Response to ISRP 2002-13 Mainstem and Systemwide Preliminary Review

ProjectID: 35045

Modeling and Information Management System to Assess Effectiveness of Alternative Actions

Pacific Northwest National Laboratory

Comment:

"The goal of building a complex model to allow alternative actions to be evaluated without experiments is laudable."

"Models may predict cause-effect relationships but to establish such relationships requires links to empirical data."

Response: While the proposed project will not explicitly conduct experiments to establish cause-effect relationships, it is designed to fully exploit studies that will elucidate physical and biologic responses to specific actions. There are 3 sets of study results that will be utilized in developing functional relationships between action effectiveness; watershed characteristics; and hydrologic, water quality, and biological responses at different scales:

- RPA 183 calls for a series of action effectiveness studies as the basis for evaluating the benefits of habitat off-site mitigation. These include at least three studies for each ESU with at least two studies focused each major category of habitat off-site mitigation (e.g., screen irrigation diversions, remove blockages, reduce sediment, etc.).
- Results of status monitoring, related action effectiveness studies, and critical uncertainties research underway across the Basin, and
- Ongoing work in the two proposed pilot demonstration watershed selected by the proposed project team (i.e., South Fork Salmon River and Methow River).

The modeling and data management system will be constructed to maximize the "value of information" (VOI) gained from these studies. Accordingly, the proposed system is designed around three essential functions:

1) Scaling (e.g., ability to capture understanding in the form of cause-effect relationships that can be transferred spatially and temporally across reaches, basins and provinces).
2) Hypothesis Testing (e.g., facilitate the adaptive management process of explicitly implementing actions with postulated outcomes, monitoring effects, revising conceptual and numerical models, etc.)
3) Information Exploitation (e.g., facilitate via the internet the sharing of information and knowledge).

Comments:

"The difficulty is in the details of building a model that is realistic enough to be useful, yet tractable for solution."

"The proposal does not provide a clear picture of how the model building will be done."

Response:

One of the goals of this project is to enable broader and more effective use of existing models within the region - not to build "a model." By providing a framework to link models and share data, the region will be able to use models and data more efficiently. Currently, many state-of-the-science models and data sets are underutilized in regional analyses. The underutilization is not a result of limited computational resources, but of limited accessibility and experience with the models and associated data sets. Additionally, analyses are often repeated (i.e., redone) due to a lack of regional "memory." The proposed framework will provide a platform for capturing, archiving, and sharing analysis results as they are completed.

The model/data linkage is managed via a Metadata Database. While the data and models remain distributed (decentralized), the metadata database provides a centralized directory to the models and data. This database established the input and output requirements of each individual model. A Task Scheduler ensures any inconsistencies in the cascade of inputs-outputs are identified and manages the sequences and execution of the models.

While the proposed assessment framework will link and execute models, it does not guarantee that the existing state-of-the-science models adequately capture all relevant physical and biological processes. However, the framework does explicitly support identifying and prioritizing future model development and data needs.

Models and data provide only a limited understanding of reality. But more importantly, they provide a documented and explicit representation that can be adapted to new understanding. This project does not attempt to create the "ultimate" model but instead will provide a framework populated with an initial, but easily adapted suite of models that encourages the overall assessment capability of the region to improve rapidly and efficiently.

Comment:

"The response should provide more details such as exactly how neural networks and fuzzy logic will be used to obtain models."

Response:

Artificial neural networks, fuzzy logic, and genetic search algorithms are part of the information technology field called evolutionary computing. These techniques have been used for over 30 years in industrial applications because of their reliability and robustness. While we expect to utilize evolutionary computing methods, we will also opportunistically utilize other established approaches, particularly statistical classification methods.

To achieve the 'information exploitation' function, the proposed framework must provide tools to identify patterns and anomalies within and between various datasets and model results. By identifying patterns, data and process knowledge developed in one watershed can be transferred to other watersheds with similar characteristics. For instance, patterns in topography often can help define the stream channel properties. By identifying anomalies, the limitations of models and the limitations of our ability to transfer knowledge between watersheds can be established. For instance, identifying two nearly identical watersheds with high and low fish survival rates, suggests some additional process or attribute needs to be identified to resolve the anomaly.

Artificial neural networks (a.k.a. neural nets) are able to rapidly identify patterns in very large, sparse (i.e., incomplete), multifaceted datasets that are typical in natural resources. Similar to regression analysis, neural nets are model free (i.e., they impose no prior structure on relationships between data). For example, a neural net can be trained with existing data sets to discern cause-effect relationships between physical process changes (e.g., altered agricultural practices) and performance of particular restoration strategies under given watershed conditions.

Fuzzy logic provides tools to incorporate ambiguous or vague process understanding (due to data gaps or conceptual model limitations) into cause-effect analyses. This is particularly useful when interfacing physical process models and ecological models such as EDT where the structural differences (i.e., deterministic continuum mechanical processes vs. rule based categorical states as inputs, respectively) are so significant. PNNL researchers have recently developed a prototype fuzzy logic-based tool (FuzzyHab) that estimates habitat diversity using fuzzy rules where the inputs are multi-valued (spatial and temporally varying) inputs from physical process models. The fuzzy rules were developed from a set of rules (Bio-Rules) developed earlier by a group of aquatic biologists. The fuzzy rules provide a compact and flexible representation consistent with intuition and more accommodating to pervasive imprecision. These fuzzy rules were able to efficiently generalize the Bio-Rules and flexible enough to accept the lack of sharp boundaries typical in categorical statements about biological systems.

Comment:

"The response should describe how the project could contribute to the CBFWA (or other) systemwide design-based M&E projects by providing modeling aspects making predictions based on data."

Response:

The proposed project couples the development of an advanced model and data management system with comprehensive, site-specific pilot implementations in the South Fork Salmon and Methow watersheds. The pilot implementations will establish cause-effect relationships needed to assess action effectiveness at multiple temporal and spatial scales. These relationships will be based on physical and biological processes as reflected in study results and physical/biological process models. Using available

information, these results will then be transferable to other locations based on site-specific data (e.g., watershed and habitat characteristics). Products of this project will include the fully operational system, complete transfer of MADIMS to the WDFW and IDFG along with appropriate documentation and training. In addition, the project will further develop and document methods for scaling (up and down), information transfer across basins, as well as fully operational Internet-based data management tools to facilitate data analysis, visualization, and model input/output stream processing.

Comment:

"This is an expensive project with a budget that is suspiciously rounded."

Response:

The proposed project's goals are ambitious. However, the benefits to the region in having a capability to consistently and adaptively capture and deploy cause-effect relationships between alternative actions and responses will pay increasing dividends over time in identifying the most promising actions for particular conditions and avoiding those that are less effective.

Commensurate with the ambitious goals, the complexity and required effort are also significant. One of the approaches used to control costs was to develop a diverse team wherein particular tasks could be allocated to those best suited within the region to accomplish them. In addition, leveraging ongoing work and associated cost sharing was a priority.