Coastal Protection and Restoration Authority
150 Terrace Avenue, Baton Rouge, LA 70802 | coastal@la.gov | www.coastal.la.gov

2017 Coastal Master Plan

Attachment C5-1: Predictive Models Technical Advisory Committee (PM-TAC) Report

Report: Version I

Date: November 2016

Prepared by: Predictive Models Technical Advisory Committee (John Callaway, Scott Hagen, Courtney Harris, Wim Kimmerer, and Michael Waldon)

Coastal Protection and Restoration Authority

This document was prepared in support of the 2017 Coastal Master Plan being prepared by the Coastal Protection and Restoration Authority (CPRA). CPRA was established by the Louisiana Legislature in response to Hurricanes Katrina and Rita through Act 8 of the First Extraordinary Session of 2005. Act 8 of the First Extraordinary Session of 2005 expanded the membership, duties, and responsibilities of CPRA and charged the new authority to develop and implement a comprehensive coastal protection plan, consisting of a master plan (revised every five years) and annual plans. CPRA's mandate is to develop, implement, and enforce a comprehensive coastal protection and restoration master plan.

Suggested Citation:

Callaway, J., Hagen, S., Harris, C., Kimmerer, W., and Waldon, M. (2016). *2017 Coastal Master Plan: Attachment C5-1: Predictive Models Technical Advisory Committee Report* (pp. 1-43). Baton Rouge, Louisiana: Coastal Protection and Restoration Authority.

Table of Contents

List of Abbreviations

ADCIRC-SWAN	Advanced Circulation-Simulated Wave Near-Shore model
CASM	Comprehensive Aquatic Systems Model
CPRA	Coastal Protection and Restoration Authority
ESMF	Earth System Modeling Framework
EwE	Ecopath with Ecosim model
HPC	High Performance Computing
HSIs	Habitat Suitability Indices
ICM	Integrated Compartment Model
PM-TAC	Predictive Models Technical Advisory Committee
MDT	Modeling Decision Team
LIDAR	Light Detection and Ranging
CRMS	Coast-wide Reference Monitoring System
MEM	Marsh Equilibrium Modeling
TroSim	Trophic Simulation Model
USACE	United States Army Corps of Engineers

1.0 Introduction

The Predictive Models Technical Advisory Committee (PM-TAC) for Louisiana's 2017 Coastal Master Plan was formed in 2013 to provide input on the use of models throughout the development of the 2017 Coastal Master Plan. This Final Report for the PM-TAC is designed to provide feedback on the overall modeling methodology and to provide advice and recommendations to improve both the modeling approach and the review process for future master plans. In this regard, the PM-TAC envisions the audience for this report as: the Modeling Decision Team (MDT), which includes staff at both the Coastal Protection and Restoration Authority (CPRA) and The Water Institute of the Gulf (the Institute), current and future modeling teams, and future PM-TAC members. Those interested in strategic planning for coastal management and restoration in other regions may also find the report of interest.

Members of the PM-TAC:

- John Callaway (Chair, also member of the 2012 PM-TAC) - University of San Francisco
- Scott Hagen - Louisiana State University
- Courtney Harris - Virginia Institute of Marine Science
- Wim Kimmerer - San Francisco State University
- Michael Waldon - U.S. Fish and Wildlife Service (retired)

Section 2 provides a brief background, while Section 3 lists the key recommendations from the PM-TAC. The remainder of the report (Sections 4 and 5) provides additional comments and suggestions to identify both strengths and weaknesses of the current approach and to improve future modeling efforts. Throughout the report, recommendations are given in ***bold italics***.

2.0 Background

The PM-TAC for the 2017 Coastal Master Plan built on previous review and input from the PM-TAC for the 2012 Coastal Master Plan; much of the guidance and recommendations from the 2012 PM-TAC were incorporated into the approach that was taken for the 2017 effort. In particular, the 2012 PM-TAC met primarily by phone with only one in-person meeting and additional in-person meetings were recommended. Based on this input, the 2017 PM-TAC had a small number of phone webinars and multiple in-person meetings. Meetings with the MDT and the modeling team occurred in December 2013, July and December 2014, March and September 2015, and March 2016, with an additional meeting of the PM-TAC in September 2016 to finalize this report. Individual meeting reports are provided in the appendix at the end of this document.

In order to improve the focus for PM-TAC input, the MDT provided a series of questions for the PM-TAC to address for each meeting, and brief meeting summary reports were required from the PM-TAC to recap comments and evaluations. These modifications improved the input from the PM-TAC as discussed below. The individual PM-TAC meeting reports, included as attachments to this document, provide additional detail on topics and recommendations from earlier meetings.

Overall, the PM-TAC believes that the modeling effort for the 2017 Coastal Master Plan substantially improved the 2012 approach. The modeling team made good progress in integrating model components, addressing uncertainty, incorporating scenarios, and more. The modeling team was also very responsive to PM-TAC suggestions, requests, and questions.

3.0 Key Recommendations

This section provides the PM-TAC's major recommendations, which include overarching issues that cut across various components of the modeling system, and those which the committee members agreed were most pressing to consider in future efforts.

Models of this level of complexity should be poised to evolve in ways that take advantage of advances in data availability, computing capacity, emerging technologies, improved understanding of relevant processes, and evolving user needs. For this reason, the ***models developed for the 2017 Coastal Master Plan should be expected to continue to evolve with a long-range goal of a fully integrated modeling system***. Features such as modularity and reliance on open-source code will help the modeling team as they continue to develop the modeling system.

Immediate future development should focus on sections of the models that were identified as most important in the 2017 model sensitivity and uncertainty analysis, or where current model dynamics appear to be overly simple. Focused studies to improve the model in these areas might include:

- Methods of collection, ground-truthing, and interpretation of Integrated Compartment Model (ICM) topographic data,
- Dynamics of marsh edge erosion,
- Dynamics of organic matter accretion in marsh soil,
- Analysis of any collection bias or other challenges in use of available total suspended solids data, and
- Population processes of important species for use in developing species-specific models.

In preparation for future work, the PM-TAC recommends that the modeling team closely examine the quality and quantity of data that are used to configure and calibrate the models, and identify which emerging types of data would be most useful for further improvement or testing of the models. For example, it is critical to consider the spatially varying uncertainty (i.e., in marsh vs. developed regions) of Light Detection and Ranging (LIDAR) data for storm surge calculations in particular and, more generally, in other aspects of the master plan. In addition, data records for calibration of sediment transport were limited and improvements to this data for model development should be evaluated. In cases when relevant field data are limited, parameters from other modeling studies or derived from laboratory experiments may augment information currently used in model development.

Considering both spatial and temporal variability in future scenarios is critical. The combined effect of sea level rise and subsidence was shown to significantly affect land loss. Spatial variation in model output appears to be driven largely by patterns of subsidence, and ***future model scenarios should be sure to use the best and most up-to-date information for subsidence***. Relative sea level rise is the major factor driving overall model response and marsh sustainability. ***Further, future model analyses should consider how accelerations in the rate of sea level rise may affect marsh sustainability***. This does not mean changing the sea level rise scenario, but evaluating how the system responds to accelerations in the rate of rise.

Although there are many problems inherent in the use of Habitat Suitability Indices (HSIs), their use is unavoidable in this modeling process because the output of the physical models is land area, which also implies habitat area. ***Future work should investigate the robustness of some of***

the HSI models to alternative formulations and to areas of uncertainty in data. The team doing Ecopath with Ecosim (EwE) modeling improved the model and the documentation. It is unclear, however, how well the EwE model output tracks historical trajectories, or how EwE is used in the overall modeling process. ***Future planning should consider a long-term effort to develop species-specific population models as an alternative or a complement to EwE and HSIs.***

The flow of water is the fundamental control on virtually all aspects of the master plan modeling efforts. Therefore, it is necessary to improve the integration of all aspects of water flow into the ICM. Towards that end goal ***the PM-TAC recommends that rainfall and runoff processes be more fully integrated into the tide, wind-wave, and surge model***. In addition, ***it is recommended that investigation of sensitivity of project evaluations to the impact of individual storms be compared to other factors (e.g., sea level, subsidence is undertaken).***

Uncertainty analysis for complex models is a maturing field. The modeling team is therefore encouraged to keep particularly alert to new research findings in model uncertainty analysis. Uncertainty analysis is closely related to model calibration, validation, and sensitivity analysis. ***The PM-TAC recommends that a combined strategy for all of these analyses be devised early during future model planning.*** While the model development team should be given the flexibility to adapt as the model materializes, early planning for uncertainty analysis may eliminate some unnecessary model runs and contribute to the effective development of future models.

Much remains to be learned from application of the current model, but analysis of model results is time consuming. ***Effort put into developing more efficient means of analyzing and post-processing model output could streamline synthesis of model results and enable the future modeling team to complete many more model runs for scenarios and uncertainty analysis.***

The PM-TAC review has benefited from multiple in-person meetings, and overall the review process has been efficient. ***The PM-TAC recommends including a more intensive meeting schedule earlier in the process and additional preparatory materials such as a briefing package.***

4.0 Specific Recommendations

Sections 4.1 – 4.8 list items that did not rise to the level of the Key Recommendations above. They are organized as related to individual components of the modeling framework. These recommendations include issues that were important but secondary to the Key Recommendations, or those for which the PM-TAC did not reach consensus on their advice.

4.1 Integrated Compartment Model (ICM)

Model planning, development, testing, and application for the 2017 ICM were daunting. The modeling team, with oversight from the MDT, developed an extremely ambitious modeling plan at the initiation of the project. The PM-TAC recognizes and commends all involved for the success of this complex integrated modeling project.

It is clear that the objectives of the ICM project were well understood by the modeling team. Successful completion of modeling followed from a set of well-defined and explicit objectives and a well-considered initial design. Project management supported the success of the project by avoiding "mission creep" through the stages of model construction and testing. ***In future modeling every decision should be traceable to clear objectives.*** Explicitly stating objectives helps modelers, users, and managers as they consider options and extensions during model development and application.

We concur with the modeling team's decision to integrate, as far as feasible, the ICM model code. ***The goal of model code integration should continue to guide future modeling designs.*** Code integration facilitated automating interaction among the sub-model components, eliminating programmer interaction during model runs, reducing the likelihood of errors, and simplifying quality assurance.

The PM-TAC supports the 2017 model design goal of maintaining modularity within the integrated code and recommends that this coding strategy should continue in the future. Modularity refers to programming and designing input databases in a structure and style that provide for separate testing and incorporation of alternative or revised modules for specific parts of the model. It can simplify programming, testing, documentation, and quality assurance. Individual modules can be modified without requiring wholesale changes to the complete modeling system and facilitates object-oriented conceptualization of model structure and dynamics.

The 2017 ICM has a spatial layout and resolution that was well designed and appears adequate for meeting many, if not all, of the modeling objectives. It remains an open question, however, whether all components of the model are appropriately resolved, and whether resolution-driven errors in one component propagate into subsequent modules. For example, suspended sediment fluxes and deposition likely have large spatial and temporal variability that are not resolved by the current model. ***Future efforts should consider the sensitivity of critical model calculations to uncertainties and errors in intermediate modeled values (e.g., sediment fluxes, deposition) in order to inform subsequent modeling efforts as to the appropriateness of the approach.***

The modeling team should exercise restraint in adding spatial resolution in future models, considering tradeoffs between desired model accuracy and computational costs. Greater numbers of compartments and increased spatial and temporal resolution will increase computational demand, and smaller compartments can exacerbate numerical instability. Conversely, model sensitivity tests or forthcoming data streams may indicate that certain model components are either over- or under-resolved. Some model components are not as data-rich as others. It would be useful to use sensitivity analysis of the model to evaluate how to design future data collection based on what has been learned from the current model calibration and development. For example, because data on suspended sediments were sparse and not collected across the full range of conditions and locations, it seemed difficult for the modeling team to fully assess the model skill in this area. ***The PM-TAC recommends that modelers prepare a brief stand-alone report or appendix to identify and report specific, critical data gaps, and needs.*** Reporting these gaps along with suggested monitoring revisions as they are recognized would assist future research and monitoring design not only for the Coast-wide Reference Monitoring System (CRMS) network, but also to provide justification for funding agencies and the research community to begin to fill those data gaps.

The dynamics of coastal marsh edge erosion are not well understood. Mechanisms of marsh edge erosion were explored in detail by the 2017 modeling team. The final 2017 ICM used a simplified formulation to describe marsh edge erosion that does not include dynamic mechanisms and local variability that may play an important role in land loss. ***Future modelers should continue to evaluate the importance of marsh edge erosion in overall land loss dynamics and consider the potential benefit of incorporating more mechanistic erosion dynamics into future model iterations.***

Future modeling should attempt to develop improved linkages among organic accretion dynamics, vegetation growth, and nutrient concentrations within the ICM model. Analysis of the 2017 modeling sensitivity and uncertainty determined that the organic matter accretion rate was an important factor for marsh sustainability (Attachment C3-24: Integrated Compartment Model Uncertainty Analysis). While modeling details of organic matter accretion (e.g., as in Marsh Equilibrium Modeling (MEM) and similar models: Schile et al., 2014; Alizad et al., 2016) is not feasible on this large spatial scale, it would be useful to evaluate whether some of the concepts of these process-based models could be incorporated into the ICM. For example, vegetation is a trap for inorganic sediment and a source of organic sediment (Kirwan and Murray, 2007). Vegetative standing crop and productivity are linked to soil nutrient concentrations, which are affected by water column nutrient concentrations and nutrient loading. The dynamic homeostatic relationship between plants and accretion may therefore depend in part on nutrient loading (Morris et al., 2002).

Calibration performance measures should be well defined and represent variables or properties of variables that are important to users and managers; the modeling team should consider the spatial and temporal scale of the available calibration data and the scale of the modeled processes and model output. For example, if a HSI for a species of fish is calculated using average monthly salinity from January to July, then it is important to know the calibration performance of the salinity model for this output. In this case, the model's performance for seasonal average salinity may differ markedly from calibration performance based on instantaneous comparison to field measurements. As a second example, if the number of days of inundation is a critical parameter for a habitat model component, then the model should be calibrated to days of inundation. In some instances, the model grid covers much larger spatial domains than the sampling program that provided the data. These concerns seemed especially apt for sediment data (including accretion rates and total suspended solids). Accretion data from single points were compared to modeled output for relatively large areas. While the model provided a reasonably good fit for accretion data when compared across all locations, sites with the highest accretion rates (i.e., deltas) could be driving the model fit. The model fit for non-deltaic locations should be further evaluated. Similarly, sampling of total suspended solids generally occurs during quiescent field conditions and from water column samples; yet, suspended sediment concentrations peak near the bed during storm conditions. This probably leads the calibrated model to underestimate sediment fluxes, and may have influenced sensitivity and uncertainty estimates.

Coastal land changes may have profound impacts on adjacent areas, including areas outside the current model domain, and these changes may affect the future trajectory of the coast. For example, how will projected future coastal land changes be affected by, and how will they affect the Gulf Dead Zone and flooding driven by stream flow and local rainfall? These questions are of economic and ecological importance to southern Louisiana but may be deemed beyond the scope of the master plan. ***The PM-TAC recommends that, where possible, CPRA and the Institute make an effort to collaborate with other institutions or agencies that have planning responsibility for these areas to promote information and data exchange.***

4.2 Use of Scenarios

In diverse modeling applications (e.g., financial, electric power grid, flood risk, epidemiology, and environmental), *scenario analysis* is commonly used to assess uncertainty and risk associated with the numerous contingencies, i.e., events that individually have a small likelihood but can influence future outcomes. In the case of the Louisiana coast, future contingencies include hurricanes, climatic change, and relative sea level rise. In scenario analysis, models are used to project the impact of an ensemble of possible contingent events. It is desirable to include a wide range of possible events within this selected ensemble.

Scenario selection for the 2017 modeling was tightly constrained by the number of model runs that could feasibly be completed and analyzed and was limited to three scenarios that were selected based on sensitivity testing (*Appendix C: Chapter* 2). In the future, increased computer speed or use of High Performance Computing (HPC) technologies might relax this constraint and allow a greater number of scenarios to be considered. Additionally, automating and streamlining the post-processing of model results might allow for many additional runs (see Model Implementation section). Future modeling will need to consider tradeoffs between adding model complexity and resolution, and expanding the number of feasible model scenario runs. ***The PM-TAC recommends prioritizing the capability to produce more model scenario runs.***

To be evaluated and compared within the ICM, projects will be implemented at some given year and month, and each model run will use the same time series of discharge and weather because storm event (e.g., track, magnitude, and timing) scenarios were not varied in the 2017 master plan modeling. The storms and storm characteristics that were chosen for the ICM represented a small set in terms of storm track, magnitude, and timing (small compared to the natural variability of storms). The PM-TAC remains concerned that a restoration project's likelihood of success, as evaluated using the ICM, may depend on whether a modeled storm passes nearby between construction and the end of the planning period. The interplay between the timing and location of a project relative to the storm tracks may create differences in project outcomes that are an artifact of the model framework. The approach used was justified by findings that project results were less sensitive to storm occurrence than to other factors, such as sea level rise. Future implementations of the ICM might incorporate a more stochastic approach to representing storm occurrences, in which numerous storm tracks are modeled and an ensemble probability of success of individual projects could be calculated. ***Future modeling studies might reconsider how to use modeling tools to investigate response within the ICM to a more widely varying range of storm tracks, magnitude, and timing for evaluating the likelihood of project success.***

Preliminary results have indicated that relative sea level rise is the big driver of future coastal land loss within the ICM. The sea level scenarios are also limited to a small number, and predictions of eustatic sea level rise have changed substantially since 2014 when the 2017 Coastal Master Plan decisions on sea level rise were made. ***Future model planning and design should consider the degree to which the system responds to sea level rise compared to other drivers, and for the next round of the master plan it will be critical to again review the literature to determine rates of sea level rise for the next round of modeling.***

4.3 Ecological Models

Ecological modeling for the 2017 Coastal Master Plan used a combination of HSIs and off-the-shelf EwE. These choices reflected the time and resources available in this effort. Here this selected modeling approach is discussed and recommendations are made for future master plan cycles. The PM-TAC assumes EwE and HSIs will be employed in future modeling efforts.

4.3.1 Habitat Suitability Indices (HSIs)

A HSI is a model that hypothesizes a quantitative relationship between one or more habitat variables and the capacity of the habitat to support the species of interest.[1] The HSI for a given location is assumed to be proportional to the carrying capacity for a unit of habitat in that location for the species of interest; as such, it does not necessarily predict abundance or catch, although catch or abundance data are usually used to develop HSIs. Furthermore, important aspects of habitat may be omitted from the analysis because data or knowledge are lacking. The modeling team should be aware of these limitations of HSIs, both in terms of calibration and in interpretation of results.

Calibration performance of the HSI components should ideally be evaluated and reported in a manner similar to calibration of other model outputs (e.g., stage and salinity). Therefore, whenever feasible, HSIs should be defined using habitat variables that are monitored or can be derived directly from monitoring habitat parameters and species abundance or success data. Defining HSIs using habitat variables that can be directly observed allows quantification of the calibrated model's capability to predict each HSI through comparison of historical observations to modeled HSI values.

ICM modeling team members should participate in future HSI design. Impact of HSI structure on model performance as well as model capability to accurately predict HSI values should be considered. Participation of modelers provides the HSI designers with expert opinion from the modeling team concerning the credibility of specific modeled habitat variables at the time that the HSIs are conceived. The PM-TAC did not find any evidence of lack of communication between teams during 2017 HSI development, but it is recommended that early participation by modelers in the future might avoid later delays or uncertainties in model application. Further, future participation of modelers insures that the modeling team is informed of what hydrologic and constituent model output variables are most important in model application.

The fish and shellfish HSIs were defined as polynomial curves fitted to the available data for each species (see individual HSI reports). Early in the 2017 process, the PM-TAC commented on this selection of methods, recommending in particular the use of piecewise linear (e.g., trapezoidal) functions instead of polynomials. Polynomials can be useful when an underlying process has a second-order functional form; however, the functional forms of most habitat use patterns are unknown. Polynomials therefore impose a particular shape that is unlikely to be related to the underlying habitat use, and may tend to fit poorly in areas of sparse data and to lack robustness at the edges of the data distribution. ***The PM-TAC recommends use of the simplest possible functions for HSIs and, in particular, to eliminate polynomials from these models.***

4.3.2 Ecopath with Ecosim (EwE)

The modeling team spent considerable effort investigating alternatives and developing and adapting the selected program to their uses. In this effort, the team evaluated several off-the-shelf models to determine which would be most suitable. The team investigated Trophic Simulation Model (TroSim) and Comprehensive Aquatic Systems Model (CASM) before selecting EwE (and also Ecospace) for modeling biological responses to changes forecasted by other

[1] http://www.nwrc.usgs.gov/wdb/pub/hsi/hsiintro.htm

model elements. The EwE complex of models[2] is a mature modeling platform which has been widely applied (Heymans et al., 2016). The modeling team made a strong case for using EwE over TroSim and CASM based on the earlier analysis by Rose and Sable (2013), a consideration of modeling objectives, and a comparison of the results of simulations using both EwE and TroSim. The PM-TAC agrees with EwE as applied in the above report, and including the enhancements made by the modeling team, it is the most suitable available tool for supporting development of the master plan.

Many of the available time series of catch used in calibrating and validating EwE had rather small interannual variability. This may have limited the ability of the calibration process to capture important links between habitat and biomass. This limitation may impair the accuracy of forecasts with EwE for conditions not previously observed (the main point of the modeling exercise). ***As additional data become available in the future, modeling should use this extended time series of calibration data and determine the improvement obtained from use of a longer calibration period. If these data are found to be insufficient for calibration, the PM-TAC recommends exploring use of some alternative methods (e.g., experimental analysis) for detecting and representing relationships between environment and biomass for some important species.***

The EwE calibration output revealed a range of fits of the model to the time-series data. It was difficult to reconcile the discussion and evaluation of the fits with the graphical presentations (e.g., Fig. 8 in EwE report). Some of these graphs suggested a systemic problem in how the model captured the variability. These were most noticeable in some of the freshwater fish, such as sunfish, largemouth bass, and killifish, all of which declined in abundance over the time series relative to model predictions. ***A better calibration may not be possible, but the report should address confidence in model predictions by discussing points where the model did poorly as well as those where the fit was reasonably good.***

The results of the 50-year simulation were analyzed to identify potential problems with model elements. Mortality of catfish was clearly too low in the initial run, and some of the potential additional predators were identified and either added or flagged for a future revision of the model (e.g., alligators). Also a crash of the largemouth bass population was traced to apparently excessive sensitivity of adults to habitat fragmentation, and allowing adults more freedom to move solved the problem. These findings somewhat undermine confidence in other aspects of the model simulations for which no problems were obvious. ***Future effort is recommended to identify additional, subtler anomalies and their causes.***

It was unclear what use was made of EwE output in conjunction with ICM and other model output or how forecasts using EwE matched those made using HSIs. These aspects of model application should be included in the reporting for the 2017 Coastal Master Plan. This is a critical piece of output for the public and for the 2017 Coastal Master Plan metrics (Reed et al., 2016). Thus, the modeling team should clarify the benefits of EwE and what information it is providing to support the master plan.

4.3.3 Recommendations for Future Ecological Modeling

Although the PM-TAC was satisfied with the progress on the HSI and EwE components of the modeling package for the 2017 Coastal Master Plan, such a modeling effort should never be

[2] http://ecopath.org/

considered complete. The fundamental EwE assumption of a time- and space-varying mass balance was reasonable, but models describing habitat use are likely to be inaccurate, as revealed by the two cases in the 50-year simulation. The focus of the HSI modeling effort was appropriately on physical habitat, since that is the principal way that forecasted changes would affect the various species. However, the actual use of a habitat by a biological population depends on factors besides the extent of that habitat, such as predation risk and the availability of alternative habitat. These factors are unlikely to be well represented for many of the species in a modeling effort as comprehensive as this. ***Therefore, future efforts should revisit the habitat descriptions and examine the basis for each, particularly for those species whose trajectories are poorly modeled by EwE. In addition, the modeling team should consider developing new tools that focus on particular species, either as supplements or replacements for the use of EwE and HSIs in the planning process. The form these models take would depend on the availability of data and the specific inputs expected from the physical models and outputs needed from the biological models.***

4.4 Storm Surge and Wave Models

The storm surge and wave model development and application to the 2017 Coastal Master Plan were built from a sound basis that spans over the past two decades. The calculations include contributions from waves from the Simulating WAves Near-Shore (SWAN) model, and tidal forcing and storm surge from the Advanced Circulation (ADCIRC) model. As a result, the intricate system of marshes, waterways, levees and related infrastructure, roadways, railroads, etc. for coastal Louisiana in general and in New Orleans in particular was well represented at the start of the first master plan. That discrete representation of topography and bathymetry has been updated, along with improvements to the description of surface characteristics to maintain the 2017 Coastal Master Plan storm surge and wave model as state-of-the-art for tides, waves, and surge simulations in coastal Louisiana. The calibration and validation of the surge and wave models were adequately explained and are well documented.

The wave and surge model as developed for the 2017 Coastal Master Plan presents opportunities for expansion because it is likely that the ADCIRC-SWAN model for the 2017 Coastal Master Plan has untapped potential that can be realized in future master plans. To realize their full potential may require more than just building on what has been done, but also re-examining some of the approaches to the wave and surge modeling. After all, the basis for all biogeophysical processes needed for the master plan is the flow of water, whether the water originates from the land mass or the Gulf of Mexico. The more directly the physics of flow (both hydrologic and hydraulic) are integrated into the ICM the better. Below are some possible paths forward.

It was recognized in early stages of surge model development for coastal Louisiana that the system of defense from storm surge and wave energy is dynamic. For example, new levees are continually being added to the defense system, and existing levees are subsiding. It should also be noted that much of the historical focus of the ADCIRC model development has been placed on southeast Louisiana, perhaps in part because data describing levee characteristics are more difficult to obtain for southwestern Louisiana. ***The PM-TAC recommends that the natural and human modifications to the system that are relevant to the master plan be regularly updated in a database and incorporated in the ADCIRC model mesh.***

LIDAR was used to determine surface elevations over vast extents of the model domain. Compared to forests and developed areas, however, LIDAR is less reliable in thickly vegetated (e.g., marshes) and uninhabited areas where ground-truthing of bare earth elevations is limited

(Rosso et al., 2006; Wang et al., 2009; Medeiros et al., 2015). ***The PM-TAC recommends that topographic data should be ground-truthed with spot-checks throughout coastal Louisiana in marsh and other regions where LIDAR data may be suspect.***

In addition, surface characterization has largely depended on land use and land cover data while finer-scale variations in characteristics (e.g., health or maturity of vegetation, fine scale patchiness, etc.) within a classification have not been distinguished. Wave and surge models are especially sensitive to classifications of land use and land cover, and misclassification can lead to substantial errors (Medeiros et al., 2012). In addition, not considering the full impact of climate change on the vegetation may limit predictions of future conditions of surface characteristics. It is recommended that future alterations of vegetative cover for coastal protection should be explored and included within the storm surge modeling to investigate their ability to attenuate waves and reduce surge propagation (e.g., if coastal forests were planted on a large scale, would they produce effective attenuation?). ***The PM-TAC recommends that the sensitivity of surface characterization of surge model parameters from land use and land cover is evaluated and more directly coupled to remote sensing data (e.g., LIDAR).***

Within the context of surge modeling, the major motivation for short-wave modeling is with respect to momentum transfer from the breaking short wave to the shallow-water long wave, which usually induces a higher surge. That physical process is well described by the surge and wave model and has been validated and documented. However, waves generated over the land mass when overland areas are intermittently flooded and the impact they have on the system have not been a focus of the ADCIRC-SWAN model application to the master plan. ***The PM-TAC recommends that the modeling of overland waves and the impact to the overall system should be considered.***

Credibility of surge modeling under future conditions was limited by the quality of subsidence estimates. This was due to scientific limitations and is not a fault of the modeling team, who are using reasonable approaches; however, subsidence remains a major source of uncertainty. Note that a recent example of the complexity of subsidence can be found in Jones et al. (2016). Uncertainty in the temporal and spatial variation of subsidence limited the ability of the models to describe future conditions of the land surface and the future levee elevation variations. ***The PM-TAC recommends that modelers run sensitivity tests that cover the range of projected subsidence rates to determine the degree of uncertainty that this adds to modeled output variables.*** Such sensitivity tests should not be limited to surge but should span all modeling efforts as appropriate.

As developed, the surge and wave models were extremely computationally expensive which limited a more extensive and direct employment for the master plan. ***The PM-TAC recommends that a simplified version of the surge model could be developed that can be run with less computational cost.*** For an example, see MacWilliams et al. (2016). The existing high-resolution surge and wave model that has been well validated could serve as a "true" solution upon which to compare more simplified models derived from the same topographic and bathymetric and surface characteristics. Development of a simplified (e.g., in terms of mesh resolution) surge and wave model would enable a greater number of storm scenarios to be run for future projects. It is feasible that a computationally efficient wind-wave and surge model could offer the opportunity to more directly incorporate surges and waves into the master plan or inform sediment and salinity transport for the next version of the ICM.

The PM-TAC recommends that episodic rainfall and runoff processes be more fully integrated into the tide, wind-wave and surge model to enable simulation of all aspects of flooding at the coastal land margin. While the state-of-the-art to incorporate overland and riverine flooding into large-scale surge models is limited, there are ongoing efforts to integrate hydrologic processes into surge models that would benefit the next coastal master plan. Such integration could provide a more rigorous evaluation of flood risk in coastal Louisiana and form a basis to integrating all of these flow processes into the modeling effort.

4.5 Risk Modeling

The risk model was an innovative application within coastal sciences, and carried some exciting possibilities with it, including the incorporation of the ICM results with land use and infrastructure projections. The juxtaposition of these model components, however, added complexity. Both components (i.e., the storm surge model and infrastructure projections) carried with them some uncertainty and error terms. The treatment of the complexity of the model, including both consideration of model sensitivity and communication of model assumptions and parameterizations, is challenging. The uncertainties in the model products may be obscure to all but the most sophisticated end user. ***Care should be taken to communicate the methods and results to the scientific and management community so that the end users understand the strengths of the approach as well as the assumptions that go into the risk calculations.***

The sensitivities to modeling constructs should be explained and explored. The results of the risk study seem especially sensitive to model constructs and assumptions. It is unlikely that end users would delve into the details of the risk model to the degree necessary to evaluate the appropriateness of these assumptions, so it remains the job of the modeling team to explore the degree to which these assumptions add uncertainty to the model result, and to communicate them effectively. Additionally, because the use of these models is relatively new within coastal sciences, future research should evaluate how well the projections perform and provide guidance for refining the methods.

For future model versions, it may be useful to incorporate human population migration following storms. The population projections within the modeling framework were assumed at the outset, and were static in that they did not incorporate feedbacks from other model results. In actuality, population densities and land use will respond to gradual increases in flooding (i.e., both coastal surge and rainfall-induced) and to damage by extreme storms (Qiang and Lam, 2016). Incorporating feedbacks between population densities and land use to flooding intensity and storm impacts may increase the realism of the risk model.

4.6 Model Uncertainty

Model uncertainty analysis attempts to provide a probabilistic description of the difference between model output and the true value (Guzman et al., 2015). Model error and uncertainty derive from several sources including data error, data availability, model structure including assumed mechanisms of action, assumptions of causality, the form of model equations, and the values of parameters used in equations. Future scenarios for this modeling effort rely on understanding of parameters and mechanisms as limited to present and historical knowledge and experience. Similarly, there is uncertainty in project outcomes, as projects may not perform exactly as expected. Parametric uncertainty in modeling is the uncertainty in model output that arises from inexact knowledge of the value of parameters used. This section focuses on parametric uncertainty, which will be referred to as uncertainty within the remainder of this report section. Researchers, users, and readers should remain mindful of all potential sources of error**.**

Terminology and methods for the study of uncertainty are not standardized (Moriasi et al., 2015). It is therefore important to clearly define terms and justify methods when presenting results showing model uncertainty. This has been done in documentation of the 2017 modeling; ***future modeling documentation should clearly define terms and methods and discuss interpretation of uncertainty prior to presentation of findings.*** Calibration error is different from uncertainty. Calibration error calculated, for example, as mean absolute error, gives a conceptually straightforward measure of how well the model recreates historical observations. Uncertainty analysis provides statistical descriptions of future model projections. Although different, both calibration error and uncertainty analysis are useful in supporting management decisions.

For large, complex models, uncertainty analysis is constrained by the number of model runs that can practically be completed. A complete and robust uncertainty analysis of all model parameters and inputs is infeasible for models with a level of complexity comparable to the ICM. The decision as to which sources of uncertainty to investigate was carefully considered in the current effort. Despite anticipated gains in computer speed and the hope that future model code will be optimized for faster execution, similar decisions concerning which parameters to analyze will likely be required in future modeling. ***Based on knowledge of model sensitivity, the modeling team should identify the modules of the ICM that carry with them uncertainties or potential errors that propagate into subsequent modules used to make key predictions.*** As a result, the modeling team will then be able to identify key components where modeling techniques (e.g., higher spatial resolution) would mitigate error propagation. ***The modeling team and CPRA should identify any modules for which uncertainty or error does not significantly limit the utility of final predictions.*** Those modules may then be replaced with more simplistic models without sacrificing the reliability and usefulness of the overall modeling system.

4.7 Model Implementation

4.7.1 Computational Approach

Within the current phase of the project, a large effort has been devoted to integrating model components, and the modeling team developed an implementation plan that allowed them to complete the integration, produce multiple model runs, and analyze results within the time allowed. However, some of the approaches that were necessary during this phase of the implementation should be revisited in planning future modeling work.

During 2017 model development, the PM-TAC recommended evaluating use of open-source software and working within the Unix operating system. These approaches would allow for the ICM to take advantage of parallel computing technologies, which should reduce model runtime. In addition, Unix and HPC are standards within modeling research communities that deal with complex Earth system models (e.g., ESMF Joint Specification Team, 2011; Peckham et al., 2013). Aligning the ICM with these research community standards would have benefits, including using community tools for model visualization, analysis, and debugging. Additionally, migration toward community modeling standards might facilitate future coupling between the ICM and other model components. ***Future modeling should attempt to use open-source software and work within the Unix operating system.***

Computation time is not the only limitation for model runs; analysis of model results is also time consuming. As the ICM matures, and the model runs becomes faster, the ***modeling team should work to develop tools that take advantage of automated model review to expedite analysis of model runs.*** Use of community tools may facilitate efficient analysis of model output.

4.7.2 Project Implementation within the Model

The approach to evaluate projects assumed that projects would work as designed (e.g., shoreline protection will reduce erosion by a specified amount); however, ***restoration outcomes themselves may be variable and difficult to predict. The modeling team should develop methods for exploring this source of uncertainty in evaluating model outcomes, as it could be an important component of the overall uncertainty.*** Evaluating performance through model hindcasting of past projects provides one source of quantitative information on project uncertainty. The modeled project performance can then be compared to performance anticipated in past management planning. This application of modeling not only provides insight into project uncertainty, but also can be used to support adaptive management for future projects.

To evaluate individual projects, model results were quantified by a set of metrics, including total land loss during targeted time horizons. Total land loss appeared to be most sensitive to the sea level rise, and this factor may have overshadowed the effect of individual projects. ***For comparing different individual projects, it may make more sense to assume a single sea level rise scenario and do additional model runs for different wave and storm scenarios.***

4.8 PM-TAC Review Process

The role assigned to the PM-TAC was designed to include some functions of a completely independent outside review group and some functions of an internal advisory panel. This was an efficient way of organizing the process, although the parameters of this role could have been made clearer at the outset. Particularly, the permissible degree of contact with modeling team members outside of the formal meeting process needed a clear and early definition. The opportunity for two-way communication both during and outside of the regular meetings focused the PM-TAC's advice and made for efficient review. This communication allowed for quick clarification arising from the unfamiliarity of PM-TAC members with model details. The advice and review process thereby became collegial with open communication throughout.

PM-TAC meetings occurred approximately twice per year, which was appropriate for the pace of the project and needs of the project for technical advice and review. ***It would be most effective to continue to organize the meetings around critical points in the modeling process, rather than based on the calendar.*** A higher frequency of meetings earlier in the project would allow the PM-TAC to get up to speed on the modeling framework, linkages, and expected uses. This timing would also provide the modeling team with more technical discussion and advice from the PM-TAC during development of the modeling system, before critical decisions are made. Likewise, an in-person meeting, or an additional conference call between the modelers and the PM-TAC as the Final Report is developed might streamline communications regarding the PM-TAC's recommendations.

Although the efficiency of having a small number of meetings with targeted input was appreciated, some members of the PM-TAC found it challenging to get back to the details of the model on such an infrequent basis. It may be useful to have a conference call between in-person meetings, and a follow-up phone call after the in-person meeting to organize the meeting summary report***. The PM-TAC recommends that the managers and the PM-TAC discuss trade-offs between meeting frequency and efficient use of time, and determine a schedule that provides optimal and efficient advice and review without placing an excessive burden on PM-TAC and modeling team members, as well as the modeling budgets.***

On some occasions, travel schedules of PM-TAC members permitted a brief meeting of the committee members on the afternoon or evening prior to the main meeting with the modeling team. This enhanced preparation for the discussion at the main meeting. While these may not be possible for all meetings, ***it would be useful to schedule these preliminary committee meetings for dates with particularly complex issues, or if the PM-TAC had not met for a prolonged period.*** A brief conference call prior to the meeting (but after preparatory materials and questions were available) would also serve this purpose.

It was useful for the PM-TAC to receive briefing material and focused discussion questions specific to each meeting. These made for more effective preparation for the meeting and for more targeted input and discussion. ***The PM-TAC recommends continuing the use of targeted read-ahead material along with focused questions for each key meeting topic.*** These materials should be available at least a week ahead of in-person meetings and should include a very brief synopsis of the current status of each modeling component and any papers, reports, or other products produced. In the future, it could also include a link to an online status board displaying the current status of each element of the project.

The meeting format was focused and effective. In some cases, the agenda for the meeting was crowded and the review of each of the model components was brief. In cases such as this it may be useful to have a ***brief webinar prior to the in-person meeting. This would allow the modeling team to update the PM-TAC in advance of the meeting and may provide for more productive input at the in-person meetings.*** Breakout sessions at a few meetings were productive and efficient when the needs for review were model-specific. It may also be useful to schedule time for the PM-TAC to meet separately during the meeting day to consolidate understanding, prepare clarifying questions, and discuss the meeting report.

The frequency and scope of the summary and final reports were on-target with the PM-TAC's input and time commitment, in general. The PM-TAC does not recommend any additional reports. The PM-TAC provided a brief summary report within a few weeks of each in-person meeting, which were organized around the focus questions for each meeting. These individual meeting reports were a useful way to finalize input from the PM-TAC and to wrap up thoughts from the meeting discussion. The notes that were provided from the meetings by Alaina Owens Grace or Joao Pereira were very helpful in preparing the PM-TAC's summary meeting reports and allowed the PM-TAC to focus on the discussion rather than taking notes during the meeting.

The preparation of this PM-TAC Final Report has also been an effective way to sum up overall inputs on the entire process, and as above, the hope is that it will be useful in directing input for the future master plans. The PM-TAC developed its Final Report before some of the final reports of the modeling team were complete. While most reports were available and most questions were answered, ***it would have been useful to have access to all of the final reports of the modeling team before preparing the review.*** Alternatively, perhaps an efficient use of the PM-TAC would be to serve as reviewers of the modeling team's final report.

The areas of expertise of PM-TAC members were complementary and covered most of the modeling components. A small committee made it easy to coordinate activities and reports for the PM-TAC. However, the PM-TAC lacked expertise in risk analysis because the member initially engaged for that topic (Brian Harper from USACE) was unable to participate. ***If the risk-modeling component will continue to have a substantial role, a subject-matter expert should be included in a future PM-TAC.***

5.0 Best Practices for Modeling

In addition to the recommendations above, the PM-TAC provides the following more general input that reiterates key points that came up during the meetings and review.

Concerns of the future user community and the public should be considered early in model planning. While an extensive public outreach effort is not suggested, at the commencement of future modeling, providing a channel for feedback is important.

Terminology used in model calibration is often inconsistently defined and applied, and confusion resulting from lack of standardized terms is an obstacle to advancing modeling standards (Moriasi et al., 2015; Zeckoski et al., 2015). Here, a calibration period is defined for which model parameters are adjusted so model output matches data as closely as feasible, and validation period of simulation for which no parameter or model adjustments are permitted except setting model boundaries and inputs to match observed values for that period. Model performance during the validation period then reflects expected model performance in projecting the future scenarios. Various concepts about calibration/validation strategies determine how these periods should be selected. Some would support the concept of including the widest available range of conditions in calibration; others support selecting less variable periods (e.g., a series of unusually wet or dry years) for calibration. In practice, other considerations may also affect the selected ranges. Future modeling teams should set up an explicit calibration and validation strategy early in the model development process (Moriasi et al., 2015). Model documentation should include definitions of the terms applied to these activities.

Care should be taken to select appropriate statistics for summarizing model uncertainty. In particular, the r-squared statistic should be dropped from all reporting of dynamic model performance, as it is misleading in evaluating the uncertainty in predictions of dynamic models. Such predictions could actually be strongly biased and inversely related to observations, but still have a high r-squared. Model bias and mean absolute error are straightforward measures that are easy to understand and should be considered as principal calibration objective measures. Moriasi et al. (2007) suggest guidelines and alternative calibration objectives for assessment of models.

The data available for calibration for any model have inherent limitations, errors, and problems. The PM-TAC recommends that the modeling team continue to evaluate data reliability and incorporate this into the calculations and presentations of model calibration and assessments of the model's suitability for forecasting.

When interpreting findings in the context of uncertainty, it is important to apply informed judgment. When reporting uncertainty, the level of understanding of the audience is an important consideration to avoid misinterpretation. Some users, for example, may conflate common statistical tests with reported model uncertainty. Particularly as future uncertainty analyses provide more extensive findings, it will be important to describe similarities and differences between statistical analysis and uncertainty estimates. Perhaps more important, model results that are comparative (e.g., between a suite of projects and the no-action alternative) should be interpreted using the uncertainty associated with the comparison, not that associated with the individual predictions. For example, the uncertainty associated with the impact of a project (e.g., land area with project minus land area without project) will be smaller than the uncertainty of land area for either individual case.

In complex models like the ICM, sub-models or modules pass output to dependent sub-modules. Model uncertainty propagating through such a sequence of modules may increase or decrease. The propagation of uncertainty from one module to the next should be studied and characterized to gain understanding of how uncertainty in final model results arises in calculations made in previous modules in the sequence.

It is important for modelers to have some first-hand contact with the modeled system. The PM-TAC therefore urges future modeling team managers to encourage members of the modeling team to have some participation in field data collection relevant to their specific modeling responsibilities. Such experience is particularly valuable to junior modeling team members.

The model could be made available for use by other researchers by providing training and technical support. This would allow for additional development of the model outside the current users and developers.

After the 2017 plan is complete, the PM-TAC recommends a continuation of some level of effort among a selected team to review and catalog on-going research progress in unpublished reports and published literature relevant to the model's future development.

6.0 References

Alizad, K., Hagen, S., Morris, J., Bacopoulos, P., Bilskie, M., Weishampel, J., and Medeiros, S.. (2016). A coupled, two-dimensional hydrodynamic-marsh model with biological feedback. *Ecological Modelling*, *327*, 29-43.

ESMF Joint Specification Team. (2011). Earth System Modeling Framework (ESMF) Web Site: http://www.earthsystemmodeling.org/.

Guzman, J., Shirmohammadi, A., Sadeghi, A., Wang, A., Chu, M., Jha, M., Parajuli, P., Harmel, R., Kbare, Y., and Hernandez, J.. (2015). Uncertainty considerations in calibration and validation of hydrologic and water quality models. *Transactions of the ASABE*, *58*, 1745-1762.

Heymans, J., Coll, M., Link, J., Mackinson, S., Steenbeek, J., Walters, C., and Christensen, V. (2016). Best practice in Ecopath with Ecosim food-web models for ecosystem-based management. *Ecological Modelling, 331*,173-184.

Jones, C., An, K., Blom, R., Kent, J., Ivins, E., and Bekaert, D.. (2016). Anthropogenic and geologic influences on subsidence in the vicinity of New Orleans, Louisiana, *Journal of Geophysical Research. Solid Earth*, *121*, 3867-3887. doi:10.1002/2015JB012636.

Kirwan, M., and Murray, A. (2007). A coupled geomorphic and ecological model of tidal marsh evolution. *Proceedings of the National Academy of Sciences*, *104*, 6118-6122.

MacWilliams, M., Bever, A., and Foresman, E. (2016). 3-D Simulations of the San Francisco estuary with subgrid bathymetry to explore long-term trends in salinity distribution and fish abundance. *San Francisco Estuary and Watershed Science*, *14*(2). Article 3. http://escholarship.org/uc/item/5qj0k0m6

Medeiros, S., Hagen, S., Weishampel, J., and Angelo, J. (2015). Adjusting LIDAR-derived digital terrain models in coastal marshes based on estimated above ground biomass density. *Remote Sensing*, *7*, 3507-3525. http://dx.doi.org/10.3390/rs70403507

Medeiros, S., Hagen, S., and Weishampel, J. (2012). Comparison of floodplain surface roughness parameters derived from land cover data and field measurements. *Journal of Hydrology*, *452–453*, 139-149. http://dx.doi.org/10.1016/j.jhydrol.2012.05.043

Moriasi, D., Arnold, J., Liew, M., Bingner, R., Harmel, R., and Veith, T. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*, 885−900.

Moriasi, D., Zeckoski, R., Arnold, J., Baffaut, C., Malone, R., Daggupati, P., Guzman, J., Saraswat, D., Yuan, Y., Wilson, B., Shirmohammadi, A., and Douglas-Mankin, K. (2015). Hydrologic and water quality models: Key calibration and validation topics. *Transactions of the ASABE*, *58*, 1609-1618.

Morris, J., Sundareshwar, P., Nietch, C., Kjerfve, B., and Cahoon, D. (2002). Responses of coastal wetlands to rising sea level. *Ecology*, *83*, 2869-2877.

Peckham, S., Hutton, E., and Norris, B. (2013). A component-based approach to integrated modeling in the geosciences: The design of CSDMS. *Computers and Geosciences*, *53*, 3-12.

Qiang Y., and Lam, N. (2016). The impact of Hurricane Katrina on urban growth in Louisiana: an analysis using data mining and simulation approaches. *International Journal of Geographical Information Science*, *30*, 1832-1852. http://dx.doi.org/10.1080/13658816.2016.1144886.

Reed, D., Hijuelos, A., Hemmerling, S., White, E., Fischbach, J., Green, M., and Saucier, M. (2016). 2017 Coastal Master Plan: Attachment B1, Metrics. Version I. (pp. 1-47). Baton Rouge, Louisiana: Coastal Protection and Restoration Authority.

Rose, K., and Sable, S. (2013). 2017 Coastal Master Plan: Strategy for Selecting Fish Modeling Approaches. Baton Rouge, Louisiana: Coastal Protection and Restoration Authority.

Rosso, P., Ustin, S., and Hastings, A. (2006). Use of LIDAR to study changes associated with *Spartina* invasion in San Francisco Bay marshes. *Remote Sensing of Environment*, *100*, 295-306.

Schile, L., Callaway, J., Morris, J., Stralberg, D., and Parker, V. (2014). Modeling tidal marsh distribution with sea level rise: evaluating the role of vegetation, sediment, and upland habitat in marsh resiliency. *Plos One*, *9*, e88760. doi:88710.81371/journal.pone.0088760.

Wang, C., Menenti, M., Stoll, M., Feola, A., Belluco, E., and Marani, M. (2009). Separation of ground and low vegetation signatures in LIDAR measurements of salt-marsh environments. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 2014-2023.

Zeckoski, R., Smolen, M., Moriasi, D., Frankenberger, J., and Feyereisen, G. (2015). Hydrologic and water quality terminology as applied to modeling. *Transactions of the ASABE*, *58*, 1619-1635. http://dx.doi.org/10.13031/trans.58.10713.

Appendix – Meeting Reports

Meeting reports were prepared by the PM-TAC following each of the six in-person meetings. These meeting reports are provided below, as they were received (i.e., unedited except for formatting for this document).

Meeting 1 – December 19, 2013

Overall Comments

- This is a very ambitious task.
- Impressed with the overall effort.
- The team seems on track to make very useful predictions.
- The work RAND is doing on the risk assessment component is very good.

General Modeling Approach

- Look at SFWMD approaches for examples of coding style, coding requirements, QA/QC, etc.
- Build the ICMs as modules so that each component can be pulled out and replaced. This is important to allow changing out models in the future or for different types of analyses.
- Suggest using “IRF” format, “Initialize, Run, Finalize” when coding each subcomponent which can help with swapping components and possibly parallelization in the future.
- A conceptual model/flow chart of how the modeling components fit together would be useful, including time steps, data transfer protocols, inputs/outputs.
- Take advantage of computing power. May not be possible to create parallel code, but the team could still take advantage of multiprocessor architectures by having various components running on different computer nodes.
- Tap into ADCIRC model for additional data needs.
- Would be useful to see the resolution and compartments for other two regions (saw Pontchartrain/Barataria at meeting).
- Care is needed in planning how water control structures will be modeled. Historical management is not appropriate in modeling scenarios which would reasonably result in altered structure management. For example, locks would likely be operated differently under varying seal level rise and salinity conditions.
- If feasible, structure management rules should be defined in input rather than hard coded. This will reduce the need for basin-specific program source codes.
- Although the Mississippi River will not be explicitly modeled, a river flow balance and stage-discharge relationships may be needed at a minimum to provide an approximation of the interaction of multiple diversions operating within a single scenario.

Modeling Components

- Need to start talking about what storms to run.
- Need to address precipitation and runoff in CLARA.
- Suggest that the team streamline and standardize the use of wave models across model components.
- Consider an adaptive grid for compartments to capture marsh erosion and accretion on the landscape (although this may not be feasible).

- Need to figure out details to handle marsh erosion and sediment redistribution (e.g., resuspension, how sediment will move into the marsh vs. move off-shore, etc.).
- Consider including sediment loss (e.g., marsh erosion) from storms to compliment modeling of sediment addition from storms. This could be done with a look up table.
- Just because HSIs predict favorable habitat suitability, it doesn't mean the species are there (especially if they are mobile); therefore, the team should caveat these analyses with "even though habitat suitability for species X is great, they are not necessarily there…"; another option is to capture the potential range of abundance of each species given a particular HSI value (e.g., at HSI value of X, there could be from Y to Z individuals/biomass there).
- In modeling system operations, try to engage operators in the modeling process to see how they learn and adapt to new conditions and constraints, and to provide a reality check on how the actions of operators are described in the model. Modeling present-day operations may be a good way to do that, if it has not already been done.
- In habitat modeling, it might be useful to include other variables known or suspected to affect populations as "nuisance variables" (i.e., variables that influence the population but are not directly of interest because they cannot be controlled).

Model Calibration and Validation

- Need to identify and factor in the end use of the model outputs; calibrate specifically for these end uses.
- Maintain communication between HSI development/selection task members and model developers to ensure that final HSIs can be reliably modeled.
- Long-term validation is important, so it would be useful to split data across sites rather than temporally when possible. Many long-term datasets can be valuable, even if the data are not available coast wide or if other simplifications have to be made. Beware of spatial autocorrelation; make sure the validation data are independent.
- Compare outputs from grid and compartment models whenever possible as a way to spot check outputs.
- Be aware of potential false trends in input datasets (e.g., there might be big drought or storm in year 1 or toward the end – so it's not really a "trend").
- Be diligent with QA/QC and managing input datasets.
- Set up a protocol whereby modelers are notified of relevant data corrections in a timely manner.
- Maintain a model version control system which includes model code and datasets.

Addressing Uncertainties

- Focus on model uncertainties that have the largest effects or are most important for decision making.
- Be mindful of structural uncertainty in the models (i.e., incomplete models or inaccurate depiction of the system), particularly related to biological models.
- Be mindful of uncertainty within each model and propagation of uncertainty through the models; testing scenarios and parameters is a good way to get at this.

Other Issues

- Start expectation management now so end users do not expect this effort to provide more than it can.
- Let end users know that models are not replacement for good professional judgment; people still need to review and consider the model outputs.
- Be careful of mission creep – can get too many requests, etc. and then get to a point where the team is beyond the possibility of doing everything.
- Prioritize on-going data collection needs and identify those that may be collected in the short-term to improve models.
- Be very careful of conclusions drawn from MODIS and Landsat turbidity estimates.

Meeting 2 – July 1, 2014

2017 Coastal Master Plan Modeling Approach

- There is some really great work here, and with further development, some of it should be peer-reviewed and published in journals.
- There are still lots of unknowns, and the team may have to change tracks as they go; it will be difficult for the modeling team to determine the best strategy until they try some different approaches.
- If the team tries something new and realizes it may not be possible in the timeframe that is available, they should quickly identify an alternative approach (Plan B) and they should not be discouraged to go with Plan B.
- There are concerns with the schedule, considering how many decisions still need to be made (e.g., software and operating system).
- The team needs to evaluate which output parameters are likely to be the key decision variables.

Integrated Compartment Model (ICM) Vision and Path Forward

- A decision needs to be made regarding the operating system to be used; the options are Windows and Linux. Most of the model components could be run on either platform, except for the GIS. The two alternatives would be to continue with ESRI (Windows only; and only on licensed computers) or switch to GDAL (runs on Linux). There are pros and cons to each alternative.
- If the team ever wants to parallelize the code, or take advantage of HPC (i.e., high performance computing) they need to use Linux. For that reason, if the team goes with ESRI, it will be limited to runs on Windows machines, which could limit production run times (i.e., have to use licensed computers). Alternatively, if they use GDAL, they could use cloud computing services, which would allow a large number of model instances to run simultaneously; or use HPC for parallel runs.
- The PM-TAC's advice is to let computational efficiency drive decisions. There was some talk that keeping the system in Windows made it easier for others to use it, but whoever uses the model later would just need to get the input, software, etc.
- If there are enough resources, the PM-TAC recommends that the team consider pursuing a dual track development, wherein they would explore recoding for GDAL/Linux while also moving forward with ESRI/Windows. They could then make the software and platform selection after sufficient experience clarified the better choice.

Uncertainties and Uncertainty Propagation

- Remember this is modeling, and although this is the best we can do right now, the team needs to make sure the public and end users know that outcomes are not 100% certain, as there are uncertainties propagating through the model layers. This needs to be communicated clearly, in terms of limitations/uncertainty. One option is to provide outcomes in terms of high/med/low or %-chance, without giving explicit values; another option is to put confidence bounds on every outcome to show how each project might affect them.
- Due to averaging, sometimes outputs farther down the chain of models may have less uncertainty than some of the initial/primary parameters. However, it is more common for model uncertainty to be greater in models that are dependent on other models (e.g.,

modeled salinity is typically more uncertain than the hydrologic model that drives the salinity transport model). Combining multiple models is a non-linear process.
- Inaccuracies of input datasets (e.g., LIDAR) may propagate (and may grow) throughout the model results, especially over model runs of 50 years.
- Another source of uncertainty lies in meteorological and source data. Even if the models represent the processes exactly, the future is still unknown in terms of the source terms (e.g., river discharge and loads) and forcing (e.g., wind and wave time series, temperature).
- Model structural uncertainty is more difficult to estimate, but should at least be acknowledged.
- Specific recommendations will depend on how the various outputs will be used. The PM-TAC may have more to add regarding uncertainty as the models develop. Definitions of ecological performance measures must take into account model calibration, performance, and uncertainty so that the credibility and uncertainty of model-calculated performance measures are within an acceptable range.

Challenges of Calibrating an Integrated Model
- A model with many layers might need adjustments and rerunning of various layers along the way.
- Could fine tune the models too much and create the appearance that everything fits well. However, over-calibration can lead to the masking of models' inability to describe the physics, and the models may not fit other datasets.

Storm-related Sediment Deposition
- Use of Tweel and Turner, 2014:
    - Use deposition data from Tweel and Turner to calibrate spatial patterns of sediment mobilization/transport from the model.
    - Including crude storm inputs from Tweel and Turner regressions would likely improve current approach (storm inputs are not incorporated in current modeling).
    - Because this is the only available source for storm deposition calibration, it would be prudent to carefully review the study's details, including all calculations, before it is incorporated into model calibration.
    - A lookup table based on the Tweel and Turner regressions could be Plan B in terms of a simpler approach for storm sedimentation; however, this would not address issues related to the limited scope of using a single data set for calibration and validation.
- It was noted that the sediment model being used (i.e., based on ECOMSed) seems to produce large suspended concentrations during times of extreme energy.
    - Can further calibration of the model based on historic storm inputs improve the current issues with over estimation of sedimentation during storm periods?
    - Other sediment entrainment functions could be used to enhance storm sediment dynamics within the model; some of these attempt to limit resuspension during energetic periods. There may be limited data available to parameterize these dynamics, but it seems worth looking into.

From the Breakout Groups:

Marsh Edge Erosion

- It was noted that attempts to correlate observed marsh edge retreat (i.e., from aerial photographs over decadal timescales) to wave forcing had so far been inconclusive. Work to-date is commendable and presents an intriguing dataset, but further work is advised.
- Previous efforts have included a range of approaches (e.g., binning wave power, splitting into groups of high, medium and low erosion). We recommend looking at other ways to tease out potential relationship(s) that might be useful in the ICM, including:
    - Keep looking for outliers and carefully evaluate their effects on analysis
    - Take care in statistical analysis of trends which may be affected by artifacts from data classification.
    - Base the trends on storm power
    - Reduce to fewer vegetation types (i.e., use habitat types)
    - Analyze data based on elevation
    - Analyze data by region
    - Do initial analysis on a subset of the data (e.g., 1 year in one region) to see if any trends are more obvious
    - Remove locations where the fetch is unlimited.
    - Try looking only at erosion of first period compared to erosion of third period; doing this creates two independent measures paired by site.
- Maybe the finding is "retreat rate measured at these spatial and temporal scales does not scale directly with modeled wave power…"
- Fitting a wave power / marsh retreat relationship over such large spatial scales, but with pixel-resolution, may produce difficulties. The lack of correlation in marsh retreat and estimated wave power could come from a number of sources. Local effects (e.g., vegetation, subsidence) could be important (i.e., though the PI said they already tried to control for those factors). The wave estimates could be inaccurate enough to contaminate the correlation. The wind field could be inaccurate over the spatial and temporal scales used; those inaccuracies propagate into the wave field.
- Alternatively, perhaps shoreline erosion exhibits a threshold effect with a very low threshold level[3]. If this is the case, the threshold might be more apparent if wave power were plotted on a log scale.
- Search for factors other than wave power which may result in erosion
- Plan B could be to treat marsh edge erosion as an uncertainty with high and low plausible range.

Barrier Islands

- The methodologies being used for evaluating barrier island retreat and marsh edge erosion are different – similar questions seemingly are being examined with two very different

[3] Dr. RE Turner recently co-authored a paper on shoreline erosion which claims to identify a threshold effect for oil impacts (McClenachan, G., Turner, R. E., and Tweel, A. W. (2013). Effects of oil on the rate and trajectory of Louisiana marsh shoreline erosion. *Environ. Res. Lett.*, *8*(044030), p. 8, http://iopscience.iop.org/1748-9326/8/4/044030).

approaches. It's not clear if one method is 'better' or 'more productive,' but it may be useful to consider pros and cons of each.

- There seems to be a disconnect between monthly average wave characteristics (e.g., height and period) for barrier island modeling and the processes driving barrier island shore changes, which may be more tied to the extreme storm values and may not be reflected in the monthly averages.
- The approach taken seems to capture spatial variability by using models reasonably well resolved in space (SWAN), but the use of monthly averages does not capture the energetic times that may be most important for transport.

CLARA Parametric Uncertainty

- Fragility of levees is difficult to assign, as there is not a strong scientific basis for developing probability of failure.
- Encourage CLARA team to be more conservative.
- The TAC would benefit from a featured presentation on CLARA at the next face-to-face meeting.

HSIs and EwE

- HSIs – some based on data, others based on expert opinion. There are issues with using data - by fitting a polynomial equation, the data tails can get distorted. There are more modern ways to fit non parametric data and get more realistic fit to the end data.
- EwE – better established, seems to be doing quite well. Need reviewers who can assess 1) does model adequately represent the inputs/system; and 2) does the model adequately represent the fish.
- See additional detailed comments regarding HSIs and EWE below.

<u>PM-TAC Comments:</u>

Future Scenarios (Subtask 4.7)

- CASCADE uses only a few downscaled IPCC scenarios to evaluate potential climate effects on SF Bay.
- It would be worth using off-the-shelf downscaled data for the master plan, since the work has already been done.
- Take outcomes of future scenarios with a grain of salt; we do not have a crystal ball.
- Clarify the difference between sensitivity analyses and future scenarios. Use the sensitivity analyses to reduce the number of scenarios that are ultimately run.
- Need to make sure modelers are in the discussion regarding which model metrics to be used in sensitivity analysis/future scenarios; they will know if a desired model output is an appropriate metric to be used

Upcoming Webinars, Meeting

- Including a late afternoon meeting and a dinner for the PM-TAC prior to the next scheduled meeting may be useful for the PM-TAC to prepare for the meeting. This would allow time for more reflection and discussion within the PM-TAC, and it would improve input from the PM-TAC at the meeting with the modeling team. Given the timing of travel, this should not add substantial time to the PM-TAC's commitment (i.e., arriving slightly earlier the day before the scheduled meeting rather than in the evening).

- Online access to the more in-depth reports would be useful in advance of PM-TAC engagements, even if reports are not finalized.
- The reports would not be provided for 'review' but for context to aid in discussions.
- TENTATIVE Webinar Topics – Sept 22
    - Experimental design for scenario sensitivity analyses
    - Storm selection for landscape
- TENTATIVE Meeting Topics - Dec 10
    - How to capture uncertainty
    - Propagation of uncertainty
    - Presentation on CLARA in general with focus on Parametric Uncertainty
    - Using the Planning Tool to sort through many model runs to look at uncertainty associated with key decision outcomes

Specific Comments on Ecological Models:

General

Because these models can be run quickly and encompass high uncertainty, it would be a good idea to run them many times for each set of inputs (e.g., T, S, marsh configuration) produced by the suite of linked models. This way the uncertainties can be propagated to the final output. This should include structural uncertainty – e.g., if experts A and B disagree about the shape of and HSI for a particular species, both should be used as alternatives. The modelers should consider how these sources of uncertainty will be propagated.

The following comment may be moot. The largest uncertainty more broadly may be in the climate projections themselves, but this should not be part of the uncertain cloud at the end of the model sequence. Rather, the climate effects should be entered as discrete scenarios, so that the accumulation of uncertainty around certain projections (e.g., shrimp catches), can be made clear and the differences among scenarios can be accompanied by their respective uncertainties.

Since the main point is to try to figure out the impacts of the projects under alternative futures, the models should be run to discover the differences between project and no-project alternatives, which implies paired comparisons in which the actual parameters are varied the same for each alternative. Then the uncertainty will be that due to differences rather than the individual projections.

We did not hear much about how these models would be linked to the larger suite of models. It would be worth considering the way model output will be passed to the ecological modelers.

We also did not hear much about specific technical reviews of the individual models.

HSIs

The Habitat Suitability Index models include some based on expert opinion that generally take rather simple forms, reflecting the lack of resolution beyond rather simple relationships. In addition to the propagation of structural uncertainty, it would be helpful if the experts could provide ranges or other estimates of their uncertainty in the parameters of these models to help with the propagation of uncertainty.

For some species the HSIs were developed by fitting statistical models to CPUE data. The models presented included polynomial regressions:

ln (CPUE + 1) ~ f (salinity, temperature)
where CPUE is catch per unit effort, and the function f is quadratic in both temperature and salinity but also includes interaction terms that bring it to a fourth-order polynomial.

This model almost certainly distorts whatever relationship underlies the response of CPUE to salinity and temperature, and should be replaced by a model that uses modern statistical approaches. The first problem is the use of log (CPUE+1), which is a rather crude method of log-transforming data having some zeros. This is appropriate only if the CPUE is in counts (e.g., fish per trawl), and in any case adding a number like this distorts the relationship. A better approach would be to use a zero-inflated lognormal, or a zero-inflated negative binomial if the CPUE is reported in catch per trawl (i.e., in numbers of fish).

The second problem is the use of polynomials to fit the data. Polynomials are useful mainly in situations where the underlying response can be shown or reasonably assumed to be second-order, in which case the parameters can usually be interpreted to represent some phenomenon. That is not the case here, and there is no underlying theory that says what the relationship should be. Therefore a curved response surface should be fitted that incorporates no previous knowledge about its shape. Several techniques are available to do that, of which a common one is generalized additive models (GAMs) with locally weighted regression smoothers.

EwE
These models are well established in the literature and in common use, and the modeling team has experience in developing the models and in working with the original authors of the models. This means that the concepts and principles in these models rest on firmer ground than those in the HSIs. However, the specific application needs a review by a handful of experts in the biota being modeled, as well as somebody with expertise in EwE (Howard Townsend?).

The earlier comment about structural uncertainty is particularly relevant here. What happens if a key species is removed, another added, or a species group split? As with the expert-derived HSIs, substantial uncertainty arising from these alternatives should be retained and propagated along with the parameter uncertainties.

Meeting 3 – December 10, 2014

In addition to specific input that was provided on the "Future Scenarios" material in the morning session, the following summary comments were given by PM-TAC members. First, all members appreciated the read-ahead materials, review questions to guide discussion, and dinner in advance of the meeting. We would like to continue incorporating this approach into future meetings. In addition, we all agreed that the modeling team is exceptionally well qualified, and was assembled with a good mix of expertise and disciplines. Coordination among team members is potentially problematic in this and other multidisciplinary projects. However, in our meetings the modeling team appears to be well-informed, well-coordinated, and making rapid progress.

Individual Summary Comments

- Met with Kristy on the EwE report; went through three sets of reviewer comments and none of the comments point out fatal flaws. There is no 'other tool' that can be used in its place except specifically developed tools. Human element is very important, in terms of fishery element changes over time, inclusion of fishing effort, etc. Opportunities are out there, but time seems limited. Timing may be off for EwE model to get integrated with the ICM. Schedule seems too ambitious.
- Schedule seems optimistic, and if the June 1 deadline is a hard deadline, it may be necessary to consider dropping the uncertainty analysis. Likely that something can go wrong and delay things. Need to have contingency plan in place upfront.
- Take advantage of the sensitivity/uncertainty runs for the future scenarios not only for developing future scenarios but just to learn about model sensitivity. Use target diagrams for calibration and validation (see reference suggestions for target diagrams below). Final project assessments might also need a review of the storm suite and storm affects, because effect of storms can be highly impactful on projects.
- This is a monumental task; the team is making good progress. Even if not explicitly including human population in the model, need to consider it because of strong ties between people and natural resources. And point out explicitly that literature inclusion had to stop by date X and that new literature will be included in the future MP iterations.
- Would be good to have calibration contingency plan. Storms – given so many variables and the fact that the group will not be able to please everyone, the team just needs to develop a defensible method for applying storms. Consider storm impacts on project performance.

Target Diagram Papers - papers that explain target diagrams and provide examples of using them to evaluate model – data agreement, and model sensitivity.

- Jolliff et al. (2009): the paper that you would cite for target diagrams. It explains them, and has some good examples.
- Bever et al. (2013): Target diagrams used to evaluate methods of estimating hypoxic volumes from numerical models.
- Friedrichs et al. (2009): Compares primary productivity estimates from 30 different models.
- Hofmann et al. (2011): Compares monthly averaged SST between models with low-resolution and high-resolution forcing fields (the paper focuses on modeling carbon in coastal areas).

Meeting 4 – March 20, 2015

Outreach Presentations

- *Did the **modeling update presentations** [from the outreach meeting; day 1 afternoon] raise any thoughts or questions?*
    - The presentations on Thursday included a good level of detail for the intended audience; if anything, they could have been slightly more technical; the audience (e.g., NGOs, local consultants, CPRA staff who are not directly involved in the Master Plan) appeared to appreciate the material and presentations.
    - It is useful to share information (e.g., model details, calibration, etc.) even before you have results to get people educated and onboard/supportive early in the process and to let them weigh in on issues before things are fully developed – even if it is just for one basin.

- *Does the PM-TAC have any suggestions for improving how the information is conveyed?*
    - It could be useful to post presentations online to reach the broadest audience.
    - Showing one or two components of the ICM in detail would have been useful.
    - There were not many academics at the presentations. If a goal is to publicize the modeling effort to the academic researchers then perhaps it would be useful to pursue department seminars at interested university departments, or organize conference workshops or presentations (i.e., reach out to all personal/professional research contacts directly). This could be done with an overview presentation of all the components followed by concurrent sessions or poster session with more details on specific components. Poster sessions would allow people to interact directly with the subject matter experts.
    - Researchers/academics may be drawn in if they had the opportunity to hear research needs, have their data used in the modeling, be able to be involved and publish papers, and have the modeling effort cite their research.

ICM Calibration/Validation

- *Is the **ICM calibration/validation strategy** appropriate for a long-term, coast wide planning* model?
    - The approach of calibration and validation is a standard modeling approach and is appropriate in the relative short-term (i.e., coming decades), but it becomes less certain in the longer term future. It is important to also note that there are constraints in terms of available data, both for calibration and validation. It would be useful to explicitly identify and document which parameters are based on empirical data/direct measurements, literature values, model calibrated, etc.

- *Are there other calibration considerations we can/should pursue?*

<u>Approaches to Calibration</u>

- "Tuning" of the model can cause concerns, particularly for the unknown conditions of the future. The challenge is in calibrating and validating to "present day" conditions, or recent historical past conditions; yet the model needs to be applied to evaluate conditions over a 50-year predictive timeframe. It will be problematic if the model is "tuned" to present-day or recent historical past, but needs to be applied to 50-year future window.

- The modeling team should identify the best calibration and validation strategy, given data limitations. One alternative to minimize tuning would be to select calibration and validation periods that are very different (e.g., wet during calibration, dry during validation); a second alternative would be to calibrate to as wide a range of conditions as possible.
- Identify parameters that would be most useful for evaluation across such a large-scale model; be mindful of calibrating to the values that "matter" for other model components. For example, if "days of inundation" is a critical parameter for the habitat model component, then calibrate to "days of inundation".
- Need to look at time step to be used in model (e.g., average annual stage) for calibration; perhaps use longer-term averages (e.g., monthly salinity) for comparing model-to-observed records.
- For the sediment module: are long-term averages appropriate, particularly if storm impacts are desired? Is data limitation too great for consideration of storm impacts?
- For the morphology module: it is important to use spatially aggregated statistics (e.g., basin-wide) for long term statistics; if you are looking at specific locations at the pixel-level, error will be much greater.
- Spend time calibrating the things that are really important for the outcomes of the modeling. For example, when the observed depth goes below 10 cm, cap the model drawdown and consider it calibrated. Do nottry to capture the extreme low stages because they might not be real.

<u>Using Metrics/Assessing Performance</u>

- Need to define model performance measures based on anticipated uses. An explicit list for each model module of the outputs that are anticipated to be used for decisions or other model inputs should be created by the modeling team and reviewed by users. This list will assist in the definition of calibration performance measures.
- Drop R^2 altogether and use bias and standard deviation of residual error instead. R^2 is a useful PM in regression analysis because regression models have zero bias, but is misleading in dynamic model calibration and performance evaluation.
- You should also consider reporting Nash-Sutcliffe calibration values.
- Keep the statistics in original units as much as possible. For example, do notdivide RMSE by normalized output; gives higher weight when observed values are small (e.g., shallow areas).
- Do a full model skill assessment and look at overall model metrics in target diagrams, which take into account bias and variance in the error. These can help synthesize all outcomes into one figure (see below for specific citations).
- Explore the use of non-parametric methods.
- Use scatter plots of model values vs. observed values. If the error is smaller at lower stages and greater at higher stages, it tells you one thing; if the error stays the same, then it tells you it's another kind of error.
- You might also try plots analogous to double mass curves with integral/sum of observed values on x-axis and integral/sum of corresponding model values on the y-axis. For a perfect model fit, this should fall on a 1:1 line. A value of the double mass curve is that small errors in timing result in only small excursions on the graph.
- Examine residuals on model vs. observed and see if they have a normal distribution.
- A map showing the spatial distribution of errors could be useful to identify where the model is performing well and where it isn't (e.g., inland vs. closer to the gulf).
- Identify the steps of calibration and considering interactions across model components (i.e., where do you need to check back for feedback between parameters?).

- Do not neglect calibrating "slow changing" variables (e.g., vegetation change, morphology change).
- Some proposed target metrics in Table 2 (e.g., 10% of historic land change rates, 20% RMSE long-term accretion) seem overly ambitious.
- Could be overly optimistic to get salinity within 20-30% RMSE.

General issues

- Need to explicitly lay out how the experts interpret the output and how it can be used for decision making. It is up to the modeling team experts to interpret the outcomes and to understand and communicate the model limitations.
- In general, provide more details on validation, in addition to calibration methods.
- Use model output to check observed data; it will be easy to identify outlier data points that should be removed from the calibration effort; use robust statistics such as quantiles, or trimmed means to reduce the influence of erroneous or extreme data points; you want the model to reflect long-term trends and averages; it doesn't have to pick up every single bump and dip.
- Consider correcting the bias before generating final outputs. Some model uses may be very sensitive to model bias error. One possible example would be counting days exceeding a criterion for salinity. In application, it may be necessary to correct for model bias before interpreting model projections.
- Is audience for Table 2 a wider group? If so, more details and documentation may be needed.

Suggested articles for model skill metrics:

Friedrichs et al. (2009). Assessing the uncertainties of model estimates of primary productivity in the tropical Pacific Ocean, *Journal of Marine Science*, *76*,113-133. Uses Target diagrams for bias and (I think RMSE).

Hetland, R. (2006). Event-driven model skill assessment. *Ocean Modelling*, *11*, 214-223. http://pong.tamu.edu/~rob/pubs/hetland_skill_ocemod.pdf.

Warner et al. (2005). Numerical modeling of an estuary: A comprehensive skill assessment. *Journal of Geophysical Research*, *110*, C05001, doi:10.1029/2004JC002691.

ICM Uncertainty Analysis

- *Is the rationale and approach of the [draft]* ***uncertainty analysis*** *clear and appropriate for this type of modeling effort?*
    - The uncertainty analysis (UA) needs to be more explicitly clarified; as it stands the UA is more of a sensitivity analysis. It makes sense to evaluate model uncertainty, and to compare this variability to sensitivity analysis based on parameters, and also to the calibration error. However, a true UA requires calibrating all the parameters, changing individual parameter values sequentially, and recalibrating everything. What the team has been talking about is a robust sensitivity analysis. A full UA is not feasible for this effort, but quantification of uncertainty under present conditions will be useful.
    - It is only possible to evaluate model uncertainty around the present/known conditions. The range of future model output reflects more than uncertainty; it is a range of output generated by an uncertain model run with different scenarios. Evaluating all of the various components of uncertainty (e.g., model, future scenarios, etc.) is very complex and difficult; it is rare to do this sort of analysis (i.e., from either an academic or applied setting).
    - Need to clarify the goals of the UA, and explain why you are doing each piece of the analysis.

  - Start with identifying which questions you are trying to answer. Because you cannot do it all; focus on key questions and design experiments around these questions. Is the first question: what is our confidence in the outcome of a project, and our estimates as to the range of outcomes of a project?
  - The second fundamental consideration is computational constraints – how many runs can you feasibly complete and analyze with the time and computational resources on hand?

- *Are there any suggestions to streamline or otherwise improve the suggested methodology?*
  - Need to identify how to select the projects which will be used in the UA. Perhaps consider the most expensive projects; alternatively consider UA for different types of projects, or for different regions.
  - Consider starting all projects in 2006, so that you have that start up time with and without projects to evaluate project effects (i.e., a comparison over 10 year period with recent data rather than with out and without projects in the future).
  - Consider modifying intermediate values (e.g., water level or sediment accumulation) as they are passed from one model to another, rather than the model parameters themselves (like roughness or resuspension coefficient). This will prevent having to recalibrate the models for each alternative set of intermediate values, (e.g., what if water level is actually 20 cm higher than we think?). Such analyses should be done for the entire set of runs for the scenarios vs. no/with project alternatives. Save those runs for use with individual projects to assess reliability of decisions about those projects.

QA/QC for Model Output

- *Does the TAC have suggestions regarding methods for enhancing and/or streamlining the **QA/QC** process for hundreds of production runs (models include: ICM, EwE, ADCIRC, and CLARA)?*
  - Documentation of model edits and changes over time must be captured clearly and accurately. Keep documentation of exactly what went into each model run, so that you can check exact inputs for every single output.
  - Like EwE, need to document that the ICM has done simplified runs (i.e., shut off inputs and ensure model runs) – document this. Check mass balance and show that it is close to 0.
  - Focus on the basics (e.g., land area/creation). Need to determine the most critical outputs, time steps and spatial scales. Then can look at other details as needed.
  - Automate things as much as possible. Set triggers to throw flags if values change more than a certain amount; to reduce noise, set a threshold, under which "project effect" does not get counted as change.
  - Animate things at much as possible. Our eyes are very good at quickly picking out data anomalies. Ensure multiple people provide a redundant review of output.
  - Use target diagrams; this can help you view a lot of the output at once.
  - For vegetation, strategically select regions at transition zones so the evaluation of modeled changes does not get masked in a basin-wide summation.

Meeting 5 - September 23, 2015

ICM calibration and validation: *Are the results of the ICM calibration and validation clearly presented in terms of the different model outputs, and are the model limitations due to available data and other constraints sufficiently identified? How could the process for calibration and validation of the integrated model be improved in future planning processes?*

- It is clear that substantial progress has been made in terms of overall model development, calibration, and validation.
- The level of documentation within the model development is impressive, and it is important that documentation is being done simultaneously with model development rather than left to the end of the process.
- The calibration data that are available have inherent limitations and problems (i.e., as is always true), and you are trying to fit the model to these potentially problematic data. In other words, it is possible that in some instances the model output may reflect the system state as accurately as the available data. We recommend that you evaluate data "reliability" and incorporate this into the presentations of model calibration and assessments of the model's suitability for forecasting. Specifically, we submit the following recommendations be implemented if feasible within constraints of project schedule and resources:
    - Be aware of the potential mismatch of the scale of available calibration data and the scale of the modeled processes and model output. For example, the accretion comes from a single point/core, but it is compared to modeled output for a large area, which encompasses a wide range of accretion rates.
    - Some calibration data are not representative of the entire spatial and temporal range of the model coverage, due to field limitations and spatial considerations, etc. In such cases you might compare the observed values to a subset of the model estimates that has been sampled with similar constraints.
    - It may be useful to restrict the comparison of modeled TSS values to those that come from the time periods or locations when and where actual data are collected.
    - Comparisons of point-based data to model output may indicate less model skill than what is achieved. Try to establish trend lines for the data and for the model output. Especially with respect to TSS, if the model can be shown to be within factors of the data and following the trend, model skill is indicated.
    - Also consider spatial issues in terms of comparing model output, which is probably depth averaged, to water sampling data that may have been collected at the surface, or may not have included samples near the bed.
    - In comparing 30-day averages from model, it would be useful to show a range of modeled output in addition to the average since the model output is usually compared to a single data point rather than an average value.
- Incorporating confidence intervals into data presentations would be useful in depicting the fit of the model to data. Add error bars where possible or some other indication of variation. As above, this should address both temporal and spatial variation in data.
- At the meeting, you primarily presented output from the conclusion of the model calibration runs, but you did not present much data throughout the course of the model runs to illustrate temporal dynamics. It would useful to see temporal dynamics from the model vs. observed data, as well as the endpoints. These comparisons would likely give the reader more confidence in the model predictions.
- It would be useful to present target diagrams, which indicate both bias and variability on the same graph, and can synthesize within one figure the model skill from multiple model

runs and / or model locations. These could be combined with presentations of various skill metrics to give an overall picture of model suitability for forecasting.

- It would be useful to evaluate model fit for land area changes across different types of wetlands/habitats since land building is a critical component of project evaluation and variation across wetland types could affect project assessment.
- The HSIs have to be grounded in data (e.g., population observations) correlated with environmental conditions. Be sure to acknowledge that populations may fluctuate spatially and temporally for reasons not captured by these covariates. This also suggests alternative fitting techniques, perhaps including upper percentiles or other methods that emphasize what population size is possible under a given set of conditions. Whether you can do that depends on how much data you have.
- The TSS modeling results seem problematic. PM-TAC discussed factors that may affect the lack of fit for TSS including sampling issues such as a mismatch in the time periods and sampling locations for modeled vs. field data collection. Phytoplankton also may be a bias in lower TSS levels in eutrophic areas.
- While the model provided a reasonably good fit for accretion data when compared across all locations, the delta locations could be driving the model fit. Without these few data points, which had the highest values of accretion, it looked like data were more of a cloud, so that the model appeared to have a much lower skill for replicating accretion rates the non-deltaic locations. The model fit for non-deltaic locations should be further evaluated.
- Given the importance of suspended sediment and deposition to the modeling, it could be useful to have continuous monitoring stations for suspended sediment and erosion / deposition in the future. The model might be useful for helping to identify suitable locations and monitoring design for this sort of data collection.
- It is important that you are aware of error propagation, even though it has not been directly addressed; it will be necessary to systematically address error propagation across the model components and within the model workflow in the near future.
- Clearly identify that error is inherently likely to increase as you move through the models. For example, models are likely to fit better for hydrology than "downstream" models in the ICM.
- For future modeling improvements (i.e., beyond the current Master Plan), we have the following additional recommendations.
    - It would be useful for future model development to highlight how you would design future data collection based on what's been learned from the current model calibration and development.
    - Some model components are not as data rich as others. We recommend that you identify and critical data gaps for future model development. In addition, it would be useful to identify how future models could be improved in terms of modeling approaches, separate from improved data availability.

Environmental Scenarios: *Are the potential ranges of values for environmental variables considered for the scenarios reasonable? Do the modeling results presented support the selection of values to be used in the three environmental scenarios?*

- The proposed future scenarios seem reasonable, but the framework for the environmental scenarios was more clearly presented at the meeting than in the read-ahead materials. Incorporate the material and approach for the environmental scenarios that was presented at the meeting into future written descriptions of the scenarios.
- Be sure to state upfront why you are running the different scenarios so that the context and role of the scenarios are clear.

- In the written materials describing the models, you need to more clearly justify why only three scenarios will be used.
- Better justification is needed to support scenarios for sea-level rise, as they seem tilted towards the medium to high end of likely predictions by 2050 (although justification may be in the Appendix – these files were not accessible on-line after the meeting).
- Identify the importance of spatial variability more clearly in relation to the scenarios. It was not clear if the scenarios were designed to address potential spatial variation in processes and responses. More explicitly identify how the overall effectiveness of individual projects would be evaluated if they perform very differently under different scenarios.
- Deleting and combining some of the axes of future change (e.g., nutrients) is helpful.
- It may be useful for future model runs to combine precipitation and evaporation/transpiration into a single factor, in order to reduce the number of variables for the scenarios.
- In future model revisions, you could correlate some of the model components based on likely future climate change scenarios (e.g., carbon scenarios and resulting impacts). This may result in better connection to Mississippi River watershed inputs and local conditions

ICM QA/QC process: *Does the QA/QC process being used to track the ICM model runs seem reasonable and sufficiently thorough to support the use of model results? Are there improvements that could be implemented in the near term or considered in future?*

- The QA/QC approach is very thorough and methodical.
- The proposed graphs are very useful and should be maintained; it is powerful to review the outputs visually.
- If possible, identify additional, simpler QA/QC checks that could be done automatically (e.g., producing simple tables of modeled vs. actual differences or other indices of relative changes that could be incorporated into the model runs).
- Identify more specifically what would raise a flag for the QA/QC questions that were presented in the excel spreadsheet.
- It would be useful to produce warning messages when a problem occurs during a run and to have these printed for future reference. For example, provide a warning in the run log giving the time and location where concentrations fall far outside those observed in calibration. This will not necessarily indicate an error, but will expedite QA/QC of individual model runs.
- It is important to document all changes to the computer code for different runs.
- Be sure to train whoever does the QA/QC very well. In training for QA/QC, run some cases that have problems to be sure that potential problems are identified. Involve as many duplicate sets of eyes in QA/QC as possible.
- It would be useful to identify upfront how much time will be necessary to do this level of QA/QC, especially relative to other aspects of the modeling effort.
- It would be useful to put together a simple flow chart/summary that shows steps for QA/QC. Some steps are implicit in the document, but need more explicit description. For example, the criteria for QA/QC: some steps/components may be evaluated based on “best professional judgment”, while some may have more specific criteria, based on specific data variations. It would be useful to identify these differences up front.
- For future modeling efforts, it would be worthwhile to step back and consider if an important parameter or key process has been left out. The current QA/QC methods obviously focus on the current model framework and it would not necessarily identify missing parameters or processes. It would be useful to generate a list of additional parameters and processes to consider for the next modeling round.

Project Implementation Within Models: *Are there are any key shortcomings of the ways in which different project types are represented in the modeling that potentially limit the utility of the results? Which of these are especially important for CPRA to be aware of as the results are used to develop alternatives?*

- If just one scenario is used to evaluate projects, be aware of potential bias that this scenario may have towards a particular marsh types or restoration approach (e.g., would evaluation of projects using a scenario with high rates of SLR favor shoreline erosion projects over other project types, or salt marsh over tidal freshwater marsh?—this is just a hypothetical example).
- Be aware that the timing/sequencing of project implementation could affect flooding. For example, modeling construction of a levee prior to construction of a stormwater pump could cause unacceptable modeled flooding.
- The models assume that projects will work as designed (e.g., shoreline protection will reduce erosion); however, restoration outcomes may be more variable. Be aware of this assumption in evaluating model outcomes.
- Be sure to mention in model documentation that use of models in planning does not preclude the need for adaptive management as the projects are implemented.

Non-structural Methodology: *Are the key assumptions underlying the methodology for non-structural project development, evaluation, and prioritization clearly defined and defensible, including the use of population and asset growth scenarios to evaluate projects and the evaluation criteria that will be used to prioritize nonstructural projects? Are there ways in which this methodology could be strengthened in the future?*

- The population growth approach and the other non-structural methodology seems well detailed and thought out [although this model component is furthest afield from the expertise of the PM-TAC].
- It would be useful to specifically identify improvements to the 2012 approach.
- Here and elsewhere, an executive summary would be very useful to set up the general issues/key points prior to the details of the main report.
- The documentation for this component is somewhat detailed, but some important material was difficult to find. For example, it was not clear how the cost function was set up. Perhaps including a concise executive summary that outlines the model framework, and then a table of contents that points the reader toward specific sections of the report may help with the ease of access.
- For future model versions, it would be useful to incorporate population migration following storms.

Additional Comments

- Within the model, projects will be implemented at some given year and month, while the model is being run using a historic time series of discharge and weather. The interplay between the timing and location of a project relative to the historic time series of storm tracks may create differences in project outcomes that are an artifact of the model framework rather than an indication of the likely success of particular projects.
- The HSIs generally use polynomial functions to fit data, but polynomials are notorious for wild oscillations at the edges of the data range. Alternative fitting procedures (e.g., GAMs using robust fitting methods) would downplay importance of outliers and would not incorporate assumed data shapes for habitat characteristics vs. suitability.

Meeting 6 - March 15, 2016

Model Output and Communication: *Does the PM-TAC have recommendations regarding the communication of modeling results? What is an appropriate spatial/temporal scale to display the result, e.g., 5 year intervals vs. annual, coast wide vs. regional? Does the PM-TAC have any suggestions to ensure the model assumptions/limitations are understood as model results are displayed?*

Communication and display of results:

- The maps of predicted shifts in vegetation types are very useful output.
- Use the output to highlight the comparisons that you'd like readers/observers to see: there are many potential comparisons within the output (i.e., different scenarios, with and without projects, different time scales, etc.). Use the output to emphasize the comparisons that are of most interest, and keep comparisons as consistent as possible across different output.
- Be aware of different audiences, and simplify output for more general audiences; emphasize the 50-year change, coast wide for more general summaries. The spatial and temporal details are very useful and should be presented for targeted audiences, especially where they identify important shifts, but this detail could overwhelm the message to more general audiences.
- When possible, use common terminology and layout with output (e.g., only the surge and risk teams include the "initial condition" landscape, but this would likely be useful for comparison on other output).
- Overlay the present-day coastline on land change maps over time. The slides from the storm surge analysis were good because they showed present-day coastline and a range of different data on one slide.
- For the high resolution maps, include an inset or some other spatial orientation of the specific location along the coast.
- Choose the type of color maps that are used carefully. "Gradient" color maps are useful for showing magnitudes that span from a low to a high value across a simple trend (e.g., sediment concentration, population size). The simplest gradient map uses a gray scale that goes from white to black, but gradient maps could also shift in color, hue, or brightness. "Divergent" color maps are useful for showing net changes when both high and low values (typically positive or negative) are of interest, but not mid-point (typically zero) values (e.g., tidal water velocity, land loss or gain). In this case, white would indicate no change, blue might indicate a land loss, and green might indicate land gain. See http://matplotlib.org/cmocean/ for examples.
- When possible, use consistent colors across maps/output from different models (e.g., white for no change, blue for erosion, green for deposition on divergent maps; or a consistent color pattern for gradient maps).
- Stress that future conditions are scenarios and not exact predictions; they represent a range of plausible futures based on a semi-quantitative simulation process.
- Use the screen/graph space efficiently to display the regions and points to be shown while minimizing extraneous detail. Some graphs were difficult to see because the screen space was wasted. In particular, enlarge the maps for the barrier islands.
- For future analysis of the ICM, it may be useful to do more sensitivity analyses to see effects of different marsh collapse thresholds.
- For the EwE model, it would be useful to see a more thorough analysis of what is driving the responses that are being observed in the model (e.g., salinity, TKN, or other factors).

Assumptions:

- Highlight what the model includes and does not include; acknowledge model limitations up front (countered with the large number of things the model does include).
- Show model domain upfront. Include a slide upfront showing what the model does/does not include.
- Need to explicitly identify that the ADCIRC storm surge model does not include riverine flooding.
- Clarify how assumptions may affect project evaluation and prioritization (e.g., are there any assumptions that apply to one type of project but not another).

Use of Model Outputs: *Are the uses to which the model outputs are being applied, appropriate given the assumptions or limitations of the modeling approach? Are there any key limitations of the models that should be more explicitly considered as the results are used to help develop the Master Plan?*

- The Planning Tool (PT) seems to be very powerful, but it's also complex. It could be a challenge to narrow down so much complexity into making decisions. Clearly state how the PT will be used, what factors will be considered, etc., so that it is not viewed as a black box.
- Mapping out the PT page would be helpful since so many different components are included.
- More clearly identify how uncertainty will be incorporated into the use of the PT and decisions that it facilitates.
- Project effects are assumed to be additive; identify this in the assumptions of the PT or explain more specifically how interactions/combinations will be analyzed. If they are not additive, identify how you can learn from evaluating a mix of alternative combinations.
- It will be useful to use the PT to identify not just the best projects but also the very poor projects, and how this dichotomy may change over time (i.e., as in the example presented of the river diversion being successful in the short term and having a negative effect in the long term).
- In using the PT, use the analysis of targeted projects to help make prioritization for future projects. Make the best plan now and also learn from it for developing future priorities.
- The modeling team should further evaluate the manner in which storms were implemented in the model (e.g., at specific times and locations) and how this impacts the model results. For example, the model showed a perceptible drop in predicted wetland area in a particular year due to one storm. The team should evaluate whether effects such as this would be significant for assessing the likelihood of success for specific projects and whether any modifications in the approach could address this artifact in future analyses.

Model Improvements: *Does the PM-TAC have recommendations for additional improvements that should be considered prior to the start of alternative model runs?*

- The model improvements that were presented are on target. The PM-TAC doesn't have the knowledge of the model details to suggest additional specific improvements. More general recommendations on model improvement will be provided in the PM-TAC's final report. That said, it is helpful that the modeling team is thinking now about what needs to be improved while their experiences are fresh. Some of the improvements may take some time to develop and it is probably not too early to get started on them. Think broadly: are there model components that are not really informative? Are there better ways to model some aspects of the system? What is not being modeled that could be because of improvements in understanding or in technology?

ICM Uncertainty Analysis: *Does the approach used for the ICM uncertainty analysis seem consistent with previous PM-TAC discussions? Is the methodology appropriately described? Are the results being interpreted appropriately by the team? What level of agreement between individual runs and composite runs should be used to validate the assumption of linearity? Based on the Phase 1 output, are there suggestions to optimize the design of Phase 2?*

- The uncertainty analysis of the model is cutting edge (i.e., very few people have ever done something like this); add a paragraph highlighting this analysis, its benefits and difficulties.
- Use absolute percentiles and drop the use of the log transformation for the distribution or Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). Clarify that MRSE is the same as RMSE; if they are equivalent, RMSE is the more commonly used term.
- Consider making perturbations based on percent absolute error not mean absolute error. This will be useful with some variables such as salinity or fish biomass in which the variance can be expected to scale with the mean.
- Clarify why RMSE is consistently greater than MAE in Table 1. Consider removing RMSE from the table and analysis.
- Add additional uncertainty runs and adjust TSS and salinity by percentage changes. This way in areas with high turbidity, you would perturb TSS more than in an area of low turbidity; this should provide more meaningful spatial variation in TSS and salinity across regions.
- The combination of the ten model variables into the 4 groups is reasonable.
- Add paragraph clarifying how perturbations are combined and potential links between variables that are more reflective of actual changes in variables that are likely to be tied together.
- Consider revising Question 2 to more clearly identify the comparison of interest, for example: "Is the scale of the model uncertainty greater than the scale of the project outcomes being predicted?" [the question currently is: "Is land area produced by large-scale restoration projects (e.g., sediment diversions or marsh creation) more/less uncertain than land area under FWOA?"] To analyze this, you could run FWOA and run project 1. Compare land area. Then run FWOA with X salinity added and project 1 with X salinity added, and compare the projected land area.
- Clarify uncertainty of future predictions (e.g., addressed with multiple scenarios) vs. uncertainty of the difference with and without projects (e.g., run projects with and without uncertainty and compare the results)
- Look at ecoregions to understand spatial differences in uncertainty and to identify if there any patterns in uncertainty across regions.
- More clearly identify how the uncertainty analysis will be used with the PT. How will this information be given to CPRA and decision makers?
- When projects are picked for Phase 2 of the uncertainty analysis, it will be important to carefully pick the types of projects that will be evaluated to be sure that they give insight into decisions about other projects. Look at enough project types so have enough understanding overall.
- Add more explanation of how the ten "key model parameters" where chosen.
- Make clear that perturbing only one parameter at a time could result in unrealistic values in other parameters.
- Clarify that perturbations are annual but that the model continues to run; e.g., morphology perturbations continue to feedback to the model, but vegetation perturbations do not have a 'memory.'