Anomaly Correlation "standard error"

The estimate of mean anomaly correlation at each forecast projection has an associated uncertainty due, in part, to the small number of runs going into the sample. We do not show AC's from the underlying individual runs (although they can be seen on the AC time series page), but we can estimate the uncertainty in the mean AC's by considering the variations in AC over *the time series of the matched model runs in the sample*. The uncertainty on the mean ("standard error") is estimated as:

Standard Error = $\sigma / \sqrt{(n\text{-}1)(1\text{-}\varphi)}$

where σ is the standard deviation, n is the number of AC values in the sample, x is the set of AC's, and φ Is the lag one autocorrelation derived from the time series x. This is empirically derived from the AC values with the approximation:

$\varphi \cong cor\ (\ x\_(1 \ldots (n-1)\ ), x\_(x \ldots n)\ ).$

The estimate of the standard error on the mean is distinct and separate from the standard deviation from the sample. The standard error is an estimate of how well we understand the underlying, fundamental AC at each forecast projection. The standard deviation is an estimate of how far off the mean value any one AC may be (Weatherhead et al., 1998). Thus, were our sample of runs to be repeated in a similar season and for a similar duration, we can say that the mean AC has a 67% likelihood of being within one standard error of our results, and a 95% likelihood of being within two standard errors of our results.

This approach at least partially accommodates the fact that the individual AC's are auto-correlated. Physically, this implies that weather situations, which can be more or less forecastable, can last for more than one day, thus the sequential forecast AC's are not independent estimates, but represent an over-sampling of the system. The standard error equation above accounts for these. It should be noted that the lag 1 Auto-Regressive assumption generally refers to a 12-hour lag (for runs that occur every 12 h) or 24-hour lag (for daily runs) as the most significant approximation to the autocorrelation. It should also be noted that aggregation of 00Z and 12Z results allows for a larger sample size, but may result in combining different physical causes of differences as well as different statistical properties of the time series.

We generate the errors on the AC *differences* in the same way. That is, at each forecast projection, we generate a time series of AC differences, and calculate the mean and standard error for the time series of differences using the methods described above.

In the figures generated, standard errors are indicated by boxes. Differences of 1 standard error are significant at the 67% confidence level; differences of 2 standard errors are significant at the 95% confidence level.

Reference:

Weatherhead, E. C., G. C. Reinsel, G. C. Tiao, X. Meng, D. Choi, W. Cheang, T. Keller, J. DeLuisi, D. J. Wuebbles, J. B. Kerr, A. J. Miller, S. J. Oltmans, and J. E. Frederick, 1998: Factors affecting the detection of trends: Statistical considerations and applications to environmental data. Journal of Geophysical Research, 103, 17149-17161.