NOTES ON NWP AND THE HISTORY OF LONG-RANGE FORECASTING

Don Gilman (ret) and Huug van den Dool
NOAA/NCEP/CPC

Corresponding author:
H. M. van den Dool email:huug.vandendool@noaa.gov
Climate Prediction Center / NCEP
WWB, W/NP51 Rm 604 phone 301 763 8000X7570
5200 Auth Road fax 301 763 8395
Camps Springs MD 20746
U S A

Introduction and summary

In its early years Numerical Weather Prediction affected only one aspect of long range forecasting, the five day "extended" product. The basic synoptic techniques had been developed during and after World War II under the leadership of Jerome Namias, a former student and associate of Carl Rossby at MIT. Unlike the approaches through analogs, local time series and periodicities, sunspot effects, and daily surface chart extrapolation that prevailed elsewhere, Namias' procedure centered on predicting the hemispheric mid-tropospheric long waves in the mid-latitude westerlies, identified by five-day averaging of the 700 mb heights. Surface mean temperature and precipitation and the tracks of daily highs and lows were then to be inferred from the abnormalities of the 700 mb circulation. The 700 mb prog chart itself depended on a a subjective combination of Petterssen-type kinematics, constant absolute vorticity trajectories (CAVT), synoptic climatology, and vigilance toward possible explosive cyclogenesis over warm waters or the onset and retrogression of Atlantic blocking.

This procedure proved to be ideally suited to the incorporation of even the earliest barotropic model output and benefitted from every succeeding NWP improvement, as models went to full hemisphere, multi-level baroclinic, primitive equations, finer mesh, longer runs, and so on. Fig.1 clearly shows the improvement over the years and its relation to NWP.

Forecasters made adjustments to model output according to their experience of the model's weaknesses and the other tools, then fed their 700 mb predictions to statistical screening-regression equations (developed in the late 50's and thereafter by Bill Klein) to estimate the surface mean temperature and rainfall, again adjusting the output. As this "man-machine mix" system gained strength through the 1960's and early 70's, the forecast interval was moved out to days 3 – 7, the frequency from three to seven times per week, and daily weather sequences added to the synoptic isobar charts. At this point, extended forecasting was re-assigned to NMC's Forecast Division, and soon Jim O'Connor, Jim Andrews and Fran Hughes began issuing a supplementary 6-10 day forecast, the forerunner of present-day medium-range predictions.

Monthly forecasting had been started by Namias in the late 1940's in much the same technical framework as 5-day forecasting, with mean 700 mb patterns and surface temperature and precipitation, but with no implications of synoptic weather sequences inside the period. Because forecasts were prepared and issued as each month began, improvements in NWP over the first few days could and did add something to the skill of the product, which was therefore not truly "long-range" – certainly not a "climate prediction" --- but a useful hybrid for calendar-month planners. The true "long-range" product was the 90-day seasonal forecast, first issued in 1972. It was constructed by empirical techniques and was unaided by NWP until the 1990's.

For all intents and purposes, NWP did not play a role in the seasonal predictions issued by CAC/CPC until late 1994. Before that time the seasonal forecast was empirical as well as largely subjective. The big change, in terms of procedures, came in Dec 1994 when NCEP/CPC changed to the so-called Long Lead forecasts, and introduced several formal

prediction tools one of which was called the Coupled Model Forecast (CM). Although the CM has not truly outperformed other tools (such as Canonical Correlation Analysis and Optimal Climate Normals, or ENSO composites, when relevant), over the US it certainly attracted a lot of attention (and absorbed the most resources by far.)

The CM in real time forecasting was an outgrowth of building an ocean data assimilation system by Leetmaa and Ji in the late 1980's to early 1990's. To this end they started with the earliest version of the Modular Ocean Model (MOM) from GFDL, applied to the tropical Pacific only. Once the ocean data assimilation was up and running the model could be coupled to the best available and affordable version of NMC's MRF. Given computer power at the time, an atmospheric model resolution of T40L18 was affordable. Testing and tuning was performed by Kumar, Ji and Leetmaa.

Because of a large drift a so-called two-tiered system was designed, at NCEP and elsewhere, that is to say, the SST (after heavy statistical correction) as a function of time was the only thing saved from the truly coupled run, and kept underneath an atmospheric model (as prescribed lower boundary) for a 2nd integration of some 9 months. The 2nd tier was ran many times with the same SST, hence leading to an ensemble of solutions in the atmosphere. Outside the Pacific global SST was prescribed from such simple strategies as (damped) persistence, making the anomalies practically zero after a few months. Two tier systems are the main approach at many centers today, except at ECMWF, and NCEP where a new 1-tier truly coupled system is implemented in summer 2004. The NCEP atmospheric model is now a recent GFS T62L64 atmospheric model coupled to a recent MOM(3) version (global) from GFDL. This is due to hard work by Saha, Pan, Wang, Beringer and many others .

Another development before 1994 were so-called AMIP runs. W. Gates organized around 1990 what was called the Atmospheric Model Inter-comparison Projects in which participating models from dozens of centers were subject to a 10 year plus integration while specifying global SST from observations. The first such run at NMC was made by Ebisuzaki and Van den Dool, a revolution considering that weather forecast skill is gone in 2 weeks and runs beyond one year had never been made. With perfect SST these AMIP runs were like coupled model runs assuming a 'perfect' ocean, providing an upper limit for skill. The number of articles written about AMIP runs is truly astounding, and much of what we claim to know about global ENSO impacts through modeling is from AMIP runs .

A third hugely important development for seasonal prediction has been the NCEP/NCAR Reanalysis (Kalnay and Kistler and many others). While one may argue that empirical methods, like CCA, have nothing to do with NWP as such, they (CCA) would be impossible were it not for the existence of global initial states used in weather prediction. Hence Reanalysis created a goldmine for empirical work.

A final point of historical interest is to note that the idea of an a-priori skill assessment was applied to CM in late 1994. For the shorter range forecast this practice never existed and remains underdeveloped, but the introduction of NWP methods in the seasonal arena necessitated an assessment of CM over as many cases in the past as possible. Similar requirements applied to empirical-statistical

methods where testing a method under cross-validation comes more natural. AMIP runs were used as a quick and dirty procedure to assess an upper limit of forecast skill for the CM. (One could argue this is cheating because SST is never know perfectly ahead of time.) But times and computer power have changed. For the summer 2004 introduction of the new coupled model (now called CFS) a total of 2500 years of coupled model runs have been made thus allowing a fair assessment of this model's capability before it is used in subsequent real time forecasts.

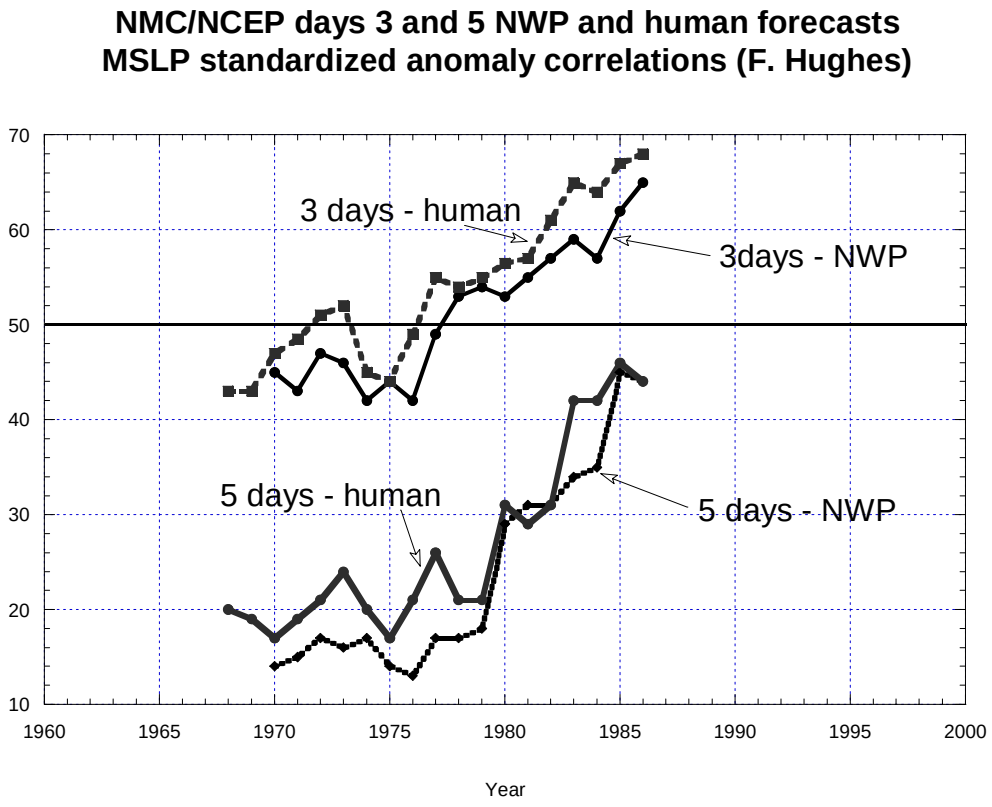
Seasonal prediction through NWP (or any other) methods is not a solved problem. Fig. 2 shows the skill (a correlation) of seasonal mean temperature from initial condition in February. The model used (2-tier CM) was operational from 2001-mid-2004. Skill was evaluated over 1979-1999 assuming perfect SST globally. In spite of this unrealistic assumption skill in excess of 0.3, 0.4 etc covers only a small portion of NA. Fig.3, similar to Fig.2 shows a skill evaluation for seasonal precipitation for the new CFS from October initial conditions (global ocean and atmosphere). Again skill assessment is modest, but at least there has been no cheating. The rather high correlation in Florida in winter is due to strong ENSO years, a fact confirmed by other methods and models as well.

While skill of seasonal prediction in general is modest over the US, there are a few areas and seasons where the news is better and moderately high skill seasonal forecasts are possible. Outside the US many areas in the tropics appear predictable to some degree. Apart from global SST other sources of predictability addressed in various models include continental soil moisture (another area where Namias was a pioneer) and global scale trends, the causes of which are not well understood, but may include anthropogenic activities.

Performance for SST forecasts is more stellar, certainly in the tropical Pacific Nino3.4 area (170W-120W, 5S-5N), the leading index for monitoring ENSO. Fig. 4 shows anomaly correlations of 0.75 at a lead of 6 months. These estimates are based on real time forecasts made by the old coupled model (CMP14), and three empirical methods (CCA (Canonical Correlation Analysis), CA (Constructed Analogue), and Markov) as well as a real time consolidation. The new model, run in hindcast obviously, over 1998-2003, shows a great improvement over the old model. But note that empirical methods are not far behind. In fact over the longer record, 1981-present, empirical methods are hard to beat.

Acknowledgment The authors thank Eugenia Kalnay for supplying an electronic version of Figure 1, and for much editorial assistance.

Fig. 1: Anomaly correlation of mean sea-level pressure over the years as maintained by Fran Hughes. The improvement at day 3 and day 5 is clearly due to improved NWP.



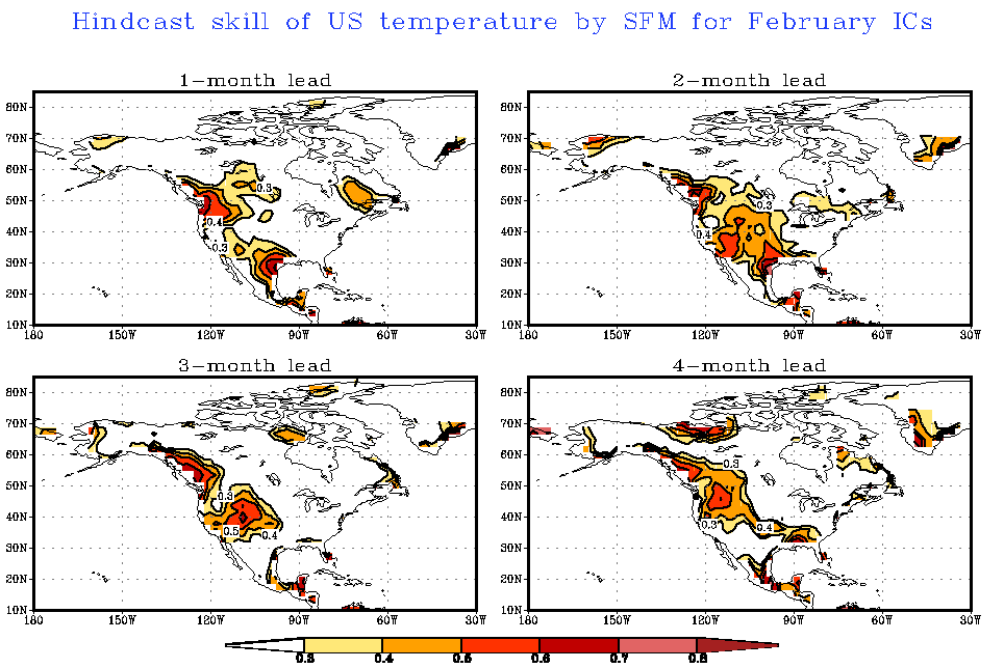


Fig.2: The skill of seasonal mean temperature forecasts over the North America, as evaluated by assuming perfect global SST, using a model of year 2000 vintage. The correlation rises above 0.4 only in a very limited area. Data courtesy of W. Wang

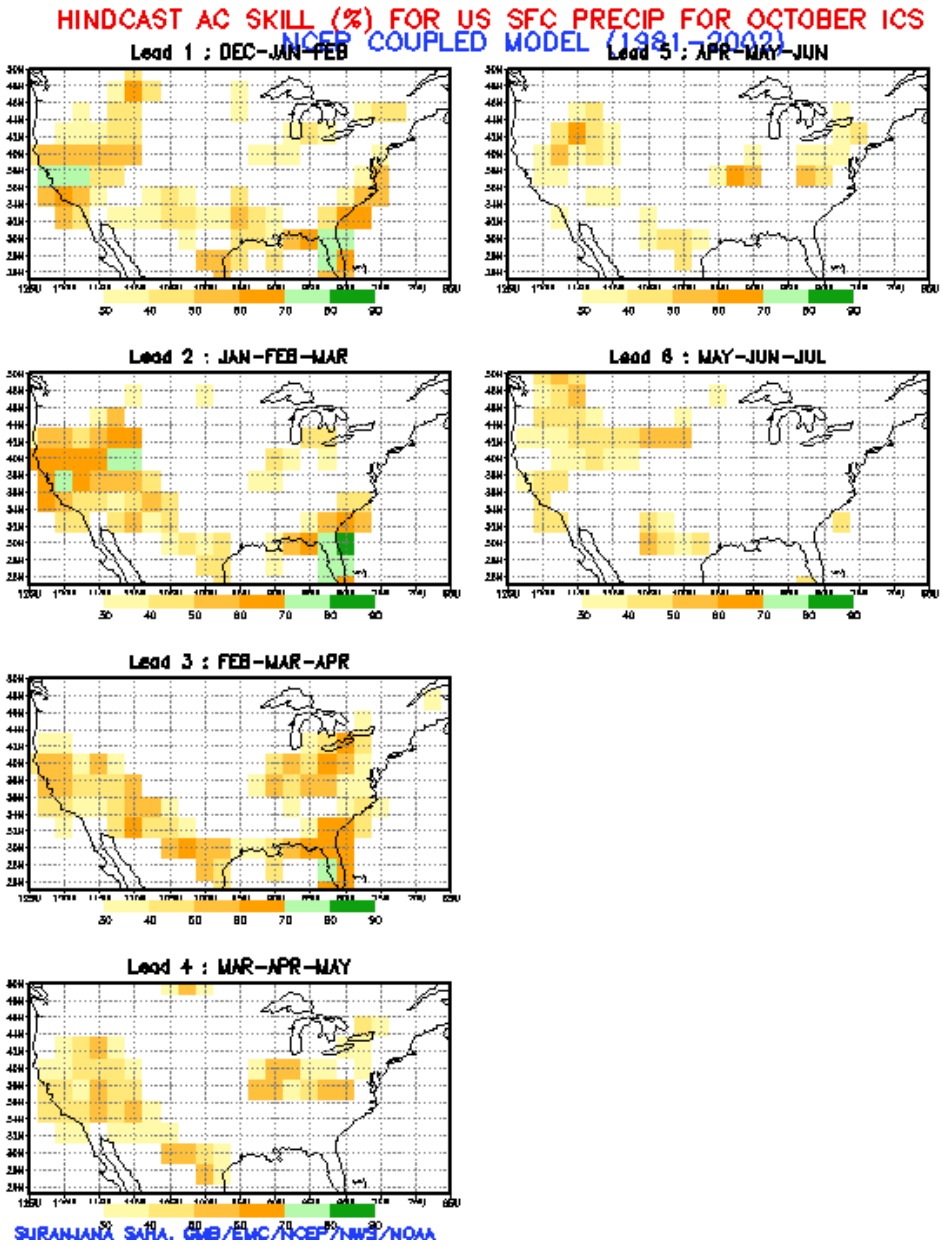


Fig.3: The skill of seasonal mean precipitation forecasts over the US, as evaluated by running a 1-tiered global coupled model of year 2003 vintage over hindcasts for 1981-2002. The correlation rises above 0.4 only in a very limited area. Some of the higher values are undoubtedly ENSO related. Data courtesy of S. Saha

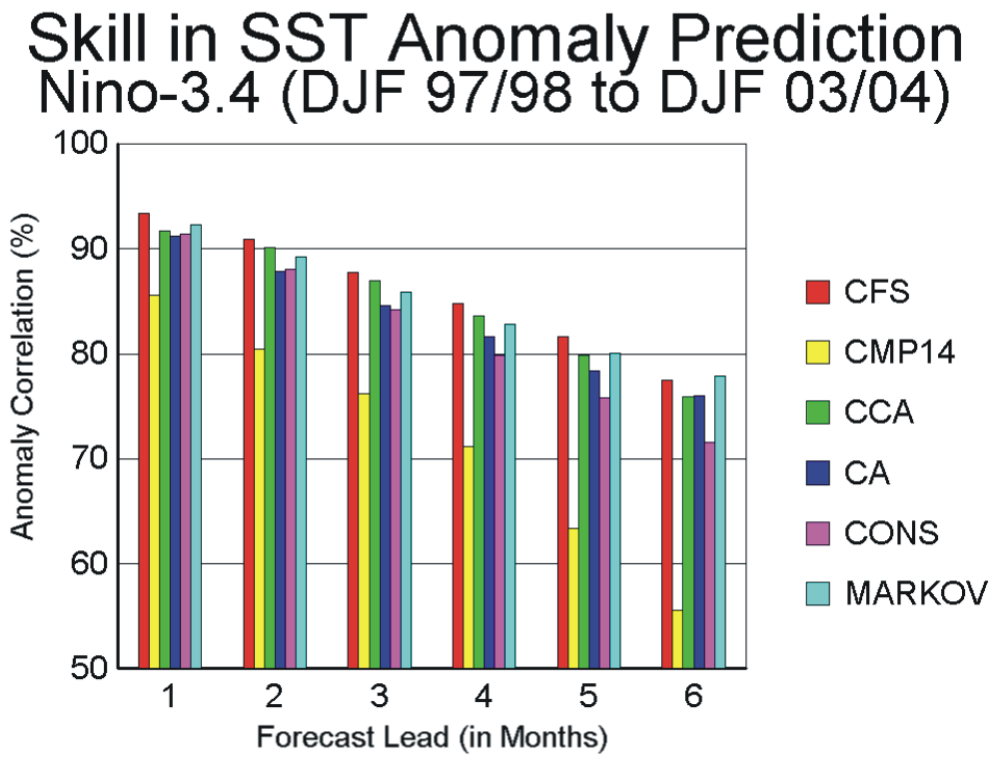


Fig.4 The anomaly correlation of predicting the Nino3.4 index in the Tropical Pacific as a function of lead for all 73 target seasons from 1998 through 2003 (not just DJF). Each histogram represents a different method. CFS is the new (summer 2004) coupled model (hindcasts). CMP14 is the older coupled model, CCA, CA and Markov are empirical methods, and CONS is a consolidation made in real time. Data courtesy D. Unger, S. Saha, L. He, H. van den Dool and Y. Xue