Summary of the November 2011-initialized experimental forecasts for 2012 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 – 8-November-2011.
This 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.

1. Summary: The experimental forecast for the 2012 season with GFDL-HyHuFS initialized in November 2011 gives an expected value of 5.4 hurricanes. That is, HyHuFS is predicting that the 2012 season should be neutral to inactive relative to the 1982-2010 climatology, below the 1995-2010 mean, and comparable to the 1982-1994 average. These forecast values arise because the coupled GCM used in the system predicts that the tropical Atlantic should be slightly warmer than the long-term (1982-2011) average, but the remote tropics should also warm enough to offset 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). A series of forecasts are generated using a hybrid model, in which a statistical model of hurricane frequency is applied 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 from output from 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.

HyHuFS exhibits positive skill from as early as October of the previous year; the forecasts summarized here are based on the forecasts initialized through November 2011. 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").

3. Detailed forecast values: **Table 1** (below) presents the mean forecast values based on HyHuFS, along with the explicit 50% and 75% ranges. **Table 1** also indicates the values from HyHuFS 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 – verifications against 2011 tentatively assume that it will close at its current (as of November 8, 2011) count of 6 hurricanes (*i.e.* an average hurricane season). It is noteworthy that inclusion of 2010 and 2011 (the first "non-retrospective" forecasts for HyHuFS) nominally increases the correlation coefficient for the November-initialized forecasts of this system (from 0.51 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** illustrates the retrospective skill of the system and the experimental forecast for 2012.

Table 3 at the end of the document provides the ensemble-mean and median forecasts for each season (1982-2012) 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).

The forecast values and ranges for the 2011 hurricane season remained relatively stable since the November 2010-initialized forecasts, and called for an average to slightly below average North Atlantic hurricane season for 2011. The 2011 forecasts stood in contrast to the system's forecasts for an active 2010 hurricane season, which were – again – relatively consistent since the November 2009 initializations. However, consistency between forecasts need not be regarded as a source of confidence.

Forecast source	Mean count	Median count	50% range	75% range	1982-2011 Retrospective Correlation of Mean	Retrospective RMS error of Mean
November 2011 initialized forecast for 2012 Season	5.41	5	3-6	2-8	0.55	2.57
1982-2011 Average for November initialized forecasts	6.43	6	3-8	2-10		
1995-2011 Average for November initialized forecasts	7.32	7	4-10	3-11		

Table 1: Summary of the forecast for the 2012 North Atlantic hurricane season initialized November 2011, based on HyHuFS. Top row indicates the expected value, median and selected uncertainty ranges for the number of Atlantic hurricanes in the 2012 season, along with the retrospective correlation and RMS error of the system initialized in November. The last two rows summarize the system's statistics when initialized in November for the whole record and for the recent active era.

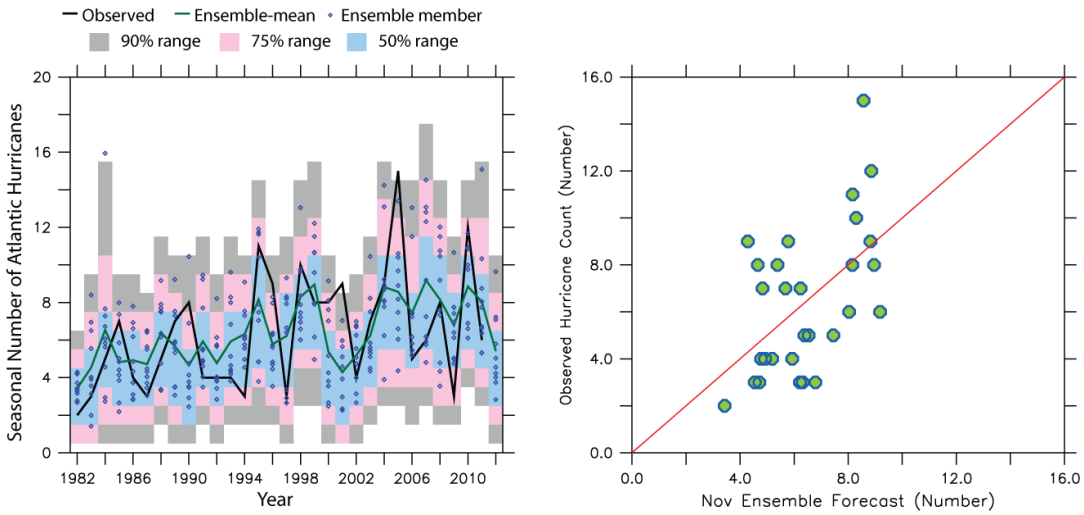

Figure 1: Retrospective and actual forecasts of North Atlantic hurricane frequency using an experimental hybrid seasonal hurricane forecast system (Vecchi et al. 2011). Left panel shows the observed seasonal North Atlantic hurricane frequency each season 1982-2012, along with forecasts based on HyHuFS using the GFDL-CM2.1 GCM, initialized in November of the previous year. Right panel shows a scatter plot between the ensemble-mean forecast (horizontal axis) and the observed (vertical axis) North Atlantic hurricane count for each year. In the left panel, 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.

In addition to the ranges, medians and means, HyHuFS forecasts the entire probability density function (PDF), which allows us to compute certain "exceedence 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 1982-2011 statistics in **Table 3**). HyHuFS indicates a heightened probability that 2012 will be a "very inactive" season, and a reduced probability of an "active" or "very active" season, relative both the long-term (1982-2011) and recent (1995-2011) 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)**
November 2011 initialized forecast for 2012 Season	31%	29%	6.6%
1982-2011 Average for November initialized forecasts	43%	22%	14%
1995-2011 Average for November initialized forecasts	53%	16%	19%

Table 2: Summary of the exceedance probabilities for an experimental forecast of the 2012 North Atlantic hurricane season initialized November 2011, 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 November 2011. The last two rows summarize the system's statistics when initialized in November 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 was finalized and its results submitted for publication, there have been two "real" forecasts to evaluate: the 2010 and 2011 hurricane seasons. For November initializations, GFDL-HyHuFS was reported to have a correlation coefficient of 0.51 and an RMS error of 2.57 hurricanes (observed standard deviation of hurricane frequency of 2.99 hurricanes) over 1982-2009. The inclusion of 2010 and 2011 (assuming that the 2011 season will close with the current 6 hurricanes) changes the evaluated correlation coefficient of 0.55 and an RMS error of 2.57 hurricanes (observed standard deviation of hurricane frequency of 3.07 hurricanes) over 1982-2011. Including the two "real" forecast years does not reduce the retrospective performance of the system.

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 be expected to lie on the diagonal. Thus far, the forecast PDFs do not show any gross departures from the diagonal, although there is a slight (non significant) tendency for more forecasts to verify below the median than above (18/30). As one would expect, there are times that the verification occurs at the extremes of the forecast PDF. For example, the forecasts of 2010 (orange symbol) indicated only a 15% chance of a value as large as was observed; but these extremes have occurred at a rate comparable to what one would expect from a Uniform sampling of the forecast PDFs. Meanwhile, the tentative verification for 2011 (green symbol) was well within the predicted 50% confidence intervals.

If this 2012 forecast verifies and the 2011 season closes at the current 6 hurricanes, the seven- year averaged hurricane count centered on 2009 will be below 7 hurricanes (6.5) for the first time since that centered on 1996, giving hints that the multi-decadal period of

enhanced Atlantic activity that has dominated the Atlantic since the mid-1990s may have begun to wane in the late 2000s.

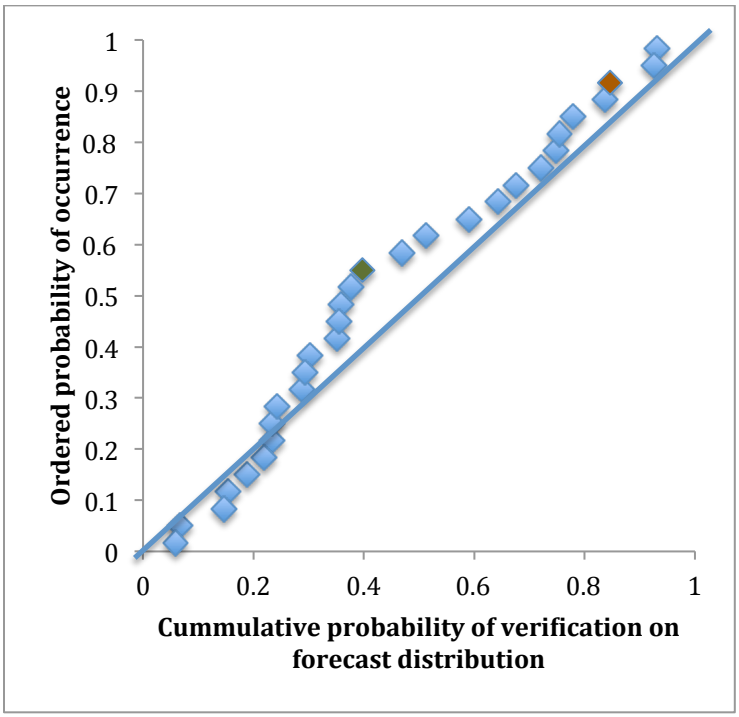

Figure 2: Graphical assessment of probabilistic skill of HyHuFS. Horizontal axis indicates the cumulative probability of the observed number of Atlantic hurricanes based on the PDF predicted from November of the previous year (*verification exceedance probability*), vertical axis shows the order of the 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 Nov. 2009 forecast of 2010, and the green symbol the Nov. 2010 forecast of 2011.

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 November of previous year	Median predicted hurricane count from November of previous year	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	3.43	3	55.6%	6.7%	0.1%	2
1983	4.54	4	42.5%	24.0%	4.2%	3
1984	6.54	6	24.2%	40.7%	13.9%	5
1985	4.83	4	35.9%	25.2%	3.8%	7
1986	4.92	5	35.1%	26.6%	4.0%	4
1987	4.71	4	34.3%	22.0%	2.0%	3
1988	6.38	6	19.3%	45.7%	10.4%	5
1989	5.68	5	25.9%	35.6%	7.1%	7
1990	4.65	4	41.9%	21.5%	5.9%	8
1991	5.93	5	23.4%	37.3%	9.1%	4
1992	4.77	4	35.6%	23.2%	3.3%	4
1993	5.93	5	23.4%	37.5%	9.0%	4
1994	6.32	6	15.3%	43.6%	7.9%	3
1995	8.16	8	12.5%	61.1%	27.9%	11
1996	5.78	5	21.7%	36.2%	6.0%	9
1997	6.21	6	24.3%	44.5%	11.3%	3
1998	8.29	8	5.8%	67.0%	23.7%	10
1999	8.96	9	5.7%	72.4%	32.3%	8
2000	5.38	5	33.0%	32.9%	7.8%	8
2001	4.29	4	43.5%	19.1%	2.0%	9
2002	5.19	5	29.3%	29.5%	3.7%	4
2003	6.24	6	18.5%	41.0%	9.6%	7
2004	8.82	8	7.8%	69.1%	30.6%	9
2005	8.57	8	7.7%	67.8%	28.6%	15
2006	7.45	7	11.9%	53.5%	18.8%	5
2007	9.18	8	8.6%	66.6%	35.9%	6
2008	8.15	8	13.5%	61.7%	28.4%	8
2009	6.79	6	13.4%	49.3%	11.9%	3
2010	8.85	9	4.6%	72.8%	30.3%	12
2011	8.03	7	8.9%	60.2%	22.1%	6*
2012	5.41	5	29.2%	31.3%	6.6%	

Table 3: Summary of forecasts and observed hurricane counts in North Atlantic from 1982-2012 initialized on the 1st of November of the previous year using HyHuFS (Vecchi *et al.* 2011). 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. *The "observed" value for 2011 is tentatively taken as 6 hurricanes, the number that occurred by the time of preparation of this report (8-November-2011), and could change as the current (2011) season continues or when the hurricane season is reassessed after its end.