

Joint probability methods to storm surge studies

Pat Fitzpatrick

Mississippi State University

- Motivation
- Concept – frequency analysis
- Concept – joint probability
- Steps in creating “brute force” joint probability method (JPM) for surge frequency analysis
- Optimal Sampling techniques for JPM
- Methods for synthetic tropical cyclone tracks
- Example – New Orleans post-Katrina levees

Additional provided material

FEMA, 2012: Operating Guidance 8-12 – Joint Probability-Optimal Sampling method for tropical storm surge frequency analysis

Joint Probability Distributions (class lecture)

Jacobsen, R., 2012: Concepts in flood return frequency analysis

Toro, G. R., D. T. Resio, D. Divoky, A. W. Niedoroda, and C. Reed, 2010: Efficient joint probability methods for hurricane surge frequency analysis. *Ocean Engineering*, **37**, 125-134.

Niedoroda, A. W., G. R. Toro, D. Divoky, H. S. Das, and C. Reed, 2010: Analysis of the coastal Mississippi storm surge hazard. *Ocean Engineering*, **37**, 82-90.

U.S. Nuclear Regulatory Commission, 2012: Guidance for performing a tsunami, surge, or seiche hazard assessment

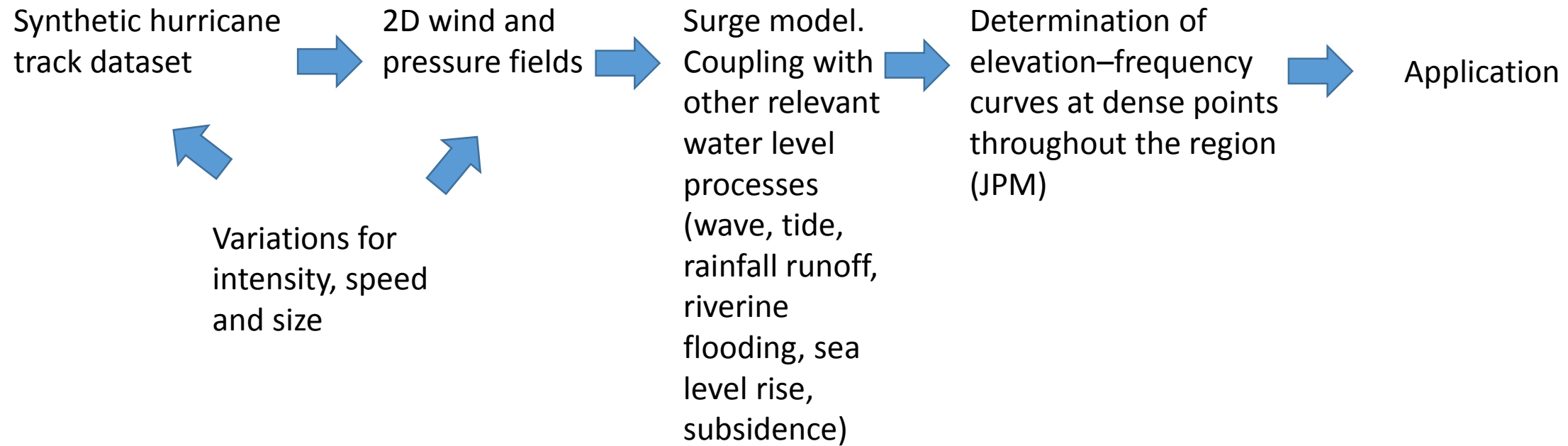
Toro, G. R., S. C. Hagen, J. Atkinson, and C. Reed, 2011: Production runs for the Big Bend region of Florida. *The Florida Watershed Journal*, 28-35.

What is a 100-year flood?

Motivation

Applications for flood or surge frequency analysis:

- Flood insurance studies. The NFIP requires return period inundation values on Flood Insurance Rate Maps (FIRMS), which determine insurance rates.
- Flood control studies. An example is the determination of southeast Louisiana's new 100-year levee heights during reconstruction after Katrina
- Flood mitigation studies. An example is the ongoing reassessment on whether coastal nuclear plants meet 1 in 10,000 year flood protection criteria.
- Flood impact studies. Determine if return levels change due to urban runoff changes, floodplain restoration, channel modification, etc.

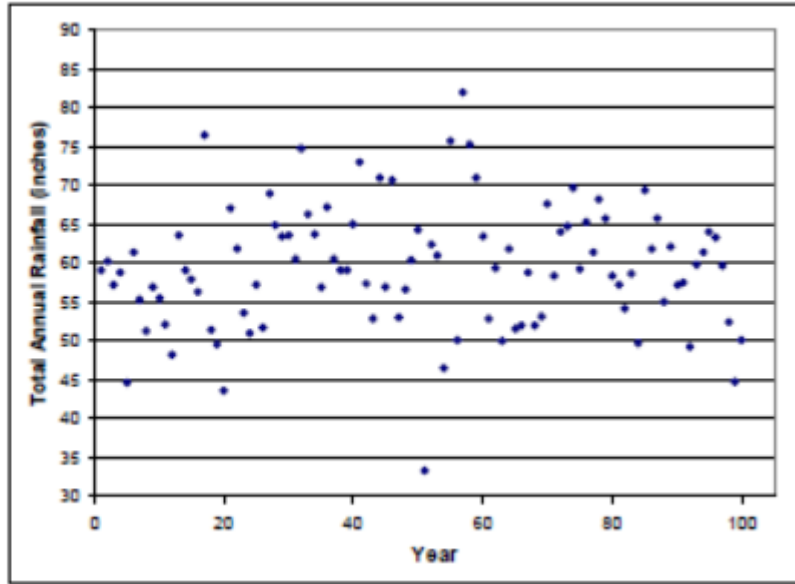


Application examples

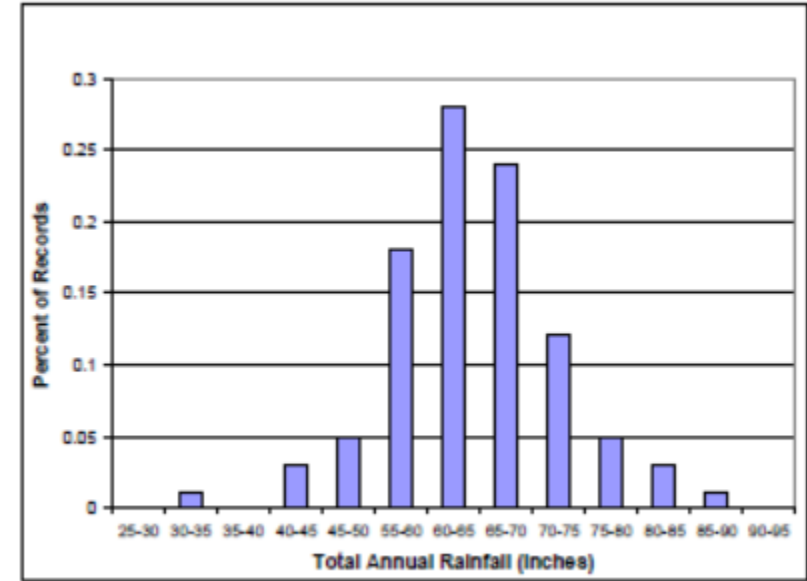
- Flood Insurance Rate Maps (FIRMs)
- Levee height design
- Elevation or protection design for nuclear plants

Concept - Frequency analysis

a) Data Record



b) Data Ranges versus Percent of Records



c) Data Ranges versus Cumulative Percent of Records

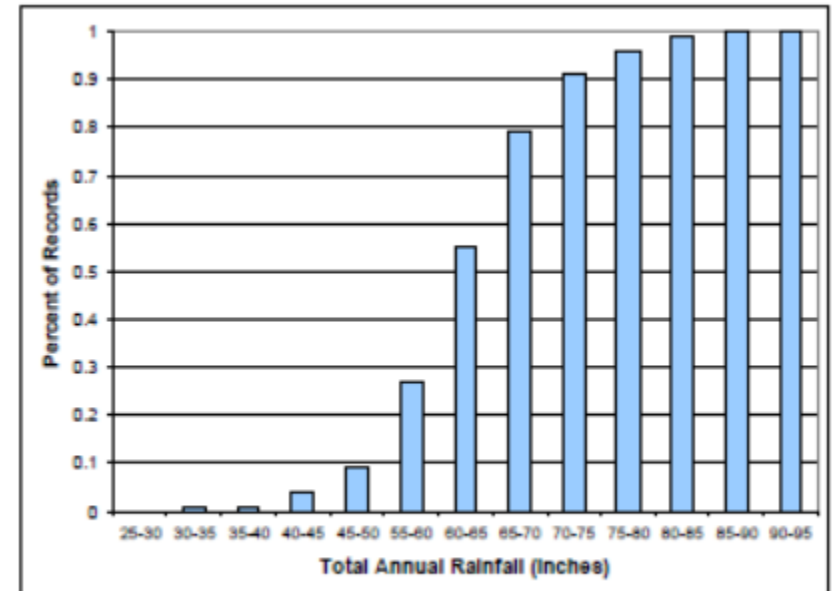
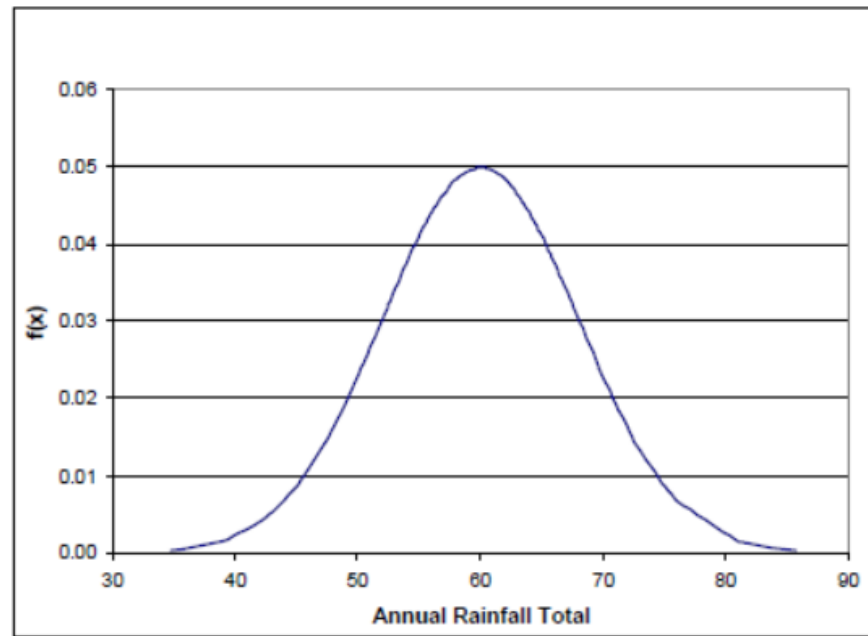
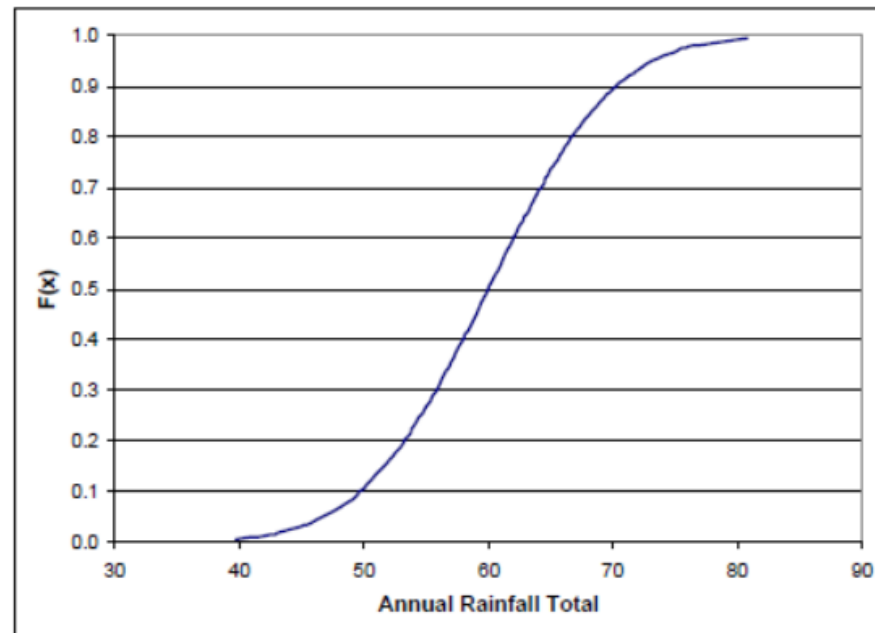


Figure 1. Variability in Total Annual Rainfall



a) Probability Density Form



b) Cumulative Distribution Form

Figure 3. Normal Distribution Function

Return frequency

- For example, if daily rainfall has exceeded 6 inches ten times in a thirty year record, the *return frequency* $F_R = n/t$, or $10/30 = 0.33$, or 33.3%
- The *return period* is $1/F_R$, or 3 years. This is the average interval between the events of a particular magnitude
- A 1% return frequency is a 100-year return period
- A 0.2% return frequency is a 500-year return period
- A 1E-4% return frequency is a 10,000-year return period

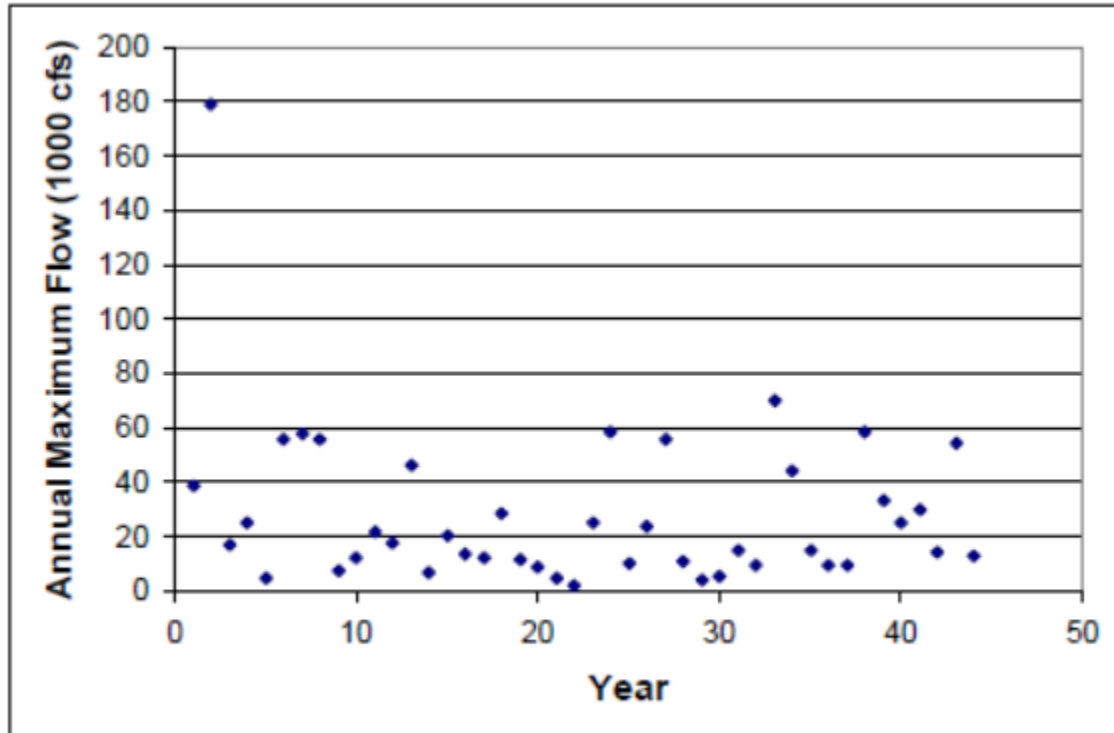
Does not mean a flood occurs every “n” years!

For example, the probability of a 100-year flood occurring in 30 years (the lifetime of the average home mortgage) is 26.0%.

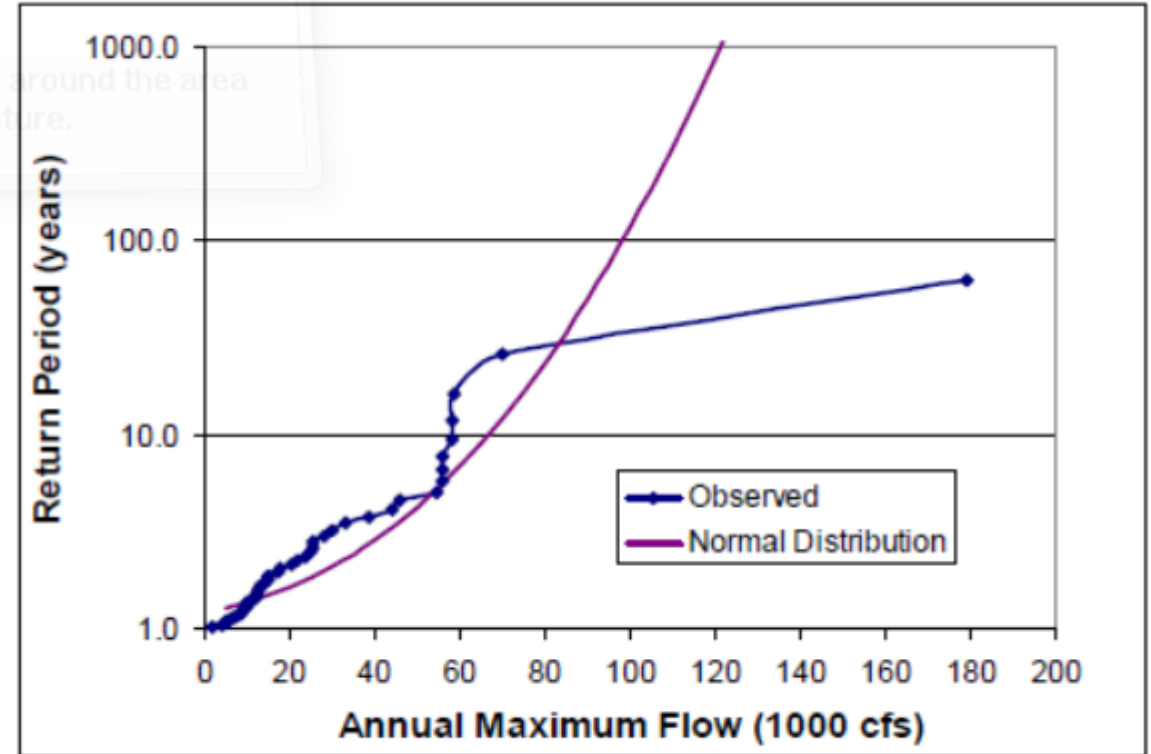
It could even occur multiple times in a century.

Chances of Being Flooded				
	Flood Level			
Period of Time	10-yr Flood	25-yr Flood	50-yr Flood	100-yr Flood
1 year	10%	4%	2%	1%
10 years	65%	34%	18%	10%
20 years	88%	56%	33%	18%
30 years	96%	71%	45%	26%
50 years	99%	87%	64%	39%

Extreme return frequencies



a) Data Record



a) Observed versus Normal Distribution Return Period

Figure 12. Example Annual River Flood Data

Need to use a different probability distribution. Examples are log-normal, log-Pearson Type III, and Generalized Extreme Value (GEV). GEV Type I is called a Gumbel distribution, GEV Type II is a Frechet distribution, and GEV Type III is a Weibull distribution.

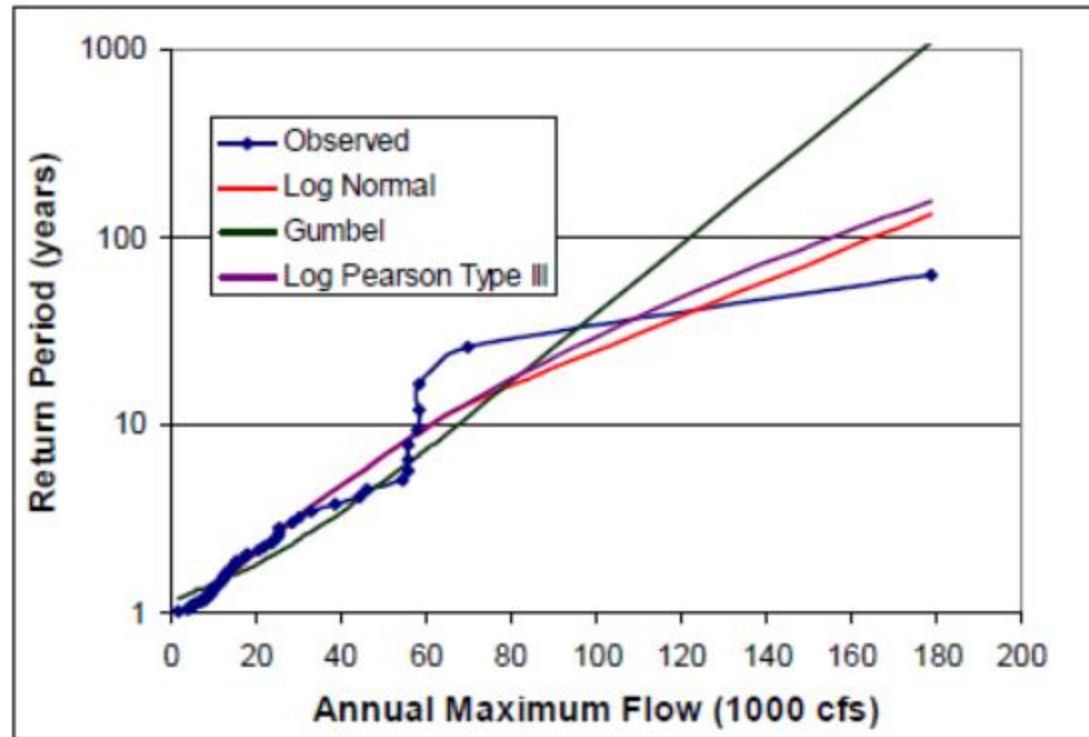
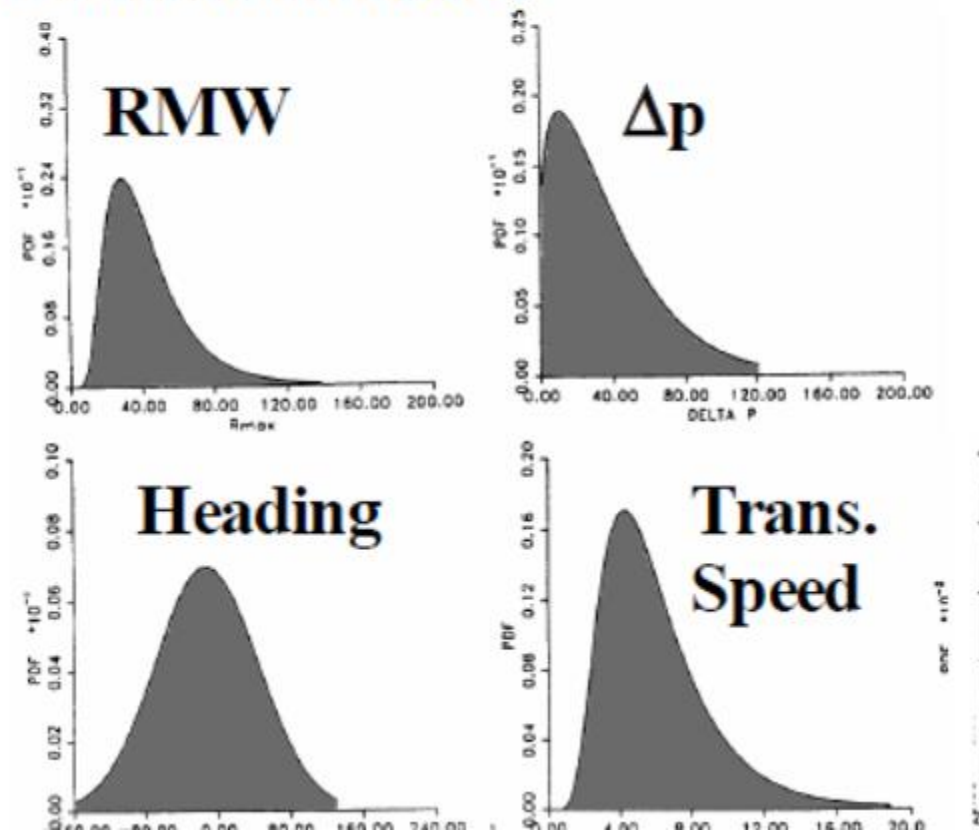


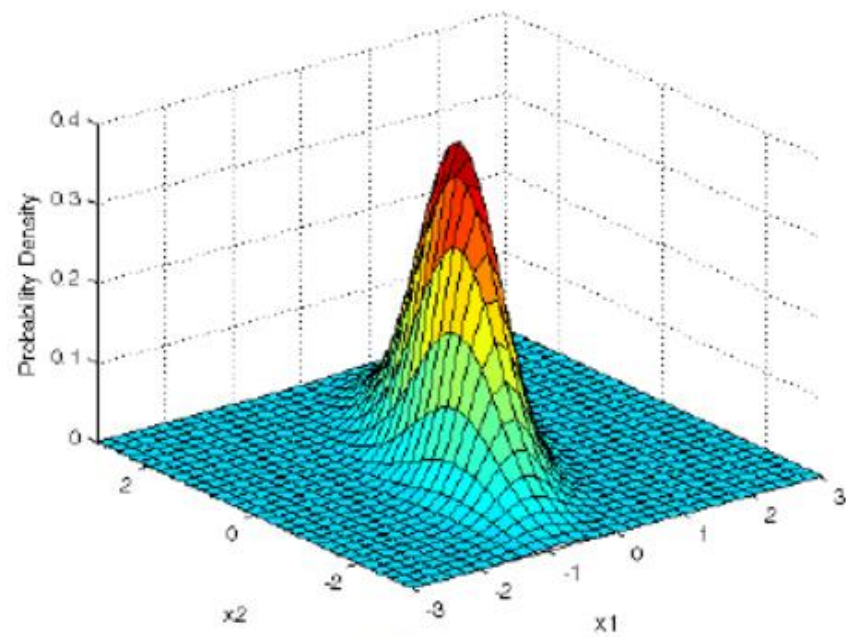
Figure 14. Frequency Distributions for Annual River Flood Data

Site-Specific Probability Distributions

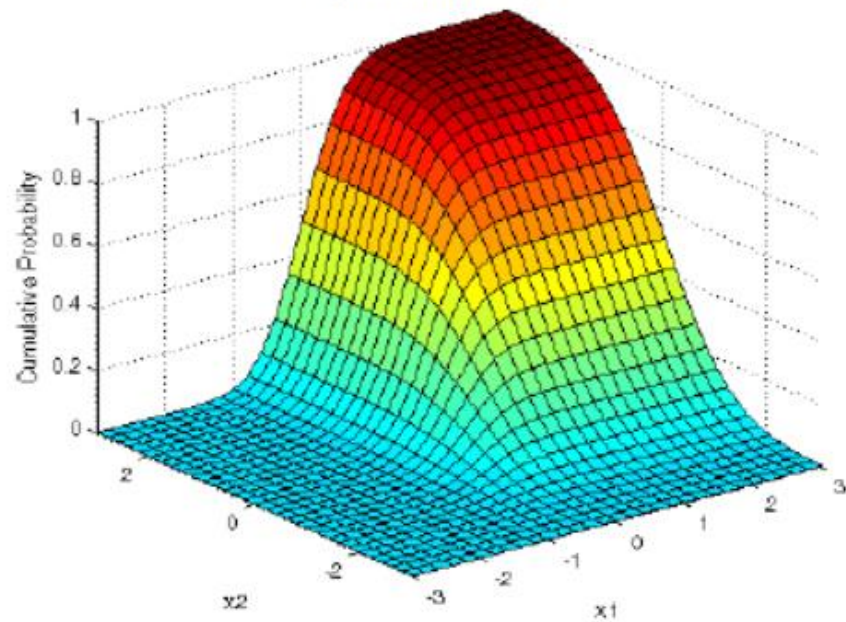


Concept - Joint probability

- Joint probability refers to the likelihood of two or more conditions occurring at the same time.
- The joint probability for two events, A and B, is expressed mathematically as $P(A,B)$. It is calculated by multiplying the probability of event A, expressed as $P(A)$, by the probability of event B, expressed as $P(B)$.
- The probability of two rolled dice simultaneously being the number five is $(1/6) \times (1/6) = 0.02777$
- However, the dependence between the two or more conditions should be non-trivial, i.e. neither independent nor fully dependent.
- It is a useful statistic to use when two or more observable phenomena can occur simultaneously



a) Probability Density



b) Cumulative Probability

Figure 17. Illustration of Joint Probability for Two Events

<http://www.mathworks.com/help/toolbox/stats/brn2ivz-89.html>

JPM usefulness for storm surge

- Earlier techniques (tide gauge analysis or the Empirical Simulation Technique) are sensitive to sample size in tide or surge simulation datasets, and cannot capture the range of storm possibilities capable of producing, for example, the 1% annual flood event.
- The JPM approach, however, has the conceptual advantage of considering all possible storms consistent with the local climatology, each weighted by its appropriate rate of occurrence.
- The most basic JPM approach assumes a parametric storm description involving several hurricane descriptors, such as:
 - Central pressure
 - R_{\max}
 - Storm direction
 - Storm speed
- Appropriate probability distributions are determined for each parameter and discretized
- All possible parameter combinations (each defining a synthetic storm) are simulated using a storm surge model

Math formalism (from Toro paper)

$$P[\eta_{\max(1\text{ yr})} > \eta] = \lambda \int \dots \int_{\underline{x}} f_{\underline{x}}(\underline{x}) P[\eta(\underline{x}) > \eta] d\underline{x} \quad (1)$$

where λ is the annual rate of storms of interest, vector \underline{x} represents the storm characteristics and $f_{\underline{x}}(\underline{x})$ represents the joint probability distribution of these characteristics, $P[\eta(\underline{x}) > \eta]$ is the probability that a storm of characteristics \underline{x} will generate a flood elevation in excess of an arbitrary value η .

For numerical calculations, the multiple integral in Equation 1 is approximated by a summation over a discrete set of storm-parameter values, as in:

$$P[\eta_{\max(1\text{ yr})} > \eta] \approx \sum_{i=1}^n \lambda_i P[\eta(\underline{x}_i) > \eta] \quad (2)$$

where each term in the summation corresponds to one combination of storm parameters (i.e., one synthetic-storm), with annual rate λ_i

Sometimes uncertainty is included through an ε term

$$P[\eta(\underline{x}) > \eta] = P[\eta_m(\underline{x}) + \varepsilon > \eta] \quad (3)$$

η_m is the model calculated surge height

The results from the production runs provide the peak flood elevation at each grid point for each of the synthetic storms. The first step in the surge frequency calculations is to construct the complementary cumulative distribution function (CCDF) of surge elevation (sometimes called the hazard curve) at each grid point. The process is simple and involves the evaluation of Equations 2 and 3, which may be combined into Equation 4 as follows:

$$P[\eta_{\max(1 \text{ yr})} > \eta] \approx \sum_{i=1}^n \lambda_i P[\eta_m(\underline{x}_i) + \varepsilon > \eta] \quad (4)$$

where $\eta_m(\underline{x}_i)$ is the peak flood elevation obtained by the production run for storm i at the grid point being considered. This calculation is performed for many values of the flood-elevation of interest η (using an increment of 0.1 ft), obtaining a CCDF

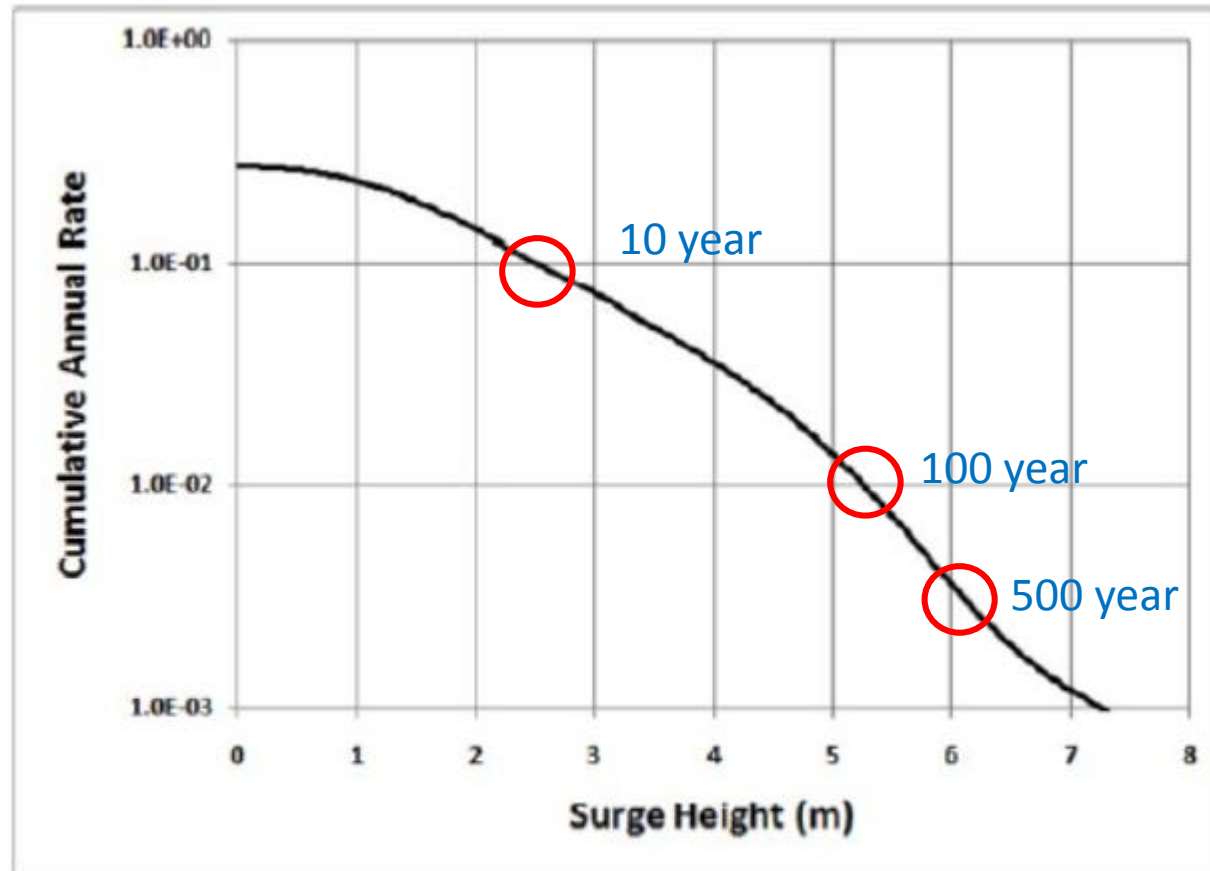


Figure 10. Example of complementary cumulative distribution function (CCDF) of maximum annual surge height at one grid point.

How “100-year” surge event is determined (full JPM)

- Develop probability distributions for each storm parameter (R_{\max} , intensity, etc.) from observations
- Establish rate of storm occurrence in space and time
- Subdivide each distribution into a small number of discrete pieces (i.e., 6 values)
- Construct all possible hypothetical tracks by taking all possible combinations of the storm quantities. For example, with six values for four parameters one constructs 1296 “storms.” (=6 pressure X 6 R_{\max} X 6 direction X 6 speed)
- Conduct hydrodynamic simulations (surge model, wave coupling, sometimes hydrology) with multiple tracks for each storm type sufficiently spaced for shoreline influence (landfall and bypassing). Track spacing is typically one R_{\max} , or about ten tracks per site (12,960 simulations)
- For each storm, compute highest surge for locations of interest, tag it with rate of occurrence
- Construct a histogram of rate versus surge height
- Find the 1% surge elevation for each location

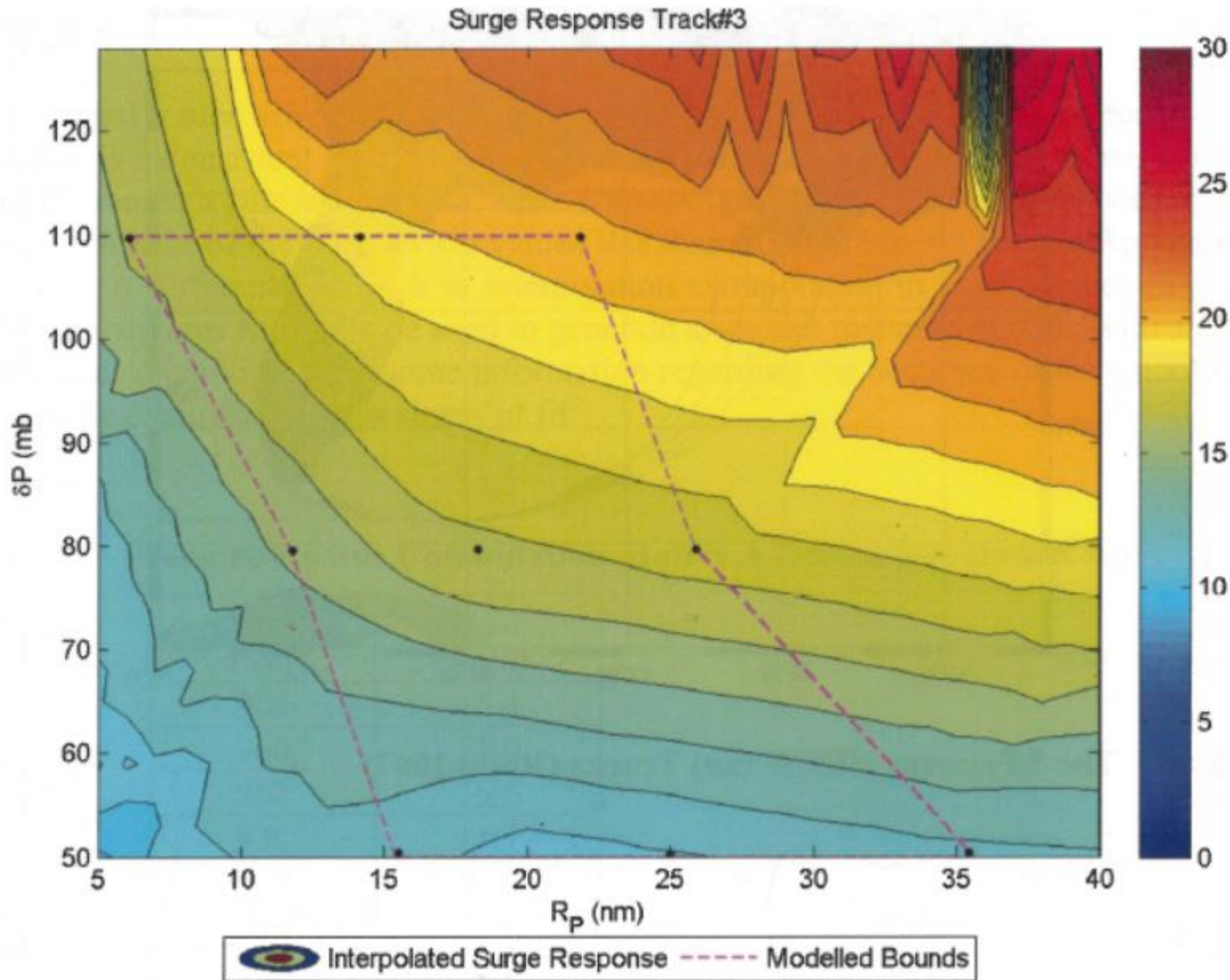
“Optimized sampling” (OS)

- JPM fine if only SLOSH is needed, but not for high-resolution runs using ADCIRC
- JPM-OS techniques seeks to reduce the number of simulations in an intelligent way (fewer combinations, tracks) while maintaining accurate frequency return values
- JPM-OS are sometimes tested with JPM SLOSH runs to see if the JPM-OS assumptions roughly match on open coast sections
- One possibility is a Monte Carlo JPM which randomly selects a sample from all possible combinations but enough to hopefully sufficiently capture the proper range of possibilities. However, sample size questions still exist.
- Two of the established techniques (next two slides) are:
 - Response Function JPM-OS (JPM-OS-RS)
 - Quadrature JPM-OS (JPM-OS-Q)

Response Surface Method

- Restricts parameters based on sensitivity response experiments (i.e. only three pressure values chosen). It is found certain combinations are linear, some responses stronger than others, and “smooth”
- Carefully choosing parameters limits combinations, and reduces simulations
- Typical example of steps:
 - Step 1: Start with ~5 tracks roughly perpendicular to landfall region and a few values of p and R_{\max} . Conduct the simulations. Interpolate or extrapolate other surge values in the p - R_{\max} plane
 - Step 2: Add a few more oblique angles ($\pm 45^\circ$), simulate on a reduced p - R_{\max} combination (compared to Step 1), interpolate/extrapolate
 - Step 3: Vary by a few storm speed parameters, simulate on a further reduced p - R_{\max} combination (compared to Step 2), interpolate/extrapolate
 - Step 4: Interpolate/extrapolate in track space for one p and R_{\max}
- This process can yield over 50,000 storms.
- Problems are in choosing the proper parameters restrictions (needs expert judgment) which can also be arbitrary; the accuracy of the interpolation; and the use of extrapolation. But results compare well to JPM-OS-Q.
- An alternate interpolation scheme is known as the *sparse grid method*.

Example of interpolation/extrapolation in Step 1 for one track



Interpolated/extrapolated surge response function in Δp - R_{\max} plane for one track. Δp is central pressure minus environmental pressure.

Black dots indicate 9 simulated storms for this track. The magenta dashed polygon indicated where bilinear interpolation is performed.

Below 110 mb, and to the right and left of the polygon, the response surface is extrapolated by maintaining a constant Δp - R_{\max} gradient from the edge of the polygon.

Above 110 mb, the surge response function is extrapolated by maintaining a constant p gradient.

Quadrature Method

- Seeks parameters and annual rates for several hundred runs by minimizing the mean square error of integration
- This minimization algorithm selects the optimal parameter combinations and assigns representative weights to each of the combinations
- This turns the multi-dimensional JPM integral into a weighted summation with specific weights for the optimal parameter combinations.
- Various assumptions are made to simplify integral. Need to test against SLOSH JPM. This also helps determine the optimum number of tracks.

Optimal Methods for extremely rare events (1 in 10,000 year)

- 1) Procedures are similar, but conduct thousands of SLOSH runs to determine optimal combinations for the peak surge events, then refine with several dozen ADCIRC runs. Derive response frequency curves.
- 2) Then do a “check” with deterministic ADCIRC runs with realistic but extreme hurricanes and quasi-perpendicular angles, and see if the results roughly match the probabilistic results. If not, there may be an issue with the synthetic tracks or the pdf assumptions.

Generation of synthetic storm datasets

- 1) Intensity, R_{\max} , direction qualitatively based on local climatology; spaced by R_{\max}
- 2) Empirical equations based on climatology, and include random error (Vickery et al. 2000)
- 3) Model downscaling
 - a) Climate model
 - b) Beta and Advection Model (BAMS) coupled to a balanced vortex model with 1D ocean coupling (Emanuel 2006a,b). Steering currents (850 and 200 mb) are from a general circulation model.

References

- Vickery, P. J., P. F. Skerlj, and L. A. Twisdale, 2000: Simulation of hurricane risk in the U.S. using empirical track model. *J. Structural Engineering*, **126**, 1222-1237.
- Emanuel, K., 2006a: A statistical-deterministic approach to hurricane risk assessment. *Bull. Amer. Meteor. Soc.*, **87**, 299-314.
- Emanuel, K., 2006b: Climate and tropical cyclone activity: A new model downscaling approach. *J. Climate*, **19**, 4797-4802.

Empirical track generation equations

$$\Delta \ln c_i = \ln c_{i+1} - \ln c_i = a_1 + a_2 \times \Psi_i + a_3 \times \lambda_i + a_4 \times \ln c_i + a_5 \times \theta_i + a_6 \times \theta_{i-1} + \varepsilon_{i1} \quad (1)$$

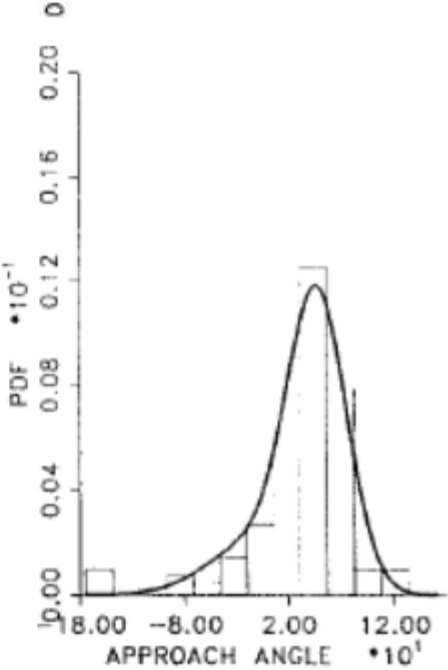
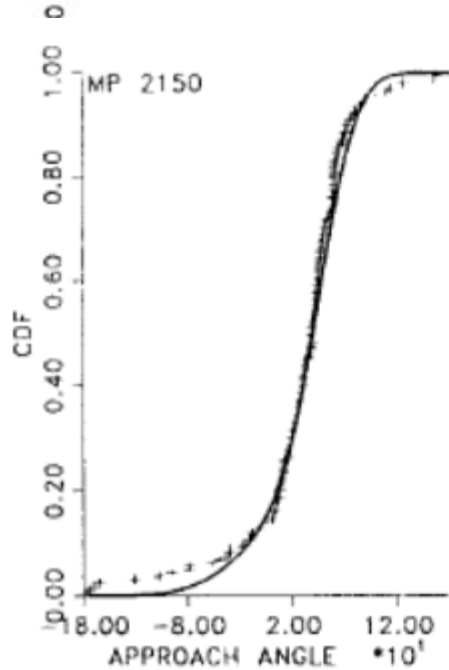
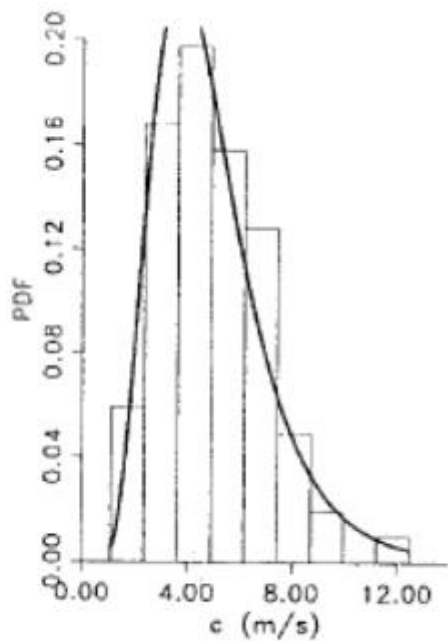
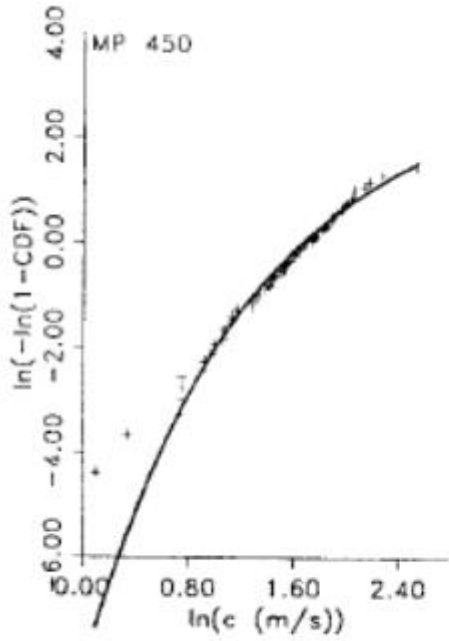
$$\Delta \theta_i = \theta_{i+1} - \theta_i = b_1 + b_2 \times \Psi_i + b_3 \times \lambda_i + b_4 \times c_i + b_5 \times \theta_i + b_6 \times \theta_{i-1} + \varepsilon_{i2} \quad (2)$$

$$\Delta u_i = u_{i+1} - u_i = c_1 + c_2 \times T_i + c_3 \times T_{i-1} + c_4 \times u_i + c_5 \times u_{i-1} + \varepsilon_{i3} \quad (3)$$

Pressure more common, but wind can be used too

where Ψ and λ = latitude and longitude of typhoon center, respectively; c = translation speed; θ = heading; T = sea surface temperature; u = maximum sustained wind speed; ε = random error.

The symbol Δ situated before each parameter denotes the change of this quantity over the current period; the subscripts i and $i + 1$ specified for each parameter express the value at the corresponding time step; and a_1, a_2, \dots is a set of constants for each rectangular grid cell within the research area. When the hurricane travels from one grid cell to another, these values are changed accordingly. The constants are computed using a multiple linear regression solution. Because this process is repeated until the synthetic typhoon makes landfall or final dissipation over the sea, a full track is created, along with all main parameters, at each time step.



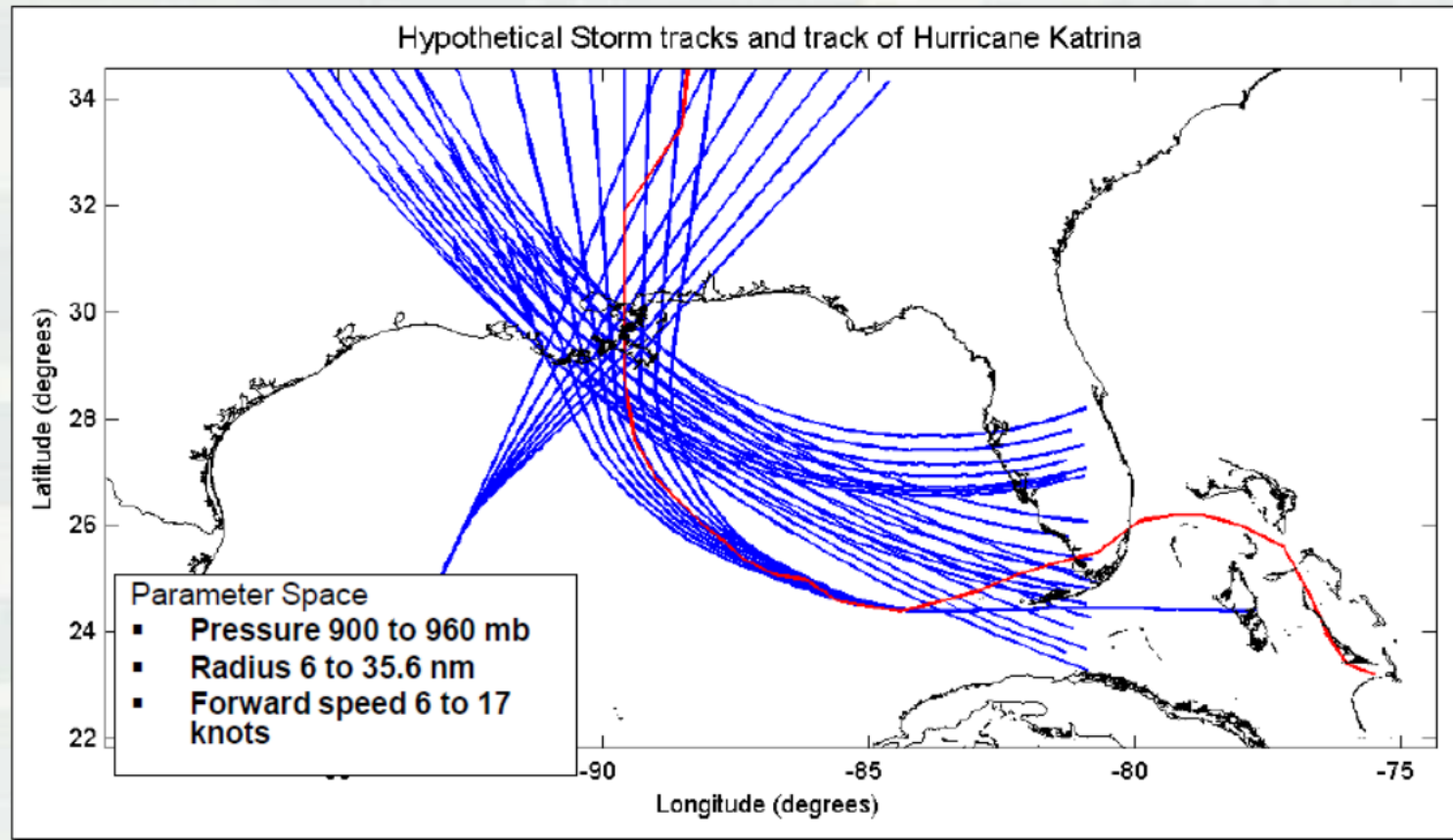
Validation examples from Vickery et al. (2000)

Example applied to the post-Katrina
New Orleans levees reconstruction

Table 1. Summary of the 152 HSDRRS JPM-OS hurricane tracks, stratified by central pressure, radius of maximum winds, translation speed, track direction, primary and secondary plus intensity (Saffir-Simpson scale), and number of storms in each group. From Jacobsen (2013).

GoM CP mb	GoM R_{max} miles	Landfall V_f mph	θ direction from	Track Set (Number)
960	40.9	12.7	Central	P (5)
	28.3	12.7	SE	P (4)
			SW	P (4)
	24.2	12.7	Central	P (5)
	20.9	12.7	SE	P (4)
			SW	P (4)
	20.4	12.7	Central	S (4)
			SE	S (3)
			SW	S (3)
		6.9	Central	P (5)
12.7	12.7	Central	S (4)	
12.7	12.7	Central	P (5)	
930	29.7	12.7	Central	P (5)
	20.4	19.6	Central	P (5)
				S (4)
			SE	P (4)
			S (3)	
			SW	P (4)
			S (3)	
	12.7	Central	P (5)	
	6.9	SE	P (4)	
			S (3)	
	SW	P (4)		
	S (3)			
9.2	12.7	Central	P (5)	
900	25.1	12.7	Central	P (5)
	21.2	12.7	SE	P (4)
			SW	P (4)
	20.4	12.7	Central	S (4)
			SE	S (3)
			SW	S (3)
	6.9	Central	P (5)	
			S (4)	
	17.1	12.7	Central	P (5)
	14.4	12.7	SE	P (4)
SW			P (4)	
6.9	12.7	Central	P (5)	
3 CP	15 CP- R_{max}	19 CP- R_{max} - V_f	30 CPD- R_{max} - V_f - θ	152 Storms

Selection of Synthetic Storms



Example 100-year surge curves for southshore



Example 100-year surge curves for southshore

