Summary of the March 2013-initialized experimental forecasts for 2013 North Atlantic seasonal hurricane frequency using the GFDL hybrid (statistical-dynamical) hurricane forecast system (GFDL-HyHuFS; *Vecchi et al.*, *2011, MWR*):

Prepared by Gabriel Vecchi, NOAA/GFDL – 29-March-2013.
This, past and future summaries, as well as discussion of HyHuFS are available at:
http://www.gfdl.noaa.gov/HyHuFS

Note: The results described in this document are **not** an official outlook. This is a research product on the continued verification and evaluation of an experimental forecast system. We make these experimental forecast results available in order to facilitate and motivate research and discussion on the topic of long-lead seasonal hurricane forecasts.

Special Note: The March 2013 experimental forecast comes from an updated version of the GFDL-CM2.1 initialized climate forecast system. Taking advantage of a computer upgrade – which required us to re-run the retrospective forecasts – we moved the GFDL-CM2.1 experimental seasonal forecast system to the latest version (v3.1) of the Ensemble Kalman Filter initialization system. In this summary we present the forecasts using the v3.1 system and present a comparison of the retrospective forecast skill using initialization from the old (v1.0, as published in Vecchi et al. 2011) and new initialization systems. Overall, we do not see any significant change to performance, although there is a nominal increase in retrospective correlation and decrease in retrospective RMS error for the March initialized forecasts from the new system – but nominal degradations & enhancements in other months.

1. Summary: As of the March 2013 initialization, HyHuFS is predicting that the frequency of Atlantic hurricanes in the 2013 season should be more active than the 1982-2012 climatology, and slightly more active than the recent 1995-2012 mean. HyHuFS indicates a reduced probability of an extremely inactive 2013 Atlantic hurricane season, an enhanced probability of above-average hurricane frequency and a slightly enhanced probability for an extremely active hurricane season.

The experimental forecast for the 2013 season with GFDL-HyHuFS initialized in March 2013 gives an expected value of 8.4 hurricanes. These forecast values arise because the coupled GCM used in the system predicts that the tropical Atlantic should be substantially warmer than the long-term (1982-2011) average (by about 0.44°C), with warm anomalies in the remote tropics (0.14°C) slightly offseting the impact of a warm Atlantic.

2. Forecast system description: This is a brief description of the experimental hurricane forecast system, HyHuFS, further details are available in Vecchi *et al.* (2011). The forecast system is hybrid statistical-dynamical, applying a statistical model of hurricane frequency to the output of initialized GCM forecasts (for this forecast the GFDL-CM2.1 system was used; Delworth *et al*. 2006, Zhang *et al.* 2007). The statistical hurricane frequency model is built to emulate a high-resolution atmospheric general circulation model (Zhao *et al.* 2009, 2010), using Poisson regression (Villarini *et al.* 2010) and two SST-based predictors: tropical Atlantic SST and tropical-mean SST. The statistical model shows a positive sensitivity to Atlantic warming and a negative sensitivity to tropical-mean warming, reflecting the strong correlation between hurricane frequency and the warming of the Atlantic relative to the tropics. The system generates explicit probabilistic ranges based on

the ensemble spread of the GCM forecasts ("climate noise") and the uncertainty explicit in the statistical model ("weather noise").

The forecasts summarized here are based on the forecasts initialized in March 2012.

3. Detailed forecast values: **Table 1** (below) presents the mean forecast values based on HyHuFS using v3.1 of the initialization system, along with the explicit 50% and 75% ranges. **Table 1** also indicates the values from HyHuFS-v3.1 averaged over the long-term (1982-2011) and the recent "active" period (1995-2011), along with retrospective correlations and RMS values, to guide interpretation. These retrospective skill measures are over the 1982-2011 period.

The bottom rows of **Table 1** show the retrospective skill of Version 1 of HyHuFS (HyHuFS-v1.0, which differs from v3.1 in the initialization scheme, slight changes to the coupled model and having been run on a different computer.) The inclusion of 2010 and 2011 (the first "non-retrospective" forecasts for HyHuFS-v1.0) leads to a small nominal increase of the correlation coefficient for the March-initialized forecasts of this system (0.49 to 0.55) relative to the correlation skill reported in Vecchi *et al.* (2011) – more discussion on the continued performance assessment of HyHuFS is provided in Section 4. **Figure 1** shows the retrospective skill and the experimental forecast for 2013 using HyHuFS-v3.1.

Forecast source	Mean count	Median count	50% range	75% range	1982-2011 Retrospective Correlation of Mean	Retrospective RMS error of Mean
V3.1 March 2013 initialized forecast for 2013 Season	8.37	8	6-11	5-12	0.58	2.54
V3.1 1982-2012 Average for March initialized forecasts	6.39	6	4-9	2-11		
V3.1 1995-2012 Average for March initialized forecasts	7.35	7	5-10	3-13		
V1.0 1982-2011 Average for March initialized forecasts	6.31	6	4-9	2-11	0.55	2.57
V1.0 1995-2011 Average for March initialized forecasts	7.27	7	5-10	3-13		

Table 1: Summary of the forecast for the 2013 North Atlantic hurricane season initialized March 2013, based on HyHuFS using v3.1 of the EnKF initialization. Top row indicates the expected value, median and selected uncertainty ranges for the number of Atlantic hurricanes in the 2013 season, along with the retrospective correlation and RMS error of the system initialized in March. The second and third rows summarize the statistics of HyHuFS-v3.1 when initialized in March for the whole record and for the recent active era. The last two rows summarize the statistics for the defunct HyHuFS-v1.0 when initialized in March for the whole record and for the recent active era.

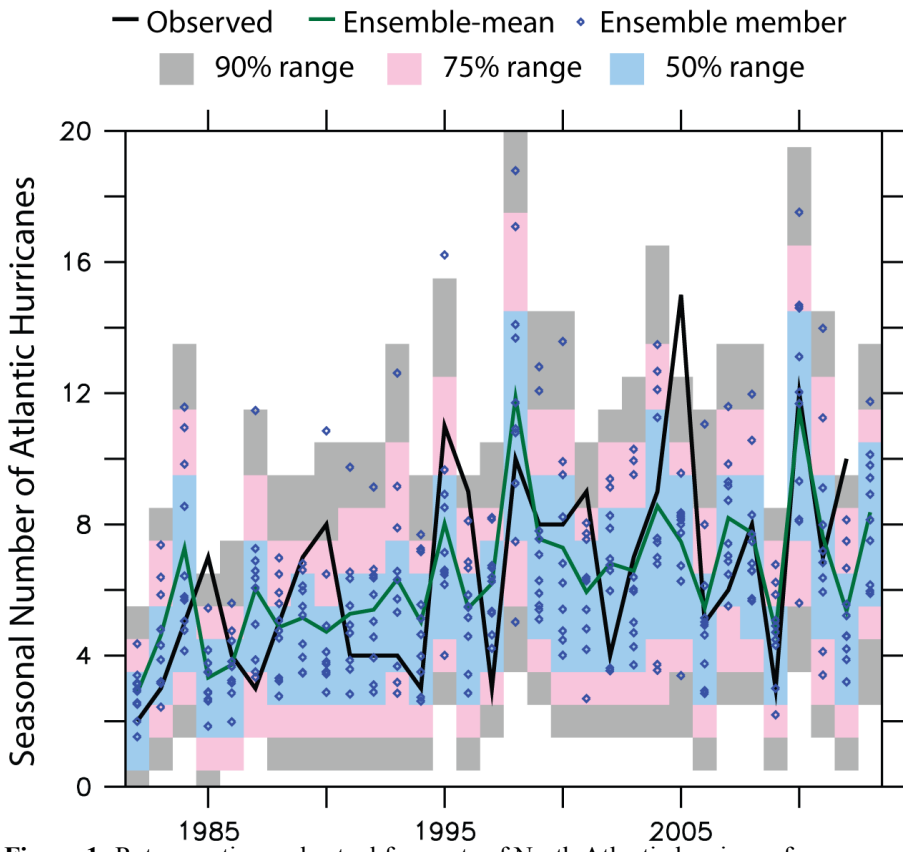

Figure 1: Retrospective and actual forecasts of North Atlantic hurricane frequency using an experimental hybrid seasonal hurricane forecast system (Vecchi et al. 2011). Both panels show the observed seasonal North Atlantic hurricane frequency each season 1982-2011, along with forecasts for the 1982-2011 seasons initialized in March based on HyHuFS using the GFDL-CM2.1 GCM – the upper panel includes forecasts for the upcoming 2013 season. The panel shows output from the new version of the system (HyHuFS-v3.1), which replaced the now-defunct version (HyHuFS-v1.0) that was the basis of Vecchi et al. (2011). The black line indicates observed hurricane counts, green line indicates the mean forecast value, shading indicates the confidence intervals computed by convolving inter-ensemble spread and statistical model uncertainty.

Table 3 at the end of the document provides the ensemble-mean and median forecasts for each season (1982-2013) in order to facilitate comparison to other forecast systems and extended evaluation of the reliability of the experimental forecast system (*i.e.,* using different statistical tests than those presented in Vecchi *et al.* (2011) or here).

In addition to the ranges, medians and means, HyHuFS forecasts the entire probability density function (PDF), which allows us to compute certain "exceedance probabilities" (the probability that North Atlantic hurricane frequency will exceed a certain number). This allows the forecast system to give the probability that the season will be "above average" (more than 6 hurricanes), "very inactive" (3 or fewer hurricanes) or "very active" (than 10 hurricanes), see **Table 2** (full 1982-2013 statistics in **Table 3**). HyHuFS indicates an enhanced probability for both a "very active" and "active" season relative to the entire 1982-2012 period, and somewhat heightened compared to the recent active era (1995-2012). Meanwhile, HyHuFS indicates a reduced probability of a "very active" season relative both the long-term (1982-2012) and recent (1995-2012) performance of the system.

Forecast source	Probability of an "above average season" (>6 hurricanes)	Probability of a "very inactive" season (≤3 hurricanes)	Probability of a "very active" season (>10 hurricanes)
V3.1 March 2013 initialized forecast for 2013 Season	68%	6%	25%
V3.1 1982-2012 Average for March initialized forecasts	41%	24%	15%
V3.1 1995-2012 Average for March initialized forecasts	52%	16%	22%

Table 2: Summary of the exceedance probabilities for an experimental forecast of the 2012 North Atlantic hurricane season initialized March 2012, based on HyHuFS. Top row indicates the forecast probabilities for an above average, very inactive and very active season with respect to the total number of Atlantic hurricanes in the 2012 season initialized in March 2012. The last two rows summarize the system's statistics when initialized in March for the whole record and for the recent active era.

4. Updated analysis of past performance: When the experimental HyHuFS long lead seasonal forecast system was described in Vecchi *et al.* (2011), its skill in "retrospective forecast" mode was evaluated over 1982-2009. *Retrospective forecasting* is an attempt to estimate forecast quality by simulating how the system would have performed had it been in existence to forecast past years, by using only information that would have been available at the time that forecasts would have been performed – but with a system designed in the present (*i.e.,* not available in the past). Retrospective forecast evaluation is a necessary step to establish the potential of a forecast system, yet it is not sufficient: since retrospective forecasting is done in the present it cannot be completely free of information about the past, and past skill may not represent the true forecast skill of a system. Therefore, it is essential to continue evaluating the performance of a forecast system on "real" forecasts – that is, forecasts about the future.

Since HyHuFS-v1.0 was finalized and its results submitted for publication, there were two "real" forecasts to evaluate: the 2010 and 2011 hurricane seasons. For single-month March initializations, GFDL-HyHuFS-v1.0 was reported to have a correlation coefficient of 0.49 and an RMS error 2.63 of hurricanes (observed standard deviation of hurricane frequency of 3.05 hurricanes) over 1982-2009. The inclusion of 2010 and 2011 changed the evaluated correlation coefficient of 0.55 and an RMS error of 2.57 hurricanes (observed standard deviation of hurricane frequency of 3.13 hurricanes) over 1982-2011. Including the two "real" forecast years gives small improvements to the retrospective skill of HyHuFS-v1.0 from March. The retrospective 1982-2012 correlation and RMS error from the new version of the forecast system, GFDL-HyHuFS-v3.1, are also nominally improved over those of the now defunct v1.0.

The forecasts using HyHuFS are fundamentally probabilistic, since the fundamental predicted value is the probability density function (PDF) for North Atlantic seasonal hurricane frequency each year. The above "skill measures" (correlation and RMS error) are not necessarily sufficient to assess probabilistic skill, and work is ongoing to extend the skill assessment to more probabilistic measures. **Figure 2** shows an example of a probabilistic skill measure, which compares the probability of exceedance in the predicted PDF for the number of hurricanes that were observed in that year (*verification exceedance probability*) with the sorted ranking (normalized by total number of forecasts) of the verification

exceedance probabilities. If the forecast PDFs were reliable, for a large enough sample, the points are expected to lie on the diagonal. Thus far, both for v1.0 and v3.1 of the HyHuFS system, the distribution of the observed value on the predicted PDF do not indicate any clear deviations from a uniform distribution, with as many points above and below the median (**Figure 2**). For both forecasts, extremes in the predicted distribution seem to verify at a rate similar to what one would expect from a uniform sampling of the forecast PDF.

Even though the forecast PDFs appear reliable, as one would expect, there are times that the verification occurs at the extremes of the forecast PDF (2012 was one of those years); but these extremes have occurred at a rate comparable to what one would expect from a Uniform sampling of the forecast PDFs. The verification using v3.1 for 2010 (orange symbol) and 2011 (green symbol) was well within the predicted 50% confidence intervals. However, the verification of the March v3.1 forecasts were an extreme outlier – with 2012 verifying at the 99% level.

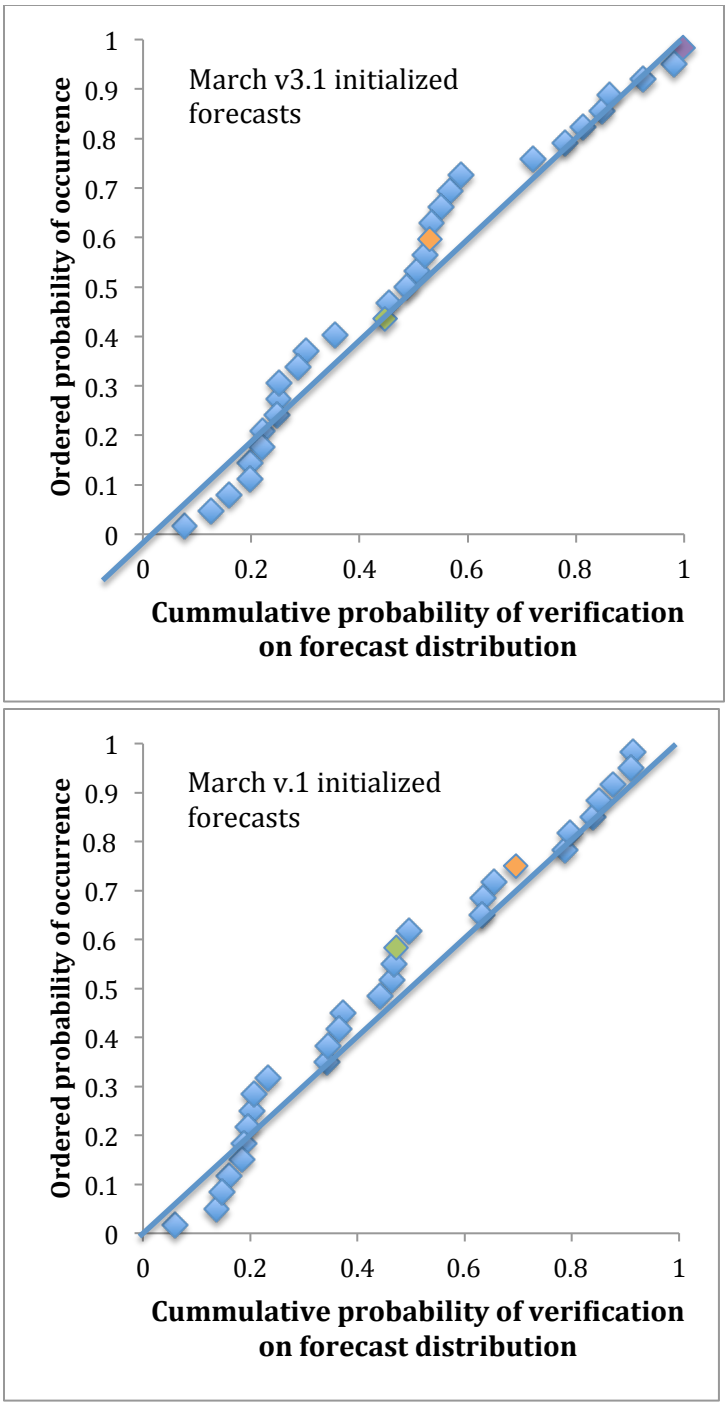

Figure 2: Graphical assessment of probabilistic skill of HyHuFS. Horizontal axis indicates the cumulative probability of the observed number of Atlantic hurricanes based on the predicted PDF (*verification exceedance probability*), vertical axis shows the order of the verification cumulative probability divided by the total number of points. For large sample size, a "perfectly" calibrated forecast PDFs are expected to result in all the points lying on the diagonal – indicating that the verification was a Uniform random draw from the PDFs that were predicted. Orange symbol highlights the March 2010 forecast of 2010, the green symbol the March 2011 forecast of 2011, and in the upper panel the violet symbol is the March 2012 forecast of 2012. Upper panel shows values from forecasts initialized with v3.1 GFDl-CM2.1 EnKF, lower panel shows values from forecasts initialized with (the now defunct) v1.0 GFDL-CM2.1 EnKF.

References:

Delworth T.R., and co-authors (2006): GFDL's CM2 Global Coupled Climate Models. Part I: Formulation and Simulation Characteristics. *J. Climate*, **19**(5), doi:10.1175/JCLI3629.1.

Vecchi, G.A., M. Zhao, H. Wang, G. Villarini, A. Rosati, A. Kumar, I.M. Held, R. Gudgel (2011): Statistical-Dynamical Predictions of Seasonal North Atlantic Hurricane Activity. *Mon. Wea. Rev*., doi: 10.1175/2010MWR3499.1

Villarini, G., G.A. Vecchi and J.A. Smith (2010): Modeling of the Dependence of Tropical Storm Counts in the North Atlantic Basin on Climate Indices. *Mon. Wea. Rev.*, **138**(7), 2681-2705, doi:10.1175/2010MWR3315.1

Zhang, S., M.J. Harrison, A. Rosati, and A.T. Wittenberg (2007): System Design and Evaluation of Coupled Ensemble Data Assimilation for Global Oceanic Climate Studies. *Mon. Wea. Rev.*, **135**(10), DOI:10.1175/MWR3466.1.

Zhao, M., I.M. Held and G.A. Vecchi (2010): Retrospective forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Mon. Wea. Rev.*, 138, 3858-3868.

Zhao, M., I.M. Held and G.A. Vecchi (2010): Retrospective forecasts of the hurricane season using a global atmospheric model assuming persistence of SST anomalies. *Mon. Wea. Rev.*, doi: 10.1175/2010MWR3366.1.

NA Hurricane Season	Ensemble-mean predicted hurricane count from March	Median predicted hurricane count from March	Probability of a "very inactive" season (≤3 hurricanes)	Probability of an "above average season" (>6 hurricanes)	Probability of a "very active" season (>10 hurricanes)	Observed North Atlantic Hurricane Counts
1982	2.85	3	67.62%	3.90%	0.08%	2
1983	4.62	4	37.59%	21.91%	2.54%	3
1984	7.28	7	14.79%	52.15%	19.37%	5
1985	3.33	3	58.50%	7.48%	0.32%	7
1986	3.72	3	50.56%	10.66%	0.50%	4
1987	6.03	6	23.73%	38.51%	9.98%	3
1988	4.86	5	33.14%	24.72%	2.63%	5
1989	5.15	5	28.21%	27.86%	3.02%	7
1990	4.72	4	40.60%	21.31%	6.11%	8
1991	5.27	5	30.07%	29.10%	5.73%	4
1992	5.40	5	28.56%	31.82%	5.71%	4
1993	6.33	6	24.85%	41.62%	13.84%	4
1994	5.03	5	33.71%	27.91%	4.39%	3
1995	8.03	7	10.35%	59.21%	21.98%	11
1996	5.45	5	26.05%	32.48%	4.72%	9
1997	6.19	6	16.78%	42.15%	7.33%	3
1998	11.88	11	3.55%	84.19%	57.07%	10
1999	7.55	7	11.36%	54.69%	19.62%	8
2000	7.29	7	15.79%	51.27%	19.45%	8
2001	5.94	6	21.42%	39.71%	6.98%	9
2002	6.81	7	15.95%	51.00%	13.29%	4
2003	6.58	6	19.04%	44.90%	13.95%	7
2004	8.56	8	13.14%	63.08%	32.05%	9
2005	7.47	7	10.36%	60.49%	16.81%	15
2006	5.44	5	31.19%	30.81%	8.32%	5
2007	8.20	8	6.20%	66.83%	23.28%	6
2008	7.74	7	8.20%	60.31%	19.34%	8
2009	4.76	5	34.18%	23.43%	2.22%	3
2010	11.48	11	2.92%	84.10%	54.75%	12
2011	7.62	7	14.75%	55.28%	21.87%	7
2012	5.35	5	27.38%	30.56%	4.79%	10
2013	8.37	8	6.13%	68.07%	25.46%	

Table 3: Summary of forecasts and observed hurricane counts in North Atlantic from 1982-2013 initialized on the 1st of March using HyHuFS-v3.1. First column lists the year, second column lists the expected value (ensemble-mean forecast), third column the forecast median, fourth through sixth columns the probability of a "very inactive", "above average" or "very active" hurricane season, respectively, and the seventh column lists the observed seasonal frequency of North Atlantic hurricanes.