

1. Mean square error factorization

Murphy (1988,1997) shows that the mean square error (MSE) of a set of forecasts can be expressed in terms of sample statistics of the forecasts and the corresponding set of observations:

$$\text{MSE}(f, x) = (\mu_f - \mu_x)^2 + \sigma_x^2 + \sigma_f^2 - 2\sigma_x\sigma_f\rho_{fx}. \quad (1)$$

In Eq. 1:

μ_f = mean forecast

μ_x = mean observation

σ_f^2 = forecast variance

σ_x^2 = observation variance

ρ_{fx} = forecast/observation correlation coefficient

The first and last terms on the RHS of Eq. 1 can be described as follows:

$(\mu_f - \mu_x)^2$ = square bias

$2\sigma_f\sigma_x\rho_{fx}$ = twice the covariance of the forecasts and observations

2. Climatology-derived baseline forecasts and their MSE

i. Static forecasts

A set of ‘static’ forecasts has, by definition, $\sigma_f^2 = 0$. Applying this to Eq. 1, the MSE for a set of static forecasts is

$$\text{MSE}_{static}(f, x) = (\mu_f - \mu_x)^2 + \sigma_x^2. \quad (2)$$

A static forecast of the sample climatological mean, $\mu_f = \mu_x$, minimizes Eq. 2. In a forecast skill calculation, using the sample climatological mean as the baseline forecast thus provides the most rigorous challenge amongst possible static baseline forecasts. A climatological mean calculated from (partially or fully) out-of-sample observations is likely to be biased ($\mu_f \neq \mu_x$), leading to a higher MSE and less of a challenge in a forecast skill calculation.

ii. Dynamic forecasts

If one suspects that the underlying climatological distribution may not be static for some reason, it is appropriate to define a time-varying climatological baseline forecast.

One way to do this is to calculate a ‘moving window’ climatological mean forecast (e.g. mean based only on the 5 previous years), but this necessitates using out-of-sample observational data that can lead to MSE-degrading bias. An unbiased alternative is to define climatological forecasts according to a linear least-squares fit of the in-sample observations. This essentially allows one more degree of freedom in the fit than the static case (a sloped forecast time series, rather than a horizontal one).

3. Application to Atlantic basin seasonal TC forecasts

The MSEs for climatology-derived baseline forecasts of the number of named tropical cyclones, hurricanes, and intense hurricanes are shown in the lower portions of Tables 1-3, respectively. All terms of the MSE factorization in Eq. 1 are included in the Tables, along with correlation coefficients. The two static baselines are ‘SC mean’, the sample climatological mean, and ‘OSC mean’, calculated from out-of-sample observational data¹. The two dynamic baselines are ‘SC trend’, the linear fit to the sample observational data, and ‘OSC moving window’, which uses the prior 5-year climatological mean (drawing partially on non-sample observational data). For the number of named tropical cyclones and hurricanes, forecasts are made for the years 1984-2005. For the number of intense hurricanes, forecasts are made for the years 1990-2005. Observational data is gathered from the spreadsheets Phil generously provided. For reference, time series of the four climatology-derived baseline forecasts and the observations are plotted in Figures 1-3.

It is clear from Tables 1-3 that SC mean and OSC mean differ exclusively because the bias in the out-of-sample climatological mean provides a positive contribution to the MSE that does not exist for the sample climatological mean. The reason that the MSE of SC trend is lower than OSC moving window (for all 3 variables) is not so clear-cut. Bias in the OSC moving window forecasts accounts for only part of the difference. The rest must be attributed to the combination of higher correlation coefficient and higher forecast variance in the SC trend forecasts. SC trend is simply a better fit to the observations (in a least square sense) than OSC moving window, even if the bias in OSC moving window is ignored.

4. Comparison to Gray et al. forecasts

Verification measures of the June and August Gray et al. forecasts (from Phil’s spreadsheets) are included in the upper portions of Tables 1-3. The MSE of the August forecasts (for all three variables) is lower than any of the four baselines, although the magnitude of the MSE difference varies substantially amongst the four baselines. The MSEs of the June forecasts are much closer to those of the baseline forecasts. In particular, the June forecasts and SC trend have similar MSEs (and in the case of hurricanes, similar MSE factorizations). In the case of the June forecasts, even the sign of the MSE difference depends on the baseline.

¹ Specifically, it is sample climatological mean for the years 1944-1983 (for named tropical cyclones and hurricanes) and 1944-1989 (for intense hurricanes), as reported by Owens and Landsea (2003). There is no common ground between these samples and the samples used to calculate ‘SC mean’.

Table 1: Atlantic basin named tropical cyclones, 1984-2005

$\sigma_x^2 = 22.59$	MSE	$(\mu_f - \mu_x)^2$	σ_f^2	$2\sigma_f\sigma_x\rho_{fx}$	ρ_{fx}
August forecast	7.23	1.51	9.69	26.56	0.90
June forecast	14.00	1.62	4.72	14.93	0.72
SC trend	15.47	0	7.12	14.25	0.56
OSC moving window	19.42	2.53	3.84	9.54	0.51
SC mean	22.59	0	0	0	0
OSC mean	30.69	8.10	0	0	0

Table 2: Atlantic basin hurricanes, 1984-2005

$\sigma_x^2 = 9.22$	MSE	$(\mu_f - \mu_x)^2$	σ_f^2	$2\sigma_f\sigma_x\rho_{fx}$	ρ_{fx}
August forecast	4.41	0.25	3.51	8.57	0.75
June forecast	6.64	0.03	2.34	4.95	0.53
SC trend	6.77	0	2.45	4.90	0.52
OSC moving window	8.61	0.56	1.47	2.63	0.36
SC mean	9.22	0	0	0	0
OSC mean	10.18	0.96	0	0	0

Table 3: Atlantic basin intense hurricanes, 1990-2005

$\sigma_x^2 = 4.36$	MSE	$(\mu_f - \mu_x)^2$	σ_f^2	$2\sigma_f\sigma_x\rho_{fx}$	ρ_{fx}
August forecast	2.19	0.47	1.87	4.52	0.79
June forecast	3.31	0.32	1.00	2.36	0.57
SC trend	2.70	0	1.66	3.32	0.62
OSC moving window	4.11	0.36	0.85	1.47	0.38
SC mean	4.36	0	0	0	0
OSC mean	5.22	0.86	0	0	0

Figure 1

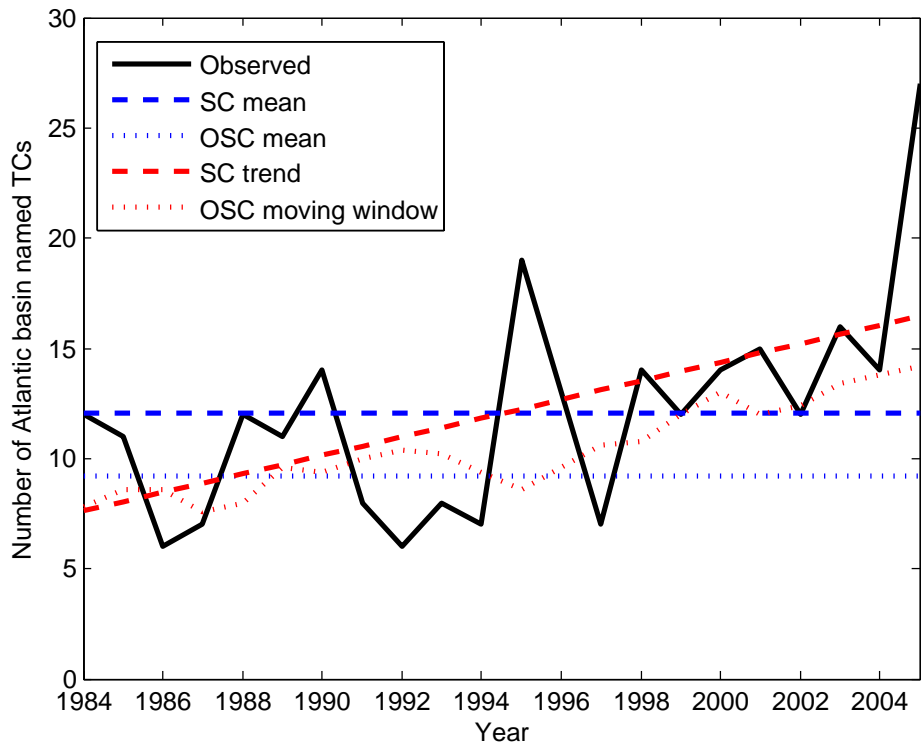


Figure 2

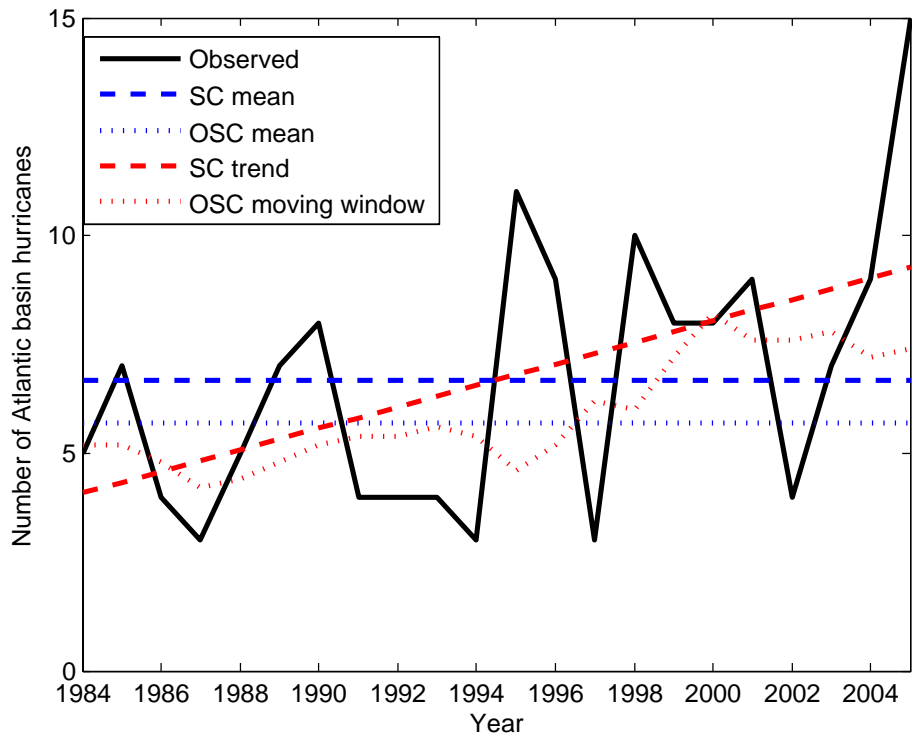
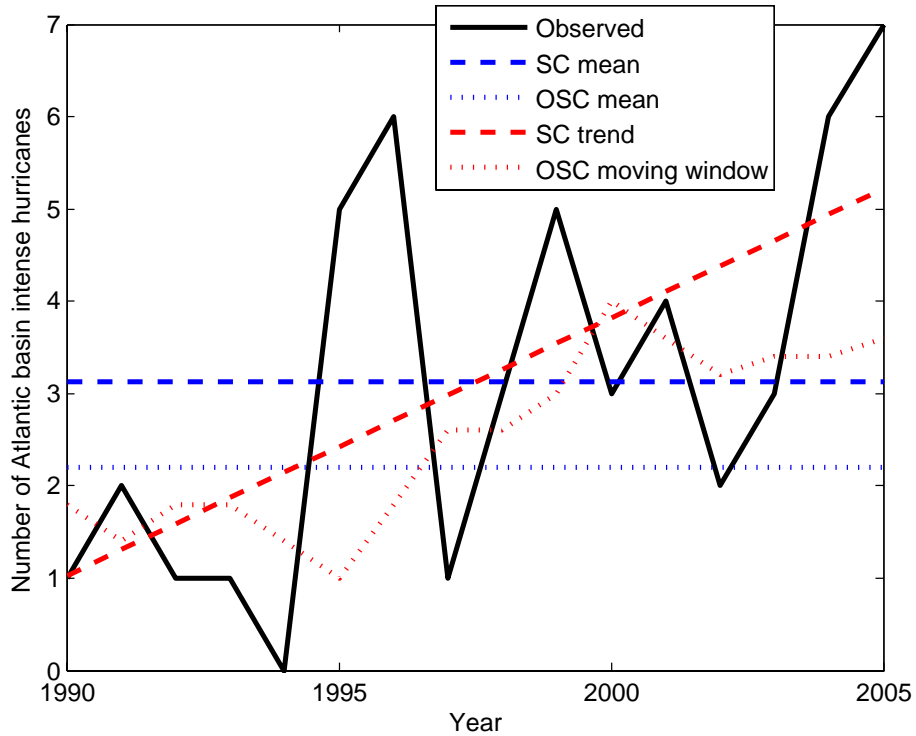


Figure 3



References

Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424.

Murphy, A. H., 1997: Forecast verification. *Economic Value of Weather and Climate Forecasts*, R. W. Katz and A. H. Murphy, Eds., Cambridge University Press, 19–74.

Owens, B. F., and C. W. Landsea, 2003: Assessing the skill of operational Atlantic seasonal tropical cyclone forecasts. *Wea. Forecasting*, **18**, 45–54.