Peer Review Panel Task 2 Report

for the

Kissimmee Basin Modeling and Operations Study

May 21, 2007

M E M O R A N D U M

TO: **Chris Carlson, SFWMD**

FROM: **Panel Members: Pete Loucks** (Chair)
David Chin
Robert Prucha

DATE: **May 1, 2007**

SUBJECT: **Peer Review Panel Report for Task 2**

Table of Contents: **Page**

Executive Summary

The overall goal of this review is to provide an unbiased expert assessment of the Kissimmee Basin Modeling and Operations Study (KBMOS) that will support the development of improved operational rules for water-control structures within the Kissimmee Basin. More specifically, this peer review assesses the quality and credibility of the science used to develop the OKISS and AFET models, and their applicability to decision-making for operational management of structures in the Kissimmee Basin.

This report summarizes the findings of the review of documents outlined in Task 2 in the Scope of Work (SOW) for the Peer Review of Kissimmee Basin Modeling and Operations Study (KBMOS). In this task, the panel assessed the process of selecting the modeling tools and their suitability for evaluating existing and proposed Kissimmee Basin hydraulic-structure operating criteria. It also assessed the proposed procedure to be used to select the preferred criteria.

Task 2 has three objectives:

A. Assess the process used to select modeling tools
B. Assess the suitability of the selected models and performance measures to evaluate existing and proposed Kissimmee Basin structure operating criteria
C. Assess the Alternative Evaluation System.

The panel's assessments contained in this report are based on documents listed in Table 1, along with the District's responses to the individual and collective comments made by the panel. Summaries of the Panel's comments and the District's responses to those comments are included in the Appendix of this report.

Overall, the panel finds that the KBMOS plan for identifying suitable alternative operating criteria for hydraulic structures in the Kissimmee Basin is sound. Panel findings related to this conclusion are:

- The process of selecting the modeling tools to be used in evaluating the impacts of alternative operating criteria was thorough, transparent, and led to a reasonable outcome.
- The panel agrees that OASIS, MIKE 11, and MIKE SHE/MIKE 11 are suitable and appropriate software codes for use at the screening, formulation, and evaluation stages of the alternative plan selection process. The performance measures to be used in evaluating alternative operating criteria are still being developed (partly in response to the Panel's comments) and include indicators that are not all easily quantified.
- The Alternative Evaluation System is entirely subjective, and will involve multiple stakeholders and interest groups. This subjective approach appears to be appropriate for selecting the best among alternative operational criteria.

In reviewing the details of the study plan, some specific concerns have been identified that should be addressed. It is the opinion of the Panel that addressing these concerns will improve the likelihood of identifying and selecting the best alternative operating criteria. The most significant of these concerns are:

1. Using the base-condition lateral inflows generated by MIKE SHE/MIKE 11 in OKISS and MIKE 11 for alternative screening and formulation needs further justification before alternatives are screened and formulated. Expanding the scope of the models to include as endogenous the lateral inflows by incorporating the water storage on the flood plain basins and its interactions with the river or channel flows seems a preferred way to model the system.
2. Further refinement is needed in the formulation and selection of some performance measures.
3. All performance measures should be taken into account in the screening and formulation process.
4. Utility functions, which are a key component of the alternative evaluation system, are yet to be developed and peer reviewed.
5. The uncertainty in modeling-tool predictions on the alternative evaluation scores should be considered and clearly documented to demonstrate that the preferred alternatives, once identified, are a statistically significant improvement over the base condition and over other alternatives that are not selected for final evaluation.
6. As presently structured, natural resources and flood control objectives are the primary bases for selecting alternatives; with water supply, aquatic plant management, and downstream ecosystems objectives have only secondary role. The adequacy of this approach requires further justification.

By addressing these primary concerns, the likelihood of achieving the project objectives using a scientifically defensible protocol will be significantly improved.

Table 1 Review Documents associated with Task 2 Peer Review Objectives

Objective	Doc #	Document Name	Content
A	1	Kissimmee Basin Assessment Report: Section 5 (Model Evaluation Report)	Summarizes the results of the model evaluation task performed in Phase I to select a fully integrated hydrologic and hydraulic modeling tool for development and evaluation of water management operating plans for the Kissimmee Basin.
A	2	Screening Tool Proof of Concept Test Plan	Describes the process used to select OASIS as a screening tool.
A	3	OASIS KCOL Model Report	Describes the implementation of the OASIS model for the Kissimmee Chain of Lakes
A	4	Screening Tool Proof of Concept Workshop Summary	Documents the evaluation of the Screening Tool (OASIS) in the Proof of Concept Workshop
B & C	5	Draft Base Conditions Summary Report	Provides an overview of the planning effort, the model tools development activities, and definition and description of the base conditions that will be simulated and used as the basis for comparison of alternative plans.
B & C	6	Draft Final Evaluation Performance Measures and Indicators	Presents the background, targets, and technical specifications for the 19 evaluation performance measures and targets developed for the comparison of alternative plans and the evaluation of operating criteria for the Kissimmee Basin
B & C	8	Alternative Evaluation System Technical Design Document	Describes the alternative evaluation system that has been developed to provide a systematic approach to comparing alternative plan results using the evaluation performance measures

Objective A:
Assessment of the Process Used to Select Modeling Tools

To assess the process used to select modeling tools the panel reviewed the following five documents:

1. Model Evaluation Report describes the process of selecting the modeling tools from many potential modeling platforms available. It identifies the general criteria used to evaluate each alternative modeling candidate in the open multi-agency multi-institutional selection process.
2. Screening Toll Proof of Concept Test Plan provides more detail on the selection of the screening model, the selection criteria, and the evaluation tests devised.
3. OASIS KCOL Model Report describes in detail the features of the OASIS model as applied to the Kissimmee Chain of Lakes, and how it can be used interactively with user groups during the screening process.
4. Screening Tool Proof of Concept Demonstration and Design Workshop outlines the results of the workshop that took place for demonstrating the features of the OASIS model as applied to the Kissimmee Basin.
5. Base Conditions Summary Report describes the creation of base hydrologic and land use conditions that will be used for comparisons to alternative management and operation policies identified by the suite of models, and also describes the suite of models used for screening, plan or policy formulation, more detailed plan or policy evaluation, and for performance measure evaluation.

Based on a review of these documents, the panel members provided comments and suggestions to District, and the District provided responses to these comments. Panel comments and District responses are contained in the Appendix to this report.

The panel agrees that OASIS, MIKE 11, and MIKE SHE/MIKE 11 are suitable and appropriate software codes for use at the screening, formulation, and evaluation stages of the alternative plan selection process. After considering the District responses to initial questions and concerns, and discussions during Workshop 1, remaining concerns identified by the Panel are as follows:

- Individual reviewer bias was a concern in interpreting reviewer evaluations of potential model codes. This concern could have been reduced by providing reviewers with a rubric relating scoring to quantifiable program features.

- MIKE SHE/MIKE 11 and MOD-HMS codes were very close in their capabilities relative to the model selection criteria, and the final selection of the MIKE SHE/MIKE 11 code was based on responses to a Request for Additional Information (RAI). Documentation of the RAI responses is minimal and does not adequately capture the rationale for selection of the MIKE SHE/MIKE 11 code for use in the alternative evaluation process.

Since the model selection process has already been completed, these suggestions might be most useful in improving future model selection processes. In the present case, the concerns identified above could be remedied mostly by improved documentation.

Objective B:
Assessment of Modeling Tools and Performance Measures

To assess the suitability of the modeling tools and performance measures for evaluating hydraulic-structure operating criteria, the Panel reviewed three documents that were considered relevant to this objective:

1. Base Conditions Summary Report describes the creation of base hydrologic and land use conditions that will be used for comparisons to alternative management and operation policies identified by the suite of models. This report also describes the suite of models used for initial screening of planning or policy formulations, more detailed plan or policy evaluation, and for performance measure evaluation.
2. Final Evaluation Performance Measures and Indicators defines performance measures and indicators that will be used to predict the extent to which any proposed plan or policy will meet flood control, water supply, aquatic plant management, and natural resource operations objectives.
3. Alternative Evaluation System Technical Design Document describes the concepts behind and the components of the Kissimmee Basin water management system and how that system will be applied.

Based on a review of these documents, panel members provided comments and suggestions to District, and the District provided responses to these comments. Panel comments and District responses are contained in the Appendix to this report. After considering the District responses and discussions during Workshop 1, remaining concerns identified by the Panel are as follows:

- The Panel considers absence of a plan to incorporate uncertainty into the assessment of Kissimmee Basin operating criteria to be a critical defect.

  - Model and data uncertainty must be taken into account in formulating performance measures and utility functions. Ultimately, it is desirable that overall alternative scores be expressed with confidence limits that can be used to show that the final alternatives are a significant improvement over both the base condition and other alternatives that are not included in the final cut. Determination of confidence limits of alternative scores will require that uncertainty in model results be quantified and propagated through the calculation of alternative scores.

  - While the panel recognizes that conducting a detailed uncertainty analysis on fully-integrated basin flow models is difficult, and that a single professionally accepted standard or guideline for conducting such an analysis has never been promulgated by the water-resources profession, uncertainty as well as sensitivity must be addressed. The Panel believes that a clear discussion of the key sources of uncertainty (i.e., input data, calibration targets, conceptual model, and numerical model) should be described at a minimum. In addition, and to the extent possible, prediction uncertainty due to the most sensitive model parameters should be assessed. This could be conducted by using valid ranges for key model parameters (that produce equally valid model calibrations based on defined target ranges) to simulate a range of preferred alternative model outputs. These results could then be used to determine the influence of parameter uncertainty on the AES scoring and

ranking. To avoid the District's concern of long run-times, a single year or two towards the end of the 1965 to 2000 simulation time could be used.

- Whereas most of the evaluation performance measures were found to be suitable for evaluating proposed Kissimmee Basin structure operating criteria, the following improvements are recommended:

  - Some performance measures might be highly correlated and duplicative. It is recommended that a correlation analysis be done to identify highly correlated performance measures and efforts be made to consolidate such measures. The existence of significant correlations between measures could bias the alternative scoring system.

  - Intra-annual variability in flows should be included in the performance measures. It is anticipated that ecological integrity is related to seasonal high water stages occurring in certain months of the year and seasonal lows in certain months of the same year.

  - Seven statistics (percentiles) are proposed to describe the Kissimmee River stage probability distribution. Using a lesser number of percentiles or moments will likely be sufficient and justified by available data. Using no more that four percentiles or moments is recommended.

The panel recognizes that some of these concerns can only be addressed when the KBMOS models have been validated and are operational.

Objective C:

Assessment of the Alternative Evaluation System

To assess the Alternative Evaluation System (AES) the Panel reviewed the following document that specifically pertained to this objective:

Alternative Evaluation System Technical Design Document describes the concepts behind and the components of the Kissimmee basin system and how that system will be applied.

Based on a review of this document, panel members provided comments and suggestions to District, and the District provided responses to these comments. Panel comments and District responses are contained in the Appendix to this report. After considering the District's responses and discussions during Workshop 1, remaining concerns identified by the Panel are as follows:

- The Panel considers it particularly important that all the subjective evaluations of the AES be transparent to all stakeholders and that this process be made easily accessible to any stakeholder so that they can enter their own subjective weights and judgments to estimate the sensitivity of the resulting operating policy to the recommended policy.

- All performance measures should be used in the screening and evaluation processes. If only some of the performance measures are used in the screening and evaluation processes, then there is an added risk that alternatives having high scores in non-considered performance measures are not promoted, even though they have higher overall scores when all performance measures are taken into account.

- Utility functions have not yet been developed. It is logical that these utility functions be developed after the uncertainty in the KBMOS suite of models is known.

- Specification of weights and utility functions are subjective. It is suggested that the sensitivity of alternative scores to weights and utility functions be identified.

- In screening and formulating alternatives, OKISS and MIKE 11 will use lateral inflows generated by MIKESHE/MIKE 11 for the base flow simulation. This assumes that the impact of alternative operating rules will not be significantly affected by using inconsistent lateral inflows at the screening and formulation levels. The validity of this assumption must be demonstrated to justify the use of OKISS and MIKE 11 at the screening and formulation stages respectively. A plan to assess the impact of using inconsistent lateral inflows was presented by the District at the Modeling Peer Review Workshop. However, this plan was found to be undesirable since it would not identify the impact of lateral flow inconsistencies during the screening process.

- It would be prudent to demonstrate that alternative operating criteria (at all levels of evaluation) are insensitive to the time series used for stages in Lake Okeechobee.

- The base condition should always be used as a benchmark for evaluating alternatives. Alternatives with performance measures less than the base condition should be ineligible for promotion

- Not all project objectives have performance measures. Specifically, the water supply, aquatic plant management, and downstream ecosystems objectives do not have performance measures and associated targets. Since the alternative plan selection process relies heavily on ranking alternatives based on a numerical scoring system in which scores are only associated with performance measures, this implies that natural resources and flood control objectives will be the primary bases for selecting alternatives and that water supply, aquatic plant management, and downstream ecosystems objectives will have a secondary role.

- Documentation of the alternative plan selection process should be revised to clarify the role of performance indicators in the promotion of alternatives. Specifically, the documentation should address the circumstances in which performance indicators will override the promotion of alternatives yielding high performance measures.

Although all of the above concerns are worthy of serious consideration, the panel recommends that the issue relating to lateral inflows be addressed with the highest priority.

Appendix

Panel Comments and District Responses

Objective A: Assessment of the Process Used to Select Modeling Tools

1. Adequate documentation and transparency of selection process and reviewer bias. For example:

    - It would have been more transparent to all readers to summarize evaluator responses to each criteria rather than just showing a compilation of scores.

    - Summarizing code capabilities in a large table would have been very useful for the reader to quickly see where some codes fall short.

    - The potential bias of each model evaluator is not clear.

    - There is minimal documented evaluation of the responses to the Request for Additional Information (RAI) in final model selection. Such information is needed to justify the selection of MIKE SHE / MIKE 11 as the more appropriate model for this project.

2. The simulated results from UKISS and OASIS should be compared to those from MIKE SHE / MIKE 11 in order to justify using one of these codes over the other.

3. Why wasn't UKISS used instead of OASIS, if indeed the test of OASIS was to duplicate the results of UKISS (Document 5).

    General Response to Comments Suggesting Document Revisions for Document # 1-4: *The model selection portion of the Kissimmee Basin Modeling and Operations Study was completed July 2005. The documentation is considered complete. The panels comments on how the results could have been better documented are appreciated and will be applied in future projects to avoid similar shortcomings in the future.*

    General Response to Comments on the Screening Tool: *Many of the questions and comments made by the panel about the Screening Tool are answered in the OKISS Technical Design Document which was not provided to the Panel and was an oversight on the part of the District. This document is now being provided to the panel.*

Objective B: Assessment of Modeling Tools and Performance Measures

Comment a) Distinguishing between 'code' and 'model' seems important here. Codes are 'selected', and 'models' are developed with the 'selected' code. The original objective indicated "selected model", but it should really be "model" as you've stated here. So are we assessing 'model' suitability and not necessarily 'code' suitability?

Response: *The panel is being asked to assess code suitability, i.e. the potential of the software to adequately simulate hydrologic and hydraulic processes critical to the KBMOS. In Workshop No. 2, the panel will be asked to assess model implementation, as described in the Technical Design Documents and Model Development Reports. Implementation focuses on the selection of appropriate time and spatial scales, an appropriate balance among the hydrologic components compatible with the available data and the objectives of the project, etc.*

Comment b) Were the proposed KB structure operating criteria ever defined? How can we assess the suitability of tools to evaluate either the proposed operating criteria without seeing these?

Response: *The study team has recently started the development and initial testing of the water control structure operating criteria. Documentation of this work is still pending. The current water control structure operations that need to be addressed in the model implementation are described in the Phase 1 Basin Assessment.*

Comment c) Do you think this objective is really to assess model and performance measure suitability for evaluating the relative effects of different operating criteria in the AES? Is it worth restating?

Response: *Objective 2B states:* Assess the suitability of the selected models and evaluation performance measures to evaluate existing and proposed Kissimmee Basin structure operating criteria. Based on the clarifications above, it may be appropriate to replace the term '…selected models…' with '… selected model codes…'.

1. It would be useful to prepare a single, comprehensive table (perhaps at the beginning of the document) that summarizes performance measures, locations, current vs. expectation by for example, natural resources of interest. It would be also be useful to include small conceptual diagrams in each performance measure that describe key flow conditions, expectations, targets etc. Interested stakeholders should be able to quickly see which performance measure(s), or expectation(s) would be used to evaluate a given system component or goal of interest to them.

   Response*: Agreed. The report will be revised to include a summary table in one of the introductory sections.*

2. What performance measures are the biggest losers through three evaluation levels? Is it possible to conduct a sensitivity evaluation at the first screening-level to see whose interest might be impacted most during the screening and formulation steps? Is there a way to track stakeholder

interests through evaluation process – so that each interest is addressed adequately? It might be useful to show how some performance measures are always dominated by others – it seems the screening tool (OASIS) would be very good at this.

***Response**: During the development of the individual modeling technical design documents, the ability of the codes to simulate the evaluation performance measures, as they existed at that time, was addressed. The issue of evaluation performance measure sensitivity to model tool will be addressed during the evaluation of Base Conditions.*

The team has developed a method to document Stakeholders input and to track how it is used in the Alternative Plan Selection Process. Further review of this is beyond the cope of this Peer Review Effort.

3. For the purposes of screening, how limiting is the one-day optimization horizon in OASIS, compared to looking further into the future in each simulation time step?

   ***Response**: The current version of OKISS is limited to current-conditions because today's operating rules are limited to current-conditions. OASIS itself is not limited to current-status information when making operating decision; it can consider past behavior and forecasts when making decisions. Incorporating near-term (2 week) forecasts into the operating rules is being considered. Incorporating long-term forecasts into the operating rules may be considered but only if forecasts are sufficiently reliable.*

4. For the purposes of screening, how limiting is the lack of modeling the surface-groundwater interaction in OASIS? The node-link based screening model should be able to include adjacent flood plains and flooding basins as well as the lakes and rivers and canals.

   ***Response**: Yes, in fact a revised version of OKISS will include nodes representing the Kissimmee River Floodplain*

5. How are lateral outflows addressed, especially during dry/low flow periods? For example, increased leakage from streams/lakes into the groundwater system may occur during dry periods and increased ET.

   ***Response**: The proposed approach has been modified. OKISS will only use MIKESHE lateral inflows*

6. Could OASIS be used to identify preferred alternatives given weighted performance measure objectives, rather than just for simulating suggested policies? If so, will it?
   ***Response**: Based on our understanding of this question, we believe the answer is yes, OASIS could be used in that way. However, the intention of the KBMOS Study is to express the Operating Criteria in terms and instructions that the Structure Operators can follow.*

7. More documentation would be useful to evaluate how the OASIS model will be implemented and validated in the LKB. For example, how will OASIS account for runoff in the lower Kissimmee Basin? The lateral inflows used in OKISS will not be consistent with the in-channel flows. Some quantitative assurance needs to be provided that this will not compromise the OKISS results.

Response: *The implementation of the model codes is not an objective for this peer review. The implantation of the OASIS mode will be part of the second peer review.*

8. What is the likelihood of valid alternative plans being eliminated based on inaccuracies of, or the limited subset of performance measures used by, the simpler water balance tool? In other words, how effective will OASIS be in screening of alternative operating policies and how has this issue been addressed?

   Response: *The KBMOS team considers the risk to be small. This issue will be discussed in more detail during Workshop #1.*

9. How will simulation errors at the screening-level be identified and then included in the alternative evaluation system? Could error-bars be added to simulated output like stage heights, flows, or velocities based on deviations between the calibrated model and observation data (at least for the 4 years used in the calibration)?

   Response*: The Screening Tool is intended to provide a simple model that will allow for the review and evaluation of numerous potential alternatives. Efforts that have been completed to-date, but not included at this level of the Peer Review Process has provided preliminary demonstration of the validity of this approach. These demonstration efforts will continue through the future Base Conditions Evaluation Efforts for the KBMOS which include a validation phase where the results of the various models are compared to quantify any significant errors between the modeling tools. Knowledge gained during this phase of the effort will be incorporated into the reporting of results.*

   a. The approach to model errors and uncertainty seems like an issue that still needs to be addressed. This seems especially true for the Evaluation tool (MIKE SHE) because results from this tool feed into the screening/formulation models, and this tool produces the most 'representative' flow conditions of the three modeling tools. The SFWMD response to this comment suggests that the calibrated model error won't be a significant issue, because simulated flows are for the base condition (not compared to historic data), and results of alternative simulations will be relative to each other. However, if the calibrated model performs poorly (for example under/over-predicts peak flows, stages, flow reversals, low flow event durations etc), one or more performance measures could still be significantly affected. The most sensitive performance measures might be related to river/lake stage elevations (i.e., E-03, E-04, E-07, E-10, E-11, E-12, and E-13). So AES scores may be lower for some of these, relative to other measures if there were no error. The fact that many of these are correlated with other flow-based performance measures suggests that calibration errors could also result in lower scores for specific groups of performance measures. It also seems possible that one alternative might be promoted over another based on a lower score for a given performance measure (or component), but both might be promoted if error were considered.

   b. At a minimum, some effort should be made to demonstrate that model error, and uncertainty (i.e., in Scenario 4 2025 AES simulations) do not adversely affect the AES process. The documentation should state how this will be addressed in the event that it is significant.

10. How will the alternative plans that will then be screened and simulated be selected and what criteria will they be based on? Are these criteria the same as used in the simulation models? Do performance measures address all operating objectives, and if not how will those objectives be included in the models?

    Response: *Please see the response to comment 14 under Objective 2C.*

11. Kissimmee Basin operating objectives include flood control (FC), water supply (WS), aquatic plant management (AP), natural resources (NR), and Lake Okeechobee ecosystems (LO). However, none of the performance measures relate to the WS, AP, and LO objectives. All operating objectives should be covered by at least one performance measure with associated target. It is acknowledged that WS, AP, and LO are covered by performance indicators, but without associated targets it is not possible to assess whether the operating objectives are being achieved.

    Response: *Please see the response to comment 14 under Objective 2C..*

12. Acceptable operating criteria are required to provide flood protection up to a 10-year storm event. In spite of this, none of the performance measures are directly related to the 10-year storm event. This issue will presumably be addressed by the U.S. Army Corps of Engineers, who is the technical lead for flood-event analysis.

    Response*: The U.S. Army Corps of Engineers (USACE) is the technical lead for the flood event analysis of this study and will be performing storm event simulations using their standard criteria and methodologies. Those are described in their DRAFT KBMOS Flood Protection Technical Memorandum that is currently being revised. That draft can be provided to the panel. However, evaluation performance measures E-07 and E-10 were developed to provide planning level information on flood control within the Kissimmee Basin for the entire 36-year period of simulation during the screening of alternative plans.*

13. Flooding in the KB will obviously bring some ecological benefits. Where is flood control to be implemented? Will the current water control structures be able to control flooding where flooding is not desired? Is there some reliability standard – like the n – year flood?

    Response*: This comment/question may be more appropriately addressed under document 6. Flood protection is handled through a Performance Indicator. All alternatives must maintain the existing level of flood protection. Defining the current level of flood protection is not within the scope of this project. Assessing flood protection will be handled in subsequent work by the COE as part of their EIS report which is needed to implement the rules generated by this project.*

14. Could the screening model be used to identify preferred plans and operating policies, or just evaluate those already proposed?

    Response*: The Screening Tool is intended to provide a simple model that will allow for the review and evaluation of numerous potential alternatives (both existing and proposed). The only limitation are those stated in the study constraints and the development of operating rules*

that can logically be implemented and monitored by the Operations Staff. Also see note on Document 3, comment 3.2 (second appearance)

15. Is there any thought given to calibrating an ANN with MIKE SHE/MIKE11 so that more runs could be made performed in less time?

    ***Response**: No. The team does not yet have enough confidence in the either the calibration data or in our understanding of the relative importance of the various model parameters to rely on ANN.*

16. What about ecological performance measures? How is hydrology to be linked to ecology, and what would be the ecological performance measures?

    ***General Response to Comments on the how hydrology and ecology are being linked in the Study:** The evaluation performance measures and indicators are intended to define desirable hydrologic characteristics within the Kissimmee Basin based on the knowledge and understanding of the potential linkages between the biology and hydrological features of the river, lakes and basin. They were developed to act as surrogates for the ecology/biology. These linkages are described in the rationale and justification sections provided for each evaluation performance measure and indicator.*

    Although ecological performance measures were not developed specifically for the KBMOS, evaluation performance measures and indicators were derived from ecological performance measures that have been developed for the Kissimmee River Restoration Project (Expectations) and that are under development for the Kissimmee Chain of Lakes Long Term Management Plan.

    Using the existing information compiled for the ecological measures as a reference, the evaluation performance measure development process focused on candidate biotic and abiotic indicators identified by the scientists and ecologists on the Interagency Study Team (see Table 3 on Page 13). Through a collaborative process, the most important aspects of hydrology that affect the candidate abiotic and biotic indicators were distilled. The measures were then translated into hydrologic terms since only hydrologic data (stages and flows) will be available from the modeling tools being implemented for the KBMOS.

17. Would not the decision making process want more than just the one best plan to be submitted to them for their acceptance and implementation. What about identifying tradeoffs among key Performance Measures and giving such information to the 'decision making process'?

    ***Response**: Although it is unclear in the report, it is the intent that the top 3 to 5 ranked alternatives will be presented to the SFWMD Governing Board for consideration. The report will be updated to better represent the intent.*

18. Creating indices by combining various performance measures and their relative weights may have to be done, but how is it to be done in a way that stakeholders can enter their own weights and see if the operating decisions are relatively insensitive to such changes in weights?

    ***Response**: The Performance Measure Evaluation Tool will have the flexibility to allow for adjustments to the weights and identify the sensitivity of the alternative scores to the objective*

weighting functions. Additional discussion of this feature and the intent is included in Document 8.

19. Why on a particular calendar date (June1) there would be such a change in zones A and B as shown on the Interim Operational Schedule for Structure S-65? Is it based on some legal requirement?

    ***Response**: The sudden change in the Operational Schedule that occurs on the structures on June 1 is not a legal requirement. This pattern is based on Operator's experience and reflects the shift from drainage for flood protection towards storage for dry season water supply needs. Different regulation patterns will be considered by this project.*

20. It appears that the MIKE 11 model will be used in Alternative Formulation with surface and ground water inflows taken from the base condition. If this is so, then there will be a disconnect between the flows in the channels and lake stages relative to the specified inflows for varying operating criteria. If this is the case, then the sensitivity of the simulations to this disconnect during the consideration of alternative operating criteria should be addressed, at least qualitatively and preferably quantitatively. Similarly, it appears that the OKISS model will be used to screen alternatives using specified inflows derived from MIKE SHE simulation of base conditions. As in the case of MIKE 11 used in alternative formulation, this is an inconsistency between channel/lake conditions and inflow conditions that needs to be addressed. As a minimum, it would need to be shown that this inconsistency is far less significant than the differences between operating criteria. The District plans to do this and provide supporting documentation of the results.

    ***Response**:* Once the base conditions runs have been created, the modeling team will perform a validation of the three modeling tools to compare the output and evaluate the ability to provide a consistent simulation of the KB. Evaluation performance measure results from the Future Base Conditions obtained with three different tools (OKISS, MIKE 11 decoupled, and MIKE SHE/MIKE 11 coupled) will be compared to evaluate the consistency between the three modeling tools.

    a. This is a good point that David made. I reviewed the SFWMD response and although it is assumed that inflows to water control catchments won't change for different operating alternatives, it seems like this should be demonstrated, at a minimum, by plotting/comparing lateral inflows and outflows (transferred to the screening and formulation tools) at different locations in the study area for the 3-5 full MIKE SHE alternatives. Documentation should include a strategy in the event that using constant lateral inflow/outflow is problematic.

    b. I was unsure, after reading the response from SFWMD whether "lateral inflows" meant "lateral outflows" as well. For example, in the MIKE 11 model, will the effects of ET and leakage from surface water (lakes, rivers) into the groundwater system be included in the 'lateral inflows'? Lateral outflows should be considered as well and documented as such.

21. It would be prudent to demonstrate that alternative operating criteria (at all levels of evaluation) are insensitive to the time series used for stages in Lake Okeechobee. The District plans to do this and provide supporting documentation of the results.

***Response**:* We agree with this statement. However, according to the results of the AFET calibration, the S-65E structure controls headwater stages independently of the downstream stage during almost 100% of the time.

22. Water extracted by public water supply wells is not returned to the system. What really happens to this water? It appears that part of the water pumped from the PWS wells is returned to the surficial aquifer via a rapid infiltration basin system, and that this partial return is being accounted for in the MIKE SHE model. This should be reflected in the documentation.

    Response: A portion of the PWS returns to the SAS through Rapid Infiltration Basin System (RIBS). This system was included to the MIKESHE module of AFET. The remainder of the water is considered as consumptive use and it is not returned to the system.

23. Although surface runoff and groundwater inflows to MIKE 11 are important – isn't ET also a big water balance component as well? Will ET loss from non-ponded surface water (i.e., flowing water bodies) also be considered in the AFET model?

    ***Response** Yes, it will.*

24. How valid is the assumption of uniform soil profiles throughout the entire KB for each soil type? What data will be used to 'calibrate' infiltration rates, runoff and AET?

    ***Response**: Being that the AFET is a regional model; the assumption of uniform soil profiles is adequate. The initial data set used in calibration process was obtained from previous calibration efforts of similar models in South Florida.*

25. There didn't appear to be any discussion on what initial conditions would be used for either the calibrated model, or any of the base condition models. It is unclear whether the 'calibrated' model (2001 to 2004) is the same as the 'Current basecase 2000' model? It seems like they should be the same, except that the calibrated model will be developed only for 2001 to 2004, and the basecase 2000 model will be developed over 1965-2000, but won't simulate land use/water use/operational changes over this period. Is this correct?

    ***Response**: The main difference between the calibration period continuous simulation model and the Current Base conditions model will be the period of simulation, rainfall input, and water control structure operating rules.*

26. For the calibrated model, will model input, such as landuse/cover, water supply, diversions remain unchanged during 2001 to 2004, or will they vary? It seems there may be some benefit to also simulate an Interim Base Condition using land-use/cover/water use for future time period, rather than 2000 year.

    ***Response**: The main difference between the calibration period continuous simulation model and the Current Base conditions model will be the period of simulation, rainfall input, and water control structure operating rules.*

***Response**: The Base Conditions Summary report suggests the best use for the Interim Base Condition is to evaluate how the preferred alternative performs during the transient periods of Kissimmee River Restoration construction. The model may have many such uses outside of this Study. Further clarification of the potential benefit is requested.*

27. Although 2005 data aren't available, if 2001 Kissimmee River conditions are specified as input, why aren't land use/cover and water use for 2001 also used? Would it be better to call this model a 2001 model? That way, it also correlates with the start of the calibration model.

***Response**: The name applied to the referenced period is the Current Base model. As stated in the document, Current Base includes year 2000 documented land use data, 1999 water use and infrastructure representative of the post Phase-1 KRR. The modeling team has refrained from placing a specific date name to any of the specified base conditions.*

28. For the calibration, will things like diversions be simulated in MIKE 11 using automated control strategies at key structures (that release, for example when simulated levels change), or will structures be operated exactly how they were during 2001 to 2004? Are there any examples of structures not operating as planned that might affect calibration, if automated controls are used?

***Response**: There are no automated controls in the KB. The calibration/verification period models manage structure to meet observed headwater conditions at each C&SF water control structure. These data are discussed in more detail in the AFET Technical Design Document which is part of the Peer Review Workshop No. 2 agenda.*

29. It is unclear how you can verify the calibrated flow model (current infrastructure of KRRP) for a period prior to the restored Kissimmee River, at least within the LKB? I thought the main point to verification (or model validation, since verification is typically used with 'code-verification') was to demonstrate that the parameterized model performs adequately for a different set of external stresses (i.e. precipitation, temperature, diversions, groundwater pumping etc).

***Response**: Most of the available post–Phase I KRR period-of- record was used in the calibration. Although the hydraulics of the floodplain will not be represented during the selected verification period, it was decided that the selected verification period will provide enough information about the rest of the basin hydrologic processes.*

30. It is unclear how well the 'calibrated' model will reproduce more extreme climate periods during the 1965-2000 time period. This seems important, because many of the performance measures are sensitive to more extreme climate conditions (i.e. flooding, high-low lake/river stages etc.). This doesn't seem to be captured by using the storm event verification time period 8/2004 to 10/2004 (too short).

***Response**: The event period verification run represents an extremely active hurricane season where three major storms passed over the KB within a 45 day period. The long term continuous calibration also includes this same time. However, the long term continuous calibration period uses daily rainfall generated from gauge data while the event period verification uses 15 minute NEXRAD data.*

31. It is unclear what 'both dry and wet periods' refers to? Does this imply that extreme wet and dry periods occur during a 4 year period?

    ***Response**: Yes, The Verification period includes the year 1994 which has the highest annual rainfall in both the Lower and Upper Basin in the late (>1990) Pre-Phase I period and 1996 that has the lowest annual rainfall in the Lower Basin during the late Pre-Phase I period. Due to land use and water use changes in the last 15 years, years prior to 1990 where not included in the selection.*

32. It is not clear which models (or codes) will be used to evaluate base conditions? What quantitative methods will be used to assess the model output to make sure it 'makes sense on a grand scale'? What happens if it doesn't, for example with the screening-level tool?

    ***Response** The Base Conditions are going to be evaluated with both tools (OKISS and AFET). Results obtained with both tools for the Base Conditions will be used to evaluate the Performance Measures. Scores obtained with both tools will be compared and checked for compatibility.*

33. How is uncertainty to be handled? Will there be a sensitivity and uncertainty analysis performed, and if so how? Uncertainty should be quantified where possible, rather than described qualitatively. For example, the uncertainty in flows and stages can be estimated from the accuracy of the instruments and methods used to measure or derive these data. The precision of target values should quantitatively reflect the uncertainty.

    ***Response**: The Earth Tech team is currently working on an evaluation performance measure sensitivity analysis.*

34. This discussion in the text on uncertainty as it ultimately relates to the process of alternative evaluation and modeling seems unclear. For most performance measures, the discussion of uncertainty seems more like an indemnification clause, rather than a discussion about how to quantify uncertainty in either the data or simulated output that will later be addressed in the alternative evaluation.

    ***Response**: The uncertainty section will be revised to improve clarity and the evaluation performance measure sections will be modified to eliminate "indemnification clauses".*

35. How will stakeholders know how well the model performs for the majority of years not used in calibration, especially if more extreme hydrologic conditions are encountered?

    ***Response**: The calibration and verification periods include a diverse range of hydrologic conditions that are representative of the 36-year period of simulation (1965-2000) proposed for use with the evaluation performance measures. The use of 1965-2000 "climate conditions" does not necessarily mean that those years are going to be explicitly simulated with the Study tools.*

36. The suggested data to be collected should be driven by their need and importance (sensitivity) in the decision making process. Are such data based on such preliminary modeling and analysis, or are all the data to be obtained and then someone will figure out how to use it?

***Response**: The information that will be provided to the decision making process will come from the Alternative Evaluation System (AES). That system is intended to provide a means to rank alternative plans relative to their ability to meet the targets and indicators described in this document. The AES will assign importance to the measures and provide information to be considered for indicators. The package of information that will be submitted for the 3 to 5 alternative plans will show how well those plans score for natural resources and perform relative criteria defined in the indicators for flood control, water supply, Lake Okeechobee, and aquatic plant management.*

37. There seems to be no water quality criteria. Why? Or will meeting water quantity criteria ensure satisfactory water quality conditions?

***Response**: While water quality is a major driver that controls the features of aquatic systems, no water quality models or evaluation performance measures addressing water quality are being developed for KBMOS. However, the modeling tools were selected based on their ability to incorporate water quality modeling in the future. KBMOS is not assuming that water **quantity** criteria will ensure satisfactory water **quality** conditions.*

38. The stated rationale for maintaining a minimum flow of 250 cfs is to attain desirable levels of dissolved oxygen. Relating the flow rate to dissolved oxygen is questionable since conventional wisdom is that reaeration rates depend on both velocity and depth. If sediment oxygen demand (SOD) is the controlling variable on DO levels, then SOD reduction might be more directly related to the channel velocity. Is there a graph or diagram that could clarify how DO is related to velocity, flow, or their rate of change in time?

Response needed

39. Is the constraint that velocities not exceed 0.5 m/s and flows be at least 7 m3/s to preserve fish communities solely based on DO levels becoming problematic, or are there other reasons (other than 0 flow)?

Response needed

40. Consideration should be given to specifying target levels that are commensurate with the accuracy of the data used to determine them and the accuracy of the model that is to be used to generate the performance measure. For example, it is doubtful that the percent of days with flows less than 250 cfs during June – October is historically equal to 0.78% (given the uncertainty with which flows are calculated), and it is certainly unlikely that the hydrology model will be able to resolve the flows to this level of accuracy. This target could be rounded to 1%. A similar consideration applies to other targets. The District will be making these changes.

***Response**: Agreed. Similar comments have been received from the Environmental Panel and the targets will be revised.*

41. Evaluation components characterize the probability of the river channel stage exceeding the average ground elevation. How was the average ground elevation calculated? Is it the average elevation across a specific channel profile? It might be just as effective and less cumbersome to

characterize the distribution by its first three or four moments (i.e. mean, standard deviation, skewness, and kurtosis). Percentile values could be used instead of moments; however, using seven percentiles to characterize a probability distribution needs to be justified. Similarly with respect to the probability distribution of stage fluctuation per water year.

***Response**: Skewness and kurtosis were considered during development of this evaluation performance measure, however, percentiles seemed to be less abstract for stakeholders and the public. Based on input from the Environmental Peer Review Panel, E-03 is dependent on E-01 and E-02 and may be dropped completely or E-03 Components A to N will be consolidated into a simplified consolidated component that will evaluate the average number of days above average ground and the average river channel stage fluctuation range.*

42. Why wouldn't the timing of the river stage and floodplain hydroperiod be considered as another component in this performance measure, or a separate performance measure? This seems important for lake stages, and it is not considered over the hydroperiod duration.

***Response**: Is reference to E-02 in the second sentence an error? Timing of river stage and floodplain hydroperiod is considered in all components of E-04. Kissimmee River Stage Recession / Ascension and timing of river flows is considered in Component A-D and H of E-02. Seasonality and Variability of Kissimmee River Flows. Stages in the river/floodplain are directly related to flows and this is a redundancy that the environmental panel has identified. E-03 is dependent on E-01 and E-02 and may be dropped completely or E-03 Components A-N will be consolidated into a simplified consolidated component that will evaluate the average number of days above average ground and the average river channel stage fluctuation range.*

43. The primary concern addressed by the Kissimmee River Energy Grade Line performance measure is the so-called scour potential at two critical locations: the downstream terminus of the backfilled canal and the point of return of the restored river channel to the C-38 canal. The scour potential is fundamentally related to boundary shear stress, which is proportional to the flow velocity squared. Therefore, the product of velocity difference squared times cycle length is a more appropriate measure of scour potential.

***Response:** This could be rewritten in terms of scour potential with the target being set to "zero," but this EPM provides more than scour protection, it also provides that the desired wetting and drying of the floodplain in the reach immediately upstream of the downstream terminus will match the intent of the restoration objectives in this area as in others.*

44. It would be more useful to calculate the confidence limits of the historical flow duration curve and then characterize the excursions of the simulated flow duration curve beyond the confidence limits of the historical flow duration curve. In this way, only statistically significant deviations from the historical flow duration curve are identified and noisy deviations are filtered out.

***Response:** This EPM may benefit from further discussion and this suggestion may be a better way to evaluate flow duration concerns. There has been a long-standing debate as to whether it is better to meet more of the low flows or high flows when there is a deficit of flow in the simulation period as compared to the reference period or is it better to distribute the deficit along the entire probability range? Repeated attempts to have biologists answer this for engineers and planners have failed. It seems intuitive that this is something one would want to*

know in evaluating alternative plan performance, so this EPM was set-up in this way; however, the five evaluation components may be collapsed into one for the entire probability range and the suggestion here would be very applicable.

45. The Frequency, Duration and Timing of High and Low Lake Stage targets do not align with pre-regulation data and reflect today's realities. Based on the supporting data presented, these targets are based more on professional judgment than on historical observations. Such targets should be flagged as somewhat less reliable than targets based on documented historical conditions.

    ***Response**: Full restoration implies restoration of all hydrologic patterns. As noted, development in the watershed prevents full restoration for the lakes of the Kissimmee basin. The idea of partial restoration assumes that significant environmental benefits will occur even though only select hydrologic signals are restored. These critical signals are our targets. Defining critical patterns is a matter of professional judgment but environmental monitoring programs allow scientists to revisit and improve these "less reliable" targets and to modify structure operating rules through adaptive management..*

46. The Lake Littoral Zone Inundation evaluation components will be significantly influenced by lake bathymetry and will provide a very weak measure of the duration and frequency of littoral zone inundation.

    ***Response**: Agreed.*

47. The sub-watershed runoff volume indicator is not affected by structure operating criteria and could be called something other that a performance indicator.

    ***Response**: Agreed.*

48. No evaluation components are specified for water supply for consumptive use.

    ***Response**: The Water Supply performance indicator was still under development at the time this document was provided to the Panel. Attached is the draft indicator.*

49. The Palustrine wetland hydroperiod/surficial aquifer water budget indicator should be affected by structure operating criteria. Does the phrase "Palustrine wetland hydroperiods and surficial aquifer water budgets may not be affected by the modification of structure operating criteria…" imply a constraint?

    ***Response**: Component A.3 was intended to calculate, for each grid cell and for each water year, the number of days in a water year that water depth is > 0.1 ft. This calculation would result in 35 values (since there are 35 water years in 1965-2000) for each grid cell. Component A.4 was intended to calculate the mean of the 35 values from Component A.3 so that a single value is obtained for each model grid cell for each simulation. The report will be revised to improve clarity.*

50. The difference between pre-channelization and pre-regulation is not clear. It is not clear in all instances that pre-channelization conditions are most suitable for achieving all of the desired expectations. There may be pre-channelization conditions that are not ideal, for example, for fish populations at a given time of year.

***Response**: In the report, pre-channelization and pre-regulation refer to approximately the same period of time (prior to May 1962). Pre-channelization was used to specifically describe that period of time in the lower basin and pre-regulation was used to describe that period of time in the upper basin. The report will be revised to improve clarity.*

51. It would be very useful to show in each performance measure section, pre-channelization versus post-channelization (current) conditions. It is difficult to visualize problems with current condtions, and how these conditions will improve by going back to the pre-channelization state. It would also be helpful to see a comparison between pre- and post-channelization here to better visualize the expected change.

***Response**: Agreed. The report will be revised to include summaries of both pre- and post-channelization/regulation hydrologic data.*

Objective C: Assessment of the Alternative Evaluation System

1. Can a plan be generated from this method that is a mix of other plans that have been identified? In other words, can this method identify the best plan according to some specified criteria?

***Comment #1 Response**: Yes, from examination of the graphics that demonstrate the application of the AES in the planning process, you will notice a 're-cycle bin' icon. This is to represent that nothing is totally discarded. It is the intent of the study team to keep the plan components that perform well. This is also represented in the planned re-formulation phase during the later stages in the alternative plan selection process and during presentations to the public, stakeholders and the SFWMD Governing Board.*

2. How can stakeholders be involved in this ranking process? Will there be public hearings on the outcomes of analyses? What if they don't agree on, or don't understand, the weightings of various criteria?

***Comment #2 Response**: A public meeting will be used to collect input for the development of system to rank and promote alternative plans (the AES). Other public meetings have been and will continue to be used to vet the AES and the results of the Alternative Planning Process. It is the intent of the study team to develop the AES within the spreadsheet environment of the Scorecard Utility. In this environment, changes to the weighting system can be implemented to demonstrate how the ranking of an alternative may change based on a change in the objective weighing in the AES. (also see response to question 12).*

3. Won't the weights depend on the value of the criteria as well as on the criteria themselves?

***Comment #3 Response**: The application of Utility Index functions is intended to normalize the scores of the components and measures, therefore the values of individual measure should not dominate the function. Additional clarification from the author of this question may be required.*

4. What if different plans satisfy different criteria best and there is a disagreement on their aggregation?

   Comment #4 Response*: The Alternative Evaluation System is designed to aid the study team members in finding the best performing alternative plan which will have the maximum "total weighted composite score" and the best combination of opportunities for water supply and aquatic plant management while minimizing undesired flows into Lake Okeechobee and avoiding any violation of the defined flood control level of service.*

5. The use of principles from the decision science discipline will not guarantee a process that is 'right' or 'scientific'. Decision making is a political process. The context (environment) in which decisions are made and the process in which they are made also matters.

   Comment #5 Response*: The use of decision science methods are intended to provide a systematic way to evaluate alternative plans and add to the transparency as well as provide skeptical stakeholders and decision makers that stable, documentable, and repeatable process is used; however, the use of these methods will not supersede the political or other human "hands on" aspects of decision making in the public arena. The AES scores and ranks alternative plans by applying a consistent set of criteria to the modeling results produced from each alternative plan simulation. The Alternative Plan Selection Process promotes alternative plans for further consideration based on the resulting scores and the associated ranking. this process is intended to assist decision makers by reducing the set of alternatives to be considered to a manageable number. Ultimately, the SFWMD Governing Board will select the preferred alternative and can redirect the Study Team to reconsider other alternatives than those promoted for their consideration.*

6. The "combining of evaluation components and locations to create a composite performance score for each Evaluation Performance Measure within an alternative plan" will be a BIG chore! Perhaps the use of color coded maps will help (similar to the HSI study).

   Comment #6 Response: *The Study Team recognized early in the project that it would be very difficult to identify an alternative plan that stakeholders could agree represents an improvement to current structure operations. We needed a system to evaluate alternative plans that was unbiased, transparent, repeatable, documentable, and implementable. A system that meets these criteria would allow us to build stakeholder consensus, confidence, and trust in the results from modeling the different alternative plans. The resulting system is the Alternative Evaluation System (AES). This system works in combination with the evaluation performance measures and indicator measures and the three phased modeling approach (screening, formulation, and evaluation) to systematically produce and preserve information that stakeholders have identified as being critical to their consideration of alternative plans. The Performance Measure Evaluation (PME) tool that is being developed will automate the AES will report the model results for each alternative plan relative to the content described in the evaluation performance measures and indicator measures. These results will be posted to the web and provided in report format with summary scoring to assist stakeholders with digestion of the content. See Figure #1 and comment #14 response for additional detail.*

7. Using utility index functions and then weights, involves two levels of abstraction. Will this confuse the issue or make it harder to be transparent and understandable to the public stakeholders?

It would be useful to include a more detailed description on how the components of performance measures at different locations will be combined into a single measure. This will provide reference documentation when the needed weights are determined, and avoid confusion when this step is coded into software.

Comment #7 Response: *Agreed. This task is part of the supplement to the Evaluation Performance Measure and Indicator Measure document that was postponed until the second part of the Review. This document presents the Utility Functions and the Study team weights for Evaluation Components and Locations.*

8. It would be useful to assess alternatives by considering both their objective score and their overall performance score. Whereas it is logical to promote an alternative based on its overall score, the objective score will be particularly useful in tweaking alternatives since consideration of the objective scores will identify the alternatives that are strong in particular objective dimensions.

   Comment #8 Response: *The objective weights will be recommended by the Consultant Team to the District based on results from a stakeholder survey and other relevant information associated with the KBMOS study objectives. Once adopted the objective weights will remain fixed throughout the alternative screening, formulation, and evaluation to ensure consistent scoring and ranking across all alternative plans. If objective weights are modified, results from all levels of evaluation would need to be recalculated relative to those weights to ensure equivalent consideration of all proposed alternative plans.*

9. Clearly the weighting will be subjective and prone to some level of bias (for each objective, location, and performance measure evaluation component). Would it be possible to evaluate how sensitive the composite performance score is to various weightings? This way, for at least some alternative plans, it might be possible to demonstrate or justify that it is a combination of performance measures that cause them to be omitted, and not just the weighting – thus reducing bias. Maybe use @risk within Excel to conduct this evaluation. Each weight could be treated as a PDF.

   Comment #9 Response*: As was mentioned in response to an earlier comment, it is the intent of the study team to develop the AES within the spreadsheet environment of the Scorecard Utility. In this environment, changes to the weighting system can be implemented to demonstrate the sensitivity (or insensitivity) of the evaluation components and locations. Testing of evaluation performance measures is also planned during the evaluation of base conditions and early screening to evaluate the potential to eliminate or combine evaluation components.*

   It seems in response to the question "Does the alternative evaluation system meet its objectives of being unbiased, transparent, repeatable, documentable, and implementable?", that this is true except for "unbiased". Is it possible to assess the level of bias, without seeing the specific values for performance measure and objective weights, and utility indices?

10. It would be nice to know for each alternative evaluated, which performance measures dominate the composite performance score.

***Comment #10 Response**: Agreed. This will be provided in the "Evaluation Report for Alternative Plan X" (see example below). The details of this report (one report per alternative plan) have not been finalized, but the concept behind this comment will be addressed. It is foreseen that each report will closely follow a template so that comparisons of performance of two or more plans will be easy. Tabular formatting will be used and some graphical representation may also be used. It is envisioned that the set of reports from each level of model simulation will be bound together as part of the modeling results documentation. In this compilation of individual reports, a set of summary tables and graphs is expected to be provided to aid in reviewing alternative plan performance efficiently.*

11. Perhaps a lot of meta data will be needed associated with each analysis to find the best solution. Who judges what is good or bad or best?

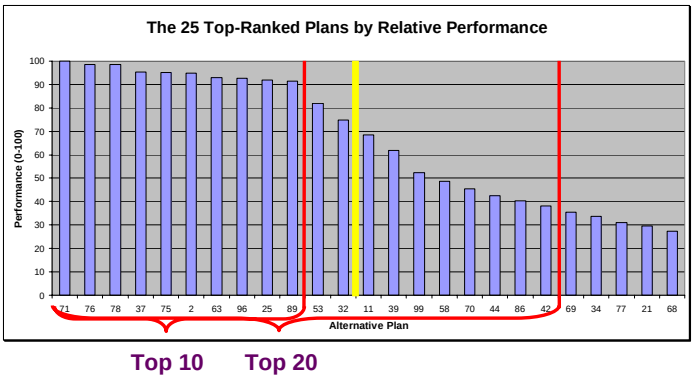
***Comment #11 Response:** The Evaluation Performance Measures and Indicator Measures describe the meta data that stakeholders have defined as essential to their decision making process. The model will produce output for each of the alternative plans and the Alternative Evaluation System (AES) will combine these meta data into a framework that can be used to score and rank alternative plans based on how well a given alternative plan meets the targets and/or conditions defined by the performance measures and indicators. The drivers for the AES scoring and ranking are the weights assigned by the Study Team to evaluation components and locations and the objective weights that will be derived from stakeholder survey results. The AES is a recommended methodology to narrow down the range of potential alternatives that meet the goals and objectives of the study. Ultimately, the SFWMD Governing Board will decide, through an open process with input from the public and stakeholders, what is Good or Bad or Best.*

12. AES will need to be very transparent and suitable for stakeholders to change assumptions to see just how sensitive the 'best' solution is to those changes in assumptions.

***Comment #12 Response:** Agreed. This has been explained and demonstrated with example data during Study Team meetings and two public meetings. Three figures from the example are included here. In the first figure, the top 25 of some greater number of alternative plans evaluated by the screening tool are plotted by rank based on relative performance. It was explained that out of 50 to 100 alternative plans evaluated with the screening tool that 10 to 20 will be promoted to the next level of evaluation.*

Evaluation Report for Alternative Plan *X*

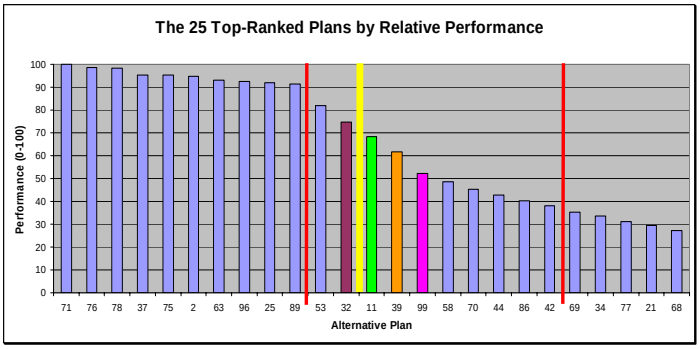
Qualitative Interpretation of Simulation Results	Quantitative Interpretation of Simulation Results
• **Flood Control** Plan does not violate the flood control constraint. Plan shows significant margin between peak flooding of with and without plan and in the S-59 subbasin. • **Water Supply** 12,400 acre-feet of excess water is available with plan in year 16 in the S-61 subbasin. • **Aquatic Plant Mgt** Opportunities for aquatic plant management occurred on Lake Kissimmee in years 9 and 21. • **Lake Okeechobee** Discharges made to Lake Okeechobee exceeded the desired volume by 26% during year 16 and 17% in year 29. (examples of qualitative interpretations) [more details and graphics as required]	• **Natural Resources** Weighted Composite Scores EPM# Score 1 6.3 2 4.4 3 3.6 4 4.5 5 5.1 6 2.5 7 6.7 8 8.2 9 4.9 10 5.5 11 3.4 12 6.8 13 5.2 Total for Plan 67.1 (examples of quantitative interpretations) [more details and graphics as required]

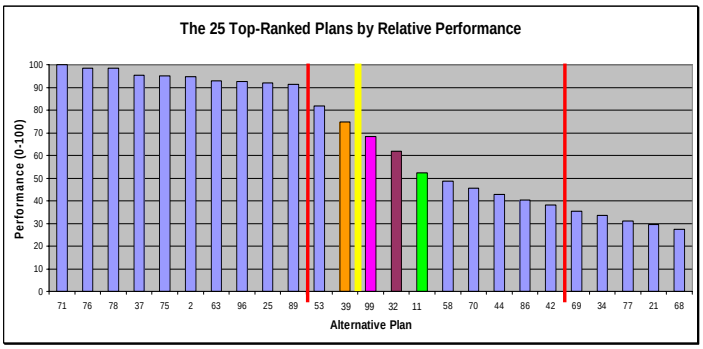


Two red bars are drawn to show the range where the promotion cut-off will occur. A rationale for making the cut between the 12th and 13th ranked alternative plans (plans 32 and 11) is indicated by the yellow bar.

For demonstration purposes, it is explained that one stakeholder group who is most supportive of plan 39 believes the ranking and cut-off may be biased or at least unfairly eliminates that plan from further consideration. To test this sensitivity, the scoring and ranking is re-calculated (without need for re-running models) with customized weights that would clearly avoid all perception of bias by the stakeholder groups as they are invited to provide their weight preferences for this sensitivity test.

After the sensitivity test is conducted, it is shown by comparing the second figure with the third by comparing relative change in the rank and positions of the plans that under this potentially extreme weighting, plan 39 does move up in rank three places and there is some minor shifting about the proposed cut-off of 12 plans. It is noted that no wholesale shift occurred where high performing plans traded places with low performing plans.





The example is ended with a reflection and suggestion that in a case like this, the original ranking with the more rigorously developed (hence, less biased in theory) weights will be used and the original rankings will be restored. However, the cut-off will be moved to include two more plans, promoting plan 39 to the next level of evaluation. Then it is explained that in this three tiered evaluation process, great effort is placed on reducing the chance that bias will unfairly eliminate a plan from consideration. The concept that inclusion is the rule rather exclusion is reinforced.

While this is simple theoretical example, it is the intent of the Study Team to evaluate the sensitivity of different weighting schemes and be able to address any similar concerns that arise among stakeholder groups.

13. The base condition should always be used as a benchmark for evaluating alternatives. Therefore, the base condition should be shown, and it should be noted that alternatives with performance measures less than the base condition are ineligible for promotion.

Comment #13 Response: *The process identified in the Base Conditions Summary Report is meant to demonstrate how the base conditions will be scored using the Alternative Evaluation System in the same manner as the alternatives. This will allow the evaluation of system improvements relative to the various alternative plans developed for the respective time periods (i.e. future base conditions will be compared for the future alternatives). The suggestion to set the Base Condition as the minimum level for alternative promotion based on each evaluation performance measure will be forwarded to the Interagency Study Team for consideration.*

14. One should discriminate between performance measures and performance indicators since these metrics are handled differently assessing alternatives. It will also become apparent that only two of the five objectives have associated performance measures.

Comment # 14 Response : *The Evaluation Performance Measures (EPM) and Indicator Measures (IM) are components of the Alternative Plan Selection Process that will be used to promote alternative plans through the screening, formulation, and evaluation phases of the study. Figure 1 is an illustration of the Alternative Plan Selection Process.*

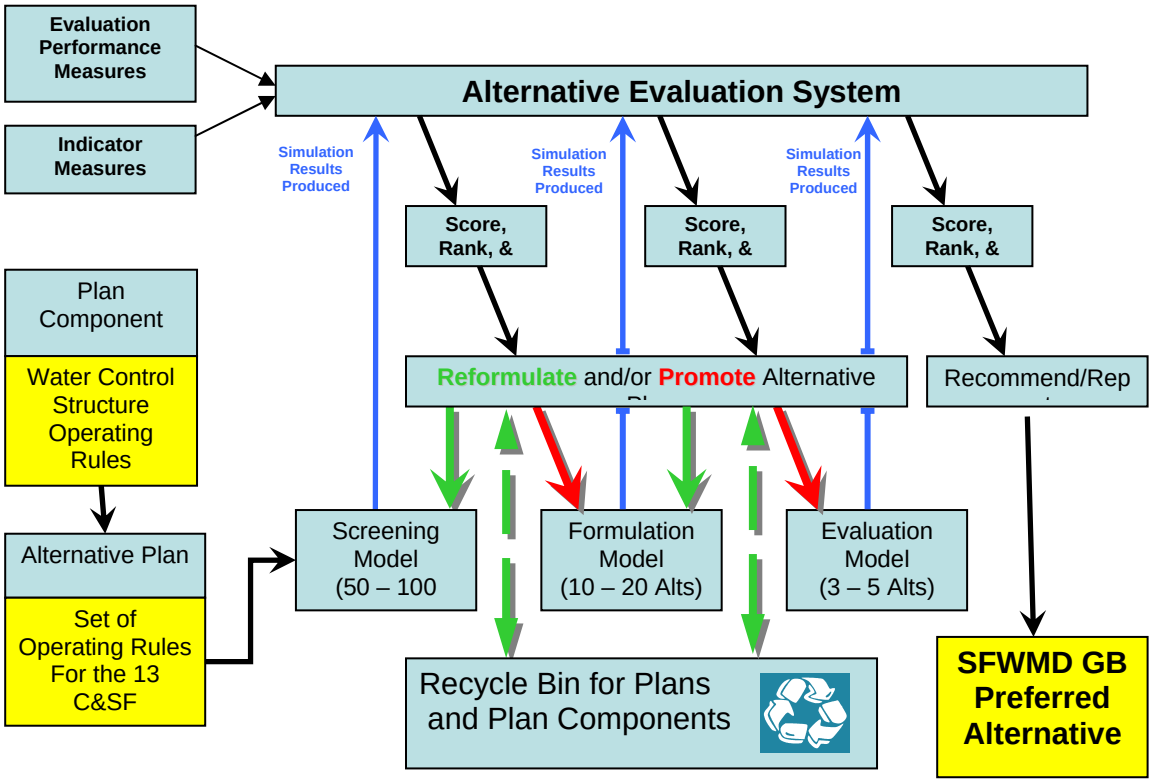


Figure 1: Overview of Alternative Plan Selection Process

The EPMs and IMs identify desired characteristics of the system hydrology. They have been defined by the Study Team to represent the desired criterion for flood control, water supply, aquatic plant management, Lake Okeechobee, and the natural resource requirements of the Kissimmee River Restoration and the Kissimmee Chain of Lakes Long Term Management Plan. The AES was originally conceived in parallel with evaluation performance measure development before the concept of indicator measure was introduced. As the evaluation performance measures matured, the concept of indicator measures emerged as the study team recognized that quantitative targets could not be developed for all aspects of hydrology related to the KBMOS Operating Objectives. At the same time, the AES was forced to evolve to include a qualitative methodology that would allow the indicator measure results to be combined with the quantitative comparison of evaluation performance measures. The Alternative Evaluation System (AES), as presented in Document #8, defines the proposed methodology for quantitatively differentiating between alternative plans using the evaluation performance measures that define natural resource requirements. Missing from this document is the methodology that will be used to qualitatively differentiate between plans using the indicator measures. The AES TDD should have been updated to reflect this methodology but was not. Below is an overview discussion of that methodology. The AES document will be amended to include a detailed discussion of this methodology.

The following is a more detailed explanation of how the KBMOS objectives will be balanced using the AES.

Together the Environmental Performance Measures and Environmental Performance Indicators represent the expected performance of a given alternative plan. When Environmental Performance Indicators for two alternative plans predict very similar outcomes,

the two plans can be ranked according to their Performance Measure scores. When Performance Measure scores for two alternative plans are essentially equal, the two plans can be ranked according to differences in their Indicator Measure predictions.

Policies, laws, mandates, and related logic have been incorporated into the Environmental Performance Indicators. The Flood Control Indicator Measure represents a planning constraint; it is treated as "pass-fail" criterion for each alternative plan. The other Environmental Performance Indicators represent planning objectives; these do not include targets or minimum or maximum limits due to lack of specific knowledge to define them and set incremental values for them.

There are no policies, laws, mandates, logic, or other guidance that provide a means to quantify predicted alternative plan performance from the Environmental Performance Indicators nor to compare it to the quantifiable scores computed for the Natural Resources Performance Measures developed for the KBMOS.

In Systems Analysis or Operations Research terms, the Alternative Evaluation System is designed to aid the study team members in finding the best performing alternative plan which will have the maximum "total weighted composite score" and the best combination of opportunities for water supply and aquatic plant management while minimizing undesired flows into Lake Okeechobee and avoiding any violation of the defined flood control level of service.

*The **best alternative plan** will be the plan that:*

1. *has the maximum "total weighted composite score" for natural resources,*
2. *has the best combination of opportunities for water supply and aquatic plant management,*
3. *has the minimum undesired flows into Lake Okeechobee, and*
4. *does not violate the defined flood control level of service.*

These are the "Quantitative and Qualitative Criteria" represented by the EPMs and EPIs and are referred to again in the following paragraphs.

If an alternative plan performs well for criteria 1-3 but fails criterion 4, it may be a candidate for reformulation with the objective of modifying it to meet criterion 4 while keeping the best features that allowed it to perform high in the other three criteria. This approach has the purpose of avoiding discarding otherwise high performing alternative plans that narrowly fail to meet criterion 4.

Based on policies and laws that direct the KBMOS, if for two plans, plan one has a higher total weighted composite score than plan two, and they both meet criterion 4, and plan two provides more opportunities under criterion 2 and/or has fewer undesired flows under criterion 3, plan one will be ranked higher than plan two. Criterion 4 must be met. There are no metrics to rate or construct tradeoffs for performance for criteria 2 and 3 among each other or with criterion 1. Criterion 1 is given precedent over criteria 2 and 3 because the primary focus of the KBMOS is to improve the natural resources of the KRR and KCOL without violating flood control level of service and to secondarily provide for water supply, aquatic plant management, and mitigation of Lake Okeechobee inflows.

There is no explicit formulation of the objective function for KBMOS, so these five study objectives/four criteria are balanced with the degree of subjectivity described above. Finding a balance between study objectives means finding a "satisfactory, subjective balance" which requires a hierarchical application of four criteria with criterion 4 being absolutely required, criterion 1 being an objective quantification of the total weighted composite score and is of primary importance, and criteria 2 and 3 being qualitative measures of performance and are of secondary importance.

15. Figure 4.4-1 in Document 8 (see below) shows how performance measures can be distributed among objectives. This figure indicates that performance measures can only be associated with a single objective. In general, performance measures can be associated with more than one objective. It might be useful to include this in the figure and subsequent illustrations of the calculation of performance scores.

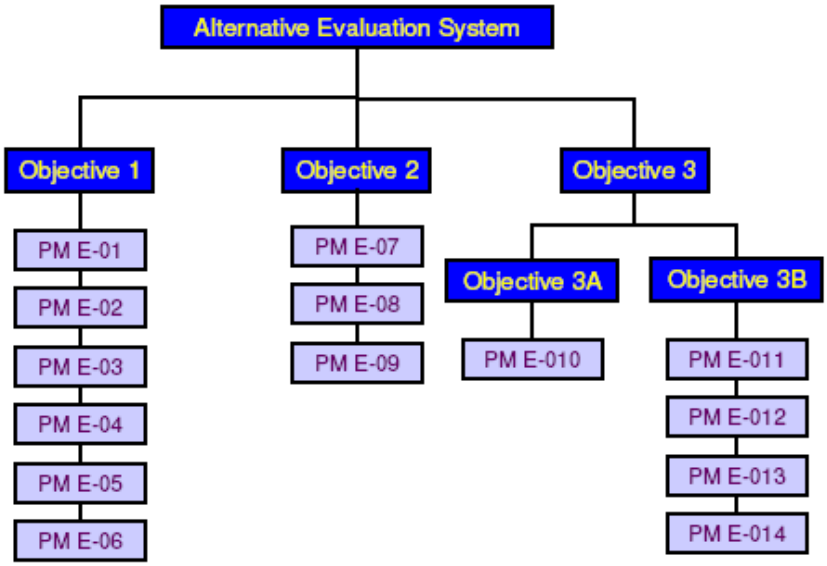


Figure 4.4-1 Example of logically organizing performance measures within a decision hierarchy.

Comment # 15 Response: *Agree that this is possible and is sometimes included in an evaluation system. For example, annual floodplain peak stage, per se, could be included in a decision hierarchy under flood control as well as natural resources and the dual set of corresponding utility index functions and weights would be configured as necessary to make this single performance measure serve two parts of the scoring scheme. However, earlier in the process of developing the performance measures for the KBMOS, it was determined that this capability did not need to be implemented and some on the Study Team believed including it as a hypothetical would cause undue confusion among some stakeholders who are being exposed to a scoring system such as this for the first time.*

District responses on plan selection process:

General Response to Comments on the Alternative Plan Selection Process: *Based on panel questions and comments related to Document #2 and Document #5 and those received from the Environmental Peer Review Panel, the District recognizes the need to produce an "Alternative*

Plan Selection Process" document that integrates content from the Alternative Evaluation System Technical Design Document and the Evaluation Performance Measures and Indicators Document into a single framework that puts into context the entire process that will be used to develop and promote alternative plans through the screening, formulation, and evaluation portions of the Study.