Mid-year progress report (Year 1)

Title: Integration of an Objective, Automated TC Center-fixing Algorithm Based on Multispectral Satellite Imagery into NHC/TAFB Operations

Principal Investigators: Anthony Wimmers, Christopher Velden (UW/CIMSS)

Reporting period: 1 August 2013 – 28 February 2014

The following is a summary of the project's timeline for Year One:

1. **Dataset development**: Compile an archive of NHC forecasts, satellite imagery, aircraft data and manual position fixes for a 2005-2012 dataset of North Atlantic tropical cyclones
2. **Algorithm refinement**: Adapt the center-fixing algorithm, ARCHER, to the test period dataset in a simulated real-time mode
3. **Algorithm evaluation**: Verify/validate the ARCHER center-fixing performance
4. **Prepare results for publication**
5. **Real-time testing:** Ready the updated ARCHER algorithm for real-time testing in the 2014 North Atlantic hurricane season (this was added to the revised timeline)

Work progress:

<u>1. Dataset development</u>

We have finalized the 2005-2012 archive of NHC forecast tracks, aircraft position-fixes and satellite imagery (85-92 GHz microwave, GOES Vis/IR/Near-IR, and ASCAT).

We were unable to find a comprehensive archive of Windsat 37 GHz imagery to complement the dataset for ARCHER calibration. However, we expect to have access to three higher-resolution 37 GHz Level-2 datastreams in real-time during the 2014 test season, from Windsat, AMSR2 and GCOM. Our plan is to apply a preliminary calibration when including this data into ARCHER, and to revise the calibration as the season progresses.

The work remaining on this task is to compile the 2005-2012 TAFB/SAB manual position fixes for comparison to ARCHER center-fixes. This will be completed at the end of May 2014.

<u>2. Algorithm Refinement</u>

In addition to the center-fix position, ARCHER has several additional outputs, most importantly the center-fix *certainty*. We express this in terms of the radius of 50% center-fix certainty and the radius of 95% certainty, both of which come from a probability density function tied to the center-fix error distribution in a calibration dataset. The center-fix certainty value allows the algorithm to automatically evaluate the sensor with the best center-fix within a given time period, and select that position for the output. Also, in order to compare geostationary observations with regular temporal spacing to polar-satellite observations with irregular spacing, we further calibrated the error of center-fixes after being "nudged" in the direction of the NHC forecast track for a certain number of hours (Figure 1).

Figure 1. Calibration of center-fix certainty as a Poisson distribution to 85-92GHz ARCHER center-fix error. The histograms are the distance error of the ARCHER center-fix compared to the NHC Best Track. Each row is for a quartile of ARCHER confidence scores (lowest scores at the top row, highest scores at the bottom row). Each column calibrates the error of this center-fix projected out in time with guidance from the NHC forecast track (forecast time = 0, 3, 6, 9 hours, from left to right). The green line is a best-fit Poisson distribution. Given this robust match to the Poisson function, the certainty of any image's center-fix can be expressed as a function of 1) the ARCHER score and 2) the time difference between the satellite observation and the required time of the center-fix.

The center-fix certainty has been calibrated for each available data source (85-92 GHz; Geostationary Visible, IR, Near IR; ASCAT) using the 2006-2011 dataset. As demonstrated in the 85-92 GHz example in Figure 1, the error distribution corresponds very well to a Poisson-curve function of the ARCHER confidence score (a low-level output value) and the nudging time.

Individual center-fixes are combined together into a storm track by using the source with the lowest expected error (highest certainty) at three-hour intervals. This produces a fully automated, satellite-based storm track with visual center-fix certainty indicators (Figure 2).

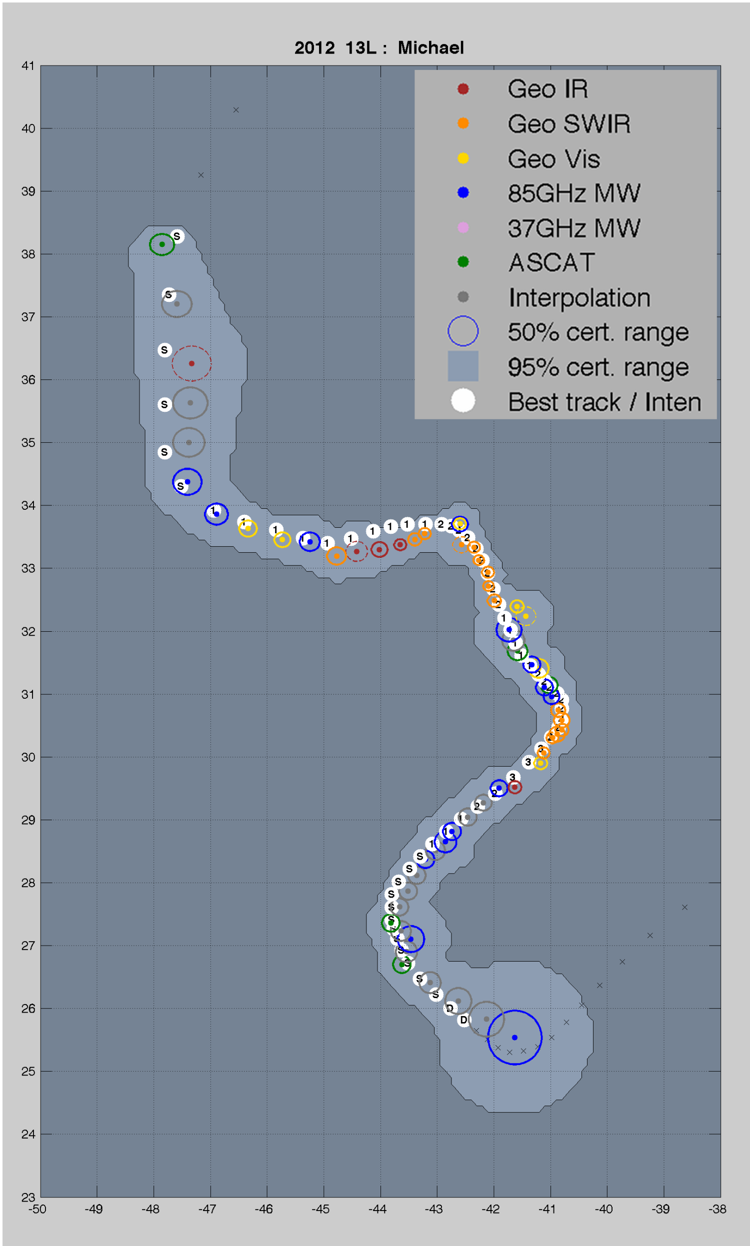

Figure 2. ARCHER-derived “track” result for Hurricane Michael (2012). Data is spaced at three-hour intervals. Data sources are indicated in the legend. Colored circles indicate the 50% certainty range for the associated center-fix, and the light shaded zone indicates the combined 95% certainty range of the satellite-based track. The NHC best track, for independent validation, is shown as white disks with the intensity written inside as D, S, 1, 2, 3, 4 or 5.

3. Algorithm Evaluation

The 2012 time period was reserved for an independent validation of the ARCHER performance. We summarize the results here, categorized by performance during tropical depression / tropical storm intensities, Category 1, and Category 2-5. The tropical depression / tropical storm column (TD-TS) is the most important, because this is where the most assistance can be offered to analysts and forecasters. Performance is also separated between “Real-time” (using only the data available at the valid time), and “Near-real-time” (using data available several hours before and after the valid time). Both versions are important for interpreting the performance of an algorithm that updates the recent track history with new information as it arrives.

The most straightforward summary of error is the average position error for all center-fixes (Table 1).

Table 1. Average error of ARCHER combined-sensor center-fix w.r.t. the NHC Best Track.

	TD - TS	Cat 1	Cat 2-5
Real-time	0.47 deg (51 km)	0.29 deg (32 km)	0.17 deg (19 km)
Near-real-time	0.38 deg (42 km)	0.27 deg (30 km)	0.16 deg (17 km)

Although the average error for the TD-TS group is reasonably low, this analysis does not account for the role of the center-fix certainty information. Specifically, the ARCHER center-fix will be used differently when the certainty is high than when it is low. Assuming ARCHER’s certainty values are unbiased, then a high certainty value means the expected error is significantly lower than the values in Table 1, and alternatively, a center-fix with a low certainty value will not be relied upon very heavily in practice. To that end, Table 2 shows the validation of the 50% certainty radius, in order to evaluate its bias. In this scheme, a bias of -10% means that a 50% certainty radius actually validates as the 40% certainty radius.

Table 2. Average bias of the ARCHER 50% certainty radius (see text for definition of bias).

	TD - TS	Cat 1	Cat 2-5
Real-time	-6.1%	-8.0%	-8.8%
Near-real-time	-13.7%	-16.7%	-11.7%

These results are quite good, because the errors in the Best Track itself (due to temporal smoothing, consensus averaging, and other reasons) should skew any well-performing algorithm toward slightly negative bias values. We interpret this table to mean that the ARCHER certainty estimates are very robust. This is confirmed by visual inspection of the track results as well.

Finally, in order to understand the relative merit of the ARCHER accuracy, we compare the error to the error of the real-time NHC analysis positions (Table 3).

Table 3. Percentage of cases in which ARCHER had lower error than the corresponding NHC analysis position.

	TD - TS	Cat 1	Cat 2-5
Real-time	29%	28%	25%
Near-real-time	36%	36%	39%

The frequency of ARCHER improvement is around 30%, which shows that although it is usually less accurate than a skilled forecaster, it has results that perform at the level of the operational forecast, and would very likely add skill when included in operations. Later in Year One we will perform a similar comparison to manual SAB and TAFB center-fixes.

<u>4. Prepare results for publication</u>

We have organized our results, such as those above, to be shared in a journal publication. When we have finished adding the remaining tasks such as comparison to manual center-fixes, we will begin to prepare our draft.

<u>5. Real-time testing</u>

A prior version of ARCHER is running at CIMSS in real-time. From April to June 2014 we will prepare this new version of ARCHER to run real-time in a similar fashion on a new server. The three output products that will be shared online will be 1) Diagnostic imagery for each satellite source's storm-centered image, 2) A tabular, chronological list of ARCHER results from each source for each TC, and 3) A composite track (as in Figure 2) for each TC.

The tabular list will be formatted such that NHC will be able to automate a system for reading ARCHER center-fix positions and certainty radii for local displaying in the AMAP2 framework.