Summary of the March and April 2014-initialized experimental forecasts for 2014 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-April-2014.
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: Starting 2012 these experimental forecasts come 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 April 2014 initialization, HyHuFS is predicting that the frequency of Atlantic hurricanes in the 2014 season should be less active than the 1982-2012 climatology, and the recent 1995-2012 mean. In its April predictions HyHuFS indicates a reduced probability of an extremely active 2014 Atlantic hurricane season, a reduced probability of above-average hurricane frequency and an enhanced probability for an extremely inactive hurricane season.

The experimental forecast for the 2014 season with GFDL-HyHuFS initialized in April 2014 gives an expected value of 5.1 hurricanes. These forecast values arise because the coupled GCM used in the system predicts that the tropical Atlantic should be somewhat warmer than the long-term (1982-2005) average (by about 0.25°C), the predicted warm anomalies in the remote tropics (0.28°C) – enhanced in the April 2014 predictions in part due to a predicted El Niño event - offset the impact of a warm Atlantic.

The April 2014 initialized predictions are different from the March 2014 initialized forecasts, which had indicated a near-normal hurricane frequency. This tendency in the April 2014 initialized predictions for a reduced number of hurricanes reflects a stronger tendency for the dynamical model to predict an El Niño in its April 2014 predictions.

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").

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 (1982-2013) and the recent "active" period (1995-2013), along with retrospective correlations and RMS values, to guide interpretation. These retrospective skill measures are over the 1982-2013 period. The inclusion of 2010-2013 leads to essentially no change in the correlation coefficient for the March-initialized forecasts of this system (0.49 to 0.51) relative to the correlation skill reported in Vecchi *et al.* (2011) using v1.0.

The bottom rows of **Table 1** show the retrospective skill of Version 3.1 of HyHuFS initialized 1-April (Vecchi et al. 2011). The retrospective RMS error initialized 1-April is considerably larger than that initialized 1-March, a difference that largely reflects the extremely active season predicted for 2010 from 1-April-2010 (**Figure 1**).

Forecast source	**Mean count**	**Median count**	**50% range**	**75% range**	**1982-2013 Retrospective Correlation of Mean**	**1982-2013 Retrospective RMS error of Mean**
V3.1 March 2014 initialized forecast for 2014 Season	8.38	8	6-11	5-12		
V3.1 1982-2013 Average for March initialized forecasts	6.44	6	4-9	2-11	0.51	2.74
V3.1 1995-2013 Average for March initialized forecasts	7.36	7	5-10	3-13		
V3.1 April 2014 initialized forecast for 2014 Season	5.12	5	3-7	2-8		
V3.1 1982-2013 Average for April initialized forecasts	6.41	6	4-8	3-11	0.47	3.18
V3.1 1995-2013 Average for March initialized forecasts	7.41	7	4-10	3-12		

Table 1: Summary of the forecast for the 2014 North Atlantic hurricane season initialized March and April 2014, 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 2014 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 bottom three rows are like the top three for the April-initialized forecasts.

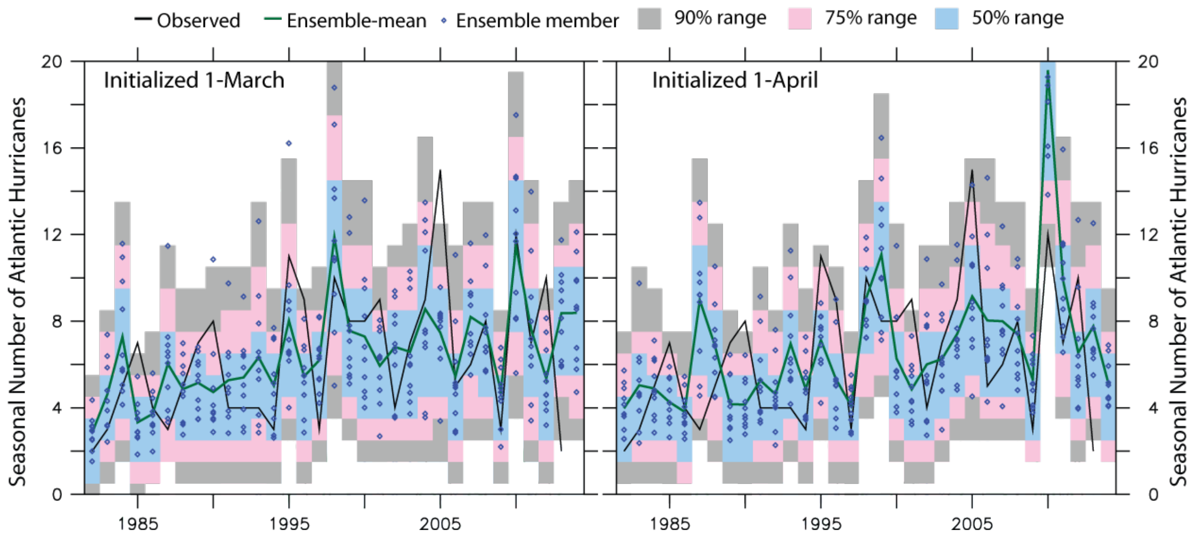

Figure 1: Retrospective and actual forecasts of North Atlantic hurricane frequency using an experimental hybrid seasonal hurricane forecast system (Vecchi et al. 2011). Shown is the observed seasonal North Atlantic hurricane frequency each season 1982-2013, along with forecasts for the 1982-2014 seasons initialized in March (left panel) and April (right panel) based on HyHuFS using the GFDL-CM2.1 GCM. 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 (1982-2014) 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 1982-2014 statistics in **Table 3**). The April 2014 experimental predictions from HyHuFS indicate a reduced probability for both a "very active" and "active" season relative to both the entire 1982-2013 period and to the recent active era (1995-2013). Meanwhile, April 2014 initialized experimental predictions HyHuFS indicates a increased probability of a "very inactive" season relative to both the long-term (1982-2013) and the recent (1995-2013) 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 April 2014 initialized forecast for 2014 Season	26%	28%	2.6%
V3.1 1982-2012 Average for April initialized forecasts	44%	22%	16%
V3.1 1995-2012 Average for April initialized forecasts	54%	15%	22%

Table 2: Summary of the exceedance probabilities for an experimental forecast of the 2012 North Atlantic hurricane season initialized March 2014, 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 2014 season initialized in March 2014. 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.

Predictions with HyHuFS for North Atlantic hurricane frequency over 2013, like those from practically all seasonal prediction systems (Vecchi and Villarini 2014), indicated normal to above-normal activity. We now know that 2013 was one of the least active hurricane seasons in the Atlantic, measured by hurricane frequency, with two hurricanes. The causes of the low hurricane frequency in 2013 are currently under investigation, in order to see the extent to which it reflects a deficiency in current seasonal prediction systems, and the extent to which it reflects the inherent uncertainty in seasonal prediction. From Vecchi and Villarini (2014):

> *In disentangling the causes of the low hurricane activity of 2013, we must ask ourselves whether our prediction systems neglected something foreseeable, and then account for this in future predictions. But the predictability of the climate system has limits, and it may be that the causes of the inactivity in 2013 were inherently unpredictable. Although extreme failure is improbable in any given year, over many years its likelihood at some (unknowable) point can become substantial even in the best possible prediction 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 expected to lie on the diagonal. Thus far the distribution of the observed value on the predicted PDF do not indicate any clear deviations from a uniform distribution (**Figure 2**). 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. For March-initialized forecasts both 2012, violet symbol, and 2013, red symbol, verified in the extremes of the distribution; for April-initialized forecasts both 2010 and 2013 verified in the extrmees of the distribution; but these extremes have occurred at a rate comparable to what one would expect from a Uniform sampling of the forecast PDFs. **(Figure 2)**.

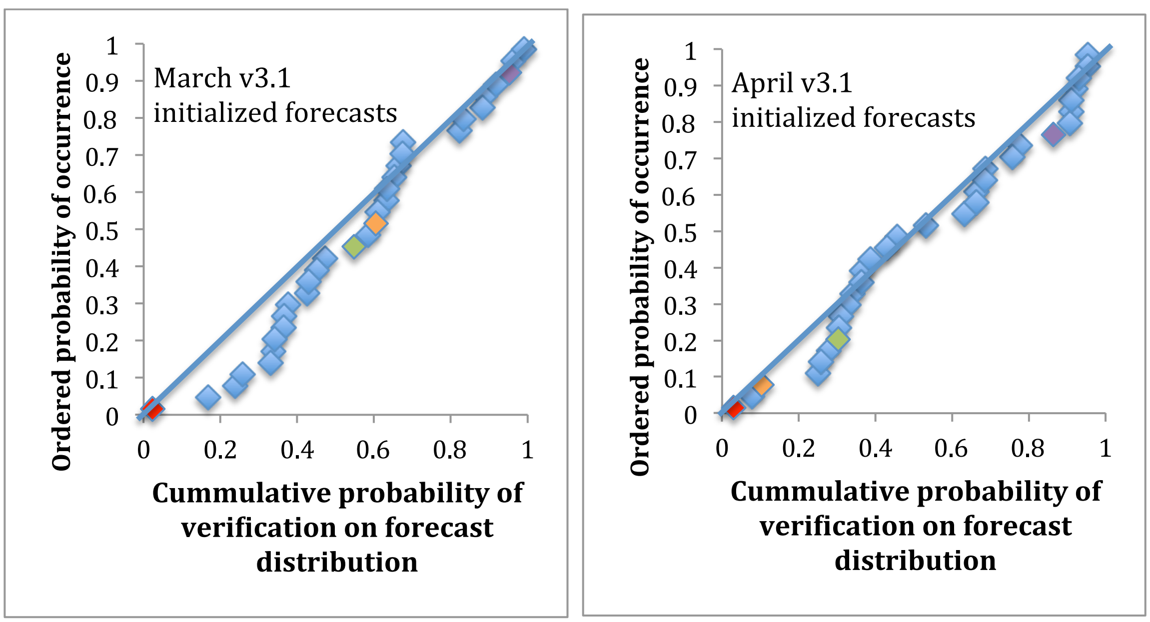

Figure 2: Graphical assessment of probabilistic skill of HyHuFS. Horizontal axis indicates the cumulative probability of the observed number of Atlantic hurricanes based on the predicted PDF (*verification exceedance probability*), vertical axis shows the order of the verification cumulative probability 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 2010 forecast, the green symbol the 2011 forecast, the violet symbol is the 2012 forecast, and the red symbol is the 2013 forecast. Left panel shows the March-initialized forecasts, right panel shows the April-initialized forecasts.

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

Vecchi, G.A. and G. Villarini (2014): Enhancing Seasonal Hurricane Predictions. *Science*, doi:10.1126/science.1247759

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 April	Median predicted hurricane count from April	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	3.98	4	46.05%	13.43%	0.74%	2
1983	5.04	5	34.03%	26.77%	5.36%	3
1984	4.89	5	32.42%	24.63%	2.64%	5
1985	4.26	4	40.95%	16.20%	1.03%	7
1986	3.81	4	48.70%	11.12%	0.51%	4
1987	8.97	9	7.86%	70.00%	33.62%	3
1988	6.87	7	13.39%	50.54%	12.85%	5
1989	4.16	4	43.85%	15.96%	1.29%	7
1990	4.15	4	43.48%	15.18%	1.12%	8
1991	5.27	5	30.69%	29.27%	5.94%	4
1992	4.62	4	37.45%	21.79%	2.67%	4
1993	6.96	6	14.24%	49.94%	14.98%	4
1994	4.86	5	32.94%	23.97%	2.69%	3
1995	7.13	7	10.29%	55.85%	13.06%	11
1996	5.32	5	29.11%	29.74%	5.61%	9
1997	4.10	4	43.02%	13.92%	0.67%	3
1998	9.07	9	3.10%	76.65%	31.32%	10
1999	11.08	11	2.01%	84.96%	51.15%	8
2000	6.28	6	20.59%	40.75%	11.30%	8
2001	4.85	5	32.50%	23.63%	2.46%	9
2002	6.01	5	26.57%	39.31%	11.26%	4
2003	6.25	6	19.51%	42.75%	9.77%	7
2004	7.41	7	10.72%	55.80%	17.41%	9
2005	9.14	9	6.11%	72.48%	33.91%	15
2006	8.03	7	10.81%	58.10%	24.04%	5
2007	8.00	8	8.67%	63.66%	22.41%	6
2008	7.40	7	11.00%	56.05%	17.65%	8
2009	5.29	5	25.80%	29.31%	3.34%	3
2010	19.59	19	0.01%	99.73%	95.72%	12
2011	9.89	9	2.98%	79.32%	39.37%	7
2012	6.54	6	19.43%	42.62%	13.58%	10
2013	7.76	7	7.78%	61.01%	18.98%	2
2014	5.12	5	27.37%	26.45%	2.62%	

Table 3: Summary of forecasts and observed hurricane counts in North Atlantic from 1982-2013 initialized on the 1st of April 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.