ECOSYSTEM RESTORATION PLANNING MODEL DEVELOPMENT

RECOMMENDED FUTURE CONSIDERATIONS

LESSONS LEARNED FROM SIXTEEN PLANNING MODEL QUALITY ASSURANCE REVIEWS

PREPARED BY
BATTELLE

MAY 2012

This Page is Intentionally Left Blank

The following lists generally recommendations based on common and/or notable comments from 16 quality assurance reviews of 34 ecosystem restoration planning models. This list is not intended to be comprehensive, but rather to focus on potential issues or concerns for consideration so that model developers can build better models and methods from the start; reduce the amount of time developing, reviewing, and revising methods and models; and minimize the potential need to redo project analyses if revisions to a model that has already been used are significant.

Model Documentation

Model documentation should include clear, specific, and detailed descriptions of:

- ✓ *The intended purpose of the model and how model outputs will be used by decision-makers* – This may seem intuitive, but it has not always been clear.
- ✓ *The ecosystem or habitat being modeled* – Clearly describe the ecosystem represented by the model when models are developed for very specific habitats within a specific region. For example, floodplain forest ecosystems can consist of several distinctly different communities, including the riparian forests, which are contiguous to and affected by surface and subsurface hydrologic features of the perennial and intermittent flowing streams throughout a watershed, and the flatwoods, which are usually associated with poorly drained shallow depressions located between the river and its tributaries. The user needs to understand which habitat is being represented. The description of the ecosystem should be detailed enough to support the "reality" of the model. Species lists are helpful for understanding the diversity associated with the habitats represented.
- ✓ *The spatial resolution and geographic boundaries of the model* (e.g., resolution of raster or elevation data; project boundaries) – The spatial resolution and geographic boundaries of the models need to be clearly understood by users to prevent misapplication. For example, the resolution of elevation data is critical for evaluating the quality of flooded habitat for feeding [birds] and breeding [fish]. Although a resolution of +/- 1 foot may mean the presence or absence of critical habitat, documentation should explain and support whether this resolution is acceptable considering whether this affects the presence of suitable habitat at any one location (i.e., there may just be a shift in the location of the habitat) vs. whether the habitat exists at all. Also, some models have been developed for specific projects, and although the models may be able to be applied outside of the project boundaries, the documentation should clearly explain that it was intended for the project area.
- ✓ *How the model was developed* (e.g., variables selected, variable weights, model architecture).
- ✓ *The link between performance measures and habitat condition* – Performance measures are the basis for evaluating habitat condition. What is lacking in several of the model documents that have been reviewed is an explanation for values of 0 – 1.0 for performance measures reflect habitat quality. For example, how does the value of 0.7 translate to number of species present, or wetland function, or number of fish nests.

- ✓ *Major assumptions and limitations* – It is critical that users understand the major assumptions of any model so that users are clear on the model's limitations to minimize potential misapplication of the model and provide a complete understanding of how model outputs can and cannot be used. Scientific support should be provided for assumptions. For example, a specific assumption might be that past and current hydrologic conditions can be used to represent future hydrologic conditions. A major assumption of nearly all of these models is that Habitat Suitability Index values accurately reflect the ability of the habitat to support a given ecological community or population. This assumption needs to be clearly stated, supported, and tested. The ability of the models to accurately reflect habitat condition and model sensitivity should be also be tested and documented to ensure that informed decisions can be made.
- ✓ *How model variables are measured* – Understanding how the how ranges of variable measurements for a given ecosystem condition influence model outputs is key to understanding how large a difference between model outputs is meaningful and reflects a change in ecosystem condition (i.e., the model precision). Results will need to be sufficiently accurate to be able to distinguish between alternatives.
- ✓ *Definition and meaning of index or sub-index values* – Documentation should explain how index values are assigned to variables and performance measures, how raw field data are used to calibrate Suitability Index (SI) curves, and how the index values correlate to predicted changes in habitat quality. For example, what percent vegetation is considered to be ideal (i.e., index value = 1.0) and why, what percent vegetation cover is considered to be worse case (i.e., index value = 0) and why, and what is the shape of the curve between those two conditions and why.
- ✓ *How to address climate change issues, when appropriate.*
- ✓ *Definition and consistent use of terms*
- ✓ *Literature references in the model documentation where appropriate and when possible*
- ✓ ***<u>Separate</u>*** *model documentation, tutorials, and spreadsheet/software user documentation* – Model documentation should be stand-alone from any project decision documents (e.g. Feasibility Studies, Environmental Impact Statements, General Re-evaluation Reports) that includes:
    - o Model Version
    - o Developers
    - o Technical Support Contacts
    - o How to use the spreadsheet, prepare input data, model inputs, etc.

Model/Method Testing and Validation

- ✓ *Perform* ***<u>and document</u>*** *model testing and validation* – Perform and document model testing and validation to ensure and support that:
    - o The model performs as expected.

- Model variables and outputs are sufficiently sensitive to measure differences between alternatives and sufficiently robust to yield meaningful results.
- Sensitivity of the model outputs to variability in the model inputs is thoroughly understood.
- The precision and accuracy of model outputs are determined.
- How well measures of habitat suitability correlate with population or community abundances (i.e., performance measures) is determined.

✓ *Link model outputs to ecologically significant conditions* (e.g. species richness).

✓ *Field-verify that model/method results accurately reflect habitat suitability/quality* – Field-verify that index values accurately reflect habitat suitability/quality and are sufficiently robust to small variations in inputs that do not reflect a difference in habitat suitability/quality and sensitive to variations in inputs that do reflect a difference. Index values should correlate well with expert opinion and or literature on ecological condition, and if they do not (i.e., low R-square) figure out why and how to fix it.

✓ *Review and update models regularly* (e.g., every five years).

Analytical Requirements

✓ *Use the highest reasonable level of precision* – Use the highest reasonable level of precision when measuring model variables for the greatest accuracy and sensitivity of model outputs. This relates to the balance between data collection and risk. What level of precision is needed to effectively determine habitat quality? The level of precision should be determined based on the expected sensitivity of model outputs for determining differences between proposed project alternatives, and the level of precision needed should be clearly documented (e.g., the level of precision for measuring water depth or elevation). For example, the best level of precision available for determining land elevations, which can translate to flood levels across a floodplain, may be +/- 1 foot. To most wading birds, this level of precision could mean the presence or absence of feeding habitat. However, across relatively flat landscapes, this also means that is the habitat isn't present in one place, it will be present in another (i.e., same habitat and area of habitat will be available either a little closer to or farther from the river), and that level of precision is sufficient to effectively determine ecosystem loss/gain amongst project alternatives. Keep in mind that models should not be overly sensitive to small differences in field data measurements, as differences in results may then be a matter of differences in data collection rather than actual differences in habitat quality.

✓ *Assign quantitative boundaries to qualitative variables* – Quantitative boundaries should be assigned even to qualitative variables to improve model sensitivity to differences in variable scores and reduce subjectivity, which can lead to differences in results between users. Even when Suitability Index values assigned to field variables are categorical, quantitative ranges should be assigned to each category. For example, "relatively sparse vegetation" could easily be assigned as a range of densities rather than leaving it to the discretion of the user. The same is true for defining “relatively short” and “relatively tall” crop cover.

- ✓ *Use sufficient data to represent all sets of reference conditions* – More than one state, or set of conditions, can be considered “optimal” for the habitats being evaluated and all sets of reference conditions should be characterized for comparison with restoration sites.

Model Spreadsheets and Software

- ✓ *Include the version number and date, as well as the names and contact information for the developers* – It is also important to make sure that the software/spreadsheet version matches the documentation version.
- ✓ *Document all revisions to the model* – When revisions are made to spreadsheets or software, document what changes were made and why.
- ✓ *Keep separate and protect model calculation worksheets* – Either protect model calculation worksheets from accidental or intentional user alteration, or, in cases where exploratory analysis may be desirable, have the system indicate when outputs are results produced by an altered state of the model.
- ✓ *Keep model spreadsheets/software as simple as possible:*
    - o Eliminate redundancies (i.e., don’t have the same calculation in more than one cell of a spreadsheet. This could result in errors when spreadsheets are revised or updated if the calculations are changed in one cell but not another, and makes review more complicated than it needs to be.)
    - o Remove any vestigial inputs and calculations that aren’t being used by the model (i.e., provide users with a clean copy). This improves transparency by allowing a user to more clearly understand how outputs are being calculated.
    - o Make use of defined names, avoiding explicit cell references, especially when switching sheets, to allows user, developer, maintainers, and reviewers to clearly and easily understand what is going into a calculation. This makes calculations more readable (e.g., the defined name “hydroperiod” has more meaning than B$8).
    - o Use model worksheet architecture that is aligned with the input data worksheet format (either both horizontal or both vertical).
    - o Avoid computational modularity violations (i.e., compute a quantity in one place only and make multiple references to the computation rather than having the same computation in multiple places).
    - o Make use of programming libraries (collections of pre-written code and subroutines, classes, values or type specifications). For example, rather than naming each month in a Get Month statement, use Java’s already-defined month string.
    - o Design the user interface in a way that is familiar to most users (i.e., similar to Microsoft products) and makes use of visual cues. Model

users will generally be most familiar with the design of Microsoft products, and most familiar with Excel. It has been suggested that models be fully implemented in Excel, which supports development of closed environments (an organizational requirement), rather than importing from and exporting to Excel from Microsoft Access. Visual cues such as different colors and fonts can help users find input data and model outputs more easily. Also, have graphics and report windows behave as expected, for example, be able to be resized and closed using resize and close buttons rather than double-clicking.

- Use a template engine for developing model code to make the programming language more compact and flexible and making code review and revision easier.
- Add comments to the code explaining the developer’s intentions.
- Develop the model as a spreadsheet or relational database when possible (this can be done as a closed environment).
- Make use of spreadsheet links to ensure that changes made to the data or formulas are updated throughout the spreadsheet.
- Provide clear instructions within the spreadsheets/software to guide the users IN ADDITION TO the user guide to make it easier for users to know what to do.
- Provide users with options for input variables that have specific values or sets of values rather than allowing a user to enter any value.
- Build in error checks and set up warnings to alert the user when inputs or erroneous or out of range.
- Design the model spreadsheet or software to accommodate as many alternatives as desired.
- Allow the user to define filenames and file locations so that the user does not have to navigate to the project folder after each model run. This also prevents accidentally overwriting files that are assigned names by the system.
- Make model outputs transparent and easy to find and understand (e.g., make use of colors or different fonts and that units are included in outputs.

Unique but Notable Recommendations

The following are recommendations that were unique to one review but are noteworthy for consideration in the development of future models.

- ✓ For Habitat Suitability Index (HSI) type models, carefully consider:
  - Which species, guilds or life history stages should be represented?
  - What are the most appropriate parameters to represent habitat quality?
  - What is the level of precision needed for model parameters?
- ✓ Develop base Suitability Indices on raw field data. Do not average data across locations for developing base Suitability Indices, but rather develop multiple Suitability Indices for locations within a given habitat to capture variability in conditions and then average.
- ✓ Provide a detailed technical rationale for the number of years and range of years used to characterize hydrologic conditions.
- ✓ Develop performance measures that consider changes in critical model variables for highly dynamic environments (e.g. fluctuations in salinity that can affect wetland condition).
- ✓ Include variables that address the importance of habitat patch size and shape, habitat corridors, and habitat connectivity.
- ✓ Be cognizant of rounding issues.
- ✓ Account for differences in feeding guilds and season in waterfowl, shorebird, and fish HSI models.
- ✓ Make results from alternative model assumptions tamper-evident.
- ✓ Create a Developers' Guide.
- ✓ Develop a Test Plan and Release Plan