Summary of the December 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-December-2011.
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.

1. Summary: As of the December 2011 initialization, HyHuFS is predicting that the frequency of Atlantic hurricanes in the 2012 season should be comparable to or slightly less than the 1982-2010 climatology, and below the recent 1995-2011 mean. HyHuFS indicates a reduced probability of an extremely active 2012 Atlantic hurricane season.

The experimental forecast for the 2012 season with GFDL-HyHuFS initialized in December 2011 gives an expected value of 6.0 hurricanes. That is, HyHuFS is predicting that the 2012 season should be comparable to the 1982-2010 climatology, but below the recent 1995-2011 mean. The "lagged ensemble" forecast, which combines the forecasts initialized in November and in December 2011, also indicates that North Atlantic hurricane frequency in 2012 should be comparable to or slightly below the 1982-2010 climatology, and inactive relative to the recent 1995-2011 active period. 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. 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 through December 2011. For December initializations we can use the GCM's ten-member ensemble initialized in December or a 20-member "lagged ensemble" from combining November and December initializations. In both probabilistic and deterministic measures of retrospective "skill" the November-December lagged ensemble outperforms the December-initialized forecast.

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 December 8, 2011) count of 7[1] hurricanes. The inclusion of 2010 and 2011 (the first “non-retrospective” forecasts for HyHuFS) leads to a small nominal reduction of the correlation coefficient for the December-initialized forecasts of this system (from 0.44 to 0.43) 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. The inclusion of 2010 and 2011 leads to a small nominal increase of the correlation of the November & December lagged ensemble forecasts (from 0.50 to 0.51) relative those in Vecchi et al. (2011). **Figure 1** illustrates the retrospective skill of the system and the experimental forecast for 2012.

Tables 3 & 4 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).

<table>
<tr><th>Forecast source</th><th>Mean count</th><th>Median count</th><th>50% range</th><th>75% range</th><th>1982-2011 Retrospective Correlation of Mean</th><th>Retrospective RMS error of Mean</th></tr>
<tr><td>December 2011 initialized forecast for 2012 Season</td><td>6.01</td><td>6</td><td>4-7</td><td>3-9</td><td rowspan="3">0.43</td><td rowspan="3">2.87</td></tr>
<tr><td>1982-2011 Average for December initialized forecasts</td><td>6.54</td><td>6</td><td>4-9</td><td>2-11</td></tr>
<tr><td>1995-2011 Average for December initialized forecasts</td><td>7.47</td><td>7</td><td>4-10</td><td>3-12</td></tr>
<tr><td>Nov. & Dec. 2011 lagged-ens. forecast for 2012 season</td><td>5.72</td><td>5</td><td>3-7</td><td>2-8</td><td rowspan="3">0.52</td><td rowspan="3">2.62</td></tr>
<tr><td>1982-2011 Average for Nov.&Dec. lagged-ens forecast</td><td>6.5</td><td>6</td><td>4-9</td><td>2-11</td></tr>
<tr><td>1995-2011 Average for Nov.&Dec. lagged-ens forecas</td><td>7.39</td><td>7</td><td>4-10</td><td>3-12</td></tr>
</table>

Table 1: Summary of the forecast for the 2012 North Atlantic hurricane season initialized December 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 December. The last two rows summarize the system’s statistics when initialized in December for the whole record and for the recent active era.

[1] The original December 2011 discussion incorrectly assumed a verification of 6 hurricanes, but was revised to 7 at 6pm 8 December 2011 to reflect the post-season upgrade of Nate. EB, thanks for the correction.

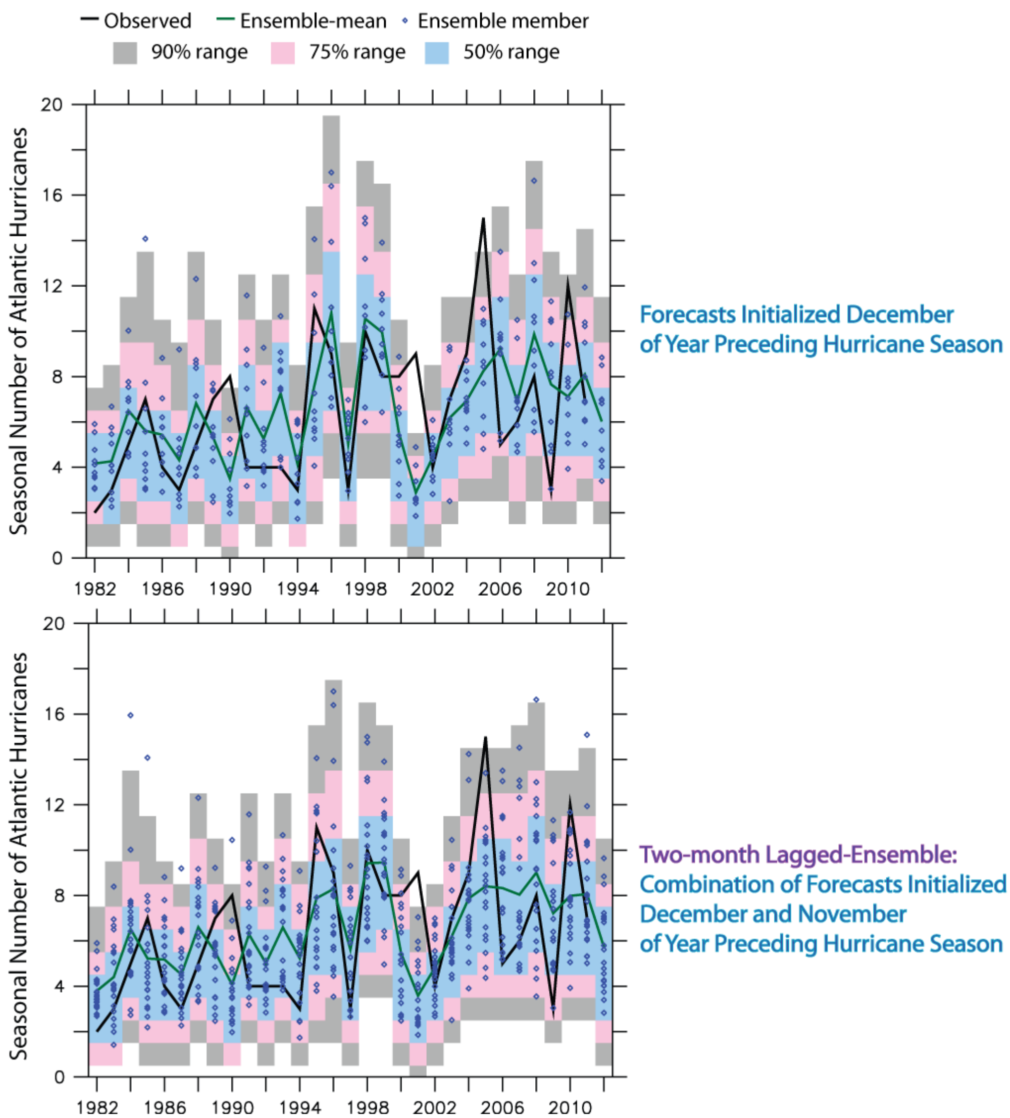

Figure 1: Retrospective and actual forecasts of North Atlantic hurricane frequency using an experimental hybrid seasonal hurricane forecast system (Vecchi et al. 2011). Both panels show the observed seasonal North Atlantic hurricane frequency each season 1982-2011, along with forecasts for the 1982-2011 seasons based on HyHuFS using the GFDL-CM2.1 GCM. The upper panel is for forecasts initialized in December of the previous year, the lower panel is for two-month "lagged ensemble" forecasts that uses information from forecasts initialized in November and December of the year preceding the hurricane season. In both panels, 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 "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 1982-2012 statistics in **Tables 3 & 4**). HyHuFS indicates a probability for a "very inactive" season that is similar to that over the entire 1982-2011 period, but that is heightened compared to the recent active era (1995-2011). Meanwhile, the probability of an "active" season is comparable to that over 1982-2011, but reduced relative to the 1995-2011 period. HyHuFS indicates a reduced probability of a "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)
December 2011 initialized forecast for 2012 Season	40%	22%	8.4%
1982-2011 Average for December initialized forecasts	44%	23%	15%
1995-2011 Average for December initialized forecasts	54%	17%	21%
Nov. & Dec. 2011 lagged-ens. forecast for 2012 season	36%	25%	7.5%
1982-2011 Average for Nov.&Dec. lagged-ens forecast	44%	23%	14%
1995-2011 Average for Nov.&Dec. lagged-ens forecas	53%	16%	20%

Table 2: Summary of the exceedance probabilities for an experimental forecast of the 2012 North Atlantic hurricane season initialized December 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 December 2011. The last two rows summarize the system's statistics when initialized in December 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 single-month December initializations, GFDL-HyHuFS was reported to have a correlation coefficient of 0.44 and an RMS error 2.82 of 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 7 hurricanes[1]) changes the evaluated correlation coefficient of 0.43 and an RMS error of 2.87 hurricanes (observed standard deviation of hurricane frequency of 3.07 hurricanes) over 1982-2011. For two-month

November & December lagged-ensemble forecasts, GFDL-HyHuFS has a correlation coefficient of 0.50 and an RMS error 2.59 of 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 7 hurricanes[1]) changes the evaluated correlation coefficient of 0.52 and an RMS error of 2.62 hurricanes (observed standard deviation of hurricane frequency of 3.07 hurricanes) over 1982-2011. Including the two "real" forecast years gives small changes to the retrospective skill of HyHuFS from December, with some scores improving slightly and others decreasing slightly. The retrospective correlations and RMS errors of the November & December lagged ensemble indicate nominally higher deterministic skill than that of the single-month December initialized forecasts.

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, the forecast PDFs are indicating a tendency for the forecast system initialized in December to verify at the low end of the distribution, with more forecasts verifying below the median than above (20/30). For the November & December lagged-ensemble there is also a weak tendency for more forecasts to verify below the median than above (18/30) (**Figure 2, lower panel**). 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.

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. For the 1982-2011 period, it appears that the November-December lagged ensemble has a verification exceedance probability that is closer to a Uniform distribution than that of the single-month December-initialized forecasts.

If this 2012 forecast verifies and the 2011 season closes at the current 7 hurricanes[1], the seven- year averaged hurricane count centered on 2009 will be below 7 hurricanes 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.

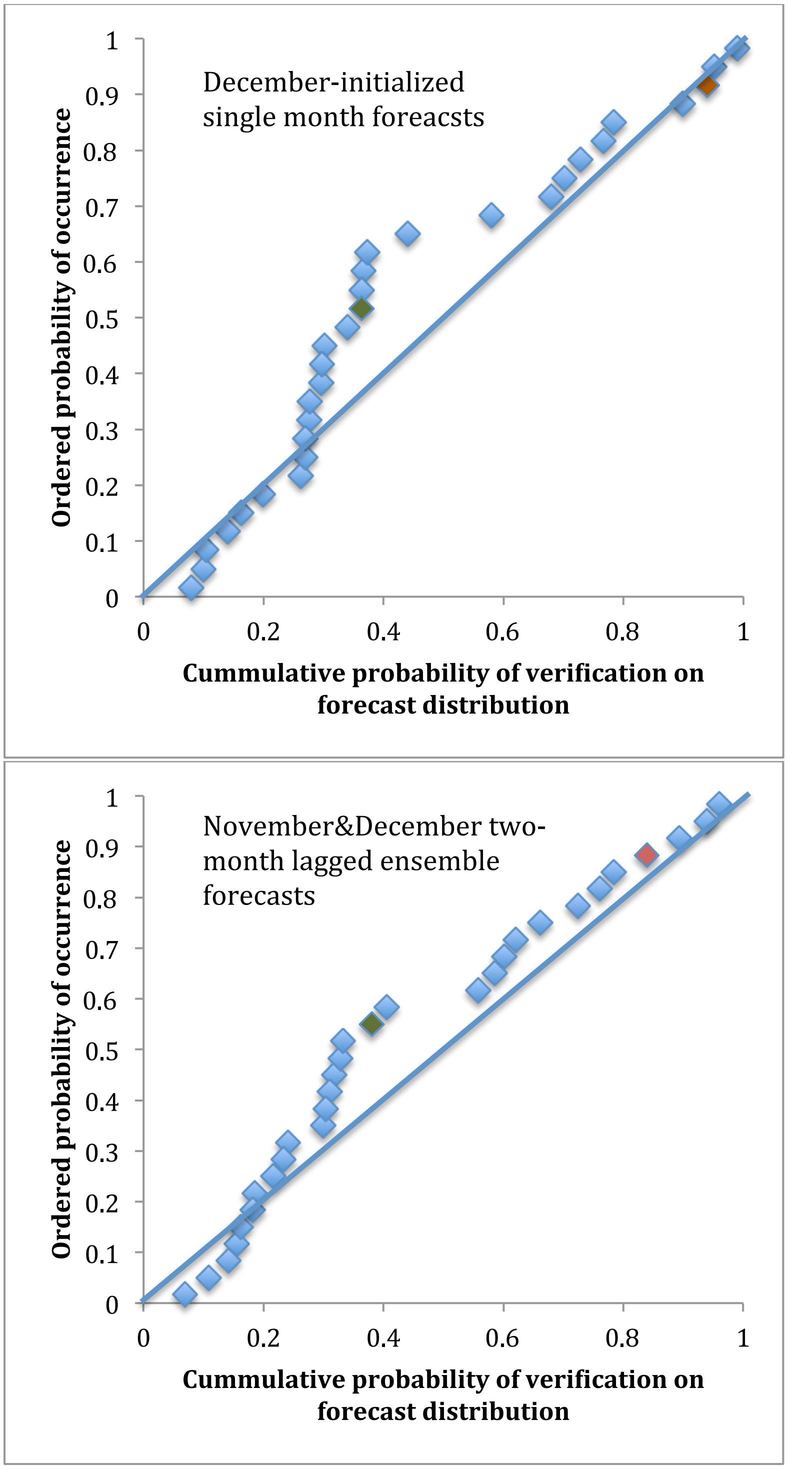

Figure 2: Graphical assessment of probabilistic skill of HyHuFS. Horizontal axis indicates the cumulative probability of the observed number of Atlantic hurricanes based on the predicted PDF (*verification exceedance probability*), vertical axis shows the order of the verification cumulative probability 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 Nov. 2009 forecast of 2010, and the green symbol the Nov. 2010 forecast of 2011. Upper panel shows values from single-month December-initialized forecasts, lower panel shows values from two-month November&December lagged ensemble.

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 December of previous year	Median predicted hurricane count from December 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	4.16	4	42.4%	14.7%	0.9%	2
1983	4.26	4	42.6%	17.9%	1.6%	3
1984	6.50	6	15.9%	45.7%	10.3%	5
1985	5.61	5	33.6%	29.9%	11.0%	7
1986	5.43	5	26.9%	31.8%	5.3%	4
1987	4.32	4	44.7%	18.3%	3.6%	3
1988	6.82	6	18.6%	47.6%	15.8%	5
1989	5.33	5	29.2%	32.0%	5.1%	7
1990	3.49	3	56.8%	10.7%	0.8%	8
1991	6.65	6	19.9%	46.4%	14.6%	4
1992	5.25	5	30.1%	28.7%	5.7%	4
1993	7.26	7	14.1%	55.5%	17.6%	4
1994	4.14	4	45.4%	18.0%	1.6%	3
1995	7.61	7	13.6%	53.3%	21.6%	11
1996	10.80	10	2.8%	80.9%	46.1%	9
1997	5.02	5	30.3%	26.3%	2.8%	3
1998	10.55	10	3.2%	81.5%	47.1%	10
1999	9.96	10	2.4%	82.1%	41.3%	8
2000	5.51	5	27.1%	33.4%	6.1%	8
2001	2.88	3	67.6%	4.8%	0.2%	9
2002	4.44	4	37.3%	17.9%	1.1%	4
2003	6.17	6	20.2%	42.0%	9.3%	7
2004	6.92	7	10.3%	52.9%	10.9%	9
2005	8.25	8	7.8%	66.3%	25.4%	15
2006	9.20	9	5.5%	74.5%	34.8%	5
2007	6.90	7	12.9%	50.3%	13.1%	6
2008	9.86	9	5.2%	74.7%	40.2%	8
2009	7.65	7	14.9%	58.0%	23.2%	3
2010	7.12	7	12.9%	53.6%	15.3%	12
2011	8.09	8	8.1%	63.6%	23.8%	7*
2012	6.02	6	21.7%	39.9%	8.4%	

Table 3: Summary of forecasts and observed hurricane counts in North Atlantic from 1982-2012 initialized on the 1st of December 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 7 hurricanes[1], the number that occurred by the time of preparation of this report (8-December-2011), and could change as the current (2011) season continues or when the hurricane season is reassessed after its end.

NA Hurricane Season	Ensemble-mean predicted hurricane count from December of previous year	Median predicted hurricane count from December 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	4.12	4	43.8%	14.7%	1.1%	2
1983	4.74	4	38.1%	25.2%	3.9%	3
1984	6.52	6	20.0%	43.2%	12.1%	5
1985	5.22	5	34.8%	27.5%	7.4%	7
1986	5.18	5	31.0%	29.2%	4.6%	4
1987	4.52	5	39.5%	20.1%	2.8%	3
1988	6.60	6	18.9%	46.6%	13.1%	5
1989	5.51	5	27.6%	33.8%	6.1%	7
1990	4.07	4	49.4%	16.1%	3.3%	8
1991	6.29	6	21.7%	41.8%	11.9%	4
1992	5.01	6	32.9%	25.9%	4.5%	4
1993	6.59	6	18.7%	46.5%	13.3%	4
1994	5.23	5	30.4%	30.8%	4.8%	3
1995	7.88	7	13.6%	53.3%	21.6%	11
1996	8.29	7	12.2%	58.6%	26.0%	9
1997	5.62	5	27.3%	35.4%	7.1%	3
1998	9.42	7	4.5%	74.2%	35.4%	10
1999	9.46	9	4.1%	77.3%	36.8%	8
2000	5.45	5	30.1%	33.2%	6.9%	8
2001	3.58	3	55.5%	11.9%	1.1%	9
2002	4.82	5	33.3%	23.7%	2.4%	4
2003	6.21	6	19.3%	41.5%	9.4%	7
2004	7.87	7	9.1%	61.0%	20.8%	9
2005	8.41	8	7.7%	67.0%	27.0%	15
2006	8.33	8	8.7%	64.0%	26.8%	5
2007	8.04	7	10.7%	58.5%	24.5%	6
2008	9.01	9	9.4%	68.2%	34.3%	8
2009	7.22	7	14.1%	53.7%	17.6%	3
2010	7.99	8	8.8%	63.2%	22.8%	12
2011	8.06	8	8.5%	61.9%	23.0%	7*
2012	5.72	5	25.4%	35.6%	7.5%	

Table 4: Summary of forecasts and observed hurricane counts in North Atlantic from the lagged ensemble of 1982-2012 forecasts initialized on the 1st of November and December 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 7 hurricanes[1], the number that occurred by the time of preparation of this report (8-December-2011), and could change as the current (2011) season continues or when the hurricane season is reassessed after its end.