TROPICAL CYCLONE CONSENSUS
DETERMINISTIC FORECASTS
LECTURE 1A: TROPICAL CYCLONE TRACKS

Russell L. Elsberry

Materials provided by Lixion Avila, Jim Goerss and Andrew Burton

Outline

Definition of terms
Basis for consensus forecasting
Consensus track spread and track errors
Practical issues with consensus forecasting
Selective consensus forecasts
Motion vector consensus forecasting

1st TRCG Technical Forum, Jeju, Korea     12-15 May 2009
DEFINITIONS OF TERMS - I

DETERMINISTIC MODEL (as contrasted to an ENSEMBLE PREDICTION SYSTEM) is a single numerical model that is typically the highest horizontal and vertical resolution model from a numerical weather prediction center.

ENSEMBLE PREDICTION SYSTEM produces *multiple forecasts* either by integrating a numerical model many times in which:

(i) initial conditions have been slightly modified to represent the uncertainty in the *initial conditions*;
(ii) the *model physics* representations (e.g., convection, friction) are different; or
(iii) both the initial conditions and the model physics are varied.

To complete the ensemble forecast in approximately the same time as a deterministic model for operational use, the numerical model typically has *less horizontal resolution* than the corresponding deterministic model.

1st TRCG Technical Forum, Jeju, Korea  12-15 May 2009
Encompassing Forecast Uncertainty

(consensus) An ensemble of likely analyses leads to an ensemble of likely forecasts

(Consensus)
Ensemble Forecasting:
-- Encompasses truth
-- Reveals uncertainty
-- Yields probabilistic information
DEFINITIONS OF TERMS - II

A CONSENSUS POSITION track forecast is simply an average of the initial and forecast latitudes and longitudes of a group of deterministic model tropical cyclone tracks at each forecast interval.

CONSENSUS TRACK SPREAD is either the average or the root mean square position deviations of the various deterministic model track forecast positions from the consensus (average) track position at each forecast interval.
Excellent example of GUNA consensus:
HURRICANE ISABEL, 1200 UTC 11 SEP 2003
Not-so-excellent example of GUNA consensus:
HURRICANE KATE, 1800 UTC 29 SEP 2003
This is a case where forming a selective consensus can be effective.
Katrina Track Forecasts
1200 UTC 24 August
Katrina Track Forecasts
0000 UTC 27 August
Ophelia Track Forecasts (nightmare)
1200 UTC 9 September
DECREASING NUMBER OF SMALL SPREAD/LARGE ERROR (SSLE) CASES WILL REQUIRE BETTER MODEL 120-H TRACK GUIDANCE

LARGE SPREAD/LARGE ERROR (LSLE) CASES HAVE COMPENSATING ERRORS THAT FORECASTER NEEDS TO RECOGNIZE

1st TRCG Technical Forum, Jeju, Korea  12-15 May 2009
Each deterministic model forecast track has *systematic* and *random errors*. As the dynamical/numerical models have been improved, the systematic errors have become smaller, and multiple skillful numerical models are now available.

A consensus (average) of *multiple skillful* track forecasts will tend to result in a *cancellation of the random errors*. The larger the number of model tracks in the consensus, the more effective the cancellation if the errors are random.
Tropical Prediction Center

yearly-average official track forecast errors and trend lines
Atlantic basin, without depressions

error in n. mi.

year

Western North Pacific (1992-2006)
72-h TC Track Forecast Error (nm)
TROPICAL CYCLONE MOTION

PATHWAYS TO IMPROVE TRACK PREDICTION

SINCE THE TEN-MEMBER CONSENSUS HAS PROVEN TO BE MORE SKILLFUL, ADDING MORE SKILLFUL MODEL TRACKS IS A VIABLE PATHWAY TO IMPROVED TRACK PREDICTION.

SINCE GFS AND COAMPS HAVE NOT IMPROVED, AND THE SMALL IMPROVEMENT IN THE CONSENSUS DURING 2005 AND 2006 MAY BE ATTRIBUTED TO IMPROVED NOGAPS, GFDN, AND UKMO, SOME FURTHER IMPROVEMENT IN TRACK IS POSSIBLE FROM MODEL IMPROVEMENTS.

JTWC ACHIEVED LARGE IMPROVEMENTMENT FROM 1999 TO 2002 FROM ADOPTING CONSENSUS FORECASTING. ONLY IN 2004 WAS JTWC ABLE TO IMPROVE SLIGHTLY ON THEIR TEN-MEMBER CONSENSUS AT 72 H.
Because the consensus track lies near the middle of the consensus spread, no matter whether the actual track is to the left or right (faster or slower) than the consensus track, at least one of the individual model tracks will always be worse.

Summed over a sufficiently large sample (e.g., a storm or a season), the track errors for a consensus formed from multiple (> 5) skillful models will be smaller than the errors for any of the individual models making up the consensus.
TROPICAL CYCLONE MOTION
IMPROVEMENT IN 120-H FORECASTS

MAJOR ADVANCE SINCE IWTC-V HAS BEEN ISSUANCE OF 120-H TRACK FORECASTS THAT ARE AS ACCURATE AS 72-H FORECASTS LESS THAN A DECADE AGO.

ONLY FOUR MODEL TRACK FORECASTS ARE AVAILABLE BEYOND 72 H. NOTE THAT JTWC DID NOT IMPROVE ON THE BEST MODEL OR THE CONSENSUS AT 120 H.

1st TRCG Technical Forum, Jeju, Korea   12-15 May 2009
DISADVANTAGES OF CONSENSUS TRACK FORECASTS

A deterministic model used in the consensus may have a systematic track error, especially in some synoptic situations.

Insufficient skillful models may be available to ensure cancellation of errors, especially in the early and late stages of the tropical cyclone.
COMPARISON WITH ENSEMBLE PREDICTION SYSTEMS

Ensemble prediction system typically displayed as ensemble mean track plus the spread in the various ensemble member tracks, which provides a measure of track uncertainty due to the initial condition uncertainty, model physics uncertainty, or both.

An average of the track errors for the higher resolution “mother” deterministic model will generally be slightly smaller than the average ensemble mean track errors because of lower horizontal resolution “daughter” ensemble numerical model.

Whereas the spread of the consensus models has been shown to be related to likely track forecast error (spread-error relationship), a demonstration is needed that an accurate spread-error relationship exists for the ensemble prediction system.

1st TRCG Technical Forum, Jeju, Korea 12-15 May 2009
Consensus Track Forecasting
Predictors of consensus forecast error must be quantities that are available prior to the time when official forecasts must be issued.

Consensus model spread is defined to be the average distance of the member forecasts from the consensus forecast.

The possible predictors are consensus model spread; initial and forecast TC intensity; initial TC location and forecast displacement of TC location (latitude and longitude); TC speed of motion; and number of members available (for CONW and TVCN/CONU).
Predicting TC Track Forecast Error

- Using stepwise linear regression and the aforementioned pool of predictors for the previous 3-4 seasons, regression models are found each season to predict consensus TC track forecast error for each combination of forecast length, consensus model, and basin.

- The regression models are then used to determine the radii of circular areas drawn around the consensus model forecast positions within which the verifying TC position is expected to be contained approximately 75% of the time.

- These circular areas are graphically displayed on the ATCF for use by the forecasters at NHC and JTWC. This graphical predicted consensus error product is referred to as GPCE (“gypsy”).
48-h Predicted Consensus Error
Hurricane Katrina - 12Z 27 August 2005
GPCE Validation
2005-2008 Western North Pacific

GPCE Validation (Percent)

CONW Forecast Error (nm)

WP2005
WP2006
WP2007
WP2008
GPCE Validation
2005-2008 Atlantic

GPCE Validation
(Percent)

CONU/TVCN Forecast Error
(nm)

AL2005
AL2006
AL2007
AL2008
Sharing Experiences in Operational Consensus Track Forecasting

Rapporteur: Andrew Burton

Team members: Philippe Caroff, James Franklin, Ed Fukada, T.C. Lee, Buck Sampson, Todd Smith.
Consensus Track Forecasting

Recent performance of a model does not guarantee success/failure next time.
Consensus Track Forecasting

Recent performance of a model does not guarantee success/failure next time.
Consensus Track Forecasting

Optimising consensus tracks

Accuracy depends on:
1. Number of models
2. Accuracy of individual members
3. Independence of member errors

Including advisories in the consensus
JTWC, JMA, CMA.
Would you add WBAR to your consensus?

A Question of Independence

Fig. 1. The 24-, 48-, and 72-h errors (km) of WBAI (triangles) and the operational CLIPER (squares) for the period 9 Jun 2003–7 Jul 2004. The number of forecasts is shown in parentheses.
Would you add WBAR to your consensus?
Practical Considerations

- Access to models?
- Where to get them from? (JMA e.g.?)
- Can we organise a central repository of global TC tracks?

Standard format and timely!
Consensus Track Forecasting

Practical Considerations contd.

- Access to software?
- Access to model fields
- Pre-cyclone phase – less tracks
- Capture/recurvature/ETT
Consensus Track Forecasting

- Multi-model approaches – ‘simple’ example.

Process:

- Acquire tracks
- Perform initial position correction
- Interpolate tracks
- Geographically average

![Consensus Track Forecasting Diagram]
Consensus Track Forecasting

- Non-selective multi-model consensus
  - Low maintenance
  - Low training overhead
  - Incorporate ‘new’ models ‘on-the-fly’
  - Robust performance
  - If many members, less need for selective approach
  - Widely adopted as baseline approach
Selective vs. non-selective approaches

- Subjective selection common place and can add significant value.
- Semi-objective selection: SAFA – implementation encountered hurdles.
- How to identify those cases where selective approach will add value?
Consensus Track Forecasting

Selective (SCON) Vs Non-selective (NCON)

How to exclude members?
Consensus Track Forecasting

Selective (SCON) Vs Non-selective (NCON)

SCON – How to exclude members?
Consensus Track Forecasting

Selective (SCON) vs non-selective (NCON)

- SCON – How to exclude members?
- Requires knowledge of known model biases (this changes with updates)
Consensus Track Forecasting

Selective (SCON) Vs Non-selective (NCON)

- SCON – How to exclude members?
- Requires knowledge of model run
- Eg analyses differs from observed BEWARE
ASSESSMENT OF WESTERN NORTH PACIFIC 96-AND 120-HOUR TRACK GUIDANCE AND PRESENT FORECASTABILITY*

Kathryn A. Payne
Russell L. Elsberry
Mark A. Boothe

*2007 Weather and Forecasting, 22, 1003-1015

Four models evaluated:
Navy Operational Global Atmospheric Prediction System (NOGAPS)
Navy version of Geophysical fluid Dynamics Laboratory (GFDN)
United Kingdom Meteorological Office (UKMO)
NCEP Global Forecast System
Motivation

• Consensus of four forecast models, (ten models available for 3-day forecasts)
  – Large uncertainty exists
• Joint Typhoon Warning Center (JTWC) errors
• How can we improve 5-day forecasts?
  – Forecasts need to know how much to trust guidance
  – Selective Consensus can be created
• Document model characteristics and recurring meteorological error mechanisms described in Carr and Elsberry (2000 a, b)
• Kehoe (2005) examined 2004 NW Pacific season for NOGAPS and GFDN
Methods

• Looked at large error cases for 2005 NW Pacific season
• Examined all 96-h and 120-h forecasts with large track errors (>400 n mi at 96 h, >500 n mi at 120 h)
• Studied forecast fields for four models: NOGAPS, GFS(NCEP), UKMO, GFDN, to determine frequently occurring error mechanisms
• Examined role of model TC structure in creating large error
• Estimated 120-h predictability through the creation of a selective consensus
Model Error Evaluation

- Models have similar performance at 72 h, greater disparity in performance at 120 h
- Models prone to large errors at 96 h and 120 h
- JTWC not beating consensus—room for improvement
Table 1. 96-h and 120-h error mechanisms for NOGAPS, GFDN, GFS, and UKMO occurring in 2005. The first (second) number listed is the number of times the phenomenon occurred excessively (insufficiently). Two “Other” GFDN errors occurred due to the model TC failing to decay over land. Fourteen “Other” GFS errors occurred when the TC excessively dissipated with no vertical shear involved.

<table>
<thead>
<tr>
<th>ERROR MECHANISM</th>
<th>Acronym</th>
<th>NOGAPS</th>
<th>GFDN</th>
<th>GFS</th>
<th>UKMO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Large</td>
<td>Due</td>
<td>To</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Errors</td>
<td>Due</td>
<td>To</td>
<td></td>
</tr>
<tr>
<td>Large Errors Due To Tropical Influences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Cyclone Interaction (tropical)</td>
<td>DCI-t</td>
<td>2-0</td>
<td>7-0</td>
<td>6-0</td>
<td>4-0</td>
</tr>
<tr>
<td>Indirect Cyclone Interaction (West TC)</td>
<td>ICI-W</td>
<td>0-0</td>
<td>0-0</td>
<td>1-0</td>
<td>0-0</td>
</tr>
<tr>
<td>Ridge Modification by TC</td>
<td>RMT</td>
<td>1-0</td>
<td>0-0</td>
<td>0-0</td>
<td>0-0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Errors Due To Midlatitude Influences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct Cyclone Interaction (midlatitude)</td>
<td>DCI-m</td>
<td>1-0</td>
<td>0-0</td>
<td>13-0</td>
<td>4-0</td>
</tr>
<tr>
<td>Response to Vertical Wind Shear</td>
<td>RVS</td>
<td>6-0</td>
<td>0-0</td>
<td>8-0</td>
<td>0-0</td>
</tr>
<tr>
<td>Midlatitude Cyclogenesis</td>
<td>MCG</td>
<td>21-9</td>
<td>4-1</td>
<td>0-2</td>
<td>2-14</td>
</tr>
<tr>
<td>Midlatitude Cyclolysis</td>
<td>MCL</td>
<td>4-0</td>
<td>0-1</td>
<td>0-1</td>
<td>0-0</td>
</tr>
<tr>
<td>Midlatitude Anticyclogenesis</td>
<td>MAG</td>
<td>7-5</td>
<td>31-0</td>
<td>0-0</td>
<td>7-0</td>
</tr>
<tr>
<td>Midlatitude Anticyclolysis</td>
<td>MAL</td>
<td>0-0</td>
<td>4-0</td>
<td>0-0</td>
<td>0-0</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td>0</td>
<td>2</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Fields not available</td>
<td>4</td>
<td>16</td>
<td>9</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>66</td>
<td>54</td>
<td>45</td>
<td></td>
</tr>
</tbody>
</table>
GFS case study - Kirogi (21 W)
Initial intensity: 90 kt

GFS MSLP 0 h

GFS 500mb GPH 0 h

GFS MSLP 24 h

NOGAPS Anal MSLP 0 h

NOGAPS Anal 500mb GPH 0 h

NOGAPS Anal MSLP 24 h
95% of GFS large errors involved incorrect TC structure
TC 18W
Saola

Actual position in 4 days

GFS forecast position in 4 days
a) GFS 500mb Forecast (valid in 18 h)

c) GFS MSLP Forecast (valid in 18 h)

e) GFS 500mb Forecast (valid in 54 h)

g) GFS MSLP Forecast (valid in 54 h)
TC 14W
Nabi

Actual position in 108 h

GFDN forecast position in 108 h
Error vs. Spread - Elsberry and Carr 2000

- Small Spread/Small Error: Ideal Case
- Small Spread/Large Error: Nightmare
- Large Spread/Small Error: Opposing errors cancel each other out
- Large Spread/Large Error: Largest opportunity for improvement
2005 Error vs. Spread

- 48% of cases with large spread (< 300 ni mi)
- 25% LSLE cases investigated for SCON
- Forecasters Beware—23% LSSE cases

Consensus (Elsberry and Carr 2000) Error vs. Spread

<table>
<thead>
<tr>
<th></th>
<th>SSLE</th>
<th>LSLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consensus Error</td>
<td>20 (13%)</td>
<td>38 (25%)</td>
</tr>
<tr>
<td>Consensus Spread</td>
<td>59 (39%)</td>
<td>34 (23%)</td>
</tr>
</tbody>
</table>
### TOPIC 3 SUMMARY REPORT

**OPTIMAL USE OF GUIDANCE - II**

**POTENTIAL REDUCTION IN 120-H ERRORS FROM OPTIMUM USE OF GUIDANCE FOR 38 (33) LSLE CASES DURING 2004 (2005)**

<table>
<thead>
<tr>
<th></th>
<th>SCON vs CONW</th>
<th>SCON vs JTWC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td>2004</td>
<td>2005</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>239</td>
<td>222</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td>626</td>
<td>629</td>
</tr>
<tr>
<td><strong>Minimum</strong></td>
<td>13</td>
<td>64</td>
</tr>
</tbody>
</table>

**SELECTIVE CONSENSUS (SCON) HAS POTENTIAL TO IMPROVE TRACK FORECASTING WHEN ONLY A SMALL NUMBER OF MODEL TRACKS ARE AVAILABLE AND ONE OR TWO TRACKS ARE CLEARLY ERRONEOUS**

IWTC-VI   21-30 November 2006
Conclusions

• Major problems exist with some 5-day model forecasts

• Key is for forecasters to know when to trust the consensus and when they might be able to improve upon the consensus with an SCON

• Results show a 222 n mi improvement over NCON in 20% of 120-h forecasts
A Weighted Consensus Approach to Tropical Cyclone 96-h and 120-h Track Forecasting

Russell L. Elsberry
James R. Hughes
Mark A. Boothe


Acknowledgments: Buck Sampson, NRL: Figure and ATCF guidance
JTWC: ATCF data
Overview

• Weighted Position Consensus Concept
• Methodology
• Weighted Position Consensus Example
• Validation Study Results
• Sensitivity Studies
• Weighted Motion Vector Consensus
  – Methodology
  – Example
  – Results
• Conclusions
• Future Work
Weighted Consensus Technique for Tropical Cyclone Track Forecasting

• **Concept:** Improve long-range track forecasts through weighting consensus members according to their consistency with the consensus track at earlier forecast times.
  
  – The consensus at 72 h has as many as 11 models while the 96-120 h consensus has 6 models or less.
  – Gain value from models that drop out of the consensus after 72 h by weighting remaining members toward the 60-72 h consensus
  – Avoid erratic track changes when a model track drops out of consensus
Weighted Position Consensus

120 h
NOGAPS
AVN (GFS)
EGR (UKMet)
GFDN
~72 h
COAMPS
AFWA MM5
JGSM
JTYM
TCLAPS
WBAR
CONW (Operational)
UAVE
Weighted Position Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>37.5</td>
</tr>
<tr>
<td>B</td>
<td>37.5</td>
</tr>
<tr>
<td>C</td>
<td>12.5</td>
</tr>
<tr>
<td>D</td>
<td>12.5</td>
</tr>
</tbody>
</table>

120 h
NOGAPS
AVN (GFS)
EGRR (UKMet)
GFDN
~72 h
COAMPS
AFWA MM5
JGSM
JTYM
TCLAPS
WBAR
CONW (Operational)
UAVE
Weighted Position Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>37.5</td>
</tr>
<tr>
<td>B</td>
<td>37.5</td>
</tr>
<tr>
<td>C</td>
<td>12.5</td>
</tr>
<tr>
<td>D</td>
<td>12.5</td>
</tr>
</tbody>
</table>
Methodology

• Data: Automated Tropical Cyclone Forecasting System (Sampson and Schrader 2000) 2006 western North Pacific Storms 1-24

• Assume only NOGAPS, GFS, UKMO, and GFDN after 72 h

• Model Precedence:

<table>
<thead>
<tr>
<th>Precedence</th>
<th>Track Description</th>
<th>Example (NOGAPS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Primary interpolated track</td>
<td>NGPI</td>
</tr>
<tr>
<td>II</td>
<td>Fiorino interpolated track</td>
<td>JNGI</td>
</tr>
<tr>
<td>III</td>
<td>12-h old interpolated track</td>
<td>NGP2</td>
</tr>
<tr>
<td>IV</td>
<td>12-h old Fiorino interpolated track</td>
<td>JNG2</td>
</tr>
</tbody>
</table>
Methodology: Model Guidance

• Four Primary Models
  – Navy Operational Global Atmospheric Prediction System (NOGAPS)
  – National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS)
  – U.K. Meteorological Office (UKMO) global model
  – Geophysical Fluid Dynamics Laboratory – Navy model (GFDN)

• Additional Models
  – Japan Meteorological Agency Global Spectral Model (JGSM)
  – Japan Meteorological Agency Typhoon Model (JTYM)
  – Coupled Ocean-Atmosphere Mesoscale Prediction System (COAMPS)
  – Air Force Weather Agency MM5 (MM5)
  – Australian TC-Local Area Prediction System (TC-LAPS).
  – Weber Barotropic Model (WBAR)
  – European Center for Medium range Weather Forecasting (ECMWF).
Methodology

• Weighting Procedure
  – Calculate normalized weights inversely proportional to distance from consensus track at 60 h, 66 h, and 72 h for each model track
  – Apply weights to 96-h, 108-h, and 120-h model track positions

• Case selection criteria
  – Must have at least 4 additional models at 72 h
  – At least two models at 96-h, 108-h, 120-h
Methodology

• Definitions
  – CONW: Consensus used operationally at JTWC
  – WCOW: Weighted position consensus using CONW
  – UAVE: Position consensus of selected models
  – WAVE: Weighted position consensus using AVE
  – UVAE: Unweighted motion vector consensus
  – WVVAE: Weighted motion vector consensus
### Model Weights

<table>
<thead>
<tr>
<th>Model</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGPI</td>
<td>0.75</td>
</tr>
<tr>
<td>AVNI</td>
<td>0.06</td>
</tr>
<tr>
<td>EGRI</td>
<td>0.04</td>
</tr>
<tr>
<td>GFNI</td>
<td>0.15</td>
</tr>
</tbody>
</table>

### Time Summary

<table>
<thead>
<tr>
<th>Time</th>
<th>CONW error (n mi)</th>
<th>WCOW error (n mi)</th>
<th>WCOW Imp (n mi)</th>
<th>WCOW % Imp</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 h</td>
<td>381.6</td>
<td>181.0</td>
<td>200.6</td>
<td>52.6</td>
</tr>
<tr>
<td>108 h</td>
<td>477.9</td>
<td>230.5</td>
<td>247.5</td>
<td>51.8</td>
</tr>
<tr>
<td>120 h</td>
<td>537.5</td>
<td>245.4</td>
<td>292.1</td>
<td>54.4</td>
</tr>
<tr>
<td>Model</td>
<td>Weights</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGPI</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVNI</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGRI</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GFNI</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>CONW error (n mi)</th>
<th>WCOW error (n mi)</th>
<th>WCOW Imp (n mi)</th>
<th>WCOW % Imp</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 h</td>
<td>85.1</td>
<td>140.5</td>
<td>-55.4</td>
<td>-65.1</td>
</tr>
<tr>
<td>108 h</td>
<td>153.7</td>
<td>232.3</td>
<td>-78.6</td>
<td>-51.2</td>
</tr>
<tr>
<td>120 h</td>
<td>225.3</td>
<td>330.9</td>
<td>-105.6</td>
<td>-46.9</td>
</tr>
</tbody>
</table>
Weighted Position Consensus
Results: Validation Study

- The weighted position consensus improved on the unweighted position consensus for both UAVE and CONW when averaged over the 24 storms from 2006.

- The weighted position consensus varies in performance on a case by case basis.

- Selectively eliminating some of the worst performing cases could significantly improve the seasonal performance of the weighted position consensus.

<table>
<thead>
<tr>
<th></th>
<th>Mean Imp</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 h WCOW Imp (n mi):</td>
<td>6.0</td>
<td>222</td>
</tr>
<tr>
<td>96 h WCOW Imp (%):</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td>96 h WAVE Imp (n mi):</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>96 h WAVE Imp (%):</td>
<td>2.8</td>
<td></td>
</tr>
<tr>
<td>108 h WCOW Imp (n mi):</td>
<td>1.7</td>
<td>193</td>
</tr>
<tr>
<td>108 h WCOW Imp (%):</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>108 h WAVE Imp (n mi):</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>108 h WAVE Imp (%):</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>120 h WCOW Imp (n mi):</td>
<td>5.3</td>
<td>168</td>
</tr>
<tr>
<td>120 h WCOW Imp (%):</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>120 h WAVE Imp (n mi):</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>120 h WAVE Imp (%):</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>
Validation Study: Performance Graphs

% Cases Improved: 46

% Cases Improved: 49

1.8\% mean improvement

1.9\% mean improvement
Sensitivity Test: Weighted Consensus without COAMPS and MM5

• In 2007, the COAMPS and MM5 will be omitted from the operational consensus CONW, which will affect the weightings that depend on the distance from consensus at 60 h, 66 h, and 72 h.

<table>
<thead>
<tr>
<th>Time</th>
<th>UAVE Error (n mi)</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>W/O C &amp; M</td>
</tr>
<tr>
<td>60 h</td>
<td>111.4</td>
<td>107.9</td>
</tr>
<tr>
<td>66 h</td>
<td>124.3</td>
<td>120.6</td>
</tr>
<tr>
<td>72 h</td>
<td>138.9</td>
<td>135.2</td>
</tr>
</tbody>
</table>

• Unweighted position consensus (UAVE) improved at 60 h, 66 h, and 72 h, which are used for the weightings.

• Weighted consensus (WAVE) improved due to the superior weighting.

• These results are further justification of JTWC’s decision to eliminate COAMPS and MM5 from CONW.

<table>
<thead>
<tr>
<th>Time</th>
<th>Mean Improvement (n mi)</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 h</td>
<td>Control: 6.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 7.8</td>
<td>222</td>
</tr>
<tr>
<td>96 h</td>
<td>Control: 2.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 3.5</td>
<td>222</td>
</tr>
<tr>
<td>108 h</td>
<td>Control: 5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 8.7</td>
<td>193</td>
</tr>
<tr>
<td>108 h</td>
<td>Control: 2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 3.3</td>
<td>193</td>
</tr>
<tr>
<td>120 h</td>
<td>Control: 5.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 9.4</td>
<td>168</td>
</tr>
<tr>
<td>120 h</td>
<td>Control: 1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W/O C &amp; M: 3.1</td>
<td>168</td>
</tr>
</tbody>
</table>
Performance Graphs: Without COAMPS and MM5

% Cases Improved: 51

3.9% mean improvement
Sensitivity Test: Weighting Optimization - JGSM

- Japanese Global Model (JGSM) was occasionally available at 96 h.

**Control:** UAVE (5 models) – WAVE (original 4 models) = WAVE Improvement

**Test:** UAVE (5 models) – WAVE (5 models) = WAVE Improvement

<table>
<thead>
<tr>
<th>Mean Improvement</th>
<th>Cases weighted with JGSM</th>
<th>All Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>W/ JGSM</td>
</tr>
<tr>
<td>96 h WAVE Imp (n mi):</td>
<td>0.5</td>
<td>7.5</td>
</tr>
<tr>
<td>96 h WAVE Imp (%):</td>
<td>0.2</td>
<td>3.1</td>
</tr>
</tbody>
</table>

- It is clear that all skillful guidance should be included in the weighting scheme even if available in a fraction of the cases.
Performance Graphs: Weighting Optimization - JGSM

With JGSM in Weighting Scheme

% Cases Improved: 61

0.2% mean improvement
Sensitivity Test: Weighted Position Consensus with ECMWF

- This test evaluates the impact of adding ECMWF to UAVE and including it in the weighting.

### UAVE Error

<table>
<thead>
<tr>
<th></th>
<th>96 h</th>
<th>108 h</th>
<th>120 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>205</td>
<td>251</td>
<td>293</td>
</tr>
<tr>
<td>ECMWF</td>
<td>181</td>
<td>221</td>
<td>262</td>
</tr>
<tr>
<td>SS</td>
<td>59</td>
<td>55</td>
<td>52</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>96 h</th>
<th>108 h</th>
<th>120 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>221</td>
<td>262</td>
<td>302</td>
</tr>
<tr>
<td>ECMWF</td>
<td>214</td>
<td>253</td>
<td>293</td>
</tr>
<tr>
<td>SS</td>
<td>222</td>
<td>193</td>
<td>168</td>
</tr>
</tbody>
</table>

### WAVE Improvement

<table>
<thead>
<tr>
<th></th>
<th>96 h</th>
<th>108 h</th>
<th>120 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>6.2</td>
<td>5.7</td>
<td>5.7</td>
</tr>
<tr>
<td>ECMWF</td>
<td>5.8</td>
<td>5.2</td>
<td>4.5</td>
</tr>
<tr>
<td>SS</td>
<td>222</td>
<td>193</td>
<td>168</td>
</tr>
</tbody>
</table>

- Adding ECMWF significantly decreases the mean consensus error at 96 h, 108 h, and 120 h, but does not have a large additional impact on the performance of the weighted position consensus.
Sensitivity Test: Weighted Consensus with ECMWF

Average Track Errors (n mi) for Homogeneous Sample (2006)

<table>
<thead>
<tr>
<th></th>
<th>00 h</th>
<th>12 h</th>
<th>24 h</th>
<th>36 h</th>
<th>48 h</th>
<th>72 h</th>
<th>96 h</th>
<th>120 h</th>
</tr>
</thead>
<tbody>
<tr>
<td>JTWC</td>
<td>10.5</td>
<td>39.7</td>
<td>67.7</td>
<td>93.8</td>
<td>118.3</td>
<td>160.7</td>
<td>225.5</td>
<td>320.8</td>
</tr>
<tr>
<td>ECMI</td>
<td>10.6</td>
<td>54.0</td>
<td>77.2</td>
<td>101.0</td>
<td>120.9</td>
<td>164.8</td>
<td>195.3</td>
<td>240.2</td>
</tr>
<tr>
<td>TESB</td>
<td>10.6</td>
<td>36.2</td>
<td>58.9</td>
<td>79.6</td>
<td>97.2</td>
<td>134.0</td>
<td>188.2</td>
<td>277.2</td>
</tr>
<tr>
<td># Cases</td>
<td>289</td>
<td>277</td>
<td>261</td>
<td>244</td>
<td>221</td>
<td>169</td>
<td>120</td>
<td>90</td>
</tr>
</tbody>
</table>

- The ECMWF has greater errors than TESB (experimental consensus) at 72 h, one of the times when the weighting is computed.
- The ECMWF performs better relative to TESB at longer forecast times (96 h, 120 h) than at 72 h.

Table provided by Buck Sampson, NRL
Performance Graphs: Including the ECMWF

% Cases Improved: 58

4.3% mean improvement
Weighted Motion Vector Consensus

- Test if applying the weighting scheme to motion vector consensus will yield improvements over a weighted position consensus

Consensus forecast tracks generated from the tracks shown in (a) (with initial position correction applied) using (b) average geographical position and (c) average vector motion. (From Burton, 2006)
Weighted Vector Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.375</td>
</tr>
<tr>
<td>B</td>
<td>0.375</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
</tr>
</tbody>
</table>

120 h
NOGAPS
AVN (GFS)
EGRU (UKMet)
GFDN
~72 h
COAMPS
AFWA MM5
JGSM
JTYM
TCLAPS
WBAR
CONW (Operational)
UAVE
WAVE
Weighted Vector Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.375</td>
</tr>
<tr>
<td>B</td>
<td>0.375</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
</tr>
</tbody>
</table>

Unweighted
0.25 to all four

120 h
NOGAPS
AVN (GFS)
EGRR (UKMet)
GFDN
~72 h
COAMPS
AFWA MM5
JGSM
JTYM
TCLAPS
WBAR
CONW (Operational)
UAVE
WAVE
Weighted Vector Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.375</td>
</tr>
<tr>
<td>B</td>
<td>0.375</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
</tr>
</tbody>
</table>

Unweighted
0.25 to all four
Weighted Vector Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.375</td>
</tr>
<tr>
<td>B</td>
<td>0.375</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
</tr>
</tbody>
</table>

120 h
- NOGAPS
- AVN (GFS)
- EGRR (UKMet)
- GFDN

~72 h
- COAMPS
- AFWA MM5
- JGSM
- JTYM
- TCLAPS
- WBAR
- CONW (Operational)
- UAVE
- WAVE
- UVAE

84-96 h
108-120 h
96-108 h
84-96 h
108-120 h
96-108 h
84-96 h
Weighted Vector Consensus

<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.375</td>
</tr>
<tr>
<td>B</td>
<td>0.375</td>
</tr>
<tr>
<td>C</td>
<td>0.125</td>
</tr>
<tr>
<td>D</td>
<td>0.125</td>
</tr>
</tbody>
</table>

120 h
NOGAPS
AVN (GFS)
EGRR (UKMet)
GFDN
~72 h
COAMPS
AFWA MM5
JGSM
JTYM
TCLAPS
WBAR
CONW (Operational)
UAVE
WAVE
UVAE
WVAE
<table>
<thead>
<tr>
<th>Model</th>
<th>Weights</th>
<th>Time</th>
<th>UAVE error</th>
<th>WAVE error</th>
<th>WAVE Imp</th>
<th>WVAE error</th>
<th>WVAE Imp</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGPI</td>
<td>0.31</td>
<td>96 h</td>
<td>252.4</td>
<td>245.3</td>
<td>7.1</td>
<td>167.4</td>
<td>84.9</td>
</tr>
<tr>
<td>AVNI</td>
<td>0.53</td>
<td>108 h</td>
<td>283.1</td>
<td>275.8</td>
<td>7.4</td>
<td>196.4</td>
<td>87.0</td>
</tr>
<tr>
<td>EGRI</td>
<td>NA</td>
<td>120 h</td>
<td>378.0</td>
<td>370.9</td>
<td>7.1</td>
<td>293.5</td>
<td>84.6</td>
</tr>
<tr>
<td>GFNI</td>
<td>0.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Weighted Motion Vector Results

<table>
<thead>
<tr>
<th></th>
<th>WAVE</th>
<th>UVAE</th>
<th>WVAE</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>96 h Imp (n mi):</td>
<td>6.2</td>
<td>18.5</td>
<td>21.9</td>
<td>222</td>
</tr>
<tr>
<td>96 h Imp (%):</td>
<td>2.8</td>
<td>8.4</td>
<td>9.9</td>
<td></td>
</tr>
<tr>
<td>108 h Imp (n mi):</td>
<td>5.7</td>
<td>16.2</td>
<td>19.7</td>
<td>193</td>
</tr>
<tr>
<td>108 h Imp (%):</td>
<td>2.2</td>
<td>6.2</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>120 h Imp (n mi):</td>
<td>5.7</td>
<td>12.5</td>
<td>17.0</td>
<td>168</td>
</tr>
<tr>
<td>120 h Imp (%):</td>
<td>1.9</td>
<td>4.1</td>
<td>5.6</td>
<td></td>
</tr>
</tbody>
</table>

- Most of the improvement over an unweighted position consensus is simply achieved by using a motion vector consensus
- Further improvement is achieved by a weighted consensus of the motion vectors
Performance Graphs: Weighted Motion Vector Consensus

% Cases Improved: 63

5.6% mean improvement
Conclusions

• This study shows that there is value in using a simple weighted position consensus in improving long-range TC track forecasts

• A motion vector consensus yields a large improvement over an position consensus