REVIEW OF TERRESTRIAL MONITORING PROTOCOLS FOR THE GRAND CANYON
Report of the Protocol Evaluation Panel (PEP)

Submitted February 4, 2008

INTRODUCTION

Panelists

Loretta Battaglia, Assistant Professor, Department of Plant Biology, Southern Illinois University

Darold Batzer, Professor, Department of Entomology, University of Georgia

Geneva Chong, Research Ecologist, US Geological Survey, Northern Rocky Mountain Science Center

David Cooper, Senior Research Scientist, Department of Forest, Rangeland and Watershed Stewardship, Colorado State University – *Panel Chairman*

Andrew Hudak, Research Forester and Landscape Ecologist, US Forest Service, Rocky Mountain Research Station, Moscow, Idaho

Michael Young, Research Fisheries Biologist, US Forest Service, Rocky Mountain Research Station, Missoula, Montana

Charge to Panelists

"Review and critique the existing effort/findings and recommend long term core monitoring protocols/methods that will meet the established Core Monitoring Information Needs associated with terrestrial biological resources (Goal 6)."

For this report, the PEP provides an assessment of past monitoring efforts and recommends long term core monitoring protocols to meet the Management Objectives and Core Monitoring Information Needs associated with terrestrial biological resources including vegetation and animal communities. Because the PEP identified some management objectives and monitoring needs that were incompatible, the PEP also included critiques of these sections where relevant; these are addressed in the introductory portions of the vegetation and animal monitoring sections. The section on vegetation monitoring focuses on remote sensing, in situ sampling, and analyses of both, whereas the section on animal monitoring reviews monitoring by taxonomic group and notes an assortment of monitoring design and analysis concerns. We urge that information obtained within Grand Canyon be compared to reference river reaches elsewhere in the southwest to improve the strength of inferences about the effects of dam operations. We conclude by encouraging the program to strengthen the research culture and develop a more open science process in which basic, peer-reviewed research guides monitoring efforts. The fluvial geomorphology research program is an excellent model to follow as it produces competition, collaboration and excellence in research vital to management.

VEGETATION MONITORING

Goal 6 and associated Management Objectives (MO) and Core Monitoring Information Needs (CMINs) Related to Vegetation Monitoring

Below we reiterate the stated goal, management objectives, and information needs provided to the PEP before the July 2007 meeting that we used as the basis of our evaluation. The PEP, however, extended its critique to the management objectives and information needs; these issues are listed parenthetically. We recommend the development of a research program that supports and guides the monitoring program, similar to what was recommended by Urquhart et al. (2000). Activities ranging from sample site selection to creation of vegetation maps and models and analyses of remotely sensed data all benefit from execution within a research-driven framework. For example, one of the questions to the PEP was how frequently should vegetation monitoring occur. We feel that question needs to be answered using existing data in an adaptive management framework (i.e., as data are acquired and analyzed the recommended frequency could change based on improved understanding of the system. Alternatively, the frequency could be altered based on hypothesized responses to event such as experimental high flows, wildfire, drought, non-native species removal, etc.).

> GOAL 6: Protect or improve the biotic riparian and spring communities within the Colorado River ecosystem, including threatened and endangered species and their critical habitat. *(PEP comment: The spring communities are not directly influenced by dam operations and management, and we recommend the removal of this community from discussion. The threatened and endangered species appear to be monitored by other entities, so they should also be removed from this discussion, but the GCMRC should collaborate with these other entities for needed data or analyses.)*
>
> MO 6.1: Maintain marsh community abundance, composition, and area in the Colorado River ecosystem in such a manner that native species are not lost.
> CMIN 6.1.1: Determine and track the abundance, composition, distribution, and area of the marsh community as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
> (*PEP comment: Many marsh communities in the Grand Canyon are largely an artifact of river regulation. The reduced peak flows and higher base flows have created relatively stable hydrologic conditions favoring marsh development, and their overall abundance has likely increased relative to the pre-dam era. The PEP recommends that managers reconsider maintenance of these anthropogenically created habitats as a priority*.)

MO 6.2: Maintain New High Water Zone (NHWZ) community patch number, distribution, composition, and area to be no lower than values estimated for 1984.
CMIN 6.2.1: Determine and track the patch number, patch distribution, composition, and area of the NHWZ community as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
(*PEP comment: the maintenance of a number of patches and distribution does not appear realistic given the current nature of the riparian system. For example, variation in magnitude of management flows may not be sufficient to maintain a dynamic riparian system; hydrologic stabilization and flow reduction have allowed dominance of non-native species such as tamarisk and annual grasses. These non-native species may pose significant risks to the structure and function of the native riparian community (e.g., increased risk of fire, competition for space, and stabilization of soil features that were historically more dynamic). Thus, the PEP suggests that tamarisk is not a desirable component to be maintained in the NHWZ. Detection, eradication, and monitoring for non-native species should be incorporated into the vegetation sampling and monitoring scheme. The PEP recommends vegetation restoration, including tamarisk (and other invasive non-native species) removal. If actions are taken to restore native species composition and structure, monitoring should be directed to evaluate the effectiveness of such actions for use in an adaptive management framework.*)

MO 6.3: Maintain Old High Water Zone (OHWZ) community abundance, composition, and distribution in the Colorado River ecosystem.
CMIN 6.3.1: Determine and track the abundance, composition, and distribution of the OHWZ community as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
(*PEP comment: Under the present flow regime, this goal is not attainable without planting, irrigation, and land management. The older mesquite and tamarisk in this zone will die within the coming decades and there is no recruitment of these species. This zone could only be maintained "naturally" by managed floods exceeding 90,000-100,000 cfs. If actions are taken to restore native species composition and structure (again, emphasis on native species), monitoring should be directed to evaluate the effectiveness of such actions for use in an adaptive management framework.*)

MO 6.4: Maintain sand beach community abundance, composition, and distribution in the Colorado River ecosystem at the target level.
CMIN 6.4.1: Determine and track the abundance, composition, and distribution of the sand beach community as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
(*PEP comment: Historically the sand beaches were essentially bare, maintained by high spring flows, and lacked developed vegetation. The PEP recognizes that determining an appropriate "target level" is challenging in a system that has*

moved to an alternative, stable state. Maintaining some of the features of the pre-dam system is desirable; where feasible, some vegetation reduction may be warranted on beaches. If actions are taken to reduce/remove vegetation, monitoring should be directed to evaluate the effectiveness of such actions for use in an adaptive management framework.)

MO 6.5: Reduce invasive non-native species abundance and distribution.
CMIN 6.5.1: Determine and track the abundance and distribution of non-native species in the Colorado River ecosystem as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
(*PEP comment: Because of the large effort necessary, natural resource managers should determine which nonnative species can be controlled or eradicated, and prioritize where such efforts will be directed. Tamarisk is highly invasive in the Grand Canyon and can change ecosystem functioning; the PEP recommends that reduction of this species should be a high priority. Tamarisk eradication efforts are ongoing in the Grand Canyon but have not targeted the Colorado River corridor. If actions are taken to reduce and control non-native species, monitoring should be directed to evaluate the effectiveness of such actions for use in an adaptive management framework.*)

MO 6.6: Maintain seep and spring habitat in the Colorado River ecosystem
CMIN 6.6.1: Determine and track the abundance, composition, and distribution of spring and seep communities as measured at 5-year or other appropriate intervals based on life cycles of the species and rates of change for the community.
(*PEP comment: These habitats are largely independent from dam operations, and it appears that other agencies are now targeting seeps and springs in the Grand Canyon. We recommend that sampling for this CMIN be discontinued by GCMRC, and instead rely on information from other agencies.*)

MO 6.7: Maintain riparian habitat in the Colorado River ecosystem capable of supporting southwestern willow flycatcher.
CMIN 6.7.1: Determine and track the abundance, distribution, and reproductive success of southwestern willow flycatcher in the Colorado River ecosystem. (*PEP comment: It appears that other agencies are now targeting the southwestern willow flycatcher in the Grand Canyon. We recommend that sampling for this CMIN be discontinued by GCMRC, and instead rely on information from other agencies. Please also see the comment under Animal Monitoring.).*

Terrestrial Monitoring Overview

We evaluated efforts by Kearsley et al. (2006) to address CMINS associated with vegetation monitoring related to Goal 6. Work conducted by Kearsley et al. from 2001 through 2003 was positively influenced by an earlier review panel (Urquhart et al. 2000). Urquhart et al. (2000) made the major recommendation to develop a comprehensive monitoring strategy with four components: (1) endangered species assessments; (2) monitoring for model development; (3) inventory of plants and animals; and (4) long-term monitoring of the main-stem corridor. Seven additional, specific recommendations were: (1) provide scientific direction to the monitoring program; (2) require integration/cooperation across disciplines through the RFP process; (3) ensure long-term comparability of data through protocol development and training; (4) incorporate a long-term perspective into monitoring plans and methods; (5) develop a terrestrial ecosystem model with the same detail as the current aquatic model; (6) determine appropriate level of activity related to the supposed Kanab Ambersnail; and (7) prioritize completion of a GIS and a land cover map, a random sampling approach, cost effective response indicators and protocols, and protocols to ensure inter-disciplinary cooperation.

From one perspective, Urquhart et al.'s (2000) recommendations were made to the Biological Resources Program overall, and we did not receive sufficient information or requests to evaluate the Program as a whole. However, some recommendations appear to require attention, particularly in relation to developing a research approach to monitoring. For example, major recommendation (2), monitoring for model development, still needs to be addressed most fundamentally through a research-driven monitoring program, which was specific recommendation (1), above. Other specific recommendations (e.g., 5: develop a terrestrial ecosystem model) would be greatly facilitated by focusing on specific recommendation (7), completion of a GIS. Therefore, even though we were mainly tasked with reviewing Kearsley et al. (2006), we feel it is critical that the recommendations from the 2000 PEP not be forgotten.

Kearsley et al. (2006) have responded to the majority of the recommendations in the 2000 PEP. For example, Kearsley et al. (2006) were unable to select their sampling locations based on a vegetation base map, as recommended by Urquhart et al. (2000), because that work did not commence until 2002, and it has not been finalized at the writing of this review (Ralston et al. in preparation). In the following discussion related to terrestrial vegetation monitoring, we discuss recommendations to address three related needs for CMINs 6.1.1 – 6.5.1:

1) Vegetation mapping: develop a repeatable, accurate approach to create vegetation maps that can be compared over time to monitor changes in area and location of vegetation types and to provide a base layer for designing terrestrial monitoring and analyzing terrestrial monitoring data.

2) Terrestrial vegetation monitoring: develop a repeatable, accurate approach to quantify changes in native and non-native plant species composition and structure (cover and

height) and their interactions with other taxa, and provide ground-truth data for vegetation mapping.

3) Habitat analyses: develop a repeatable, accurate approach to quantify interactions between abiotic and biotic habitat characteristics and associated vertebrate and invertebrate animals or taxonomic groups of interest.

Urquhart et al's (2000) and Kearsley et al.'s (2006) recommendations for improved terrestrial inventory and monitoring methods included the development of a complete vegetation base map that would facilitate their second recommendation of random sampling. Ralston et al. (in preparation) describe the objectives for use of one such base map: "The primary objective of our current mapping project was to develop a digital inventory map of vegetation to enable patch-scale and landscape scale change detection, and to establish randomized sampling points for ground surveys of terrestrial fauna (principally but not exclusively, birds)." Once this initial map is finalized, we agree that future vegetation data should be collected to allow comparisons with previously collected data, but we do not believe that the same methods should be used if better and more widely used ones exist.

Vegetation Mapping

A system for creating vegetation base maps is needed to quantify changes in location and area of vegetation types over time. We believe the process and database (map) described in Ralston et al. (in preparation) provide an adequate baseline to guide near-term sampling efforts and quantify the extent of several vegetation types at the time the data were collected. However, changes in available technology since that work was initiated in 2002, and changes expected in the future, remind us that the vegetation mapping process is not static, and improvements should be made continuously. For example, it may be more appropriate to sample in the same way as the National Park Service and map canyon vegetation as part of the Grand Canyon National Park vegetation mapping project (e.g., http://www1.nature.nps.gov/im/units/scpn and http://biology.usgs.gov/npsveg/index.html). In addition, other methods may be more suitable for correlations with and corrections of remotely sensed data (see remote sensing discussion, below) or for analyses of vegetation communities (see statistics discussion, below). Improvements can even be made in how existing data are analyzed and interpreted. The following review provides guidance regarding the remote sensing components of a vegetation mapping program.

Remote Sensing Component

The Grand Canyon presents a challenging environment for remote sensing applications, given the extreme topography. As such, the utility of remote sensing for monitoring must first be tested within the research realm. The logistical difficulty of monitoring terrestrial ecosystem processes on the ground makes remote sensing an appealing tool for decreasing reliance on field work, although no remote sensing tool should ever be considered a complete substitute for field-based monitoring. The goal should be to leverage the remote sensing data to add value to the field data, to the extent possible.

The two types of remotely sensed data that have been tested for monitoring capability in the Grand Canyon have been airborne high resolution digital imagery and airborne discrete-return light detection and ranging (lidar) data. Airplane pilots can plan their flight lines to coincide with times that the canyon will be sunlit. A minor complication with sensors on fixed-wing aircraft, as with passive optical imagery, is in having to mosaic multiple straight flight lines arrayed along the twisting path of the river (e.g., Ralston et al., in preparation). The idea of using a helicopter-based platform was suggested during the panel discussion, and should be explored. A helicopter could efficiently follow the winding route of the canyon, or hover above monitoring sites of interest for higher density sampling. This would efficiently reduce the flight time and data required to produce complete coverage, and probably cost as well. We do not see much potential for current space-borne platforms in the Grand Canyon because the spatial resolution of most satellite sensors is too coarse, and, even the more fine resolution sensors such as IKONOS and Quickbird, are constrained by the narrow time window of satellite overpass, which would result in large portions of the canyon being in shadow. We recommend evaluating the use of helicopter-mounted lidar with simultaneous acquisition of digital imagery during optimum lighting conditions.

The use of hyperspectral imagery was also mentioned during discussions, but hyperspectral data tend to be expensive to acquire and analyze. As with any passive optical imagery, data collected along flight lines obtained at different times of the day would have varying orientations with respect to the sun angle and thus different albedos. This would be a formidable problem in the Grand Canyon that would require advanced techniques to correct (Kennedy et al. 1997). If this source of variation between flight lines can be effectively removed in the calibration-to-reflectance process, then hyperspectral imagery, with upwards of 200 contiguous bands spanning the visible, near infrared, and short-wave infrared (SWIR) regions of the electromagnetic spectrum, would allow more subtle discrimination of vegetation types than is possible with only the blue, green, red, and near infrared bands that typify most digital imagery. This could be useful for producing a more detailed vegetation map because varying amounts of lignin in plant tissue causes variation in the SWIR that can allow species-level differentiation of some species in some systems. Distinguishing tamarisk from mesquite, marshes, and the other vegetation types should be easy, while species-level discrimination would be hard. Spectral mixture analysis is probably the most robust method available for estimating the fractional cover of the response variables. It is attractive because it is applied at the subpixel level, providing a direct estimate of fractional cover that is a biophysical variable, which is more intuitive, interpretable, and useful to the user than, say, a vegetation index (e.g., NDVI) that must be related to the cover attribute of interest, with an added error component.

Airborne Imagery

The Ralston et al. (in preparation) vegetation communities selected as classes for mapping seem reasonable based on the data collected from the field, and we liked the cascading series of criteria, including fuzzy logic, for assessing classification accuracy. However, we think it is important to decide and define what level of classification accuracy is required to meet monitoring objectives. While the fuzzy logic may increase accuracies for the purpose of a report, it may not be helpful for satisfying actual monitoring requirements, and we address preferred alternatives for plot size and location and data analyses in the vegetation sampling section.

We had difficulty interpreting Figure 1 in Appendix A: Image Processing (Ralston et al., in preparation). We are skeptical that the “synthetic NIR” band can realistically replace the unfortunately saturated NIR band. Leaf cell water content is the primary factor causing high reflectance in the NIR, while chlorophyll absorption drives reflectance in the red and green channels. We would expect this relationship is nonlinear, not linear as is the NIRpan transformation equation at the bottom of p. 52. Is there a reference that could be cited? Is there other imagery available that could be employed besides these data with the miscalibrated NIR channel? We have the impression (from our visit, not from Ralston et al.) that quite a bit of remote sensing data have been accumulated by GCMRC (or related entities) within the past decade, and it would be beneficial to evaluate if any historical data are useful given currently available analysis techniques.

We liked the use of texture images as additional bands. We think there is good potential to improve classification accuracies further by incorporating topographic, geomorphic, and vegetation zone (OHWZ, NHWZ) variables as predictors in a classification scheme. RandomForest is a powerful classification tool that bootstraps the data to produce a more robust model. The topographic predictors could be extracted from lidar data if a better lidar data processing stream was followed (which we’ll discuss next). So we think there is a great opportunity here for integrating image and lidar data into producing a baseline vegetation classification map and formulating a remote monitoring strategy.

Airborne Lidar

The extreme topography in Grand Canyon makes any remote sensing problematic, but increases the appeal of lidar, especially to preclude the effects of shadowing from the sun. The positional accuracy of the 2005 lidar data is less than desired, or required for geomorphological monitoring. We believe some of these issues are related to processing. First, the ground height estimated by lidar appears lower (Fig. 2; Ralston et al in preparation) than estimated by the total stations used for independent validation of ground height. We believe this discrepancy may be the result of a vertical datum mismatch, since it says in the Methods that the NAD83 GPS observations were not converted to NAVD88. This is not something that can be accomplished in ArcInfo. Two programs that can be used for converting vertical datums are CORPSCON from the Army Corps of Engineers (http://crunch.tec.army.mil/software/corpscon/corpscon.html) and VERTCON from the National Geodetic Survey (http://www.ngs.noaa.gov/TOOLS/Vertcon/vertcon.html). Another issue is the classification of points as ground or vegetation returns. The deconvolution of ground and vegetation in topographically complex terrain is difficult. For instance, where dense vegetation cover occurs, there may be over-classification of vegetation returns as ground returns. This would lead to erroneously high estimates of the ground surface, and thus underestimation of vegetation height. Thus, poor estimation of ground height is inextricably linked to poor estimation of vegetation structure. However, we doubt that the vegetation is too thick for there to be adequate ground returns, because sufficient canopy penetration has been demonstrated in high-biomass forests with much higher leaf area index than could be approached in the arid Grand Canyon ecosystem, even in the thickest tamarisk stands.

Different lidar vendors use different proprietary algorithms to accomplish the necessary task of filtering vegetation returns, which is prerequisite to producing a bare-earth digital elevation model. The GCMRC apparently did not receive a bare-earth DEM from the vendor, which is very surprising, and strongly suggests poor classification of the ground returns (the recently published Evans and Hudak (2007) algorithm (which runs in ArcInfo) may be worth using). We suspect that the existing lidar data are adequate for creating a more useful bare earth surface model, for example, with improved analysis approaches, and we recommend exploring these options further to benefit more fully from the initial cost of the data acquisition.

Summary and Recommendations Regarding Remote Sensing

It is unrealistic to expect that, if only the spectral resolution were finer, individual species could be identified, or if only the spatial resolution were finer, individual plants could be quantified. The major emphasis for monitoring should remain on the field data component, which we discuss in the next section. Topographic (e.g., elevation, slope, aspect), geomorphic, and other predictor variables (e.g., solar insolation) explain much of the variability in observed species distributions, as quantified in the field data. There are a variety of predictive modeling and mapping tools (e.g., RandomForest, Breiman 2001, Breiman et al. 2006), general additive models (GAMS; Moisen and Trescino 2002), and imputation (Moeur and Stage 1995, Ohmann and Gregory 2002) that are well-suited for exploiting these relationships between dependent response variables gathered at discrete points in the field and independent predictor variables derived from spatially continuous remote sensing band layers, topographic layers, etc. It is critical that the field data be accurately geolocated, to minimize errors due to misregistration between the field and remotely sensed data. Time of acquisition of field and remotely sensed data should also match as closely as possible. However, since phenological variation is minimal in the Grand Canyon, temporal misregistration should not be as problematic as spatial misregistration.

Legacy data. We recommend that the GCMRC make a concerted effort to build upon available legacy information for establishing monitoring sites in the field (for both remotely sensed and ground samples), and for developing vegetation and topographic maps (reanalysis of the lidar data). Historical aerial photo data that predate the Glen Canyon dam do exist. The earliest photos were acquired in the mid-1930s. Many of these are too shadowed to be of any use for even visually resolving vegetation at the bottom of the canyon. However, a few vegetated sites are visible, and may even indicate tamarisk invasion at this time (we know that tamarisk established prior to Glen Canyon Dam construction). There are also 1952 photos from Glen Canyon that should be leveraged when establishing monitoring sites within this reach. Dr. Jack Schmidt and his students and staff already have orthorectified the historical aerial photos at select sites within each reach. These sites were selected to meet geomorphic research or monitoring objectives, meaning they should not be expected to bias any vegetation sampling (or arthropods, terrestrial vertebrates, etc.). Biological monitoring sites should be linked to geomorphological monitoring sites to the largest extent possible to facilitate ecosystem-level study. Otherwise, it is much more difficult to relate the biological and ecological processes to the physical processes that have historically driven the functional dynamics of the Grand Canyon ecosystem. Choosing sites based on existence of legacy data need not compromise the landscape-level inference afforded by the generalized random-tessellation stratified (GRTS) sampling strategy (recommended by Urquhart et al. 2000), as geomorphic reach is an important stratification variable that should be employed, and priority should be given to sites within each reach where historical data exist.

While we approve of scanning legacy photos at the native resolution of the scanner (10 microns) for archiving purposes, it is probably overkill to map or monitor at such a high resolution. Extremely high resolution can create as many or more problems than it solves. From an analytical standpoint; the finest resolution is often not the appropriate resolution. Semivariogram analysis can objectively inform the decision of what is the practical and useful spatial resolution for mapping and monitoring. Field plots will be sufficiently spread out along the canyon that spatial autocorrelation will not be a concern. However, spatial autocorrelation is a property in the imagery that can be exploited to improve vegetation mapping. It would be useful to extract the pixel values from within the original TWINSPAN polygons representing the vegetation classes of interest for monitoring (because the TWINSPAN vegetation types have been verified in the field), and generate semivariograms from them. The range (lag distance of the sill) informs the user of the scale of spatial structure captured by the imagery. If the image resolution is sufficiently fine, then the spatial scale of the structural attribute of interest can be characterized from the imagery, and related to that attribute as measured in the field, for mapping. For example, if you can see individual tamarisk trees in the imagery, then the range in a semivariogram generated from these pixels will indicate at what scale tamarisk trees are spaced or clumped. This knowledge helps guide the choice of texture filter window sizes to use to highlight structure attributes of interest (Hudak and Wessman 1998).

There are known problems associated with TWINSPAN (see statistical section below), and the PEP cautions that the continued use of polygons derived from these previous analyses could present problems. We recommend that an appropriate alternative analysis be identified and the data used in previous analyses be reanalyzed and compared to the TWINSPAN output. We strongly recommend that this step be taken before proceeding with products from analyses that may be flawed. If there are no major discrepancies between the sets of output, then switching over to the new technique(s) should be relatively simple. If the different methods do suggest different results, the program may need to re-evaluate the original delineations and conclusions they may have drawn from these methods.

These issues with data analysis highlight the need for a more active research arm of the monitoring program. If the research program is run in parallel with monitoring, there can be a positive feedback loop whereby monitoring is both guided by and contributes to research. Analysis of data (especially those already collected), peer-review, and timely publication of results are necessary steps in this process. This synergism would benefit the overall program and would help to ensure that methods employed, analysis, and conclusions of the monitoring efforts result from periodic feedback from the larger scientific community.

Tom Gushue, who handles the GIS and historical aerial photo library, showed us the polygon coverages from the TWINSPAN plots of Kearsley. Another coverage should be built showing the polygons where 1930s or 1952 historical aerial photos of vegetated sandbars exist along the canyon. The locations of Webb's comparative photo points should also be assembled into a GIS coverage, and incorporated into the GRTS sampling design to the extent possible. Jack Schmidt's geomorphic sites overlap 5 of the 9 TWINSPAN sites, upstream of river mile 90, with just one site per geomorphic reach. Exact study site or plot locations would need to be randomly placed within the legacy polygons to preclude sampling bias. The cumulative number of all sites with some form of image legacy data will be small and thus easy to weave into the larger pools of sites selected following the GRTS sampling design. There are substantially more legacy sites with ground plot data, and new sites (where there is no legacy information) could be viewed as supplementary plots to fill out the GRTS design once all existing plots are located in a GIS.

Plot size (mapping). While a vegetation plot may be representative of the polygon sampled, it is another matter to assume that vegetation structure and composition do not vary spatially within the polygon. Image pixel variation within a polygon can approach, but should not exceed, the pixel variation between polygons. When it comes to empirically relating field plot data to remotely sensed data, the plot scale of measured data inputs for the attributes of interest should be commensurate with the scale of outputting those attributes into a map. Otherwise, you are faced with a scale mismatch that will compromise the accuracy of the vegetation classification. Let the size of the plot be determined by considerations on the ground. It is advantageous if the plots are large enough to include a number of image pixels. Now, you have a number of pixels (and their associated variability) to relate to your field data. The mean of these pixels aggregated within the plot footprint will correlate better with the field attribute of interest than just the pixel value at plot center would.

GRTS Sample Design

Many monitoring programs are based upon investigator chosen sites that are felt to represent the resources being analyzed. However, Urquhart and Kearsley (unpublished manuscript) argue that there are a number of problems with this approach. A probability based sample may yield a better picture of the status and trends of the resource. The distribution of sample plots should be spatially balanced to incorporate a broad range of measurement conditions. One problem that arises with this approach is that the sample frame will include sites that cannot be sampled due to safety, inaccessibility or other reasons. In addition, sample locations must be tied to the major stressors or drivers of resource abundance or distribution.

For the GCMRC Urquhart and Kearsley used a one dimensional generalized random-tesselation stratified (GRTS) design (Stevens and Olsen 2004) to choose transect locations, with the goal of producing a spatially balanced probabilistic sample to collect vegetation cover and species richness data. The GC was divided into 702 segments, as opposed to a continuous length. Sample sites were chosen from the 702 segments, and then a random point was selected within each segment for sampling. A total of 150 points were selected for use, and transects were established in these locations. In Urquhart and Kearsley (unpublished manuscript) a total of 100 sample transects was identified, and transects were classified as being within "wide" or "narrow" canyon reaches based upon gross geologic characteristics of the Grand Canyon. Narrow reaches have granite or limestone bedrock which is resistant to widening, while the wide reaches have shale and other softer bedrock that is more erosive. Urquhart and Kearsley also acknowledge that on a reach scale the geomorphic characteristics of the river and fluvial deposits are controlled by debris fan and eddy complexes. However, the sample sites were not chosen with stratification by geomorphic feature or landform. It is unclear whether the sample transects are identified as being within debris fans, pools, eddys or bars, and what the distribution of transects are within these four major fluvial landforms.

The panel encourages the close integration of riparian ecological monitoring and geomorphic research and knowledge about the Grand Canyon. The importance of fan-eddy complexes for controlling plant establishment and overall vegetation patterns is paramount in river canyons with fan-eddy complexes (Stevens et al. 1995, Cooper et al. 2003, Birken and Cooper 2006). It is unclear whether the "narrow" and "wide" dichotomy captures the larger landscape scale variation in geomorphic, climate and other processes that would affect vegetation composition. This should be tested with the data. Likely the canyon bottom or canyon rim to rim width is a continuum through the Grand Canyon. In addition, the local stream gradient that affects the riparian vegetation is controlled by the transect position within the debris fan-eddy complex, and the geomorphic environment of each transect should be fully understood. This could be quantified by measures of canyon width in the area of each study transect, and also an analysis of river gradient, or geomorphic landform. We recommend that the geomorphic setting within fan-eddy complexes be verified for each transect and an effort made to have an adequate sampling of the vegetation in debris flows, eddy's, pools and riffle environments. In addition, future analyses of vegetation composition and change should be stratified by geomorphic setting as well as elevation above the river. The vegetation of pools will be more similar to each other between wide and narrow canyon reaches, or at different points in the Grand Canyon, than the vegetation of pools and debris flows within the same river reach.

Terrestrial Vegetation Monitoring

Vegetation monitoring provides an assessment of the impacts of Glen Canyon Dam on downstream terrestrial systems (flora and fauna) of the Grand Canyon. Vegetation sampling plots will be located at study sites, selected in accordance with the GRTS scheme, but further distributed within additional strata (e.g., water zones described below) and in conjunction with sampling other taxa (e.g., arthropods, small mammals, and birds). We commend Kearsley et al.'s (2006) approach, and we recommend that their study sites and legacy data continue to contribute to future sample site location selection. In the review below we present options for improving future sample site and plot location, and provide a specific evaluation of Kearsley et al.'s (2006) statistical analyses of their vegetation data and some recommended alternatives to TWINSPAN for Ralston et al. (in preparation). Where feasible, the PEP encourages use of vegetation monitoring data as additional ground-truthing information to support vegetation mapping efforts.

Sampling Frequency

The decision of what is the appropriate temporal sampling frequency should be defined by objective analyses of the available historical data, as well as what is practical and useful. At the minimum, there needs to be some consideration of the rate of vegetation change since dam operations commenced, relative to the ability to detect that change at a given sampling interval. If, for example, woody cover can be estimated in the field to an accuracy of +/- 5%, while the rate of detectable woody cover change is 1%/year, then it makes little sense to monitor woody cover change annually. More could be gained by resampling 100 reference plots every 5 years than by resampling 20 reference plots every year, for the same effort. Thus, the PEP recommends a sampling cycle guided by research. For example, research could demonstrate a fixed appropriate frequency (e.g., 5 years), however, if events occur that accelerate changes (e.g. prolonged drought, flood release, etc.), then sampling should be conducted opportunistically. Further, the PEP strongly recommends that planned high flow events be bracketed before and after by sampling. Additional off-cycle sampling in the first and second post-event years will also be required to capture lags in response, for example. Determination of sampling frequency is a critical component for research.

Plot Location

New study sites should be located based on the GRTS discussion, above. If the GCMRC should permanently mark GRTS study sites, they should be marked with two pins; one at the highest elevation and a separate benchmark. Relocations should be made using surveying. At each study site, sample sites should be located based on existence of legacy data (spatial and plot) if any exist. For example, Kearsley et al. (2006) established sampling sites based on three zones: the old high water zone (OHWZ, ≥ 90,000 cfs stage elevation), the new high water zone (NHWZ, approximately 25,000 - 90,000 cfs), and shoreline area (water's edge to the 25,000 cfs stage elevation). These zones are meant to represent the hydrologic gradient imposed by stream flow. These sites and strata are still useful for future sampling because of their legacy data.

However, it is unclear whether designation of these zones (and the 15k, 25k, 35k, 45k, and 60k cfs stage elevation points for the Kearsley et al. (2006) vegetation dynamics study) takes into account stage elevation changes that may occur from scouring and sediment deposition during high flow events. If this is indeed an omission, it may be a problem for interpreting patterns of change. Particularly for stage elevation classes (e.g., shoreline and NHWZ) where the elevations may change dramatically between consecutive sampling periods, the contributions of sediment inputs/outputs should be captured and related to any shifts (or non-shifts) in the vegetation. Study sites should be randomly selected x-y points, but the actual ground location of plots in a study site (e.g., the distance from the water's edge - the z dimension) should be allowed to vary as sediment is eroded or deposited and the shoreline moves; however the same x and y points should be sampled each year. The Kearsley et al. (2006) approach sampled points for vegetation based upon an elevation, which with sediment deposition could shift the points up or down the slope, resulting in a different location for plots in subsequent years. For long-term monitoring we suggest monumenting study site centers with rebar or PVC so that they need not be relocated at each visit. Where it is feasible for plot data to be used in conjunction with remotely sensed data, plots need to be accurately geolocated (e.g., accurate to within half the diameter of the plot). Plots do not need to be monumented for purposes of ground-truthing, but plots selected for repeated sampling as part of the GRTS design do.

The hydrologic gradient is often steep in the canyon and therefore the zones may be spatially compressed at some sites but more spread out at others. There is some concern that the vegetation plots may encompass more than one zone at some sites. For example, in the vegetation structure and composition chapter of the Kearsley *et al*. report (2006), the plots were of 3 m radius around the pitfall traps. Can we be assured that the plots did not straddle more than one zone? This sampling issue is less of a concern in the vegetation dynamics study where a 1m x 1m sampling frame is positioned at a given stage elevation class using surveying techniques. Vegetation within the frame is then sampled at each stage class, and the frame is moved along at the same elevation parallel to the river such that two samples are taken on the upstream side and two are taken on the downstream side of the river to OHWZ transect line. The resulting 4m strip of $1m^2$ plots is used to characterize that particular stage elevation class. Although these strips seem appropriate for sampling vegetation in narrow stage elevation classes (see additional Recommendations Regarding Vegetation below) characteristic of steep riparian gradients in the corridor, a summary of the elevation changes along each of the river to OHWZ transects would be useful for interpretation. Some evidence should be given that the same elevation classes are being sampled across sites and that the plots are not encompassing different elevation ranges.

Statistical Analysis

Statistical analysis by Kearsley et al. (2006) using ANOSIM is an appropriate technique for comparing species composition across years or among zones (or among major fluvial landforms if desired). The main limitation of this technique and others similar to it (*e.g.*, Multi-response Permutation Procedures- MRPP) is that interactions between the factors cannot be tested. In this case, interactions between zones and years may very well be important to assess for repeat panels. An alternative is to use the permutational dissimilarity-based method called NPMANOVA (Anderson 2001). The approach is similar to ANOSIM in that a variety of dissimilarity measures can be used as the foundation, but it enables testing of more complex factorial and nested models when using multivariate (compositional data). Programs for implementing this method are free and available for downloading on Dr. Marti Anderson's website (http://www.stat.auckland.ac.nz/~mja/). An add-on will soon be available as part of the user-friendly software package PRIMER.

Comparisons between zone assemblages over time are needed to fully evaluate compositional trends. There may be differences among zone assemblages, even if long-term changes in the environment are occurring. For example, if there is a drying trend, we might expect all zones to have a consistently more xeric species composition, although the NHWZ may become more similar to the historic OHWZ with time. *Directionality* in composition should be investigated as part of the long-term monitoring effort, in addition to simply testing for differences among the zones/years. For example, is the OHWZ becoming more xeric and increasingly dissimilar to the other zones and its historic condition? Is the NHWZ becoming less dissimilar to the historic OHWZ (*i.e.*, convergence)? We recommend coupling the tests for differences among zones/years with an evaluation of the trajectories of these zones through time using non-metric multidimensional scaling.

The diversity analysis appears to include both native and non-native plant species. As one of the management objectives is to reduce non-native species, a separate diversity analysis with just natives is warranted. Non-native species richness should not be included in measures that might be considered positive indicators of ecosystem integrity. While it is recognized that tamarisk is considered to be an important component of this somewhat novel community, especially in terms of habitat structure for birds (e.g., willow flycatcher), including non-natives in measures of community health conflicts with established management objectives.

Two-way Indicator Species Analysis (TWINSPAN), a divisive hierarchical clustering technique, is being used for the initial classification of plant communities for the vegetation mapping. The foundation of TWINSPAN is Correspondence Analysis (CA), and the TWINSPAN classification is derived by dividing up the CA ordination over and over. The CA ordination technique is fraught with problems, particularly with complex datasets that have more than one underlying gradient (McCune and Grace 2002). The "arch effect" is an artifact of the method whereby positions of sampling units (SUs) are distorted along axis 2 in CA ordination space, and distances between SUs are compressed at the ends of the arch but stretched in the middle (McCune and Grace 2002). The underlying dissimilarity measure in CA is chi-squared distance, which has been shown to exaggerate differences between SUs containing rare species (Minchin 1987). For these reasons, CA and TWINSPAN are not recommended (McCune and Grace 2002). Therefore, we recommend that all use of this technique for the vegetation mapping and ground-based plot sampling be terminated.

If a hierarchical clustering technique is desired for the vegetation mapping work, we recommend using agglomerative cluster analysis with Sorenson's distance measure and Ward's method for linkage (McCune and Grace 2002). An ordination technique such as non-metric multidimensional scaling (Minchin 1987, McCune and Grace 2002) could be used as an alternative tool to explore compositional patterns in the vegetation and to determine assemblage groups without the assumption of hierarchical structure.

Recommendations Regarding Vegetation

Vegetation maps should be developed based on remotely sensed data and other spatial layers (e.g., topography and geomorphology). Resolution will need to be fine enough to capture patches of a minimum size of all vegetation types of interest, and this resolution will need to be decided on by the users/stakeholders (e.g., Stohlgren et al. 1997). The use of TWINSPAN for determining vegetation classes should be discontinued in favor of one of the alternative multivariate methods suggested above. We understand that this work is ongoing, and the use of a new statistical approach may pose challenges at this point, but it is highly recommended.

For ground-truthing the current vegetation map, B. Ralston (personal communication) recommends the use of 5 x 20 m transects for assessing broad vegetation classes (locations based on appearance, not surveying). For vegetation types that occur in larger patches (e.g., OHWZ, NHWZ, marshes and sandbars) consider using the multi-scale, circular plot developed in conjunction with the USDA Forest Health Monitoring Program (e.g., Barnett et al. 2007). Plots should be placed randomly within the vegetation zone at the study site, and plot center should be GPS located for correlation with remotely sensed data. If it is difficult to obtain GPS locations, survey the location from any known benchmark.

Once the first vegetation map is completed (e.g., Ralston et al. in preparation), the timing of subsequent mapping efforts should be planned to coincide with events expected to result in changes (e.g., sufficiently post-flood to expect changes to have occurred). If there are no suitable events, then mapping could be repeated on a cycle based on expected changes in vegetation (e.g., 5 years mentioned in the CMINs and confirmed by B. Ralston, personal communication).

For vegetation monitoring, we accept the methods in Kearsley et al. (2006) and advise continued use of the GRTS scheme and study sites already selected. These sites should be used to sample vegetation (plots), fauna (methods are taxa specific), and abiotic variables (e.g., bare soil, rocks, litter, wood in the vegetation plots; see Chong et al. 2001). When it is feasible to utilize these data for ground-truthing remotely sensed imagery, data from these 3 plots could be averaged to describe the larger area and correlated with pixels. Any future changes in plot size and/or layout should be determined by research with maintenance of the minimum sampling unit (1-m^2) allowing comparability across sample times if overall plot types are changed.

We strongly suggest that in the future terrestrial vegetation plots be distributed based upon a stratification of the Grand Canyon using fan-eddy complexes. The locations and geomorphic processes of these complexes are relatively well understood, likely are mapped, and as described previously have a significant influence on riparian vegetation composition and dynamics. Use of the fan eddy complex as the basic unit of stratification and study would provide a direct link and integration of terrestrial biota with geomorphic processes and studies. Another important stratification should be based upon bedrock type at river level, which in part determines canyon width and influences geomorphic processes of hillslopes and tributary streams that form fan eddy complexes. Simply stratifying based upon distance and canyon width (two classes) does not provide a direct link to geomorphic processes.

Sufficiently fine-scaled vegetation maps can be used in resource selection analyses related to animal monitoring (e.g., Manly et al. 2002). If even finer resolution is needed (e.g., for ground-dwelling arthropods), the 1-m^2 (sub) plot data can be used in correlation analyses, for example (also see Kearsley et al. 2006).

ANIMAL MONITORING

The PEP first evaluated whether past animal monitoring (arthropods, herpetofauna, birds, and mammals) from 2001 through 2003 met existing management objectives and core monitoring information needs for Goal 6. Our comments are based largely on the presentations by Kearsley and Brantley and the Kearsley et al. (2006) report. References to page numbers are from that document.

CMIN 6.1.1. The abundance, composition, distribution, and area of animals in the marsh community were not addressed directly. It does not appear that marshes were identified as a distinct habitat type for sampling, although there is some notion that these are subsumed by a "shoreline" category. We recommend explicitly addressing this ambiguity. We do note, however, that the origin of many marsh communities is recent and attributable to dam construction and operations because flows have been stabilized relative to historical annual fluctuations. The addition of periodic BHBFs is likely to alter the distribution of these newly formed marshes through sediment scour and deposition. The present sampling regime is unlikely to account for this degree of temporal variation. If marsh communities are considered distinct habitats, more intensive sampling will be necessary to monitor animal community composition and dynamics.

CMIN 6.2.1 and 6.3.1. For the NHWZ, this involved the composition and area of the animal community and for the OHWZ, the abundance, composition, and distribution of the animal community. Birds, arthropods, amphibians, reptiles, and small mammals were the target of focused sampling in both zones. Further comments on sampling of these taxa are given below.

CMIN 6.4.1. The composition, abundance, and distribution of animals in sand beach habitats were not completely addressed because the differences among sand beaches, marshes, and NHWZ habitats are unclear. Monitoring was directed at a shoreline zone, the boundaries of which were defined by diel fluctuations of flow during the time of sampling (p. 35). It appears that this zone often overlapped with sand beaches (p. 44), but also included long stretches of shoreline that would not be considered beach habitat. In addition, difficulties in assigning birds to this zone or the NHWZ led to pooling of observations from both habitats. As was noted above for marshes, sampling directed specifically at beaches will be necessary to address this CMIN.

CMIN 6.5.1. Some aspects of the distribution and abundance of non-native animals were assessed as part of the overall effort. In particular, six species of non-native arthropods were detected (p. 75) and a House Sparrow was observed (p. 250). Whether the sampling as currently conducted can detect changes in the distribution and abundance of these species is unclear (see below).

CMIN 6.6.1. The composition, abundance, and distribution of seep and spring animal communities were not sampled. However, the PEP questions why these communities would be sampled by the GCMRC because these habitats are largely independent from dam operations, and it appears that other agencies are now targeting seeps and springs in the Grand Canyon. We recommend that sampling for this CMIN be discontinued by GCMRC, and instead rely on information from other agencies.

CMIN 6.7.1. Aspects of the abundance, distribution, and reproductive success of the Southwestern Willow Flycatcher were monitored. Monitoring of this federally listed species requires a different protocol and amount of effort than other species being tracked by the GCMRC. Because counts are extremely low (from none to at most a few pairs in recent decades), it is virtually impossible to link dam operations to variation in the bird's presence or abundance. Given that Grand Canyon National Park has decided to monitor this species, we recommend that sampling for this CMIN be discontinued by GCMRC. Any observations of Southwestern Willow Flycatchers in routine GCMRC sampling can still be relayed to personnel at Grand Canyon National Park to supplement their work.

In summary, we recommend that for animals CMINs 6.1.1 to 6.5.1 be pursued in any future monitoring program, and sampling for CMINs 6.6.1 and 6.7.1 be discontinued as these needs are being met by other agencies. The overlap between monitoring needs of Grand Canyon National Park and the GCMRC still appears to be substantial for many taxa, particularly birds and mammals. Partitioning these efforts may increase overall efficiency.

Monitoring of Particular Animal Taxa

We concur with many of the conclusions reached by Kearsley et al. (2006) with respect to recommendations for monitoring of various taxa (p. 200-209). Below, we offer additional comments.

Arthropods

Whereas insects and other arthropods may not be organisms of special concern, the decision to include them in the overall monitoring effort was probably sound. Their abundance and diversity permits rigorous statistical tests to be employed. They also provide the primary link between plants and higher animals, and impacts on arthropods may be an early signal to broad ecological change.

The past arthropod sampling program used a variety of sampling methods (plant sweep netting, pit-fall traps, light traps, malaise traps, targeted sampling) to attempt to describe all arthropods along the riparian corridor. This was a laudable goal, but cannot be sustained over the long-term (material collected several years ago is apparently still being worked up) and the site-intensive nature of the sampling meant only a relatively small number of locations could be sampled (1 per day). However, the data set collected thus far should provide the means to develop an efficient long-term monitoring program.

The Lightfoot-Brantley group indicates that voucher collections have been developed to assist future efforts. Those collections should be completed (possibly requiring some wrap-up funding), and the collections need to be curated in a form that makes them readily accessible to future researchers (not stored piece-meal in museums). While those collections should be continually up-graded, we concur with the recommendation in the Kearsley report (p. 204) that the taxonomic level available in those collections serve as the benchmark for future efforts, and we recommend that the past focus on inventory aspects be de-emphasized. Workers in aquatic systems (where invertebrate bioassessment has been best developed) find that in many cases the genus or even family level is sufficient to detect biotic change (Bailey et al. 2001, but see Lenat and Resh 2001). Since many taxa have already been classified to species, future workers should use that information. However, we recommend that the taxonomic level now available serve as the basis for future work. (Collaborating with proposed ATBI efforts may be an efficient way to further develop collection taxonomy without additional cost, and as mentioned all material collected in GCMRC efforts should be made available to ATBI researchers.)

Analyses by the Lightfoot-Brantley group identified indicator arthropods for specific habitat types (Shore, NHWZ, and OHWZ) and detected several non-native arthropods; these taxa will be valuable for future work. The existing data set should be analyzed further using a wider variety of multivariate techniques (e.g., nonmetric multidimensional scaling) to enable community metrics to be used for monitoring. It appears that most indicator taxa were identified from plant sweep netting and pit-fall trapping, so those techniques should be the focus of future monitoring efforts. The plant sweep netting, in particular, seems appropriate because patterns between plants and arthropods can be directly assessed. Unless the other sampling techniques (light traps, malaise traps, targeted sampling) can be used without burdening the overall effort (in terms of both field logistics and laboratory time), we suggest that they be discontinued. The use of light traps and malaise traps in particular habitat zones seems of limited value because organisms collected could have originated across a broad range of habitats, not just near the trap.

By truncating the sampling intensity at any one site, it might be possible to include more sites, and develop better statistical rigor. Ideally, arthropod data would be available for every plant sample, as links between future plant change and arthropods will provide mechanistic clues. However, an arthropod sample at every plant plot would only be possible with plant sweep netting, as all other techniques require equipment to be left in place for many hours and permit collection of only a single sample per day. Thus, plant sweep netting should probably be the primary basis for any future monitoring. Dr. Brantley and the Kearsley report (p. 203) indicated that their group was concerned that the standardized sweeps did not reach plant surfaces > 2 m high, and some organisms might be missed. Sampling tall vegetation only up to 2 m resulted in mostly woody stems being sampled, instead of the foliage where many arthropods occurred. However, because baseline samples from 2001-2003 were collected only in areas < 2 m high, we recommend that that effort be duplicated in future years to enable direct statistical comparisons over time rather than developing a completely new methodology. However, it seems logical that when vegetation exceeds 2 m then supplemental (but separate) samples should be added (sweep nets with extension poles are commercially available). Biases of the original design could then be assessed empirically, and appropriate modifications designed, and methods to back calibrate 2001-2003 data explored.

Because pit fall trapping produced some unique and valuable results (p. 203), we recommend that it be retained. However, because logistically only a single site per day can be sampled, only a subset of sites in the overall study could be included. In this subset, other time intensive sampling might also be conducted.

Birds

The panel was provided little information on the bird sampling program, so our ability to make recommendations is limited. We were also informed that Grand Canyon National Park has initiated an avian monitoring program. If an avian program is retained by GCMRC, we recommend that, as before, bird sampling be directly integrated with the other efforts.

The number of sites used for bird monitoring was based in part on the power analysis by Spence (p. 33), who concluded that surveying approximately 64 sites three times during the breeding season would be necessary in order to detect trends (but it was not specified whether this was in abundance, distribution, or presence; see further comments below). This sampling, however, relied on single visits to each of the sites during the breeding season, and how this influenced estimation of bird abundance or presence was not addressed. At each site, individuals used walking surveys, rather than point-centered methods, to gauge relative abundance of species. We agree that walking surveys are more appropriate because long, narrow bands of vegetation are characteristic of each zone. Avian sampling must also incorporate diurnal and seasonal variation in bird presence and behavior, so bird sampling might have to be restricted to a subset of the plant plots (much like pit fall trapping for arthropods). The indicator analysis used for arthropods would also be useful for birds

Despite the challenges, including birds seems useful to a monitoring program because declines in charismatic birds would better bolster public support for modifying management than decreases in either plants or arthropods. Any avian program should focus on overall communities rather than specifically target endangered species. As mentioned, it appears that the National Park Service is taking the lead on monitoring of endangered birds in the Grand Canyon.

Response of waterfowl and waterbirds was mentioned as a possible priority. If these birds could be monitored in a rigorous fashion from boats as trips progressed, and not detract significantly from other priorities, their inclusion in the overall monitoring program seems appropriate. Perhaps waterfowl surveys could be focused solely in the upper reaches of the corridor where they apparently concentrate due to the water clarity.

Amphibians and Reptiles

From the information presented to the panel (p. 103-119), it appeared that standardized baseline data for amphibians and reptiles was not obtained in the 2001-2003 effort. Sampling herpetofauna at the optimal time was difficult. Some species are most readily detected while basking in sunlight, and delaying sampling until this time while staying with the rest of the surveying teams was logistically impossible. Moreover, some of the sites selected for vegetation and avian sampling were almost never in sunshine and thus unsuitable for amphibian and reptile monitoring. Nonetheless, the species list generated by this effort does serve as an initial estimate of the number of species present.

In contrast, the survey protocol made it difficult to estimate relative abundance, and thus these data are of limited use for future comparisons. Because amphibians are in global decline and their association with aquatic habitats suggests that they may be vulnerable to dam operations, they (and to a lesser extent, reptiles) appear to be excellent candidates for long-term monitoring. It may be difficult, however, to integrate surveys of these taxa with other monitoring because of the particular demands of sampling this faunal group. Reconciling this issue may require either abandoning herpetofaunal sampling or tailoring monitoring to more specifically address the abundance and distribution of these species. We favor the latter course, and recommend consulting the U.S. Geological Survey's Amphibian Research and Monitoring Initiative website (http://armi.usgs.gov/index.asp). Of particular interest is that the proportion of area occupied (http://armi.usgs.gov/PAOEstimator.asp) has been proposed as a national standard for amphibian monitoring, and it may also be applicable to other species.

Small Mammals

Small mammals (9 of the 29 species detected) were the subject of intensive sampling associated with the integrated monitoring effort. These species were amenable to quantitative sampling, and showed responses to the different vegetation zones that may make them useful subjects for detecting the effects of dam operations (perhaps primarily change in the OHWZ vegetation).

This monitoring helped establish up- and downstream boundaries on the distribution of beaver, but the detection methods used for this and other large mammals are of uncertain value for detecting trends. Other researchers have assessed associations between beavers and riparian plants (e.g., *Salix* and *Tamarix*) in the Grand Canyon, and have submitted a manuscript for journal review. If that paper is accepted by peer reviewers and proceeds to publication, GCMRC might consider funding follow up work on beavers because the ecosystem engineering capabilities of beavers in the Grand Canyon (in this case vegetation rather hydrologic impacts) might be inter-related with dam management. However, it would be difficult to integrate beaver monitoring with an overall riparian monitoring effort.

For other mammals, it appears Grand Canyon National Park is monitoring many of these populations. Because of their reliance on flying arthropods, often in association with watercourses, bats may be an important group to monitor by the park or GCRMC.

Additional Recommendations on Animal Sampling

We were impressed with the effort involved in the overall sampling program for riparian zone animals, and we also found the past attempts to link animal and plant patterns was a useful exercise in terms of exploring ecological mechanisms. However, it appears that the past level of animal sampling may not be sustainable because many projects remain incomplete, and even four years after sampling ended, journal articles have not been published. We encourage the scientists involved with the past 2001-2003 sampling effort to complete their analyses because those data sets could serve as foundations for future efforts (GCMRC might consider funding wrap up efforts, with the contingency that journal articles are produced by any continued funding).

Past efforts tried to be all-encompassing, attempting to inventory virtually every riparian organism along the Colorado River. This type of effort is consistent with an “All Taxa Biotic Inventory (ATBI)” but beyond the scope of monitoring biotic change associated with operation of Glen Canyon Dam. We recommend that GCMRC efforts now be focused on monitoring, not inventories (although any data collected on riparian animals can still be shared with the personnel charged with the ATBI proposed by Grand Canyon National Park). Moreover, to sustain a decades-long effort to monitor riparian animal communities, we believe that some choices will need to be made because it is not feasible to monitor the presence, distribution, and abundance of all animal species; rather, indicator, keystone, or umbrella species, despite their shortcomings as representatives (Lindenmeyer et al. 2002), will need to be selected from the various taxonomic groups. To the extent practicable, we further recommend that the vegetation monitoring program remain as a framework for animal monitoring to facilitate understanding their ecological connections and joint responses to dam operation.

NEED FOR A RESEARCH PROGRAM TO DIRECT MONITORING ACTIVITIES

Statistical Analyses

One of the key research questions that needs to be addressed in a monitoring effort such as this is detectability. How large must changes be before they are statistically detectable? Given the background variability in the Grand Canyon, is the design sufficiently powerful to have the desired level of detection? Finally, should there be a distinction between statistical and ecological change?

We believe that monitoring plans should address whether the desired measurements are of indices or true values of presence, distribution, or abundance. Presently, the metrics derived from sampling of animal species in the Grand Canyon should be regarded as indices because none of the sampling accounts for detectability of individuals in different habitat types, with different sampling schemes, or under different environmental conditions. In most cases, the detectability of species by sampling methods is less than one; i.e., not all animals (or species) present are observed. This leads to the underestimation of animal distribution and abundance, and failure to detect a species when it is present. Moreover, variability in detection in different habitats and in different weather conditions can confound attempts to establish linkages between species abundance or presence and environmental variables. More important is that failure to account for detectability introduces a source of error of unknown magnitude. This error leads to greater variation in the metrics being monitored, and can increase the time to detect statistically significant trends that may require a management response (MacKenzie 2005). The key problem, however, is that estimating detectability requires repeat sampling and much greater effort, at least initially. If biases in detectability are low and constant, indices offer a much more cost-effective method of monitoring populations. Characterizing the biases and variation in detectability should be a priority because it will dictate whether monitoring of particular groups can rely on indices. For those taxa that are inconsistently detected or counted, the GCMRC will have to decide whether to conduct more intensive sampling or to discontinue monitoring. A rapidly developing literature (summarized in MacKenzie et al. 2006, and specific to particular groups e.g., amphibians, Bailey et al. 2004; birds, Alldredge et al. 2007) offers extensive guidance on this issue.

The point of many monitoring schemes is to detect temporal variation in terms of population trends, overall fluctuation, or synchrony. A vital concern is the period of monitoring that may be necessary to detect a change in population size, which is contingent on the precision of population estimates, the temporal fluctuation in abundance, the magnitude and direction of change to be detected, and the level of confidence one hopes to have in the prediction (Peterman 1990, Thompson et al. 1998). For example, in a recent assessment on detecting trends in bird populations in North America, Bart et al. (2004) recommended a monitoring scheme that yielded 80% power to detect a 50% decline within 20 years based on annual monitoring, noting that detecting smaller declines at shorter intervals was impractical. Specification of each of these values is largely absent from the monitoring plan we have examined. Analyzing the data already obtained (similar to the suggestions by Spence), in light of issues involving detectability, could result in estimates for these parameters for many of the surveyed taxa (much as was done for plants [p. 202] although detectability may also be an issue there when multiple observers are used), and help guide decisions about which taxa can be monitored with the effort and funding that is available.

Sampling Frequency

Currently, recommendations are for sampling to occur every 5 years, with the most recent samples being collected in 2003. In general, a 2 year duration of sampling per cycle (rather than the 3 used in the past) seems reasonable. Within that 2 year period, sampling of individual metrics must account for life cycle patterns, with sampling coinciding with peak seasons of growth (plants) or activity (animals). Sampling frequency should be determined both by expected rates of change and expected drivers of change (e.g., natural droughts, management high-flow events). For example, baseline sampling should occur pre- and post-managed-flood event. To capture peak phenology in vegetation, that may mean sampling the growing season before and the growing season after, whereas, for animals there may be opportunities to get multiple seasons within one or more years before and after. Whenever animal data are collected, corresponding habitat data must be collected simultaneously even though the full suite of vegetation and habitat data would not be collected.

An example of sampling frequency based on a March year 1 release (e.g., proposed March 2008 event) could be: (1) pre-release terrestrial data provided by Kearsley et al. 2006; (2) post-release 2008 spring, mid-summer and fall animal (e.g., selected arthropods, small mammals and birds) sampling and summer vegetation sampling; and (3) full suite sampled in 2009 (1 year post-event). If significant changes are found post event (either 2008 or 2009), then a year 3 (2010 in this case) sample should be considered to capture any rapid rates of change. Depending on rates of change, sampling could be repeated year 5 then at intervals (e.g., 5 years) defined both by observed rates of change and/or needs to sample based on the occurrence of other drivers of change (e.g., drought). Additionally, subsets of samples could be collected opportunistically related to events or other management actions (e.g., fire, non-native species removal, restoration, etc.). Those data would contribute to long-term monitoring and to adaptive management applications.

The Need for A Reference Standard

A final concern, and one that applies to nearly all of the monitoring within the Colorado River below Glen Canyon Dam, is that the effects of dam operations cannot be fully appreciated unless sampling is conducted in portions of the river system that are not affected by the dam. All downriver study sites are influenced by dam operations, and analyses of data from there are prejudiced by this bias and perhaps by an overly narrow geographic perspective. In preparing for this protocol evaluation panel, we reviewed the recent research and documentation on humpback chub below Glen Canyon Dam as well as information on the now-extinct big-river fish once found there. Our conclusion was that the current research program would greatly benefit from consideration of how populations of these fish are faring elsewhere in the Colorado River basin, which might lead to redirecting some of the current research. Similarly, the understanding of changes in the OHWZ and NHWZ vegetation is incomplete without consideration of riparian vegetation unaltered by flow regulation, especially if one must also account for large-scale ecological drivers such as climate change and nonnative species invasions. Thus, upriver sites for the study of aquatic food webs, large wood distribution, vegetation patterns, bird relations to vegetation structure, and other ecological patterns and processes are the necessary controls that underlie reliable inferences about the consequences of the dam's presence and its current and projected operations. To some degree, studies conducted adjacent to tributaries may also serve as controls, but for patterns associated with the main stem, locations above Glen Canyon Dam will be essential.

Other topics

Determining the appropriate and accepted analyses is an important area where research can inform the monitoring. Keeping up with the latest and best techniques for data analysis is a time-consuming, yet essential piece of a monitoring program. Continuous input from someone with this kind of expertise is needed.

Below, we provide some more specific possibilities for future workers to consider.

1) For plants, pre-dam aerial photos of the canyon may indicate baseline conditions for plant cover, and planned remote sensing efforts might detect on-going temporal change. For animals, it could be assumed that they will track changes in vegetation. An important study to establish baseline condition could be developed from the analysis of the pre-dam air photos. This baseline, along with reference river analyses, would provide an important framework for understanding effects of Glen Canyon Dam. This has been done to large extent for geomorphic processes, and should be done for riparian vegetation.

2) The Colorado River through Cataract Canyon likely provides the most suitable reference site for riparian vegetation and geomorphic processes to the Grand Canyon. This river section has similar fan-eddy dominated reaches, with abundant talus, and only small patches of channel-margin deposits and eddy bars, to the Grand Canyon (J. Schmidt, personal communication). While the Colorado River through this reach is partly regulated by Flaming Gorge and numerous dams on the Colorado River, and its tributaries, it is the only river reach of similar size and with similar landforms as the Grand Canyon, and is as "pre-dam" as can be found in the Colorado River basin.

3) Streams tributary to the Grand Canyon could also serve as reference streams. The advantages of these streams include logistic proximity and climatic conditions for the

Colorado River and its tributaries will be similar. The major drawback is that the main stem and its tributaries (or springs) are not ecologically similar.

4) Other minimally regulated or unregulated rivers in the Southwest already being monitored might also be used for reference. These smaller rivers must occur in arid region canyons dominated by fan eddy complexes, and support intact vegetation and animal communities. An example is the Yampa River canyon within Dinosaur National Monument. This option has the advantage of using data that are already being collected, and the fact that the habitats are largely similar. However, as with tributaries, these rivers will be ecologically different from the Colorado because of their smaller size.

SUMMARY RECOMMENDATIONS

Specific recommendations are provided throughout this document. Below we itemize some of the most important suggestions.

1. The Management Objectives (MOs) and Core Monitoring Information Needs (CMINs) developed to direct GCMRC actions need re-evaluation as some are unrelated to dam operations, others seem unrealistic, and others need clarification.
2. We recommend that CMINs 6.6.6 (seeps and springs) and 6.6.7 (SW willow flycatcher) be discontinued by the GCMRC, and that the GCMRC arrange to receive data being collected by other agencies on these topics.
3. Vegetation maps need to be finalized.
4. A schedule and strategy for repeat mapping and change detection analyses should be developed.
5. Sample plots should be located randomly within the randomly selected sample sites, and plot design should be modified as discussed in the vegetation summary.
6. The suite of biotic variables chosen for monitoring can be truncated to include only organisms conducive to rigorous sampling and methods conducive to rigorous analyses. Data on native and non-native plant species should be collected simultaneously, but separate analyses should be conducted for each.
7. Monitoring large wood abundance and location should be considered.
8. The absence of monitoring sites outside of Glen Canyon reduces the strength of inferences about the effects of dam operations. Reference sites not subject to the flow regime imposed by Glen Canyon Dam are necessary, but their location may vary depending on the suite of monitoring variables under consideration.
9. Publication of results in peer reviewed journals should be a priority, as there is a large audience interested in how flow regulation affects Grand Canyon habitats, and the journal review process is a valuable mechanism for quality control.
10. Sampling frequency should be determined both by expected rates of change and expected drivers of change (e.g., natural droughts, management high-flow events).

LITERATURE CITED

Alldredge, M.W., K.H. Pollock, and T.R. Simons. 2007. Estimating detection probabilities from multiple-observer point counts. The Auk 123:1172-1182.

Anderson, M.J. 2001. A new method for nonparametric multivariate analysis of variance. Austral Ecology 26: 32-46.

Bailey, L.L., T.R. Simons, and K.H. Pollock. 2004. Estimating detectability parameters for plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68:1-13.

Bailey, R.C., R.H. Norris, and T.B. Reynoldson. 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. Journal of the North American Benthological Society 20:280-286.

Barnett, D.T., T.J. Stohlgren, C.S. Jarnevich, G.W. Chong, J.A. Ericson, T.R. Davern, and S.E. Simonson. 2007. The art and science of weed mapping. Environmental Monitoring and Assessment 132:235-252.

Bart, J., K.P. Burnham, E.H. Dunn, C.M. Francis, and C.J. Ralph. 2004. Goals and strategies for estimating trends in landbird abundance. Journal of Wildlife Management 68:611-626.

Birken, A. and D. Cooper. 2006**.** Processes of *Tamarix* invasion and floodplain development during the 20th century along the lower Green River. Ecological Applications 16:1103-1120.

Breiman, L. 2001. Random forests. Machine Learning 45:5-32.

Breiman, L., A. Cutler, A. Liaw, and M. Wiener. 2006. Breiman and Cutler's Random Forests for Classification and Regression. http://CRAN.R-project.org/ [R package version 4.5–16].

Chong, G.W., R.M. Reich, M.A. Kalkhan, and T.J. Stohlgren. 2001. New approaches for sampling and modeling native and exotic plant species richness. Western North American Naturalist 61:328-335.

Cooper, D.J., D.C. Andersen, and R.A. Chimner. 2003. Multiple pathways for woody plant establishment on floodplains at local to regional scales. Journal of Ecology 91:182-196.

Evans, J.S. and A.T. Hudak. 2007. A multiscale curvature algorithm for classifying discrete return lidar in forested environments. IEEE Transactions on Geoscience and Remote Sensing 45:1029-1038.

Grand Canyon Monitoring and Research Center (GCMRC). 2006. Monitoring and research plan to support the Glen Canyon Dam adaptive management program, fiscal years 2007-2011.

Gregory, S., K. Boyer, and A. Gurnell (editors). 2003. The Ecology and Management of Wood in World Rivers. American Fisheries Society, Symposium 37, Bethesda, Maryland.

Hudak, A.T. and C.A. Wessman. 1998. Textural analysis of historical aerial photography to characterize woody plant encroachment in South African savanna. Remote Sensing of Environment 66:317-330.

Kearsley, M., N.S. Cobb, H.K. Yard, D.C. Lightfoot, S.L. Brantley, G.C. Carpenter, and J.K. Frey. 2006. Inventory and Monitoring of Terrestrial Riparian Resources in the Colorado River Corridor of Grand Canyon: An Integrative Approach. Final Report Cooperative Agreement / Assistance Awards 01 WRAG 0044 (NAU) and 01 WRAG 0034 (HYC).

Kennedy, R.E., W.B. Cohen, and G. Takao. 1997. Empirical methods to compensate for a view-angle-dependent brightness gradient in AVIRIS imagery. Remote Sensing of Environment 62:277-291.

Lenat, D.R. and V.H. Resh. 2001. Taxonomy and stream ecology – the benefits of genus- and species-level identifications. Journal of the North American Benthological Society 20:287-298.

Lindenmeyer, D.B., A.D. Manning, P.L. Smith, H.P. Possingham, J. Fischer, I. Oliver, and M.A. McCarthy. 2002. The focal-species approach and landscape restoration: a critique. Conservation Biology 16:338-345.

MacKenzie, D.I. 2005. What are the issues with ‘presence/absence’ data for wildlife managers? Journal of Wildlife Management. 69:849-860.

MacKenzie, D.I., J.D. Nichols, G.B. Lachman, S. Droege, J.A. Royle, and C.A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248-2255.

MacKenzie, D.I., J.D. Nichols, J.A. Royle, K.H. Pollock, J.E. Hines, and L.L. Bailey. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego, USA.

Manly, B.F., L. McDonald, D.L. Thomas, T.L. McDonald, and W.P. Erickson. 2002. Resource Selection by Animals: Statistical Design and Analysis for Field Studies. 2nd Edition. Kluwer Academic Publications, Boston, Massachusetts, USA.

McCune, B. and J.B. Grace. 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon.

Minchin, P.R. 1987. An evaluation of the relative robustness of techniques for ecological ordination. Vegetatio 69:89-107.

Moeur, M. and A.R. Stage. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. Forest Science 41:337-359.

Moisen, G.G. and T.S. Trescino. 2002. Comparing five modelling techniques for predicting forest characteristics. Ecological Modelling 157:209-225.

Ohmann, J.L. and M.J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A. Canadian Journal of Forest Research 32:725–741.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2-15.

Thompson, W.L., G.C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California.

Stevens, D. and A. Olsen 2004. Spatially-balanced sampling of natural resources. Journal of American Statistical Association 99: 262-278.

Stevens, L. E., J. C. Schmidt, T. Ayers, B. Brown. 1995. Flow regulation, geomorphology, and Colorado River marsh development in Grand Canyon, Arizona. Ecological Applications 5: 1025-1039.

Stohlgren, T.J., G.W. Chong, M.A. Kalkhan, and L.D. Schell. 1997. Multiscale sampling of plant diversity: effects of minimum mapping unit size. Ecological Applications 7(3):1064-1074.

Urquhart, N.S., G.T. Auble, J.G. Blake, D.T. Bolger, T. Gerrodette, S.G. Leibowitz, D.C. Lightfoot, and A.H. Taylor. 2000. Report of a peer review panel on terrestrial aspects of the Biological Resources Program of the Grand Canyon Monitoring and Research Center. June 23, 2000. U.S. Department of the Interior, Grand Canyon Monitoring and Research Center. Flagstaff, AZ. 49 p.

Urquhart, N. S. and M. Kearsley. Unpublished manuscript. Executing a spatially balanced sampling design along the Colorado River in Grand Canyon National Park, AZ.