Model Evaluation Tools – Tropical Cyclone (MET-TC)

Kathryn M. Newman
Introduction
Introduction

• WHAT is MET-TC?
  • A set of tools to aid in TC forecast evaluation and verification
  • Developed to replicate (and add to) the functionality of the National Hurricane Center (NHC) verification software
  • Modular set of tools which utilize the MET software framework
    • Allows for additional capabilities and features to be added to future releases

• WHY use MET-TC?
  • Provides Tropical Cyclone (TC) verification statistics consistent with operational centers
  • Easily parse and subset TC datasets
Compile & build

- Must use **METv4.1 or METv5.0** for MET-TC
- MET-TC specific code and tools:
  - `bin/` : executables for each MET-TC module (`tc_dland`, `tc_pairs`, `tc_stat`)
  - `share/met/config/` : configuration files
    (`TCPairsConfig_default`, `TCStatConfig_default`)
  - `share/met/tc_data/` : static files used in MET-TC (`*land.dat`, `wwpts_us.txt`)
  - `doc/` : contains the MET-TC User’s Guide
  - `src/tools/tc_utils/` : source code for three MET-TC modules
  - `scripts/Rscripts/` : contains the R script (`plot_tcmpr.R`) which provides graphics tools for MET-TC
Getting Started...

• The **best track analysis** is used primarily used as the observational dataset in MET-TC.
  • May use any reference dataset in ATCF format

• The input files must be in Automated Tropical Cyclone Forecasting System (**ATCF** format).

• Model output must be run through an internal/external **vortex tracking algorithm**
Observations

- Observations are an important consideration for TC verification
  - Quality and quantity of observations available
    - Typically sparse or intermittent

- The **best track analysis** is used primarily used as the observational dataset in MET-TC.

All operational model aids and best track analysis can be found on the NHC ftp server: ftp://ftp.nhc.noaa.gov/atcf/archive/

The Best Track is a subjective dataset – not consistent across basins!!
Observations

- **Best track analysis**
  - Subjective assessment of TC’s center location and intensity (6 hr) using all observations available
  - Includes center position, maximum sfc winds, minimum center pressure, quadrant radii of 34/50/64 kt winds
  - Subjectively smoothed representation of storm’s location and intensity over its lifetime
Getting Started...

- Automated Tropical Cyclone Forecasting System (**ATCF** format)
  - First developed at Naval Oceanographic and Atmospheric Research Laboratory (NRL)
  - Currently used for NHC operations

- Must adhere to for MET-TC tools to properly parse the input data (first 17 columns must exist - missing values ok)
  - To ensure proper matching input data must contain:
    - Basin, cyclone number, initialization time, forecast hour, model name
  
- MET-TC User’s Guide outlines these 17 columns and necessary fields

- For detailed information on ATCF format:
Getting Started...

- Model output must be run through an internal/external vortex tracking algorithm
- Any algorithm that obtains basic position, maximum wind, minimum sea level pressure information from model forecasts (in ATCF format) may be used

- Fully supported and freely available: GFDL Vortex Tracker

For more information (includes code and documentation):

MET-TC components

- Primary functions of the code are:
  - Compute pair statistics from ATCF input files
  - Filter pair statistics based on user specifications
  - Compute summary statistics
TC-dland

- Aids in quickly parsing data for filter jobs:
  - Only verify over water
  - Threshold verification based on distance to land
  - Exclusion/inclusion of forecasts within a specified window of landfall

- **Input**: ASCII file containing Lon/Lat coordinates of all coastlines/islands considered to be a significant landmass. (aland.dat, shland.dat, wland.dat)

- **Output**: gridded field representing distance to nearest coastline/island in NetCDF format

![Distance to Land Map](image)
TC-dland

• Usage: `tc_dland`  
  `out_file`  
  [-grid_spec]  
  [-noll]  
  [-land file]  
  [-log file]  
  [-v level]

  ➢ This exe only needs to be run once to establish the NetCDF file!
  ➢ If running over the AL/EP and desire NHC land/water determination OR 1/10th degree grid global coverage: NetCDF file provided in build

<table>
<thead>
<tr>
<th><strong>out_file</strong></th>
<th>Indicates NetCDF output file containing the computed distances to land</th>
</tr>
</thead>
<tbody>
<tr>
<td>-grid_spec</td>
<td>Overrides the default 1/10th grid</td>
</tr>
<tr>
<td>-noll</td>
<td>Skips writing to reduce size of NetCDF file</td>
</tr>
<tr>
<td>-land file</td>
<td>Overwrites the default land data file</td>
</tr>
<tr>
<td>-log file</td>
<td>Outputs log messages to the specified file</td>
</tr>
<tr>
<td>-v level</td>
<td>Overrides the default level of verbosity (2)</td>
</tr>
</tbody>
</table>
TC-pairs

- Produces pair statistics on independent model input or user-specified consensus forecasts
- Matches forecast with reference TC dataset (most commonly Best Track Analysis)
- Pair generation can be subset based on user-defined filtering criteria
- ASCII pair output allows for new or additional analyses to be completed without performing full verification process

This tool is similar to **point_stat**: matched pair information!
# Tc_pairs

- **Input**: NetCDF gridded distance file, forecast/reference in ATCF format

- **Output**: TCSTAT format
  - Header, column-based ASCII output

- **Usage**: `tc_pairs`
  
  - `-adeck source` ATCF format file containing TC model forecast
  - `-bdeck source` ATCF format file containing TC reference dataset
  - `-config file` Name of configuration file to be used
  - `[-out base]` Indicates path of output file base
  - `[-log file]` Name of log file associated with pairs output
  - `[-v level]` Indicates desired level of verbosity
Tc_pairs

- Configuration file determines filtering criteria

<table>
<thead>
<tr>
<th>MODEL</th>
<th>VALID_MASK</th>
</tr>
</thead>
<tbody>
<tr>
<td>STORM_ID</td>
<td>CHECK_DUP</td>
</tr>
<tr>
<td>BASIN</td>
<td>INTERP_12</td>
</tr>
<tr>
<td>CYCLONE</td>
<td>CONSENSUS</td>
</tr>
<tr>
<td>STORM_NAME</td>
<td>LAG_TIME</td>
</tr>
<tr>
<td>INIT_BEG/END</td>
<td>BEST_BASELINE</td>
</tr>
<tr>
<td>INIT_INC/EXC</td>
<td>OPER_BASELINE</td>
</tr>
<tr>
<td>VALID_BEG/END</td>
<td>MATCH_POINTS</td>
</tr>
<tr>
<td>INIT_HR</td>
<td>DLAND_FILE</td>
</tr>
<tr>
<td>INIT_MASK</td>
<td>VERSION</td>
</tr>
</tbody>
</table>

➢ Take care not to over-subset!
Can perform additional filters with tc_stat tool

```c
// Model initialization time windows to include or exclude
// init_beg = "";
// init_end = "";
// init_inc = [];
// init_exc = [];
// Valid model time window
// valid_beg = "";
// valid_end = "";
// Model initialization hours
// init_hour = [];
// Lat/lon polylines defining masking regions
// init_mask = "";
// valid_mask = "";
// Specify if the code should check for duplicate ATCF lines when building tracks
// check_dup = FALSE;
// Specify whether special processing should be performed for interpolated models.
// interp12 = REPLACE;
// Specify how consensus forecasts should be defined:
// e.g.
// consensus = [
//   {
//     name = "CON1";
//     members = ['"MOD1", "MOD2", "MOD3"'];
//     required = [TRUE, FALSE, FALSE];
//     min_req = 2;
//   }
// ];
// consensus = [];
//```
### Tc_pairs

**Output in ASCII space delimited columns with header information**

<table>
<thead>
<tr>
<th>LEAD</th>
<th>VALID</th>
<th>INIT_MASK</th>
<th>VALID_MASK</th>
<th>LINE_TYPE</th>
<th>TOTAL</th>
<th>INDEX</th>
<th>LEVEL</th>
<th>WATCH_WARN</th>
<th>INITIALS</th>
<th>ALAT</th>
<th>ALON</th>
<th>BLAT</th>
<th>BLON</th>
<th>TK_ERR</th>
<th>X_ERR</th>
<th>Y_ERR</th>
<th>ALTK_ERR</th>
<th>CRTLK_ERR</th>
</tr>
</thead>
<tbody>
<tr>
<td>360000</td>
<td>20091106_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>12.60000</td>
<td>-82.30000</td>
<td>91.04526</td>
<td>46.95995</td>
<td>-78.00001</td>
<td>-80.83007</td>
<td>41.86590</td>
</tr>
<tr>
<td>480000</td>
<td>20091107_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>14.10000</td>
<td>-84.00000</td>
<td>170.34669</td>
<td>52.66857</td>
<td>-162.00005</td>
<td>-169.03902</td>
<td>20.81721</td>
</tr>
<tr>
<td>600000</td>
<td>20091108_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>15.40000</td>
<td>-83.90000</td>
<td>228.12263</td>
<td>52.49704</td>
<td>-221.99999</td>
<td>-221.96036</td>
<td>52.48623</td>
</tr>
<tr>
<td>720000</td>
<td>20091109_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>16.80000</td>
<td>-84.00000</td>
<td>315.20052</td>
<td>75.60019</td>
<td>-305.99997</td>
<td>-305.94534</td>
<td>75.58471</td>
</tr>
<tr>
<td>840000</td>
<td>20091110_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>18.70000</td>
<td>-84.10000</td>
<td>443.02103</td>
<td>121.62077</td>
<td>-426.00002</td>
<td>-435.51454</td>
<td>80.77160</td>
</tr>
<tr>
<td>1080000</td>
<td>20091112_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>21.90000</td>
<td>-86.20000</td>
<td>683.36364</td>
<td>315.83208</td>
<td>-605.99997</td>
<td>-678.00027</td>
<td>84.45908</td>
</tr>
<tr>
<td>1320000</td>
<td>20091114_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>27.90000</td>
<td>-88.60000</td>
<td>963.87555</td>
<td>425.51444</td>
<td>-920.00000</td>
<td>-907.04431</td>
<td>-190.24182</td>
</tr>
<tr>
<td>1560000</td>
<td>20091116_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>34.70000</td>
<td>-91.80000</td>
<td>1194.03580</td>
<td>557.51444</td>
<td>-1260.00000</td>
<td>-1247.04431</td>
<td>-350.24182</td>
</tr>
<tr>
<td>1800000</td>
<td>20091118_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>41.50000</td>
<td>-96.00000</td>
<td>1444.03580</td>
<td>689.51444</td>
<td>-1620.00000</td>
<td>-1607.04431</td>
<td>-510.24182</td>
</tr>
<tr>
<td>1920000</td>
<td>20091119_060000</td>
<td>NA</td>
<td>NA</td>
<td>TCMPR</td>
<td>11</td>
<td>1</td>
<td>TD</td>
<td>NA</td>
<td>NA</td>
<td>11.00000</td>
<td>-81.00000</td>
<td>44.90000</td>
<td>-98.10000</td>
<td>1569.03580</td>
<td>755.51444</td>
<td>-1800.00000</td>
<td>-1787.04431</td>
<td>-590.24182</td>
</tr>
</tbody>
</table>
TC Metrics

- **Track Error**: great-circle distance between the forecast location and the actual location of the storm center (nmi)
- **Along-track Error**: indicator of whether a forecasting system is moving a storm too slowly/quickly
- **Cross-track Error**: indicates displacement to the right/left of the observed track
- **Intensity Error**: Difference between forecast and actual intensity (kts)
**Tc_stat**

- Provides summary statistics and filtering jobs on TCST output

  ✓ **Filter job:**
  - Stratifies pair output by various conditions and thresholds

  ✓ **Summary job:**
  - Produces summary statistics on specific column of interest

- **Input:** TCST output from `tc_pairs`
- **Output:** TCST output file for either filter or summary job

---

**This tool is similar to stat_analysis: summarizes pairs (filter/summary jobs)!
**Tc_stat**

- **Usage:** `tc_stat`  
  `-lookin source`  
  [ `-out file` ]  
  [ `-log file` ]  
  [ `-v level` ]  
  [ `-config file` ] | [JOB COMMAND LINE]

- **-lookin source**  
  Location of TCST files generated from tc_pairs

- **-out file**  
  Desired name of output file

- **-log file**  
  Name of log file associated with tc_stat output

- **-v level**  
  Verbosity level

- **-config file**  
  Configuration file to be used

**Job command line**  
specify joblist on command line

» Configuration file options will be applied to every job, unless an individual job specifies a configuration option – joblist options will override
Tc_stat

Configuration file will filter TCST output from tc_pairs to desired subset over which statistics will be computed.

<table>
<thead>
<tr>
<th>AMODEL/BMODEL</th>
<th>INIT_MASK/VALID_MASK</th>
<th>LANDFALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>STORM_ID</td>
<td>LINE_TYPE</td>
<td>LANDFALL_BEG (END)</td>
</tr>
<tr>
<td>BASIN</td>
<td>TRACK_WATCH_WARN</td>
<td>MATCH_POINTS</td>
</tr>
<tr>
<td>CYCLONE</td>
<td>COLUMN_THRESH_NAME</td>
<td>EVENT_EQUAL</td>
</tr>
<tr>
<td>STORM_NAME</td>
<td>COLUMN_STR_NAME</td>
<td>EVENT_EQUAL_LEAD</td>
</tr>
<tr>
<td>INIT_BEG/INIT_END</td>
<td>INIT_THRESH_NAME</td>
<td>OUT_INIT_MASK</td>
</tr>
<tr>
<td>INIT_INC/INIT_EXC</td>
<td>INIT_STR_NAME</td>
<td>OUT_VALID_MASK</td>
</tr>
<tr>
<td>VALID_BEG/VALID_END</td>
<td>WATER_ONLY</td>
<td>JOBS [ ]</td>
</tr>
<tr>
<td>VALID_INC/VALID_EXC</td>
<td>RAPID_INTEN</td>
<td>VERSION</td>
</tr>
</tbody>
</table>

// Stratify by the ADECK and BDECK distances to land.
water_only = FALSE;
// Specify whether only those track points for which rapid intensification/weakening of the maximum wind speed occurred in the previous time step should be retained.
rapid_inten = {
  track = NONE; (NONE, ADECK, BDECK, BOTH)
  time = 24;
  exact = TRUE; (exact or max int. diff)
  thresh = >=30.0;
}
// Specify whether only those track points occurring near landfall should be retained, and define the landfall retention window as a number of seconds offset from the landfall time.
landfall = FALSE;
landfall_beg = -86400;
landfall_end = 0;
// Specify whether only those track points common to both the ADECK and BDECK tracks should be retained.
match_points = TRUE;
// Specify whether only those cases common to all models in the dataset should be retained.
event_equal = TRUE;
// Specify lead times that must be present for a track to be included in the event equalization logic.
event_equal_lead = ["12","24","36"];
Tc_stat

- TC_stat output similar to TC_pairs for filter job (TCSTAT)
- Summary job output
  - "-column" option produces summary statistics for the specified column
  - "-by" option can be used to search each unique entry in selected column

<table>
<thead>
<tr>
<th>Column number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SUMMARY: (job type)</td>
</tr>
<tr>
<td>2</td>
<td>Column (dependent parameter)</td>
</tr>
<tr>
<td>3</td>
<td>Case (storm + valid time)</td>
</tr>
<tr>
<td>4</td>
<td>Total</td>
</tr>
<tr>
<td>5</td>
<td>Valid</td>
</tr>
<tr>
<td>6-8</td>
<td>Mean including normal upper and lower confidence limits</td>
</tr>
<tr>
<td>9</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>10</td>
<td>Minimum value</td>
</tr>
<tr>
<td>11-15</td>
<td>Percentiles (10th, 25th, 50th, 75th, 90th)</td>
</tr>
<tr>
<td>16</td>
<td>Maximum Value</td>
</tr>
<tr>
<td>17</td>
<td>Sum</td>
</tr>
<tr>
<td>18-19</td>
<td>Independence time</td>
</tr>
<tr>
<td>20-23</td>
<td>Frequency of superior performance</td>
</tr>
</tbody>
</table>
Graphics tools

• Graphical capabilities are included in the MET-TC release
  • `plot_tcmpr.R`

• **Input:** TCSTAT tc_pairs output
• **Output:** R graphics, tc_stat logs/filter job TCSTAT (optional)

• **Usage:** `Rscript plot_tcmpr.R -lookin`
  • `-filter` (specify filter job)
  • `-config` (run filter job w/ configuration file)
    • Default Rscript configuration file included in release
Graphics tools-examples

Boxplots of Track Error by ADeck Model

Mean error w/ CIs

Rank frequency

Frequency of superior performance

Mean of ADeck Maximum Wind Speed – BDeck Maximum Wind Speed by ADeck Model

GRAPHICS

Mean error w/ CIs

Rank frequency

Percent of Cases

Lead Time (h)