PERFORMANCE OF THE JPM AND EST METHODS IN STORM SURGE STUDIES

David Divoky
Watershed Concepts
2835 Brandywine Road, Atlanta, Georgia 30341 USA

Donald T. Resio
U.S. Army Corps of Engineers, ERDC-CHL,
3909 Halls Ferry Rd., Vicksburg, MS 39180, USA

1. INTRODUCTION

This paper compares the performance of the Joint Probability Method (JPM; Myers, 1970) and the Empirical Simulation Technique (EST; Scheffner, et al, 1999) for estimation of hurricane surge frequency. These two statistical approaches have been widely used in past coastal flood studies by both the U.S. Army Corps of Engineers (Corps) and the Federal Emergency Management Agency (FEMA). Statistical simulation methods such as JPM and EST are required for coastal flood frequency analysis primarily because there is an insufficient historical record from which to derive frequencies by more conventional means, such as gage analysis. Hurricanes, for example, are both sporadic and of limited spatial extent, contributing to a great deal of sample variation (sample error) in local tide gage records.

Used in its original form, an EST estimate for a site is based entirely on the historical storms and flood levels observed at that site. Alternate life cycles are simulated by assuming that storm occurrence rate follows a Poisson distribution, and by implementing a bootstrap resampling from the set of observed events to construct synthetic records. Flood frequency and variability estimates are then derived from this synthetic data. More recent applications of EST have also considered hypothetical storms obtained by parallel displacement of the historical tracks so as to provide more uniform coastal coverage.

JPM, on the other hand, seeks to consider all possible storms, not just those particular combinations of storm characteristics that have been observed within the historical record. To achieve this, a number of assumptions regarding storm description and behavior must be made. For example, it is assumed that storms can be characterized by a small set of parameters that can be combined in numerous ways, such that the probability of a given combination — a given storm — is determined by the joint probability of the several parameters.

It is evident that both approaches have conceptual strengths. For example, EST is based entirely on storms that actually occurred within the study area (so that the sample is known to be consistent with the local population; this is not necessarily true of a JPM study). On the other hand, it is evident that some coastal regions may have been either lucky or unlucky in their recent history, experiencing too few or too many severe events. This sample variation appears more troublesome for EST than for JPM, since JPM considers all possible combinations of storm parameters thought to be consistent with local conditions, not just the random small observed set.

Considerable interest in the merits of the two approaches developed during recent post-Katrina/Rita work undertaken by the Corps and FEMA. The LaCPR, IPET, and FEMA studies in Louisiana and Mississippi all adopted the JPM approach in the belief that it is the better approach for *hurricanes*, if not necessarily for other sorts of storms. This difficult choice was based on the best interpretation of the available data and on the perceived strengths and limitations of the two methods, but remained a matter of considerable uncertainty and some controversy. The work reported here was undertaken in order clarify the relative performance issues.

2. APPROACH

The key idea of the approach is to first posit forms of the *hidden rules* of nature, then to use those rules to generate appropriately random synthetic "flood" events over periods of "history" and lengths of "coast," and from those results to exercise *both* EST and JPM to estimate flood frequencies. The frequencies estimated in this way can be compared to the *true* values which are implicit in the postulated hidden rules. A second key element of the approach is to adopt a very simple but adequate proxy for coastal floods in order to permit very large numbers of simulations to be made with minimal computational effort.

2.1. BASIC ASSUMPTIONS

Many factors are critical to real-world storm surge simulation that are of little or no significance to the essential features of this study. In a real application, one is concerned with the complex response of the sea over an arbitrary basin and terrain, including the effects of wind, pressure, waves, coriolis force, tides, and more. However, such complexity can only obscure the underlying features of interest here, so that a simple proxy idealization of the environment is preferred. Since a key factor is the size of the flood footprint compared to the spacing of storm tracks, the concept of *space* must be certainly be preserved, but it is sufficient to consider only the most general such space: a long, straight, uniform coast, intersected by random straightline storm tracks. We have assumed a straight coast 1000 miles in length.

Hydrodynamic considerations have been dispensed with entirely. Instead, simple functions of wind speed have been adopted as indicator functions, replacing surge calculations. Storm surge certainly mimics the onshore winds in the only way essential to our purposes: surge tends to be higher where the winds are greater, so that the shape of the alongshore surge footprint can be sufficiently well approximated by some function of windspeed; the details of the transfer function between winds and surge are of little interest beyond this simple fact. Similarly, the absolute magnitude of the response is not relevant, just as, for example, the problem is independent of the choice of unit used to measure surge. Representative wind speeds and directions can be computed very efficiently using any of a number of idealized wind field representations.

The approach requires adoption of a parameterized storm description. For this, the simple windfield model used in many past FEMA coastal flood insurance studies was chosen (FEMA, 1988). The model characterizes a hurricane windfield in terms of the usual five parameters: the central pressure depression, ΔP; the radius to maximum winds, R_{max}; forward speed of the storm, V_f; track direction, θ; and shoreline crossing point, X. The top curve in Figure 1 shows the shape of this windfield measured in units of R_{max} from the eye, and normalized by the peak wind speed. The bottom curve is the shape of the squared winds, discussed below, and the middle curve is representative of actual surge variation on straight coasts.

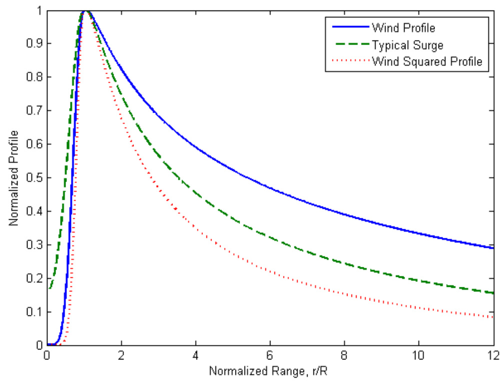

Figure 1 Wind and Surge Profiles

The unknown population was defined so as to be generally consistent with the realities of the northern Gulf of Mexico as characterized during recent post-Katrina studies (Resio, 2007; Toro, 2007). Each parameter was described by a probability distribution similar to, but not necessarily identical to, a distribution used in one of those studies (differences have been assumed deliberately, to ensure that the findings given here will *not* be construed as actual Gulf estimates). In particular: the pressure deficit ΔP has been taken to be Gumbel; the storm radius, R_{max}, is assumed gaussian and correlated with pressure; the forward speed of the storm, V_f is gaussian and independent of other parameters; the track direction, θ, is also independent and gaussian. The shoreline crossing point, X, has been assumed to be uniformly distributed.

The correlation postulated for R_{max} is a linear dependence of the mean radius on central pressure, such that greater central pressure depressions are associated on average with smaller storms; the coefficient of variance of R_{max} is taken to be constant so that the spread around the pressure-dependent mean is greater for weaker storms. Finally, the rate of storm occurrence was taken to be 5E-04 storms per kilometer per year, consistent with the rate of occurrence of the stronger storms in the northern Gulf.

The basic proxy for surge elevation was taken to be the square of the normal component of the wind at the center of the shoreline (the conceptual study site), normalized by a fixed constant so as to give response values on the order of ten to twenty, in accordance with one's expectation for surge magnitude; it must be remembered, or course, that the scale is entirely arbitrary and the results to be shown below should not be interpreted in any way as surge elevations for Gulf regions. The square was chosen since the wind stress is proportional to the square of the wind speed; the normal component was adopted as the obvious choice. This basic profile is illustrated by the bottom curve in Figure 1. Owing to all the ignored hydrodynamic processes, this profile is actually somewhat narrower than is typical of real surge along a straight coast. In order to reveal the effect of the width of the surge footprint, therefore, a wider profile was also considered, corresponding to the first power of the wind speed; this is shown by the upper curve in Figure 1. Typical surge profiles tend to lie in between these two, as shown by the middle curve (the upper curve inside the radius of maximum winds). This footprint variation is of interest since it is anticipated that EST performance may improve with increasing storm size (additional tests were made using still broader profiles corresponding to the 0.5 and 0.25 powers of the basic profile).

2.2. SIMULATION SETS

Within these idealized assumptions, a basic simulation set is defined by a selected period of record, T, and a conceptual coast length, L; as noted, the conceptual coast is 1000 miles long. A basic period of record of 65 years was chosen. This choice was made since it is approximately the same as the record used in the post-Katrina studies, and corresponds to the period since WWII during which adequate data exists. Two other periods were considered. A short record length of 30 years was investigated, since this is the common rule of thumb for the minimum length needed for a 100 year frequency analysis. A longer record of 140 years was also adopted, corresponding roughly to the entire hurricane record, and standing in the same proportion to 65 as 65 does to 30.

Multiple sets, *S*, of each period of record were simulated (50 and 10 sets of 65 year records, and 10 sets of 30 and 140 year records). Although these repetitions correspond to as many as 3,250 years, they are still inadequate to establish precise quantitative estimates of performance owing to the slowly-converging effects of sample variation. However, as will be seen, they are fully adequate to display the qualitative features of primary interest.

The study then proceeded by looping over the *S* periods of record, then over the *T* years within a period. For each year, the hypothetical number of storms was determined from the adopted Poisson distribution, and for each of those storms, parameters were randomly chosen from the postulated parent distributions.

2.3. JPM ANALYSES

Given the storms defined in the several simulation sets, the JPM analyses proceeded just as they do in practice. That is, for a given simulation set, the first order of business is estimation of the storm parameters from the sample. This includes the annual rate of occurrence and the distribution of each of the storm parameters except the crossing point which, as usual, was assumed uniform. A basic set of JPM analyses was performed using the same distribution types as chosen for the parent distributions, but fitted to the samples and without consideration of correlation between pressure and radius; that is, pressure and radius were treated as entirely independent in the basic set of JPM determinations.

Additional JPM cases were run using alternate assumptions regarding the form of the pressure distribution. These included gaussian, generalized normal (subsuming log normal), generalized logistic, and Pearson III representations. The specific forms adopted for these distributions followed Hosking and Wallis (1997). Finally, additional simulations were run incorporating the sample correlations between pressure and radius. In this case, each sample was used to determine the best linear fit of radius vs pressure, and the coefficient of variance was similarly estimated from the samples.

It might be objected that parameterization using the same families as used for the parent distributions would give an unwarranted advantage to the estimates, but this is not the case, owing to the large amount of sample variation. Furthermore, it was not thought appropriate to assume unrealistic distribution shapes, as might be justified were nature thought to be somehow malicious in hiding her cards. In fact, the broad nature of the distributions is known from observation, so that accepting them is reasonable. Note that the experiments using the gaussian, logistic, generalized normal, and Pearson III forms for the pressure samples were done blind; that is, no attempt was made to see which fit the sample best, and it is likely that the three-parameter distributions would perform better than the basic Gumbel, as far as apparent fit is concerned. But as will be seen below, the choice of the form of the sample pressure distribution was essentially irrelevant to the results.

Finally, all JPM estimates were made by a brute-force approach, chopping each parameter distribution into pieces, and simulating all combinations. No more sophisticated JPM-Optimal Sampling techniques were used (see Toro, 2007, and Resio, 2007, for a description of such methods) . For the basic JPM evaluations, 2016 storms were constructed from 8 pressures, 4 radii, 3 forward speeds, 3 track angles, and 7 crossing points. Sensitivity to the number of storms was tested using 5,600, 8,400, 22,400 and 23,520 storms in different cases.

2.4. EST ANALYSES

The EST analyses also followed precedent, based upon prior Corps studies for example, to determine stage frequency relationships from the same samples used with JPM. The EST effort was considerably less than for JPM, since it was not necessary to first estimate parameter distributions – the simulation sets provide all the information required by EST.

No effort was made to exercise the many flexible options which can be found inside the EST program. Instead, the intent was to perform a straightforward analysis using EST as it would usually be used (and has been used) in such an application. For this, we followed many helpful suggestions of Scheffner (2007) regarding the art of actual EST practice.

Only in one way did we depart from a simple EST evaluation of the bare historical record. That was in a set of additional EST simulations to investigate the effect of adding hypothetical storms by translating the historical tracks parallel to themselves to alternate shoreline crossing points, and considering those storms to be part of the historical record. A few words about how that was done are in order.

In some past EST studies, tracks have been arbitrarily added by replicating a real track some distance to both the left and right of itself, and dividing the probability mass of the storm accordingly. For example, replicants might be added 1 degree to each side, with the weight divided by three. The justification for this approach (suggested in the EST documentation) is not evident. Clearly, the result would be different if the division were made into eleven storms, say, and if the ten replicants were translated to positions spaced by, say, multiples of five degrees to the left and right. In such an extreme case, 10/11 of the probability weight of the storm would be lost from the study area. This is clearly unreasonable, but is not in principle different in any way from a division by three with nearby spacing.

In this study, an approach with some reasonable justification was used. Every storm in a given sample fell at some coastal crossing point, but could just as well have fallen at any other point, by the assumption of uniformity. Furthermore, each storm had a “rate” of once per T years (typically, once per 65 years). Now, imagine the coast of length L divided into equal segments of length dx. Then the chance that the storm might have passed through any such segment would be dx/L, with a rate of $dx/(LT)$. One imagines, then, including *all* such storms in the EST calculations. Of course, many of them would fall far from the site, not producing local surge, and so would not need to be included in the simulations. It is recognized that the rate associated with each replicant goes to zero by this method as the coastal length, L, goes to infinity, but this is not a problem since a larger L would bring more storms into the sample in a precisely compensating way. Ideally, the spacing dx would be taken to be about equal to the storm radius in order to smooth out alongshore variations. However, for reasons of simplicity, a fixed spacing of 20 miles was used in our tests.

3. RESULTS

3.1. ESTIMATES OF THE *TRUE* VALUES

The simulation sets included different assumptions regarding nature, such as the width of the surge footprint. For each such variant of the parent, there is a corresponding set of *true* values for, say, the 50, 100, and 500 year surge values.

True values have been determined by simply performing a JPM analysis using the postulated parent distributions, including the pressure-radius correlation. A very large number of simulations were done, in order to ensure that the true values were accurately evaluated. Both 8,400 storms and 22,400 storms were simulated, with good agreement between the two estimates. In each of the tables shown below, the appropriate true values are shown along with the JPM and EST estimates.

Attention has been focused on the 50, 100, and 500 year values. It had been intended to include 10 year values (of interest for FEMA flood insurance studies) , but to do so requires simulation of many more (weaker) storms and so was not done. In each case to be presented, we show the means, standard deviations, and coefficients of variation (CV) of the estimates. These are usually based on 10 simulation sets of 65 years each, but other values are shown as noted. Tests were performed with 50 sets, and with periods of record of 30 and 140 years as previously discussed. In some cases, we show values for each of the ten simulation sets, so as to display the magnitude of the sample variation most vividly.

3.2.1. BASIC CASE, 10 SIMULATION SETS

This case included 10 sets of 65-year records, with the following findings for JPM and EST, as identified in each table. The JPM estimates do not account for correlations between pressure and radius found in the samples; complete independence of parameters is assumed. The *true* values, of course, reflect the assumed parent correlation.

JPM	50 yr	100 yr	500 yr
1	6.3	7.8	11.0
2	5.9	7.5	10.4
3	5.6	7.0	9.8
4	6.2	7.6	10.4
5	6.1	7.7	10.5
6	6.3	7.9	11.3
7	5.8	7.3	10.3
8	6.5	8.0	11.1
9	6.0	7.6	10.8
10	6.0	7.3	10.1
Ave	6.07	7.57	10.57
SD	0.27	0.31	0.47
CV	0.04	0.04	0.04
True	5.7	7.1	10.0

All estimated values are somewhat high, a characteristic which was observed throughout this work. Despite this, the variation is relatively small from set to set, as reflected in the coefficients of variation which are only 4% of the means. The 100 year estimates for these ten cases range from 0.1 unit less to 0.9 units more than the true value of 7.1 units.

The EST results were as follows for the same data sets:

EST	50 yr	100 yr	500 yr
1	8.3	9.6	12.8
2	---	---	---
3	8.1	10.9	18.9
4	6.0	7.4	11.0
5	5.3	8.3	16.5
6	10.0	14.4	26.8
7	5.3	6.8	10.4
8	6.6	7.9	11.2
9	6.0	8.5	15.0
10	5.4	6.6	9.3
Ave	6.77	8.95	14.6
SD	1.66	2.44	5.52
CV	0.25	0.27	0.38
True	5.7	7.1	10.0

The most prominent feature of these EST results is the very large variability of the estimates for the ten 65 year records. The 100 year estimate, for example, varies between 6.6 and 14.4 units, with an average of 8.95. No results are shown for Set 2, owing to an insufficient number of storms in the sample for that 65 year period.

3.2.2. BASIC CASE, 50 SIMULATION SETS

In this case, 50 sets of 65 year periods were simulated. We show only the summary data for the JPM estimates.

JPM	50 yr	100 yr	500 yr
Ave	6.11	7.67	10.82
SD	0.36	0.39	0.60
CV	0.06	0.05	0.06
True	5.7	7.1	10.0

It can be seen that the results are not greatly different from the JPM findings with ten 65 year samples. The EST equivalents are also not significantly different from the 10 set EST case.

3.2.3. BASIC CASE, 140 YEAR RECORDS

In this case, based on 10 sets, each record consists of 140 years, but is otherwise determined in the same manner as the preceding cases. Both the JPM and EST estimates are shown; the true values, of course, are unchanged from before. Only the summary values are shown (variability can be inferred from the CV values).

JPM	50 yr	100 yr	500 yr
Ave	6.15	7.81	10.90
SD	0.18	0.23	0.39
CV	0.03	0.03	0.04
True	5.7	7.1	10.0

Interestingly, longer periods of record did not improve the accuracy of the JPM means, but did reduce the variability. This implies that the shorter records do a reasonably good job of capturing the distribution shapes.

The EST results were:

EST	50 yr	100 yr	500 yr
Ave	6.13	8.13	13.36
SD	1.29	1.28	2.79

CV	0.21	0.16	0.21
True	5.7	7.1	10.0

In each case, the EST estimates are improved with the longer records.

3.2.4. BASIC CASE, 30 YEAR RECORDS

Again, there were 10 sets of simulations in this experiment, but each consisted of only 30 years of data. The results were as follows:

JPM	50 yr	100 yr	500 yr
Ave	6.19	7.74	10.89
SD	0.62	0.68	1.09
CV	0.10	0.09	0.10
True	5.7	7.1	10.0

The JPM averages were degraded only slightly, although the variability was substantially increased for these shorter records. For EST, the results were:

EST	50 yr	100 yr	500 yr
Ave	6.33	8.47	13.80
SD	1.74	2.21	4.51
CV	0.27	0.26	0.33
True	5.7	7.1	10.0

Interestingly, the means were not much degraded from the 140 year case, although the variability is greater. Even more interestingly, the 65 year results are somewhat worse than these, a finding that apparently must be attributed to the dominant effects of sample variation, and the need for much larger simulation sets to reach more definitive conclusions. However, the results clearly demonstrate the quandary of an analyst who must work with only a single data set of perhaps as many as 65 years.

3.3.1. WIDER SURGE FOOTPRINT (1.0)

In this case, we repeated the basic case of 10 sets of 65-year records, but adopted a wider surge profile (the top curve in Figure 1) based on the 1.0 power of the wind shape factor. The expectation was that EST would improve in this case, since there is a greater chance of the study site being affected by larger surge. The summary results were:

JPM	50 yr	100 yr	500 yr
Ave	7.04	8.25	10.81
SD	0.23	0.28	0.43
CV	0.03	0.03	0.04
True	6.7	7.9	10.7

Note that the *true* values are different for this case. The EST results were:

EST	50 yr	100 yr	500 yr
Ave	8.14	10.22	15.31
SD	1.52	2.27	4.79
CV	0.19	0.22	0.31
True	6.7	7.9	10.7

3.3.2. WIDER SURGE FOOTPRINT (0.25)

The preceding EST results are, in fact, better than the EST estimates made with the narrower surge profile. In view of this, further widenings were tested, using the 0.5 and 0.25 powers of the windfield shape factor (not shown in the figure). The following summary results are for the more extreme of these cases (0.25). Again, note that the *true* values are different from those before.

JPM	50 yr	100 yr	500 yr
Ave	8.30	9.68	11.64
SD	0.23	0.24	0.42
CV	0.03	0.03	0.04
True	7.9	9.2	11.9

For the EST determinations, the corresponding findings were:

EST	50 yr	100 yr	500 yr
Ave	9.83	12.01	16.87
SD	1.35	1.84	3.36
CV	0.14	0.15	0.20
True	7.9	9.2	11.9

In this case, the relative errors are comparable to the 1.0 power case, but the variability is substantially reduced, as might be expected since the chance of missing a site is less with a wider footprint.

3.4. HYPOTHETICAL TRACKS (EST)

As a final EST experiment, the basic case, discussed first, was re-done with *all* storms replicated at 20 mile intervals along the shoreline, and with the storm weights divided accordingly. There is no JPM equivalent for this case.

EST	50 yr	100 yr	500 yr
1	6.2	8.4	15.2
2	5.9	8.0	14.2
3	5.6	7.8	14.0
4	6.3	8.5	15.3
5	5.9	8.0	14.9
6	6.4	8.7	15.3
7	5.4	7.3	13.9
8	6.4	8.8	16.1
9	5.8	8.1	14.7
10	5.8	7.9	13.9
Ave	5.95	8.13	14.75
SD	0.35	0.45	0.73
CV	0.06	0.05	0.05
True	5.7	7.1	10.0

These results are substantially better than the basic EST estimates made without hypothetical tracks, especially in terms of the variability of the result, which now shows a CV of only about 5% of the mean vs 30% without hypothetical tracks. The means are slightly improved (except for the 500 year estimate which is negligibly worse; however, see a concluding comment about the 500 year values).

3.5.1. JPM WITH GAUSSIAN PRESSURES

The remainder of the results to be shown concern only JPM, with no EST counterparts. Since the findings are not strikingly different from earlier material, only summary results will be shown. The first result is for a JPM estimate using the basic 10 sets of 65-year records, but using a gaussian distribution to fit the sample pressure data. One might expect this to give a poor quality result, owing to the lack of similarity between a symmetrical gaussian and a more reasonable Gumbel distribution.

JPM	50 yr	100 yr	500 yr
Ave	6.16	7.69	10.48
SD	0.26	0.27	0.34
CV	0.04	0.03	0.03
True	5.7	7.1	10.0

The result, however, is seen to confound the expectations. The averages are only slightly worse than the fits obtained using a Gumbel distribution, and the variability is actually slightly improved. This result also goes to support the argument made earlier that adopting the same distribution types for both the parent and the samples is not an unreasonable choice.

3.5.2. JPM WITH GENERALIZED NORMAL PRESSURES (LOG-NORMAL)

This experiment duplicates the previous case, but with the generalized normal distribution chosen rather than a simple gaussian. The generalized normal is a three parameter distribution including both gaussian and log-normal distributions as special cases (see Hosking and Wallis , 1997).

JPM	50 yr	100 yr	500 yr
Ave	6.06	7.63	10.61
SD	0.25	0.28	0.47
CV	0.04	0.04	0.04
True	5.7	7.1	10.0

The results are negligibly different from the preceding case, or from the basic case using Gumbel sample fits.

3.5.3. JPM WITH GENERALIZED LOGISTIC PRESSURES

This continues the ideas of the previous two tests, but this time with the generalized logistic distribution for the fits (see Hosking and Wallis for details of the distribution).

JPM	50 yr	100 yr	500 yr
Ave	6.04	7.61	10.44
SD	0.23	0.26	0.46
CV	0.04	0.03	0.04
True	5.7	7.1	10.0

As before, the results change little with a different distribution.

3.5.4. JPM WITH PEARSON III PRESSURES

Finally, the following table shows a summary of the findings with JPM and the basic 10 sets of 65-year records, but with pressures fit using the Pearson III distribution.

JPM	50 yr	100 yr	500 yr
Ave	6.07	7.63	10.63
SD	0.25	0.28	0.47
CV	0.04	0.04	0.04
True	5.7	7.1	10.0

The lack of any significant change is remarkable.

3.6. JPM WITH P:R CORRELATION

As a final experiment, we show the JPM results for the basic case of 10 sets of 65-year records, but with the sample correlations between pressure and storm radius accounted for. This was done in a manner similar to that used in the post-Katrina Gulf studies: a regression relationship was first determined between the sample radii and pressures, in order to establish a mean radius as a function of pressure. The distribution of the radius around this mean was assumed to be gaussian, with a fixed coefficient of variation; that is, larger means imply larger standard deviations. The results were:

JPM	50 yr	100 yr	500 yr
Ave	6.02	7.52	10.38
SD	0.29	0.36	0.55
CV	0.05	0.05	0.05
True	5.7	7.1	10.0

One should compare these results with the very first table, which was for the same conditions except that full independence was assumed. At each recurrence level, these new results accounting for the sample correlations are negligibly better for the averages than the case assuming independence, and are negligibly worse in terms of variability.

In order to understand the reason for this lack of improvement, the correlation data for 50 sets of 65-year records was inspected. It was found that in 24 of those 50 simulated records, the sample correlation was of *opposite sign* to the *true* correlation postulated for nature's distributions. In other words, in nearly half of the simulated data sets, the effect of sample variation was to show an *increase* of mean radius with an increase of ΔP, contrary to assumption. On reflection, this is not too surprising, although it remains very troublesome as a practical matter. *All* samples will exhibit a correlation, even if the parent distributions are entirely *uncorrelated*. One can see this easily by tossing a pair of coins ten times, creating two ten-element series; the sample correlations between the two series will almost always be found to be significant, but are entirely spurious. This observation merits careful reflection and some additional investigation since the real-world analyst is provided with only one relatively short data set from which to infer the nature of the correlations. Arguments based on storm physics become crucial to the justification of sample estimates of the correlations.

4. DISCUSSION AND CONCLUSIONS

The two primary conclusions of the study are:

1) JPM was found to be remarkably robust, even with relatively short data sets, and is surprisingly insensitive to choice of distribution forms and to parameter correlations (within the ranges tested, which do, however, approximate what is known of nature). In the results shown here, JPM estimates were typically biased high, but only by about 5%, and CVs were usually only a few percent of the mean. To some degree, this robust performance at the 100-500 year level indicates that the more uncertain tails of the assumed distributions do not dominate the results (although they become increasingly important for more rare events).

2) EST was found to be extremely sensitive to sample variation of the sort associated with hurricane surge (but not necessarily with other less-sporadic and more wide-spread processes such as northeasters). Variability of an EST estimate over ten replications of the period of record from a fixed parent population often exceeded a factor of two, and CVs of 20% and more were typical. As with JPM, the averages were biased somewhat high.

Interestingly, inclusion of the sample P:R correlations did not improve the JPM estimates. Contrary to the underlying correlation posited for nature (based on real data), the simulations produced opposite P:R correlations nearly half the time, based on 65-year samples. A major sort of correlation not considered in this work is associated with track angle in the event that storms can both enter and exit a coast (the Florida peninsula, for example). In such a case, the correlation between angle and all other storm parameters is strong, and has always been recognized and accounted for in past JPM studies by the simple artifice of dividing the storm population into entering and exiting families.

The variability found for the EST tests is of particular concern for an individual coastal flood study. Even if a method performs extremely well on the average, this is of little comfort if results for a particular period of record can vary by a factor of two. The overall mean performance in a national program might not be much affected by such variability, with errors at many sites averaging out; still, the expected error for any/every particular study might be unacceptably large.

The work reported here is illustrative, only, and cannot be read as a final statistical evaluation of the JPM:EST merits. Both approaches can probably be implemented in better ways than the simple methods used here. For example, no advantage has been taken of recent advances in JPM methodology (Resio, 2007, Toro, 2007). JPM fits were all blind, with no effort to choose distributions exhibiting best-fit performance. Similarly, EST gives the user wide latitude in choice of internal assumptions and methods, none of which have been exercised here. For example, all EST tests shown were done using four nearest neighbors in the random walk interpolations (although experiments with three to six neighbors produced indistinguishable results). Only 500 years were considered in a single EST life-cycle (although 100 such runs were made in each case). In a typical test, doubling the length of an EST period to about 1000 years improved the 500 year mean estimates by about 2% (and the more frequent levels by lesser amounts). Rather, the goal has been to show what a coastal engineer would be likely to encounter with each approach in a typical case. The findings seem clear, and confirm the anticipation of the LaCPR, IPET, and FEMA post-Katrina/Rita teams that EST estimates would suffer excessively from sample error. The commonly asserted JPM deficiency regarding poor handling of storm parameter correlations was not observed in our tests of the pressure-radius correlation (in fact, the limited experiments done here showed that JPM seems quite robust in this regard) although additional investigations of this matter would be useful.

It is noted that the degree of EST improvement that might be gained by properly replicating historical tracks along the coast (as opposed to the sort of arbitrary replication that has been done in the past) appears to be very small. The primary improvement is in reduction of alongshore variability rather than in improved mean estimates at a point. This is because the tracks being replicated are still only derived from the limited sample, and do nothing to account for the storm possibilities not included in that sample. It is also noted that increasing the number of storms through track replication largely negates the perceived advantage of EST regarding study economy. For each of the ten 65-year sets simulated here with hypothetical tracks, the number of storms in a set was on the order of 400-500, each of which would require expensive hydrodynamic modeling in an actual application (whereas recent JPM applications appear to have been successful with about half that many model simulations).

Finally, we emphasize that these findings apply only to hurricane surge, and not to other applications for which EST may perform well, and may be the tool of choice. Numerous examples of such other applications are summarized in Scheffner, et al (1999).

5. ACKNOWLEDGMENTS

The authors wish to thank Dr. Norman Scheffner for his advice and assistance with EST and its application. Numerous helpful discussions with Drs. Robert Dean, Gabriel Toro, Alan Niedoroda, and Peter Vickery, during the course of several post-Katrina studies, are also gratefully acknowledged. This work was supported in part by the U.S. Army Engineer Philadelphia District under Purchase Order W912BU-07-P-0208.

6. REFERENCES

Federal Emergency Management Agency, (1988). Coastal Flooding Hurricane Surge Model; 3 Volumes. Washington, D.C.

Hosking, J.R.M. and J.R. Wallis, (1997). *Regional Frequency Analysis*, Cambridge University Press, Cambridge, U.K. See also J.R.M. Hosking (2000). Fortran Routines for use with the method of L-moments. Research Report RC 20525, IBM Research Division, Watson Research Center.

Myers, V.A. (1970). Joint Probability Method of Tide Frequency Analysis Applied to Atlantic City and Long Beach Island, N.J. ESSA Technical Memorandum WBTM Hydro-11. US Dept of Commerce, Washington, D.C.

Resio, D.T. (2007). White Paper on Estimating Hurricane Inundation Probabilities. Version 12a, May 27.

Scheffner, N.W. J.E. Clausner, A. Militello, L.E. Borgman, B.L. Edge, and P.J. Grace (1999). Use and Application of the Empirical Simulation Technique: User's Guide. U.S. Army Corps of Engineers, Engineer Research and Development Center, Technical Report CHL-99-10.

Scheffner, N.W. (2007). Personal communication.

Toro, Gabriel (2007). Joint Probability Analysis Of Hurricane Flood Hazards For Mississippi. Report prepared by Risk Engineering for URS Group in support of the FEMA-HMTAP flood study of the State of Mississippi. October.