

Article

Bias and Efficiency Tradeoffs in the Selection of Storm Suites Used to Estimate Flood Risk

Jordan R. Fischbach ^{1,2}, David R. Johnson ^{2,3,*} and Kenneth Kuhn ^{2,3,†}¹ RAND Corporation, 4570 Fifth Ave., Ste. 600, Pittsburgh, PA 15213, USA; jordanf@rand.org² Pardee RAND Graduate School, 1776 Main St., Santa Monica, CA 90401, USA; kkuhn@rand.org³ RAND Corporation, 1776 Main St., Santa Monica, CA 90401, USA

* Correspondence: djohnson@rand.org; Tel.: +1-765-494-7122; Fax: +1-765-494-7693

† These authors contributed equally to this work.

Academic Editor: Rick Luetlich

Received: 12 July 2015; Accepted: 28 January 2016; Published: 15 February 2016

Abstract: Modern joint probability methods for estimating storm surge or flood statistics are based on statistical aggregation of many hydrodynamic simulations that can be computationally expensive. Flood risk assessments that consider changing future conditions due to sea level rise or other drivers often require each storm to be run under a range of uncertain scenarios. Evaluating different flood risk mitigation measures, such as levees and floodwalls, in these future scenarios can further increase the computational cost. This study uses the Coastal Louisiana Risk Assessment model (CLARA) to examine tradeoffs between the accuracy of estimated flood depth exceedances and the number and type of storms used to produce the estimates. Inclusion of lower-intensity, higher-frequency storms significantly reduces bias relative to storm suites with a similar number of storms but only containing high-intensity, lower-frequency storms, even when estimating exceedances at very low-frequency return periods.

Keywords: flood risk; joint probability methods; storm selection; computational efficiency

1. Introduction

Flooding is the most frequently occurring weather-related natural disaster. In the United States, for example, 90 percent of all natural disasters involve flooding [1]. While not all flooding is caused by severe storms, coastal communities are particularly at risk of flooding from storm surge. Sea level rise and, in some areas, land subsidence will compound the problem, making many coastal communities significantly more vulnerable to storm surge hazards in future decades [2–4]. In addition, multiple studies have concluded that the average intensity of Atlantic hurricanes is likely to increase in the future due to projected increases in sea surface temperatures and other factors [5–8].

The Joint Probability Method with Optimal Sampling (JPM-OS) is one approach for estimating the probability distribution function of flooding in a region [9–13]. Storms are characterized using a set of parameters such as their minimum central pressure deficit, forward velocity, and location at landfall. A hydrodynamic model is used to simulate storms whose parameters span the range of plausible values that could occur in nature. A joint probability distribution function (PDF), fit to the historical record of observed storms, is used to estimate the relative likelihood of observing each storm in the suite of simulated, “synthetic” storms. The form of the joint PDF is based on assumptions about conditional relationships between the storm parameters and structural assumptions about how parameters are distributed; for example, the potential intensity of a hurricane is related to its size (radius of maximum windspeed) [14]. The probability masses associated with each synthetic storm are combined with their simulation results to build a cumulative distribution function (CDF) for quantitative metrics of flood risk like surge elevations, flood depths, or direct economic damage.

Long-range planning for flood risk management can involve evaluation of a large number of possible flood control or risk management strategies, such as different combinations of levee and floodwall projects, hazard mitigation, or broader coastal resiliency investments. Benefit-cost analysis entails predicting risk reduction benefits in multiple future time periods over a project's useful lifespan. Uncertainty about climate change, land subsidence, or other deeply uncertain drivers may necessitate that a range of future scenarios should be modeled. All of these factors multiply the number of scenarios, and thus model runs, needed to conduct a thorough and complete comparison of the benefits and costs of different flood risk reduction investments. Computational, budget, and time constraints can thus put downward pressure on the number of storms that can be run per case. We refer to the resulting problem of deciding what storms should be included in the analysis storm suite as the "storm selection" process.

This problem is exacerbated in diverse and complex wetlands ecosystems, such as those present in coastal Louisiana. In these areas, lower-resolution hydrodynamic models may produce biased results, and simplifying assumptions regarding future climate or sea level rise impacts may be invalid [15,16]. A high-resolution hydrodynamic simulation such as the ADvanced CIRCulation (ADCIRC) model is needed to accurately reproduce storm surge behavior from coastal storms [17].

This study has produced estimates of flood depth and damage at different exceedance probabilities associated with storm surge hazard in coastal Louisiana using the Coastal Louisiana Risk Assessment model (CLARA), a quantitative simulation model of storm surge flood risk developed by researchers at the RAND Corporation [18–20]. CLARA was developed to better understand how future coastal changes could lead to increased risk from storm surge flooding to residents and assets in Louisiana and assess the degree to which investments in risk mitigation could reduce this risk.

We have expanded upon the original storm selection approach described in Fischbach *et al.* [18] by examining flood depth and damage exceedances rather than surge elevation exceedances. The storm selection analysis was performed as part of a larger model development effort in support of the State of Louisiana's 2017 Coastal Master Plan. This paper draws largely from Fischbach *et al.* [20], a technical report describing the larger project; interested readers may refer to that document and Fischbach *et al.* [18] for additional details about CLARA.

Flood depth exceedance and expected annual damage estimates were produced using a reference set of 446 storms—the complete corpus of storms developed for use in recent JPM-OS studies of the Gulf coast—and with various subsets of the reference set, ranging from 40 to 304 storms. This paper examines the biases in flood depth and damage exceedances introduced by reducing the storm suite relative to a full implementation of JPM-OS using 446 storms. We also discuss the estimated parametric uncertainty associated with results from different storm suites. The results provide guidance for the types and size of storm suites that can be used to produce a good approximation of the results from the reference case when a full implementation of JPM-OS is computationally infeasible or prohibitively expensive for regions like coastal Louisiana.

2. Methods

The original storm selection problem was to choose a suite of storms to simulate that, when used to fit a surge response surface, produces a probability distribution of surge that closely replicates the true underlying distribution [9]. Prior analyses developed a suite of 304 idealized, synthetic storms that are defined parametrically by their central pressure c_p , radius of maximum windspeeds r_{max} , forward velocity v_f , longitudinal location at landfall x , and angle of incidence at landfall θ_l . For the Louisiana coastline, landfall was defined as the point when the eye of the storm crosses 29.5° N latitude [21,22]. These studies focused on representing storms that produce Category 3 or greater wind speeds on the Saffir-Simpson scale (minimum central pressure of 960 mb or lower); later work increased the number of available storms to 446 by adding intermediate storm tracks and less extreme storms with a minimum central pressure of 975 mb.

The 446 storms were run through a dynamically coupled set of storm surge and wave models: ADCIRC and Simulating WAVes Nearshore (SWAN), respectively [17,23,24]. These models were adapted and calibrated to current environmental and landscape conditions by other researchers working with the State of Louisiana [20]. The storms were also run through these models using coastal landscape and sea level conditions in the year 2065 as projected for the “Less Optimistic” scenario used in Louisiana’s 2012 Coastal Master Plan [16,25]. An initial screening analysis, described in the Experimental Design section, used results from the current conditions model outputs. The final, detailed analysis focused on the Less Optimistic future scenario, rather than current conditions. The higher level of risk (particularly in areas enclosed by ring levees) associated with the Less Optimistic future allows more information about bias to be drawn from a wider range of return periods, as there are fewer return periods where no flooding occurs.

Using this approach, we estimate flood depth and damage statistics at a large number of spatial grid points. Estimates of flood statistics derived from the full 446-storm suite—referred to as the reference set—are used as a reference standard for comparison. We ran CLARA’s modified JPM-OS procedure for a relatively large number of subsets of the reference set to investigate the use of smaller suites of synthetic storms. While we cannot assert that the 446-storm results are themselves unbiased or a true representation of the underlying storm risk, they represent the most complete set currently available for this region and are the best available basis for comparison when evaluating subsets.

In some cases, subsets were formed by eliminating storms from the reference set in ways that we hypothesized would introduce minimal bias. In other cases, subsets were formed by reducing the number of parameters that vary within the storm subset. For example, the storm tracks represented in the reference set are based on analysis of historic storm landfalls, in which a mean angle of incidence θ_l was calculated for each landfall point x . The reference set includes storms that follow the mean historic landfall angles, as well as storms that proceed along tracks 45 degrees to either side of the mean angle; some of the subsets we tested were formed by eliminating these “off-angle” tracks from the reference set. Further details on the subsets of storms considered are provided in the experimental design section of this manuscript.

For each subset of the reference set we proceeded as follows.

1. Fit response surfaces for peak storm surge and peak significant wave heights at each location not enclosed by a levee represented in the CLARA model (the “unenclosed grid points”) and at special points (“surge and wave points,” or SWP) 200 meters in front of levees and floodwalls along the boundary of enclosed protection systems [20]. Separate response surfaces are fit for surge hydrographs (surge elevations over time) and wave periods at the SWPs. The wave heights and periods are calculated at the time of peak surge in the surge hydrograph; the response surfaces for these quantities are fit using these values, and they are assumed to be constant over time (Estimated wave heights are also limited by the depth of the underlying surge. Holding wave heights and periods constant throughout the hydrograph reduces computational complexity and provides a first-order approximation of wave behavior during the points in the surge hydrograph where surge levels are high (*i.e.*, when overtopping may occur).).
2. Estimate peak surge and significant wave heights, using the response surface fits to obtain the predictions; also predict the surge hydrographs and wave periods at SWPs. In unenclosed areas, the effective flood elevations resulting from each synthetic storm are calculated as the sum of peak surge and the free-wave crest height. At SWPs, the predicted surge and wave characteristics are used as inputs to the flood module (as described in Step 4). The response surface fits are based upon the ADCIRC and SWAN outputs for each storm included in a given subset; these storms are sometimes referred to as the “training set” in the ensuing discussion. In some cases, the response surfaces can be used to predict surge and wave characteristics for a larger number of synthetic storms than those used to fit the response surface. For example, Set D3 includes storms on each track with a central pressure of 960 mb and r_{max} values of 11 and 35.6 nautical miles. The response surface is fit using these storms to estimate the effect of r_{max} on surge and waves. The

predictions also include a storm with an intermediate r_{max} of 21 nautical miles. Predictions are made for all storms in the 446-storm set that can be estimated using a response surface based on a given subset. For example, if a subset only includes storms with $v_f = 11$ knots, the response surface cannot identify the marginal effect of v_f on surge and waves, so synthetic storms with other values for v_f are excluded from the set of synthetic storms used to generate predictions. A general rule is that for any combination of track and angle represented in a training set, the set of predicted synthetic storms run through the rest of the steps listed below includes all storms from the 446-storm corpus corresponding to that track and angle, excluding storms with $v_f = 6$ or 17 knots if the training set does not also include storms with variation in the forward velocity.

3. Partition the JPM-OS parameter space by the synthetic storms used for predictions and estimate the probability mass associated with each synthetic storm. The probabilities assigned to each synthetic storm are derived using maximum likelihood methods on a data set of observed historic storms, under certain structural assumptions about the joint probability distribution of the storm parameters [26].
4. Run the predicted synthetic storms through CLARA's flood module using the predicted SWP hydrographs and wave heights to obtain final still-water flood depths on the interior of enclosed protection systems. The flood module includes analysis of overtopping from surge and waves, system fragility and the consequences of breaches, and routing of water between interior polders, also accounting for pumping capacity and rainfall [20,27]. The probability of system failure is modeled as a function of overtopping rates [28].
5. Combine the flood depths and probability masses associated with each predicted synthetic storm to build the distribution functions of flood depths at each CLARA grid point.
6. Extract the flood depth exceedance values from the cumulative distribution functions at various return periods (A flood depth value with a $(1 - 1/n)\%$ annual exceedance probability (AEP)—the chance of occurring or being exceeded in a given year—is referred to as having an n -year return period. For this study, we recorded 5-, 8-, 10-, 13-, 15-, 20-, 25-, 33-, 42-, 50-, 75-, 100-, 125-, 150-, 200-, 250-, 300-, 350-, 400-, 500-, 1000-, and 2000-year exceedances.).
7. Run the flood depth exceedances through CLARA's economic module to estimate economic damage exceedances. Also estimate the expected annual damage (EAD) associated with each storm subset.

Full details on how each of these steps are accomplished (for example, the specification of the response surface model) can be found in Fischbach *et al.* [20]. We have found that deriving flood depth exceedance curves from small numbers of storms can be problematic in enclosed areas behind levee systems, with large, sudden jumps in exceedance values; for example, one might estimate a 12-foot difference in flood depths between the 100-year and 125-year exceedances [29]. If one uses a small number of storms, leaving large gaps in the parameter space, each storm may produce very different results from its "nearest neighbor" storms. This is exacerbated by the nonlinear relationship between surge and wave heights and levee overtopping, and also by the impact of system fragility. As a result, using the response surface to generate predicted synthetic storms with intermediate parameter values not represented in the training set is necessary for "smoothing out" the exceedance curve and making it less sensitive to small changes in the probability weights assigned to each storm. For consistency, when running synthetic storms through the CLARA flood module, we use the response surface predictions for all synthetic storms tested as part of a particular test set (as opposed to using the actual hydrodynamic modeling outputs for storms, when available, and response surface predictions only for storms outside of the training set).

The performance of each storm subset was evaluated by summarizing the bias in predicted flood depths at various return periods, relative to the results generated by the reference set. We also analyzed and compared the estimated standard errors associated with each subset.

3. Experimental Section

We first conducted a screening-level analysis by evaluating a large number of storm suites—sixteen in total, inclusive of the 446-storm reference set. For the initial screening, flood depths were calculated at a total of 18,273 grid points across the coast. These grid points were chosen by selecting a limited sample of 24 watersheds that varied by (a) size, (b) proximity to the coast, (c) degree of economic development, and (d) whether the watershed is protected by some federally-accredited levee or floodwall system. (The full CLARA v2.0 study region for coastal Louisiana consists of 77,643 grid points comprising 117 watersheds (excluding points consisting of open water). Grid point resolution varies with the concentration of population and assets, with a maximum distance of 1 km between neighboring points.) Watersheds protected by a fully-enclosed ring levee system were excluded from the initial screening sample to avoid running CLARA's flood module (step 4 above) as part of the screening exercise. Damage results were also excluded from the screening analysis.

The ten JPM-OS tracks used in the 2012 Coastal Master Plan analysis are sometimes referred to as primary tracks. They were labeled E1 through E5 and W1 through W5 for tracks in the eastern and western halves of the coast, respectively. Secondary storm tracks correspond to paths in between the primary tracks and were denoted by a B at the end of the track name (e.g., track E1B). Tracks also vary by their angle of incidence made with the coastline upon landfall. Tracks following the mean landfall angle are referred to as central-angle tracks; those making landfall at angles 45 degrees less or greater than the mean angle are referred to as off-angle tracks.

The modeled subsets were chosen to be collections of storms with easily interpretable and describable characteristics, to have variation in the total number of storms, and to have variation in the types of storms represented over the subsets. They were also chosen to avoid groups of storms that could cause identifiability or other performance issues in the response surface model. Louisiana's Coastal Protection and Restoration Authority (CPRA) also identified a need for storm sets with fewer than 154 storms—and preferably fewer than 100 storms—that could be used to evaluate a range of individual structural protection projects during 2017 Coastal Master Plan model production using available computing resources. Table 1 describes what types of storms comprise each of the storm suites evaluated in the screening-level analysis. The number of storms listed in the table represents the number of storms in the training set used to fit the response surface in Step 1 of the procedure described in the Methods section; as noted in the description of Step 2, the set of predicted synthetic storms run through the rest of the procedure may be larger than the training set. However, the size of the training set is the relevant number to discuss when comparing storm suites, as it represents the number of storms that must be run through the more computationally intensive ADCIRC and SWAN models.

Table 1. Characteristics of storm suites selected for screening-level analysis.

Set	Storms	Description
S1	446	Reference set
S2	40	2012 Master Plan (MP) storm suite: 10 primary storm tracks, 4 storms per track varying c_p and r_{max}
S3	90	
S4	304	LACPR storm suite: all storms in reference set, except those with 975 mb c_p
S5	194	All storms from LACPR storm set with 11-knot v_f
S6	154	All storms from LACPR storm set with 11-knot v_f on primary storm tracks
S7	216	All storms from reference set with 11-knot v_f on primary storm tracks
S8	144	All storms from reference set on primary storm tracks with 900 mb or 930 mb c_p
S9	120	All LACPR storms on primary, central-angle tracks (includes variation in v_f)
S10	264	All storms from reference set with 11-knot v_f
S11	50	2012 MP storm suite expanded to 5 storms per track varying c_p and r_{max}
S12	60	Set 11, plus storms with 975 mb c_p and central values for r_{max}
S13	120	All central-angle, primary-track storms with 11-knot v_f (includes 975 mb storms)
S14	148	Set 13, plus all 960 mb and 975 mb storms on secondary storm tracks
S15	154	Set 3, plus all 960 mb and 975 mb storms on primary, off-angle storm tracks
S16	182	Set 3, plus all 960 mb and 975 mb storms on secondary tracks or primary, off-angle tracks

Notes: c_p —central pressure; r_{max} —radius of maximum windspeed; v_f —forward velocity; LACPR—Louisiana Coastal Protection and Restoration study [22].

The initial findings led to the development of several new test suites, ranging from 60 to 154 storms, for the detailed analysis. The final sets tested, including number of storms and a description of key characteristics, are shown in Table 2 below. Note that the storm set ID numbers for the detailed analysis do not all match up with the IDs used in the screening analysis; although some sets from the screening analysis were included in the detailed analysis, the ID numbers were designated independently. The detailed analysis included estimation of flood depth exceedances at all grid points across the coast, including in areas enclosed by ring levees. Damage exceedances and bias in EAD were also analyzed for these storm suites.

Table 2. Characteristics of storm suites selected for detailed investigation.

Set	Storms	Description
D1	446	Reference set
D2	40	2012 Master Plan (MP) storm suite: 10 primary storm tracks, 4 storms per track varying c_p and r_{max}
D3	60	2012 MP storm suite, expanded to 5 storms per track varying c_p and r_{max} , plus storms with 975 mb central pressure and central values for r_{max}
D4	90	2012 MP storm suite, expanded to 9 storms per track varying c_p and r_{max}
D5	90	2012 MP storm suite, expanded to 7 storms per track (excludes one 930 mb and one 900 mb storm), plus 975 mb storms using extremal (rather than central) r_{max} values
D6	92	Set 3, plus 960 mb and 975 mb storms on off-angle tracks only in E1-E4
D7	92	Set 3, plus 960 mb and 975 mb storms on off-angle tracks only in W3-W4, E1-E2
D8	100	All central-angle, primary-track storms with 11-knot v_f , plus 975 mb storms with central r_{max} values
D9	110	All central-angle, primary-track storms with 11-knot v_f , plus 975 mb storms with extremal r_{max} values
D10	120	All central-angle, primary track storms with 11-knot v_f , including all 975 mb storms
D11	154	Set 4, plus all 960 mb and 975 mb storms on primary, off-angle storm tracks

Notes: c_p —central pressure; r_{max} —radius of maximum windspeed; v_f —forward velocity.

The off-angle tracks for sets D6 and D7 were selected based on their paths passing near to New Orleans, and thus having the potential for significant flooding in urban centers where a large number of grid points and economic assets are densely packed. Our hypothesis was that including off-angle storms affecting the greater New Orleans region would improve accuracy in estimating how likely it is for the city to experience the catastrophic flooding associated with levee failures.

4. Results and Discussion

An example result from the screening analysis is shown in Figure 1. The depicted storm suite (Set S3 from Table 1) is comprised of nine storms from each of the ten primary tracks. Within a track, storms vary by c_p and r_{max} . The central pressure values represented are 900, 930, and 960 mb; storms with 975 mb pressures were not included. All storms have the central value for v_f of 11 knots. The figure illustrates bias in the estimated 100-year flood depths, relative to the reference set. Note that this particular suite tends to strongly overpredict 100-year depths (blue coloring), relative to the 446-storm reference set, in the far eastern portion of the state, east of New Orleans and the Mississippi River, while it tends to underpredict depths (orange coloring) in the watersheds farthest west.

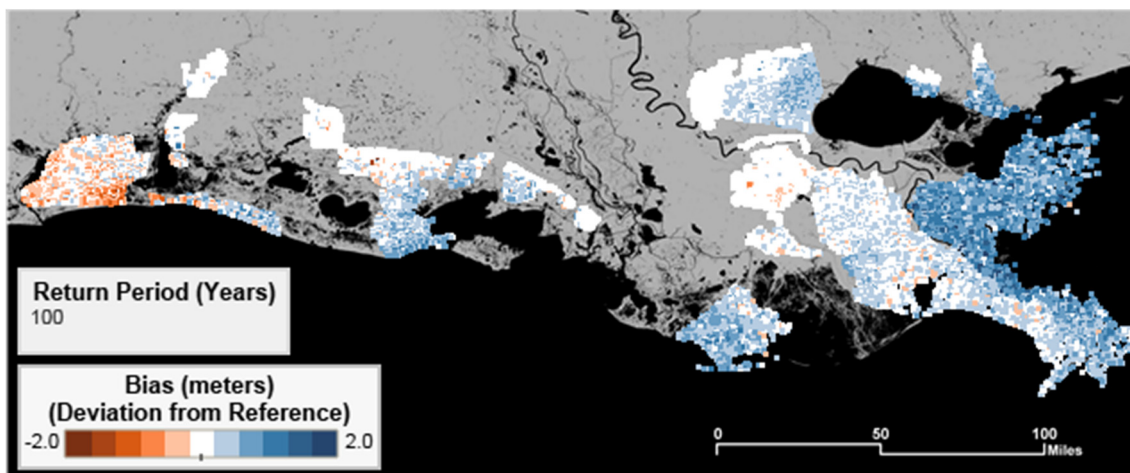


Figure 1. Deviation from reference set in 100-year flood depth exceedances in screening analysis watersheds from a suite consisting of 90 storms (screening set S3).

Figure 2 summarizes the findings of the screening-level analysis. It shows the root mean squared error (relative to the reference set), averaged over all grid points in the screening watersheds, for each modeled storm suite; the figure plots this average bias as a function of the number of storms in each suite. Blue points represent storm suites which contain at least one storm with a 975 mb minimum central pressure, the dot shape represent suites which contain off-angle storm tracks. Sets 8 and 9 (as well as the reference set) from the screening analysis also contain storms with non-central values for forward velocity.

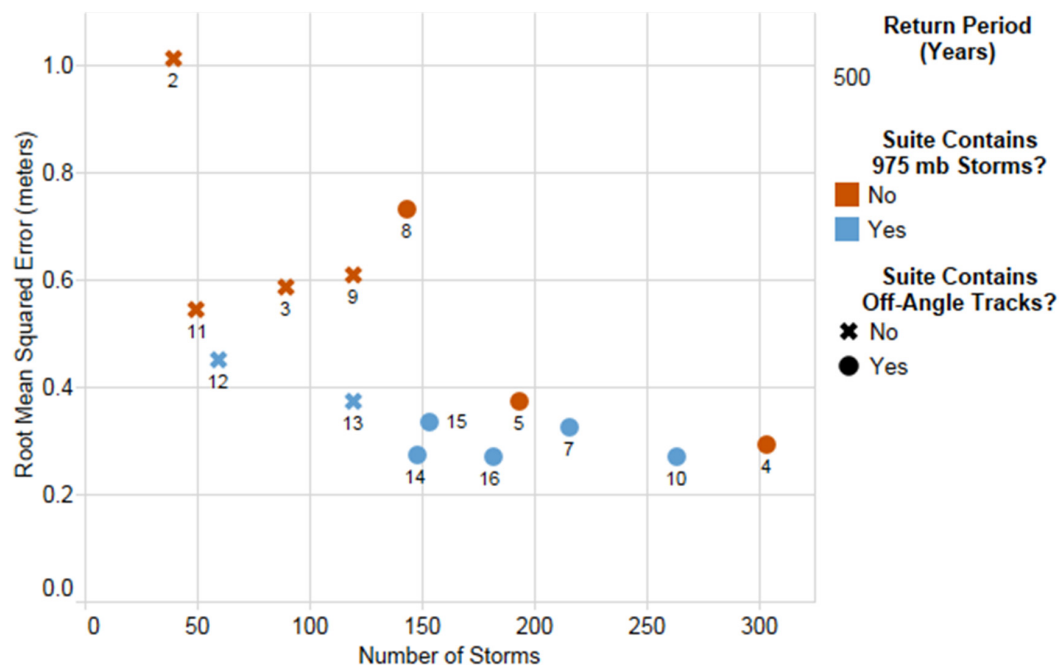


Figure 2. Average coastwide bias and variation by number of storms (screening analysis), 500-year flood depths.

In addition to the dimensions illustrated in the figure, sets S4, S8, and S9 contain storms with non-central values for forward velocity. The screening results suggest that including higher-frequency storms with 975 mb central pressure improved statistical performance. By contrast, secondary or

off-angle storm tracks and storms with non-central values for forward velocity did not yield similar improvement and were generally not included in the more detailed final testing and results.

4.1. Flood Depth Bias and Variance Comparisons

Figure 3 summarizes the average coastwide 100-year flood depth bias (root mean squared error, y-axis) and coefficient of variation (point size) for each set in the detailed analysis, plotted against the number of storms in a given storm suite (x-axis). As before, colors indicate whether the set includes 975 mb storms, and shape indicates whether off-angle tracks are included. Bias is estimated relative to the flood depth results from the reference set.

Summary results show that average bias in flood depths at the 100-year interval varies from less than 0.25 m to nearly 1 m, depending on the storm sample. Substantial bias is observed for Set D2, the 2012 Coastal Master Plan suite, compared with the other candidate sets tested. All additional sets with more than 40 storms improve upon the 40-storm results. In fact, increasing the number of storms to 60 (Set D3) leads to substantial improvement, reducing average bias by more than a half a meter. Similar results are observed at other return periods.

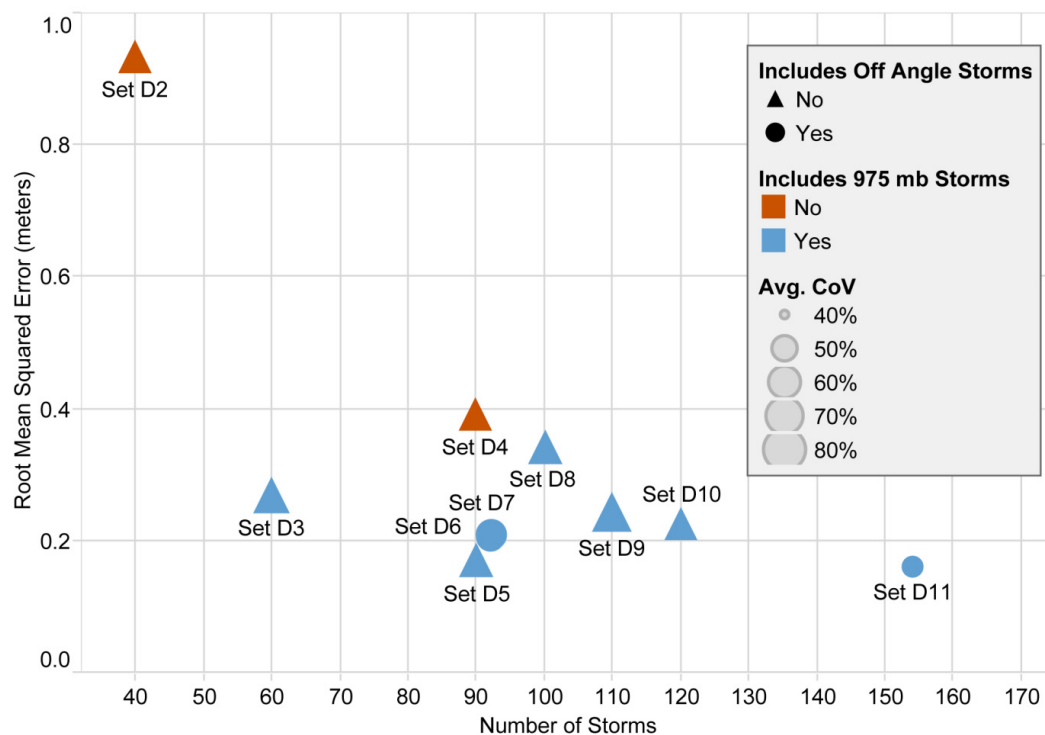


Figure 3. Average coastwide bias and variation by number of storms (detailed analysis), 100-year flood depths.

We focused primarily on suites of storms that include 975 mb storms because the screening analysis results suggested that such suites outperform suites that exclude the less intense storms. The relatively poor performance of sets D2 and D4, the suites with no 975 mb storms, suggests that this is still a valid conclusion when considering the full coastal study region. Improvement is less apparent when off-angle tracks are included, however, especially for sets in which only certain off-angle tracks were included (e.g., Sets D6 and D7).

Figure 4 illustrates how the average bias, again measured by RMSE, changes for each storm suite at different return periods of the flood depth distribution. The x-axis is the return period, on a logarithmic scale. The thickness of the lines shown in the figure reflects the coefficients of variation, a measure of the uncertainty around derived point estimates; thicker lines denote greater uncertainty.

With few exceptions, the relative order of performance across storm subsets is consistent over a wide range of return periods. Sets D2 and D4 are conspicuous in their poor average performance; neither set contains any storms with central pressures of 975 mb, further emphasizing the importance of including higher-frequency events in the training set. Interestingly, this effect is still particularly apparent in the tail of the distribution beyond the 100-year AEP interval, indicating that exclusion of higher-frequency events from the response surface model also skews the model's predictive accuracy for more extreme synthetic storms.

In general, average bias is smaller at 50-year and more frequent return periods. This is due to the large number of points in which no flooding occurs for these exceedances. The absolute error is more likely to be small when the depth estimates themselves are small. Set D11, with 154 storms, performs well across most of the distribution, and better than any smaller sets. However, several of the sets with fewer than 100 storms are not far behind in terms of average bias across the exceedance distribution.

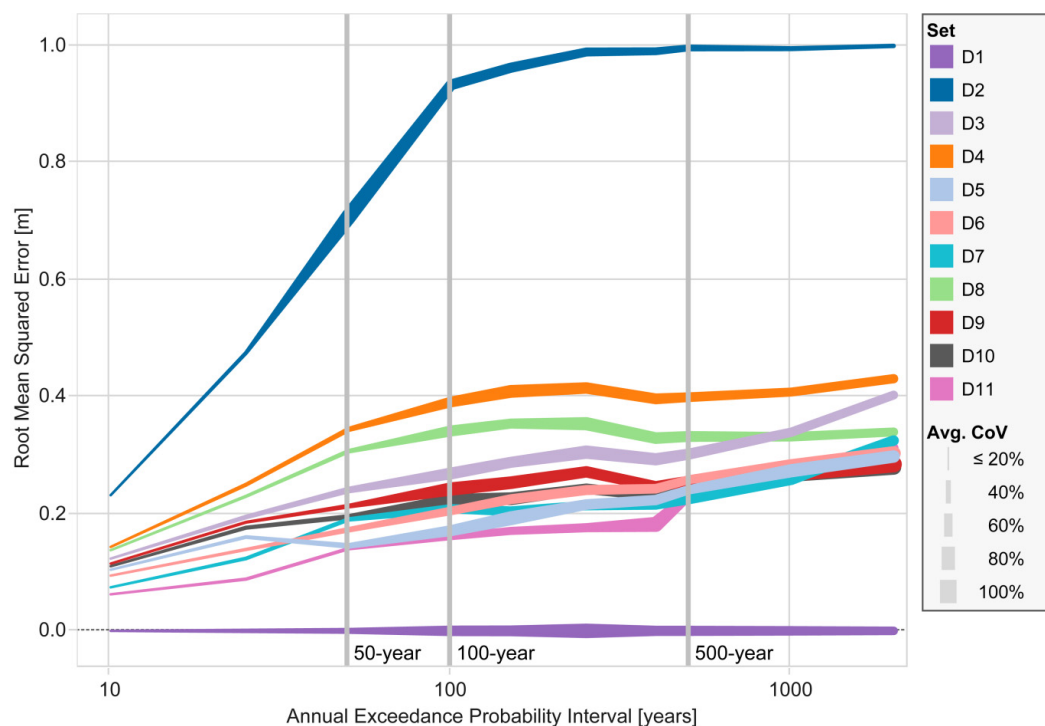


Figure 4. Average bias and variance by exceedance interval for each modeled storm suite.

Figure 5 shows a sequence of maps that display the spatial patterns of bias, relative to the reference set, for three of the storm suites tested. Flood depths at the 100-year return period, as estimated using the reference set, are shown for comparison in Figure 6; when examining the three maps in Figure 5, this gives some indication of the magnitude of the biases relative to the baseline flood depths. When examining the spatial distribution of bias, some interesting patterns emerge. The Set D2 results, representing the 2012 Coastal Master Plan's 40-storm sample, show a substantial overestimate of flood depths across nearly the entire study region. This effect occurs because of the types of storms excluded from Set D2; it only contains extreme central-angle storms on primary tracks, leaving out both 975 mb and off-angle storms. For instance, Set D2 excludes off-angle storms that pass well to the east of New Orleans. These storms would tend to lower the estimated flood depth exceedances in St. Bernard Parish, so as a result Set D2 has positive bias in that region.

Set D3, which includes 60 storms, improves dramatically on the Set D2 results. Some positive bias is still noted in the western portion of the coast and many areas east of the Mississippi River, but the magnitude is notably lower than Set D2. In addition, Set D3 actually underestimates 100-year flood

depths compared with the reference in some enclosed areas, including the East Bank of the Greater New Orleans Hurricane Storm Damage Risk Reduction System (HSDRRS) and the Larose to Golden Meadow levee system.

Set D11 yields the lowest overall bias when averaged over all points, and it shows balanced results coastwide when looking at geospatial patterns. Positive bias is still observed in the western parishes, but again with a lower magnitude than Sets D2 or D3. There are some instances where the 100-year flood depth estimates are both positively and negatively biased within the same watershed. When these differences in flood depths are translated into damage, the coastal and parish-level results are more similar to those of the reference set.

The differences between enclosed and unenclosed areas are clearly shown in Figures 7 and 8 (note the different scale on the y-axis between figures). The curves shown in Figure 4 are broken out by showing the same results for points enclosed by (Figure 7) and unenclosed points (Figure 8), respectively.

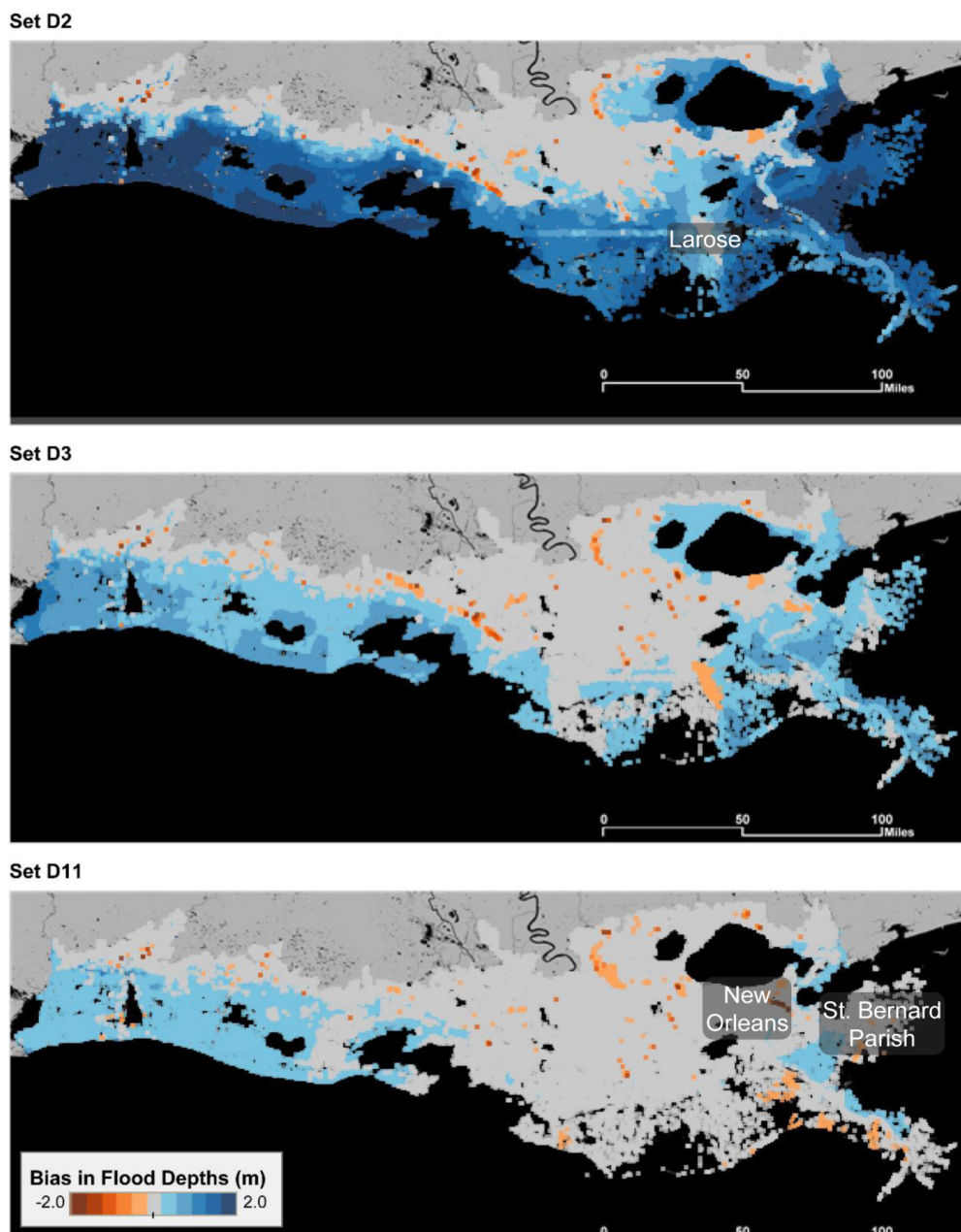


Figure 5. Maps of bias by grid point for detailed analysis sets D2, D3, and D11, 100-year flood depths.

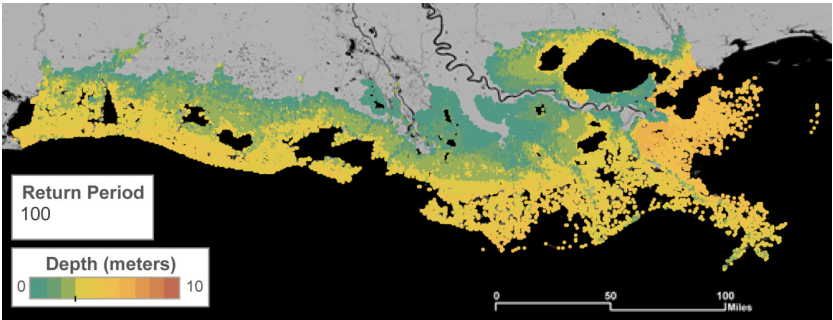


Figure 6. Median 100-year flood depths from reference set D1.

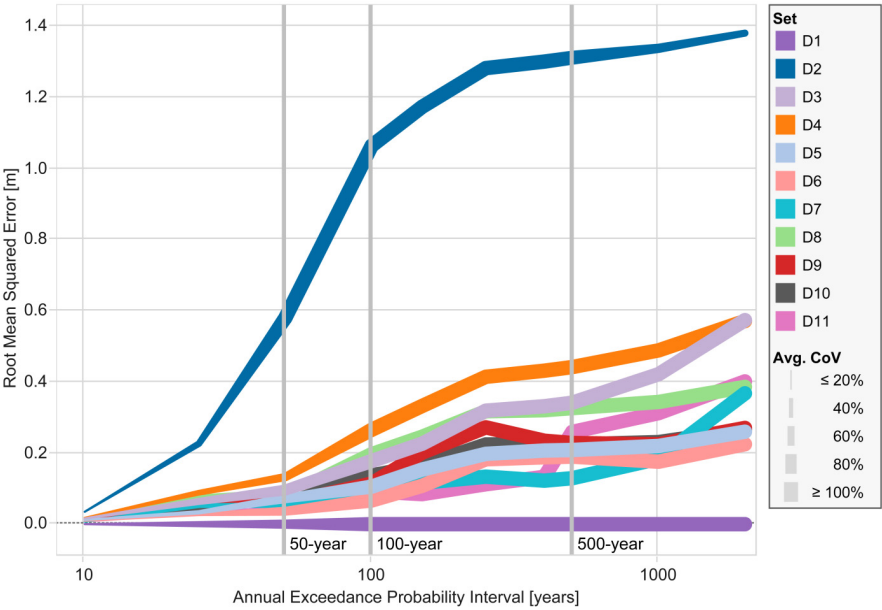


Figure 7. Average bias and variance by exceedance interval (HSDRRS Only).

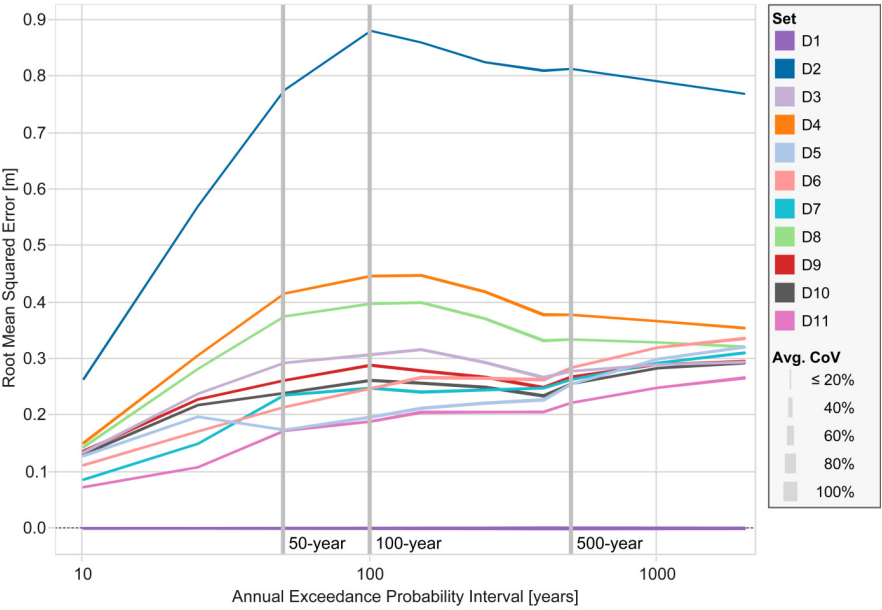


Figure 8. Average bias and variance by exceedance interval (unenclosed points only).

HSDRRS was accredited in 2014 by the United States Federal Emergency Management Agency as protecting the Greater New Orleans area against at least 100-year surge levels [30]. However, the same standard may not be met in 2065 under the Less Optimistic future scenario assumptions in a future without further mitigating actions. The precise level of protection provided by HSDRRS is uncertain [31]. As such, the average coefficient of variation is relatively large, over 100% for many storm suites over large portions of the distribution.

In unenclosed areas, the performance of all suites but Set D2 is relatively good and largely similar. With the exception of Set D2, all storm suites have an average bias of less than half a meter at all exceedance intervals, and the average uncertainty in the predicted values is also considerably smaller than for points within HSDRRS. This reflects that the exceedance curves in unenclosed areas have lower variance and are more stable, in the sense that few points experience sudden, sharp increases in flood depths at any point of the distribution.

The greater variation in performance in enclosed areas illustrates that flood depth estimates are more sensitive to the choice of storm suite when the accuracy and uncertainty in response surface predictions is further compounded by interactions with engineered flood protection systems. Maintaining accuracy with fewer storms is easier for prediction of storm surge elevation exceedances, as in the original storm selection problem faced by JPM-OS, than for flood depth exceedances when accounting for the levee and floodwall failures, overtopping, and interior drainage dynamics.

Figure 9 focuses in on differences in storm suite performance for enclosed points within HSDRRS, on the east and west sides of the Mississippi River, at the 500-year AEP interval. Consistent with what was shown in Figures 7 and 8 results are most uncertain and show the greatest variance in the East and West Bank areas. Except for Set D2, all sets considered yield an average bias of approximately 0.3 meters (1 foot) or less in enclosed locations on the East Bank. Uncertainty is greater on the West Bank. Similarly, Sets D3 through D11 also result in less than 1 meter of average bias in enclosed areas not within HSDRRS, with many sets also yielding less than 0.3 meters of average bias in these areas (results not shown).

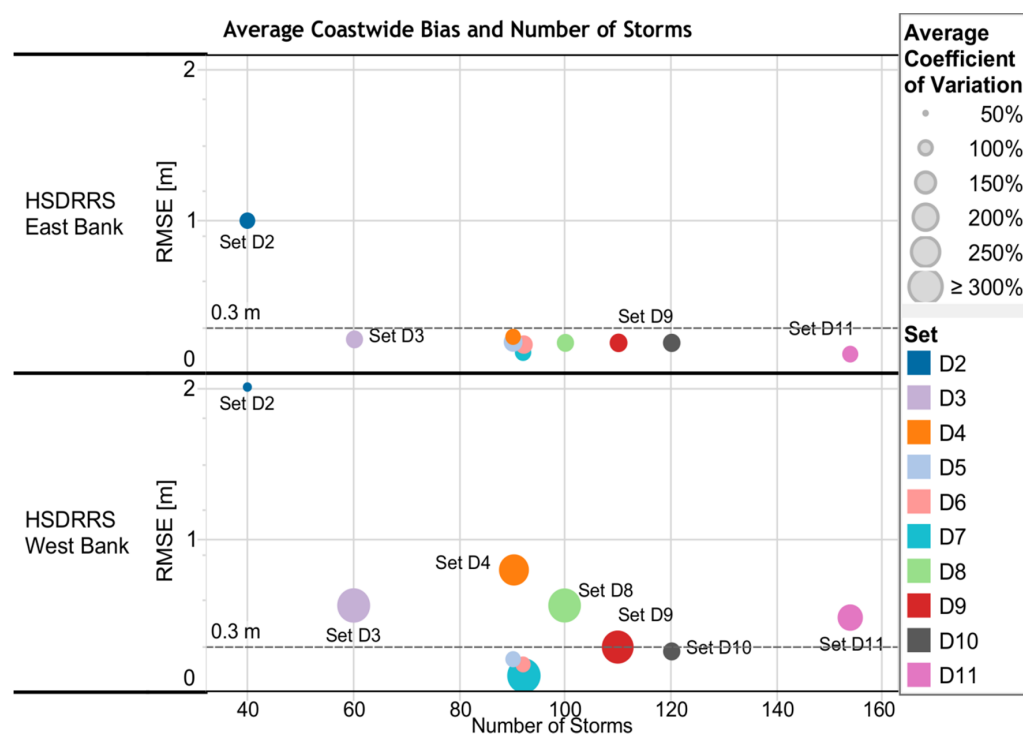


Figure 9. Average bias and variation, 500-year flood depths, east and west banks of the Greater New Orleans Hurricane Storm Damage and Risk Reduction System.

Figure 9 confirms that Set D11 is the best or near-best performer in terms of mean bias across all enclosed areas. For sets with fewer than 100 storms, performance depends on HSDRRS location. Set D3 performs substantially worse, however, on the West Bank, as do many other sets. By contrast, Set D7 was specifically designed to improve performance in the West Bank area of HSDRRS. This suite adds storms to Set D3, including off-angle tracks for the middle of the coast with landfall locations observed to have the greatest effect on this portion of HSDRRS. The results bear this out: Set D7 is the best overall set in terms of bias in the West Bank, even slightly better than the larger Set D11. However, Set D7 produces greater bias in the East Bank HSDRRS, and is among the worst performers when looking at enclosed areas other than the West Bank. It also has a significantly larger coefficient of variation on the West Bank than Sets D5 and D6, which both have nearly identical average RMSE and number of storms to Set D7. This illustrates that while it may be possible to carefully tailor a storm suite to perform well in a particular region, results may be significantly worse in other, even nearby, areas. It can be difficult to predict the overall performance of a storm set, which emphasizes the importance of conducting a storm selection analysis like this on at least one landscape scenario before embarking on a major flood risk assessment encompassing many future scenarios, time periods, and system configurations.

4.2. Damage Bias Comparisons

Next, we sought to understand how the potential bias from different storm suites translates to bias in damage estimates. Somewhat surprisingly, we found that the total bias associated with a suite of storms was typically driven by bias in unenclosed areas, although such regions are more sparsely populated and contain fewer assets than in protected areas like New Orleans. The greater protection in enclosed areas provided by federal levee systems instead meant that there was a lower projected baseline risk in those regions; most storm suites did not predict extensive flooding in enclosed areas except at lower-frequency AEPs, so this resulted in less bias.

Figure 10 below shows a similar spatial breakdown as Figure 9, for example, but instead displays bias in terms of damage (EAD, median results) estimated by the CLARA v2.0 economic damage model. Set D2, which yields very high EAD bias overall, is omitted from the figures below for clarity. Figure 10 confirms the performance noted above, with roughly the same overall ranking of storm suites by EAD bias as with flood depths. Sets D5, D6, D7, and D11 all have a cumulative bias of less than \$0.5 billion across all points. D3 shows a total bias over all points of about \$1.5 billion; D8 has bias of \$2 billion, while D4 has \$2.5 billion.

Increasing the number of storms from 60 up to 90–110 does not substantially improve EAD performance for enclosed areas; Sets D4 and D8 consistently produce worse error than Set D3, despite having 90 and 100 storms, respectively. Instead, it is the addition of 975-mb storms that appears to drive improvement. Set D4 has no such storms, while Set D8 contains half as many 975-mb storms than the other suites with 90–110 storms.

Similarly, Figure 11 shows a summary by coastal Louisiana parish (county) of bias in EAD compared to the reference set. Some storm suites, regardless of coastwide performance, produce large bias in specific parishes. For example, Set D4 produces approximately 50% more damage in Jefferson and St. Bernard Parishes than the 446-storm reference set. In turn, this leads to a substantial upwards bias in coastwide EAD because these parishes contain large concentrations of assets.

Sets D5, D6 and D11 are the highest performers when disaggregating results over the parishes; they are the only storm sets with only one parish's bias over 20% and a maximum bias of magnitude 25% or less. The maximum bias for any parish from Set D11 is 22% in Calcasieu Parish, which is in the far west of the state and includes relatively few assets. All storm suites overestimated EAD in Calcasieu, with errors ranging from 14 to 41%, indicating a general difficulty in reproducing estimates of risk so far west with a small number of storms. Each of the other storm suites produces biases as large or larger in several parishes. Few storm sets performed well, with errors of 10% or less, in both Orleans and Jefferson parishes, the two parishes with the largest value of assets at risk. EAD

performance in non-HSDRRS parishes, by contrast, is typically more similar across all sets tested in this round of analysis.

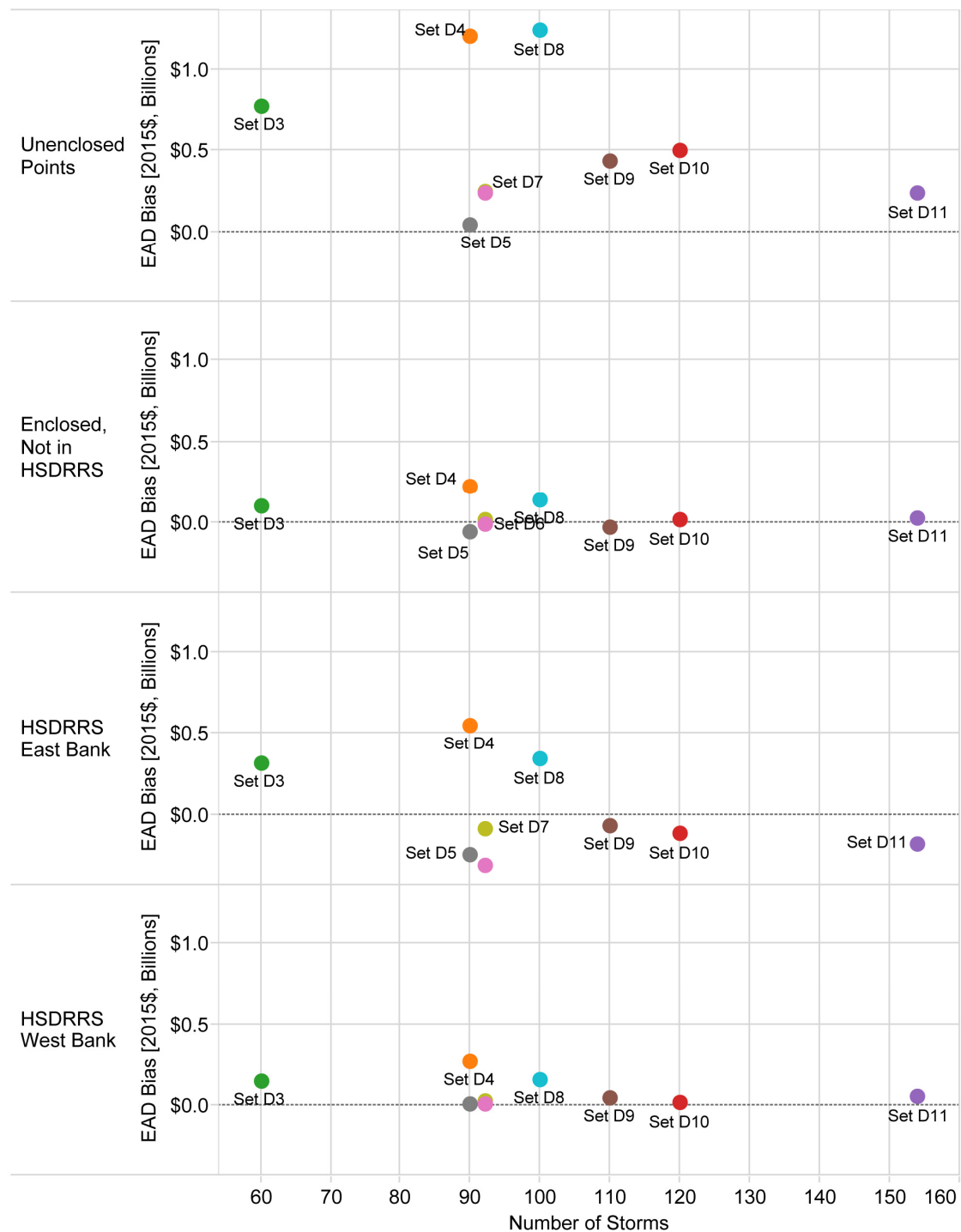


Figure 10. Bias in expected annual damage (median estimates), by location.

The results discussed so far have portrayed the median (50th percentile) outputs over the sampling design, with the coefficients of variation giving some information about the parametric variation across different suites. Figure 12 illustrates the distribution of EAD in another way. The y-axis indicates the bias in coastwide EAD relative to the full 446-storm reference set; the three points for each subset, from

bottom to top, represent the bias associated with the 10th, 50th, and 90th percentile values, respectively, of EAD.

Parish Name	Set								EAD Bias (percent)
	D3	D4	D5	D6	D7	D9	D10	D11	
Acadia	2%	3%	-2%	0%	8%	-4%	-1%	9%	
Ascension	27%	34%	14%	19%	8%	20%	21%	10%	
Assumption	2%	3%	0%	3%	0%	1%	1%	-3%	
Calcasieu	24%	24%	14%	23%	41%	17%	19%	22%	
Cameron	0%	-3%	2%	-2%	2%	-1%	0%	-1%	
Iberia	0%	2%	-5%	0%	5%	-3%	-2%	6%	
Jefferson	32%	44%	-4%	-17%	-4%	2%	-1%	-5%	
Jefferson Davis	9%	13%	1%	9%	18%	8%	8%	15%	
Lafourche	5%	11%	0%	2%	2%	1%	4%	2%	
Livingston	15%	17%	13%	10%	5%	14%	14%	3%	
Orleans	4%	22%	-21%	-15%	0%	-8%	-13%	-7%	
Plaquemines	12%	22%	2%	1%	2%	8%	7%	4%	
St. Bernard	27%	57%	7%	-8%	0%	15%	9%	7%	
St. Charles	5%	8%	2%	3%	0%	5%	5%	1%	
St. James	-4%	9%	-12%	-5%	-11%	-12%	-12%	-16%	
St. John the Baptist	24%	30%	13%	13%	7%	21%	21%	13%	
St. Martin	1%	0%	0%	1%	0%	0%	0%	-1%	
St. Mary	0%	3%	-10%	-2%	-3%	-5%	-6%	0%	
St. Tammany	12%	14%	9%	7%	7%	10%	10%	8%	
Tangipahoa	0%	0%	-1%	-2%	-7%	-1%	1%	-7%	
Terrebonne	4%	8%	-4%	2%	-2%	1%	1%	0%	
Vermilion	5%	7%	3%	6%	7%	3%	4%	7%	

Figure 11. Percentage bias in expected annual damage by parish (median).

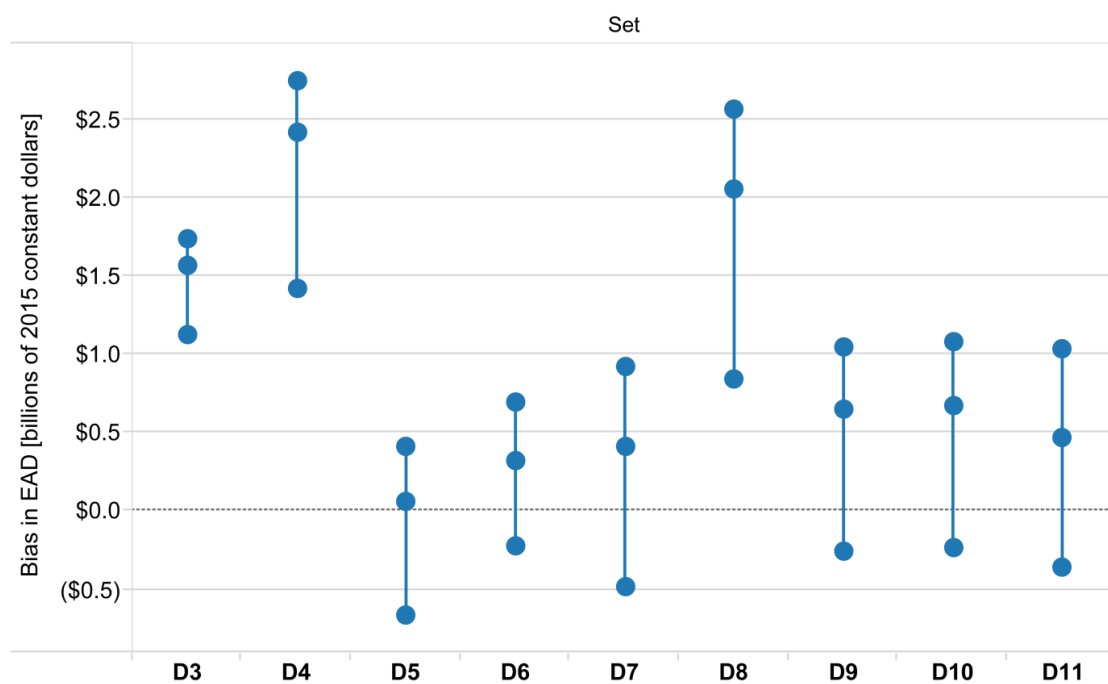


Figure 12. Coastwide bias in terms of expected annual damage (billions of 2010 US dollars).

The primary takeaway here is that many storm suites end up with similar ranges of uncertainty in terms of EAD bias relative to the reference set, when aggregating damage coastwide. Considering only suites with fewer than 100 storms, Set D5 is the best performer at the median, but includes a wider range of results across the parametric distribution than Set D6, and the 10th percentile estimate is that Set D5 may underestimate EAD more strongly than any other set. Set D3 has the smallest range of variation in bias between the 10th and 90th percentile estimates, but it has a larger median bias than most of the larger sets; Set D6 has the second smallest range of uncertainty. Six sets underestimate EAD at the 10th percentile but overestimate it at the median and 90th percentiles.

5. Conclusions

The storm selection analysis shows a tradeoff between the number of storms and the resulting bias when compared with the reference set of 446 storms. Results show that nearly all storm sets tested produce lower bias when compared with the 2012 Coastal Master Plan 40-storm suite (Set D2). Aside from the fact that Set D2 was the smallest storm suite tested, this should not be surprising, given that the 2012 Coastal Master Plan was intended to model damage associated only with Category 3 or greater hurricanes. Substantial improvement is noted when storms with 975 mb central pressure were included, as well as with the addition of off-angle storms in some cases. The inclusion of storms with variation in forward velocity was found to have little impact on performance.

Of the storm suites tested, Set D6 (92 storms) appears to yield the best balance of results. It shows relatively low bias compared with the reference set in terms of both flood depth and damage, no concerning spatial patterns of bias, and reasonable performance in enclosed areas (particularly Greater New Orleans). It also has the second smallest range of uncertainty in damage among the suites analyzed. Set D3 is notable for producing less average error in flood depths than expected given that it contains only 60 storms.

The results of this storm selection analysis suggest some insights for future planning studies using a JPM-OS approach with high-resolution storm surge and wave modeling. First, the number of storms needed to estimate unbiased flood depths or damages can vary greatly depending on the size of the geographic area of focus, landscape type and configuration, presence of hurricane protection structures, and range of exceedance probabilities targeted. In some cases—for instance, in areas without protection structures or with relatively similar landscape characteristics over a broader area—a smaller storm suite may yield acceptably low bias, with results suitable for screening or planning-level decisions.

In some areas, however, large numbers of storms are needed to avoid bias and assess a wide range of plausible storm impacts from storms with different characteristics. Larger storm suites are also needed to support flood protection engineering and design studies, where decision makers need confidence that the widest range of plausible outcomes are simulated and adequate factors of safety have been developed accordingly.

For the 2017 Coastal Master Plan, a hybrid approach will be adopted: a smaller suite of 60 storms (Set D3) will be used for preliminary evaluation and testing of individual risk reduction projects, but coastwide alternatives—including hurricane protection structures, ecosystem restoration, and nonstructural risk reduction (hazard mitigation for buildings)—will be evaluated and compared using a larger set of 92 storms (Set D6). This two-step approach will help to confirm or validate preliminary project-level results and ensure that potential risk and damage reduction benefits from the plan as a whole are relatively unbiased.

Acknowledgments: Other RAND researchers and programming staff who have contributed to development of the CLARA model include Ricardo Sanchez, Chuck Stelzner, Rachel Costello, Edmundo Molina-Perez, and James Syme. This work was funded by the Louisiana Coastal Protection and Restoration Authority under the 2017 Coastal Master Plan's Master Services Agreement. Coordination of the process has been managed by The Water Institute of the Gulf. In particular, the authors wish to thank collaborators Hugh Roberts and Zach Cobell at ARCADIS; Mandy Green, Melanie Saucier, Mark Leadon, and Karim Belhadjali at CPRA; Denise Reed at The Water Institute; and Gary Cecchine at RAND Corporation for their intimate involvement in and support of this project.

Author Contributions: Fischbach is serving as principal investigator for the 2017 Coastal Master Plan's Risk Assessment team. He led the development of the storm selection analysis methodology and results visualizations, with Johnson's assistance. Johnson is the lead technical developer of the CLARA model; he executed the model runs to support the storm selection analysis and developed the storm suites included in the detailed analysis based on the initial screening results. Kuhn assisted in carrying out the analysis and quality assurance process. All authors contributed to the technical report on which this journal article is based; Johnson led its adaptation for journal submission, with assistance from the other authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Federal Emergency Management Agency. Fact sheet: Flooding—Our nation's most frequent and costly natural disaster. Federal Emergency Management Agency: Washington, DC, USA, 2010; p. 2.
2. Karl, T.R.; Meehl, G.A.; Peterson, T.C.; Kunkel, K.E.; Gutowski, W.J., Jr.; Easterling, D.R. *Weather and Climate Extremes in a Changing Climate, Regions of Focus: North America, Hawaii, Caribbean, and U.S. Pacific Islands*; NOAA: Silver Spring, MD, USA, 2008.
3. Burkett, V.R.; Davidson, M.A. *Coastal Impacts, Adaptation and Vulnerability: A Technical Input to the 2012 National Climate Assessment*; Burkett, V.R., Davidson, M.A., Eds.; Island Press: Washington, DC, USA, 2012.
4. Neumann, J.E.; Emanuel, K.; Ravela, S.; Ludwig, L.; Kirshen, P.; Bosma, K.; Martinich, J. Joint effects of storm surge and sea-level rise on US coasts: New economic estimates of impacts, adaptation, and benefits of mitigation policy. *Clim. Change* **2014**, 1–13. [[CrossRef](#)]
5. Knutson, T.R.; McBride, J.L.; Chan, J.; Emanuel, K.; Holland, G.; Landsea, C.W.; Held, I.M.; Kossin, J.P.; Srivastava, A.K.; Sugi, M. Tropical cyclones and climate change. *Nat. Geosci.* **2010**, 3, 157–163. [[CrossRef](#)]
6. Emanuel, K. Downscaling CMIP5 climate models shows increased tropical cyclone activity over the 21st century. *Proc. Natl. Acad. Sci. USA* **2013**, 100, 12219–12224. [[CrossRef](#)] [[PubMed](#)]
7. Villarini, G.; Vecchi, G.A. Twenty-first century projections of north atlantic tropical storms from CMIP5 models. *Nat. Clim. Change* **2012**, 604–607. [[CrossRef](#)]
8. Webster, P.J.; Holland, G.J.; Curry, J.A.; Chang, H.R. Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science* **2005**, 309, 1844–1846. [[CrossRef](#)] [[PubMed](#)]
9. Resio, D.T.; Irish, J.L.; Cialone, M.A. A surge response function approach to coastal hazard assessment. Part 1: Basic concepts. *Nat. Hazards* **2009**, 51, 163–182. [[CrossRef](#)]
10. Irish, J.L.; Resio, D.T.; Cialone, M.A. A surge response function approach to coastal hazard assessment. Part 2: Quantification of spatial attributes of response functions. *Nat. Hazards* **2009**, 51, 183–205. [[CrossRef](#)]
11. Toro, G.R.; Resio, D.T.; Divoky, D.; Niedoroda, A.W.; Reed, C. Efficient joint-probability methods for hurricane surge frequency analysis. *Ocean Eng.* **2010**, 37, 125–134. [[CrossRef](#)]
12. Ho, F.P.; Myers, V. *Joint Probability Method of Tide Frequency Analysis Applied to Apalachicola Bay and St. George Sound, Florida*; U.S. Department of Commerce: Silver Spring, MD, USA, 1975.
13. Myers, V. *Joint probability method of tide frequency analysis applied to Atlantic City and Long Beach Island, New Jersey*; U.S. Department of Commerce: Silver Spring, MD, USA, 1970.
14. Shen, W. Does the size of hurricane eye matter with its intensity? *Geophys. Res. Lett.* **2006**, 33, L18813. [[CrossRef](#)]
15. Westerink, J.J.; Luettich, R.A.; Feyen, J.C.; Atkinson, J.H.; Dawson, C.; Roberts, H.J.; Powell, M.D.; Dunion, J.P.; Kubatko, E.J.; Pourtaheri, H. A basin- to channel-scale unstructured grid hurricane storm surge model applied to southern Louisiana. *Mon. Weather Rev.* **2008**, 136, 833–864. [[CrossRef](#)]
16. Cobell, Z.; Zhao, H.; Roberts, H.J.; Clark, F.R.; Zou, S. Surge and wave modeling for the Louisiana 2012 Coastal Master Plan. *J. Coast. Res.* **2013**, 88–108. [[CrossRef](#)]
17. Luettich, R.A.; Westerink, J.J. *Formulation and Numerical Implementation of the 2d/3d Adcirc Finite Element Model Version 44.Xx*; R. Luettich: Chapel Hill, NC, USA, 2004.
18. Fischbach, J.R.; Johnson, D.R.; Ortiz, D.S.; Bryant, B.P.; Hoover, M.; Ostwald, J. *Coastal Louisiana Risk Assessment Model: Technical Description and 2012 Coastal Master Plan Analysis Results*; RAND Corporation: Santa Monica, CA, USA, 2012.
19. Johnson, D.R.; Fischbach, J.R. Using cost-effective and robust strategies to assess the potential for nonstructural risk reduction in coastal Louisiana. In *Improving Flood Risk Estimates and Mitigation Policies in Coastal Louisiana under Deep Uncertainty*; RAND Corporation: Santa Monica, CA, USA, 2013; pp. 67–101.

20. Fischbach, J.R.; Johnson, D.R.; Kuhn, K.; Pollard, M.; Stelzner, C.; Costello, R.; Molina-Perez, E.; Sanchez, R.; Roberts, H.J.; Cobell, Z. *2017 Coastal Master Plan: Model Improvement Plan, Storm Surge and Risk Assessment Improvements (Subtask 4.9)*; Coastal Protection and Restoration Authority: Baton Rouge, LA, USA, 2015.
21. *Performance Evaluation of the New Orleans and Southeast Louisiana Hurricane Protection System*; US Army Corps of Engineers: New Orleans, LA, USA, 2009.
22. *Louisiana Coastal Protection and Restoration Technical Report*; US Army Corps of Engineers: New Orleans, LA, USA, 2009.
23. Dietrich, J.C.; Tanaka, S.; Westerink, J.J.; Dawson, C.N.; Luettich, R.A., Jr.; Zijlema, M.; Holthuijsen, L.H.; Smith, J.M.; Westerink, L.G.; Westerink, H.J. Performance of the unstructured-mesh, SWAN+ADCIRC model in computing hurricane waves and surge. *J. Sci. Comput.* **2012**, *52*, 468–497. [[CrossRef](#)]
24. Booij, N.; Holthuijsen, L.H.; Ris, R.C. The “swan” wave model for shallow water. *Coast. Eng. Proc.* **1996**, *25*. [[CrossRef](#)]
25. Peyronnin, N.; Green, M.; Richards, C.P.; Owens, A.; Reed, D.; Chamberlain, J.; Groves, D.G.; Rhinehart, K.; Belhadjali, K. Louisiana’s 2012 Coastal Master Plan: Overview of a science-based and publicly-informed decision making process. *J. Coast. Res.* **2013**, *67*, 1–15. [[CrossRef](#)]
26. Resio, D.T. *White Paper on Estimating Hurricane Inundation Probabilities*; US Army Corps of Engineers: New Orleans, LA, USA, 2007.
27. Johnson, D.R.; Fischbach, J.R.; Ortiz, D.S. Estimating surge-based flood risk with the Coastal Louisiana Risk Assessment Model. *J. Coast. Res.* **2013**, *67*, 109–126. [[CrossRef](#)]
28. *Final Post Authorization Change Report: Morganza to the Gulf of Mexico, Louisiana*; US Army Corps of Engineers: New Orleans, LA, USA, 2013.
29. Johnson, D.R. *Improving Flood Risk Estimates and Mitigation Policies in Coastal Louisiana under Deep Uncertainty*; Pardee RAND Graduate School: Santa Monica, CA, USA, 2013.
30. Holder, K. *FEMA Accredits Hurricane and Storm Damage Risk Reduction System (HSDRRS)*; US Army Corps of Engineers: New Orleans, LA, USA, 2014.
31. Johnson, D.R.; Fischbach, J.R.; Kuhn, K. *Current and Future Flood Risk in Greater New Orleans*; The Data Center: New Orleans, LA, USA, 2015.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).