

Time-frequency localization and long- and short-term memories in the GFS precipitation forecasts

Lu, C., H. Yuan, S. Koch, E. Tollerud, J. McGinley, and P. Schultz

NOAA Earth System Research Laboratory and CIRA, Colorado State University

1. Introduction

Precipitation possesses its own temporal variability, such as seasonal and diurnal cycles. It is important for numerical weather prediction (NWP) models to correctly capture these natural variabilities. Meanwhile, due to uncertainty in the forecast model, there are day-to-day random forecast errors. Hence, a given time series of precipitation forecast errors will most likely consist of variability on several different scales. Distinguishing these signals and variabilities from this complex dataset may provide us with crucial insight into the temporal characteristics of forecast errors.

The first issue we investigate in this study is the frequency of our current precipitation forecast errors have, and the time at which these errors occur. In another words, we want to know what time scale the NOAA/NCEP's Global Forecast System (GFS) precipitation forecast errors project to, (e.g., daily, weekly, or otherwise), and whether these errors occur on a certain day, or in a particular month, or season. The time scale of errors may provide us with the information about the type of precipitation systems that the GFS may have difficulty forecasting. The time information may help us pinpoint the possible physical reasons why the GFS generates these large forecast errors.

Once we identify different time scales of forecast errors, we would like to know if these errors possess any time memory. Even errors that appear to be random may still have some built-in persistence or correlation. This is the basic concept of forecast error memory. The different length of these memories may indicate possible model intrinsic deficiencies in different precipitation parameterizations corresponding to different physical scales.

In this study, we will use a continuous wavelet to conduct the time-frequency localization of precipitation forecast errors. We will then conduct an analysis for the multi-moment correlation function of the forecast errors. This is known as the Hurst parameter analysis (Lu and Koch 2007). From these analyses, we may gain a better understanding of temporal characteristics of precipitation forecast errors.

2. Forecast model and observational data

The NOAA/NCEP GFS daily precipitation forecasts are verified against observations considered to be verification truth. These observations are a set of global satellite precipitation estimates, produced and archived at the University of California-Irvine. This satellite data product is called

PERSIANN. Detailed information about this dataset is described in Yuan et al. (2007).

The GFS data is archived on a 1x1 latitude/longitude degree grid, while the satellite estimates are archived on a 0.25x0.25-degree grid. A 24-h precipitation accumulation from GFS's 1-day forecast and from PERSIANN estimates are used.

3. Spatially averaged time series of precipitation forecast errors

Root mean square (RMS) errors of GFS daily precipitation forecasts are computed for 2005-2006 by verifying the GFS 1-day precipitation forecasts against UC-Irvine's PERSIANN satellite observations. These errors are then averaged over different geographic areas. Figure 1 shows the definition for various geographic areas over which spatial averaging is conducted. The classification of different geographic areas is based on considerations of continents vs. adjacent oceans (e.g., CONUS vs. Pacific/Atlantic Oceans), northern and southern hemispheric counterparts of continents and oceans (e.g., North America vs. South America; the North Pacific vs. the South Pacific), and tropics vs. extratropics (e.g., equatorial vs. mid-latitude regions).

Spatially averaged time series of precipitation forecast errors for four different geographic areas are shown in Fig. 2. They are: North America (Fig. 2a), South America (Fig. 2d), the North Pacific (Fig. 2c), and the South Pacific (Fig. 2d). Daily forecast error variances superimposed on seasonal variation can be seen in these figures, except for the

South Pacific in Fig. 2d, where the winter dry season signal is very weak and brief. North America and South America have opposite seasons, as expected. In general the daily error variances are larger over the continents than those over the oceans.

4. Time-frequency localization of GFS forecast errors

Continuous wavelet (CW) analyses of these error time series are conducted. These analyses project the GFS precipitation forecast errors onto both time and frequency subspaces. Mathematically, the wavelet transformation can be expressed as

$$\hat{f}(\Omega, t) = \int f(\tau) \psi_{\Omega, t}^*(\tau) d\tau, \quad (1)$$

where $\psi_{\Omega, t}^*$ is a wavelet kernel function, the asterisk stands for a conjugate, $f(\tau)$ denotes the time series, and $\hat{f}(\Omega, t)$ is wavelet time-frequency decomposition of the original time series.

Figures 3a-d show the error power as a function of time (abscissa) and period (ordinate). The most noticeable but not surprising result from these analyses is the error scale separation: errors with a time scale of 1-10 days and seasonal errors (~150 days). As before, the forecast for the South Pacific (Fig. 3d) presents a vague seasonal change, with a time scale that extends to more than 200 days. The North American Continent (Fig 3a) peaks its forecast errors in the northern summer, while South America (Fig. 3b) peaks its errors in the southern summer. In addition, the North Pacific Ocean (Fig. 3c) peaks its seasonal error about one month (in May) earlier than

that for the North American Continent (in June).

Short time-scale forecast errors typically occur in the time periods of 1-10 days; most of them peak around two days. These errors may represent synoptic-scale or mesoscale forecast errors. Taking North America as an example, there are large signals occurring between the months of April and May likely related to severe spring storms in the Central Plains; in the months of August to October related to landfall of tropical storms and North American monsoon rainfall; and in the months of December and January related to North American winter storms.

5. Hurst parameter analysis

Finally, we compute the Hurst parameters to determine if different scales of forecast errors possess any time memories. The central idea is to compute multimoments of a correlation function (structure function), $S_q(\tau)$, for different lags τ :

$$\begin{aligned} S_q(\tau) &= \langle |\varepsilon(t+\tau) - \varepsilon(t)|^q \rangle \\ &= \langle |\delta\varepsilon(\tau)|^q \rangle, \end{aligned} \quad (2)$$

where $\varepsilon(t)$ is the error time series and angle brackets indicate an ensemble average. According to the similarity theory, if such computed correlation function can be scaled by

$$S_q(\tau) = C_q \tau^{\zeta(q)}, \quad q \geq 0, \quad (3)$$

where C_q is a proportional coefficient and $\zeta(q)$ is a scaling power function, the process is said to be self-similar or self-

affine. The Hurst parameter is defined as $H(1) = \zeta(1)$ (the first multifractal parameter). If $\zeta(1)$ is a linear function on a log-log plot of $S_1(\tau)$, then persistence (memory) exists. High moments of $S_q(\tau)$ provides an indication if such persistence will break.

Figures 4a-4d show the multimoment correlation function on a log-log plot for North America, South America, the North Pacific, and the South Pacific, respectively. First, we can see [by examining the higher moments of $S_q(\tau)$] that there exists an error scale break at about 10 days for all these geographic areas. This is consistent with the wavelet analysis results obtained in the previous section. Second, these forecast errors possess a seasonal memory of about 150 days, although errors in the Southern Hemisphere do have a tendency to correlate slightly longer. Third, a short memory of 4 days (slightly longer for oceanic areas) can be found in these forecast errors.

6. Conclusions

The GFS precipitation forecast errors possess two distinctive scales. One is a seasonal scale, with about a 150-day period. The second scale varies between 1-10 days, and may be related to synoptic or mesoscale precipitation systems.

For seasonal-scale forecast errors, North America and South America tend to peak in opposite seasons. Errors over the North Pacific tend to reach a maximum about one month earlier than those over the North American Continent. The South Pacific Ocean has

a vague seasonal signal. There is a strong presence of synoptic-scale and mesoscale forecast errors over North America in the spring, late fall, and winter that can be physically related to typical precipitation weather systems in North America.

The forecast error scale separation is further confirmed in the multimoment correlation function analysis. The Hurst parameter is calculated, which reveals that the GFS precipitation forecast errors possess a short memory of about 4 days and a long memory of about 150 days.

Acknowledgments: We would like to thank Ann Reiser, who conducted technical editing. This research is in response to requirements and funding by the NOAA Office of Oceanic and Atmospheric Research. The views

expressed are those of the authors and do not necessarily represent the official policy or position of the NOAA.

References

Lu, C., and S. Koch, 2007: Interaction of upper-tropospheric turbulence and gravity waves as obtained from spectral and structure function analysis. In the process submitting to *J. Atmos. Sci.*

Yuan, H., C. Lu, E. Tollerud, J. McGinley, P. Schultz, 2007: Analysis of precipitation forecasts from the NCEP global forecast system. *AMS 22nd conference on Weather Analysis and Forecast*, Park City, Utah.

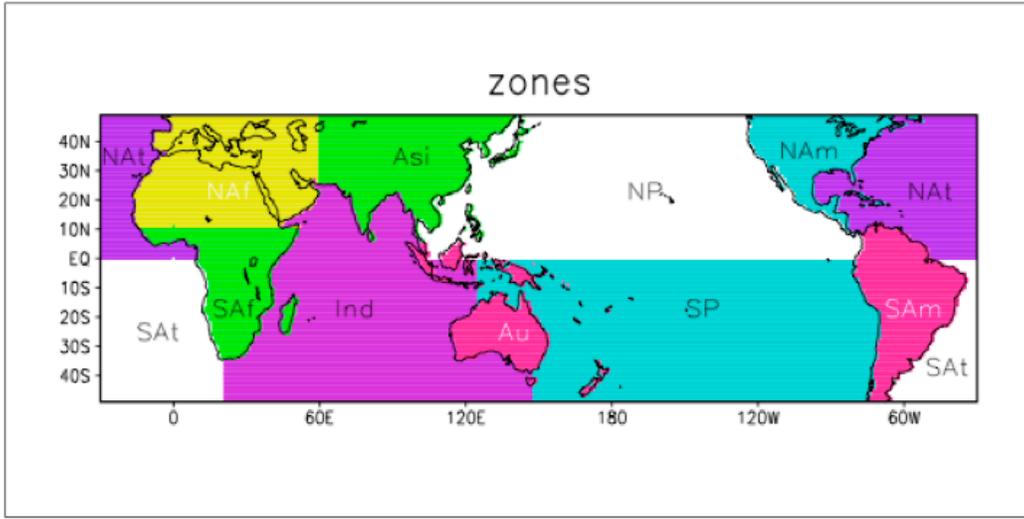


Fig. 1: Geographic classification for error comparison.

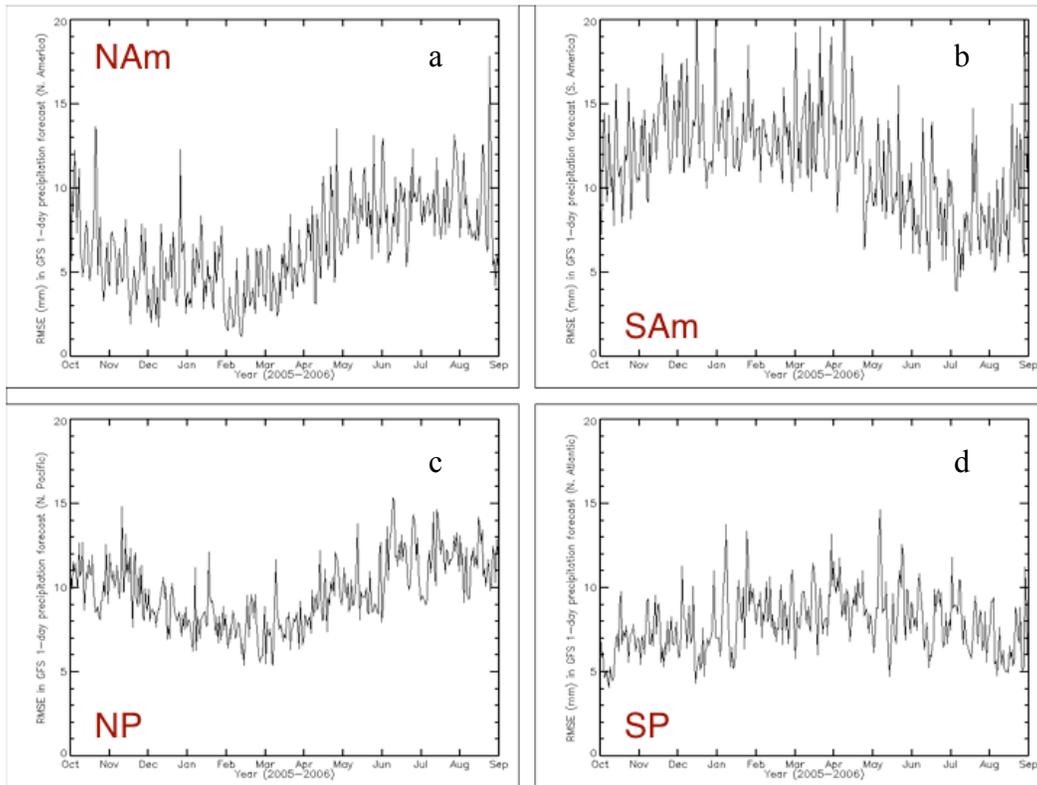


Fig. 2: Spatially averaged time series of GFS precipitation forecast errors over a) North America, b) South America, c) North Pacific, and d) South Pacific (see Fig. 1 for geographic definitions).

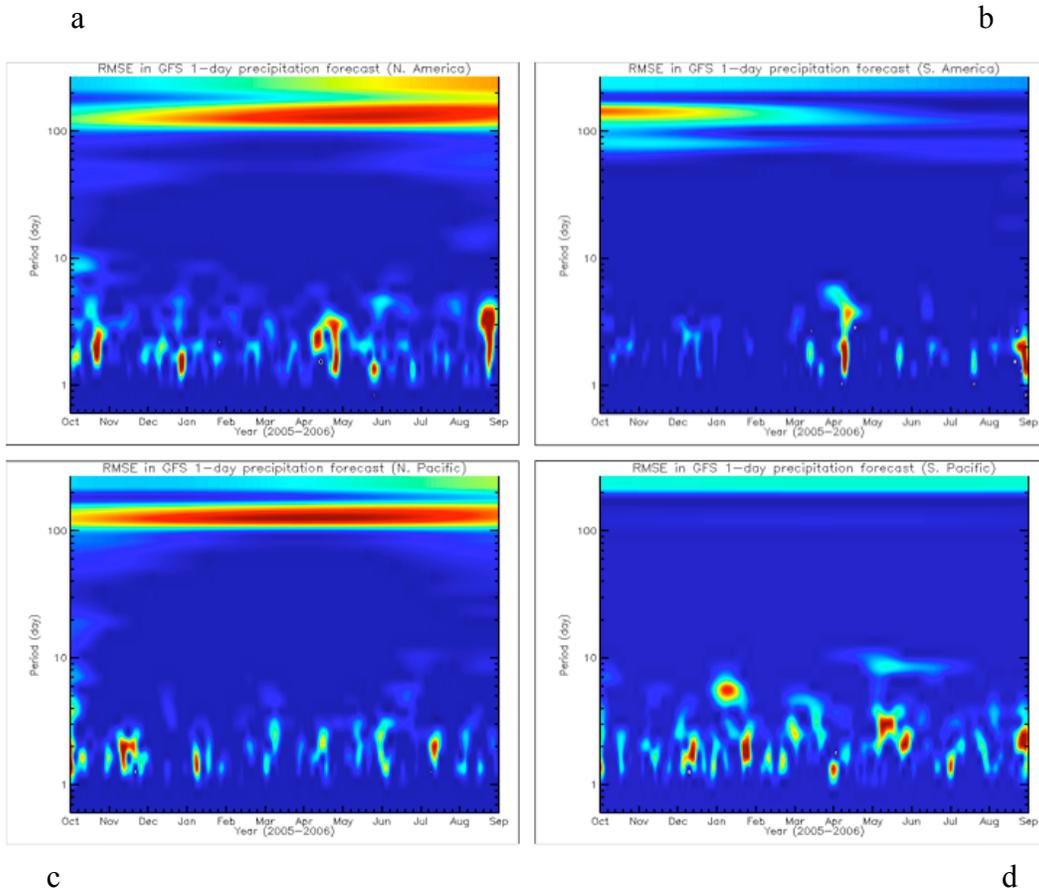


Fig. 3: Wavelet time-frequency analysis of GFS precipitation forecast errors over a) North America, b) South America, c) North Pacific, and d) South Pacific.

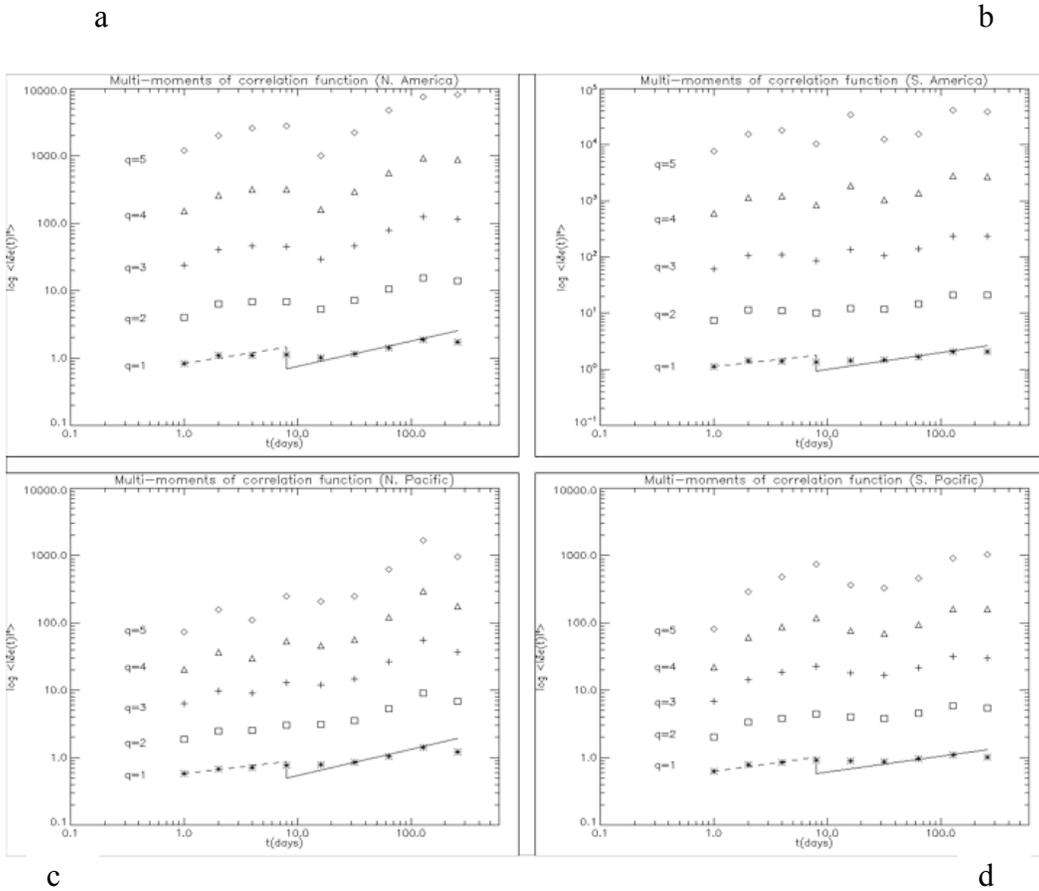


Fig. 4: Log-log plots of multimoments of the correlation function for GFS precipitation forecast errors over a) North America, b) South America, c) North Pacific, and d) South Pacific. The dash and solid lines are least-square fitting of $S_1(\tau)$, which give the range of the autocorrelation (the memory).