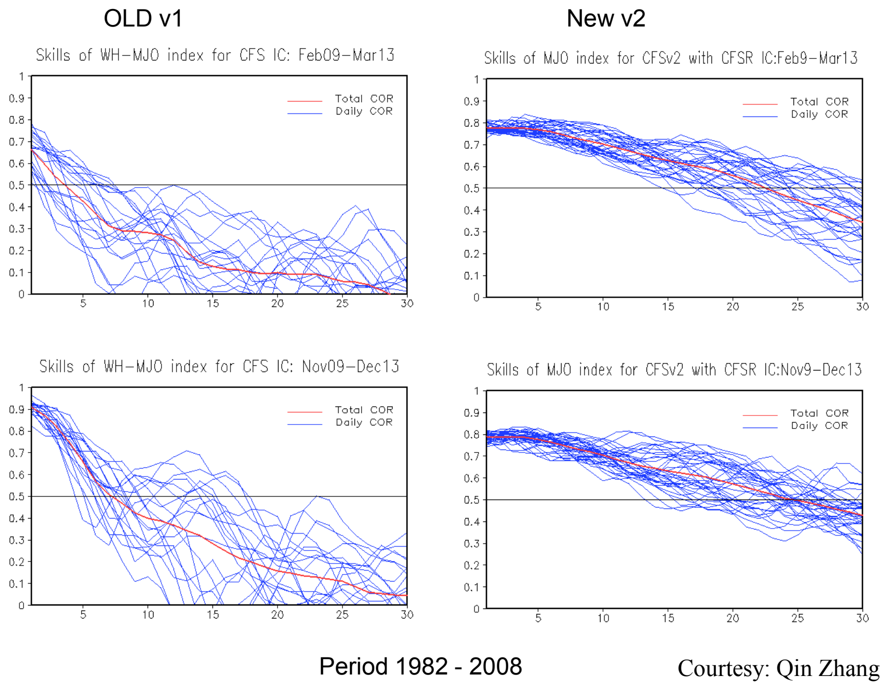

Figure 6: The anomaly correlation of the MJO (as expressed by Wheeler and Hendon Index) predictions out to 30 days by CFSv1 (left) and CFSv2 (right). The period is 1982-2008. Along the upper row initial conditions during Feb, 9 – Mar 13, and the bottom row for Nov, 9 – Dec, 13. Thin blue lines for each year, the red line an aggregate across all years. Courtesy Qin Zhang.

is the improvement of half a generation (15 years of work), taking into account that CFSv1 has rather old R2 initial conditions. The AC stays above 0.5 for nearly three weeks in the new system, while this was only one week previously. (That the AC starts at 0.8, even in the new system is caused by the fact the model OLR (in first 6 hours) and observed OLR are quite different. This can be massaged away.)

One rarely sees such a demonstration of improvement. The causes are probably very many, but especially the improved initial states in the tropical atmosphere and the consistency of model and initial state suggest themselves. Further research should bring out the importance of coupling to the ocean (if any) and its quantitative contribution to skill.

References

Fan Y. and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present, *J. Geophys. Res.*, **113**, D01103, doi:10.1029/2007JD008470.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83,** 1631–1643.

Kistler, R., E. Kalnay, W. Collins, S. Saha, G. White, J. Woollen, M. Chelliah, W. Ebisuzaki, M. Kanamitsu, V. Kousky, H. van den Dool, R. Jenne, and M. Fiorino, 2001: The NCEP–NCAR 50–Year Reanalysis: Monthly Means CD–ROM and Documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H. M. van den Dool, H.L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, and P. Xie, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.

S. Saha, S. Moorthi, H-L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y-T Hou, H. Chuang, H-M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. van den Dool, Arun Kumar, W. Wang, C.Long, M. Chelliah, Y. Xue, B. Huang, J-K Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C-Z Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W Reynolds, G. Rutledge and M. Goldberg, 2010: "The NCEP Climate Forecast System Reanalysis." *Bull. Amer. Meteor. Soc.*, **91**, 1015-1057.
http://journals.ametsoc.org/doi/pdf/10.1175/2010BAMS3001.1 &
http://journals.ametsoc.org/doi/suppl/10.1175/2010BAMS3001.1 (supplemental material)

Simmons, A. J., K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee (2010), Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets, *J. Geophys. Res.*, 115, D01110, doi:10.1029/2009JD012442.