



HRH Test Evaluation Suite

Developmental Testbed Center

May 7, 2009

Verification Measures in Test Plan - I

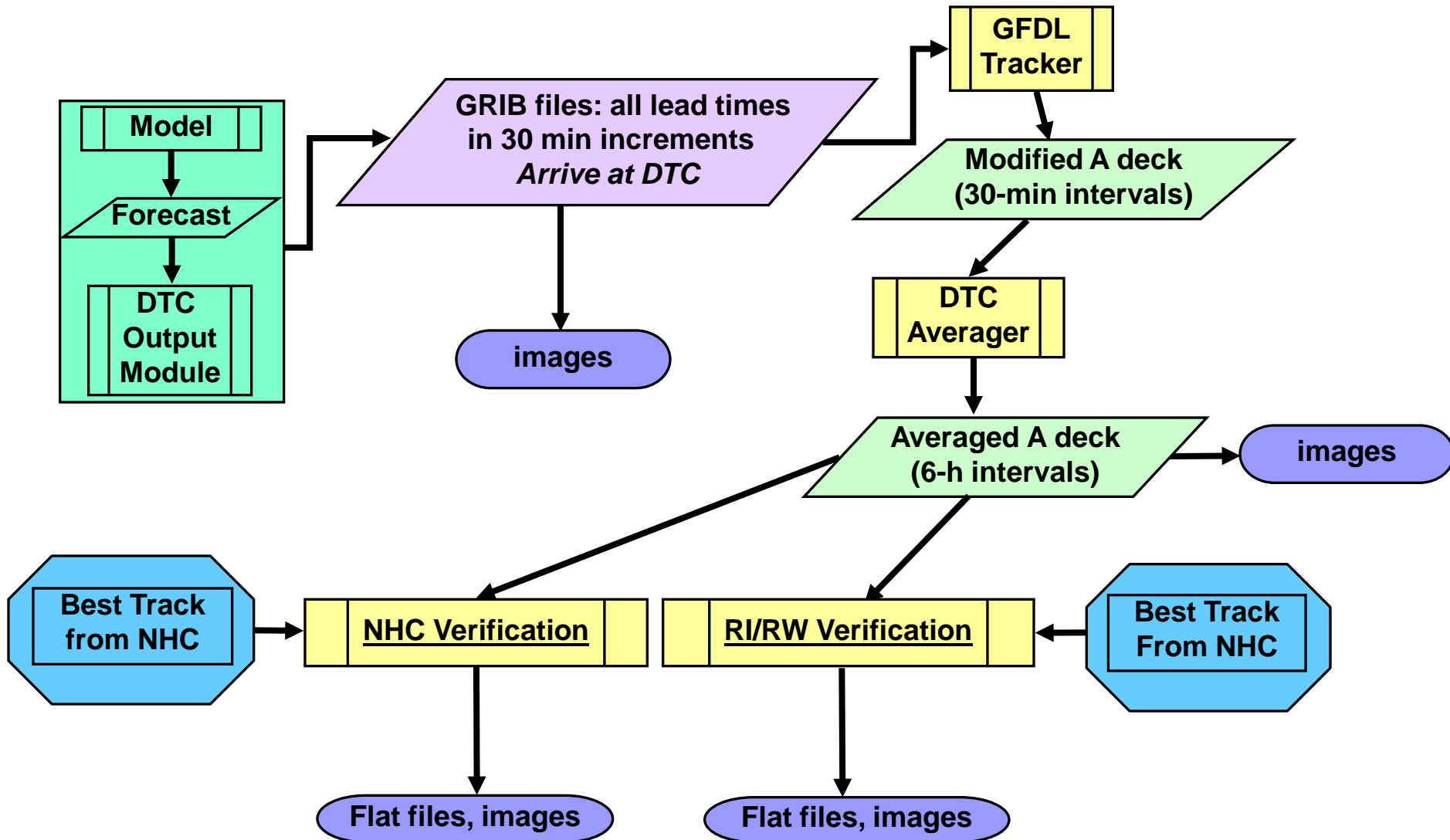
Name	Statistical methods and metrics	Feas	Prior
Center Position	MAE, error distribution, along- and cross-track MAE and bias; MAE skill scores	High	High
Intensity	MAE, bias, error distribution, MAE skill score	High	High
Occur RI	2-D contingency table, POD, FAR, CSI, skill scores	High	High
Occur RW	2-D contingency table, POD, FAR, CSI, skill scores	High	High

Green	Computed and presented here
Black	Computed but will not be presented
Purple	Computation under development
Red	Considering abandoning

Verification Measures in Test Plan - II

Name	Statistical methods and metrics	Feas	Prior
Onset RI	Timing error MAE, bias, MSE, RMSE, RMSE skill score, error distribution	High	High
Onset RW	Timing error MAE, bias, MSE, RMSE; MAE and RMSE skill scores, error distribution	High	High
Wind radii	MAE, bias, RMSE, error distribution; MAE and RMSE skill scores	High	Med
RMW	MAE, bias, RMSE, error distributions; MAE and RMSE skill scores	Med	Med
Consistency	Correlation, Distributions of changes over time	Med	High

DTC Evaluation System for HRH



Averager

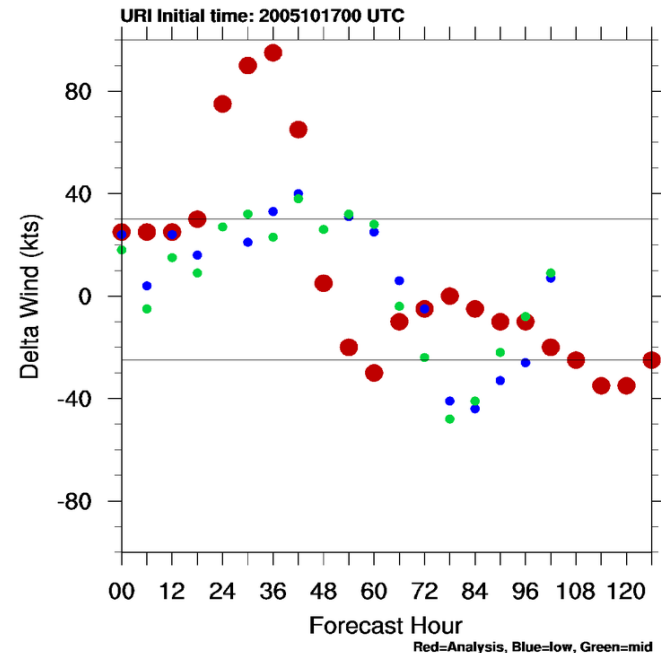
Issue: verify representative maximum wind (not instantaneous)

After discussions in entire HRH group, it was decided to:

- Output forecasts every 30 minutes
- Run tracker every 30 minutes
- Compute a running mean of the max wind over a 2h window:
$$V_{\text{mean}}(t) = [V(t+60) + V(t-30) + V(t) + V(t+30) + V(t+60)] / 5$$
- Use V_{mean} for verification of maximum wind

DTC RI / RW Verification

- RI – increase 30 kt in 24 h
- RW – decrease 25 kt in 24 h over water
- Events are identified in Best Track and forecast track
- A RI or RW match occurs when event happens at the same time in forecast and best track
- Future development: more sophisticated and informative match
- 1 contingency table for all lead times due to small sample size



Contingency Table for RI/RW

		O	
		Y	N
F	Y	a	b
	N	c	d

$$HR = (a+d)/n$$

$$POD = a/(a+c)$$

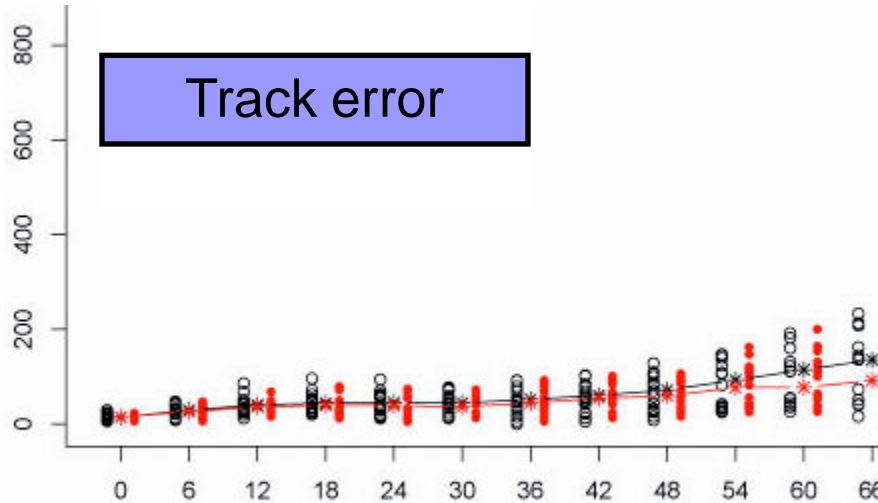
$$CSI = a/(a+b+c)$$

$$FAR = b/(a+b)$$

NHC Verification System

- Ingests Best Track and forecast track
- Processed each case separately to generate:
 - For each forecast lead time
 - Track error (vector and magnitude)
 - Intensity error
 - Wind radii error
- Output for all cases was ingested in R-language script to
 - Compute mean, median, quartiles, outliers
 - Plot distribution as a boxplot - when sample size is large enough

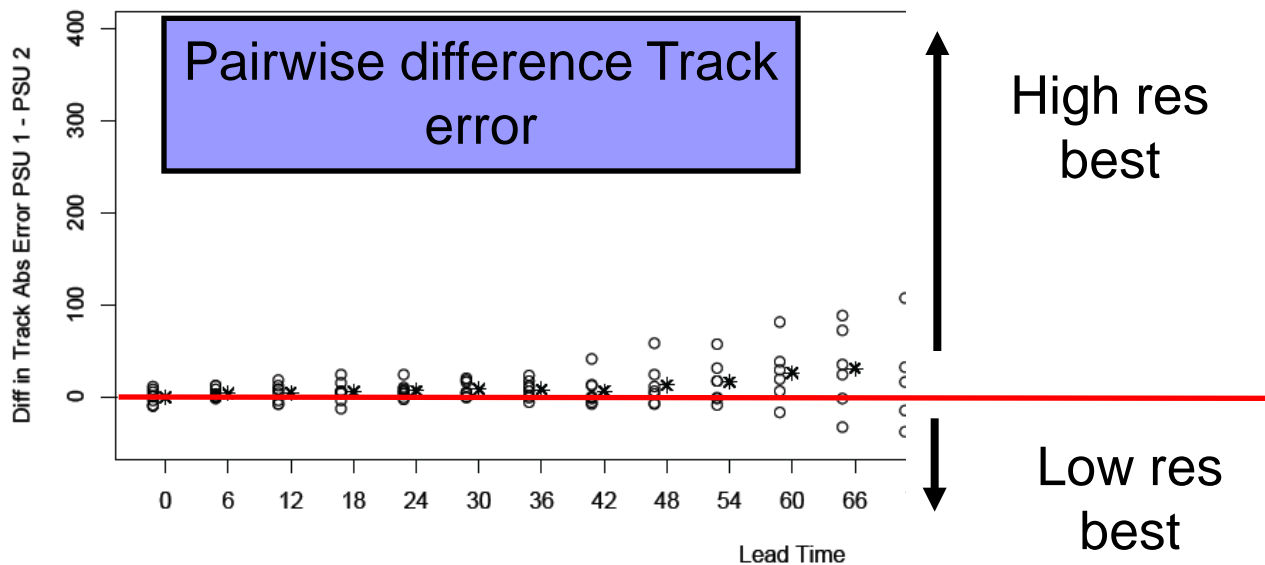
Distribution and differences



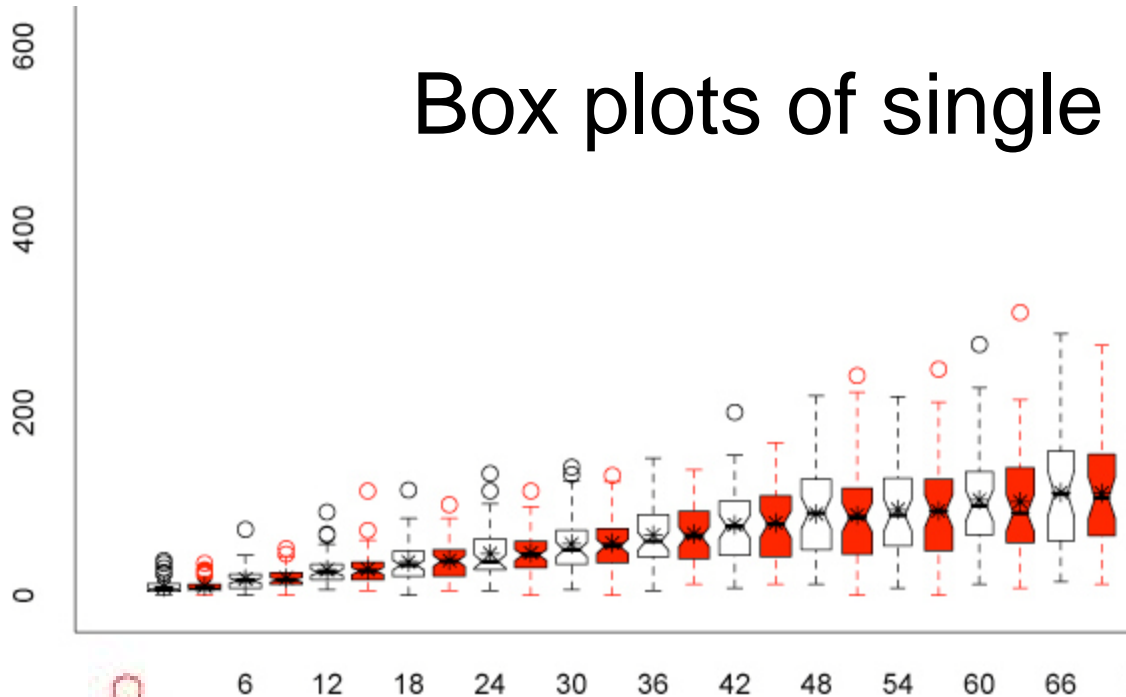
Lower res = black
Higher res = red

Pairwise difference for each case and lead time: error (low res) – error (high res)

Low res - High res



Box plots of single models



Lower res = black

Higher res = red



Median: bold waist

Mean: star

95% CI on median: notch

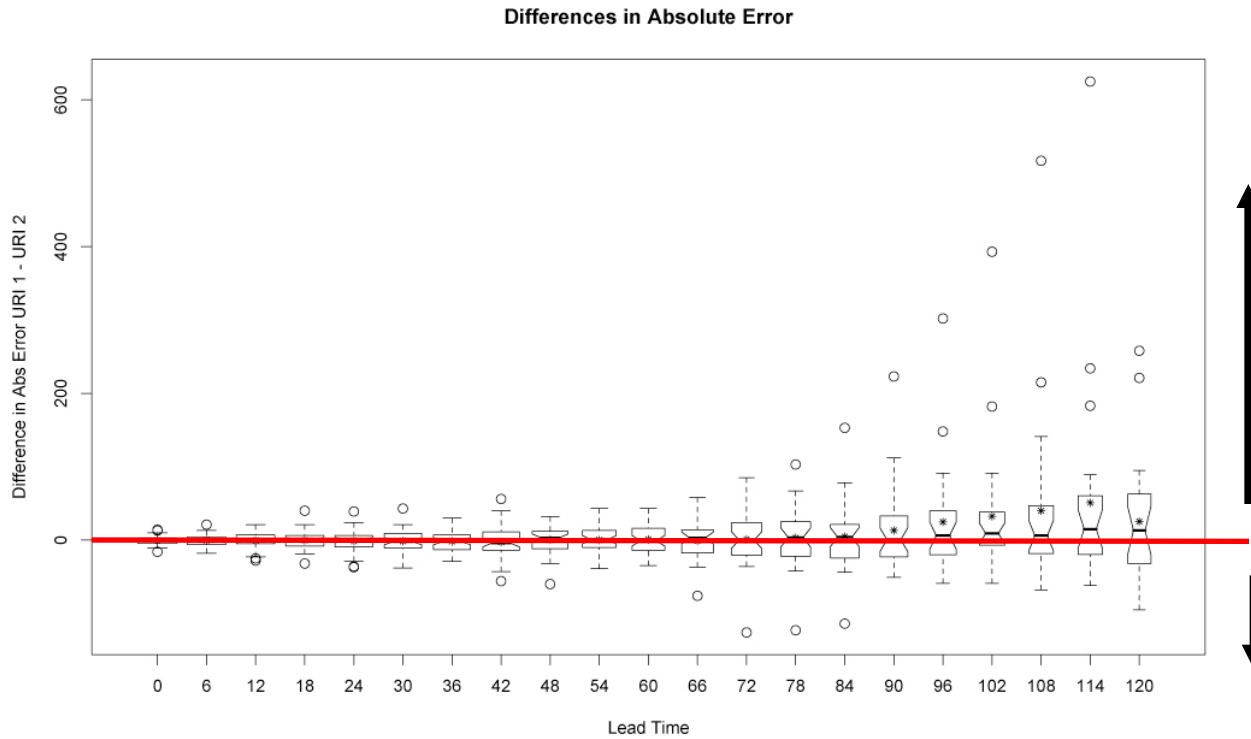
Sample size: width of box

25% and 75% quartiles: bottom and top of box

Length of whiskers: furthest point from median that is not an outlier.

Outliers: points further away from median than $1.5 * \text{IQR}$ (circles)

Box plots of differences



High Res
best

Low res
best