Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
40th NOAA Annual Climate Diagnostics and Prediction Workshop
Denver, CO, 26-29 October 2015

Intraseasonal Tropical Storm Prediction in the NCEP CFSv2 45-Day Forecasts

Lindsey N. Long[1,2], Jae-Kyung E. Schemm[1], Stephen Baxter[1]
[1]*Climate Prediction Center, NOAA/NWS/NCEP, College Park, Maryland*
[2]*Innovim LLC, Greenbelt, Maryland*

1. Introduction and motivation

The majority of tropical storm (TS) forecasts focus on either the short-term (1-5 days) or the seasonal aspect by ocean basin. Although studies have shown predictability at the intraseasonal timescale using mechanisms such as the Madden Julian Oscillation (MJO) (Maloney and Hartmann, 2000; Klotzbach, 2010), there are few products which attempt to utilize these signals to produce operational products. With the availability of the Climate Forecast System Version 2 (CFSv2) 45-Day forecasts, the ability to forecast at the intraseasonal timescale can be more thoroughly examined. Because the CFS is a fully-coupled climate system, it is well equipped to handle forecasts out to weeks 1 to 4.

The Climate Prediction Center (CPC) currently issues the Global Tropics Hazards and Benefits (GTHB) Outlook (http://www.cpc.ncep.noaa.gov/products/precip/CWlink/ghazards/), which produces forecasts for tropical precipitation and TS formation globally for weeks 1 and 2. Shaded regions indicate either high or moderate confidence of TS formation and weekly total rainfall in the upper/lower third of the historical range. This product is released each Tuesday and contains both a graphical representation of this information and a detailed discussion. During the active TS season for the Northern Hemisphere (June 1st – November 30th), the outlook is also updated on Friday for a limited region (120°E-0° and 0°-40°N) which encompasses the Atlantic and both the Eastern and Western North Pacific basins.

The shaded areas in the GTHB Outlook, which represent areas with favorable conditions for tropical cyclogenesis, are determined subjectively based on a few forecasts tools. With the creation of a year-round forecast of TSs for Weeks 1-4 by CFSv2, we hope to provide these forecasters with an objective tool for the Outlook. It may also assist in the possible expansion to include Week 3 in the Outlook.

2. Data and tracking methods

a) CFSv2 45-day forecasts

The CFSv2 is a fully coupled atmosphere-ocean-land model run operationally at the National Centers for Environmental Prediction (NCEP) since April 2011 (Saha *et al.* 2014). The CFSv2 45-Day forecasts are currently run four times daily at the 00Z, 06Z, 12Z, and 18Z cycles with four ensemble members each. The sixteen members created daily have output saved every six hours. A 14-year hindcast has also been produced for 1999-2012 with only one member for each of the four initializations. Because of the sparse number of daily ensemble members in the hindcasts, the five days prior to the forecasted day are used to create a more robust 20-member ensemble. For verification, observations from the National Hurricane Center (NHC) and the Joint Typhoon Warning Center's (JTWC) Best-Track datasets are utilized. Since the 2015 best-track data are not yet available, the advisories from operational centers are used for corresponding basins.

b) TS tracking and filtering

The detection and tracking method used in this study is based on the algorithm created by Camargo and Zebiak (2002). With this method, a point must meet seven criteria in order to be considered a TS. Many of the thresholds used in the criteria are basin and model-dependent. The seven basins used in this study are the Atlantic (ATL), eastern North Pacific (ENP), western North Pacific (WNP), North Indian (NI), South Indian

---

Correspondence to: Lindsey N. Long, 5830 University Research Court, Climate Prediction Center, NOAA/NWS/NCEP, College Park, Maryland; E-mail: lindsey.long@noaa.gov

(SI), Australian (AUS), and South Pacific (SP). Once a point is detected as a possible TS, it is tracked forward and backward in time following a vorticity maximum that must exceed 3.5×10^{-5} s^{-1}. Tracks are compared and duplicate tracks are removed.

Once TSs have been detected and tracked for each member, forecasts of storm counts and storm tracks by basin are created for each weekly period out to week 4. As discussed above, these forecasts are based on a 16-member ensemble for the operational forecasts and a 20-member ensemble for the hindcast runs. With the storm activity analysis on the 14-year hindcast data, the CFSv2 storm activity climatology is established and utilized to remove the storm activity bias from future forecast runs.

While examining the storm analysis results from the hindcast runs, it was discovered that the CFSv2 produced too many storms. These erroneous storms, or False Alarms (FA), are storms that do not occur in observations. In order to filter these FA storms, the storm tracks are converted to storm track density values, meaning each track point is converted into a grid point. Every time a storm track touches a grid box, the box value increases by one. This process is continued for each ensemble member, and the grid boxes are divided by the total number of ensembles, creating a storm track density distribution. Figure 1 illustrates how FA's are removed from forecast storm activity. Figure 1a shows an example of a storm track density distribution from a forecast ensemble suite. The FAs are then filtered by removing the weekly storm track climatology (Figure 1b), the weekly FA climatology (Figure 1c), and finally, using a 0.5 threshold on the remaining points. This threshold assures that at least one member still contains a storm. Any remaining grid points are considered likely areas for TS activity (Figure 1d). The forecast shows a high confidence for storms in the WNP and ENP basins. The observations (Figure 1e) show that one storm in the WNP and two storms in the ENP verify, although the WNP storm is closer to the coast than forecasted.

3. Hindcast storm activity evaluations

Using the 20-member ensemble, the average numbers of storms present for Weeks 1-4 are calculated for each basin for the 14-year hindcasts. The anomaly is then computed to remove the seasonal variability. Figure 2 shows the anomaly correlations for Week 1 through Week 4 with the average correlations represented by a straight, solid line. The basins with the highest Week 1 scores are the ENP, WNP, SI and SP basins with average values between 0.49 and 0.51. The average correlation for the ATL basin (0.33) is brought down by two bad years (2002 and 2003). After looking closer at the

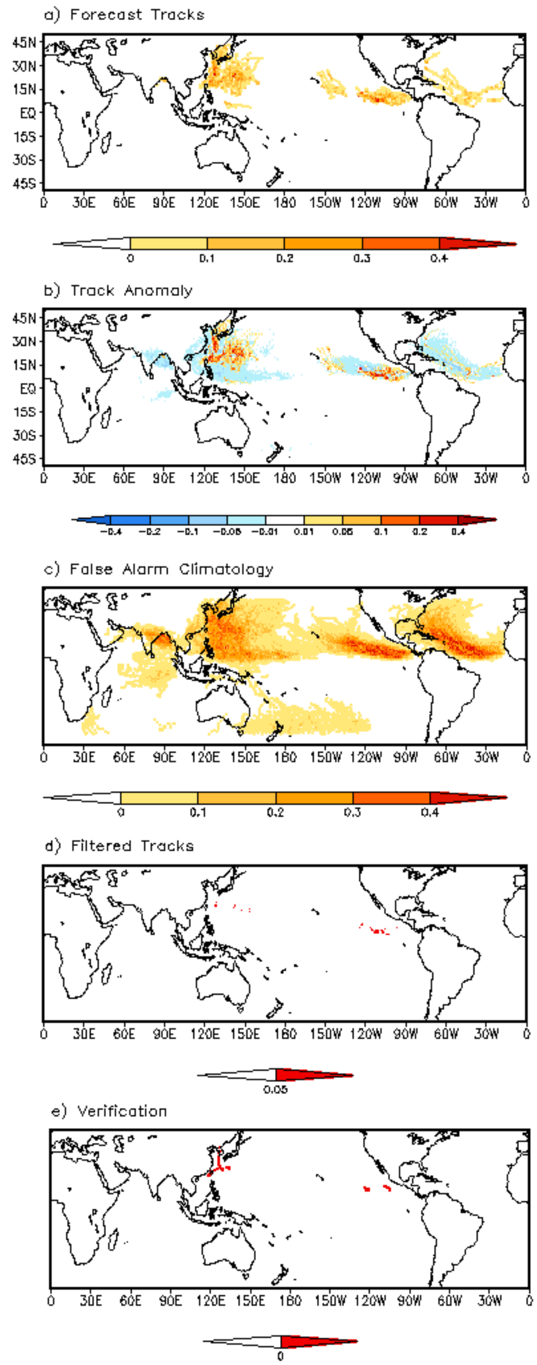

Fig. 1 An example of the storm track filtering technique by step for August 1, 1999: a) original storm track density distribution, b) storm tracks with weekly climatology removed, c) false alarm weekly climatology for July 30-Aug 5, d) final filtered tracks with weekly FA climatology removed, and e) observed storm track for verification.

storm counts in the basin's subregions and also wind shear anomalies, it was found that the forecasts in the subtropical, Northern Atlantic accounted for these low scores. As expected, skill drops with lead time, but there is still skill evident in most basins for Weeks 2-4 with scores for Week 4 remaining above 0.2.

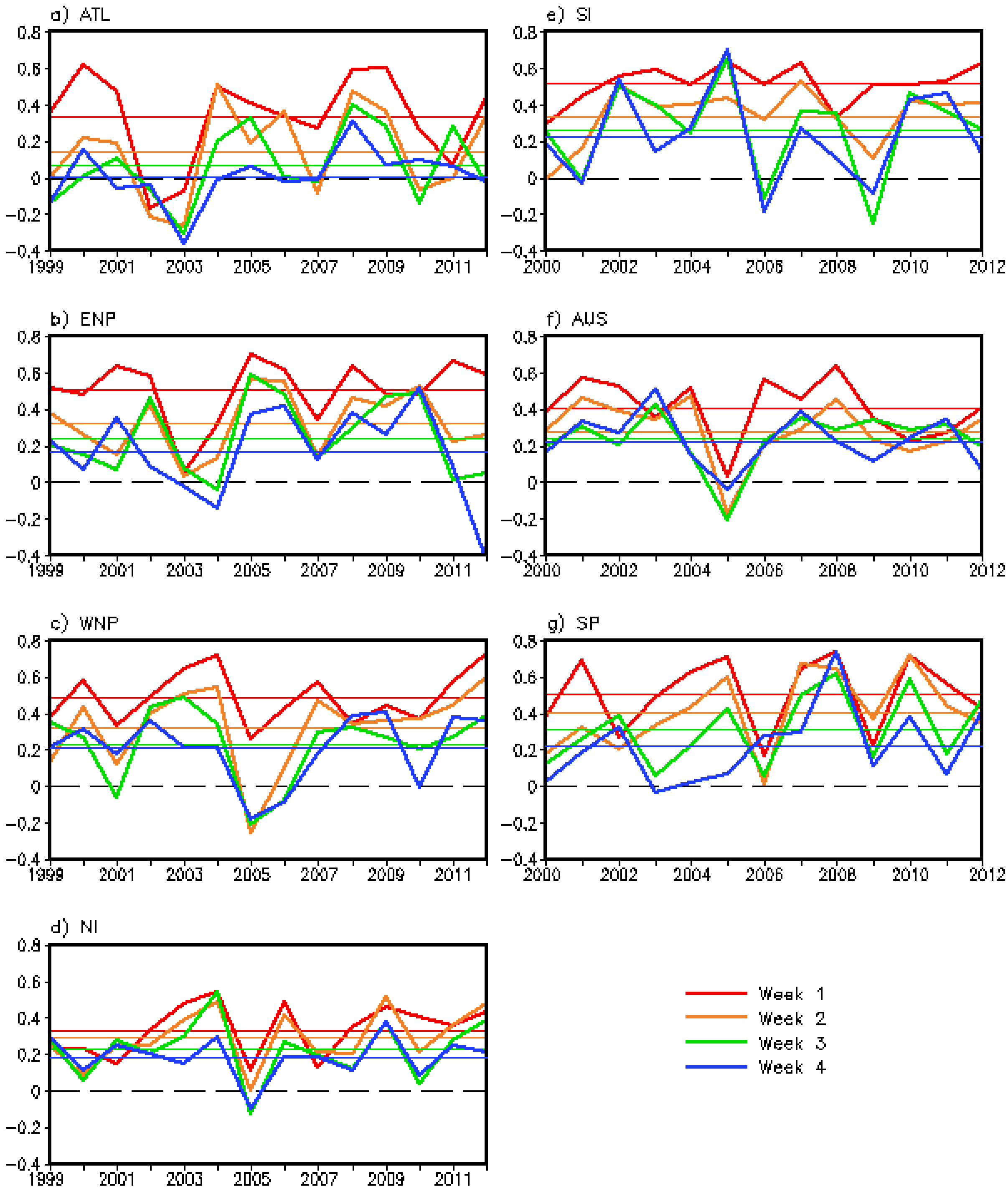

Fig. 2 Tropical storm count anomaly correlations by week for the a) Atlantic, b) Eastern North Pacific, c) Western North Pacific, and d) North Indian basins for 1999-2012, and the e) South Indian, f) Australian, and g) South Pacific basins for 2000-2012. Because the 1999 SH season begins in 1998, it is not included. The average correlation is shown using a straight, solid line.

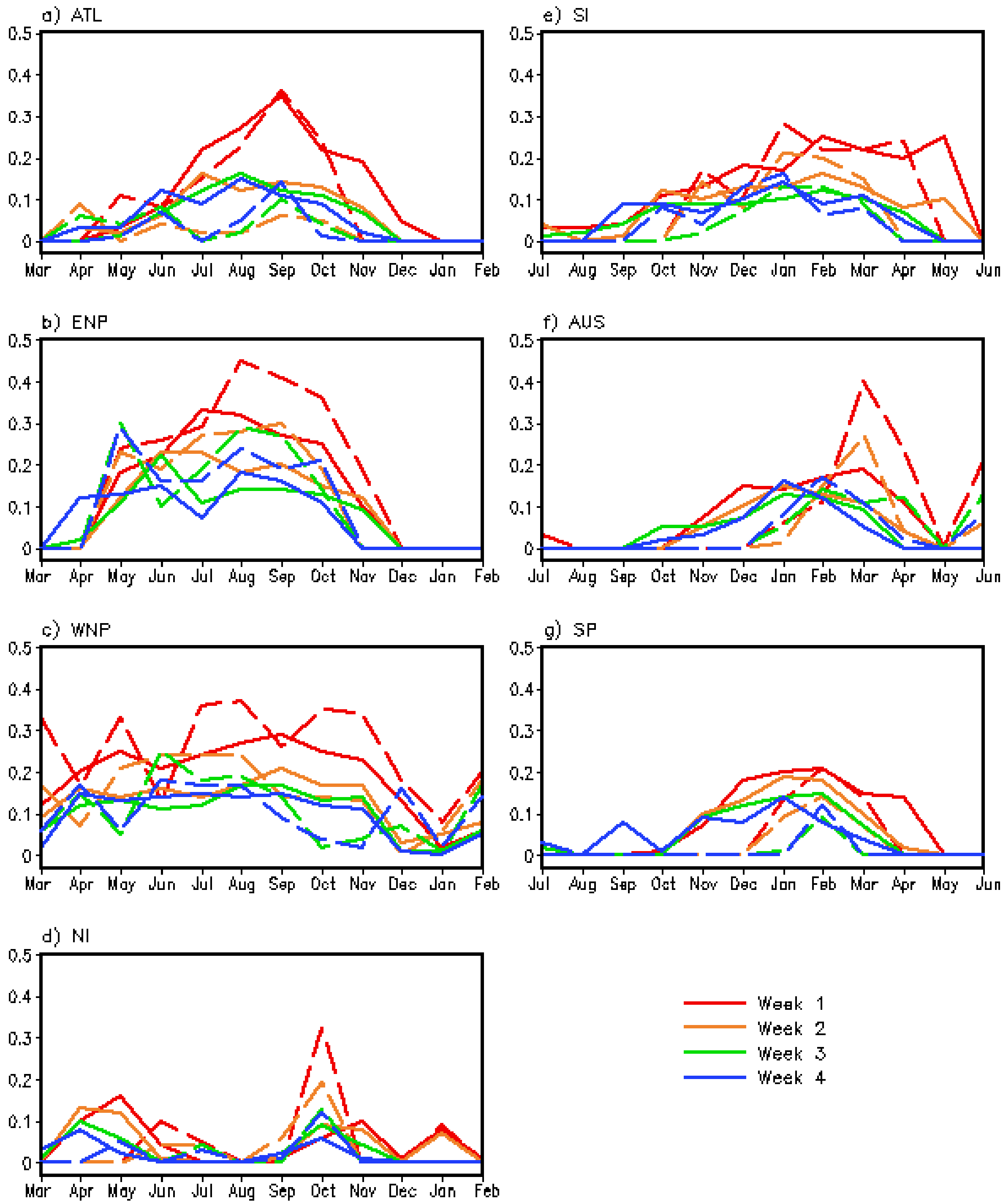

Fig. 3 Heidke skill scores for filtered storm tracks by week for the a) Atlantic, b) Eastern North Pacific, c) Western North Pacific, and d) North Indian basins for 1999-2012 (solid lines), and the e) South Indian, f) Australian, and g) South Pacific basins for 2000-2012 (solid lines). Dotted lines are for the real-time evaluation from 2014-2015.

After performing the filtering described above, the Heidke Skill Scores (HSS) are computed for the storm tracks in each basin. Because of the nature of the HSS, months with no storms will have a score of zero. No credit is given for a correct forecast of zero storms when there is also a verification of zero storms. Skill is achieved through the correct forecast of a storm track (Hits). Therefore, as expected, skill scores increase with increased seasonal activity. Figure 3 shows the HSS for each weekly lead by basin (solid lines). The highest Week 1 scores are present in the ATL, ENP and WNP basins, with scores between 0.25 and 0.35 during the most active part of the season. The SI and AUS basins show an increase in scores during the latter half of the season instead of during the peak in the seasonal cycle for these basins. This indicates the model is either missing the observed storms or producing too many FAs during these months. Skill scores for Week 2 decrease, but they then remain steady for Weeks 3 and 4.

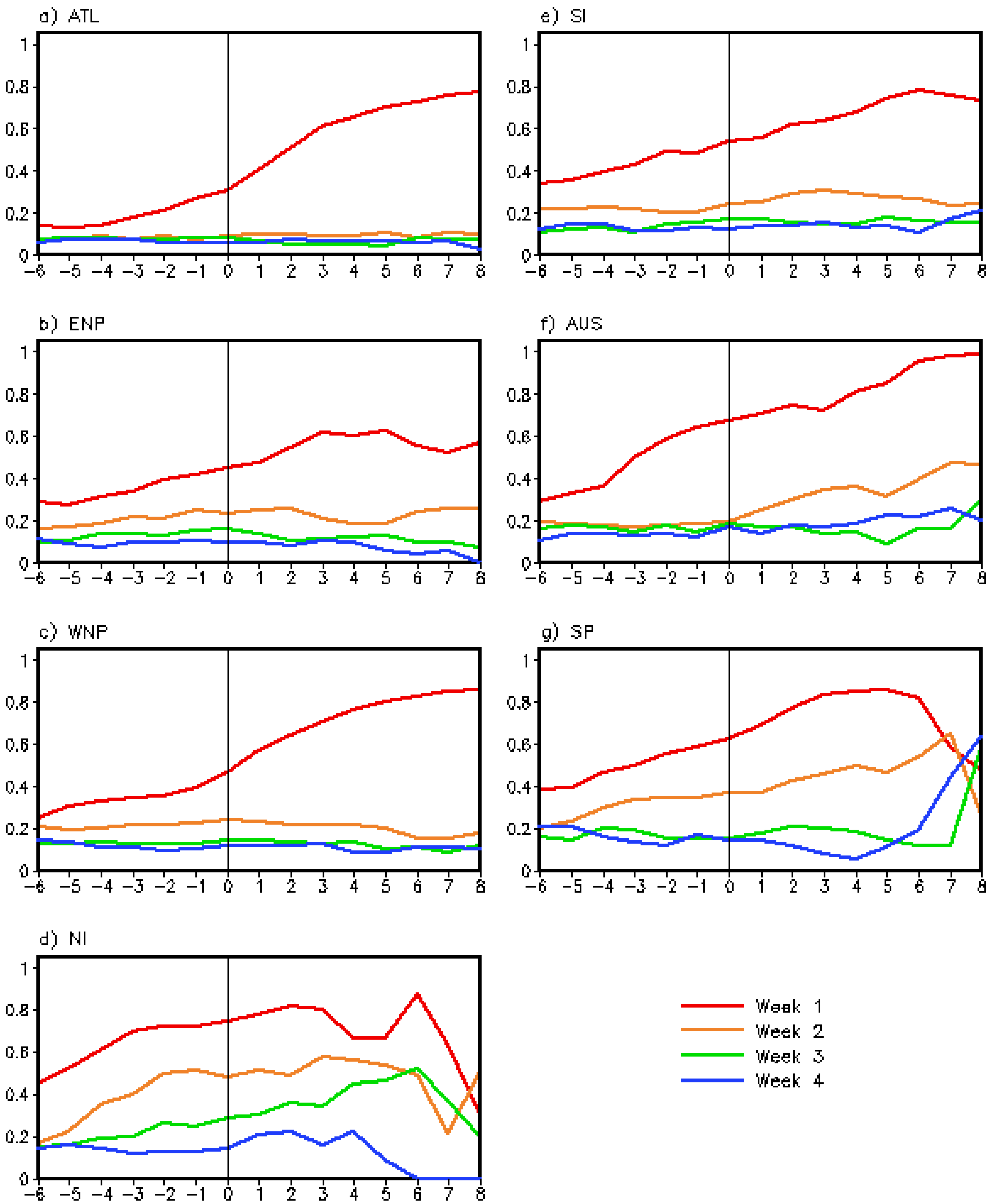

Fig. 4 Genesis lag day plots by week for the a) Atlantic, b) Eastern North Pacific, c) Western North Pacific, d) North Indian basins, e) South Indian, f) Australian, and g) South Pacific basins averaged for each storm from 1999-2012. The black vertical line indicates the day of genesis (Lag Day 0).

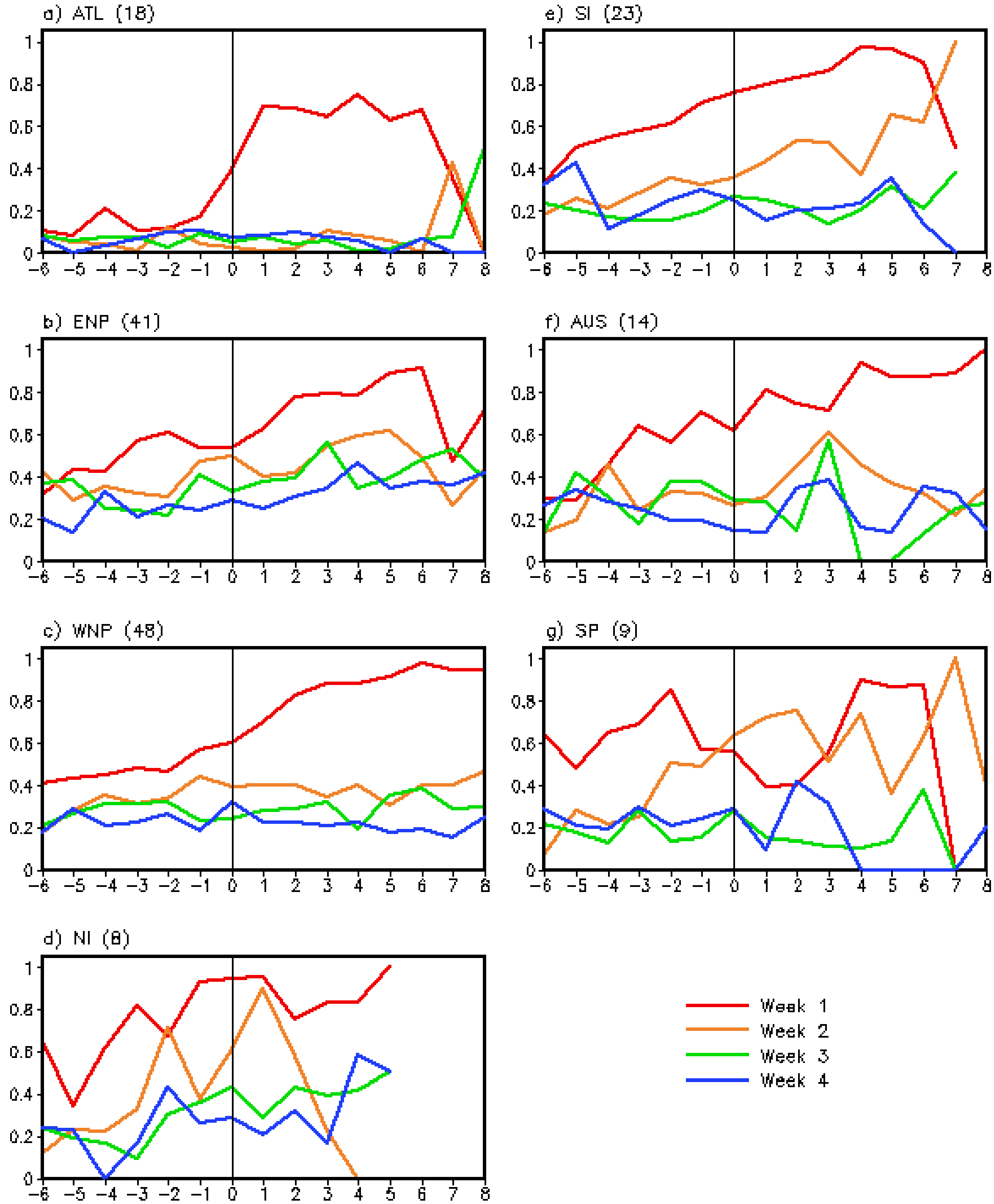

Fig. 5 Same as Figure 4, but for the 2014-2015 real-time evaluation. Numbers in parentheses indicate the total number of observed storms for the 2-year period.

Another way to view the model's skill in forecasting storm track is to compare the number of grid point hits (model and observations both showing a storm present) as a lag from the genesis point for individual storms. Unlike the HSS, this takes into account only hits or misses by the model and not FAs. Figure 4 shows this lag as a percentage of model hits versus the total possible hits for all storms in a basin during the 14-year period. The genesis day (Lag day 0) is highlighted with a vertical black line. During Week 1, the storm is included in the model's initial conditions (IC) starting at day 0. Therefore, negative lags (left of the black line) forecast both track and cyclogenesis, while positive lags (right of the black line) forecast only the track. For weeks 2-4, the entire period is forecasting both track and cyclogenesis with the exception of a storm that lasts over seven days. In this case, the storm is present in the ICs for the 7th and 8th lag day of Week 2.

Because of the influence of the ICs in Week 1, the percentage of hits, or hit rate, increases once the TS has formed (Figure 4). However, there is still considerable skill during the negative lags, most notably in the NI and Southern Hemisphere (SH) basins. This indicates that the lower skill in the HSSs for these basins is due to a high number of FA storms and not a missed forecast of observed storms. The NI and SP basins also show promise during Week 2. The opposite is true for the ATL, ENP, and WNP during the longer leads. Although the HSS shows higher scores for weeks 2-4 in these basins, the hit rate is very low, indicating more misses of observed storms and less FAs.

4. Real-time forecasting

In December 2013, ongoing, real-time prediction began using the 16-member operational runs described above. Because this began in the middle of the SH season, the results described below for the 2014 season are for January 1-May 30 only for the SH basins. The storm count anomaly correlations are given in Table 1. Numbers in bold represent those above the hindcast average seen in Figure 2. Because all sixteen ensemble members use ICs within a 24-hour period versus the five-day average of ensemble members needed for the hindcast runs, higher values are expected; however, this is not true for all basins. The WNP basin is the only basin with higher values for all weekly leads. The ATL and ENP basins also have many weekly leads above the hindcast average, with more occurring in 2014. The basins in the SH and the NI basin show mostly lower skill than the hindcasts runs, meaning there is little improvement with decreased lead time.

Table 1 Storm track count anomaly correlations for 2014 and 2015 during the active seasons. For 2014, SH basin correlations are for the shortened forecast period Jan 1 - May 31. Bold values are for those higher than the climatological value.

Year	Week	ATL	ENP	WNP	NI	SI	AUS	SP
2014	Week 1	**0.49**	0.36	**0.76**	0.24	**0.72**	0.06	0.25
	Week 2	0.43	**0.45**	**0.55**	0.14	**0.53**	-0.28	0.34
	Week 3	**0.26**	**0.36**	**0.39**	0.05	0.06	-0.37	**0.37**
	Week 4	**0.27**	**0.31**	**0.41**	-0.04	-0.13	-0.30	**0.29**
2015	Week 1	0.32	0.49	**0.63**	**0.43**	0.36	0.36	**0.52**
	Week 2	0.12	**0.33**	**0.74**	0.16	0.19	**0.30**	0.27
	Week 3	**0.10**	0.21	**0.47**	-0.04	0.13	0.05	0.03
	Week 4	**0.08**	0.12	**0.32**	-0.05	0.18	-0.19	-0.03

The dotted lines in Figure 3 represent the HSSs for the 2014-2015 operational evaluation. For most of the basins, the scores increase in magnitude, but overall show a similar pattern. The biggest score increases occur in the ENP, NI and AUS basins. The highest scores for AUS remain in the latter part of the season, peaking in March, while the highest scores for SI tend to shift more towards January, the peak in the seasonal cycle.

Although the genesis lag day plots for the 2014-2015 forecasts (Figure 5) are much noisier than those for the 14-year hindcast runs, they are overall very similar in structure. There is, however, increased skill for the real-time prediction in every basin at each weekly lead except for the ATL basin. These results are consistent with the increased HSS. An interesting point to note is that although the SI and AUS basins show relatively good skill in predicting the observed storms in the genesis lag day plots, the count correlations are relatively low except for the early week leads for the SI basin in 2014. This indicates an abundance of FAs still remaining in these basins.

5. Conclusion

With the availability of the CFSv2 45-day runs at NCEP, a new product on TS intraseasonal prediction has been developed to assist CPC forecasters. This product provides guidance on both storm count and storm location. Although skill drops with lead time, Weeks 2-4 still show skill for both storm count and storm track. Real-time experimental predictions for the 2014 and 2015 seasons show increased skill for many basins. It indicates predictability for the ATL, ENP and WNP basins, while the SH basins still struggle with FAs. This product is currently available as a non-operational product on the CPC ftp site at: ftp://ftp.cpc.ncep.noaa.gov/llong/main.html. Comments and suggestions are always welcome.

Acknowledgements. This study was partially supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program.

References

Camargo, S.J., and S.E. Zebiak, 2002: Improving the detection and tracking of tropical cyclones in atmospheric general circulation models. *Wea. Forecasting*, **17**, 1152-1162.

Klotzbach, P.J., 2010: On the Madden-Julian oscillation-Atlantic hurricane relationship. *J. Climate*, **23**, 282-293.

Maloney E.D., and D.L. Hartmann, 2000: Modulation of hurricane activity in the Gulf of Mexico by the Madden-Julian oscillation, *Science*, **287**, 2002-2004.

Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208.