

# Analysis of Hurricane Forecasting Errors

ASQ New Orleans Chapter Meeting Jan '09

Matthew Barrows

Process Engineering Mfg. Technologist

Six Sigma Master Black Belt

Monsanto Company - Luling, Louisiana

# Hurricane Forecasting

- Seasonal Forecast Accuracy
- Track Forecasting
  - Understanding track forecasting error
  - An alternate “error cone” methodology

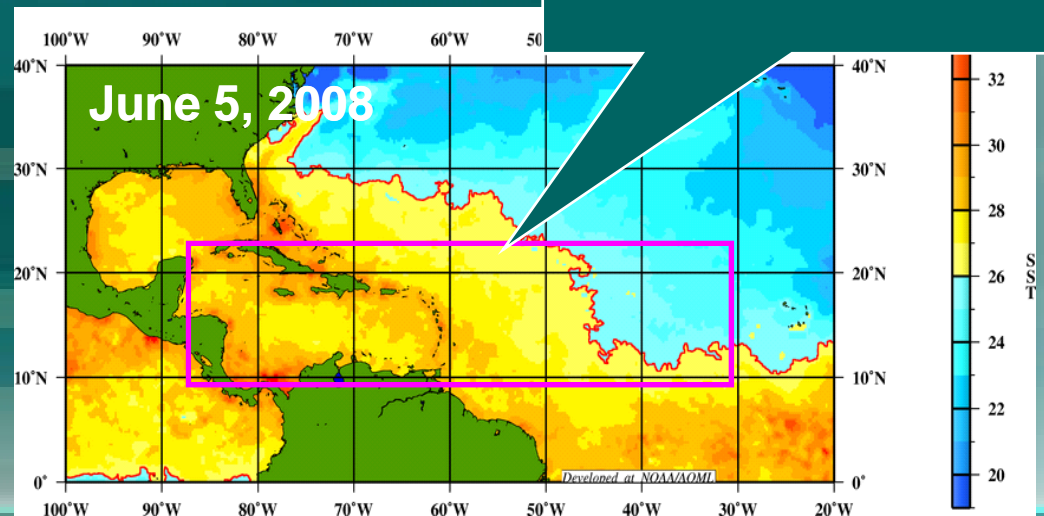
*“Never, no matter what may be the progress of science, will honest scientific men who have regard for their reputations venture to predict the weather.”*

*~ Arago’s Admonition*

# Seasonal Forecasts

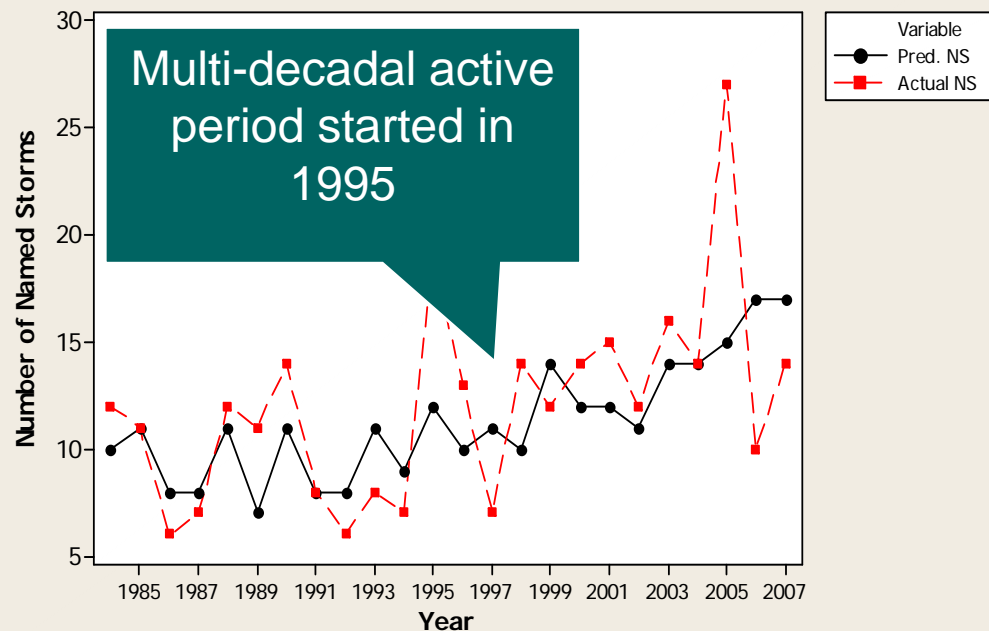
- Forecasting what the upcoming season will look like:
  - How many storms, how strong, how long.
  - Based on key hurricane climate predictors:
    - Sea Surface Temperatures (SST)
    - Wind shear (El Nino/La Nina)
    - Other (pressure, dust, etc.)
  - Virtually useless before May/June (so far)
- Issued by a few organizations:
  - Dr. Gray (& Dr. Klotzbach) – Since 1984.
  - NHC/NOAA – Since 1999.
  - Other private companies.

Sea Surface Temp in the Main Development Region (MDR) needs to be at least 26 C (80F) to support hurricane development.

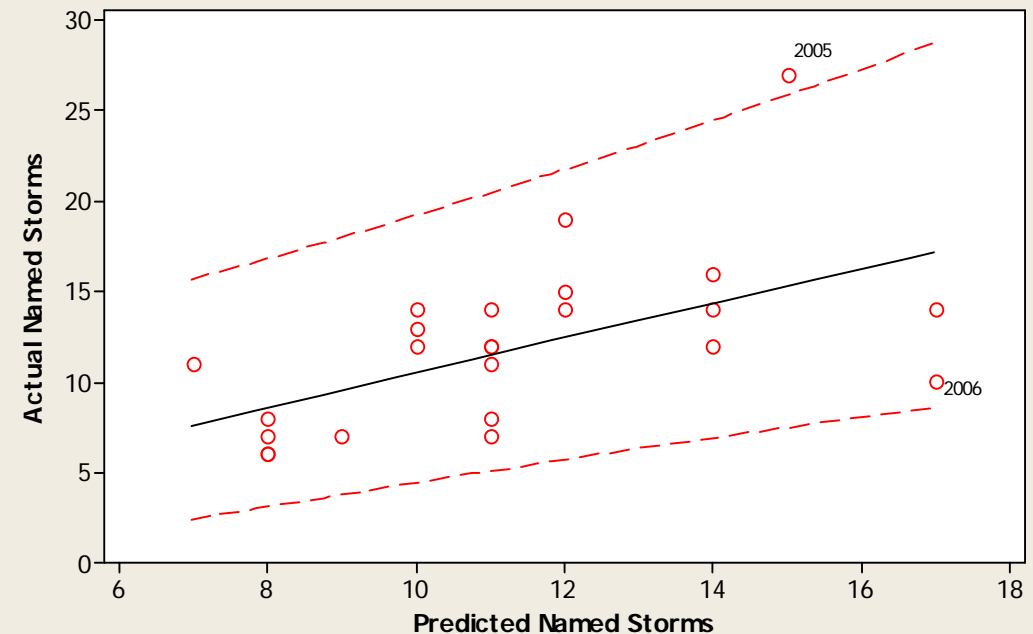


# How good are seasonal forecasts?

Dr. Gray - June Named Storm Prediction Accuracy



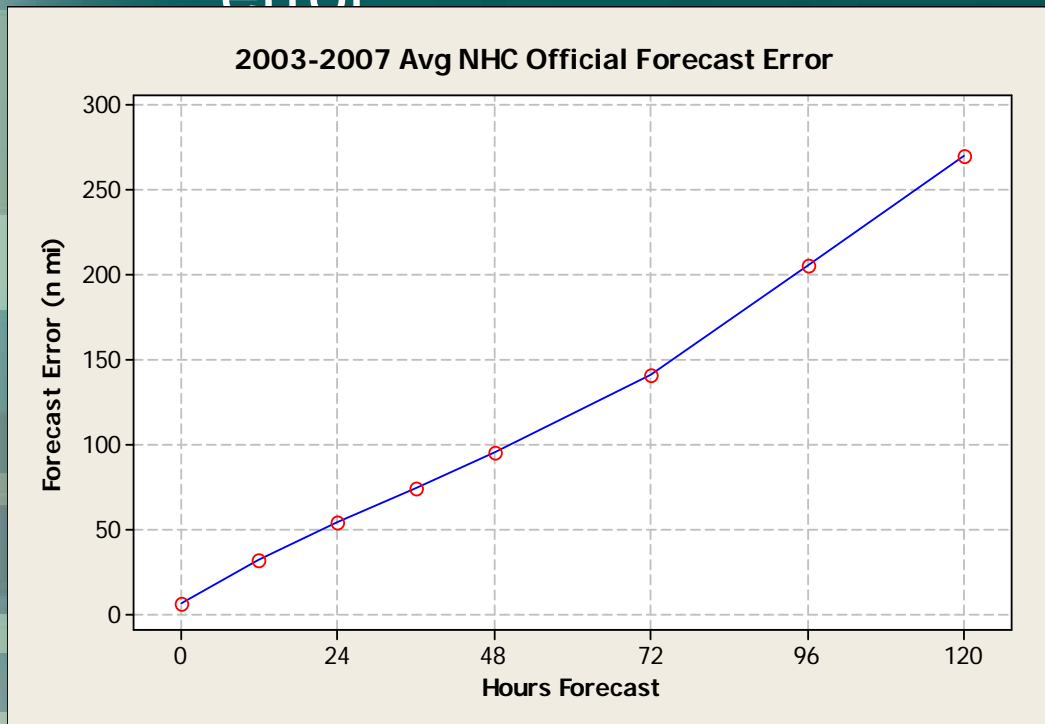
Dr. Gray - June Named Storm Prediction Accuracy



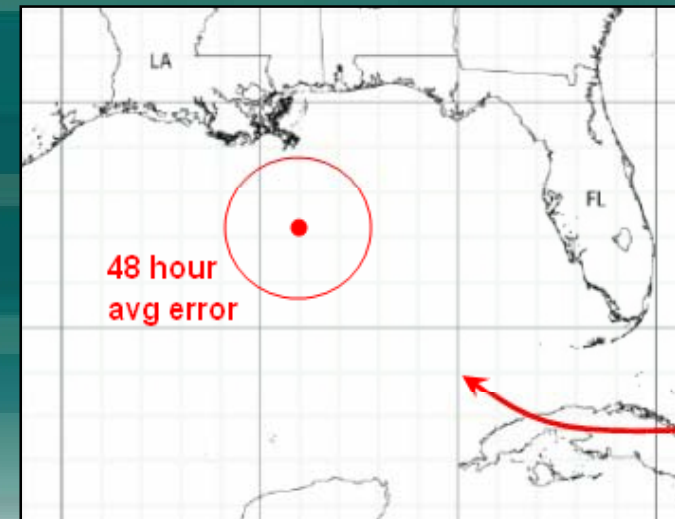
- Focus on Dr. Gray's since he has been doing it the longest and only one who reports detailed records of accuracy long enough back.
- June forecasts do demonstrate real forecasting ability.
  - But only 35% of the variability in # of storms predicted.
  - So if they predict 15 named storms, could be anywhere from 8 to 26 actual storms.
  - Best for just gauging an active vs. non-active season.

# Hurricane Track Forecasts

- How good are the forecasts tracks that the NHC reports during the season?
  - At the NHC, there is much more focus on track forecasting than seasonal forecasting.
  - The NHC publicly tracks “great circle” or “radial error”

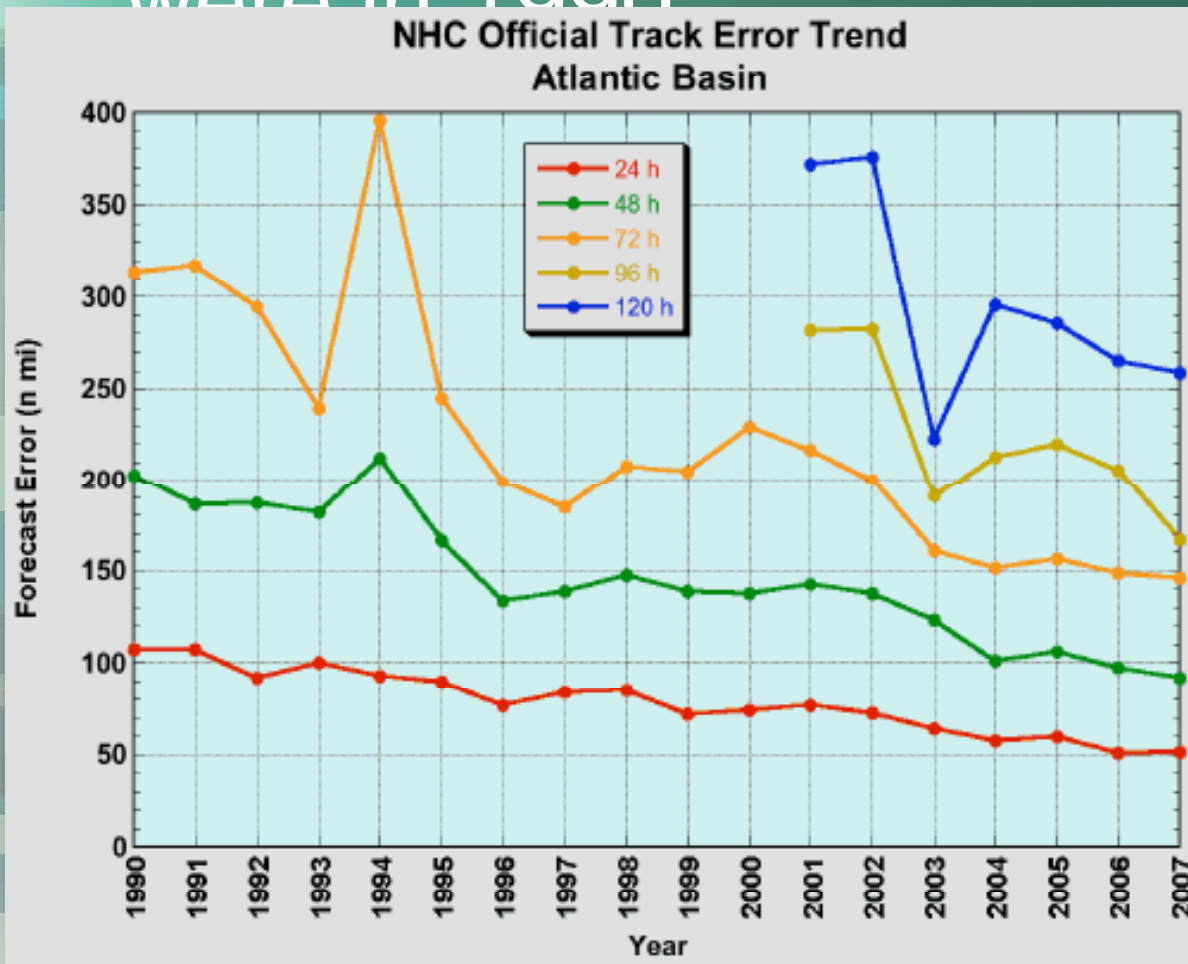


- 5 year avg errors
  - ~ 50 n mi at 24 hrs
  - ~ 100 n mi at 48 hrs
  - ~ 150 n mi at 72 hrs



# Hurricane Track Forecasts

- NHC forecast errors have been continually improving. Half of what they were in 1990



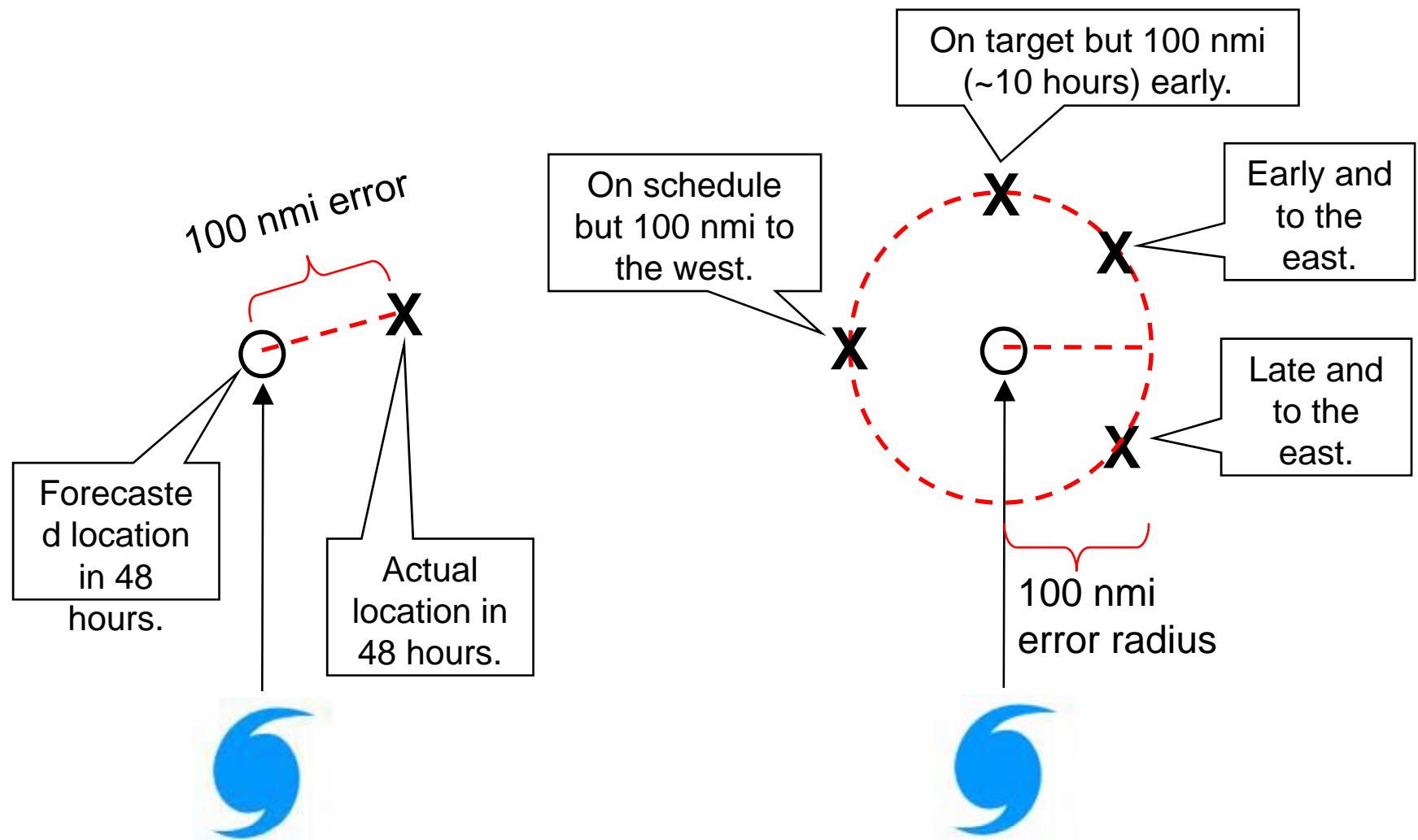
- Gained a full day of forecast ability.
- Now we are as certain 48 hours out as we were 24 hours out in the 1990's.

# Great Circle Error

- “Great circle” or “radial error” is the linear distance between the target and the actual outcome.
- Radial error does not distinguish or care about the directionality of the error from the target.
- Outcomes that are equidistant from the target are equivalent.
- All the blue outcomes are equal where the yellow outcome is better (with a smaller radial error).
- Radial error is used commonly in ballistics where all that really matters is how close the weapon came to the target.
- Also used in mechanical construction
  - Such as how close holes being drilled in a plate come to the target location.

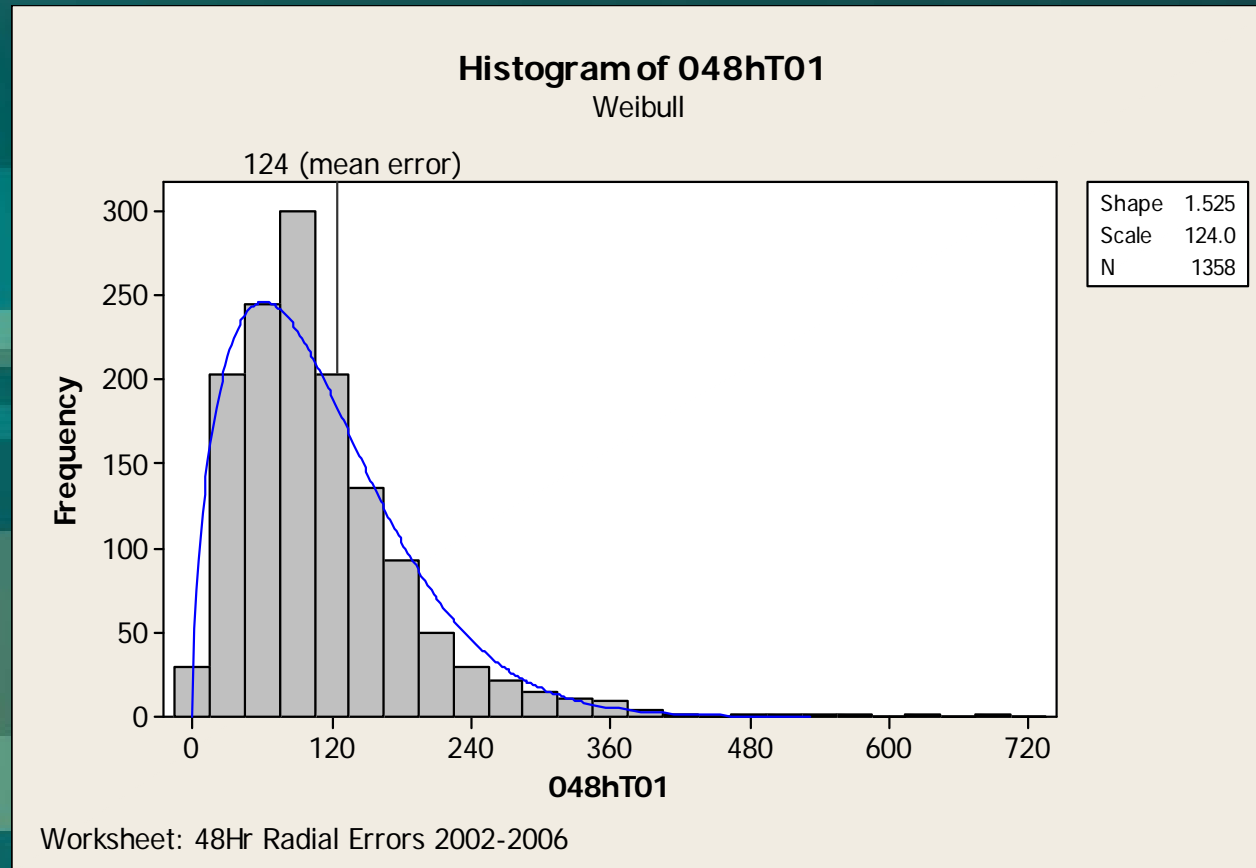


# Radial Error for Hurricane Forecasts



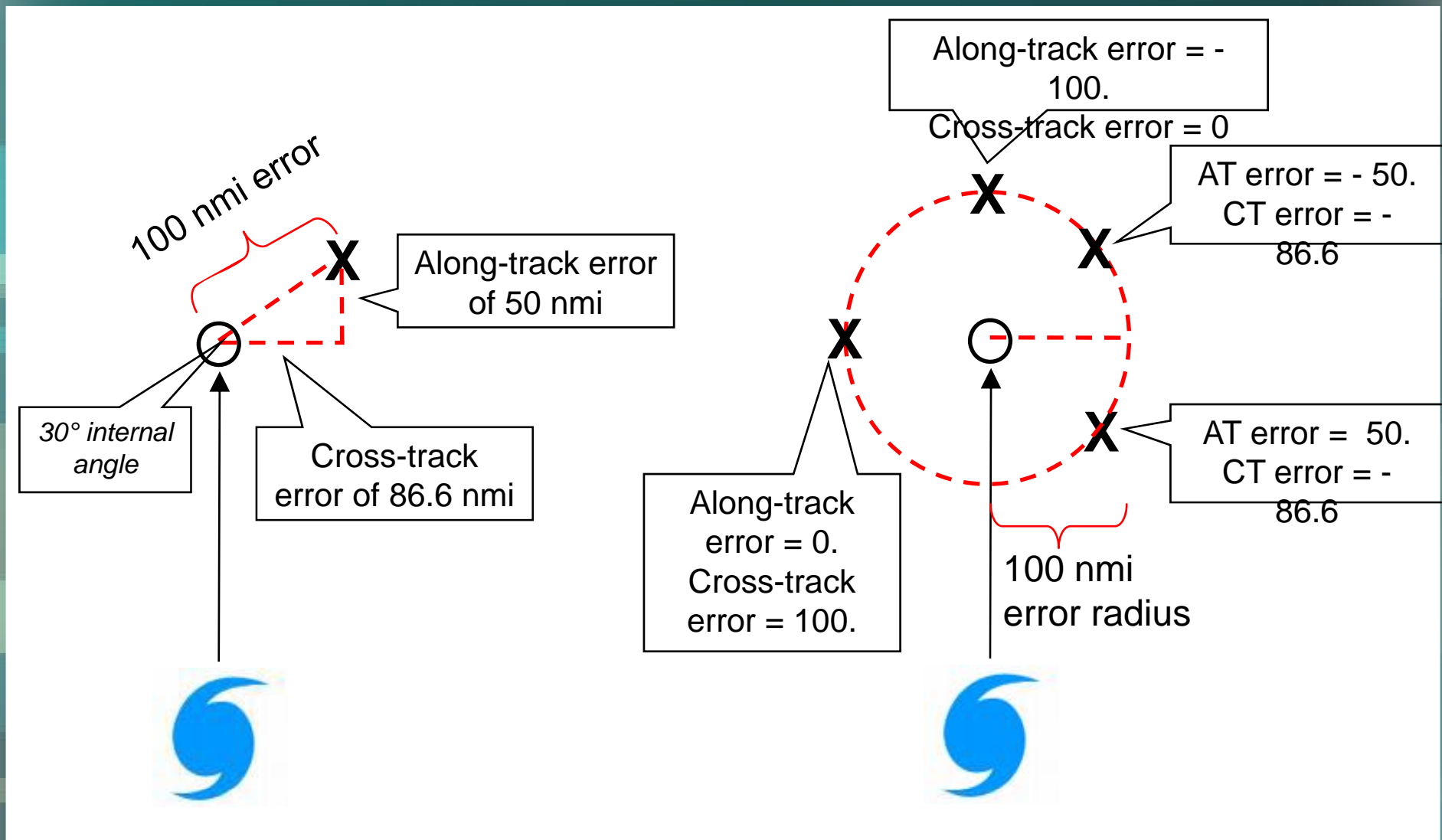
- Great circle or radial error methodology treats any error along a great circle as an equivalent error.
- Practically this is not at all true for hurricanes! To the locale that is forecast to be hit it can make a huge difference!

# Hurricane Radial Error



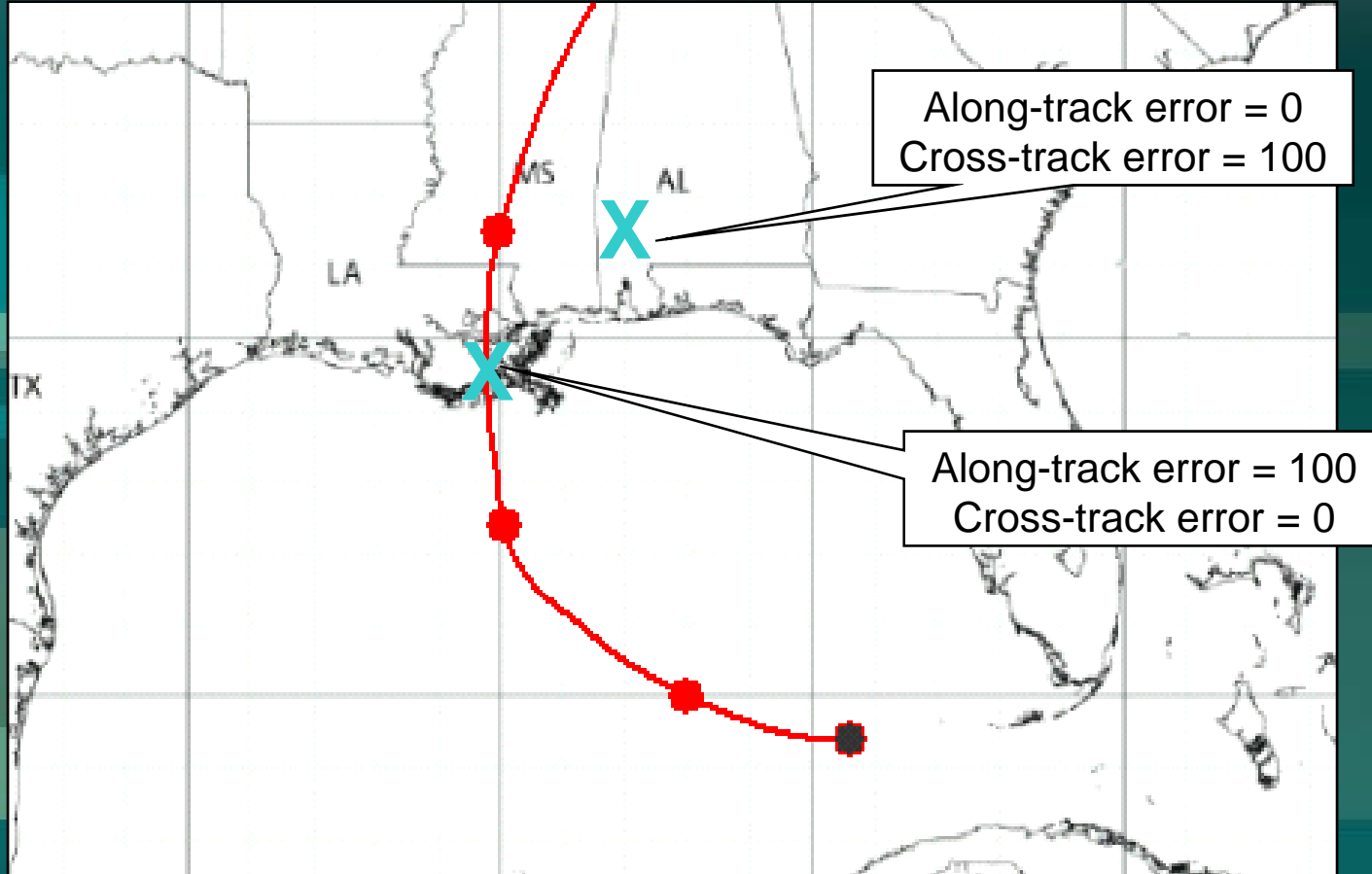
- Example of 48 hour forecast radial errors.
  - 2002-2006 forecasts
  - Mean error of 124 nmi.
  - Not a normal distribution (negative errors not possible by this definition.)
  - Can fit a Weibull distribution.

# Alternative to Radial Error



- Along-track and Cross-track error can be calculated for each forecast point.
- This results in a much more practical error quantification.

# Along and Cross Track Error



- Both X's have the same radial error.
- If you live on the gulf coast does it make a difference which outcome really occurs?
- Cross-track error has much more practical importance.

## Cross-track & Along-track Error Data Set

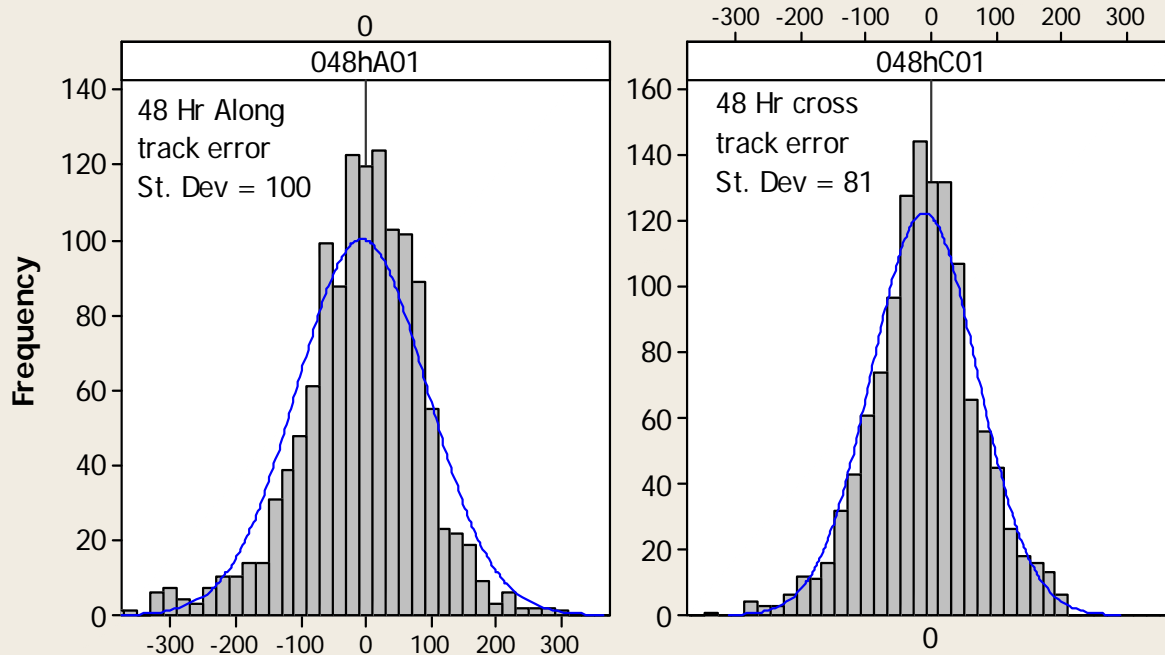
- Using NHC forecast data for all storms 2003-2007.
- Official NHC forecasts used (no individual model)
- Cross-track & along-track error calculated for each forecast period for each forecast point.
- Cross-track and along-track error is calculated as the error parallel and perpendicular to the storm motion at the time of the forecast point.
- Error data obtained from James L. Franklin, Senior Hurricane Specialist at the NHC.
  - Franklin maintains and issues hurricane forecast validation reports annually.
  - Cross-track and Along-track errors are not usually publicly discussed and shared but are available.

# My Analysis Objectives

- Understand NHC forecasting errors better.
- Any differences in Along and Cross track errors?
- Are there any variables that impact forecast accuracy of a given storm?
- Are there better ways to display NHC forecasts and errors?
  - That can lead to more clear decisions by businesses, communities, and individuals?

# Cross-track & Along-track Error

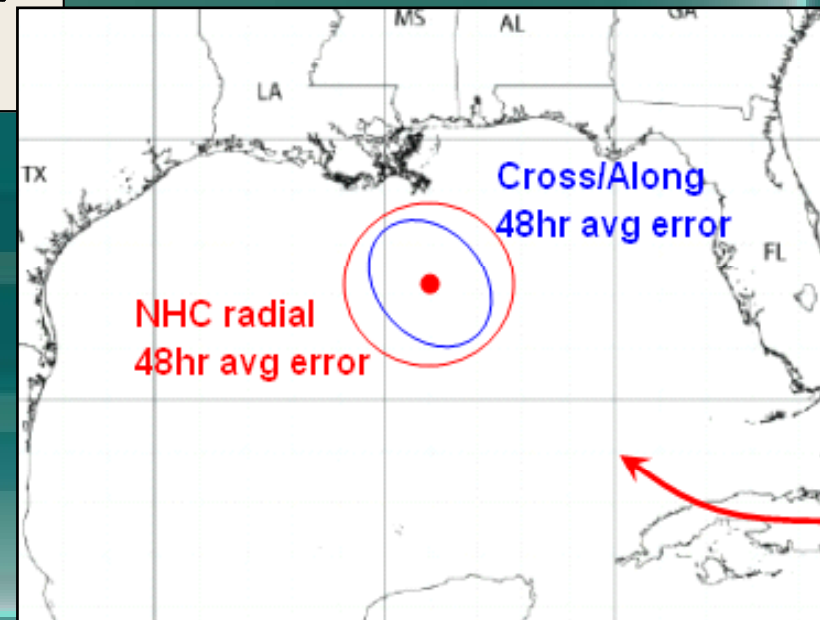
Histogram of 48 Hr Along and Cross Track Errors  
Normal Distribution Fit



Worksheet: 2003-2007

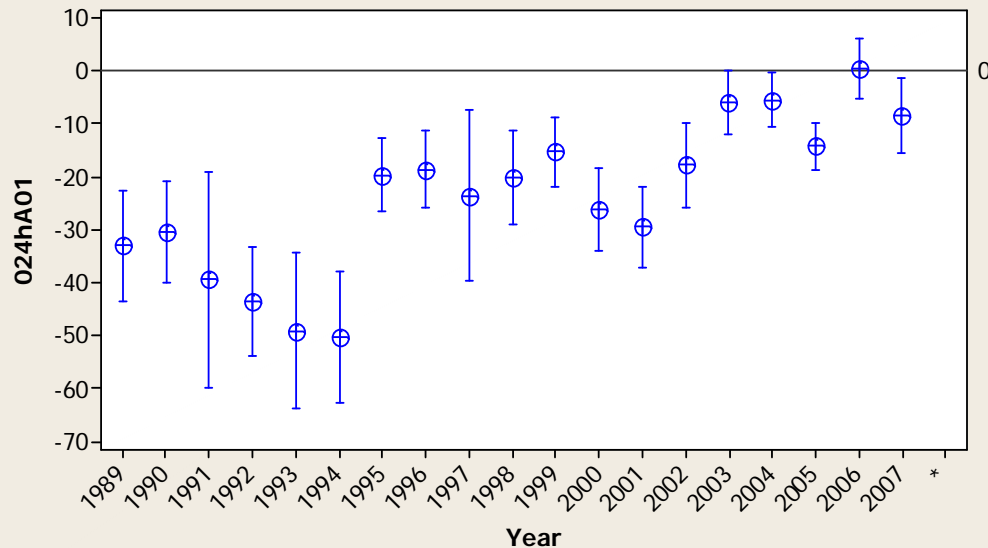
- The error envelope is really an ellipse, not a circle.
- NHC is better at forecasting where a storm is heading than forecasting when it will get there.

- Along track error std. dev is higher than cross track error.
- Avg error:
  - Along track = 80 nmi
  - Cross track = 65 nmi



# Along-track errors: forecast bias?

Interval Plot of 024hA01  
95% CI for the Mean



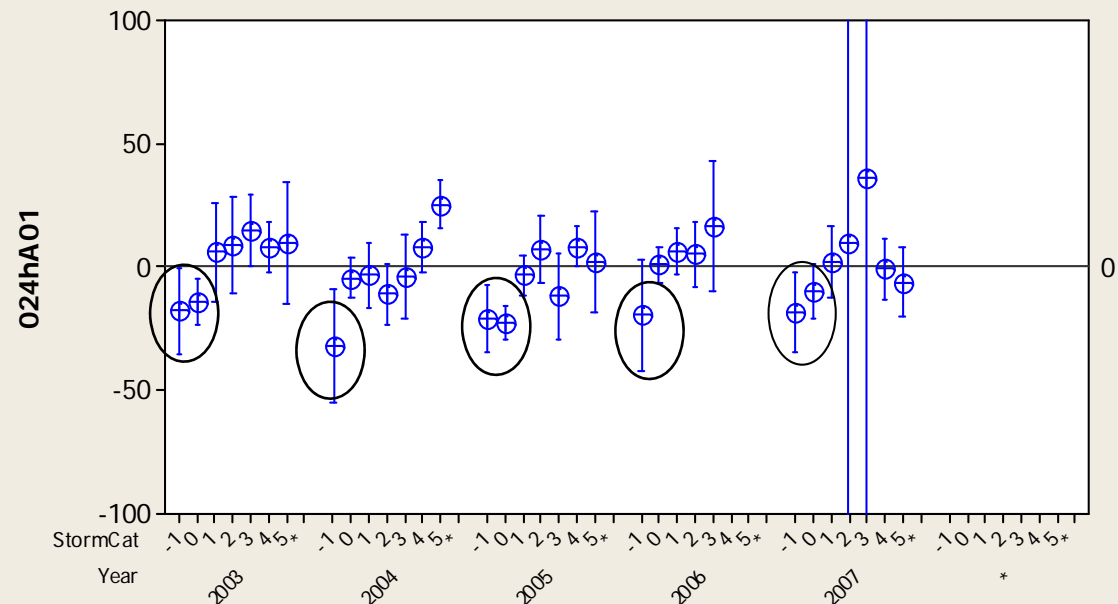
Average 24 hour error bias has generally been negative (forecast is behind actual storm). But this bias has shrunk considerably in the last 5 years.

Worksheet: ALL DATA ('89-'07).MTW

Looking at just the last 5 years by storm category, most of the negative bias is from TD's and TS, hurricanes don't show the same consistent negative bias.

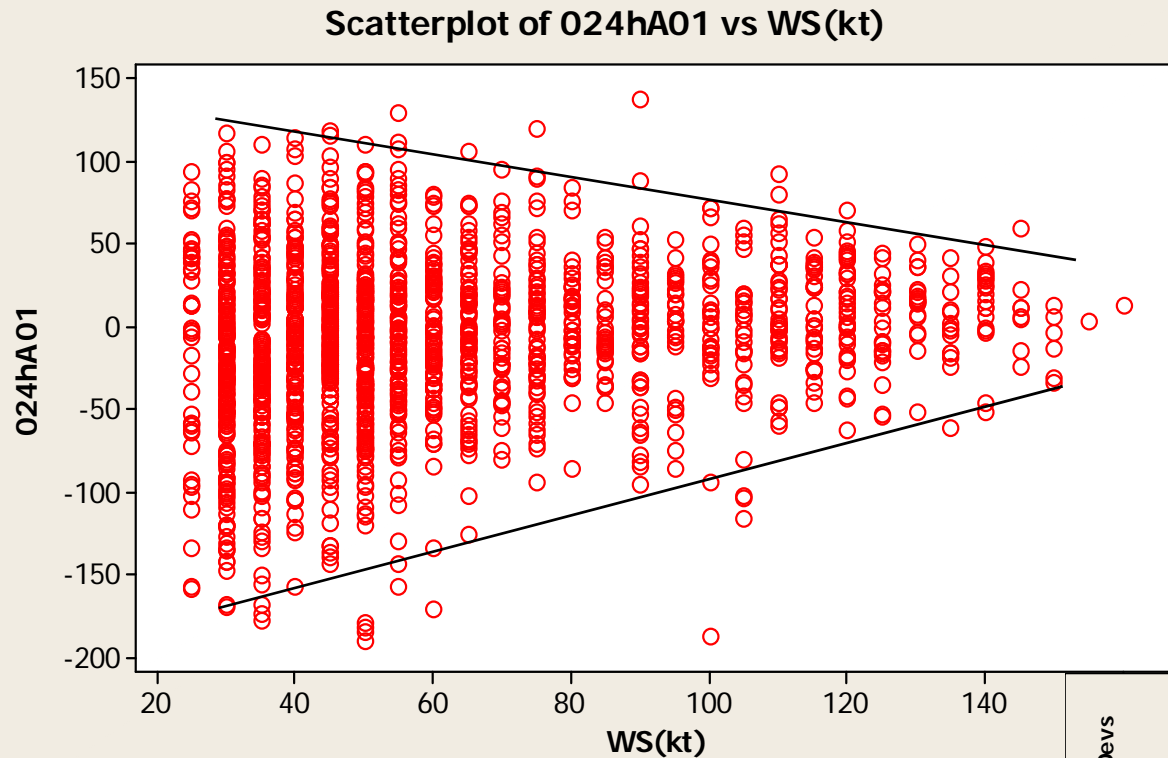
So, no bias will be used.

Interval Plot of 024hA01  
95% CI for the Mean



Worksheet: ALL DATA ('89-'07).MTW

# Along-track errors: std. deviation

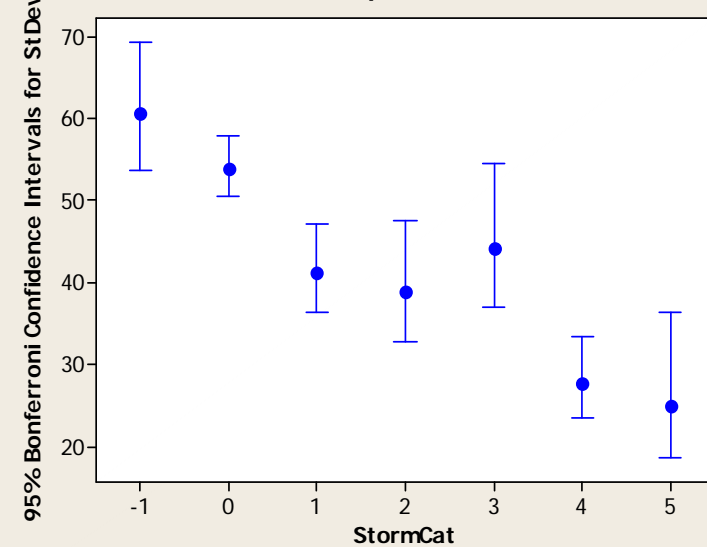


Worksheet: 2003-2007

Strong relationship of variability of along track error to wind speed of storm at time of forecast. See this for all forecast horizons. Will develop regression to fit this.

Test for equal variances by storm category confirms relationship of lower variance in forecasting error for higher category storms.

Test for Equal Variances for 024hA01

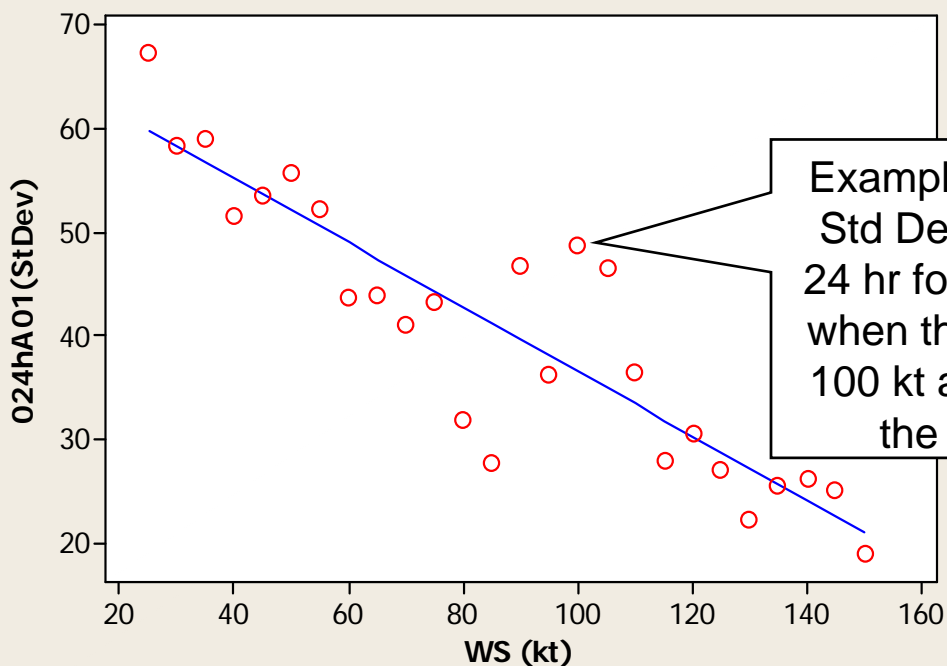


Bartlett's Test	
Test Statistic	136.54
P-Value	0.000
Levene's Test	
Test Statistic	17.19
P-Value	0.000

Worksheet: 2003-2007

# Along-track errors: std. deviation

**Fitted Line Plot**  
 $024hA01(StDev) = 67.60 - 0.3110 WS \text{ (kt)}$



- Regression showing nice linear relationship of 24hr forecast standard deviation to wind speed at time of forecast.

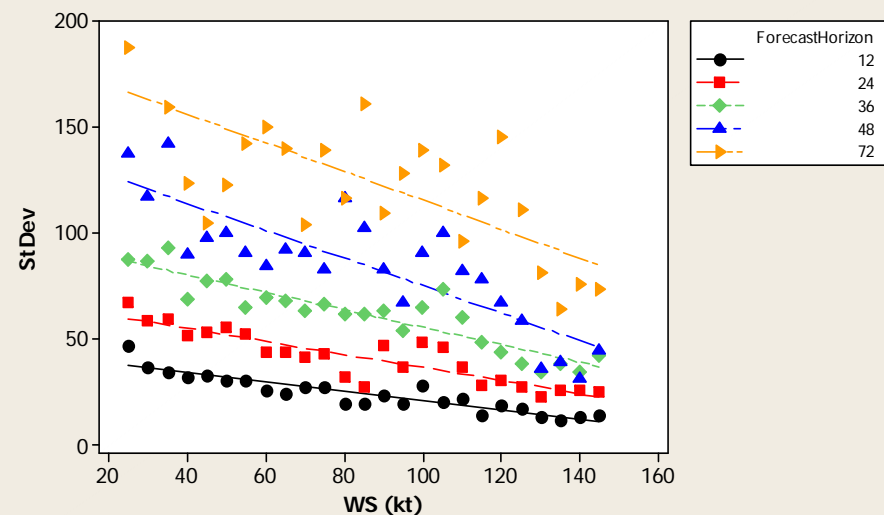
- As windspeed at the time of forecast increases, 24hr forecasting error std. deviation decreases.

Worksheet: 2003-2007

- Saw this relationship at all forecast horizons.

- Also can see obvious relationship of StDev and forecast horizon.

**Scatterplot of StDev vs WS (kt)**



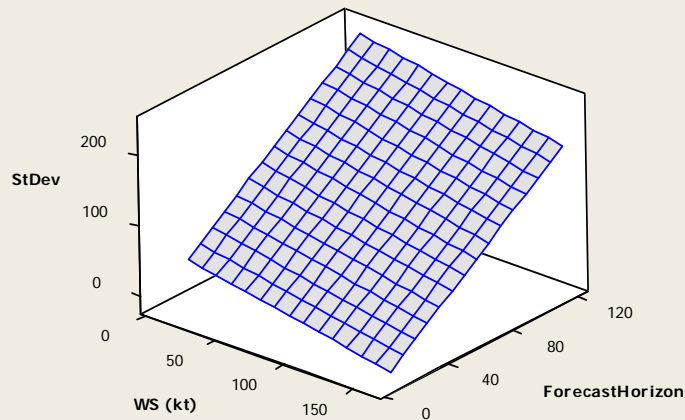
Worksheet: 2003-2007

# Along-track errors: response surface regression

Response surface regression. The interaction and square terms were not significant so removed them. End up with flat plane surface.

Looked good, but residuals shows bad homoscedasticity problem.

Surface Plot of StDev vs ForecastHorizon, WS (kt)



Worksheet: 2003-2007

## Response Surface Regression: StDev versus WS (kt), ForecastHorizon

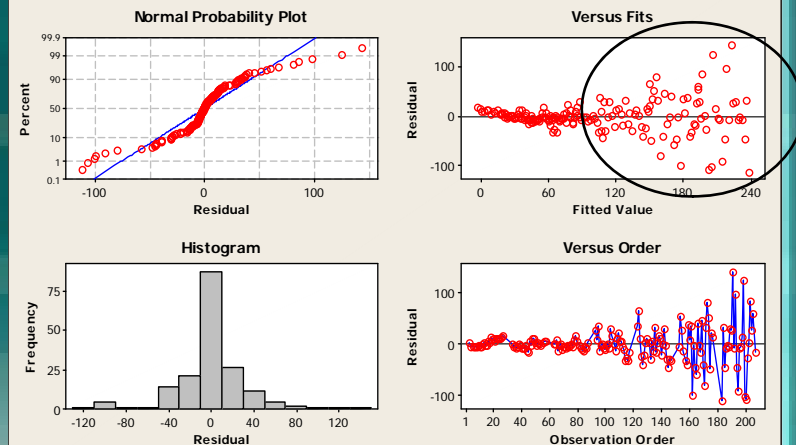
The analysis was done using coded units.

Estimated Regression Coefficients for StDev

Term	Coef	SE Coef	T	P
Constant	116.46	2.558	45.525	0.000
WS (kt)	-28.89	5.021	-5.753	0.000
ForecastHorizon	96.88	3.692	26.243	0.000

S = 33.0345 PRESS = 196386  
R-Sq = 80.76% R-Sq(pred) = 79.87% R-Sq(adj) = 80.53%

Residual Plots for StDev

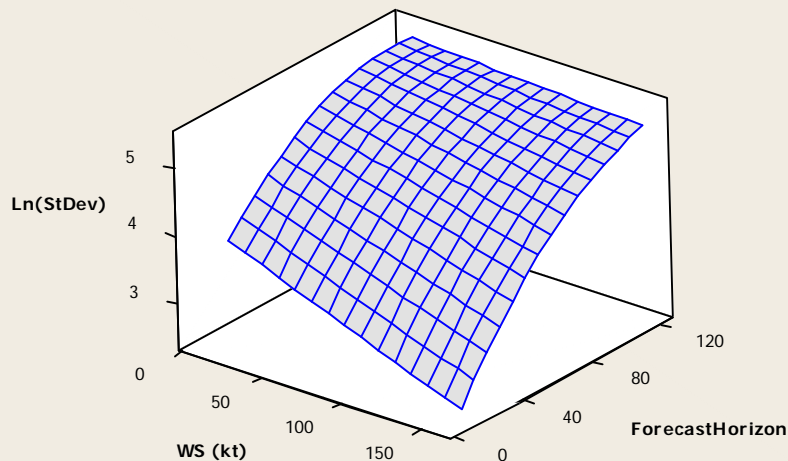


Worksheet: 2003-2007



# Along-track errors: response surface regression

Surface Plot of Ln(StDev) vs ForecastHorizon, WS (kt)



Worksheet: 2003-2007

## Response Surface Regression: Ln(StDev) versus WS (kt), ForecastHorizon

The analysis was done using coded units.

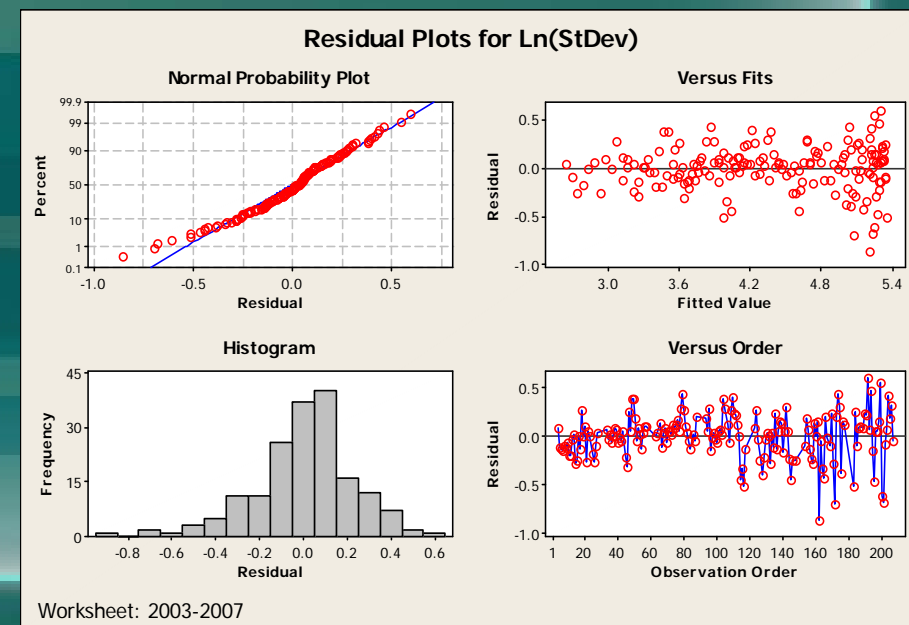
Estimated Regression Coefficients for Ln(StDev)

Term	Coef	SE Coef	T	P
Constant	4.7259	0.02908	162.510	0.000
WS (kt)	-0.3881	0.03644	-10.650	0.000
ForecastHorizon	1.0468	0.02628	39.836	0.000
ForecastHorizon*ForecastHorizon	-0.4960	0.04764	-10.412	0.000
WS (kt)*ForecastHorizon	0.2984	0.05271	5.662	0.000

S = 0.234552 PRESS = 9.96054  
R-Sq = 91.64% R-Sq(pred) = 91.10% R-Sq(adj) = 91.44%

Regression now has some interaction and square terms but remember the response is on ln(stdev) so need to convert that back.

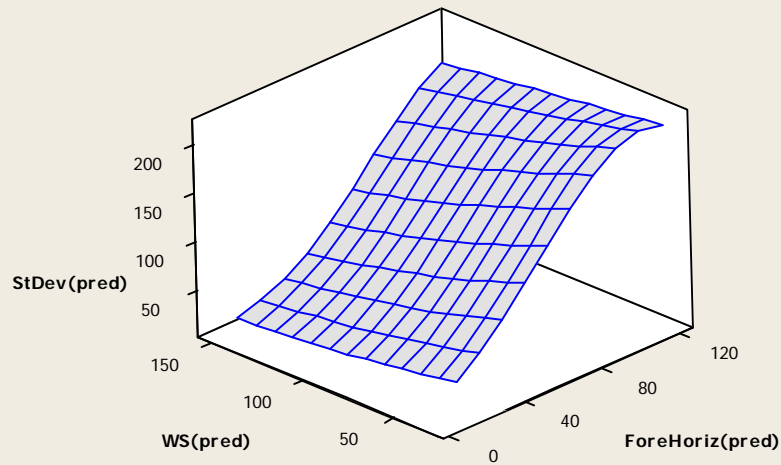
Residuals look much better although just a bit of issue still at higher fits but won't go any further as this is probably due to lower number of datapoints generating the stdev for some conditions.



Worksheet: 2003-2007

# Along-track errors: response surface regression

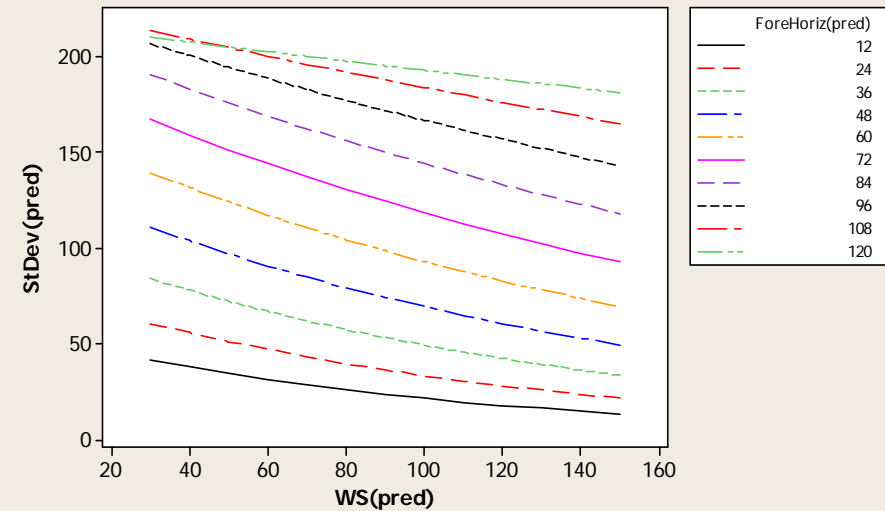
Surface Plot of StDev(pred) vs WS(pred), ForeHoriz(pred)



Worksheet: 2003-2007

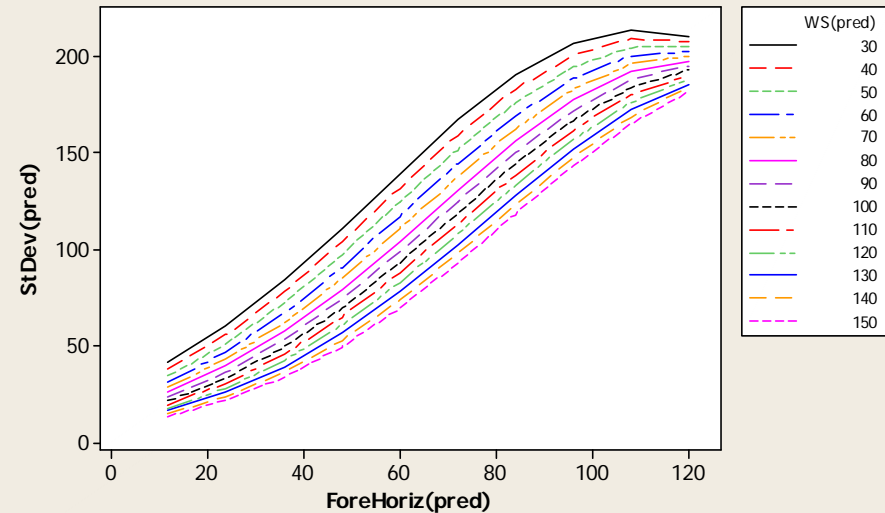
Different perspectives of the resulting model converted back to StDev units from the ln transform. Not a linear surface!

Scatterplot of StDev(pred) vs WS(pred)



Worksheet: 2003-2007

Scatterplot of StDev(pred) vs ForeHoriz(pred)

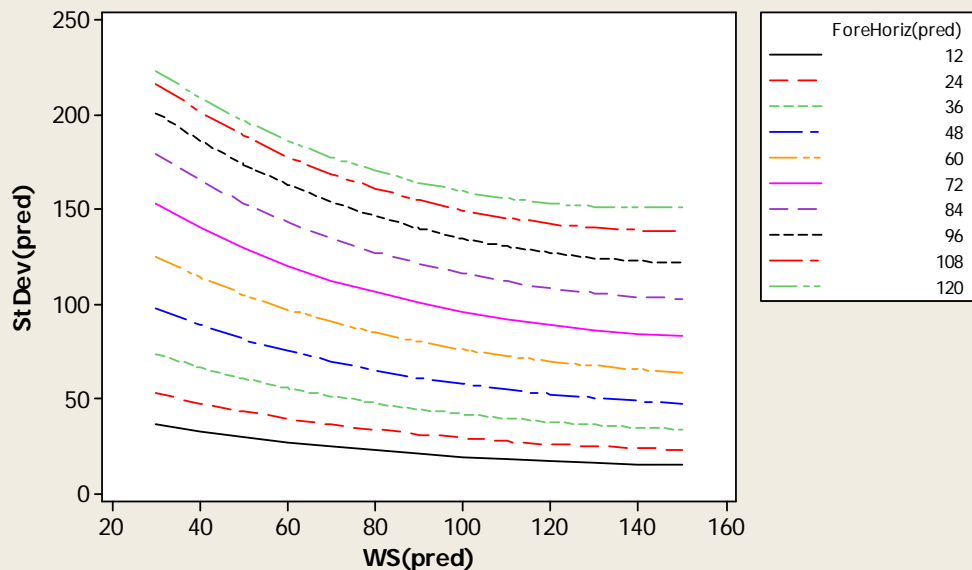


Worksheet: 2003-2007

# Cross-track errors: response surface regression

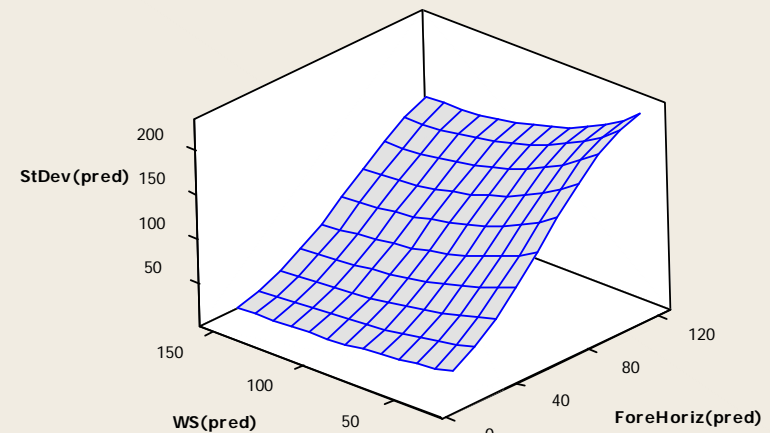
- Same approach used on cross-track errors.
- Same qualitative findings and results.
- Plots shown below in untransformed units.

Scatterplot of StDev(pred) vs WS(pred)



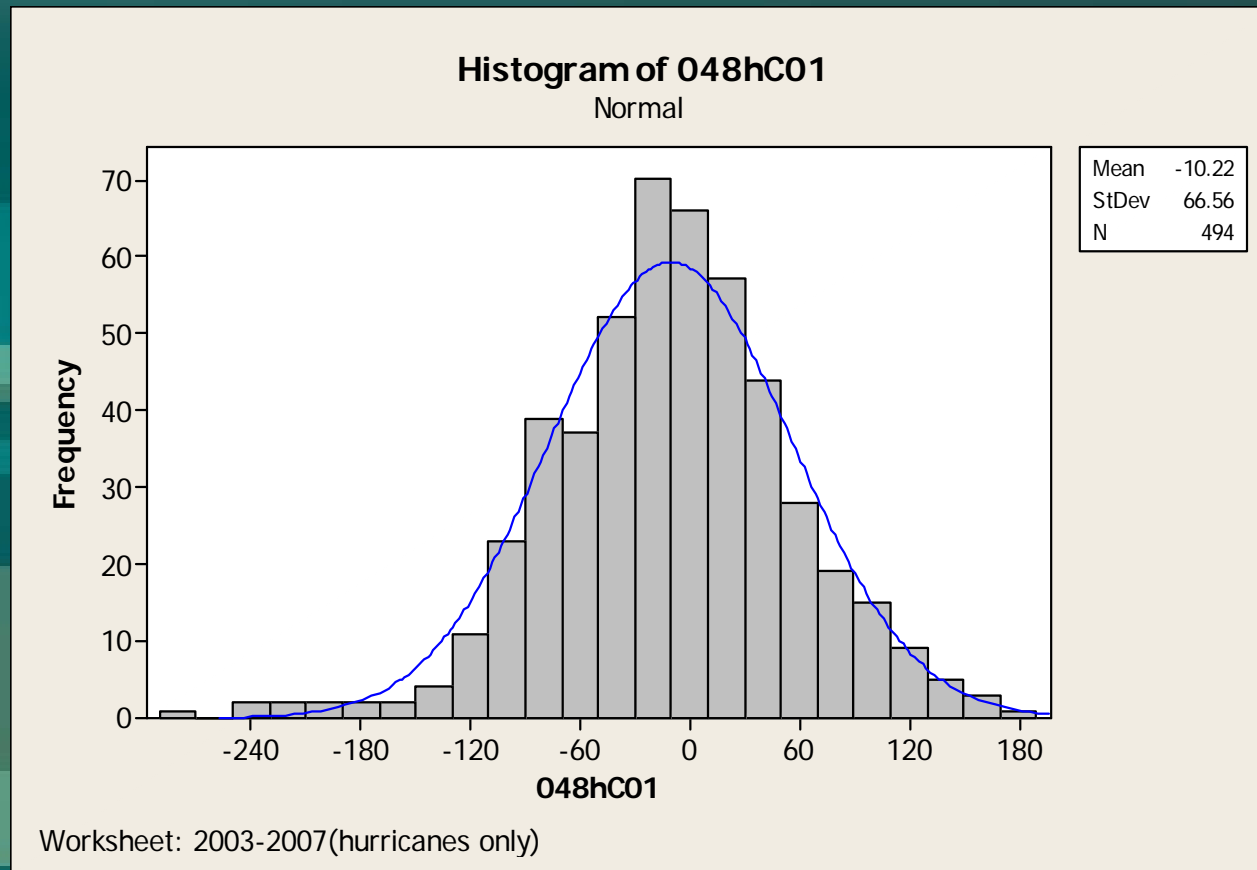
Worksheet: 2003-2007

Surface Plot of StDev(pred) vs WS(pred), ForeHoriz(pred)



Worksheet: 2003-2007

# Cross-track errors distribution

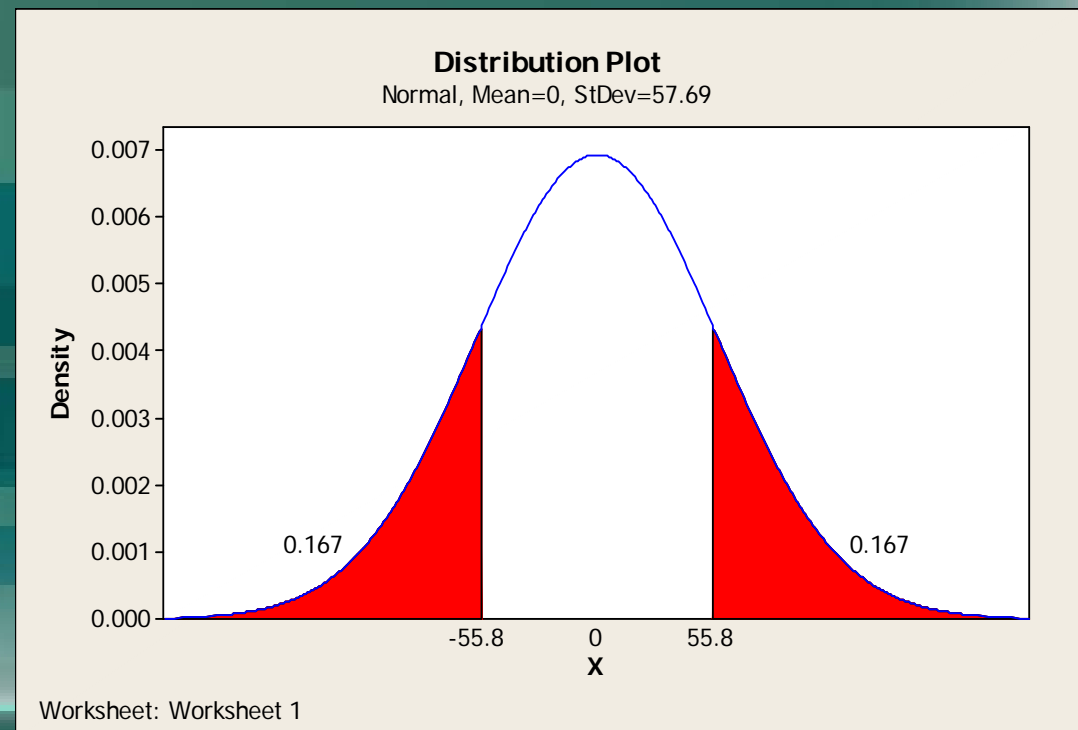


- Cross-track errors for a given storm strength and forecast horizon fit a normal distribution well.
- Small degree of extra outliers on the left tail. Probably due to the fact that most storms curve to the right.
- A normal distribution calculation will be used to estimate forecast error envelopes.
- Along track errors fit a normal distribution also (actually better than cross track errors).

# Error Regression Result

$$\sigma_{\text{crosstrack}} = e^{(3.6 - 0.013 * WS + 0.035 * Hrzn + 0.000031 * WS^2 - 0.00015 * Hrzn^2 + 0.000038 * WS * Hrzn)}$$

- Result is an equation relating std. deviation of forecasting error to
  1. windspeed (strength) of a storm at time of the forecast
  2. forecast horizon
- One equation for cross-track error and one for along-track error.
- For example: A 48 hour forecast horizon for a storm with a current strength of 100 kts (115 mpg) has a cross-track error standard deviation of 57.69 nmi.
- 67% interval is calculated using a normal inverse probability calc.
  - = +/-55.8 nmi
- Much smaller than the standard NHC ~100 nmi estimate for a 48 hr forecast.

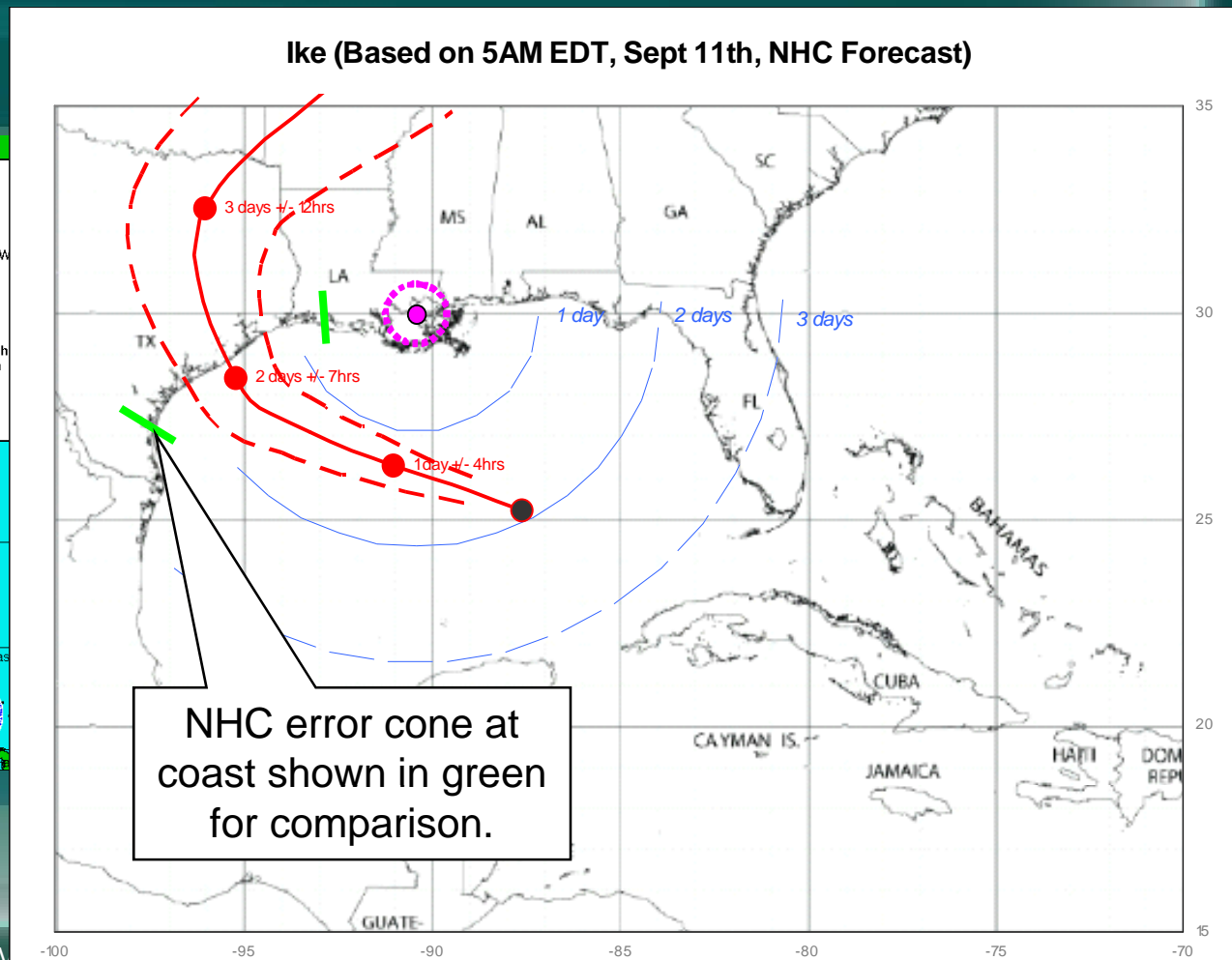
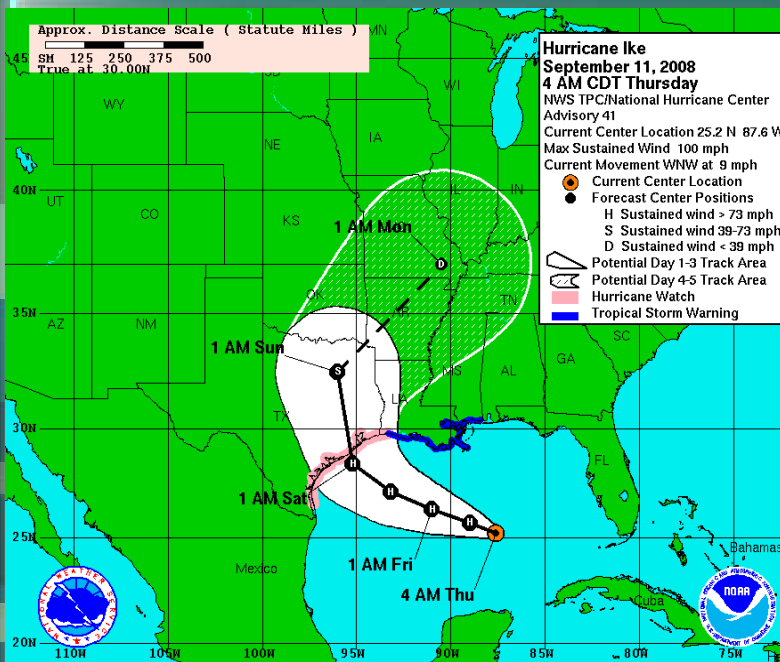


# Error Regression Conclusions

- NHC forecast errors are elliptical with the long axis being along the direction of storm movement.
- Radial error methodology employed by the NHC to generate average error statistics used for reporting uncertainty envelopes result in a conservative (too big) of an envelope.
- Forecast horizon has a significant impact on standard deviation of forecast error.
- Storm strength has a significant impact on standard deviation of forecast error.
  - Stronger the storm the smaller the average error (as much as half)
  - Makes sense since as storms get stronger they also get more organized and result in more accurate initial condition estimates when generating model runs. Stronger storms also are more dominant atmospheric structures that start to shape the environment around them which ought result in more predictable and “rational” behavior.

# Practical Use

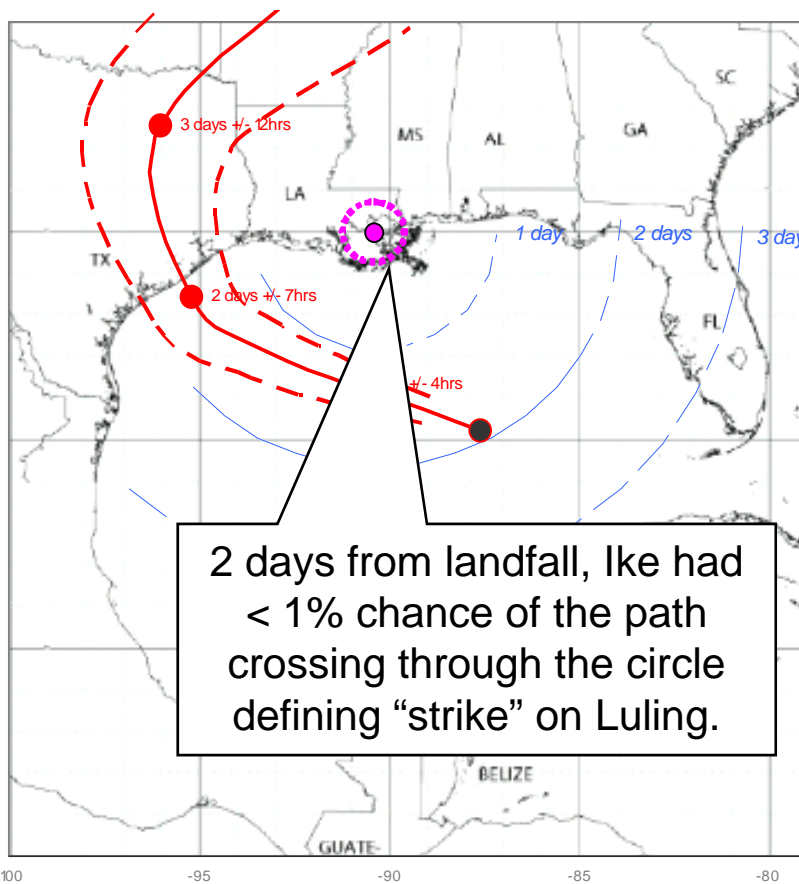
- Error equations were put into an excel based tool to plot NHC hurricane forecasts along with these cross track and along track error envelopes.
- Notice considerably narrower “error cone” as it is based on cross track error only and narrower due to Ike’s 100 mph strength at this time.
- Along track error shown as +/- hours.



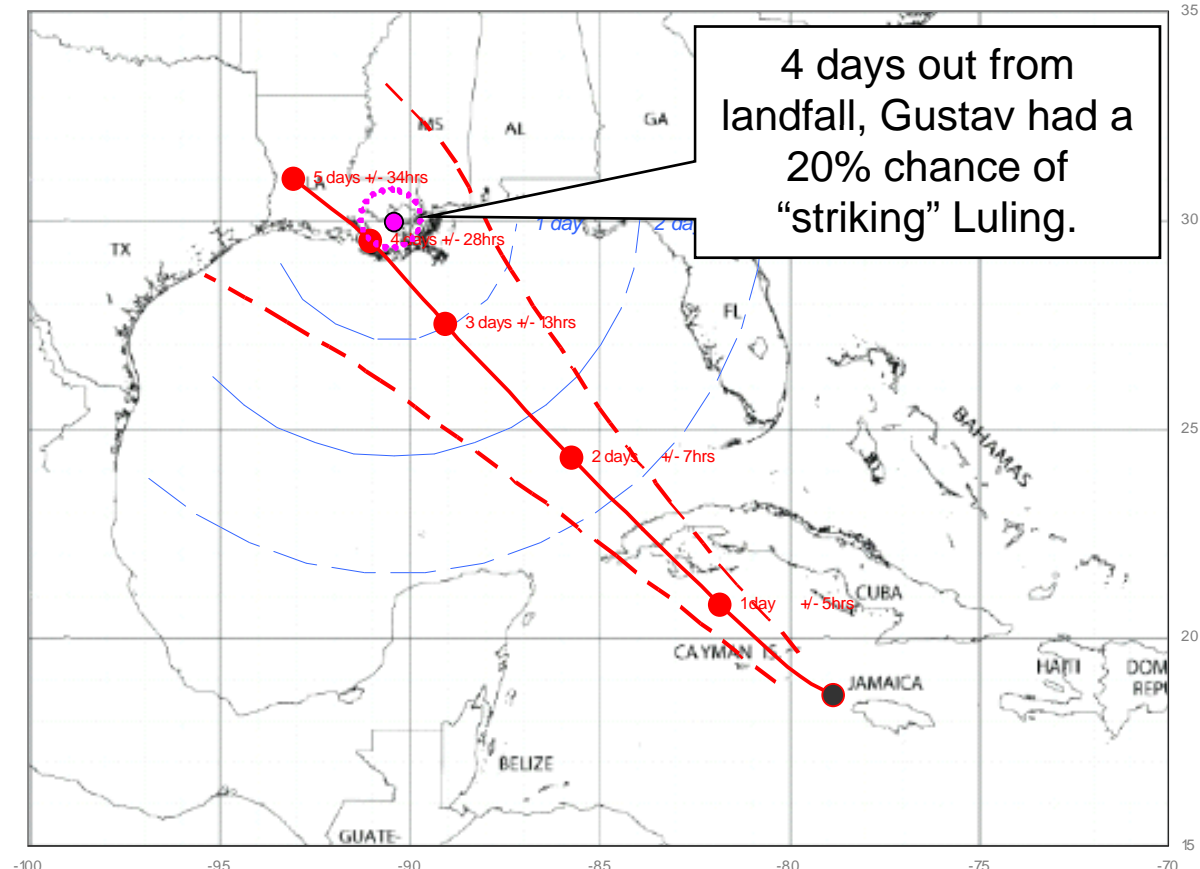
# Practical Use

- With the excel tool we can estimate a probability of a given storm's path crossing a location of interest.

Ike (Based on 5AM EDT, Sept 11th, NHC)

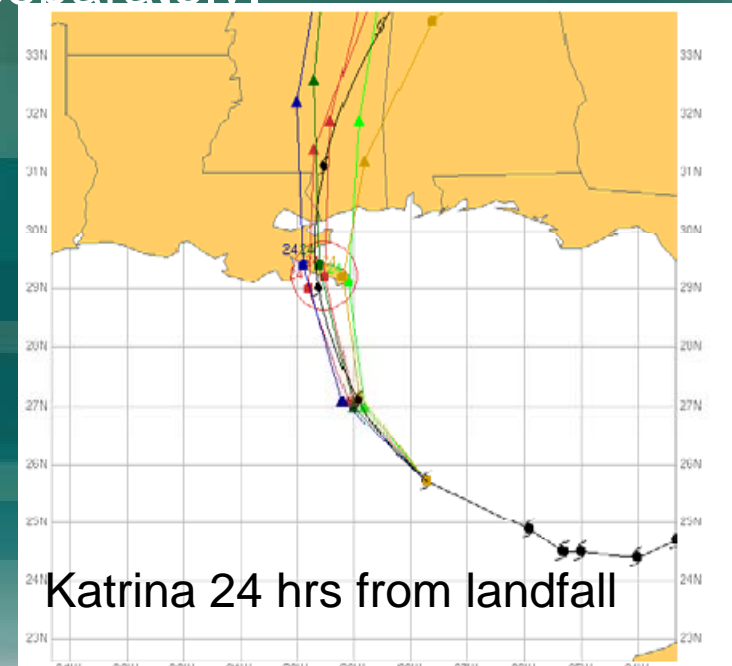
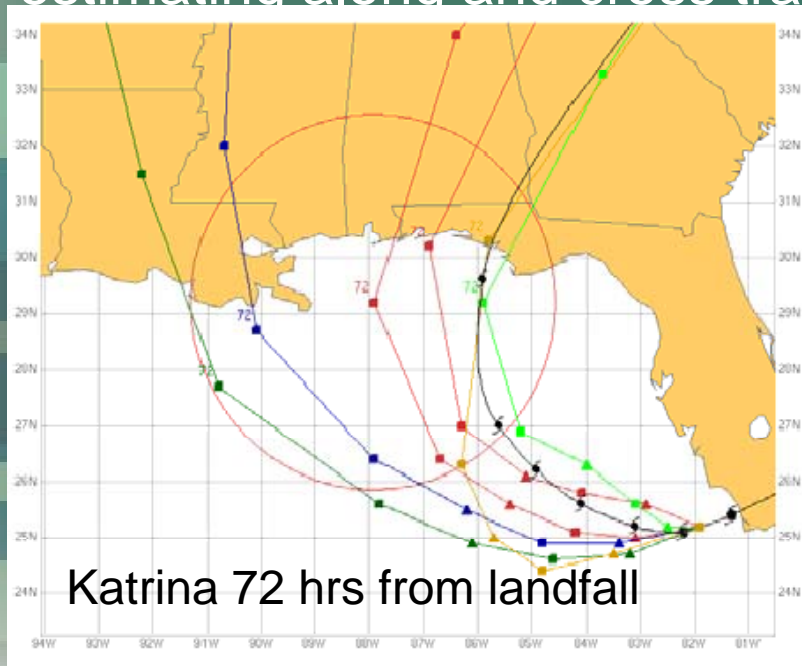


Gustav (Based on 2PM, Aug 29 NHC Forecast)



# Related work

- Dr. James Goerss @ NRL has been working on prediction of forecast error.
- His findings are consistent with this work:
  - Storm strength is an important predictor.
  - Uses model spread as a predictor of forecast error.
- GPCE (graphical predicted consensus error) shown below uses an error circle based on model spread and other variables.
- GPCE-AX is where the work is going which takes a similar approach of estimating along and cross track errors separately.



# My takeaways...

- NHC official track forecasts are quite good
  - Especially up to the 2-3 day window.
  - Especially when a storm is hurricane or greater strength.
  - The official “error cone” is usually too wide and conveys more uncertainty than there really is.
  - NHC forecast errors continue to improve.
    - So if you have a distant memory of that time they got a forecast way wrong, that memory is probably not indicative of today’s forecasting ability.

# Conclusions...

- Future plans
  - Continue to refine model and use
  - Share results
  - Possibly develop Monte Carlo simulation to optimize decision/response planning
  - Special thanks to:
    - James L. Franklin, Senior Hurricane Specialist at the NHC.
    - Dr. James Goerss at the Naval Research Lab
- Questions?