

2014 WRF for Hurricanes Tutorial

January 16, 2014

College Park, MD

Forecast Verification

Kathryn M. Newman

National Center for Atmospheric Research
Developmental Testbed Center

Boulder, CO

Outline

- Introduction to Forecast Verification

- Introduction
- Observations
- Basic verification metrics
- Uncertainty & Confidence intervals
- MET overview

- MET-TC

- What is MET-TC?
- Getting Started
- TC-dland
- TC-pairs
- TC-stat
- graphics



Introduction to Forecast Verification



Introduction

- What is Verification?

- The process of comparing forecasts to relevant observations
- Measures quality of forecasts
- Evaluation of a particular model or condition

- Why Verify?

- Help understand model biases and performance of models under certain conditions
- Help users interpret forecasts
- Identify forecast weakness, strengths, differences



Introduction

- Verification goals depend on the questions we want to answer
 - Determines attribute to measure
 - Drives choices in verification statistics, measures, and graphics
- Before starting any verification study:
 1. **Identify multiple verification attributes** that provide answers to the questions of interest
 - ✓ *Position, wind, QPF, RI, landfall ...*
 2. **Select measures and graphics** to appropriately measure and represent the attributes of interest
 - ✓ *Track (along/cross) error, Intensity error, Contingency tables ...*
 3. **Identify a standard of comparison** that provides a reference level of skill
 - ✓ *CLIPER, SHIFOR, Baseline model ...*



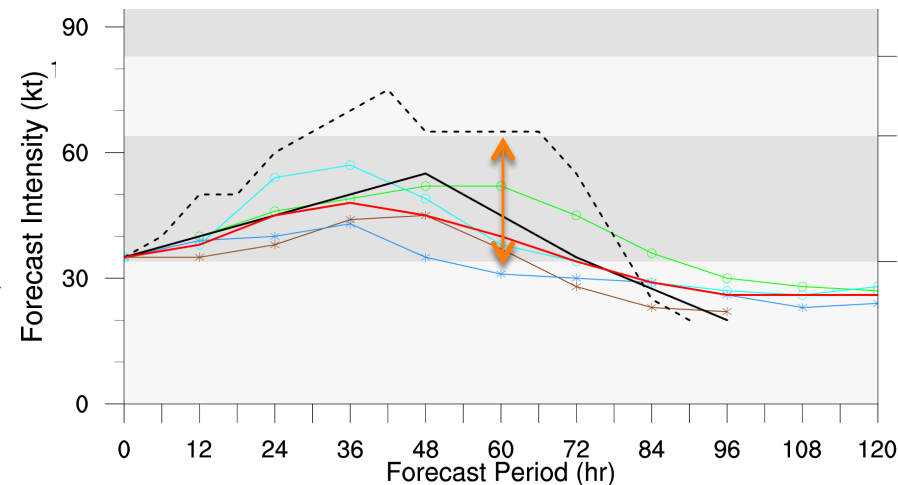
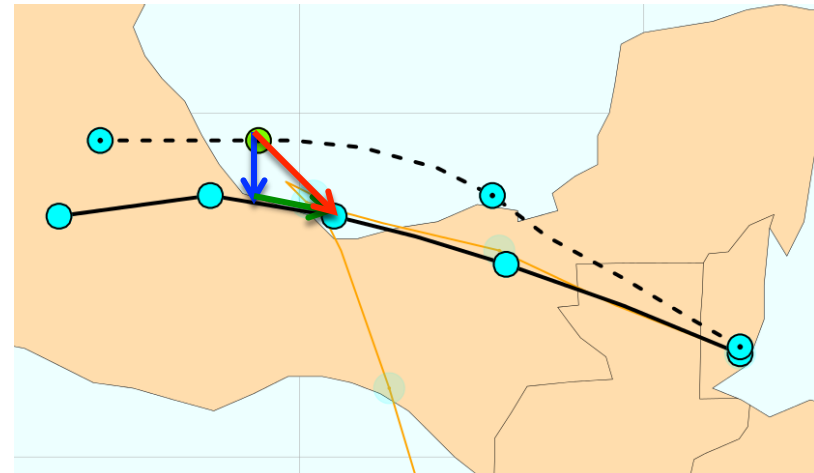
Observations

- Observations are an important consideration for TC verification
 - Quality and quantity of observations available
 - Typically sparse or intermittent
 - May infer characteristics from indirect measures (satellite)
- Refer to previous lecture for more detail on observational datasets
- 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

AL, 09, 2011082218,	, BEST,	0, 193N,	680W,	75,	988,	HU,	64,	NEQ,	25,	0,	0,	0,	1010,	250,	15,	85,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	60,	0,	120
AL, 09, 2011082300,	, BEST,	0, 197N,	688W,	80,	981,	HU,	34,	NEQ,	160,	70,	50,	100,	1010,	300,	15,	105,	30,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	60,	0,	120
AL, 09, 2011082300,	, BEST,	0, 197N,	688W,	80,	981,	HU,	50,	NEQ,	70,	30,	30,	70,	1010,	300,	15,	105,	30,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	60,	0,	120
AL, 09, 2011082300,	, BEST,	0, 197N,	688W,	80,	981,	HU,	64,	NEQ,	25,	0,	0,	35,	1010,	300,	15,	105,	30,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	60,	0,	120
AL, 09, 2011082306,	, BEST,	0, 201N,	697W,	80,	978,	HU,	34,	NEQ,	180,	120,	90,	130,	1010,	300,	15,	105,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	120,	35,	150
AL, 09, 2011082306,	, BEST,	0, 201N,	697W,	80,	978,	HU,	50,	NEQ,	90,	60,	40,	70,	1010,	300,	15,	105,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	120,	35,	150
AL, 09, 2011082306,	, BEST,	0, 201N,	697W,	80,	978,	HU,	64,	NEQ,	45,	30,	20,	35,	1010,	300,	15,	105,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	120,	35,	150
AL, 09, 2011082312,	, BEST,	0, 204N,	706W,	80,	978,	HU,	34,	NEQ,	180,	120,	90,	130,	1008,	300,	15,	105,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	120,	35,	150
AL, 09, 2011082312,	, BEST,	0, 204N,	706W,	80,	978,	HU,	50,	NEQ,	90,	60,	40,	70,	1008,	300,	15,	105,	0,	L,	0,	,	0,	0,	IRENE,	D,	12,	NEQ,	180,	120,	35,	150

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)

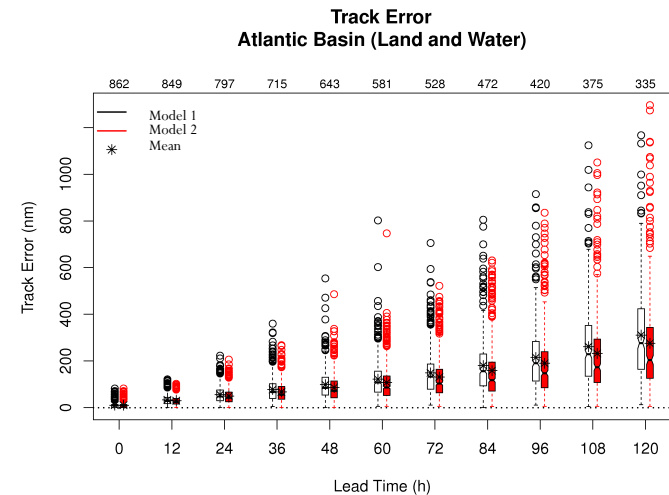
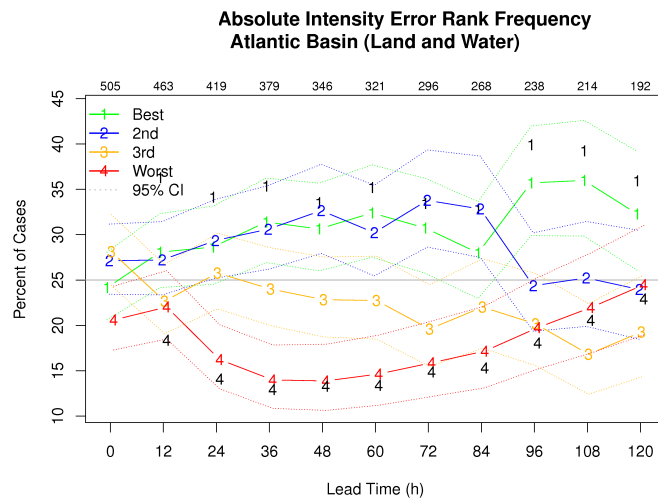


Graphics courtesy of NCAR TCMT

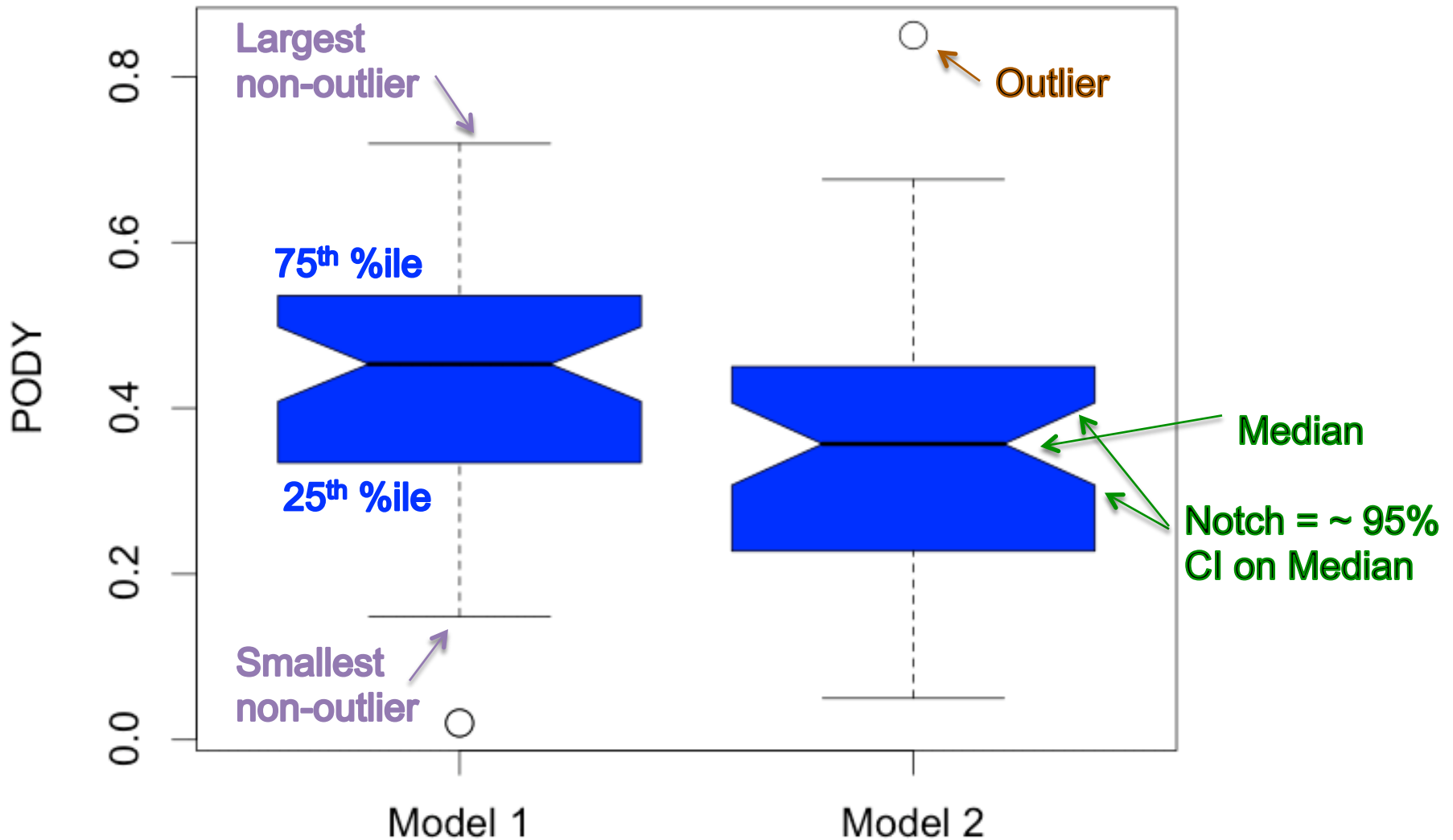


TC metrics

- **Skill Scores:** Used as a standard of comparison, skill diagrams are often used to compare model skill relative to CLIPER/SHIFOR
- **Frequency of Superior Performance:** ranking a particular model forecast relative to the performance multiple model forecasts
- **Distribution of errors:** Box plots can be used to highlight the distributions of the errors in the forecasts



Box (& Whisker) Plot



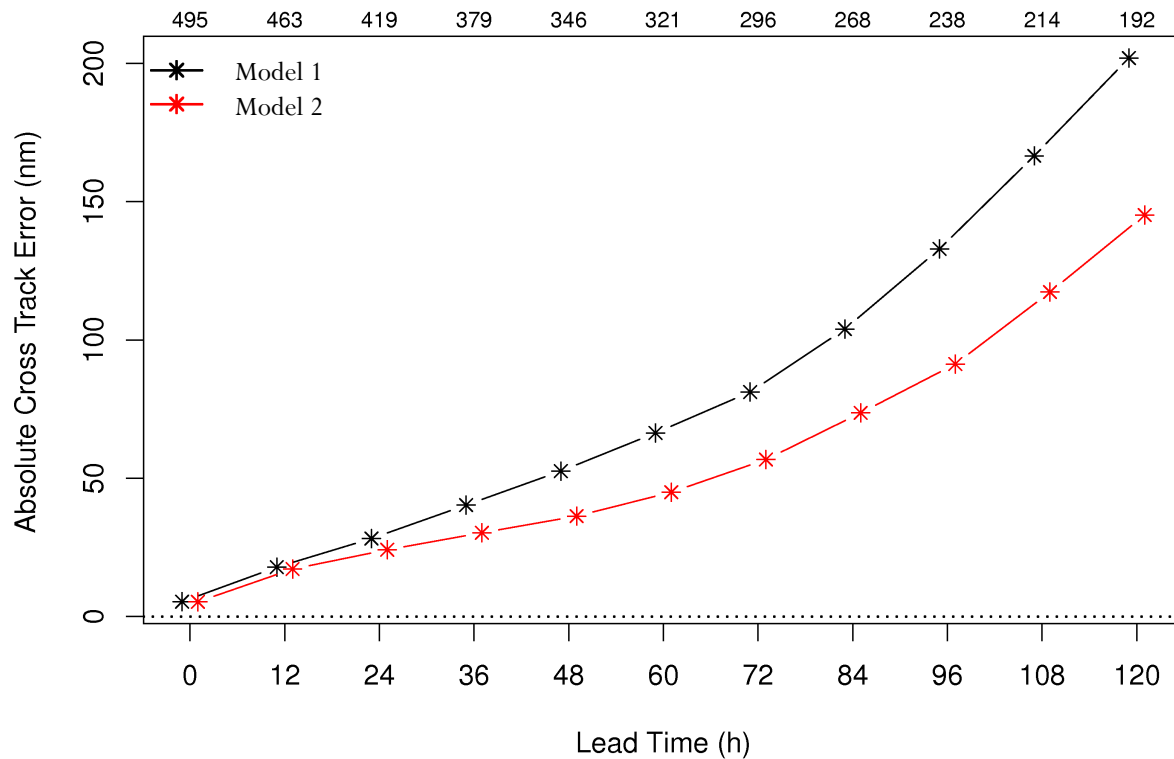
Uncertainty

- Observations and analysis products as well as models themselves are subject to uncertainty
- Need to be aware of sample size!
 - Typically smaller samples due to lower frequency of occurrence relative to other weather phenomena
- Accounting for sampling uncertainty:
 - Verification statistic is a realization of a random process
 - What if the experiment were re-run under identical conditions?
Would you get the same answer?



Confidence intervals

**Mean Absolute Cross Track Error
Atlantic Basin (Land and Water)**

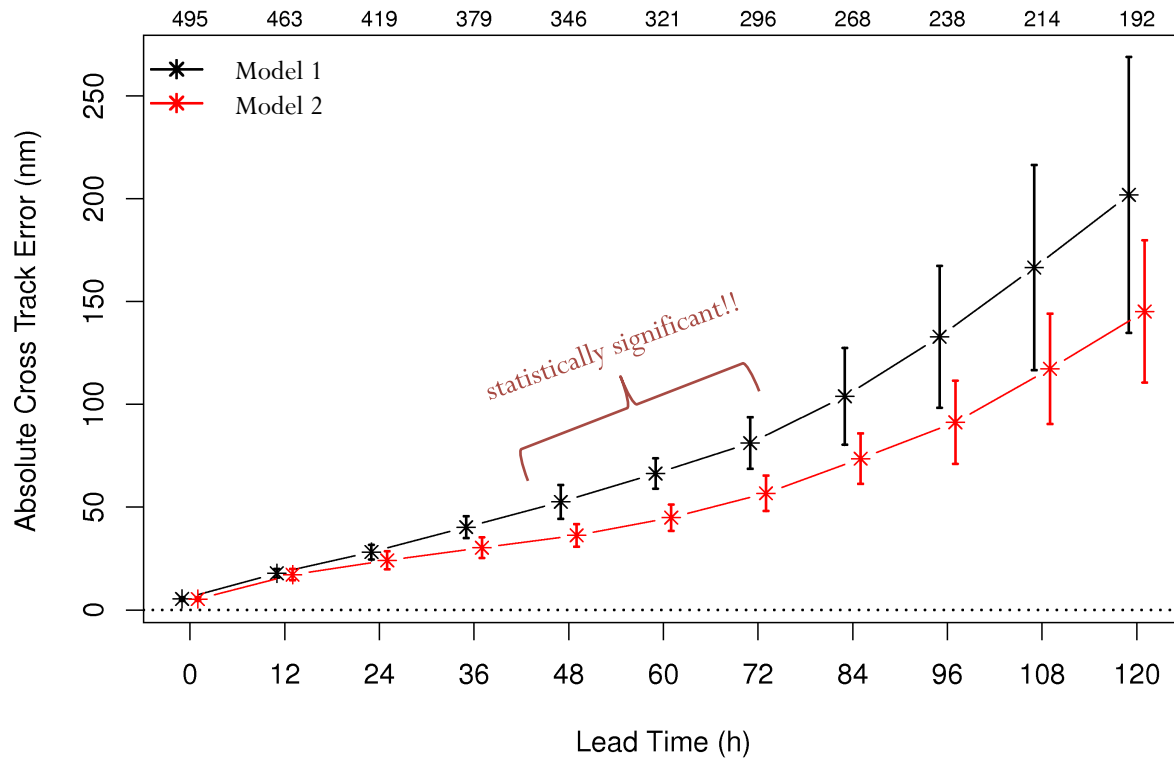


Mean absolute cross-track errors for two models.

Scores are very similar at short lead times, but seem to diverge at longer lead times

Confidence intervals

Mean Absolute Cross Track Error
Atlantic Basin (Land and Water)

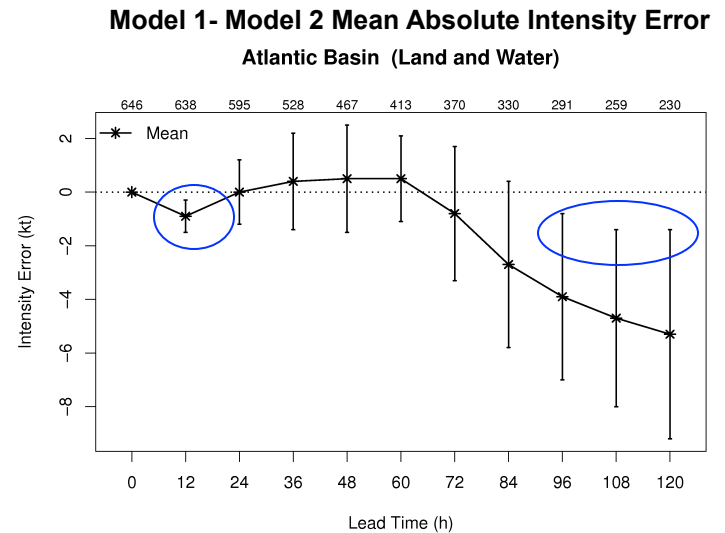
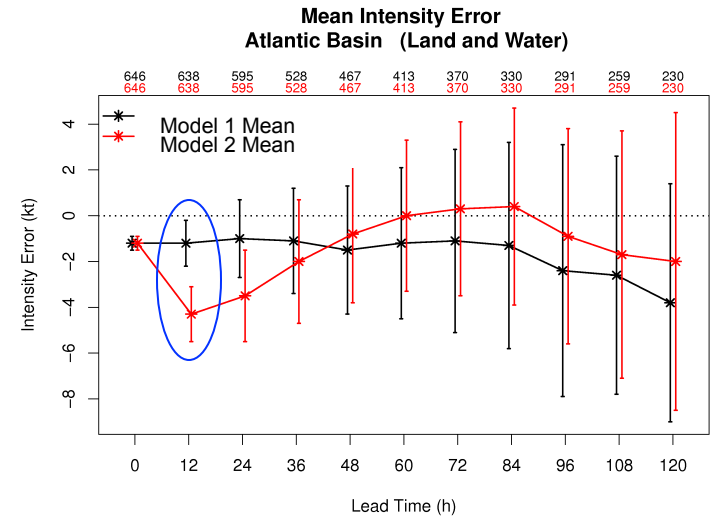


Confidence Intervals (CIs) indicate no significant difference between 0-36 h, after 84 h

Statistical significance indicated where CIs don't overlap

Confidence Intervals

- Two ways to examine scores:
 - CI about absolute scores
 - May be difficult to differentiate model performance differences
 - SS where two model CIs do not intersect
 - CI about Pairwise Differences
 - May allow for differentiation of model performance.
 - SS where CIs do not encompass 0

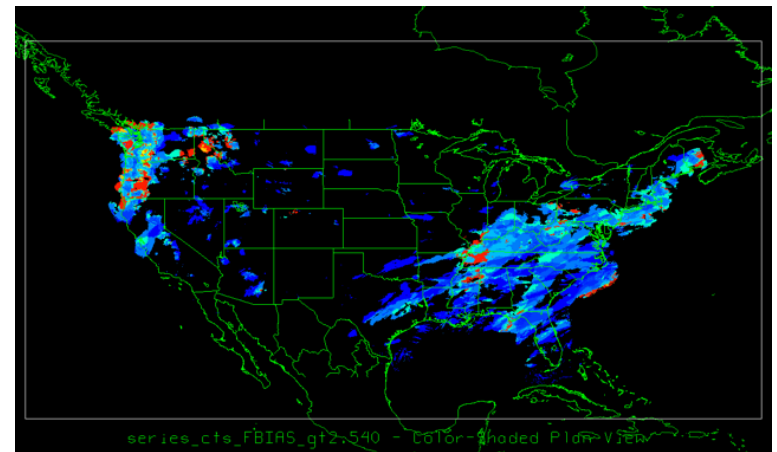


➤ For MET-TC CIs are computed using the assumption of normality for the mean



Model Evaluation Tools

- What is MET?
 - MET is a set of tools for evaluating model forecasts
- A modular set of forecast evaluation tools
 - Freely available, highly configurable, fully documented, supported
- MET includes:
 - Reformatting tools
 - Statistical tools
 - Analysis tools
- MET works directly with post-processed model output to perform a large variety of statistical analyses



Precipitation frequency bias generated from MET output

Model Evaluation Tools

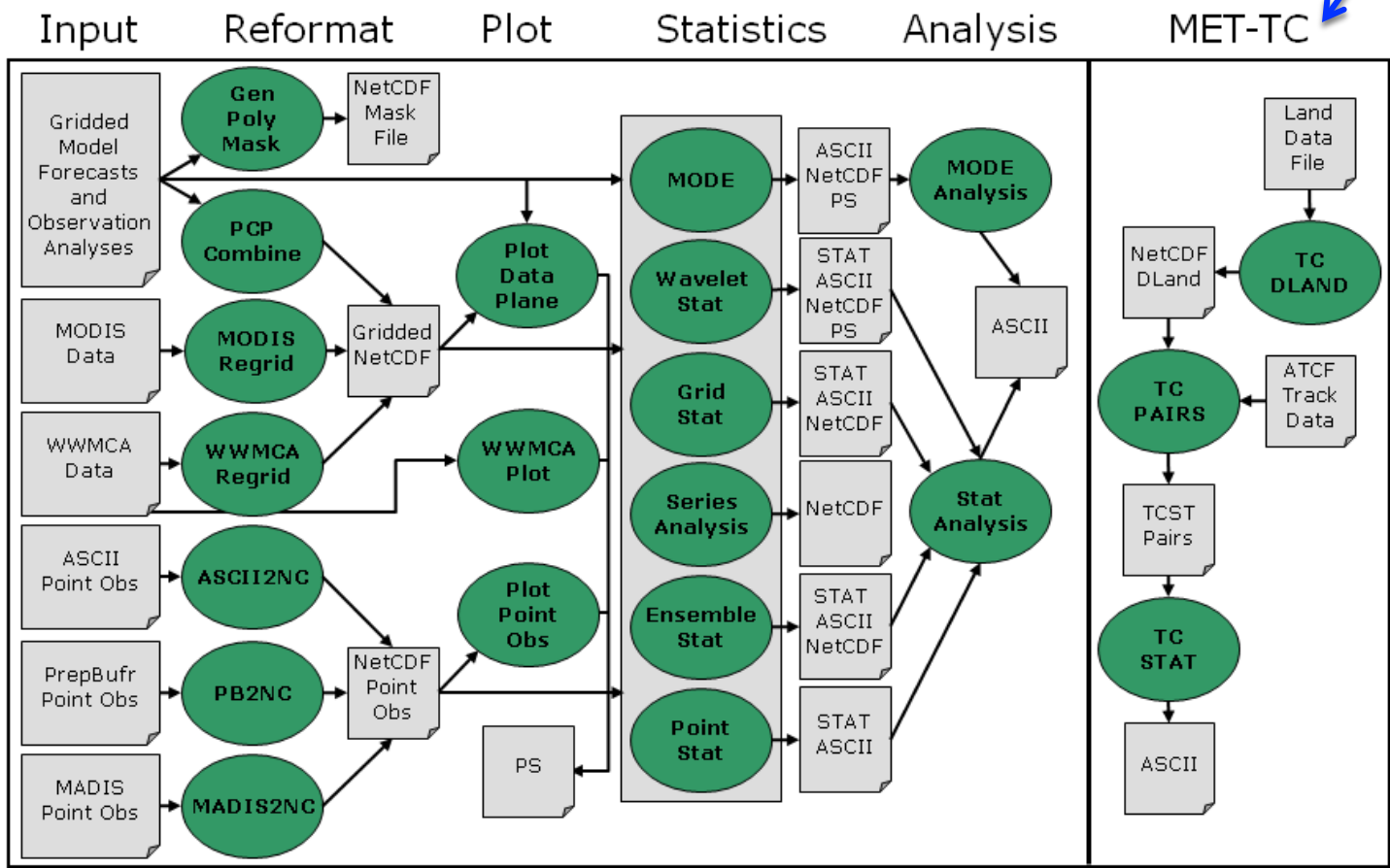
- Overview of tools
- MET provides a variety of verification techniques:
 - Gridded model data to point-based observations
 - Gridded model data to gridded observations
 - Ensemble and probabilistic verification methods
 - Aggregating output through time and space



Model Evaluation Tools

MET-TC covered in detail for this talk

- Overview of tools



References & Further Reading

- Gilleland, E., 2010: Confidence intervals for forecast verification. NCAR Technical Note NCAR/TN-479+STR, 71pp. *Available at:*
<http://nldr.library.ucar.edu/collections/technotes/asset-000-000-000-846.pdf>
- Jolliffe and Stephenson (2011): Forecast verification: A practitioner's guide, 2nd Edition, Wiley & sons
- JWGFVR (2009): Recommendation on verification of precipitation forecasts. WMO/TD report, no.1485 WWRP 2009-1
- Nurmi (2003): Recommendations on the verification of local weather forecasts. ECMWF Technical Memorandum, no. 430
- Wilks (2006): Statistical methods in the atmospheric sciences, ch. 7. Academic Press
- See also: <http://www.cawcr.gov.au/projects/verification/>

Appendix C of MET Documentation: <http://www.dtcenter.org/met/users/docs/overview.php>



Model Evaluation Tools–Tropical Cyclone



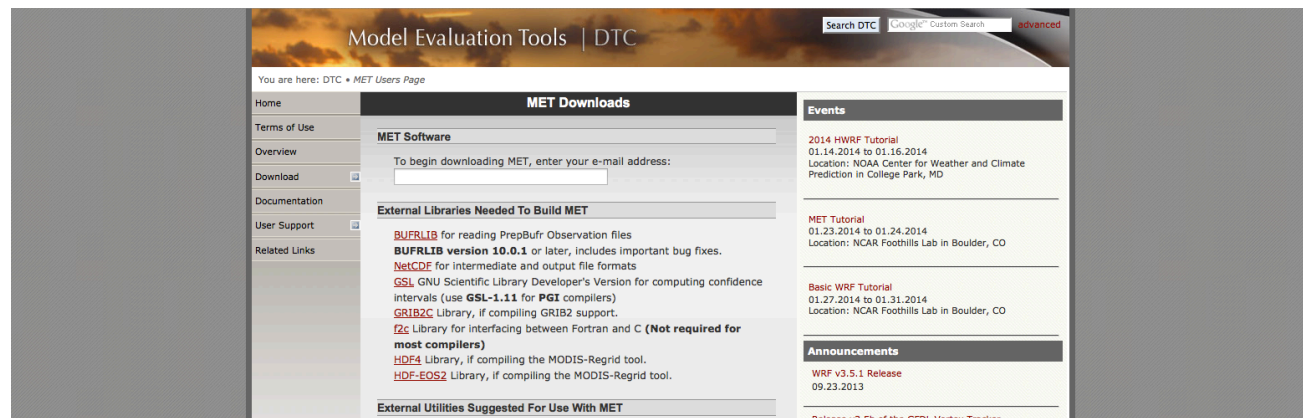
Introduction

- 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 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 a standard set of verification metrics and comprehensive output statistics
 - Available to all users
 - Enables consistent forecast evaluation studies to be undertaken across the community



Compile & Build

- Download MET (must be v4.1) release and compile locally
 - Need to register to download: www.dtcenter.org/met/users
- Supported platforms and compilers
 1. Linux with GNU compilers
 2. Linux with Portland Group (PGI) compilers
 3. Linux with Intel compilers
 4. IBM machines with IBM compilers



The screenshot shows the 'MET Downloads' page on the Developmental Testbed Center (DTC) website. The page is titled 'Model Evaluation Tools | DTC' and includes a search bar. The main content area is divided into several sections:

- MET Software:** A form to begin downloading MET, requiring an e-mail address.
- External Libraries Needed To Build MET:** A list of required libraries and their versions:
 - BUFRLIB** for reading PrepBuf: Observation files
 - BUFRLIB version 10.0.1** or later, includes important bug fixes.
 - NetCDF** for intermediate and output file formats
 - GSL** GNU Scientific Library Developer's Version for computing confidence intervals (use **GSL-1.11** for **PGI** compilers)
 - GRIB2C** Library, if compiling GRIB2 support.
 - FX** Library for interfacing between Fortran and C (**Not required for most compilers**)
 - HDF4** Library, if compiling the MODIS-Regrid tool.
 - HDF-EOS2** Library, if compiling the MODIS-Regrid tool.
- External Utilities Suggested For Use With MET:** A section for additional tools.
- Events:** A list of upcoming events, including the 2014 HWRP Tutorial and the MET Tutorial.
- Announcements:** A section for recent releases, including WRF v3.5.1 Release and Release v3.5h of the CFM1 Vortex Tracker.



Compile & build

- Dependencies
- Required:
 - GNU Make Utility
 - C++/Fortran compilers (GNU, PGI, Intel, or IBM)
 - Unidata's NetCDF Library
 - NCEP's BUFRLIB Library
 - GNU Scientific Library (GSL)
- Optional/Recommended:
 - R statistics and graphics package



Compile & build

- The build will include MET and MET-TC
- MET-TC specific code and tools:
 - **bin/** : executables for each MET-TC module (`tc_dland`, `tc_pairs`, `tc_stat`)
 - **data/config/** : configuration files (`TCPairsConfig_default`, `TCStatConfig_default`)
 - **data/tc_data/** : static files used in MET-TC (`aland.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...

- Model output must be run through an internal/external vortex tracking algorithm (*GFDL vortex tracker – previous lecture*)
- The input files must be in Automated Tropical Cyclone Forecasting System (ATCF) format.
 - Must adhere to for MET-TC tools to properly parse the input data (first 17 columns required)

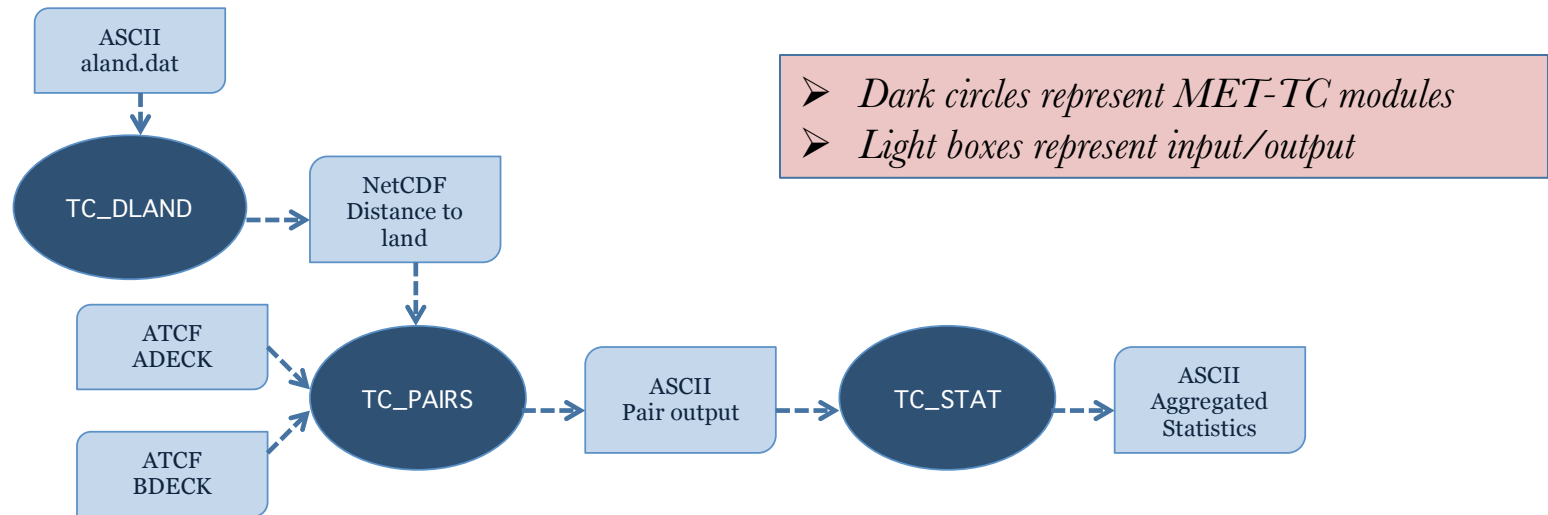
For detailed information on ATCF format: http://www.nrlmry.navy.mil/atcf_web/docs/database/new/abdeck.txt

- 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/>

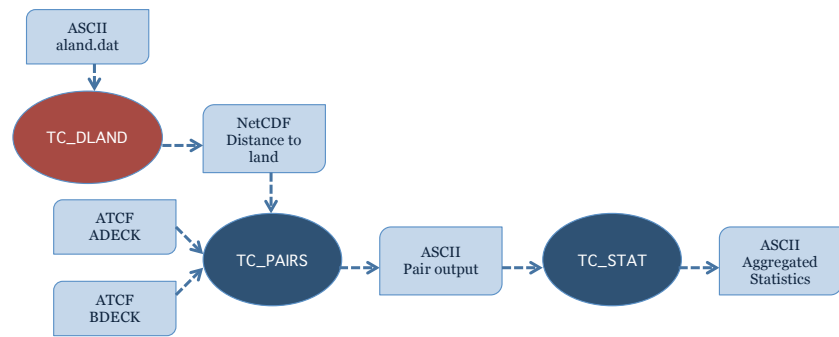


MET-TC components

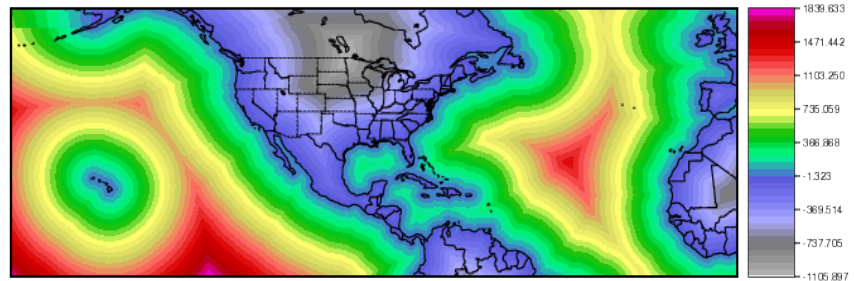


- Primary functions of the code are:
 - Compute pair statistics from ATCF input files
 - Filter pair statistics based on user specifications
 - Compute summary statistics

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



TC-dland

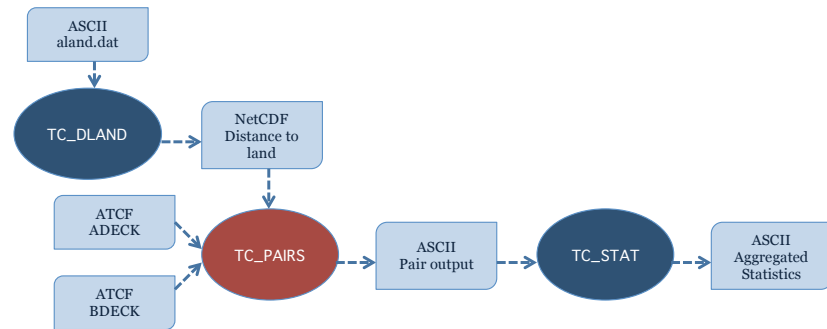
- Usage: **tc_dland**
out_file
[-grid_spec]
[-noll]
[-land file]
[-log file]
[-v level]

- 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:
NetCDF file in build

out_file	Indicates NetCDF output file containing the computed distances to land
-grid_spec	Overrides the default 1/10 th grid
-noll	Skips writing to reduce size of NetCDF file
-land file	Overwrites the default land data file
-log file	Outputs log messages to the specified file
-v level	Overrides the default level of verbosity (2)



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

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

-bdeck source

-config file

[-out base]

[-log file]

[-v level]

-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

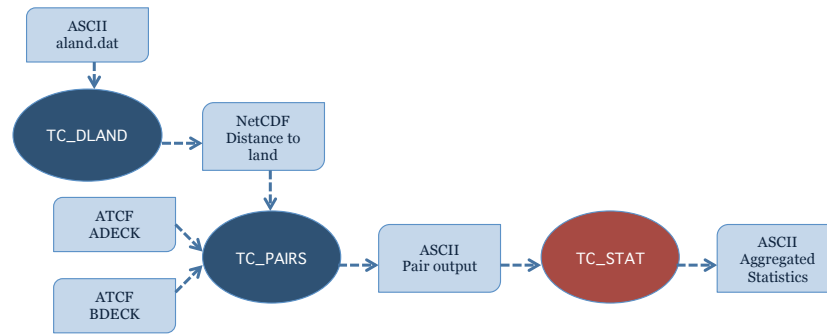
MODEL	VALID_MASK
STORM_ID	CHECK_DUP
BASIN	INTERP_12
CYCLONE	CONSENSUS
STORM_NAME	LAG_TIME
INIT_BEG/INIT_END	BEST_BASELINE
INIT_INC/INIT_EXC	OPER_BASELINE
VALID_BEG/VALID_END	MATCH_POINTS
INIT_HR	DLAND_FILE
INIT_MASK	VERSION

```
//  
// 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  
//  
check_dup = FALSE;  
  
//  
// Specify if special processing should  
// be performed for interpolated models  
// with  
// 12-hour spacing.  
//  
interp12 = TRUE;
```

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



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

Tc_stat

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

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

-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



Tc_stat

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

AMODEL/BMODEL	INIT_MASK/VALID_MASK	LANDFALL
STORM_ID	LINE_TYPE	LANDFALL_BEG (END)
BASIN	TRACK_WATCH_WARN	MATCH_POINTS
CYCLONE	COLUMN_THRESH_NAME (VAL)	EVENT_EQUAL
STORM_NAME	COLUMN_STR_NAME (VAL)	OUT_INIT_MASK
INIT_BEG/INIT_END	COLUMN_STR_NAME (VAL)	OUT_VALID_MASK
INIT_INC/INIT_EXC	INIT_THRESH_NAME (VAL)	JOBS []
VALID_BEG/VALID_END	INIT_STR_NAME (VAL)	VERSION
VALID_INC/VALID_EXC	WATER_ONLY	
INIT_HR/VALID_HR/LEAD	RAPID_INTEN (THRESH)	

```
//
// Stratify by the ADECK and BDECK
// distances to land.
//
water_only = FALSE;

//
// Specify whether only those track
// points for which rapid intensification
// occurred in the BDECK track between
// the current time and 24-hours prior
// should be retained.
//
rapid_inten      = FALSE;
rapid_inten_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. May
// modify using the "-match_points" job
// command
// option.
//
match_points = TRUE;

//
// Specify whether only those cases
// common to all models in the dataset
// should
// be retained.
//
event_equal = TRUE;
```



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

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

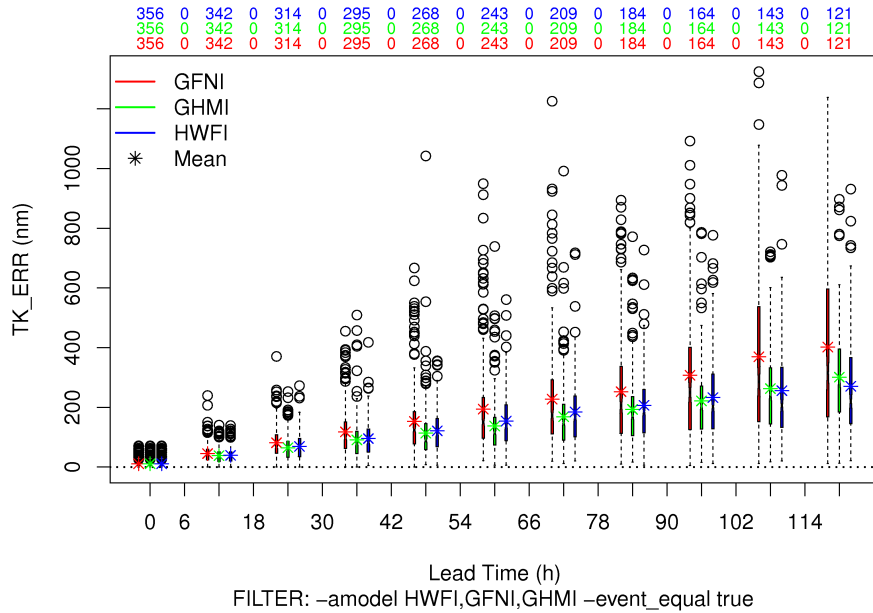
Graphics tools

- 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

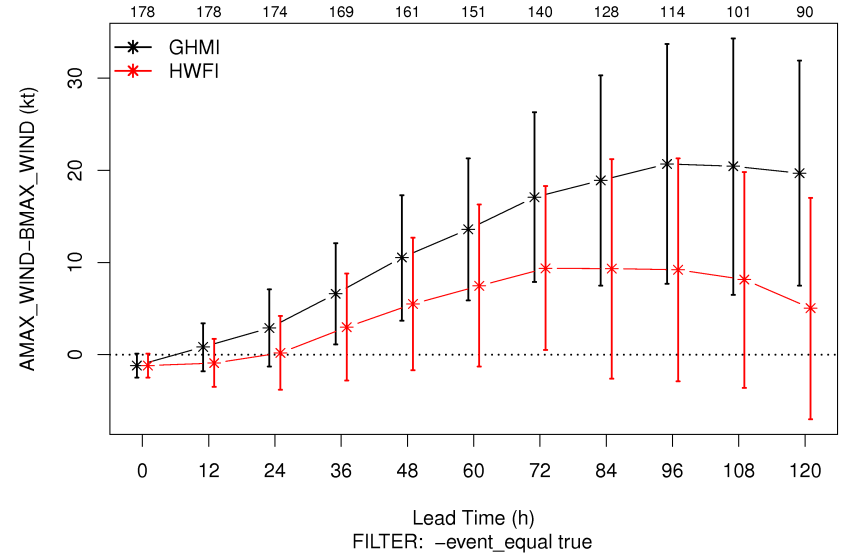


Graphics tools

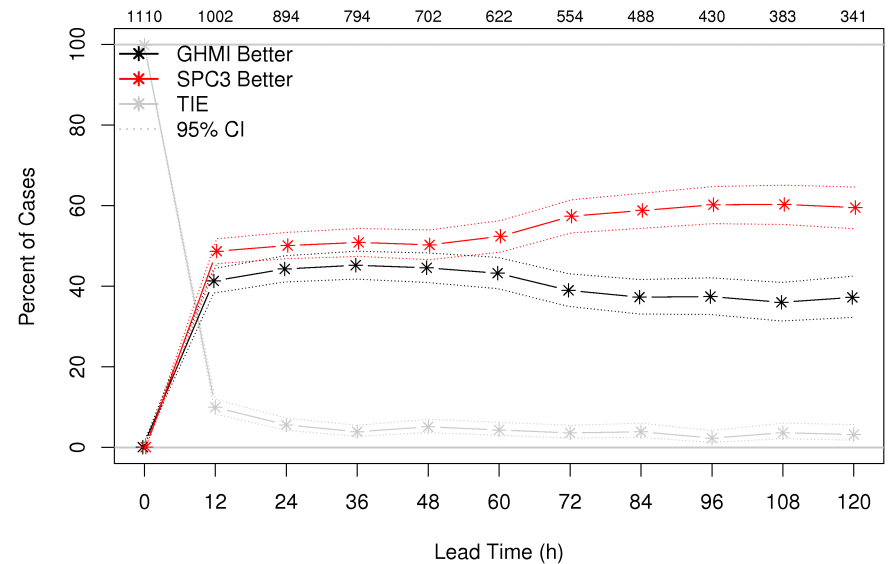
Boxplots of Track Error by ADeck Model



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



Absolute Intensity Error Difference >=1 kt Atlantic Basin (Land and Water)



- For code download and user's guide:

www.dtcenter.org/met/users

- Contact for questions, help, comments:

met_help@ucar.edu

