STAFF ANALYSIS

SCIENTIFIC PEER REVIEW OF CALIFORNIA RAPID ASSESSMENT METHOD (CRAM)

SEPTEMBER 2011

STATE WATER RESOURCES CONTROL BOARD
DIVISION OF WATER QUALITY
CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

ANALYSIS OF CALIFORNIA RAPID ASSESSMENT METHOD PEER REVIEW BY STATE WATER BOARD STAFF

Introduction

A peer review of the California Rapid Assessment Method for Wetlands (CRAM) has concluded. This peer review was conducted according to California EPA's peer review standards, and was overseen by Cal EPA Peer Review Manager, Dr. Gerald Bowes. The following analysis of the review was prepared by State Water Board Staff from the Division of Water Quality's 401 Certification and Wetland Unit.

The peer reviewers are listed in Table A below. The Staff of the State Water Resources Control Board extend our sincere thanks to the reviewers for their efforts. The reviewers' comments indicate that they understand the intent of their review, are qualified to conduct the review, and that their reviews were adequately supported by the materials they were provided.

Staff finds that the reviewers' comments demonstrate an adequate technical understanding of CRAM and the process of its continued development. Staff finds no reason to provide the reviewers with any follow-up clarifications that might alter their reviews.

Staff does not agree with all the reviewers' comments, and notes that while some comments seem to reflect minor misunderstandings of the method, other comments clearly reflect reviewers' preferences regarding elements in wetland monitoring approaches that are practiced in other regions of the U.S. In general, staff finds that these comments indicate an overall consistency of CRAM with wetland monitoring approaches across the country.

The different reviewers' comments about each statement vary significantly in detail, and in more than a few cases there are clear differences in opinion. There does not seem to have been any communication among the reviewers about their reviews or any effort to rectify each review with the others. One valuable aspect of the reviews is therefore their collective breadth of perspective. This was fostered by the different technical expertise of the reviewers plus the independence of their reviews.

Staff also notes that many comments pertain to the content or format of the CRAM manual, fieldbooks, or other supporting materials. For example, there are suggestions for additional definitions of key terms, expanded discussions of selected conceptual models, clarification of some assumptions, more scientific rationale for some metrics, and further discussion of CRAM limitations and appropriate CRAM uses. Staff agrees with most of these suggestions and recommends that they be used to guide the next revision of the CRAM manual and other supporting materials (which is in progress by scientists at Moss Landing Marine Laboratory).

The peer reviewers generally agree that CRAM is based on sound scientific knowledge, methods and practices, although some concerns regarding the method were presented. The reviewers note that, as with any environmental assessment methodology, ongoing development and refinement of the method is necessary to ensure the method's reliability and accuracy, and that application of the assessment results could be subject to misuse. As discussed in the analysis below, scientific and administrative processes are under way – indeed have been in development since before the review was initiated - to address these concerns.

This report does not generally address the policy-related comments, but staff notes that multiple reviewers expressed concern that inadequate funding or other constraints could prevent full implementation of the CRAM development process and associated quality assurance measures. These concerns do not pertain to the scientific validity of CRAM. However, policy makers should be aware that the success of CRAM or any assessment program depends on factors beyond the science. The reviewers, in their experience, identify those non-scientific factors as likely to be essential to an assessment method's success or failure.

This analysis addresses the peer review request point by point in Part 1. Certain topics which recur throughout the reviewers' responses are addressed together in Part 2. Part 3 presents staff recommendations for incorporation of the peer review results into Water Boards assessment strategies.

This analysis frequently makes reference to the "Level 2 Committee" (L2), a committee of the California Wetland Monitoring Workgroup (CWMW) which in turn is a workgroup for the SB1070-created California Water Quality Monitoring Council (Monitoring Council). L2's chair is held *ex officio* by a State Water Board staff person to help coordinate communication between the scientific and regulatory communities. L2 is charged with the task of leading the scientific development and refinement of Level 2 wetland assessments in California. At this time, CRAM is the only Level 2 assessment under development by CWMW, but other assessments for other purposes may be proposed or developed in the future.

"Level 2" assessments are rapid field assessments of wetland condition; the nomenclature is based on recommended USEPA guidelines for developing successful statewide wetland monitoring programs.[1] "Level 1" constitutes the development of comprehensive wetland mapping, based primarily on remote sensing data; "Level 3" assessments constitute a wide range of detailed studies and/or site-specific observations of factors affecting or determining wetland conditions, functions, or services, such as bird populations, botanical surveys, hydrologic studies, etc.

For more background on the topics address in this analysis, readers are encouraged to consult the peer review request and the CRAM Manual.

Staff consulted with some of the Principal Investigators for CRAM in the preparation of this analysis, and thanks these scientists for their assistance. (Those consulted are listed in Table B) but any errors or omissions are the responsibility of State Water Board Staff who compiled this report.

---

[1] United States Environmental Protection Agency (USEPA) - Office of Water & Wetlands, 2002, Technical Paper: *Wetland Monitoring and Assessment – A Technical Framework*; EPA publication 843-F-02-002(h). [http://www.epa.gov/owow/wetlands/pdf/techfram_pr.pdf]

Table A: Peer Reviewers for the California Rapid Assessment Method	
Reviewer	Reviewer's Affiliation
James T. Anderson, Ph.D.	Professor, Wildlife Ecology and Management Director, Environmental Research Center Division of Forestry and Natural Resources West Virginia University
Gerald J. Niemi, Ph.D.	Professor, Biology Senior Research Associate, NRRI University of Minnesota
Rick Rheinhardt, Ph.D.	Research Associate Professor Department of Biology East Carolina University
Denice Heller Wardrop, Ph.D., P.E.	Assistant Director, Penn State Institutes of Energy and the Environment Associate Director, Penn State Cooperative Wetlands Center Senior Research Associate and Associate Professor of Geography and Ecology, Penn State University

Table B: Principal Investigators for the California Rapid Assessment Method (CRAM) Consulted for the Staff Analysis of the CRAM Peer Review	
Principal Investigator	Principal Investigator's Affiliation
Josh Collins Ph.D.	San Francisco Estuary Institute
Chad Roberts, Ph.D., PWS	Independent Wetland Scientist
Eric Stein, Ph.D.	Southern California Coastal Water Research Project
Martha Sutula, Ph.D.	Southern California Coastal Water Research Project
Ross Clark	Moss Landing Marine Laboratory (formerly with California Coastal Commission)

Part 1
Response to Individual Comments

The Peer Review Request asked the reviewers to respond to 32 numbered statements about CRAM which asserted certain key points about the method. These statements are arranged into three parts. Reviewers were also asked to make “big picture” (BP) comments to synthesize previous remarks or introduce new topics.

In Part 1, each Reviewer’s responses are presented for each of the statements in the Peer Review Request. For example, Gerald Niemi’s comments on statement 1 are shown as **NIEMI-1**. Staff comment and analysis is shown as **S-NIEMI-1**. Some of the reviewers’ more lengthy responses address multiple topics. These are enumerated as **NIEMI-1a**, etc. Staff Response to **NIEMI-1a** appears as **S-NIEMI-1a**.

In order to facilitate the reader’s ability to discern the content sources in this report:

Statements from the Peer Review Request are presented as boxed text to help the reader discern those statements from the reviewers’ comments.

Peer Reviewers’ comments are presented almost exactly as received, but converted to black 12 point text for readability. Only minor formatting changes (e.g., removal of extra spaces, etc.) were made. In general, formatting, spelling, emphasis, italicization, bold text, etc. in these remarks are those of the Reviewers unless otherwise noted.

Staff analysis of the reviewers’ comments is presented in 11 point blue text.

GROUP 1 ISSUES

(related to CRAM Manual Executive Summary and Chapter 1; Technical Bulletin Sections 1, 2 and 3)

> **(1) After reviewing the Topics in this request, please comment on the recommended appropriate and inappropriate uses of CRAM listed in Table 1** [*of the Peer Review Request*]**. (CRAM Executive Summary, Ch. 1.3; Technical Bulletin Sec. 3B-C)**

NIEMI-1: Responses to Table 1. In general, these are reasonable, but they do not necessarily stand alone without some additional detailed knowledge of CRAM and its overall intent.

a. Appropriate uses – Table 1.b – this statement is awkwardly worded. An appropriate use might include the evaluation of wetland beneficial uses within ecological reserves, mitigation banks, etc.

b. In general, I suggest many of the "appropriate uses" could be expanded upon in this table. For instance, it is not clear what "unauthorized (enforcement) actions" or what "required targets or performance criteria" represent. Perhaps a bit more detail like f and g under inappropriate uses would improve this description.

S-NIEMI-1: Staff agrees that the table of appropriate and inappropriate uses should be clarified. L2 will consider expansion of this list. Staff suggests that the table could be replaced with official CRAM user guidelines based on the existing CRAM technical memorandum that addresses these matters in more detail. Guidelines are updated periodically through the Level 2 Committee and the first update will address these concerns.

ANDERSON- 1: CRAM like most wetland rapid assessment procedures has multiple functions and uses. The 8 recommended appropriate uses are all expected and typical of other rapid assessment procedures. I believe the documentation throughout the manual is supportive of these uses. I also believe the inappropriate use section is useful for acknowledging the limitations of the CRAM. I do not have any issues with item 1.

S-ANDERSON-1: No Response Warranted

RHEINHARDT- 1a: The list of appropriate and inappropriate uses for CRAM (Table 1) covers most possible applications of any assessment method. The inappropriate uses listed for CRAM is reasonable, but a few of the appropriate uses listed need more clarification and/or a discussion of limitations, particularly regarding items b, f, and g.

S-RHEINHARDT-1a: See S-NIEMI-1.

RHEINHARDT-1b: Evaluating success of restoration sites and mitigation banks (generally large-scale restoration sites) relative to performance standards may not be practical with CRAM. This is because typical compensatory mitigation success criteria are generally more specific than the metrics evaluated by CRAM.

S-RHEINHARDT-1b: CRAM is not intended to be used by itself to assess mitigation, but rather to be combined with Level 3 methods for these types of complex analyses.

Further, although hydrology can be restored immediately in restoration, in most cases, it takes several to many years for vegetation and some of the structural patch metrics outlined in CRAM to develop fully. For example, mounds and snags take more time to develop than the typical time frames allocated to monitoring mitigation success.

In any case, assessments of projects would not be based on a single measure. CRAM will be used repeatedly in conjunction with other methods to assess change over time. There are plans to develop “project performance curves” for each wetland class by assessing projects of different age using CRAM. This will help determine the method’s ability to resolve temporal changes within projects of varying size and duration.

Limited data for estuarine wetlands suggest that CRAM attribute scores and metric scores track temporal changes in overall structure and form of projects, but Staff agrees that more data of this kind will help improve CRAM.

Use of CRAM by the water boards should be phased in over time for more complex uses which require greater levels of confidence. For example, use of CRAM to monitor mitigation success would not be recommended without substantial additional experience with CRAM and analysis of that experience. The way to obtain that experience is to apply CRAM to less complex purposes to build experience and data sets which can be analyzed.

WARDROP-1: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The only use of CRAM that is debatable is the monitoring of mitigation banks. Some studies are questioning the use of tools developed on naturally-occurring sites for application at mitigation sites, due to a number of factors including unique landscape factors associated with mitigation sites, the similarity of a constructed ecosystem with the operating rules of a natural one, etc. A limited test of CRAM on mitigation versus natural sites is required before widespread or regulatory use.

S- WARDROP-1: CRAM scores for mitigation projects and other wetland projects will be assessed relative to cumulative frequency distributions derived from regional and/or statewide ambient surveys, and relative to “least-impacted” reference sites. This should meet the need to compare mitigation sites to ambient or “natural” conditions.

Discussions have begun about stratifying the ambient surveys and project assessments based on landscape context to improve the resolution of CRAM.

Staff notes that this is a valid concern in a regulatory context. There are some fundamental hydrological differences between many mitigation sites and almost any natural site. Scoring approaches that are tailored for “new” mitigation or restoration sites separately from ambient or “natural” sites may be appropriate.

(2) The U.S. Environmental Protection Agency has developed a three-level approach for wetland assessment. The design provides that each level builds on the previous one, and adds more specific information. Level 1 data is usually generated from aerial photos and maps for the purposes of inventory and planning; Level 2 data is collected in the field using rapid assessment methods; Level 3 includes data of a specific type to assess site specific concerns or validate methods or conclusions. CRAM is designed to be a Level 2 assessment tool to evaluate the overall condition of wetlands and to assess impacts of human activities that stress the local wetland ecology. (CRAM Executive Summary, Ch. 1.2)

This framework is summarized in a USEPA technical paper (6. USEPA, 2002). This framework is more fully explained by Kentula (7. 2007) Stein, et al (8. 2007) and in other USEPA publications (9., 2003; 6., 2006)

NIEMI-2: I agree – the EPA documents and supporting literature identify the appropriate use of CRAM as a level 2 process in wetland evaluation.

S-NIEMI-2: No Response Warranted.

ANDERSON- 2: The Level 1-2-3 USEPA wetland assessment approach has been widely adopted by state and federal agencies. I am well versed with this 3-tiered approach and believe that CRAM is well suited for Level 2 assessments.

S-ANDERSON-2: No response warranted.

RHEINHARDT-2: CRAM could reasonably be considered a Level 2 assessment approach because it is intended to be rapid and field-based. While it is true that Level 2 assessments should be based on Level 2 and Level 3 data, which are used to “validate rapid methods” and “characterize reference condition” (Stein et al. 2007), it was not clear if Level 3 data were used to validate or even help calibrate the ratings, particularly since the review material did not include background data to support the development of the method and ratings, nor were studies cited to support assumptions for the ratings. Based on the narratives for various ratings, it is clear that best professional judgment (BPJ) has to be invoked often.

S-RHEINHARDT-2: During CRAM development Level 3 data were used to evaluate how effectively CRAM index scores, attribute scores, and metric scores track variability in selected wetland functions or specific aspects of wetlands condition. Qualified L3 data were regressed on L2 data and generally supported the models (Stein et al 2009).

The distribution of L3 data from ambient surveys were used to identify natural breaks in data that could inform the CRAM rating tables. In the absence of such data breaks, the tables are based on quartiles of CRAM scores or Best Professional Judgement. The ratings can be adjusted as additional ambient data become available. The need for objectivity is addressed in Part 2.

WARDROP-2: The term “surface waters” is used repeatedly in Chapter 1; it would seem that “wetlands” would be a more appropriate term.

This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The description of Level 1 assessment in the CRAM Executive Summary incorrectly summarizes language contained in US EPA 2006, as well as other references that describe the basis of this work (Wardrop et al., 2007: Brooks et al. 2004,; Brooks et al., 2006).

US EPA 2006, Table 1, defines Level 1 Landscape Assessments as, "Use GIS and remote sensing to gain a landscape view of watershed and wetland condition. Typical assessment indicators include wetland coverage (NWI), land use and land cover". The CRAM description completely utilizes only the inventory portion of this description. Landscape Assessments are intended to be a first cut at a condition assessment of the wetland, and an inventory alone does not satisfy the intent of a Level 1 assessment.

Many authors have contributed methods that utilize landscape characteristics and metrics as a first cut at condition assessment (examples include Wardrop et al., 2007; Hychka et al., 2007; Wellar et al., 2007). Only if a Landscape Assessment provides a condition assessment can the following be true: "each level builds on the previous one, and adds more information".

The CRAM method, as stated, deviates significantly from this concept of interconnectedness between levels of assessment with refinement from one level to the one "above" it.

S-WARDROP-2: The term "surface waters" is a commonly used collective term for streams, wetlands, ponds, lakes, etc. (as opposed to ground water). The term is used in this context to communicate that CRAM is used to assess both non-fluvial wetlands and rivers/streams. While wetlands may sometimes occur in rivers or streams, the presence of fluvial processes requires different assessment approaches from those used to assess "non-fluvial" wetlands. Future revisions of CRAM manual and field books will seek to clarify the meaning of these terms.

The separation of L1 and L2 (CRAM) assessments is addressed separately in Part 2.

(3) CRAM's wetland classification system is reliable for use throughout California (CRAM Ch. 1.5 & 3.2.2; Technical Bulletin, Sec. 3.A)

NIEMI-3: This is always subject to change and on-going consideration, but the classification seems very reasonable and justifiable for use in CRAM.

S-NIEMI-3: No response warranted

Anderson-3: CRAM is designed for use throughout California. I was pleased to see the perceived limitations in assessing structurally simple wetlands, such as headwater riparian areas with simplistic vegetation structures, as mentioned in the document. CRAM is designed as a dynamic system to capture the various wetland classification regimes (HGM, NWI, and unique CA wetlands) so it should have enduring value. While I do not find these limitations overly troublesome I encourage the development of additional models suitable for these important but under represented wetlands.

S-ANDERSON-3: The process of CRAM implementation accommodates the potential for developing new modules for different wetland classes, or for modifying existing modules to address “subclasses” of wetlands that are not appropriately covered by existing modules. For example, module development is under way for arid and headwater streams, wet meadows, and seasonal lagoons. This methodology-monitoring function is assigned to the L2 Committee and is ongoing.

RHEINHARDT- 3a: Wetlands in California formed under, and are controlled by, an extremely diverse array of climatic regimes, surficial geologies, soils, and natural disturbance regimes. Natural variation is so wide, even within a hydrogeomorphic class of wetlands, that variations in hydrologic regime, species composition, and landscape connectivity attributes would make it difficult to select a set of assessment criteria that could be applicable to all wetlands even within a class.

S-RHEINHARDT- 3a: Eco-regional variations in CRAM modules are being considered for wetland classes for which CRAM demonstrates systematic bias due to regional variations in condition. However, it has been demonstrated through a statewide ambient assessment of estuarine wetlands that CRAM is capable of identifying and documenting real regional variations within wetland classes, and this element of ambient wetland assessment is considered to be a benefit of using a standardized methodology across the broad gradients that exist in California.

RHEINHARDT- 3b: Classifying wetlands by hydrogeomorphic and regional differences prior to identifying metrics that could be used to measure condition is the optimal way to minimize the amount of inherent variation among wetlands. CRAM used a hydrogeomorphic differences to classify wetlands into six classes and 12 subclasses. However, the classification seems to have been developed after data were collected in reference sites rather than before. Thus, it doesn’t appear that classification was used to its maximum advantage, which would be to separate wetlands into hydrogeomorphic classes prior to identifying appropriate metrics for each class. Instead, condition states and metrics were identified that could be used for all wetland classes, data were collected, and then the classification was used to calibrate metrics relative to the hydrogeomorphic classes.

S-RHEINHARDT-3b: CRAM was originally conceived as using one module for all wetland classes in California. The classification system, which is based on the classes in HGM, was developed prior to the development of the CRAM methodology, in order that experts for different wetland classes could be assembled to assist in developing CRAM, and so that the ability of CRAM to assess wetlands of several kinds could be tested. The methodology intentionally uses the same module structure (the same four attributes) to address different wetland classes, although the metrics used for each attribute may differ among modules for different wetland classes. The ability to use the same module structure for different wetland types is considered a desirable characteristic of the methodology.

RHEINHARDT- 3c: By doing this, some metrics that could have been very useful for differentiating condition in a particular class, but perhaps not useful at all for other classes, were missed or excluded. Consequently, the resulting narratives used for discriminating alternative states for some metrics had to rely on best professional judgment (BPJ) to obtain ratings. As a result, many condition scores are insensitive in the intermediate range in condition, the range where the most discrimination is needed.

S-RHEINHARDT-3c: The overall structure of CRAM was found to be applicable across wetland classes, and the use of the same rating categories (attributes) across classes allows a more direct comparison of wetland condition among classes than would varying the rating criteria by class. However, as noted in S-RHEINHARDT-3b, the specific metrics used for assessing each attribute do differ in different wetland modules, when scientific evidence and best professional judgement indicate the need for adjusting the method. Additional adjustments are likely to be made as experience is gained in implementing CRAM.

The need for objectivity is addressed separately in Part 2.

WARDROP-3: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The question is oddly worded, but let me state the assumptions under which I am answering:

I cannot comment on whether or not all possible wetland types in the state are covered by the classification, nor do I have the extensive field experience in California to determine whether it is reliable. Based on the extensive fieldwork done in support of CRAM's development, I believe that the above statement can be supported to the largest possible extent. However, I would raise a different question concerning whether the classification is appropriate for use in conjunction with the rapid assessment of wetland condition.

There can be situations where wetland types are present in a mosaic, and this mosaic functions in a different way, by virtue of its complexity, than the sum of its parts. For example, the mid-Atlantic region of the U.S. often contains wetland systems, along headwater streams, that have riverine, slope, and depressional elements/subsystems at small spatial scales that collectively result in a system called a headwater complex (the complex can occur at spatial scales as small as 0.5 acres). The different elements occur proximate to one another, making delineation between specific elements arbitrary.

In addition, one must ask if each were assessed separately, would the combined answer be the same as if the whole mosaic were assessed?

For example, consider the following system: a headwater stream occurring in a narrow valley between the base of opposing slopes. A depressional wetland may occur in the valley at the base of a slope, along with other slope wetlands, that grade into the riverine wetland generally present along the stream. Now assume that there is a stressor, perhaps a culvert discharge, that directly affects only the depressional wetland.

Theoretically, CRAM will only account for the presence of this stressor on one wetland (the depression), and not in the adjoining wetlands, although the stressor affects the entire mosaic of wetlands, since they are hydrologically connected. Will the collection of CRAM scores for these individual wetlands be comparable to a CRAM score that would be developed based on a mosaic wetland class? I don't know the answer, but my experience has been that it is important to test for inconsistencies in lumping or splitting wetland classes in rapid assessment techniques.

During development of HGM functional assessment models for Pennsylvania, we did indeed develop separate models for depressions, slopes, and floodplains, and a mosaic class termed headwater complexes. We found that assessment results for headwater complexes were more stable and consistent than utilizes the suite of scores provided by the individual assessments applied to the subsystem elements.

While the recommendations concerning AA features and sizes provide guidance on "splitting", they do not seem to address when “lumping” is appropriate. Fennessy et al., 2004 states: “The goal of classification is to reduce variability within a class and enable more sensitivity in detecting differences between least- impacted and impaired wetlands.” The question is whether this is best achieved, in some cases, by a lumped or mosaic class.

S-WARDROP-3: CRAM is executed with modules that have been constructed for identified wetlands classes in California. New wetland modules are only developed when the existing modules are deemed non-applicable without major revision. The need for a new module is first recognized by the L2 Committee and a new module is then developed. A new module may be recognized as necessary if the existence of a new wetland class is identified, as Wardrop describes in the mid-Atlantic region.

The classification in use in CRAM *does* need to be considered in an overall, statewide classification framework. An element in CRAM’s development methodology allows tailoring modules to “new” contexts without having to develop completely new modules (when the data dictate that such tailoring is necessary). CRAM needs a capability to merge elements of several modules when that would fit particular results better than existing modules.

As noted in S-RHEINHARDT-3, CRAM is not designed to analyze mosaics. If a mosaic is encountered, multiple modules and multiple AAs would be needed for an assessment of an area. For the theoretical example provided by Wardrop (a combination in a small area of riverine wetland – slope wetland – depressional wetland), the area would be assigned to a dominant wetland class and assessed using the appropriate module. In addition, it can be noted that riverine assessments with CRAM explicitly take into account the riparian context. Stressors affecting the slope and depressional wetlands in the small area likely would be identified in the riverine assessment.

Aggregations of wetland subtypes into new wetland types, as described by Wardrop, are possible in CRAM, and may be studied in the future. Wardrop states that: “…my experience has been that it is important to test for inconsistencies in lumping or splitting wetland classes in rapid assessment techniques.” This sort of testing has been and continues to be part of the CRAM developmental process, under the ongoing guidance of the L2 Committee. As new circumstances are encountered, the question of “lumping vs. splitting” is answered based on data collected in the field. While new wetland classes are only recognized when the existing modules are deemed non-applicable, tailoring of modules to new contexts is ongoing.

For example, in response to the numerous large alternative energy projects proposed in the deserts and inland mountains of southern California, a considerable effort is under way for development or adaptation of CRAM modules for use in arid riverine systems. This process is providing practical experience in applying CRAM to new places, evaluating how it responds, and applying that experience to real projects.

The separation of L1 and L2 assessments is a major message that is addressed separately in Part 2.

GROUP 2 ISSUES

(related to CRAM Manual Chapters 2, 5, and Appendix III; Technical Bulletin Sections 2,3, and 4)

(4) Sources of stress or pressure affecting the condition or state of wetlands are identified in CRAM. This design facilitates management *responses* to prevent or mitigate undesirable effects. CRAM assumes that the "pressure-state-response" (PSR) model applies to wetland assessment and monitoring and that this framework may be used in CRAM to evaluate the state, or condition, of wetlands (CRAM Ch. 2.2.1).

CRAM utilizes the Pressure-State-Response model (PSR) of adaptive management 11. Holling 1978, 12. Bormann *et al*. 1994, 13. Pinter *et al.* 1999).

NIEMI- 4: The concept of pressure-state-response (PSR) model is relatively well accepted by ecologists and is being applied in numerous places in various forms across the globe. Moreover, the cycle of the adaptive management approach is equally well-accepted as a framework for evaluation, planning, action and monitoring. The PSR model can be defended as a general guideline as long as one avoids the implication to the philosophical arguments of cause-and-effect. The key concern is whether these steps are actually fully followed and implemented by management agencies. This concern applies to the implementation of CRAM in California in light of monetary resource issues, in particular, but also commitment on the part of management agencies over the long term. Not a technical matter

S-NIEMI-4: No response warranted

ANDERSON-4: There is a solid foundation of literature stressing the use and value of adaptive management and Pressure-State-Response modeling. CRAM reportedly uses this Pressure-State-Response approach and best sums it up in the last sentence of 2.2.1 (although there appears to be a typo) "For the purposes of CRAM the PSR model is simply used to clarify that CRAM is mainly intended to "described" state conditions of wetlands". I assume the authors are trying to put CRAM in a theoretical framework but I don't truly see the need for this.

S-ANDERSON-4: This will be addressed with additional text as the CRAM manual is revised.

RHEINHARDT-4: Basing the CRAM condition assessment framework on PSR theory is reasonable. The approach taken in developing CRAM was to base the assessment framework on condition (or state) and use of stressors (as a checklist) to independently identify pressure or response. (An exception was made for invasive species, which is a stressor that was also used as an indicator.) However, other stressors could have been used to provide indicators of condition. For example, the presence of undersized culverts that restrict tidal exchange could just as well have been used as a straightforward indicator of hydroperiod alteration in perennial tidal marshes. Therefore, although stress and condition were separated in CRAM, it seems that stressors could have been used advantageously in more instances to help explicitly evaluate condition.

S-RHEINHARDT-4: The development of a stressor index and other methods of stressor evaluation are addressed separately in Part 2. Development work regarding the effects of stressors on wetlands through CRAM assessment is ongoing.

WARDROP-4: This statement is judged to be based upon sound scientific knowledge, methods, and practices***.***

Additional comments: The P-S-R model is a useful, adaptive management construct, and its use in the technical background documentation of CRAM is well-appreciated.

However, if you adhering to this concept strictly, and consider its application to the CRAM method itself, the inclusion of buffer and landscape context in CRAM metrics is inconsistent. They can be considered "pressure" measures, not "state" ones, and have been shown time and time again to be related to a number of condition indicators. At best, their inclusion in the overall score is circular. I would assess whether the CRAM gives the same result both with and without their inclusion. It is admirable that this issue is raised in Section 2.3.1, but quantitative assessment of the effect of the metric's inclusion/exclusion, or a sensitivity analysis, would be illuminating.

S-WARDROP-4: The separation of L1 and L2 (CRAM) assessments is addressed separately in Part 2. Additional CRAM development work regarding the effects of stressors on wetland condition is ongoing.

(5) CRAM assumes that ecological conditions can be evaluated based on a fixed set of observable indicators, and that conditions respond to variation in natural and anthropogenic stress in a predictable manner. (CRAM Ch. 2.2.2 - 2.2.3)

These assumptions are common to most wetland rapid assessment methods (14. Fennessy *et al*. 2004,).

NIEMI-5: The assumption that ecological condition can be evaluated based on observable indicators is a very reasonable assumption. Tests on whether ecological condition can be predicted in response to variation in natural and anthropogenic stress have been elusive, but some examples exist. McDonald and I reviewed the use of ecological indicators (Niemi and McDonald 2004) and clearly this is the primary assumption and use of ecological indicators. In addition, indicators are used as an early-warning signal of ecological problems and as barometers for trends in ecological resources. There is a substantial literature to support these assumptions.

S-NIEMI-5: No response warranted

ANDERSON-5: Agreed; these assumptions are common to most rapid assessment methods. Ecological conditions are evaluated based on a fixed set of observable indicators, and that conditions respond to variation in natural and anthropogenic stress in a predictable manner.

This is common in CRAM and other methods. I have looked extensively at the metrics for developing our own rapid assessment metrics and find them to be grounded in good (albeit sometimes incomplete data). However this should not be viewed negatively as obtaining all of these data would take additional substantial time and money.

S-ANDERSON-5: No response warranted

RHEINHARDT-5: The stated assumption is probably scientifically valid in that the condition of an individual indicator (metric) probably responds to stress in a predicable manner, though perhaps not in a linear manner. The only way to test condition response is to find an array of sites that have only one stressor (the same one) that varies in intensity.

Alternatively, one could conduct largescale ecosystem field experiments, but that would be impractical. In either case, one could determine how indicators of condition vary along a continuum of stress. However, in most realworld sites there are usually multiple stressors involved in controlling a wetland’s condition, and so much interaction among stressors, that is often impossible to quantify the response in condition along a gradient of alteration. One alternative would be to recognize the fact that there are multiple stressors and develop assessment criteria that use multiple indicators for a given condition state.

S-RHEINHARDT-5: No response warranted

WARDROP-5: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: This statement is full of sub-phrases incorporating undeclared nuance, the most obvious being, “in a predictable manner”. While the statement above is sound, its embodiment in CRAM is what is critical, not the statement itself. Its embodiment cannot be assessed without a detailed review of the data itself, i.e., was predictable assumed to be a linear relationship between stress and any given indicator, or were other relationship types considered?

In addition, the following statement is incorrect:

> CRAM metrics were built on this basic assumption according to the following three criteria common to most wetland rapid assessment methods (Fennessy et al. 2004):
>
> > *the method should assess existing conditions* (see Section 2.1 above), without regard for past, planned, or anticipated future conditions;
> > *the method should be truly rapid*, meaning that it requires two people no more than one half day of fieldwork plus one half day of subsequent data analysis to complete; and
> > *the method is a site assessment* based on field conditions and does not depend largely on inference from Level 1 data, existing reports, opinions of site managers, etc.

The three criteria were not used to select the individual metrics which constitute the construction of CRAM, they are used to screen which metrics can be obtained according the constraints of a rapid method, i.e., they define the pool of potential metrics, not the final list of those which are appropriate indicators that respond to stress in a predictable manner.

S-WARDROP-5: These criteria were used both ways. They were used to screen existing metrics provided by other methods and they were used to guide the development of new metrics. Revision of the CRAM manual is possible to clarify the wording of this topic.

> **(6) In CRAM, wetland condition is defined as the ability of a wetland to maintain its complexity and capacity for self-organization with respect to species composition, physio-chemical characteristics, and functional processes, relative to healthy wetlands of the same type. CRAM assumes that wetland condition is based on an evaluation of wetland location, form and structure**. **(CRAM, Ch. 2.2.3, Appendix III – Glossary)**

NIEMI-6: The definition of wetland condition is reasonable. I am not going to wordsmith the document, but the concepts of history and ecological processes should be incorporated into complexity and the capacity for self-organization.

S-NIEMI-6: No response warranted.

ANDERSON-6: The definition of wetland condition is suitable and appropriate because it includes reference to healthy wetlands and includes the concept of self-organization as it relates to biotic and abiotic characteristics.

S-ANDERSON-6: No response warranted.

RHEINHARDT-6a: This is a scientifically reasonable definition of condition, particularly the part that defines condition as being "relative to healthy wetlands of the same type." The terms "location, form, and structure" were not defined in the field manual or glossary, so it is difficult to determine how wetland condition can be determined using these three characteristics.

S-RHEINHARDT-6a: Definition of "location, form and structure" will be addressed as the CRAM manual and other supporting materials are revised.

RHEINHARDT-6b: If the three characteristics refer to geomorphic position, composition, and physical structure (habitat related structure), then the characteristics could provide a context for evaluating condition. However, a wetland's geographic location should not be used as a criterion for determining condition, except in context to its surrounding landscape, which can constrain what is possible.

The real test of a robust assessment is whether conditions intended to be measured can be done so objectively, in a repeatable manner, and can successfully differentiate wetlands of different condition.

S-RHEINHARDT-6b: The separation of L1 and L2 (CRAM) assessments is addressed in Part 2.

See also S-WARDROP-3 above and S-RHEINHARDT-25 below. The classification in use in CRAM does need to be considered in an overall, statewide classification framework, with an element in the methodology that allows tailoring modules to "new" contexts without having to develop numerous completely new modules. CRAM needs a capability to merge elements of several modules when that would fit particular results better than existing modules.

Peer review request statement 6 asserts that "CRAM assumes that wetland condition is based on an evaluation of wetland location, form and structure." The intended meaning of this statement was that condition of a wetland cannot be evaluated without consideration of the geographic, geologic and hydrologic context of the wetland area being assessed. CRAM's developers agree that the "geographic location" of a wetland, in terms of where it is in the state as a whole, does not pre-determine condition score, although variations in hydrology and other physical and biological parameters that vary geographically certainly do affect wetland condition across this highly varied state, and an aim for CRAM has been to help in understanding this variability.

CRAM's developers have found that a wetland's *landscape position* does, in fact, affect physical, hydrological, and biotic characteristics of the wetland, and therefore its condition. In this sense, CRAM is consistent with Rheinhardt's opinion that "a wetland's geographic location should not be used as a criterion for determining condition, except in context to its surrounding landscape…"

WARDROP-6: This statement is unclear and not commonly utilized in current literature.

Additional comments: Self-organization is a concept often discussed in the complex adaptive systems literature, and its use as a foundational concept here is ill-advised. While complexity could be clearly described with respect to the three factors listed in the definition, the same cannot be said of self-organization.

My suggestion is to either drop the concept from the definition, or include a narrative describing clearly how the concept is utilized in CRAM. I would offer the observation that none of the metrics in CRAM, or their scoring, utilizes the concept of self-organization; e.g., which metric or its valuation, currently in CRAM, delineates a wetland with a capacity for self-organization from one that does not have such capacity?

S-WARDROP-6: Staff agrees with this comment. The concept of self-organization is applied to the delineation of the Assessment Area (AA), but not to CRAM *per se*. This will be addressed with additional text as the CRAM manual is revised.

(7) As part of the CRAM assessment, a checklist is provided that characterizes stressors as to their likely effect on wetland condition. It is assumed that the stressor checklist can be used by researchers and managers to explore possible relationships between condition and stress, and to identify actions to counter stressor effects. (CRAM, Ch. 2.2.1, 5).

NIEMI-7a: I could not find reference to this in Ch. 2.2.1.5 unless I was looking in the wrong document. It is unclear.

There is a "Guidelines to Complete Stressor Checklists" in CRAM 5.0.2, Chapter 5. As a checklist, these are reasonable but I do have some concern with the limited area of the assessment, especially 50 m. I realize that the “Buffer and Landscape Context Attributes” include assessments up to 500 m.

There are substantial ways that the influence of the surrounding landscape of a wetland can be improved (see Hollenhorst et al. 2007). For instance, wetlands and wetland biota can be influenced by direct stress to the wetland, adjacent disturbances (e.g.., 0-500m), and well beyond 500 m for disturbances that directly affect the contributing area of the watershed of the wetland.

I would foresee a better linkage between level 1 and level 2 assessments. Level 1 assessments with good geographic information systems (GIS) analysis (e.g., using aerial photography) could provide information that is simply verified by the level 2 CRAM. Note that there is a substantial amount of publically-available, spatial information that should be available to environmental regulators in the state of California (e.g., see variable lists in Danz et al. 2005, 2007) that could be used to better assess stress within the watersheds of the respective wetlands.

S-NIEMI-7a: The measurement and interpretation of the effects of stressors on wetland condition is a focus of ongoing development by L2. The relationship between L1 and L2 is addressed in Part 2.

NIEMI-7b: For instance, one variable not included is human population density, we found this to be an important variable in the Great Lakes (Brazner et al. 2007a,b). With California’s high human population, this is an important stress variable and is available at a reasonable spatial resolution. This will likely be even better with the next US census. Moreover, these data could then be used to better quantify the gradient of disturbance among the wetlands included in the validation (see below). In summary, the checklist is a reasonable start in the CRAM, but I suggest that future improvements could be made in this portion of the assessment and should eventually be included.

S-NIEMI-7b: The refinement of the stressor index is addressed in Part 2. Staff agrees that limiting considerations to 50 meters may not fully capture the effects of many regional stressors.

ANDERSON-7: This is a reasonable assumption. There are multiple scientific papers stressing the importance and impacts of many of these stressors. Similar stressors and ratings are used in numerous other rapid assessment methodologies including Ohio, Pennsylvania, and West Virginia.

S-ANDERSON-7: No response warranted.

RHEINHARDT-7: The four assumptions outlining the usefulness of a stressor checklist (Ch. 5, Rationale) are all reasonable. The checklist itself (Worksheet 5.1) is comprehensive, although some stressors might be difficult to identify during a typical site visit. While it is true that the stressor checklist could be used “by researchers and managers to explore possible relationships between condition and stress, and to identify actions to counter stressor effects,” in CRAM the end user also uses the list to decide how to rate wetland condition. Unfortunately, the stressor checklist only provides two levels of stress for the user to identify and both choices require using BPJ: (1) “Present and likely to have negative effect on AA”” or (2) “Significant negative effect on AA.” The result is that BPJ (and hence, subjectivity) has to be invoked in rating wetland condition.

Specific suggestion/comments to the stressor list

(1) A detention basin is identified as a stressor, but a detention basin could also be viewed as a response to a hydrologic and nutrient stressor (i.e., it ameliorates runoff from impervious surfaces and eutrophication). Thus, it seems that this “stressor” could both improve and degrade wetland condition. The end user is left having to use BPJ to decide whether the effects of a detention basin are more negative than positive.

(2) Stormwater discharges were not explicitly listed as a stressor although non stormwater discharges were (line 1 under Hydrology). Stormwater discharges may be included under point sources, but if so, then this should be made more explicit.

(3) Stressors can originate within a wetland too, e.g., forest clearcutting, filling, ditching.

S-RHEINHARDT-7: Internal stressors are included in the stressor checklist. The development of a stressor index is addressed in Part 2.

WARDROP-7: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The question is unclear, and the references are incorrect (e.g., there is no Section 2.2.1.5).

Assuming that the statement refers to material that is described in Sections 2.2 and Section 5.1, my comment concerns the presence of buffer characteristics as a metric in CRAM, and the list of buffer and landscape attributes in the Stressor Checklist. Buffer characteristics have been shown to be valuable diagnostics in explaining disparities between wetland condition and landscape attributes (Wardrop et al., 2007b). Like the presence of invasive plants, which can be argued as either a metric of a stressor, a sensitivity analysis of CRAM concerning these metrics should be investigated. If they are found to have a significant role in the eventual CRAM score, it would suggest circularity in the reasoning and approach.

S-WARDROP-7: The development of a stressor index is addressed in Part 2. Staff notes that despite the difficulty caused by unclear page references, both Wardrop and Niemi correctly discerned the meaning and intent of Statement 7.

(8) CRAM accepts the primary assumption that the condition of a wetland is determined by interactions among internal and external hydrologic, biologic (biotic), and physical (abiotic) processes, as presented by Brinson,(15, 1993) and others (e.g., 16. Barbour, 1995). CRAM is based on a series of assumptions about how wetland processes interact through space and over time. First, CRAM assumes that the condition of a wetland is mainly determined by the quantities and qualities of water and sediment (both mineral and organic) that are either processed on-site or that are exchanged between the site and its immediate surroundings, Second, the supplies of water and sediment are ultimately controlled by climate, geology, and land use. Third, geology and climate govern natural disturbance, whereas land use accounts for anthropogenic stress. Fourth, biota tend to mediate the effects of climate, geology, and land use on the quantity and quality of water and sediment. Fifth, stress usually originates outside the wetland, in the surrounding landscape or encompassing watershed. Sixth, buffers around the wetland can intercept and otherwise mediate stress . *(*CRAM, Ch. 2.2,3 *).*

Most of the structure of CRAM is found in these assumptions. (17. Collins, pers. comm.., 2008e)

NIEMI-8: The assumptions are all fine and supportable via the scientific literature. Effective implementation of these assumptions is the difficult part.

S-NIEMI-8: No response warranted.

ANDERSON-8: Six assumptions are made regarding CRAM. Like in previous statements these are all mostly items that are based on common scientific knowledge and principles. There is adequate literature evidence to support all 6 of these assumptions. I see no issue here.

S-ANDERSON-8: No response warranted.

RHEINHARDT-8: All of the assumptions listed above are scientifically defensible. However, many stresses can originate within a wetland (see assumption #5), e.g., forest clearing, construction of drainage ditches, channelization of streams, etc. These examples of internal stressors can be quite common, are easily identified, and indicators of these stressors could be quantified to help evaluate condition.

S-RHEINHARDT-8: No response warranted.

WARDROP-8: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: Two minor revisions could be made to the text. Firstly, the language following the statement, “CRAM is based on a series of assumptions about how wetland processes interact through space and over time”, deals only with the interaction of processes in space; no mention is made of any temporal interactions. Either a discussion of temporal interactions needs to be added, or the phrase simplified to indicate that only spatial factors are considered. Secondly, disturbance needs to be added to the statement concerning the sixth assumption, in order to be consistent with Figure 2.2.

S-WARDROP-8: This will be addressed with additional text as the CRAM manual is revised.

(9) The metrics used in CRAM are ecologically meaningful, field-based measures that record the potential range of conditions in a wetland's hydrology, physical structure, biotic structure, and adjacent landscape, and are formulated to have a response to variations in stress that is distinguishable from natural variation. . (CRAM, Ch. 2.3.1 and 2.1).

CRAM metrics are field-based, ecologically meaningful, and have a dose-dependent response to stress that can be distinguished from natural variation across a stressor gradient (18. Fennessy, et al., 2004; 19. Fennessy, et al, 2007; 20. National Research Council, 2001; 16).

NIEMI-9: The separation of response to stress and distinguishing this from natural variation requires a long term data set to examine natural variation and responses to specific human-caused events. It is, however, a reasonable assumption if the field-based measurements of wetland quality are related to a broad gradient that defines natural, high quality wetlands from those that have been degraded by human activity. The application of CRAM to a gradient of wetland conditions is a reasonable assumption in this context.

S-NIEMI-9: No response warranted.

ANDERSON-9: I do not believe that the metrics are completely independent from natural variation but I agree with the statement that the variation in stress can be distinguished from natural variation. Wetlands are subject to tremendous natural variation that prevents simple classification and categorization of all natural variation. However the rating scale and simple checklist of stressors does allow for the natural and anthropogenic forces to be separated.

S-ANDERSON-9: No response warranted.

RHEINHARDT-9: The four wetland attributes recognized by CRAM are reasonable. Other types of attributes could just as well have been recognized, e.g., biogeochemical attributes. For the most part, the attributes identify stressors and responses to stress that could be used to differentiate variation due to stress from natural variation. However, the metrics used in CRAM to assess these variations vary tremendously from being objective, precisely measureable criteria that are ecologically meaningful, to being vague, subjective metrics that require invoking best professional judgment (BPJ). The degree of subjectivity required to rate a particular attribute affects its ability to determine condition in a repeatable manner. A couple of specific examples of subjective and objective attribute descriptions are discussed in more detail under #12.

Identifying indicators that are sensitive enough to differentiate condition might have been possible had wetlands had been pre classified and if the ratings had been customized to for each wetland type. This step appears to be missing from the design phase (see #10 below), but perhaps this could be done retroactively for the field books, each of which focuses on only one wetland class. Certainly, there is no need to include indicators and stressors in a field book that are not appropriate for the wetland class of focus. For example, there is no reason to include material on estuarine or riverine hydrology in a field book devoted to depressional wetlands. On the other hand, more detailed information on alterations, indicators, and metrics that pertain to the condition of depressions should be included in a depression field book.

S- RHEINHARDT-9: The need for objectivity is addressed in Part 2. As noted in S-RHEINHARDT-3a---c, the ability to use the same methodology across wetland types is considered a desirable characteristic of the methodology (although CRAM does incorporate class-specific variations in metric assessments).

WARDROP-9: This statement is judged to be based upon sound scientific knowledge, methods, and practices for the four wetland attributes recognized by CRAM; it does not judge whether the four recognized attributes are the most appropriate of all possible attributes.

Additional comments: the text that reads, "and are formulated to have a response to variations in stress that is distinguishable from natural variation" implies that the metrics were manipulated to have the desired relationship to estimates of stress. In many IBI techniques, a criteria for metric inclusion is that it has a **predictable** response to stress that is distinguishable from natural variation, i.e., the system is responding according to a conceptual ecological model.

S-WARDROP-9: No response warranted.

(10) CRAM has separate adaptations – called "modules" for each major wetland type. A module is developed in a nine step process organized into three phases: basic design, calibration, and validation. This developmental framework results in a valid Level 2 assessment method. (CRAM, Ch. 2.3)

This design process is discussed in greater detail by Sutula, et al. (21. 2006a; 22. 2006b;), Stein, et al. (23. 2009; 24, 2007) and in the CRAM Quality Assurance – Quality Control Plan (CQAQC; 25 Collins, et al 2005). This process is consistent with recommendations published by various authors (26. Hruby, et al., 1999; 27. Hruby, 2001; 19).

NIEMI-10: I have serious concerns regarding the verification and validation phase of the CRAM procedure. One place where these results are reported is in Sutula et al. (2006). For instance they state "*To verify CRAM, the three Regional Teams selected 118 wetlands representing high quality and low quality conditions for each of the wetland classes. The a priori classification of condition was based on consensus of the experts following their review of pertinent site-specific reports and field visits. Verification was conducted through the Regional Teams, and the regional results were complied into one verification dataset."* However, they state *"This provided **preliminary** assurances that the draft metrics were able to distinguish among wetlands of varying condition."* They further go on to state - *"In both calibration and validation, the concept is to determine whether RAM metric, attribute, and overall index scores are good predictors of wetland condition, as measured against the results of more intensive measures of wetland condition."* **However, they go on to identify cost as a prohibitive factor that these intensive measures were not really completed nor was a statistically-designed sampling process included.**

This paper was then followed up by a paper by Stein et al. (2009) which strived *to "present a case study of the validation of the riverine and estuarine modules of the California Rapid Assessment Method (CRAM)."* This paper presents a substantial amount of information to validate the CRAM procedure including data on avian diversity, benthic macroinvertebrate indices, and plant community composition. These are all important components of wetland ecosystems and provide an important step toward validation. I applaud the authors for publishing these data in the peer-reviewed literature and in an appropriate journal. My concerns are related to the following issues and the potential for broad application of CRAM. However, note I am not criticizing the paper per se – I am glad it was published but my major question is whether it is adequate as the primary basis for accepting CRAM across the broad state of California for these two wetland types?

These issues are the following. 1) The validation was confined to the riverine and estuarine classes, plus they noted that depressional wetlands, vernal pools, seeps and springs, and lake and lagoon fringe wetlands will occur in the future. 2) The total sample size was 95 riverine sites (54 with macroinvertebrate data and 41 with bird data) and 38 estuarine sites (all with vegetation data only). How does this compare to the number of wetlands of these two types in California? Also these biotic data represent a small subset of the biotic community and none of them include information about ecological processes. 3) Many of the relationships are not very convincing – e.g., see Fig 1.c; Fig. 2b or 2c. There are definitely a number of outliers in these figures. What if some of these wetlands were identified as of poor quality, but actually contained a few important wetland-dependent bird species? Wetland-dependent species should be analyzed separately as should species that are endangered, threatened, or special concern (Peterson and Niemi 2007). An analysis of those wetlands that did not "fit" the pattern is important to examine further. Overall bird species richness is not a good metric to use. It is too general – at minimum the analysis should have been completed for wetland-dependent species. It is not clear what role the individuals gathering MAPS data were involved in the interpretation of these bird data? Also it is unfortunate that amphibians were not included in the analysis, their distribution and use of the wetlands are easy to monitor. 4) Finally, I do appreciate the "*Final Thoughts"* in this article. It is clear that the authors recognize the limitations of the methodology and it is likely cost that has prohibited them from gathering the appropriate comprehensive data to fully validate CRAM. In summary, I do not think the procedure has received adequate validation. Validation should require the gathering of data unique to the CRAM process and should include a wide spectrum of the wetland characteristics including benthos, birds and plants, but also amphibians and fish as well as water quality. Several professional statisticians should also be consulted to receive their opinion on the adequacy of the validation process and interpretation of the results.

S-NIEMI-10a: The need for further validation is addressed separately in Part 2.

Statisticians from USEPA and California Sate University Monterey have participated, and continue to participate, in the design and analysis of CRAM.

Validation of CRAM is an ongoing process. Staff notes that the validation work for CRAM that has been done to date is among the most rigorous validations ever done for any rapid assessment method, and that the use of a "weight of evidence" approach provides additional certainty to the validation studies.

It is expected that continuing refinement of current and possible new validation methods will be a topic addressed by the Level 2 committee.

Funding: Niemi notes that perceived deficiencies in validation of CRAM may in part be tied to funding constraints. Although not strictly a scientific issue, staff notes that funding always constrains methods. CRAM was developed using the best available means obtainable with available funding levels. Future development is expected to continue under those same constraints. Staff finds it to be unreasonable to forestall development of CRAM until funding for all possible or desirable methods is available.

NIEMI-10b: Ideally, given the importance and breadth to which CRAM will be applied in California, a solid field-based and statistically valid sampling regime should have been completed to validate the CRAM procedure. The use of data that happened to be available for a subset of wetlands is not adequate as a validation process.

S-NIEMI-10b: This comment does not criticize the overall scientific approach to module development used in CRAM but criticizes what is considered to be insufficient underlying scientific (L3) data. Reviewer Niemi perceives a lack of data and the lack of funds to get the data.

Niemi is certainly correct that CRAM as it exists today needs additional development and validation in order to reliably serve all purposes for which it may be useful in the whole state. However, it's important to understand the planned phased introduction of CRAM as presented in CWMW Bulletin 1.

Applications of the California WRAMP, including CRAM, are being tested for various larger-scale projects at this time. These tests are contributing to site assessment and mitigation planning. It is expected that experience gained with these tests will help guide the next phases of CRAM development, including regulatory applications.

The Level 2 committee will be providing more detailed guidance for phasing, and will spell out the application steps clearly and emphatically. The Level 2 committee will also be providing more concise guidance on issues such as sampling and statistical methods.

CRAM has many useful applications now, and a requirement to begin running CRAM in a regulatory context is appropriate for those initial, less complex uses (such as ambient monitoring, which might not directly contribute to a project site assessment) and for the purpose of building the data set.

The adequacy - or inadequacy - of CRAM for various applications should not be asserted without incorporating this context.

Staff also notes that failure to continue development until all the issues are resolved would be impossible. By continuing with development and phasing-in application of CRAM, all of the issues raised by Niemi can be addressed.

ANDERSON-10: The CRAM development methodology is well-developed and well-respected. Several additional states have developed similar methodologies but with different metrics. Nonetheless, these other states used a similar methodology of design, calibration, and validation. Moreover, as pointed out in the question, the CRAM was partially based on other wetland rapid assessment literature that has been through the peer review process.

S-ANDERSON-10: No response warranted

RHEINHARDT-10: During the basic design phase of the CRAM development process (Table 2.1), all four steps would have been more effectively accomplished if wetlands had been classified first and then attributes and metrics identified in the context of specific wetland subclasses. This would have helped focus the choice of metrics, refined how metrics were to be measured, and made the descriptions of alternative states more focused on what is possible for each wetland class than what is possible for all wetlands. This focus could have reduced the reliance on using BPJ to rate condition. Therefore, although the calibration and validation phases in Table 2.1 seem reasonable, it will likely be impossible to validate the attributes that require subjective interpretation (see #4 above).

S-RHEINHARDT-10: As noted in the discussion of statement 3 above, the wetland classes were developed before CRAM (based on existing HGM categories), such that experts in different classes and from different regions could be assembled as a development team to increase the likelihood that CRAM would be applicable across wetland classes. It has been found that the overall structure of CRAM *is* applicable across classes, but that metrics or ratings may need to be class-specific. These adjustments are made as part of the CRAM development process, and will continue to be made by the L2 when necessary.

WARDROP-10: This statement cannot be assessed without accompanying definitions of design, calibration, and validation. No definitions are given, just the processes used in CRAM to address them.

Additional comments: The following references should be consulted by the authors, and a definition of each term included in the glossary:

Mitro M.G. 2001. Ecological model testing: verification, validation, or neither? *Bulletin of the Ecological Society of America*, 82:235-236.

Oreskes N, Shrader-Frechette K, Belitz K. Feb 1994. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263:641-646.

Rykiel, E.J., Jr. 1996. Testing ecological models: the meaning of validation. *Ecological Modeling*, 90:229-244.

For example, Rykiel (1996) defines calibration as "*the estimation and adjustment of model parameters and constants to improve the agreement between model output and a data set." The steps listed under calibration in the CRAM manual would not fit this definition.*

S-WARDROP-10: The definitions used in CRAM will be clarified when the CRAM manual and other supporting materials are revised. Staff notes that the term "calibration" was used in the CRAM development process to scale metrics to reflect the content of external data sets, which appears to be much the same process that Wardrop cites above.

(11)CRAM is calibrated through repeated (iterative) field evaluations of the metrics, where the descriptions of the range of potential conditions are adjusted in relation to the other. These exercises are repeated until the calibrations for all metrics work together to provide results that are replicable by any observers in any similar wetland. (CRAM Ch. 2.3.2).

The calibration phase began with the Regional Teams selecting sets of wetlands that clearly represented a broad range of conditions. Data from these assessments were used to adjust the number of metrics and the number of alternative states of some metrics, and to revise the text of the CRAM assessment forms and within the CRAM Users Manual (25).

NIEMI-11: Many of the same issues identified in response 10 above are relevant here as well. The same level 3 data (e.g., MAPS, EMAP) were used for the calibration.

S-NIEMI-11: No response warranted

Anderson-11: CRAM has undergone extensive calibration of metrics and will continue to undergo calibration. This is a sensible approach and is consistent with an adaptive resource management framework, where additional information is used to inform management decisions. These changes in metrics and methodology have declined and will continue to decline, but as mentioned this is a good approach to follow. Repeatability among observers is critical for a useful rapid assessment technique.

S-ANDERSON-11: No response warranted

Rheinhardt-11: This approach to iterative calibration is reasonable and necessary. However, experts in the Regional teams may tend agree when they invoke BPJ, but not all end users will be experts and some users may have biases or agendas that would preclude them from obtaining the same ratings as the experts who test the method. Each metric and alternative state should be examined to see if there is room for users to misinterpret the rating. For example, if there can be a disagreement about what constitutes an “intermediate mix of native and nonnative vegetation” vs. “substantial amounts of nonnative vegetation,” then there is potential for error, either accidental or purposeful, which testing by an expert Regional team won’t be able to identify. Disagreement and misinterpretation can be minimized by using the narrative to “box” the user into the correct (intended) answer by using clear, objective criteria to rate all alternative states.

S-RHEINHARDT-11: The need for objectivity is addressed in Part 2. Ongoing refinement of metrics and training methods can address the potential for misinterpretation.

WARDROP-11: This statement cannot be assessed without an accompanying definition of calibration (see above comment to Issue 10). The objectives of the calibration phase, as described in Section 2.3.2, do not meet most available definitions of calibration.

Additional comments: Section 2.3.2. states:

> The calibration phase was used to determine if the draft wetland classification scheme, the attributes, the metrics, and the narrative descriptions of alternative states were (1) clear and understandable; (2) comprehensive and appropriate; (3) sensitive to obvious variations in condition; (4) able to produce similar scores for areas subject to similar levels of the same kinds of stress; and (5) tended to foster repeatable results among different practitioners. The calibration phase was also used to test and select methods of calculating, scaling, and weighting scores for metrics, attributes, and AAs.

If we utilize the Rykiel's definition of calibration as "*the estimation and adjustment of model parameters and constants to improve the agreement between model output and a data set",* it is clear that the steps listed above would not fit this definition. Calibration should entail the adjustment of the metrics to Level 3 information, not an adjustment to metrics to make them either more stable between field personnel (a QA/QC issue), or to make sites comparable according to Best Professional Judgement. Another method of calibration would be to adjust the metrics so that the results are at least anchored at the opposing ends of the condition spectrum, i.e., that extreme values of the metrics represent the extreme ends of the range of condition.

S-WARDROP-11: The process used by the PI Team in developing CRAM included evaluations of existing data sets and modifications of the conceptual metrics and the overall methodology to reflect the range of conditions in the data sets. The PI Team named this process "calibration." Waldrop's comment indicates that other definitions exist for this term. Staff does not find that the existence of other definitions for the term "calibration" means that the use of this term in the CRAM development process constitutes a scientific error.

Currently no "gold standard" of overall condition exists for any wetland class against which CRAM can be calibrated; all that is possible is to regress L3 data on CRAM scores and see if the results are consistent with expectations based on conceptual models of how Level 2 assessments should track L3 information. Oversight for this process has been assigned to the L2 Committee.

Staff notes that "calibration" also refers to a part of the CRAM process that pertains to the QAQC element whereby different CRAM assessment teams (field teams) are independently calibrated to a set of common test sites. This is termed inter-team calibration. Staff notes that few similar efforts document taking this extra step.

The Level 2 committee will continue to strive to reconcile inconsistencies in terms used.

(12) CRAM's condition-based rapid assessments can be expected to reliably discriminate between wetlands of moderately different condition classes, after appropriate data collection, QA, calibration and validation of a sufficient body of data. (CRAM, Ch. 2.3.2; Technical Bulletin Sec. 3.J)

NIEMI-12: It is unclear to me the extent to which CRAM can discriminate "*between wetlands of moderately different condition classes.*" I do not see appropriate data within the material provided to assess this ability to discriminate. I am concerned that there is not an independent "gradient" that was derived to identify those wetlands (whether riverine, estuarine, or depressional) that are of good or bad condition.

S-NIEMI-12: Since index scores (i.e. "overall scores") are averages of attributes scores, high index scores mean all attribute scores (and all metric scores) are also high. Low index scores mean all attribute and metric scores are low. But many possible combinations of high and low attribute and metric scores can yield the same intermediate index score.

Wetlands with intermediate index scores can only be distinguished based on their different metric scores (and sometimes their attribute scores), but not their index scores; distinguishing between intermediate-condition wetlands depends on the examination of their attribute and metric scores, and the precision of these scores. With enough precision, intermediate-condition wetlands could be distinguished by a single different metric score.

CRAM generally identifies a difference of ±10 points in any (metric, attribute, or index) score between two wetlands as indicating real differences in condition between the two wetlands in that metric/attribute. This is the context of "moderately different condition classes" identified in the above question. CRAM is expected to consistently and reliably indicate condition differences when the wetlands are substantially different from one another (scores differing by >10 points), and CRAM is expected to generally and reliably indicate condition differences when the wetlands are only moderately different.

The ability – or inability - of wetland rapid assessments to discriminate between wetlands of moderately different condition classes is often cited as a problem with rapid assessments in general. Field tests of CRAM sensitivity along single – stressor gradients are planned to help address this concern. Additional tests are under consideration, and may be conducted as needed and funds permit. Tests involving independent (i.e., non-CRAM) assessment of "good" or "bad" condition using quantitative scientific approaches are possible.

Until refinement of the method's capability is accomplished as part of the method's development process, Staff notes that CRAM scores can still be reliably applied to many useful purposes as long as limitations are recognized.

ANDERSON-12: True. That's one of the reasons for doing a rapid wetland assessment technique. Small differences in condition class are very difficult to detect but a moderate target is suitable.

S-ANDERSON-12: No response warranted.

RHEINHARDT-12: CRAM is unlikely to reliably discriminate between wetlands of moderately different condition classes, since half of the metrics used to rate condition depend on subjective interpretation (see Table 1, p. 24, under "The Big Picture." CRAM may be able to discriminate the condition of metrics for which its rating can be objectively derived, but it will not always be able to reliably discriminate between ratings of metrics that require subjective interpretation.

An example of a metric where objective, ecologically meaningful rating criteria are provided is "Structural Patch Type Richness" (Table 4.16). The rating for this metric is based on the number of patchtypes found (an objective metric) where patch types (indicators) have been thoroughly defined (Section 4.3.1.1). In contrast, "Buffer Condition" (Table 4.8) is an example of a condition whose rating criteria are subjective, i.e., it relies on best professional judgment (BPJ) for scoring.

The indicators used to justify buffer condition ratings only describe what constitutes "good buffer conditions (an A rating)," i.e., "prevalence of native vegetation, absence of exotic vegetation, absence of recent substrate disturbance, and absence of trash or debris." Thus, Rating A can probably be determined objectively. However, it would be difficult to objectively justify Ratings B through C, since they rely on subjective criteria to discriminate condition, i.e., BPJ must be invoked to differentiate:

(1) "an intermediate mix of native and nonnative vegetation" (Rating B) from "substantial amounts of nonnative vegetation" (Rating C),

(2) "mostly undisturbed soils" (Rating B) from "moderate degree of soil disturbance" (Rating C) from "highly disturbed soils" (Rating D), and

(3) "little or no human visitation" (Rating B) from "a moderate intensity of human visitation" (Rating C) from "very intense human visitation" (Rating D).

S- RHEINHARDT-12: See S-NIEMI-12 above. The reviewer suggests that a move toward less subjective and more objective metrics could improve their reliability by reducing the use of "subjective" metrics. Many CRAM metrics are not quantitative, but are still not highly subjective. These metrics are categorical, and while the assignment of a particular metric score into one of the metric categories requires practitioner judgement, the categories are generally distinct from one another, and CRAM implementation has demonstrated that adequately trained field personnel are able to make consistent assignments to these categorical metrics. The need for objectivity is addressed in Part 2.

WARDROP-12: This statement cannot be assessed without accompanying definitions of design, calibration, validation, and sufficient data.

Additional comments: If the individual portions of the rapid assessment cannot be assessed for their scientific credibility in their definition or use, neither can the cumulative result.

S-WARDROP-12: This will be addressed with additional text as the CRAM manual is revised.

(13) Procedures are in place to refine CRAM metrics through calibration studies to improve discrimination between wetlands exhibiting moderate differences in condition. (CRAM, Ch. 2.3.3; 25).

The USACE review of CRAM notes that "…[the] statewide generalization approach and the combination of all metrics into a single grand condition score may not produce sufficient discrimination among wetlands of moderately different condition classes to be useful in some situations." (28. Klimas, 2008.). This concern is addressed through specified calibration and validation procedures (Technical Bulletin, Sec, 4.E), although it is understood that the precision of CRAM scores has limits.

NIEMI-13: It is good that procedures are in place to refine CRAM metrics through calibration studies. I would agree with the comments of Klimas that it can be difficult to discriminate among wetlands of moderately different condition – though the term "moderate" could be operationally defined. This level of uncertainly could also be statistically analyzed with appropriate field data and the calibration process (e.g., repeated measurements by the same individual and other individuals) could be incorporated into the analysis.

S-NIEMI-13: See S-NIEMI-12 above. A definition of "moderately different wetland condition" will be incorporated as the CRAM manual is revised.

ANDERSON-13: CRAM has undergone extensive calibration of metrics and will continue to undergo calibration. This is a sensible approach and is consistent with an adaptive resource management framework, where additional information is used to inform management decisions. These changes in metrics and methodology have declined and will continue to decline, but as mentioned this is a good approach to follow.

S-ANDERSON-13: No response warranted.

RHEINHARDT-13: Klimas (2000) made a good point in stating that many CRAM ratings are unlikely to discriminate among conditions, particularly when metrics are combined into a single grand condition score. The lack of discrimination, particularly in the middle ranges (ratings of B and C), is likely due to the subjectivity inherent in some of the rating criteria (see #9 above). It is doubtful that calibration and validation can address this problem, unless the specific, subjective criteria currently used as a basis for ratings are replaced with objectively differentiable criteria. This doesn't necessarily mean that additional metrics have to be measured. For example, there could be a list of negative indicators (or even stressors) and the rating could be based on the number of negative indicators identified, a weighted average, or some combination of indicators (a multivariate approach) present at an assessment site.

S-RHEINHARDT-13: See S-NIEMI-12 and S-RHEINHARDT-12 above.

WARDROP-13: This statement cannot be assessed without an accompanying definition of calibration (see above comments to Issues 10, 11, and 12).

Additional comments: This statement is the same as that present in Stein et. Al, 2009, which states the following (context provided to make the following point):

> In this paper, we present the results of validation of the California Rapid Assessment Method (CRAM; Collins et al. 2006) in riverine and estuarine wetlands. We define ''validation'' as the process of documenting relationships between CRAM results and independent measures of condition in order to establish CRAM's defensibility as a meaningful and repeatable measure of wetland condition. The overall validation process includes several steps designed to meet the following objectives: 1) assure that the method is producing meaningful results based on a comparison between CRAM scores and independent measures of condition (evaluation), 2) make adjustments to the method, where needed, to improve the ability of the CRAM to discern differences in wetland condition (calibration), and 3) minimize observer bias by assessing repeatability between independent assessment teams and modifying metrics that lead to inconsistencies (standardization).

In the above, calibration is defined as one step in the validation process. If this definition of calibration s used, the issue statement can be judged as consistent with sound science. However, in the CRAM manual, calibration is an undefined separate step of the process from validation. In addition, the calibration as described does not address the USACE issue. Many attempts at reconciling condition assessment scores (as the dependent variable) with some measure of stress (as the independent variable) demonstrate a large disparity or lack of correspondence in the middle region of condition or stress. This may be due to a fundamentally different basis of ecosystem response than the one assumed (e.g., alternative stable states or the presence of thresholds in one variable or another), a controlling variable that has not been measured as a part of the assessment, or a poorly conceived expression of stress. It is not clear what is encompassed, or intended, by "procedures are in place to improve discrimination".

S- WARDROP-13: See S-WARDROP-11 above. Staff notes that the process described in Stein et al. (2009) constitutes a use of output from the assessment process to increase the ability of the methodology to differentiate wetlands of differing condition. The CRAM development team termed this process "calibration." Even though the methodology does not explicitly define this term, this use comports with the definition cited by Wardrop "to improve the agreement between (model) output and a data set."

(14) The stated within-team and between-team precision of CRAM of plus or minus 10% for attribute scores and overall site scores is acceptable for Level 2 conditional assessments of wetlands. (CRAM, Ch. 2.3.2; Technical Bulletin, Ch. 4.B).

CRAM precision is calculated as the average difference in CRAM scores independently produced by different trained practitioners or teams of trained practitioners for the same wetland area and assessment period. Precision is calculated for CRAM metrics, attributes, and for the overall site or index score.

The latest version of CRAM sets the precision target for attribute and overall scores at ±10%. More recent guidance incorporates the same precision targets as part of the process for determining the number of required assessments (Technical Bulletin, Sec. 4.B).

Statistical limits of confidence are currently being calculated for estuarine wetlands in four coastal regions based on the 2007 statewide survey of estuarine wetland condition using CRAM. The results will help practitioners quantify the probability that two CRAM scores are statistically different, and will enable CRAM practitioners to compare individual scores to ambient condition with known statistical confidence (17, 23).

NIEMI-14: The 10 % for attribute scores seems like a reasonable level of within-team and between-team precision; however, this level of variation needs to be considered within the context of the precision upon which the program wants to identify “moderate” differences in condition classes. The higher the level of variation allowed in attribute scores results in less precision to discriminate among condition in wetlands. Again, consultation with a professional statistician is essential to adequately address this question.

S-NIEMI-14: No response warranted

ANDERSON-14: I think that almost all natural resources related field studies would be sufficiently happy with 10% precision. As long as training continues to be mandated for use of CRAM than these target numbers should be able to be met.

S-ANDERSON-14: No response warranted

RHEINHARDT-14: The level of precision required (10% or some alternate value) really depends on the level of precision that is considered sufficient relative to how CRAM is intended to be used.

Testing for precision is necessary to determine how confident one can be about the repeatability of scores. Scores among “trained practitioners” should be expected to be similar to one another and adjustments should e made if they are not, but it seems that tests should be conducted to determine the level of agreement between experts and trained practitioners (to test if ratings for alternative narratives are sufficient clear for users to get the intended answers) and between experts and untrained practitioners (to test to what extent training is needed for users to obtain the intended answers). If training is not required before one can legally use CRAM, then it would be useful to know how confident one can be that scores obtained by untrained users are similar to what would be obtained by experts.

S-RHEINHARDT-14: Training is required for all practitioners prior to their authorized use of the methodology. The training includes both "classroom" explanation of all metrics and field practice in measuring and inferring alternative metric states. Part of the training regime includes QA standards for evaluating the consistency of metric scoring, and the target for each training course is to achieve coherence in metric (as well as attribute and index) scores; i.e., differences in scores among trainees <10%.

WARDROP-14: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: I would expect that the USACE review has assessed the implication of this statement when utilizing CRAM in management and/or regulatory contexts, and such would be outside of the scope of this review. However, one approach that would be helpful in this regard would be the comparison of the 95% confidence interval of the CRAM score for a probabilistic-based survey of wetlands to this precision standard.

S- WARDROP-14: L2 will consider conducting this comparison using the results of the ambient survey of estuarine wetlands.

(15) The maximum CRAM score for a wetland type represents the best achievable condition for that wetland type in California. (CRAM, executive Summary; Ch. 3.5, 3.8, Technical Bulletin Ch. 3.J.).

Best Achievable Condition is defined as the state exhibited by selected sites that have been subject to the least levels of anthropogenic stress. This criteria is consistent with widely accepted recommendations on the selection of reference sites (29. Stoddard, 2006; 16). The overall score for a given wetland therefore indicates how it is doing relative to the best achievable conditions for that wetland type in the state.

NIEMI-15: This is completely logical.

S-NIEMI-15: No response warranted

ANDERSON-15: The concept of best achievable condition can be derived two ways: based on an idealized wetland taking into consideration all wetlands across types, location, etc. or the way CRAM was done which is by looking at the data and finding the best wetland based on the pool of surveyed wetlands. I agree with the CRAM approach but one must realize that the "best" achievable score could change over time.

S-ANDERSON-15: CRAM allows for both approaches to be taken. The idealized best achievable condition is represented by the "A" ratings for the metrics; all the "A" narratives and ratings compiled together represent the ideal. However, ambient surveys and local expertise are used to identify "least impacted" sites, which sometimes score high enough to fit the ideal of "best achievable."

RHEINHARDT-15: Scoring condition relative to "Best Achievable Condition," as defined, is a scientifically reasonable standard.

S-RHEINHARDT-15: No response warranted

WARDROP-15: This statement cannot be assessed without an accompanying description of how the original pool of sites for CRAM development was selected.

Additional comments: It is unclear how the original pool of candidate sites was selected. In general, pools of candidate sites that are selected as per a probabilistic survey are a better source of potential reference sites than those that are not, due to the removal of any bias about what constitutes a "good" site. For example, the best achievable condition will be interpreted by an individual as only the best site that he/she has personally visited; if the pool of candidate sites is probabilistically-based, one can put an estimate on the confidence of having at least one reference site in the pool

S-WARDROP-15: This comment confuses reference sites with best-achievable condition (BAC). BAC is represented by the "A" ratings for the metrics; all the "A" narratives and ratings compiled together represent the ideal or best-achievable condition. This constitutes an internal "reference condition" used to standardize CRAM scores.

However, ambient surveys and local expertise are used to identify "least impacted sites," which sometimes score high enough to fit the ideal "best achievable condition." These real sites that are least impacted and that may also be consistent with best-achievable condition (i.e., sites that get "A" scores for all metrics), can also be used as "reference sites," places practitioners can go to see "A" conditions.

(16) Overall performance of CRAM is validated by evaluating the relationship of metric scores and attribute scores to Level 3 data of key indicators of selected wetland services (e.g., natural values, water quality) (CRAM, Ch. 2.3.3;)

Validation is defined as "…the process of documenting relationships between CRAM results and independent measures of condition in order to establish [CRAM's] defensibility as a meaningful and repeatable measure of wetland condition." (23).

CRAM validation work has emphasized a "weight of evidence approach." CRAM data are compared to multiple lines of independent observations to make inferences or reach conclusions about environmental systems or stressors. (23).

This involves regressing CRAM metric and attribute scores on Level 3 data that are sensitive to changes in wetland condition. This agrees with the process recommended by Fennessy (18, 19) and is also consistent with Hruby's conclusions (26, 27).

NIEMI-16: I have made extensive comments on the issue of validation in response 10. The basic problem is the lack of an "independent" gradient of stressors to the wetland ecosystems. Independent evaluation as used here primarily relies on data gathered on responses on birds, benthos, and plants by MAPS, EMAP, etc. This is fine for some of the validation, but the analysis as described above is not extensive enough. In addition, I suggest that an independent gradient of the stressors to the wetlands need to be better developed. If best achievable condition is defined as indicated in Issue 15 *("least levels of anthropogenic stress")*, then this represents a reference condition. I agree with this interpretation and I contend it is possible to partially quantify this reference condition and the proportion of stress across a gradient of good to poor condition (Danz et al. 2005, 2007; Host et al. 2005).

S-NIEMI-16: Validation is addressed in Part 2. This task could be undertaken, subject to agency commitment including funding. It would involve an independent (i.e., non-CRAM) assessment of condition and stressors in randomly selected wetlands in each class throughout California, although it may be difficult to isolate the effects of one stressor when wetland condition often is affected by multiple stressors (see comment RHEINHARDT -16 below).

ANDERSON-16: Agreed. This is an appropriate way to conduct these analyses. After reading the background and referenced papers I believe their approach was suitable.

S-ANDERSON-16: No response warranted.

RHEINHARDT-16: Validation with level three data could provide evidence of scientific validity, particularly if a "weight of evidence approach" is taken. However, because wetlands are usually altered by multiple stressors, it is difficult to determine to what extent a given stressor affects a given metric or attribute. Testing would require either:

> (1) Finding array of sites that have only one stressor (the same one) that varies in intensity and then determining if the metric varies predictably relative to the amount of stress present. These situations will be difficult, if not impossible, to find for most metrics. Most wetlands are either intact or have multiple stressors controlling their conditions. With multiple stressors, there is interaction among stressors, which makes it impossible to quantify (regress) a response in condition for a single metric along a gradient of alteration.
>
> (2) Conduct field experiments in a series of unaltered wetlands wherein one stressor is introduced at various intensities. This would have to be done for each stressor/metric and each wetland type and perhaps regional subtype. Then, one could determine how indicators of condition vary along the continuum. However, conducting large-scale, ecosystem studies is extremely impractical and would probably not be permitted.

Therefore, although validation would be useful, it may be impractical to achieve for most metrics. Therefore, it would be best to validate metrics using approach (1) above, where possible. However, for most metrics, the most practical approach would be to assume that best achievable condition and the worst condition define the possible range in condition and that the associated range in metric scores reflects the continuum between least and most altered condition.

S-RHEINHARDT-16: The reviewer's latter suggestion has been a basic developmental approach for CRAM. The reviewer's list of possible validation studies could be extended indefinitely. Practical and economic considerations dictate that only a few of the possible validation strategies can feasibly be implemented. Validation is addressed in Part 2.

WARDROP-16: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: This statement is easily agreed to, by virtue of manuscripts published in peer-reviewed literature (Stein et al., 2009).

S-WARDROP-16: No response warranted

> **(17) CRAM scores are not invalidated by natural stochastic changes inherent in all natural systems.** (**CRAM, Ch. 2.2.2- 2.2.3)**
>
> An individual CRAM score represents only the condition of a site on the day of assessment. CRAM assumes that "…geology and climate govern natural disturbance, whereas land use accounts for anthropogenic stress." (CRAM, Ch. 2.3.3). Natural stochastic changes inherent in all natural systems (30. Hruby, 2006), especially riverine systems, do not invalidate or otherwise adversely affect assessment data and scores produced by CRAM, when practitioners exercise good professional judgment in analyzing scores.

NIEMI-17: This is a difficult and perhaps intractable problem. My response is similar to what I stated in response to Issue 9 above. CRAM scores are not necessarily invalidated by natural stochastic changes, but common sense or good professional judgment is required in a subjective type assessment like CRAM.

SI-NIEMI-17: No response warranted

ANDERSON-17: Again agreed. Most rapid assessment techniques follow this same procedure. One would assume that a CRAM score on one day would be similar to scores on subsequent days under most natural circumstances with a few minor exceptions (i.e., major flooding, fire, etc.). Anthropogenic impacts would be more likely to cause short term changes in CRAM scores, which are expected. The "good professional "judgment" statement reinforces the need for good training standards on use of the protocol.

S-ANDERSON-17: No response warranted

RHEINHARDT-17: CRAM seems to do a sufficient job in incorporating indicators and states that respond to anthropogenic-mediated stressors, while not being invalidated by natural stochastic changes. However, for many ratings practitioners must determine whether a wetland is "moderately" or "significantly" different from what should be expected (relative to the natural state), and so the resulting scores depend on the quality of the scorer's best professional judgment (BPJ), which can vary widely. Because geologic and climatic factors vary widely across California, users must incorporate knowledge of local constraints to determine if a condition of a wetland differs from what is expected.

Therefore, to assess all wetland types across all of California, a user would be required to have the type of "good professional judgment" (sensu Hruby) that probably few practitioners possess. An alternative approach would be to write the CRAM rating narratives in a way that forces practitioners into identifying the correct rating (the rating that an expert would obtain), rather than relying on the user's BPJ to obtain the correct answer.

S-RHEINHARDT-17: The need for objectivity is addressed in Part 2.

WARDROP-17: This statement cannot be assessed as to scientific soundness, because it is depends upon the construction of CRAM itself.

Additional comments: If invalidate is, "to weaken or destroy the cogency of", then the protection of CRAM from the invalidating impacts of natural stochastic changes would require a sampling of the same wetland sites before and after such events, and comparing the resulting CRAM scores. At that point, metrics which exhibited instability to such events could be removed from the assessment. Without demonstration of such, I cannot assure the soundness of the issue statement, as given.

S-WARDROP-17: CRAM requires identifying whether the site being assessed has been subject to a major natural disturbance (fire, flood, landslide, etc) or conversion from one wetland type to another. This is being done to allow cases of disturbance and conversion to be culled out and compared to other cases without such disturbance. As data for these case types accumulate, the sensitivity of CRAM metrics to natural disturbance can be assessed. This is part of the CRAM development process.

(18) CRAM scores do not provide measures of human value or importance. Assessments of wetland condition that quantify the wetlands capacity to perform various functions, such as CRAM, do not rate the value of the assessed wetland. Rather, value is assumed to be found in the diversity of ecological services provided by a wetland (CRAM, Ch. 2.2.4; Technical Bulletin, Sec. 3.C). As such, a low scoring wetland may have high value to man based on a number of other considerations, such scarcity of certain wetland types.

NIEMI-18: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The attributes, metrics, and scoring utilized in CRAM do not provide any such measures.

S-NIEMI-18: No response warranted

ANDERSON-18: CRAM is an ecological integrity/ecological services assessment method. Other considerations may include things such as endangered/threatened species habitat or aesthetics. Additional criteria need to be considered separate from CRAM for human value and importance considerations. This is certainly appropriate for CRAM and similar methods.

S-ANDERSON-18: No response warranted

RHEINHARDT-18: CRAM does not purposely incorporate value or importance into its rating scheme. However, the way the rating criteria are defined, smaller wetlands, less structurally complex wetlands, and headwater wetlands are less likely to obtain best achievable scores, making it seem that both types are less valuable. It seems that wetlands that have not been altered should achieve the highest rating no matter how complex or high in a stream network they occur. Size could be a rating criterion, but only in the context of best achievable wetlands of the same regional wetland type and potentially on the basis of habitat requirements for certain selected aquatic species or species guilds important to the class. However, rationale would have to be provided for choosing specific guilds or species to define optimal habitat size and hence, optimal wetland size.

S-RHEINHARDT-18: Size and landscape position can affect CRAM scores. The reviewer is correct that larger and more complex wetlands in most landscape contexts tend to achieve higher scores, regardless of class. This is by design. A basic assumption in CRAM (with which the reviewers have agreed) is that for any wetland class, bigger and more complex wetlands tend to provide higher levels of more functions.

WARDROP-18: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The attributes, metrics, and scoring utilized in CRAM do not provide any such measures.

S-WARDROP-18: No response warranted

GROUP 3 ISSUES

(related to CRAM Manual Chapters 1.3, and Technical Bulletin Chapters 2,3, and 4)

(19) CRAM scores will be used to adjust metrics as needed to remove any systematic bias against particular kinds of wetlands, or their natural settings. (CRAM Ch. 1.5, 3.2.2.1;Technical Bulletin Ch. 3A)

Validation efforts have established that CRAM is applicable throughout the range of environmental conditions commonly encountered in California (21; 23). However, since CRAM metrics give higher weight to structural complexity, there may be a bias against wetlands that naturally exhibit less complexity, such as vernal pools, mud flats, or riverine wetlands headwater areas. In addition, CRAM assessments in riverine systems that have been subject to recent channel disturbance are problematic. Instances may also exist where a positive bias might result in CRAM scores that are higher than site conditions might dictate under Level 3 assessment. Therefore, future adjustments in CRAM metrics may be necessary to remove systematic bias (Technical Bulletin, Sec. 3.A; p. 23).

NIEMI-19: As stated in Issues 10 and 16, I do not agree that validation has been established. The method has inherent biases based on structure and selected biota. The CRAM does include training, plus a selection of a diversity of personnel to complete evaluations is essential to minimize these biases.

S-NIEMI-19: Validation is addressed in Part 2. In general, statement 19 describes how CRAM seeks to avoid or remove intrinsic bias *in the methodology* that would result in spuriously low or high scores for metrics or attributes for different wetland classes. For example, if a wetland class always demonstrates low complexity for a given metric, then that class will always score poorly with respect to the scores of other classes on the same metric; this represents a bias intrinsic in the methodology. Adjusting the scoring brackets can remove such a systematic bias to more accurately reflect the range of condition demonstrated within the subject wetland class. The L2 Committee should exercise due care to remove such biases.

ANDERSON-19: I think it is important to recognize the limitations of the method so by stating this up front that is good. Future adjustments should continue to be made to the method as new biases are discovered and ways to account and compensate for these biases are created. The addition of updates is not problematic and should continue.

S-ANDERSON-19: No response warranted

RHEINHARDT-19: CRAM scores do seem to be biased in favor of wetlands with naturally more structural complexity. There also seems to be an inherent bias against wetlands associated with headwater streams and smaller wetlands. The CRAM manual identified these biases as potentially being problematic, and stated that future adjustments would be made. It is suggested that “future adjustments” be designed to correct biases against particular wetland types should be done one wetland class at a time. That is, for each wetland class, all metrics and indicators for that class should be consolidated, and then any other indicators that might be useful for describing condition for that wetland class should be added, even if the indicator might not be useful for other wetland classes. The narratives should be rewritten, based on how the metrics change with stress, in an unambiguous manner as possible in order to force users to choose the correct answer. One option would be to base ratings on the state of multiple metrics assessed together (i.e., a multivariate approach); this might be the better approach because there are usually multiple stressors affecting wetland condition. Then, the revised rating narratives could be tested first with experts, and then with trained end-users. The same wetland class should be tested across California and then metric/ratings could be adjusted on a regional basis, if necessary and as appropriate.

S-RHEINHARDT-19: See also S-RHEINHARDT-18 and S-NIEMI-19. Testing for bias is essentially already a part of the existing and ongoing process of CRAM development. However, the extent of modification suggested here (that is, developing modules with different structure for different wetland classes) would invalidate the general utility of the CRAM modules, because the resulting “sub-modules” will not be comparable to one another, much less to CRAM results from other wetland types.

WARDROP-19:
This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The methodologies to achieve this certainly exist, but attendance to them in future CRAM endeavors cannot be assumed.

S-WARDROP-19: No response warranted

(20) The selection of CRAM reference sites for all conditions of wetlands is based on "best professional judgment" of regional teams. Selected reference sites exhibit a wide range of condition, from poor to best achievable condition (Technical Bulletin, Sec. 4.C).

Reference sites for all wetland classes and conditions will continue to be established. Reference sites will exhibit a range of conditions from poor to "best achievable condition." (19, 29) Reference sites will be used for calibration and validation studies, and training.

NIEMI-20: At this point in the development of CRAM, "*best available judgment"* is perhaps the best that can be used, but refinement and improvement by incorporation of quantifiable stress information is possible with modern GIS and remote sensing technology (Host et al. 2005, Wolter et al. 2006, Danz et al. 2007).

S-NIEMI-20: No response warranted.

ANDERSON-20: Agreed. However it must be recognized that the conditions in and around a wetland will change over time, so condition at one time may not be representative at another time. Calibration and validation studies must take these changing conditions into effect.

S-ANDERSON-20: No response warranted

RHEINHARDT-20: The selection of reference sites, from least to most altered, by experts is a reasonable approach, and is the approach taken by HGM and some other assessment approaches. The selection of reference sites is one phase of protocol development where best professional judgment (BPJ) should be applied. The BPJ of experts should also be used to develop the narrative or criteria for measuring deflection from natural, best achievable condition in a way that forces users to choose the intended answer for a given condition without invoking BPJ. If practitioners have to invoke BPJ to rate a wetland, the resulting answers might vary widely, particularly in the intermediate range of scores.

S-RHEINHARDT-20: The need for objectivity is addressed in Part 2.

WARDROP-20: The selection of CRAM reference sites for all conditions of wetlands by "best professional judgment" does not represent sound scientific knowledge, methods, and practices.

Additional comments: The basis of "poor" and "best' condition should be a survey that can express, with a given measure of confidence, the probability that such has been found. In general, pools of candidate sites that are selected as per a probabilistic survey are a better source of potential reference sites than those that are not, due to the removal of any bias about what constitutes a "good" or "bad" site. For example, the best achievable condition will be interpreted by an individual as only the best site that he/she has personally visited; if the pool of candidate sites is probabilistically-based, one can put an estimate on the confidence of having at least one reference site in the pool.

S-WARDROP-20: Reference site selection does not rely on BPJ alone. Reference sites are being selected by region for each wetland class based on a combination of prior probabilistic ambient surveys (e.g., estuarine wetlands), logistics, and security criteria (e.g., public ownership and ease of access), plus consensus among local experts.

The CRAM Manual is under revision to provide greater specificity in how reference sites are selected and maintained. At the time of this report, work in progress by staff scientists of the Southern California Coastal Water Research Project (SCCWRP) is contributing to improvement of a statewide reference site system that would be used by CRAM and other assessment programs.

(21) "Best Achievable Condition" reference sites are used to define the highest standard for the type of wetland being assessed. This reference standard becomes the point against which the range of wetland conditions can be judged from highest to lowest, and thus becomes the basis of individual CRAM metric scores. (CRAM, Ch. 3.8.1) Therefore, any two scores for the same type of wetland can be compared to each other because they are based on the same statewide standard (Technical Bulletin, Sec. 3.J).

NIEMI- 21: Similar response as above to Issue 16.

S-NIEMI-21: Validation is addressed in Part 2.

ANDERSON-21: That is true as it is for other rapid assessment techniques. Many are developed on a statewide, geographic province, or other category. Statewide is generally best since it simplified the entire system. The only issue is it may reduce the number of responsive metrics in some cases. However most methods take this into account based on explanatory power of the models. I assume that is what was done in this circumstance.

S-ANDERSON-21: No response warranted

RHEINHARDT-21: "Best achievable condition" is a reasonable way to define the standard for a given wetland type. However, care has to be taken that the natural variation for a wetland type is not so wide that the method is insensitive to typical alterations. Insensitivity is most likely to occur if the wetlands have not been sufficiently pre-classified prior to calibrating metrics against the standard. For example, natural conditions vary so much in riverine wetlands relative to stream order, topography, surficial geology, climate, and elevation, that incorporating this much variation in standards may unduly restrict the types of metrics that can be used, and make the metrics that can be used too insensitive.

S-RHEINHARDT-21: No response warranted

WARDROP-21: "Best Achievable Condition" may vary by surrounding landcover types for any wetland type; under such conditions, the above statement cannot meet sound scientific standards.

Additional comments: Many studies have indicated that the "Best Achievable Condition" may vary according to surrounding land cover (Wardrop et al., 2007; Hruby, 30). Surrounding land cover may alter the very processes by which structures, and thus functions, are expressed. Thus, some comparison of CRAM scores for wetlands of the same type, but in varying land cover surroundings, would have to be provided to be able to make the last statement in the above issue.

S-WARDROP-21: CRAM recognizes that landscape context affects scores and that landscape context is relevant to reference site designation. A site is not judged to be in a state of Best Achievable Condition (BAC) if it is affected significantly by stressors originating in the landscape surrounding the site that depress the function of the site. Sites that are artificially maintained at a higher level of function, even those not requiring ongoing human intervention, would not be used as examples of BAC.

(22) The same scores for different wetlands of the same type do probably represent the same overall condition and functional capacity (CRAM, Ch. 3.8.1; Technical Bulletin Sec. 2, 3B, 3J).

NIEMI-22: I agree and support this statement.

S-NIEMI-22: No response warranted

ANDERSON-22: Yes that is the assumption and I think in this case it is accurate. The same scores should provide a similar measure of condition and functional capacity.

S-ANDERSON-22: No response warranted

RHEINHARDT-22: This assumption is probably not true because individual attributes and metrics haven't been weighted according to their influence on condition. The composite score (sum of all attribute scores) of two wetlands may not represent the same overall condition unless all the individual metric scores are similar for both wetlands, meaning that the stressors are probably similar and of similar magnitude. If a different set of stressors are impacting two wetlands of the same type, the metric scores would vary, but the two wetlands could still obtain similar scores. However, that would not necessarily mean that the overall conditions of the two wetlands are the same. For example, riverine wetland #1 could have its stream so deeply channelized that the former floodplain never receives overbank flow (making hydrologic score=0), but its other attributes (buffer, physical structure, biotic structure) might be intact, providing a composite score of 64.

In contrast, riverine wetland #2 could have its buffer and hydrology intact, but its physical and biotic structure could be reduced by 36 points, providing a composite score of 64. It could reasonably be argued that wetland #1 has a lower overall condition because hydrologic connectivity is so overwhelmingly important to the ecological functioning of riverine wetlands that eliminating this connectivity results in a state change, whereas a reduction in structure would not be severe enough to cause a change in state. This real difference in overall condition is due to the fact that individual metrics haven't been weighted according to their influence on condition. Therefore, it is unlikely that similar composite scores mean that two wetlands are similar in condition.

S-RHEINHARDT-22 Response: Statement 22 intended to express the concept that the "same scores" means the same metric scores, the same attribute scores, and the same index scores. When the metric scores are the same, the presumption would be that the two wetlands exhibit the same condition, and the sameness of condition is presumed to equate to the same functional capacities. Rheinhardt seems to agree with this clarification when he states that: "The composite score (sum of all attribute scores) of two wetlands may not represent the same overall condition unless all the individual metric scores are similar for both wetlands, meaning that the stressors are probably similar and of similar magnitude." [Staff notes that the index score is actually calculated as the average of the four attribute scores, rather than the sum.]

WARDROP-22: The same scores for different wetlands of the same type do probably represent the same overall condition, but the same cannot be said for functional capacity.

Additional comments: CRAM is, according to all the provided materials, a method of condition assessment. However, it is lacking metrics and attributes that explicitly address key variables of the recognized biogeochemical functions of wetlands (Table 2.3), such as cover estimates for carbon pool estimation, soil redox characteristics or indicators, and soil organic carbon. Therefore, functional capacity is not directly assessed, and is done in an extremely limited manner.

S-WARDROP-22: CRAM is, as Wardrop states, a condition assessment. One of the underlying assumptions of CRAM is that scores are directly proportional to functions or ability to function, even though functions or ability to function are not assessed directly (some debate continues over the meaning or propriety of the term "functional capacity"). The L2 Committee continues to evaluate this assumption; currently there are no data available that refute it. The semantic element of the comment will be addressed in the CRAM manual revision that is currently in progress.

(23) Predictions of future conditions of wetlands may be possible through statistical analysis of CRAM reference site data and other CRAM data. (CRAM, Ch. 3.8.1).

A U.S. Army Corps of Engineers (USACE, or Corps) review of CRAM notes that:

"Some of the most important potential applications of assessment approaches involve projecting future conditions to calculate specific gains or losses for with- and without-project scenarios, mitigation site development, and management effects. The data needed to develop such trajectories are best assembled as part of the reference data collection process (28. Klimas 2006), and are not currently a focus of CRAM development. However, given the stated intention of the CRAM developers to actively maintain, build, and use the database to improve the approach, it seems appropriate that one important target would be to develop recovery trajectories suitable for generating future scenarios under conditions of interest to planning and regulatory offices of the Corps, EPA, and State agencies." (31. Klimas, 2006)

Klimas also states that:

"A full evaluation of competing impact and mitigation scenarios requires projection of future conditions… No tools [are] provided for adapting CRAM for use in such situations… To be fair, most other existing assessment systems are equally unsuited to the task… To a certain extent, this weakness in CRAM may be addressed over time as the database grows and new information is applied to the refinement of CRAM…"

In light of these comments, it should be noted that: "As with any assessment method, the ability of CRAM to detect change depends on the size of the change relative to the precision of CRAM." (Technical Bulletin, Sec. 5.C)

NIEMI-23: If as stated in Issue 17 that "*land use accounts for anthropogenic stress*" then future conditions must rely on land use change as a basis for predicting future condition of a wetland. I agree that this line of reasoning is reasonable but I have not seen a good linkage between land use and wetland condition in this CRAM process nor have the appropriate studies been completed to substantiate this linkage over time. The CRAM process has the potential to do this over the long-term as long as the appropriate land use/land cover data layers are included (e.g., see Wolter et al. 2006). Additional information on point sources, human population density, invasive species, and other relevant information should be incorporated. I agree with most of the points made by Klimas.

S-NIEMI-23: The statement under discussion refers to potential capabilities of CRAM, not capabilities that are already developed. Klimas asserts that "performance curves" capable of relating CRAM scores to age of project are possible. Having such curves could allow reviewing agencies to infer future condition from present CRAM scores. Development of such curves for some wetland classes has been identified as a future task to be completed by the L2 Committee. Klimas also refers to relationships between land use context and CRAM scores. Post-stratifying CRAM scores based on land use context to test for land use effects on CRAM scores has been considered. If agencies express such a need, this could be done as a part of additional CRAM development.

ANDERSON-23 At this time it may not be possible to use CRAM to track and gage future conditions and trajectories. The time may come when this is possible if additional research, monitoring, and analysis is completed.

S-ANDERSON-23: See S-NIEMI-23.

RHEINHARDT-23: Klimas' suggestion that an important objective of CRAM "would be to develop recovery trajectories suitable for generating future scenarios" is a good one. At present, CRAM could be used for alternatives analysis to determine the degree to which condition will decline following project impact or increase following restoration. However, it might be challenging to express such change in condition in terms of change in ecological services.

HGM can express a gain or loss in ecosystem services because functional capacity could be converted to ecological services. Although CRAM may be able to predict the effects of future impacts, its ability to predict the future condition of restorations sites, or whether they are on a correct trajectory, is not currently possible unless the reference data set includes data from restoration sites (successful and unsuccessful ones) and space-for-time data from best achievable sites. Space-for-time data are derived from sites that represent intermediate stages of development toward best achievable and are probably only applicable for forested wetland types that typically undergo regeneration after forest clearcutting. In any case, trajectory analyses (sensu Minchin et al. 2005), and comparisons of restoration sites with restoration targets through time (target analysis: Rheinhardt et al. 2007), would only be useful at the level of a metric or attribute, since a composite CRAM score does not provide information about the components of the score.

Minchin, P.R., Folk, M. & Gordon, D. 200). Trajectory Analysis: a New Tool for the Assessment of Success in Community Restoration. Ecological Society of America 90th annual meeting, Montreal, Quebec, August 7-12, 2005.

Rheinhardt, R., T. Wentworth, and M. Brinson. 2008. Developing and testing multivariate approaches for evaluating success of restored vegetation communities. Report to the Ecosystem Enhancement Program, North Carolina Department of Environment and Natural Resources, Raleigh, NC, USA. http://www.nceep.net/pages/resources.htm

S-RHEINHARDT-23: See S-NIEMI-23. The methodology summarized by Rheinhardt ("space-for-time" sequences involving both impact sites and restoration sites) is the methodology that the L2 Committee has contemplated for use in developing performance curves for California.

WARDROP- 23: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The methodologies to achieve this certainly exist, but attendance to them in future CRAM endeavors cannot be assumed. To develop future trajectories, some analysis of "Best Achievable Condition" according to surrounding land cover would be necessitated (see discussion under Issue 21).

S-WARDROP-23: See S-NIEMI-23 and S-RHEINHARDT-23.

(24) The CQAPP is consistent with sound scientific data management practices. Methods specified in the CQAPP are sufficient to assure consistency in the statewide collection of data over time. (CRAM Ch. 3.8.2)

The QA goals and procedures in the CQAPP are specified for each step in CRAM developmental process (25, Technical Bulletin Sec. 4). Training and certification programs for practitioners of CRAM are under development as part of the CRAM Quality Assurance Project Plan (23, 25, Technical Bulletin Sec. 3.H).

NIEMI-24: The version of QA/QC in my packet is dated 3 May 2005 (Collins et al. – Insert 25 of Binder 3 of 3) and it is labeled Draft. It is a draft. There is some detail on a variety of issues, but it is incomplete and needs substantially more information on how QA/QC will be evaluated and how the resulting data will be statistically analyzed for uncertainty in the various methods.

S-NIEMI-24: Revisions to the CRAM QA/QC manual are in progress.

ANDERSON-24: Agreed. I think the CRAM QA/QC is a model document that others should follow. I have recently worked on a similar document and appreciate all the thought and effort that has gone into it.

S-ANDERSON-24: No response warranted.

RHEINHARDT- 24: Most of the CRAM QA/QC manual provides practical steps to help insure quality control of outputs, for example those that insure that the CRAM forms are thoroughly completed and the scores documented, requirements for training and supervised practice, seasonal constraints on assessment widows, and minimum and maximum areal requirements for assessment area size. The necessary procedures are thoroughly outlined and are scientifically sound. The plan to audit 10-15% of assessments by expert staff is a good idea, but if CRAM use becomes routine across California, then it might not be possible to keep up with the demand or the funding required to do so might not be available. A better approach would be to make sure that the assessment protocol is more objective and less prone to error in the first place, so that regular auditing in not necessary. This means reducing, as much as possible, the need for objectivity in interpreting metric or attribute scores.

S-RHEINHARDT-24: No response warranted.

WARDROP-24: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The QA/QC procedures are sound and well-documented.

S-WARDROP-24: No response warranted.

GROUP 4 ISSUES

(see CRAM Manual Chapters 2,3,4, Appendix I;
Technical Bulletin Chapters 2,3, 4, and 5)

Classification, Sampling and Attribute Scoring Procedures with CRAM

The general procedure for using CRAM consists of eight steps. Significant elements of this procedure are discussed below. Reviewers are, of course, invited to comment on any or all of the steps or parts of steps that comprise CRAM's procedure.

Step 1 - Assemble background information about the management of the wetland.

Step 2 - Classify the wetland using this manual (see Section 3.2 and Figure 3.2).

Step 3 - Verify the appropriate season and other timing aspects of field assessment.

Step 4 - Estimate the boundary of the AA (subject to field verification).

Step 5 - - Conduct office assessment of stressors and on-site conditions of the AA.

Step 6 - Conduct the field assessment of stressors and on-site conditions of the AA.

Step 7 - Complete CRAM assessment scores and Quality Assurance Procedures.

Step 8 - Upload CRAM results into regional and statewide information system.

CRAM's wetland and riparian classes were developed based on the National Research Council's (NRC) recommendations, and to be consistent with the National Wetlands Inventory (NWI) and the State Wetland Inventory (still under development). (CRAM, Ch. 3.2, Figure 3.2).

(25) In regards to Step 2 above: Further refinement and subdivision of CRAM's wetland classes could take place over time based on wider data collection and analysis. Following CRAM's developmental procedures, meaningful conditional assessment scores for these new wetland classes can continue to be obtained (CRAM, Ch. 2.3.3; Technical Bulletin Sec. 3.A).

NIEMI-25: The continued refinement of wetland classes over time is a healthy approach; especially as new information is obtained. I certainly support this perspective.

S-NIEMI-25: No response warranted.

ANDERSON-25: Agreed. Like the other methods the classification and resulting assessment scores will (and should) change over time.

S-ANDERSON-25: No response warranted.

RHEINHARDT-25: The sensitivity and specificity of CRAM metrics could be improved by customizing the assessment protocol by wetland type and subtype, so further refinement of wetland classes would be helpful. At present, separate protocols have been customized for three wetland types (estuarine, riverine, and depression), but this was done by simply cutting and pasting relevant material from the general protocol into one place, e.g., grouping all riverine-related material under a riverine manual. However, in doing so, terminology that overlapped with other wetland types was not removed, leaving extraneous material in the manuals and making the protocols somewhat confusing.

(There were also some mistakes made in transferring protocol, some of which are identified, by wetland class, in the "Specific Comments" section).

For each wetland type, there are additional metrics (indicators) that could be used to make the evaluation of condition more objective. These metrics might be robust for a particular wetland class, but would be neither sensitive nor appropriate for other wetland types. For example, degree of channel incision would be a useful metric for evaluating riverine wetland hydrologic condition. Therefore, when refining and subdividing wetland classes for customizing protocols, an attempt should be made to identify additional metrics for each wetland class that might be robust and be capable of being objectively measured.

S-RHEINHARDT-25: The need for objectivity is addressed in Part 2.

As noted above in S-WARDROP-3 and S-RHEINHARDT-6b, customizing existing modules and adding modules will occur as data-driven exercises.

New wetland classes are only identified when application of existing modules produces anomalous or inconsistent results. A process has been established by the L2 committee to guide subsequent action when this occurs.

1. A researcher finds that an existing module produces anomalous or inconsistent results that cannot be rectified within the existing module without major revisions.

2. The researcher presents the data and conclusions to L2.

3. If L2 agrees that a new module is the only way to rectify the inconsistency, that new module proposal is added to the L2/CWMW work plan.

4. Development of the module will proceed depending on need, relative importance, and availability of funding and staff.

An example of this process in action can be seen with recent developmental work for a CRAM Desert Riverine module. The existing CRAM Riverine module gave consistently lower scores to episodic desert streams. It was noted, for example, that the riverine module's metrics rely on typical riparian vegetation and indicators of typical fluvial process, and that these were just not applicable in the desert. This result was reported to L2. L2 agreed, and work commenced on a desert riverine module. This work was deemed a priority because of the extensive development projects under construction in California's deserts.

L2 does not foresee a need for a great proliferation of modules. Many wetlands function in similar ways, and a more realistic expectation is that the use of existing modules may be extended to additional categories of similarly functioning wetland. As an example of this, the L2 Committee has developed a draft module for "Wet Meadows" in the Sierra Nevada, based on a demonstrated need for such a module. The draft module is derived from general ecological patterns included in the existing preliminary Depressional module and the "Slope Wetland/Seeps and Springs" working draft module. The draft Wet Meadow module has subsequently been found to be useful in assessing condition in several types of "flowthrough" wetlands in the northern Coast Range. The L2 Committee expects that such circumstances will recur frequently.

The default option is not to develop new modules or recognize new classes until those actions are supported by a clearly demonstrated need. Too much subdivision of wetland modules reduces the applicability of the CRAM modules in comparing wetlands within each class and considering condition among classes.

The L2 Committee is attentive to the offsetting concerns of identifying truly separate wetland classes while avoiding unnecessary subdivision of classes.

WARDROP- 25: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: Further refinement is always possible with the addition of new data. However, I am more concerned about the lack of consideration of "lumping" classes than I the availability of data to sub-divide. Please see the response to Issue #3.

S-WARDROP-25: See S-RHEINHARDT-25.

(26) In regards to Step 4 above: Assessment Areas (AAs), as defined through criteria listed in Chapter 3 of the CRAM Manual , are a valid sample type and size for determining wetland condition (CRAM, Ch. 3.5, Technical Bulletin Sec. 5).

The Assessment Area (AA) is the portion of a wetland that is assessed using CRAM. An AA might include all of a small wetland in its entirety, but in most cases the wetland will be larger than the AA. Steps for delineating AAs are summarized in Table 3.5 in CRAM, Ch. 3.5: Special criteria for delineation of AAs for Riverine and Vernal Pool wetlands are also described.

NIEMI-26: Wetland delineation is not my area of expertise, but the procedures look reasonable. However, note my comment under Issue 7 – I think the buffer and landscape context are too small. Among the best concepts we have learned over the past twenty years in ecology is the importance of landscapes and watersheds on receiving bodies such as wetlands, lakes, and streams. GIS has given us the capabilities of incorporating analyses hardly dreamed about even 30 years ago.

S-NIEMI-26: Generally the comment misinterprets the focus of question 26, which addresses the identification of the Assessment Area for a CRAM assessment.

CRAM uniformly incorporates a landscape-based approach to wetland assessment, an understanding widely included in agency decision-making. Current federal regulations call for a "watershed approach" in mitigation for activities affecting wetlands. State Water Board Resolution 2008-0026 "...recognizes that watershed-focused planning is the most effective strategy for maintaining and enhancing [wetland] functions..." and directs staff to develop wetland protection policy with a "watershed focus." These regulatory and policy drivers could result in continued emphasis on this topic.

ANDERSON-26: Determining assessment areas are one of the most difficult parts of the process, but it is obviously central to the entire assessment procedure. The AA procedure for CRAM is well documented and although takes some professional judgment; most of the subjectivity has been eliminated.

S-ANDERSON-26: No response warranted

RHEINHARDT-26: The minimum and maximum size criteria (Table 3.7) appear to be reasonable and the assumptions for specifying size ranges by wetland type are justified. Rules for delineating wetland assessment area boundaries (Table 3.5 and 3.6) are also reasonable. However, the adequacy of the suggested size ranges for the various wetland types should be tested using the grid sampling approach (Appendix I, Figure 1) to see if randomly chosen cells in a grid incorporate the natural structural complexity for all wetland classes.

S-RHEINHARDT-26: No response warranted.

WARDROP-26 This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: This is perhaps the best-written and conceived portion of the manual, and its comprehensiveness and use of examples is well-appreciated. The following factor provides a solid foundation and check on the methodology: "The size-frequency distribution of wetlands for each wetland type (a Level 1 analysis) was also considered when the recommendations for AA sizes were being developed."

S-WARDROP-26: No response warranted

(27) In regards to Step 4 above: The ambient assessment sampling and the project assessment sampling methods for CRAM may both be expected to provide valid, repeatable results.
(CRAM, Ch. 3.5.3, Appendix I).

Separate sampling methods are prescribed for project assessments vs. ambient assessments. Ambient assessment sampling is described in CRAM, Ch. 3.5. Project assessment sampling is discussed in CRAM, Ch. 3.5 and App. 1. The same rules for delineating AAs pertain to both of these purposes for using CRAM. However, they may require different numbers of AAs.

NIEMI- 27: This seems reasonable – though same concerns as in Issue 7 and 26. Also, the sample and project assessments need to be incorporated into QA/QC procedures.

S-NIEMI-27: No response warranted

ANDERSON- 27: Appendix I does a great job of describing the assessment protocols. Again it comes down to assessing and determining the AA. As long as training is required and protocols are followed than results should be repeatable.

S-ANDERSON-27: No response warranted

RHEINHARDT- 27: Prescribing different the methods for choosing AAs to determine assessment vs. ambient conditions is a reasonable approach. However, the guidelines provided for determining the required number of AAs (Table 3.8) is a bit confusing. For example, if the average differences in scores are more than 15%, then there could be cases where taking more samples from the wetland would not be possible without overlapping sample locations.

However, the guidelines for establishing a grid system (Appendix 1, Figure 1) makes it seem that overlap is not allowed. There is no clear guidance on what to do if all grids are sampled and the last score still varies by more than 15% from the average of the other scores (most likely to occur in small, heterogeneous areas). This lack of guidance could lead a reader to believe that perhaps randomly chosen AAs are allowed to overlap (like in a Monte Carlo sampling procedure) until scores stabilize. More guidance is needed regarding sampling overlap and what to sample after the entire wetland has been sampled.

S-RHEINHARDT-27: Guidance for project assessments is currently being addressed as the CRAM manual and other supporting materials are revised and updated.

WARDROP- 27 This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The critical concepts common to both approaches (Sample Universe and Sample Frame) are well-founded, logical, and clearly presented.

S-WARDROP-27: No response warranted

(28) In regards to Step 7 above: CRAM metric scoring, as described in Chapter 4 of the CRAM manual, is based choosing the most appropriate narrative description of the state or condition of the metric being observed in the field, ranging from low to high ecological service. CRAM yields an overall AA score based on the individual scores of the attributes and their metrics (CRAM, Ch. 3.8.1).

The metrics are organized into 4 main categories (termed “attributes”): (1) Landscape Context and Buffer, (2) Hydrology, (3) Physical Structure and (4) Biotic Structure (CRAM, Executive Summary). These four categories are commonly recognized as being “universal” attributes of wetlands (18,21). CRAM has standardized these metrics and narrative descriptions across all wetland types as much as possible (CRAM, Ch. 3.8.1).

NIEMI- 28: The basis for these measurement categories is well-founded in the literature. An important refinement possibility for the future would be to analyze the covariation among these categories. For instance, it might be possible to simplify the method later if it is found that some of the quantitative measures being gathered are highly correlated and therefore redundant. We gathered and analyzed over 200 independent stress variables in wetlands of the Great Lakes. We found that there was substantial correlation among the stress variables and they could be reduced to a much smaller number yet retain a substantial portion of the variation (see Danz et al. 2005, 2007). This can simplify the process in subsequent years in these wetland assessments. Stein et al. (2009) did this type of analysis and more of it is encouraged.

S–NIEMI-28: The L2 Committee is aware that CRAM requires continuous scientific evaluation; the potential effects of autocorrelation among metrics on CRAM assessment results have been discussed as a scientific concern to be addressed in future studies.

ANDERSON- 28: Variation in metrics and narratives among wetland types is to be expected and does not limit or reduce the value of CRAM.

S-ANDERSON-28: No response warranted

RHEINHARDT-28: The four main categories (attributes) are a reasonable way to partition condition in wetlands. However, some of the rating narratives are generalized so much that they require subjective interpretation by users. For examples where objective and subjective criteria are used in CRAM, see comment #12 above.

For field books, it would be clearer to include only examples of alterations and indicators that would be expected to occur in the wetland type under consideration. For example, to minimize confusion in assessing estuarine wetlands, only information about estuarine wetlands should be included in the field book; all references, except in the classification key, to riverine, depression, and slope wetlands should be excluded.

S-RHEINHARDT-28: This concern for metric confusion can be addressed as the CRAM manual and other supporting materials are revised and updated. Also see S-RHEINHARDT-12.

WARDROP-28: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The manual provides clear and easy-to-use narrative scoring criteria. However, a broader question is begging to be asked, and is not included in the topics to be reviewed. CRAM is an ecological model, similar to hydrogeomorphic functional assessment models. In developing a typical HGM guidebook, the following steps are followed prior to the scoring, calibration, and validation steps:

- Identification of controlling variables in the process of interest
- Construction of the conceptual model
- Selection of indicators
- Selection of measurement techniques

Uncertainty and potential error is present in each of these steps. For example, the wrong controlling variables could be identified, the variables in the functional assessment model might be combined in an inappropriate way, good indicators are not available, or measurement techniques are inherently variable.

Documentation of these steps is not included in CRAM in any explicit way. For example, the mathematical combination of submetrics in the raw attribute score for Buffer and Landscape Context implies a certain understanding of the wetland ecosystem:

$$\left(\text{Buffer Condition} \times \left(\text{\% AA with Buffer} \times \text{Average Buffer}\right)^{1/2}\right)^{1/2} + \text{Landscape Connectivity}$$

The basis of this model is not discussed or documented. Without this foundational information, the appropriateness of the compositional metrics and the resulting attribute scores cannot be assessed.

S–WARDROP-28: Documentation regarding these concerns can be addressed as the CRAM manual and other supporting materials are revised and updated.

ISSUES TO BE REVIEWED

The Readiness of CRAM Modules for Specific Wetland Types

<u>GROUP 5 ISSUES</u>: (related to CRAM Estuarine and Riverine Wetland Field Books)

A. <u>CRAM for Estuarine and Riverine Wetlands</u>

(29) CRAM reliably evaluates the condition of estuarine and riverine wetlands in California.

The CRAM Field Guides for Estuarine and Riverine systems, as presented in the respective Field Guides for those modules (3,4), provide details on application of CRAM to those ecosystems. Recent validation studies have shown that Estuarine and Riverine CRAM scores correlate well with a variety of more intensive Level 3 data sets (23).

NIEMI-29: In consideration of the value of wetland ecosystems, I do not think this material provides conclusive evidence that CRAM can be reliably used to evaluate condition of estuarine and riverine wetlands in California**.** Wetland ecosystems provide enormous environmental benefits to humans (flood control, nutrient processing, etc.) and to biota. The ***precautionary principle*** should be used. Given the questions I have raised and preliminary nature for much of the work, especially validation, further development is necessary. I would also be concerned that if the method is applied now, then further development and validation would not occur.

S-NIEMI-29: Validation is addressed in Part 2.

The issue might not be whether CRAM is ready to be used --- it is in fact being used by estuarine and riverine science teams --- but how the data will be interpreted and used. CRAM has been used to conduct a statewide ambient survey of estuarine condition and is part of an ongoing effort to assess wadeable streams statewide. Staff concedes that the fact that CRAM is being used is not evidence that the results are valid or that the data are being properly interpreted. As with any assessment, quality control of the methodology, data collection and data interpretation is necessary.

Niemi states here that: “I would also be concerned that if the method is applied now, then further development and validation would not occur.” As noted in the introduction to this report, this is not a scientific issue, but it is a valid policy issue. Policy makers should take note that scientists have observed that assessment methods fail when support for development and refinement is lost.

Niemi’s lengthy response to statement 10 includes the following: “It is clear that the authors recognize the limitations of the methodology and it is likely cost that has prohibited them from gathering the appropriate comprehensive data to fully validate CRAM. In summary, I do not think the procedure has received adequate validation.”

In both these statements, Niemi makes no distinction between the various appropriate uses of CRAM that are possible at varying stages of development. Less complex applications of CRAM data are possible at earlier stages of module development. More complex applications should not be put into practice until larger data sets allow for more confidence.

No assessment will ever be possible if delay of implementation is necessary until all of the issues are resolved. L2 and CWMW state that phasing the introduction of CRAM prudently and cautiously addresses these concerns, and State Water Board staff concurs with this position. Staff recommends that any Water Boards policy statement regarding Level 2 surface waters assessment provide for appropriate phasing of those assessments into the regulatory process, to allow time for the recommended CRAM developmental process to unfold. This approach could be integrated with the phased implementation schedule that is to be part of a proposed new wetland policy.

ANDERSON- 29: I was extremely impressed with these stand-alone field books. I think they are well-done with adequate explanation of use and limitations with the exception of a few minor issues. I think these additions are important contributions to the entire CRAM method.

S-ANDERSON-29: No response warranted

RHEINHARDT-29: As discussed in comments #12 and #28 above, the reliability of the estuarine and riverine guidebooks depends on the degree of emphasis on the subjective criteria used for rating attributes. In both field books, there is a mixture of subjective and objective ranking criteria. Both guidebooks would be more reliable if the subjective ranking narratives were revised to require more objective criteria for making ranking decisions.

The correlation of CRAM scores with Level-3 data in the validation studies probably depends more on the experience of the CRAM testers than on the validity of the method. Experts using CRAM can probably get the intended (correct) ratings using their BPJ when evaluating the subjective criteria. However, users with typical delineation experience are much less likely to get the intended ratings, at least for wetlands of intermediate condition.

S-RHEINHARDT-29: Objectivity is addressed in Part 2.

WARDROP-29: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: This statement is easily agreed to, by virtue of manuscripts published in peer-reviewed literature (Stein et al., 2009). The article includes an excellent description of the actions taken as a result of the validation process (Table 7), as well as an extensive description of the weight-of-evidence approach and its appropriateness in CRAM development and refinement.

S-WARDROP-29: No response warranted.

(30) In cases where the physical findings indicating the top of the stream bank where the stream flow would enter the active flood plain (bankfull indicators) are obscured or unreliable due to channel disturbance, CRAM adequately addresses the uncertainty of using bankfull width by applying a sensitivity test when determining the metric score in the hydrologic connectivity section (29).

NIEMI-30: No comment. This issue is not within my level of expertise. I suggest that it needs to be evaluated by a hydrologist.

S-NIEMI-30: No response warranted

ANDERSON- 30: I do not have an issue with the method although it may be somewhat arbitrary but I think it is reasonable and easily repeatable if a few details are clarified. The example indicates that for the Site Potential Vegetation Height (SPVH) if the vegetation (alders) is 5 m tall than the AA would extend 10 m from the backshore. I do not see a definition for SPVH. In particular the word potential throws me off. If the alder in this case was only 2 m tall would the AA still be 10 m? How is the backshore defined?

S-ANDERSON-30: This issue has already been identified and addressed by CRAM's Principal Investigators. These revisions will be included when the CRAM manual and other supporting materials are revised.

RHEINHARDT-30: The adequacy of the uncertainty test depends on how well it works in practice. The rationale presented for the sensitivity test is reasonable.

S-RHEINHARDT-30: No response warranted.

WARDROP- 30: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The approach appears eminently reasonable, although some sort of reference for the approach would increase its defensibility.

S-WARDROP-30: This can be addressed when the CRAM manual and other supporting materials are revised.

B. Depressional Wetlands

(31) CRAM can be expected to reliably evaluate the condition of depressional wetlands, assuming continued development according to CRAM's developmental process.

The CRAM Field Guide for Depressional Wetlands is available (5). This module has not undergone the level of calibration and validation work that has been given to the Estuarine and Riverine modules. It is at an earlier stage of development, and is expected to be validated and calibrated as the data is collected on additional assessment sites. Despite this limitation, the Depressional module is ready for use as prescribed in the Technical Bulletin and by Sutula (21), and is expected to provide reliable data if the CRAM developmental process is continued.

NIEMI-31: I cannot support its use on the basis of my answer in Issue 29. Considerable developmental, evaluation, validation, and QA/QC issues remain.

S-NIEMI-31: We agree that the depressional module is not yet ready for broad application. Funds are being secured to validate the module.

ANDERSON-31: The protocol to date seems to be based on solid information. As additional data are obtained than I would expect the protocol to change. There is nothing wrong with this but users must realize the limitations of the current protocol.

S-ANDERSON-31: No response warranted

RHEINHARDT-31: As is true for the wetland manuals for estuarine and riverine wetlands, there are many examples provided in the documentation that are not relevant to depressions, i.e., the examples are only relevant to other wetland types. For example, under Hydrology (page 14, paragraph 2), there is a list of direct sources of water that are irrelevant to depressions (tides, high riverine flow, etc.) and indirect sources that are irrelevant (tide gates, check dams, weirs). This extraneous information is distractive to the end user. When extraneous information is removed, only this narrative remains for describing depressional hydroperiod (p. 16).

Hydroperiod
Definition: Hydroperiod is the characteristic frequency and duration of inundation or saturation of a wetland during a typical year. Depressional wetlands typically have daily variations in water height that are governed by diurnal increases in evapotranspiration and seasonal cycles that are governed by rainfall and runoff.

The few indicators provided in Table 4.10 are to be used to decide, using BPJ, if filling and/or draining are of a different magnitude or rate than would be expected in comparison to natural wetlands. For example, it would be difficult to determine whether inundation patterns are of "greater magnitude or duration than would be expected" or "substantially lower magnitude or duration than would be expected", since none of the indicators is compared to what is "expected" or provides a way to differentiate condition based on alterations to inundation patterns (filling and drawdown). The indicators are all clues that could be used to come help conclude whether filling or drawdown patterns might be unnatural, and the stressor checklist provides useful information about what alteration(s) to look for, but if an alteration is identified, the user has to decide if the alteration is likely to have a" negative effect" or a "significant negative effect" and then decide if the filling or drawdown patterns are of "greater magnitude" or "lower magnitude" than expected. The rating a CRAM user obtains depends on his/her experience with the ecosystem (its hydrology and biology) and personal biases about what constitutes good or poor condition, or his/her agenda (e.g., a consultant who wants to minimize the score as much as possible). This requirement to invoke BPJ makes the ratings imprecise, particularly for the middle range of ratings (Ratings B and C).

As mentioned in earlier comments, ratings of attributes require this type of subjectivity for other wetland types as well. Experts might tend to obtain similar ratings based on BPJ, but many users with typical delineation experience probably would not.

S-RHEINHARDT-31: The need for objectivity is addressed in Part 2. Staff notes that for CRAM scores to be accepted into the CRAM database, practitioners must have passed a mandatory CWMW-approved training course led by a CWMW-approved trainer (Scores submitted by untrained practitioners would not be accepted by L2 or CWMW). Practitioners who implement and use the Depressional module will be required to undergo the same intensity of training that practitioners must obtain with respect to the Riverine and Estuarine modules.

The manner in which this is implemented into regulatory programs is undecided, but the need for guidance on this point is recognized.

WARDROP- 31: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The above statement is true if the developmental process described in the CRAM Manual is supported with the type of validation work presented for the estuarine and riverine models in Stein et al., 2009.

S-WARDROP-31: No response warranted

C. Slope, Playa, Lacustrine, and Vernal Pool Wetlands

(32) The development of CRAM for these four wetland types has not progressed as far as those listed above. If the same developmental process is applied to these four types, then a similar outcome is anticipated resulting in CRAM for these types. (CRAM, Ch. 1, 2)

Field manuals for slope, playa, lacustrine and vernal pool wetlands are under development, using the methodology described in CRAM, Ch. 1 through 4, and as discussed in Topic 1 above. Additional data collection is underway. It is expected that these modules will be ready for wider use over the next few years.

NIEMI- 32: The CRAM methodology is applicable to all of these wetland types, but all of the precautions and concerns that I have identified apply here as well

S-NIEMI-32: No response warranted

ANDERSON- 32 I wish them well with completing this task and recommend they stay on the same course they have been on.

S-ANDERSON-32: No response warranted

RHEINHARDT- 32: All comments made previously in this review apply to the development of field books for other wetland types, i.e., regarding subjective vs. objective ratings and the inclusion of distracting extraneous information in specialized field books.

S-RHEINHARDT-32: No response warranted

WARDROP-32: This statement is judged to be based upon sound scientific knowledge, methods, and practices.

Additional comments: The above statement is true if the developmental process described in the CRAM Manual is supported with the type of validation work presented for the estuarine and riverine models in Stein et al., 2009.

S-WARDROP-32: Validation is addressed in Part 2.

The Big Picture

Reviewers are not limited to addressing only the specific issues presented above, and are asked to contemplate the following questions.

> **(a) In reading the CRAM Manuals and Field Guides, the Technical Bulletin CRAM Implementation Technical Bulletin, the CRAM Quality Assurance/Quality Control plan, the supporting information at www.cramwetlands.org, and proposed implementation language, are there any additional scientific issues that are part of the scientific basis of CRAM not described above? If so, please comment with respect to the statute language given above in the first three paragraphs of Attachment 2.**
>
> **(b) Taken as a whole, is the scientific portion of CRAM based upon sound scientific knowledge, methods, and practices?**

Reviewers should also note that some proposed actions may rely significantly on professional judgment where available scientific data are not as extensive as desired to support the statute requirement for absolute scientific rigor. In these situations, the proposed course of action is favored over no action.

The preceding guidance will ensure that reviewers have an opportunity to comment on all aspects of the scientific basis CRAM. At the same time, reviewers also should recognize that the Board has an obligation to consider and respond to all feedback on the scientific portions of CRAM. Because of this obligation, reviewers are encouraged to focus feedback on the scientific issues highlighted.

(Reviewers' Responses to the "Big Picture" questions are designated "BP" in the following section)

NIEMI- BP: The scientific basis of rapid assessment methods is reasonably well-developed. I have major concerns with the extent to which CRAM has been validated and I would strongly encourage the development of stronger, independent analysis of a human stress gradient. There is a draft document that was produced for the USEPA Office of Water in 2005 that developed this concept of a Biological Condition Gradient further, but I am not sure of its status. However, many of the peer-reviewed articles that I have cited have shown that it is possible to quantify this gradient with available stressor information (Danz et al. 2005, 2007) and interpret these data in the context of wetland biota (Brazner et al. 2007a, b). It is even possible to consider samples of the wetland biota to estimate the ecological condition of wetlands (e.g., see Howe et al. 2007). For instance, it is entirely reasonable to assume that the organisms living in these wetlands can indicate their overall condition, especially since they are living in the prevailing wetland conditions.

If we consider the monetary value of wetlands which is something we have ignored for too long, the consequences of making mistakes in wetland evaluation is very high. Wetlands provide many important functions (flood control, nutrient and other contaminant processing and removal, habitat for game and non-game species; nurseries for fisheries, recreational opportunities), so the protection of their integrity and appropriate assessments are essential. This need was clearly identified by Collins et al. (2005, 2008) within this packet of material. Coordinated efforts with Oregon and Washington for the entire US Pacific Coast should also be considered, especially since financial budgets are presumed to very tight.

References Cited [by Niemi for his BP statements]

Brazner, J.C., N.P. Danz, G.J. Niemi, R.R. Regal, Trebitz, A.S., R.W. Howe, J.M. Hanowski, L.B. Johnson, J.J.H. Ciborowski, C.A. Johnston, E.D. Reavie, V.J. Brady, and G.V. Sgro. 2007a. Evaluation of geographic, geomorphic, and human influences on Great Lakes wetland indicators: a multi-assemblage approach. Ecological Indicators 7:610-635.

Brazner, J.C., Danz, N.P., Trebitz, A.S., Niemi, G.J., Regal, R.R.,Hollenhorst, T., Host, G.E., Reavie, E.D., Brown, T.N., Hanowski, J.M., Johnston, C.A., Johnson, L.B., Howe, R.B., Ciborowski, J.H. 2007b. Responsiveness of Great Lakes wetland indicators to human disturbances at multiple spatial scales: a multi-assemblage assessment. Journal of Great Lakes Research 33(3):42-66.

Collins, J., M. Sutula, E. Stein, and R. Clark. 2005. DRAFT Quality Assurance Project Plan for the Development of a Wetland Rapid Assessment Method in California: Phase II (USEPA Grant Funding Numbers: CD 97954801-6, CD-96911101 and, CD-96911301) 3 May 2005

Collins, J.N., E.D. Stein, M. Sutula, R. Clark, A.E. Fetscher, L. Grenier, C. Grosso, and A. Wiskind. 2008. California Rapid Assessment Method (CRAM) for Wetlands. Version 5.0.2. 151 pp.

Danz, N., R. Regal, G.J. Niemi, V.J. Brady, T. Hollenhorst, L.B. Johnson, G.E. Host, J. M. Hanowski, C.A. Johnston, T. Brown, J. Kingston, and J.R. Kelly. 2005. Environmentally stratified sampling design for the development of Great Lakes environmental indicators. Environmental Monitoring and Assessment 102:41-65.

Danz. N.P. G.J. Niemi, R.R. Regal, T. Hollenhorst, L.B. Johnson, J.M. Hanowski, R. Axler, J.J.H. Ciborowski, T. Hrabik, V.J. Brady, J.R. Kelly, J.C. Brazner, R.W. Howe, C.A. Johnston, G.E. Host. 2007. Integrated gradients of anthropogenic stress in the U.S. Great Lakes basin. Environmental Management 39:631-647..

Fennessy, M., Amy D. Jacobs, and Mary E. Kentula. 2007 *An Evaluation of Rapid Methods for Assessing the Ecological Condition of Wetlands.* Wetlands, 27:543–560.

Hollenhorst, T.P., Brown, T.N., Johnson, L.B., Ciborowski, J.J.H., and Host, G.E. 2007. Methods for generating multi-scale watershed delineations for indicator development in Great Lake coastal ecosystems. *J. Great Lakes Res*. 33:13-26.

Host, G. E., Schuldt, J., Ciborowski, J.J., Johnson, L.B., Hollenhorst, T., and Richards, C. 2005. Use of GIS and remotely-sensed data for a priori identification of reference areas for Great Lakes coastal ecosystems*. Int. J. Remote Sens*. 26:5325-5342.

Howe, R.W., R.R. Regal, G.J. Niemi. N.P. Danz, J.M. Hanowski. 2007. A probability-based indicator of ecological condition. Ecological Indicators 7:793-806.

Klimas, C.V. 2008. Comments on the California Rapid Assessment Method for Wetlands Response to request from USEPA for review of CRAM by the US Army Corps of Engineers, Wetland and Coastal Ecology Branch of the Environmental Laboratory, US Army Engineer Research and Development Center, Vicksburg, Mississippi, dated Jan. 11, 2008).

Niemi, G.J. and M. McDonald. 2004. Application of ecological indicators. Annual Review of Ecology and Systematics 35: 89-111.

Niemi, G.J., Kelly, J., and N. Danz. 2007. Environmental indicators for the North American Great Lakes Coastal Region: Introduction and Prospectus. Journal of Great Lakes Research 33(3):1-12.

Peterson, A. and G.J. Niemi. 2007. Evaluation of the Ohio Rapid Assessment method for wetlands in the Western Great Lakes: an analysis using bird communities. Journal of Great Lakes Research 33(3):280-291.

Stein, E.D., A.E. Fetscher, R.P. Clark, A. Wiskind, J.L. Grenier, M. Sutula, J.N. Collins, and C. Grosso. 2009. *Validation of a Wetland Rapid Assessment Method: Use of EPA's Level 1-2-3 Framework for Method Testing and Refinement*. Wetlands 29(2):648-665.

Sutula, M., E. Stein, J. Collins, A. Fetscher and R. Clark. 2006. *A Practical Guide for the Development of a Wetland Assessment Method: The California Experience*. J. Amer. Water Resources Assn. Feb. 2006, p. 157. (Paper no. 04215).

Wolter, P.T., Johnston, C.A., and Niemi, G.J. 2005. Mapping submerged aquatic vegetation in the U.S. Great Lakes using Quickbird satellite data. *Int. J. Remote Sens.* 26:5255-5274.

_____ , Johnston, C.A., and Niemi, G.J. 2006. Land use land cover change in the U.S. Great Lakes basin 1992 to 2001. *J. Great Lakes Res*. 32:607-628.

S-NIEMI-BP: Validation is addressed in Part 2.

Niemi states *"I would strongly encourage the development of stronger, independent analysis of a human stress gradient."*

Stress gradients are addressed in more detail in Part 2. An independent analysis of a human stress gradient can be done, and would probably provide some useful information. This would certainly be in the form of Level 3 studies/research that would require effort beyond that of a rapid assessment. The CRAM PI Team can undertake studies to document the applicability of these suggestions for California wetlands and their applicability to CRAM, if support for such directions in CRAM's development is received from interested agencies or entities.

ANDERSON-BP: Overall I see very few issues with the scientific merit of CRAM. Where possible they have used the best available scientific data. Best professional judgment is required in places but is kept to a minimum. I found that the CRAM methodology is provided in great enough detail to be repeatable and consistent. I believe it will improve California's wetland management capabilities.

S-ANDERSON-BP: No response warranted

RHEINHARDT-BP-a: *CRAM condition assessments are intended to provide a scientifically defensible Level-2 approach for providing "consistent, scientifically defensible, affordable information about wetland conditions throughout California." The approach taken is to use narrative descriptions of various metrics to rate alternative states of condition, calibrated by major hydrogeomorphic subclasses, and use a list of potential stressors to help identify the causes for conditions that are determined to be less than the best achievable. Although originally developed to assess all types of wetlands in California, CRAM has since been amended by providing separate field books for three wetland subclasses: estuarine, riverine, and perennial depressional. The CRAM approach has been designed to be flexible in that standards and calibrations can be revised to reflect new data and quality can be evaluated by auditing a proportion of assessments with expert teams.*

While the CRAM approach to use narrative descriptions to differentiate various condition states provides a rapid way to evaluate wetlands, some of the methods used for scoring metrics and some of the narrative used to identify alternative states rely on subjective interpretations that could lead to unreliable and inconsistent results. Table 1 shows that half the sub-metrics rely on subjective interpretations or best professional judgment (BPJ) of the user. This review provided examples of specific attributes that require objective interpretations and how individual biases could affect scoring outcomes. Although experts might interpret the subjective sub-metrics similarly, ratings among typical end users will likely vary widely, particularly for wetlands exhibiting conditions in the middle range (scores of B & C).

Table 1. List of CRAM sub-metrics indicating whether their rating criteria require a subjective or objective interpretation.

Sub-metric	**Subjective**	**Objective**
Landscape connectivity		✓
% AA with buffer		✓
Buffer width		✓
Buffer condition	✓	
Water source	✓	
Hydroperiod/Channel stability	✓	
Hydrologic connectivity		✓
Patch type		✓
Topgraphic complexity	✓	
Number of plant layers		✓
# codominant plant spp	✓	
% codominant invasive spp		✓
Horizontal interspersion	✓	
Vertical biotic structure	✓	

Given that wetlands naturally vary widely across such a geologically and climatically diverse region as California, even within a given hydrogeomorphic class, many end users will lack the expert knowledge required to obtain the same score as expert users for the seven subjectively-rated attributes. The fact that experts were used to test the repeatability (precision) of CRAM results means only that experts tend to agree. It would be preferable to test whether (a) experts obtain the ratings intended, and (b) whether the ratings of typical users agree with ratings obtained by experts.

The State has initiated a program to train CRAM users, track their assessments, and conduct random audits of CRAM assessments. Experience to date shows that trained users meet precision and repeatability objectives based on comparisons to expert results.

Since half the sub-metrics rely on subjective interpretation, particularly in the middle range, it will not be possible to differentiate condition reliably among wetlands with composite (overall) scores in the middle range. This is precisely the range where the most resolution is needed; usually, no assessment method is required to recognize unaltered and severely altered wetlands. The true test of an assessment method is whether it can reliably and accurately differentiate wetlands of intermediate condition. It is doubtful that CRAM can do this. However, if the subjective sub-metrics are revised so that appropriate, hydrogeomorphically-specific, objective narratives are used as rating criteria, then CRAM could be retrofitted to provide the sensitivity needed.

S-RHEINHARDT-BP-a: See S-NIEMI-12 and S-RHEINHARDT-12 & 22 above. Wetlands with intermediate index scores can be distinguished based on differences in their metric scores (and sometimes their attribute scores). Index scores can be used to classify the wetlands as attaining intermediate condition, but distinguishing among intermediate-condition wetlands depends on examination of their attribute and metric scores, and the precision of these scores. With enough precision, intermediate-condition wetlands could be distinguished by a single different metric score.

Objectivity is discussed in greater detail in Part 2.

RHEINHARDT-BP-b: The most likely reason that subjective criteria are so common in CRAM is that the CRAM procedure was originally developed to assess all wetlands across a widely variable geologic and climatic region without regard to wetland class. Had development of the method focused on hydrogeomorphic subclasses first, perhaps by Physiographic Province or Major Land Resource Areas (MLRA), then more indicators might have been identified that could have provided better resolution in the intermediate range of condition. The development of separate field books for selected hydrogeomorphic wetland subtypes is a step toward better resolution, but objective criteria to replace the subjective ones are still needed for each wetland subtype, particularly criteria that are specifically calibrated to each subtype. In addition, in order to minimize confusion by the end user, all extraneous information that pertains to other wetland subtypes should be removed from the field books. Regardless of the objective criteria used, the narratives should be written to minimize bias by providing unambiguous, objectively rated metrics that force the user to choose the correct (intended) rating and minimize the potential for users invoking BPJ in their evaluations.

S-RHEINHARDT-BP-b: As stated in S-RHEINHARDT-3 and 10 above, the wetland classes were developed before CRAM, so that experts from different classes and regions could be assembled as the development team to increase the likelihood that CRAM would be adequately applicable across wetland classes. Initially it was expected that a single “module” would be applied to all wetlands in California; developing and testing the initial version of the methodology clearly indicated that varying the methodology for different wetland classes (patterned on the well-known hydrogeomorphic classes) was warranted. It is likely that the manual reflects some of this history and that the presentation in the manual will be improved in future manual revisions.

The overall structure of CRAM has been observed to be applicable across classes, but the metrics may vary among wetland types. Maintaining the basic structure of the method (i.e., the four attributes) across wetland types is considered to be advantageous in allowing joint considerations of the several types, while varying the metrics allows the methodology to be adjusted to the different classes. These adjustments are made by the L2 Committee as experience is gained with the methodology.

Objectivity is discussed in greater detail in Part 2.

RHEINHARDT-BP-c:
Even if protocol is improved by making narrative descriptions of subjectively assessed alternate condition states more objective and less ambiguous, there are still problems with some of the

objectively derived sub-metrics that should be addressed. Some of the problems related to the objectively assessed metrics are listed below, but details are also provided under "Specific Comments," at the end of the review.

1. Landscape Connectivity: Habitat size is important in cases where there are spatially dependent resources (food and shelter) required by a wetland-dependent species. The Landscape Connectivity metric is already biased favoring larger wetlands, and so connectivity requirements may be accounted for if condition is calibrated by regional wetland subclass at a size that is considered critical to a species of interest. Calibration by local wetland subclass is important because the average distance between wetlands is a landscape attribute under local geologic control. If calibration is simply by wetland class and not calibrated regionally, then the automatic bias favoring larger wetlands could be a problem if natural wetland size varies regionally. Similarly, the amount of connecting or natural upland habitat that supplies supplemental resources to aquatic dependant species is also important. In this case, there will be some critical size threshold of supplemental habitat below which resources will be insufficient to maintain a given population.

Thus, the size of the AA, plus the area of connected, intact habitat of similar type is important to condition. Neither optimal habitat size, nor amount of habitat providing supplemental resources, was considered in the Landscape Connectivity metric. Both are more important to condition than simply the average distance to other wetlands.

A more scientifically sound approach would be to define landscape connectivity in terms of supplemental habitat requirements of important wetland species, based on regionally specific wetland subclasses.

S-RHEINHARDT-BP-c: The scientific relationships discussed here and in the following sections were considered in the development of CRAM, and these (and other) scientific relationships are a central concern for the management of the methodology by the L2 Committee. CRAM is intended to assess the overall condition of a wetland, and thus the potential for the wetland to support high levels of the kinds of functions or processes that are expected of the wetland, given its geomorphic position and landscape context.

CRAM therefore by design is not calibrated to the needs of particular species, even if such needs are known. Such calibration likely would vary among species, leading to some average calibration, which is essentially the status of CRAM at this time.

In other words, the metrics reflect consideration of many functions by many experts for each wetland class and region of the State. It is conceded that a narrower focus on selected species or functions could improve the resolution of the method in their regard, but at the cost of reducing the relevance of the method in many other regards.

CRAM, or any conditional assessment, must balance the generality of the method with its ability to detect levels of change or difference that can influence management decisions.

Staff considers that when management concerns are focused on specific functions (such as a single species, flood peak attenuation, or carbon sequestration), Level 3 tools designed to assess those particular functions are required. Level 2 tools such as CRAM should not be expected to provide that level of detail.

RHEINHARDT-BP-d:

2. Average Buffer Width: This sub-metric is objectively measured, but the method used does not address all the underlying rationale for evaluating the metric. The method for determining the average buffer width ignores portions of the wetland that have no buffer (i.e., average width is determined from edges where there is a buffer) and provides no information about the importance of the proximity of buffer types to the wetland edge.

As written, Buffer Width may be useful for determining the size of supplemental habitat (which would provide “a greater capacity to serve as habitat for wetland edge dependent species”), but it does not help determine the capacity to protect the wetland from exogenous stressors (i.e., “reduce the inputs of non-point source contaminants, to control erosion, and to generally protect the wetland from human activities”). The assessment of supplemental habitat is better addressed in the Landscape Connectivity variable (see above), leaving buffer width to address the capacity of the buffer zone to ameliorate exogenous stressors. The ability of a buffer to ameliorate exogenous stressors depends on the landuse of the zone abutting the wetland boundary (buffer zone) and the distance of the landuse from the wetland edge. Native, mature vegetation provides better buffering capacity than successional vegetation or non-native cover types (e.g., row crop, managed lawn, impervious surfaces). All these cover types could be ranked according to their relative capacity to ameliorate exogenous stressors and by their distances from the wetland boundary (i.e., landuse closer to the boundary has more effect on buffering capacity than landuse further away). A more scientifically sound approach toward measuring buffer capacity would be to incorporate the above discussed condition parameters as a percentage of the entire wetland or AA edge (perimeter) in combination with buffer condition and distance metrics (see Buffer Condition below). This would require developing a multivariate condition metric that incorporates proportion of buffer zone in various landuse types in relation to their distance from the wetland or AA boundary.

Buffer Condition: As currently defined, buffer condition only relates to the proportion of non-native vegetation and intensity of soil alteration within areas defined as buffer. As stated above, determining the condition of the entire buffer zone is more scientifically sound than just assessing the condition of the portion of the zone that has a vegetated buffer. A better approach would be to not only evaluate the quality of landuse in the buffer zone in relation to its proportion of the perimeter, but also each cover type should be weighted according to the inverse of its distance (1/distance) from the edge (Figure 1). This inverse distance weighted (sensu King et al. 2004) effect has been used in other assessment applications (Rheinhardt et. al. 2007).

Rheinhardt, R., M. Brinson, R. Brooks, M. McKenney-Easterling, J. Masina-Rubbo, J. Hite, and B. Armstrong. 2007. Development of a referenced-based method for identifying and scoring indicators of condition for coastal plain riparian reaches. Ecological Indicators 7:339-361.

King, R.S., J.R. Beaman, D.F. Whigham, A.H. Hines, M.E. Baker, and D.E. Weller. 2004. Watershed land use is strongly linked to PCBs in white perch in Chesapeake Bay subestuaries. Environmental Science and Technology 38:6546–6552.

[Note from Staff: Figure 1, copied below, was provided by Dr. Rheinhardt as part of his review.]

Land use by cover type	LEFT SIDE ZONES (distance from stream)							RIGHT SIDE ZONES (distance from stream)					
	0-10 ft	%	10-50 ft	%	50-90 ft	%		0-10 ft	%	10-50 ft	%	50-90 ft	%
Old Forest	20		25		5		OF	20		25		5	
Mature Forest	20		25		5		MF	20		25		5	
Young Forest	19	35	24	35	5		YF	19		24		5	
Successional Forest	19		23		5		SF	19		23		5	
Recently Harvested	18		22		5		RH	18		22		5	
Shrubs/Saplings	17		21		4		SS	17		21		4	
Perennial Herb	16		2		4		PH	16		2		4	
Low intensity pasture	15		20		4		LIP	15		20		4	
Annual rowcrop	14		18		3		AR	14		18		3	
Low density residential			15	15	3	50	LDR			15		3	
Intensely managed lawns	9	65	11		2		IML	9	85	11	5	2	
Medium density residential			7	50	1	50	MDR			7		1	
High density residential			7		1		HDR			7	50	1	50
Medium density mobile homes			6		1		MDM			6		1	
High density mobile homes			5		1		HMD			5		1	
High density buildings			0		0		HDM			0	30	0	50
Impervious	0		0		0		IP	0	15	0	15	0	
Total % (QC check)		100		100		100			100		100		100
RZC Scores by zone		6.7		14.2		2.0			0.0		4.1		0.5

Figure 1. An example of a method for using an inverse distance weighted metric for determining riparian ecosystem buffer condition. The highest intensity developments would be scored as "Impervious" if within the two 0-10 ft zones, one on each side of stream. In this example, the total riparian zone cover (RZC) score is 27 out of a possible 100. Within each zone (three on each side of stream), the proportional cover is multiplied by the maximum possible score for the landuse type in that zone, which is proportionally weighted by inverse distance, area, and biomass. If the six bufffer zones were comprised entirely of Old Forest (top row), then the maximum possible score would be 100 (sum of top row).

S-RHEINHARDT-BP-d: This comment seems to turn on how the term "buffer" is defined or perceived. In CRAM, a "buffer" is an area outside a wetland that helps to protect conditions in the wetland from stressors in the surrounding region. Areas not serving to buffer the wetland are not counted. Rheinhardt's comment suggests an alternate perspective.

Rheinhardt suggests that all of the surrounding area should be considered and as buffer, with scoring of how well different parts of the surrounding area work as buffer. In developing his argument, based on work done in stream systems, the reviewer seems to equate "buffer" and "riparian." CRAM does not conclude that these two concepts are the same, although it is true that most wetlands have riparian areas and many have buffers. The CRAM concept of "riparian" is akin to the definition developed by the National Research Council (Brinson, et al., 2002)[2], that "riparian" refers to an area where a wetland (or other aquatic area) exchanges matter and energy with surrounding non-wetland areas. In CRAM a "buffer" is an area outside a wetland that helps to protect conditions in the wetland from stressors in the surrounding region.

Whether or not a wetland's riparian zone functions as a buffer depends on its composition and structure. An area is not automatically a buffer just because it adjoins a wetland. To be a buffer, an area adjoining the wetland has to have the structure and form to provide buffer functions. In general the degree to which a wetland area will be protected from external stress by a buffer depends on how much of the wetland has a buffer, how wide that buffer is, and what the condition of the buffer is (i.e., its per-unit area ability to function as a buffer due to its structure and composition).

During CRAM's development, the idea of modifying the buffer metric to account for differences among non-buffer land use types and their distance from wetlands was tested by regressing CRAM scores on many land use metrics calculated for zones of different width around wetlands. One conclusion was that land use data of the needed spatial resolution and accuracy were not reliable across the state. Moreover, such data are not widely available for California. However, where the necessary data were available, nothing correlated to CRAM scores (index scores and attribute scores excluding those for landscape context and buffer) better than the existing buffer metric.

Staff notes that better documentation and publication of results such as this can be, and perhaps should be, a part of future CRAM management, assuming that funding is secured to undertake such activity.

Objectivity is addressed in Part 2.

RHEINHARDT-BP-e:
3. Number of plant layers. There is a scientific basis for using the number of plant strata as an indicator of habitat structure. However, as defined, the strata do not adequately characterize condition for naturally forested wetlands, such as riverine wetlands, because the minimum threshold height for the "very tall" stratum is only 3.0 m (9.6 ft.). Mature forests, which should be the condition for "best achievable" condition, naturally support a canopy of trees much greater than 3.0 m high. As calibrated, a very young riverine redwood forest with a 3.2 m tall (10 ft) canopy would be rated the same as a mature redwood forest with a 100-m tall canopy.

Obviously, the condition of these two forests is very different: habitat is entirely different and the living and detrital biomass, indicators of biogeochemical cycling and nutrient uptake capacity, differ by several orders of magnitude. A more scientifically sound approach would be to define threshold heights relative to the characteristic structure of mature examples of regionally specific wetland subclasses.

S-RHEINHARDT-BP-e: The varying relationships among biotic variables such as canopy height and ecological functions have been, are, and will continue to be concerns for the L2 Committee. The L2 Committee will address this technical issue if a need to do so is demonstrated, using the process discussed in Statement 25.

---

[2] Brinson, M., et al. 2002. Riparian Areas: Functions and Strategies for Management, National Academies Press, Washington, D.C.

RHEINHARDT-BP-f:
4. Number of co-dominant species: This metric does not evaluate whether inappropriate species are present, it just evaluates the number of species (richness), which by itself is not always a reliable indicator of quality for most wetland types. That is, co-dominance could conceivably be comprised of inappropriate native species, which could indicate poor condition even if species richness is high and invasive species are unimportant.

For example, mesic species could dominate where hydric species should dominate or upland species could dominate areas where mesic or hydric species should dominate or successional species could dominate any zone. In addition, because dominance (>10%) is estimated over a large area, rather than in smaller plots, aspect dominance could affect the estimate, leading to biased and imprecise results. A more objective approach would be to determine dominance from measurements of relative cover, relative density, or relative basal area using plots (or with a plotless method) and compare the dominate species with a list of species expected to dominate in best achievable, mature sites (relative to regional wetland subclass).

S-RHEINHARDT-BP-f: More detailed (although arguably not significantly more objective) measures for most of the CRAM metrics undoubtedly could be devised. It's unclear that those measures would satisfy the need and intent to further develop CRAM as a rapid (that is, Level 2) assessment method. The detailed biotic field measures suggested by the reviewer are well-known research techniques in ecology, but requiring them would move CRAM away from Level 2 and toward Level 3 (i.e., quantitative, but not rapid).

The L2 Committee will address the technical issue regarding vegetation characterization if a need to do so is demonstrated, using the process discussed in Statement 25.

RHEINHARDT-BP-g:
The above outline describes the problems with metrics that were considered to have objective criteria for rating condition. Problems with metrics that were considered to require subjective interpretation to obtain a rating are discussed in the "Specific Comment" section below and in the main body of the review.

S-RHEINHARDT-BP-g: The need for objectivity is a main message and is addressed separately in Part 2.

RHEINHARDT-BP-h:

Specific comments

p. xi. The Table of Contents should be rechecked and corrected. I didn't check every entry, but the "List of Worksheets," p. xi, has many mistakes, see for example pages for Worksheets 4.2, 4.5, 4.8.4, 4.8.5, and 5.1.

p. 9, Condition. Is condition defined relative to a given wetland type? If so, then "within a given wetland type" should be added to the end of the sentence.

p. 11, Section 2.2.4. The assumptions of CRAM deserve an expanded explanation. Assumption 1: Why does "the societal value of a wetland (i.e., its ecological services) matter more than whatever intrinsic value it might have?" Does this mean that CRAM only focuses on the condition of ecological services or does it mean that CRAM considers wetlands manipulated to provide maximize ecological services more valuable than wetlands with intact ecological integrity and a lesser capacity to provide services?

Some wetlands that maximize services are either not sustainable or require energy subsidies to maintain them. Intact wetlands provide free services, even if those services may not always be provided at maximal levels.

Assumption 2: Why does “the value (of a wetland) depends more on the diversity of services than the level of any one service?” This assumption may be true, but needs elaboration. Assumption 3: The “diversity of services increases with structural complexity and size. CRAM therefore favors large, structurally complex examples of each type of wetland.” This assumption is likely not valid. Structurally simple wetlands are not necessarily depauperate relative to ecological services. Salt marshes are generally considered rather simple structurally, yet they provide many very important ecological services. Large size may be important, but access to supplemental upland habitat may be just as important.

For some species, a variety of wetlands of small size has been shown to be important to amphibians (Semlitsch and Bodie 1998, Semlitsch 2000, Babbit 2005).

Semlitsch, R.D, and J.R Bodie. 1998. Are small, isolated wetlands expendable? Conservation Biology 12: 1129-1133.

Semlitsch, R.D. 2000. Does size matter: the value of small, isolated wetlands. The National wetlands Newsletter, January-February: 5-13.

Babbitt, K.J. 2005. The relative importance of wetland size and hydroperiod for amphibians in southern New Hampshire, USA. Wetlands Ecology and Management 13: 269-279.

p. 21, Figure 3.2.
The flowchart in Fig. 3.2 does not show any groundwater controlled depressional wetlands in California. If there are groundwater-controlled depressions in California, then CRAM users will not key to them using the flowchart.

Kettle holes in moraines, mentioned on p. 5 of the Depressional field book, are usually groundwater driven, but I don’t know if there are any in California. Rains et al. (2008) showed that vernal pools on hardpan soils in California’s Central Valley were fed by shallow groundwater, though groundwater recharge is probably related to recent precipitation events. There are groundwater springs in the Modoc Plateau region and at the headwaters of many blue-ribbon trout streams (Rains, pers. comm.)

Rains, M.C., R.A. Dahlgren, G.E. Fogg, T. Harter, and R.J. Williamson. Geologic control of physical and chemical hydrology in California vernal pools. Wetlands 28: 347–362.

p. 22, under 3.2.2.1.
Height of bankfull flow does not have a “recurrence interval of about 1.5 to 2.0 years” in all riverine ecosystems. In flat landscapes (e.g., Central Valley), bankfull flows may occur several times per year (Sweet and Geratz 2003).

Sweet, W.V, and J.W. Geratz. 2003. Bankfull hydraulic geometry relationships and recurrence intervals for North Carolina’s coastal plain. Paper No. 02013 of Journal of the American Water Resources Association 39 (4): 861-871.

p. 23, under 3.2.2.1.2, sentence 2, which says, “A channel can be confined by artificial levees and urban development if the average distance across the channel at bankfull is **more** than half the distance between the levees or more than half the width of the nonurbanized lands that border the stream course.” This example is confusing because it seems to be the same example used for non-confined channels. It seems that the word “more” should have really been “less.”

p. 30, Table 3.7, under vernal pool type.
If preferred size is an area, then size should be square units, e.g., 1.0 km2, rather than 1.0 km. Also, 300m x 300m does not equal 1.0 km,2 so the area and dimensions are not consistent.

p. 37.
DOQQ is an abbreviation for digital orthophoto quarter quadrangle

p. 36 and others.]
The term “percent invasion” is not as clear as using “percent of invasive species”

p. 43, Section 4.1.1. It seems that “Landscape Connectivity” should really be called “Aquatic Connectivity,” since the metric measures an AA’s “spatial association with other areas of aquatic resources.”

p. 44, Part D.
Neither optimal habitat size, nor amount of habitat providing supplemental resources, was considered in the Landscape Connectivity metric. Both are more important to condition than simply the average distance to other wetlands. A better approach would be to define landscape connectivity in terms of supplemental habitat requirements of important wetland species, based on regionally specific wetland subclasses.

p. 44, Part D.
The assumption that 500 m is a critical distance for evaluating aquatic connectivity deserves a citation.

p. 44-45.
Landscape connectivity is does not take into account the quality of “connected” wetlands. They could be sources of invasive species.

p. 45. Paragraph 1 under “Riverine wetlands.”
Give example of barriers through which wildlife cannot pass and define “it.”

pp. 45-46.
Using the term “buffer” to mean land cover types constituting connectivity in wetlands is confusing, especially since “buffer” is also used in the more conventional sense of land cover types that ameliorate potential impacts to wetlands from uplands.

p. 45.
Last paragraph says that open water is a non-buffer land cover type. If so, then open water should have been included in column 2 in Table 4.4 (p. 49). Open water isn’t included as an acceptable cover type for buffers, but is identified as an acceptable buffer cover type for connectivity. Using the term “buffer” as a cover type in both contexts (buffer and connectivity) is

confusing. At the very least, a note should be provided in Table 4.4 saying that open water is considered a buffer cover type only in regard to connectivity.

p. 46, Fig. 4.1.
The direction of flow is needed to understand the caption.

p. 46, Table 4.3. The protocol says to assess for connectivity on only one side of channels that cannot be waded and both sides of channel that can be waded. However, Table 4.2 (p. 45) says that connectivity is assessed using remote imagery. If remote imagery is to be used, then it seems that it shouldn't matter whether a channel is wadeable or not.

p. 47. “Feral pets” is an oxymoron; an animal can be either feral or a pet, but not both.

p. 48. Buffer quality and intactness are known to be important to wetlands. However, the CRAM buffer indicators do not take into account:

1. Successional status of naturally forested buffers. For naturally forested buffers, successional status affects buffer condition from water quality and habitat perspectives. Mature forests have more carbon to fuel denitrification, more ability to uptake nutrients, and provide higher quality habitat than younger forests, fields, lawns, and other non forest cover types. In CRAM, native vegetation is considered sufficient buffer for best condition regardless of age.

2. Inverse distance weighting effects. Cover types closer to the stream edge are more important than cover types further away, following an inverse distance-weighted relationship (1/distance, sensu King et al. 2004). Alternative states of buffer condition (Table 4.8) do not take into account distance from the wetland. A buffer in good condition near the wetland boundary and poor condition toward the outer edge is scored the same as a poor condition near the boundary and a good condition at the outer edge.

King, R.S., J.R. Beaman, D.F. Whigham, A.H. Hines, M.E. Baker, and D.E. Weller. 2004. Watershed land use is strongly linked to PCBs in white perch in Chesapeake Bay subestuaries. Environmental Science and Technology 38:6546–6552.

p. 48, last sentence. Provide an estimate of % upland buffer in Fig. 4.2 to improve understanding.

p. 50, Table 4.6 and Fig. 4.3. Insert “average” before “Buffer Width.”

p. 52, (Table 4.8). Rating of buffer condition leaves much room for subjectivity in interpreting the various terms, e.g., “substantial amounts,” “moderate intensity,” “intermediate mix,” undisturbed soils,” etc.

p. 53, under Definition. Water sources were identified as “inputs of water into” and “diversions of water from” the AA. I think the later should be “diversion of water into” the AA, since the diversion is a source (input).

p. 53, paragraph 3 under Definition. It is not clear if a distinction is being made between natural and man-made inputs (sources). Man-made inputs usually suggest a reduction in condition.

p. 54, under Estuarine. The rating for water source allows for much subjectivity in identifying alternate states to determine a rating. For example, using plants to infer salinity tolerance assumes the end user is an expert at using plant composition to differentiate subtle changes from expected condition due to the source and salinity of water.

How could an expert, much less a novice, distinguish a difference in condition between an A or B rating in Table 4.9 using plants? This problem is likely due to trying to subsume condition of all wetland types under one rating system. Practitioners have to invoke BPJ and answers using BPJ can vary widely among people. More precise, unambiguous language (standards) should be used in different ratings, requiring a different rating narrative for each wetland type or at least for each broad HGM category.

The advantage is that more specific, repeatable, objective standards would provide clarity for each rating category.

p. 54, other wetland types. Using plants as indicators for water source in non-estuarine wetland types is fraught with the same subjectivity problems as stated above for estuarine plant indicators, but perhaps it is even more difficult in freshwater wetlands.
Using Reed (1988) to determine the permanence or consistency of water source based on wetland hydrophytic status seems particularly problematic. The hydrophytic plant list of Reed (1988) was only meant to be used to determine, with other information, if a particular area is a jurisdictional wetland.

p. 55, Table 4.9.
Rating conditions based on terms like “occasional or small effects” and “a few small storm drains” provide a lot of room for subjectivity. An example of an objective indicator is under Rating C: “Indications of substantial artificial hydrology include developed or irrigated agricultural land that comprises more than 20% of the immediate drainage basin within about 2 km upstream of the AA.” This description leaves no room for subjectivity. All ratings should be as unambiguous.

p. 56, under Seasonality.
It seems that the method loses a lot of utility if hydroperiod condition can only be assessed within restrictive and varying seasonal constraints.

p. 57, Table 4.10.
Some of these indicators are objective, others are subjective (and thus of limited utility), such as “compressed or reduced plant zonation,” “late-season vitality of vegetation,” “extensive fine grain deposits.” End users would have trouble determining how reduced zonation has to be, to be considered “reduced,” or how widespread deposits have to be, to be “extensive,” or what exactly defines “fine-grained.”

p.57, Table 4.10. Embolden “**out of**” to be consistent.

p. 57, Table 4.11a and b. Many indicators used to rate hydroperiod are subjective. For example, many ratings depend on determining if hydroperiod is different to some degree “than would be expected under natural conditions.” This means that the user has to know what would be reasonably expected and how much the assessed site differs from expectations. These ratings are just a way to assign BPJ to a rating, which incorporates biases of users.

p. 58, under Perennial Estuarine, third paragraph. The paragraph says that an increase in tidal prism tends to cause a reduction in hydroperiod, but that “marshes tend to build upward in quasi-equilibrium with sea level rise.” Hydroperiod doesn’t change for marshes that keep pace with sea level rise; hydroperiod would increases in marshes that cannot accrete fast enough. The reduction in hydroperiod is really only relative to what it would have been if sufficient accretion hadn't occurred. Therefore, rather than reducing its hydroperiod with sea level rise, an accreting marsh maintains its hydroperiod, albeit at a higher relative elevation. (see likewise, p. 17 in the Estuarine field manual.)

p. 59, Table 4.12. Why not use field indicators of hydroperiod alterations, such as tide gates or restrictive culverts, to evaluate condition? Otherwise, to determine if hydroperiod had been altered, one would have to install a tide gage and interpret the data to determine if alterations have really occurred. Using BPJ to infer altered tidal prisms introduces a high potential for observer bias.

p. 59, Table 4.13.
The ratings imply that less frequent than normal tidal inputs are worse than more frequent than normal inputs. There should be justification provided for this assumption. It seems that indicators of an alteration in frequency should be provided as rating criteria.

p. 60, under Riverine.
There is a bias in the assessment against headwater streams and wetlands, which comprise most of a river's stream network and are the source for most of a river's water quality problems. Page 56 says that hydroperiod is relatively unimportant in riverine wetlands, but worksheet 4.4 assesses hydroperiod (in terms of stability). Instability might be difficult to envision in a deeply channelized stream, since aggradation indicators would be subtle and the most stable channel would be a concrete lined one. This is an example of where degree of incision would be a better metric for condition.

p. 61, Worksheet 4.4.
The field indicators may have a bias against natural mountain streams under geologic control. A channel bed on bedrock with natural nick points would get downgraded. However, perhaps the long list of potential indicators in Worksheet 4.4 might be sufficient to differentiate condition. Only repeated testing can ensure that. Even so, there is no indicator for channelization. Channelizatin usually leads to aggradation as the channel attempts to re-equilibrate.

p. 61, Worksheet 4.4.
The worksheet indicators are not defined well enough to prevent subjectivity and so can lead to a wide variety of interpretations. For example, under Active degradation, "characterized by" in first checkbox is undefined (natural channels have cutbanks), "abundant" is undefined in 2nd checkbox, "age structure" is undefined in the 4th checkbox, etc..

p. 62, under Rationale.
There has been a great deal of science conducted on the relationships between carbon and nutrient cycling, floodplain/channel connection, but almost nothing was cited in this section.

p. 63. For riverine wetlands, channel entrenchment can result from channelization, but in rivers where channelization has occurred, there are no indicators for channelization or levees, both of which would be more direct indicators than entrenchment ratio. It seems that these indicators should at least be provided as alternatives.

p. 63. Entrenchment ratios will likely be measurable for most riverine systems, but perhaps not for rivers/creeks that flow across flat landscapes, where flood prone width could be could be extremely wide. In such cases, the ratio of bank depth (distance from top of bank to thalweg) to bankfull depth might be more indicative.

p. 64, Table 4.15b. The rating criteria for confined channels don't provide an altered condition to artificially confined (channelized) channels that were once, natural, unconfined channels. That is, natural channels that were altered by channelization and levee construction are treated the same as unaltered, confined channels. This could have grave consequences for protecting natural riverine ecosystems, in that for natural flowing rivers that are channelized, the CRAM assessment could give the river an A rating as long as the channelization maintains a entrenchment ratio greater than two.

Channelization (usually in conjunction with levee construction) is one of the most severe types of alterations that can be imposed on natural flowing rivers (dams are another) because it disconnects the river from its floodplain wetlands. Yet, the presence of channelization and levees are not explicitly acknowledged as an alteration to connectivity.

p. 65, 4.3.1, Patch Types.
For each patch definition, there should be a list of the wetland types that would be expected to support that type of patch. For example, the "hummocks and sediment mounds" indicator provided examples of wetlands that have these patch types. However, no wetland examples were provided for cobbles, boulders, undercut banks, and debris jams.

pp. 65-67. Patch Types.
The list and descriptions of habitat patches is comprehensive and is the best list of physical indicators for use as assessment metrics. The fact that the list is customized for each wetland type will likely insure that it is sensitive to variations in condition. This is the type of sensitivity minimizes subjectivity in ratings. The only indicator missing is depressions or small oxbows (abandoned channels) on river floodplains.

p. 65, Bank slumps.
Bank slumps were considered to be a negative indicator of hydroperiod in riverine wetlands (under active degradation) in Worksheet 4.4, p. 61, but here it is a positive indicator of habitat. Bank slumps and undercutting naturally occur in riverine systems; it's the intensity of erosion that determines degree of instability.

p. 66, Islands.
Spoil piles and fill could quality as an island, as defined. Fill should be explicitly excluded from the definition of an island; otherwise, one could argue that fill improves condition by creating an island.

p. 66, Microalgae.
Microalgae, in abundance, could also indicate eutrophication, a negative attribute.

p. 67, Secondary channels. Sentence 4,
A more descriptive term for "Tributary channels" would be "Drainage channels."

p. 69, Large and small swales.
End of sentence 1, "pools" should be "swales."

p. 70, worksheet 4.16.

As long as ratings are based on data from the least altered reference (best achievable) sites of relatively uniform size, then resulting ratings would be useful. However, since ratings are dependent on the area of the assessment site, it is critical that the relationship between ratings and wetland size have been accurately calibrated.

p. 72, Fig. 4.6.
Were these diagrams based on data (x-sections) from reference sites? If not, then would be unreliable to base ratings on them.

p. 74, section 4.4, second to last sentence.
The manual states that “Plant detritus is a main source of essential nutrients.” Essential nutrients for what? This statement needs further elaboration.

p. 74, under Definition, line 2.
Does “percent invasion” mean percent of species that are invasive or the % of an area dominated or covered by invasive species? This should be clarified. Page 77 says “percent invasion” is the % of co-dominants that are invasive. In this case, the term would be clearer if it were called “percent of dominant invasive species” or “percent invasive species.”

p. 75, under Rationale, line 1.
Not all wetland types have a naturally “rich” population of native flora, particularly as dominants. Many estuarine wetlands, for example, tend to have only one or a few dominants. This section, like other Rationale sections needs more support (citations) for the statements made.

p. 75, under Number of plant layers present. There is a scientific basis for using the number of plant strata as an indicator of habitat structure. However, the lower height thresholds chosen for the strata need more justification, particularly for “tallest” stratum in naturally forested wetlands. The presence of a mature forest canopy is a robust indicator of condition in forest ecosystems, yet as calibrated, a canopy only needs to be >3.0 m tall for this indicator to attain the highest possible rating. This means that, for example, for a riverine wetland dominated by coast redwood, a very young forest with a 3.2 m tall (10 ft) canopy would be rated the same as a mature redwood forest with a 100 m canopy. Obviously, these two redwood forests are structurally very different and their conditions differ markedly. The living and detrital biomass, indicators of biogeochemical cycling and nutrient uptake capacity, are also many orders of magnitude different. No rationale was given for why strata were defined as they were. In fact, for the strata ranging from “medium” to “very tall,” most scientists would probably concur that the categories should be an order of magnitude higher for most forested ecosystems (i.e., medium= 3-7.5 m, tall= 7.5-15 m, very tall = >15 m).

p. 76, Number of co-dominant species.
This metric does not evaluate whether inappropriate species are present, it just evaluates the number of species (richness), which by itself is not always a reliable indicator of quality for most wetland types. In addition, because dominance (>10%) is estimated over a large area, rather than in smaller plots, aspect dominance can affect the outcome, giving biased and imprecise results.

p. 76, last line. Delete “have.”

P. 77, first line. Change “Ni” to “No.”

p. 77, under 4.4.1.4. Determining the number of native plant species in a vernal pool (VP) requires a botanical VP specialist, which would likely severely limit the number of people who could do the VP assessments. Are some of the native species better indicators of high quality sites than others, say a select group of 15 or 20 region-specific species? Most field-oriented people could easily learn 15-20 key species, but differentiating >100 species is much less likely. If a group or groups of indicator species could be identified, the assessment could be made more rapid and user-friendly.

p. 77, under 4.4.1.4.
The "total number of native plant species listed in Appendix V" should be changed to "total number of native vernal pool plant species listed in Appendix VI."

p. 81, worksheet 4.8.7.
Insert "from Appendix VI" between "Species" and "Observed."

p. 83, under Figs. 4.7 – 4.10.
It seems that interspersion, as modeled in the figures, would be very difficult to rate in the field. Remote imagery might provide the best way to visualize interspersion.

However, there should be data from reference sites to support the calibrations, particularly as they relate to best achievable wetlands. For example, if unaltered estuarine wetlands sometimes show degree of interspersion similar to B or C (Fig. 4.10), then the ratings would give false results (i.e., an unaltered wetland could rate low, even though it might reflect what is possible under regional, geomorphic, and climatic constraints). Reviewers don't have access to the reference data on which the calibrations are based, so it is not possible to check this, but it seems the figures were based on some philosophical ideal rather than on reference data.

p. 87, under Vertical biomass structure.
The problems identified under "Number of plant layers present" (p. 75) are the same ones identified for this indicator because the same definitions for strata were used. There seems to be much subjectivity involved with determining "interspersion of plant layers" and "vertical biotic structure," and so field testing would be needed to confirm whether or not the indicators can be measured consistently (precisely), whether they can truly (accurately) differentiate condition along a continuum of degradation, and whether unaltered wetlands are being downgraded using this metric.

p. 91, paragraph 3, sentence 1. Change "immediate hat" to immediate adjacent area that."

p. 91. The assumptions used for the stressor checklist make sense.

p. 91, Table 5.1. Insert "natural" before "disturbance."

P, 101, under Step 1, paragraph 2. Change "less that 80%" to "less than 80%" and "Area than conduct" to Area, then conduct."

p. 102, Fig. 1. The cells shown in on Fig. 1 (4 x 5 grid) do not seem to match the frames used for Figs. 2-4 (6 x 6 grid?). It is difficult to tell if the circular frames are to be placed in the center of the cells or in the corners of the cells. See also Fig. D1, p. 41 in Appendix D.

p. 107, Appendix II Flow Chart.
Under Step 3, change "Appendix IV" to "Appendix V." it

seems that in an assessment area (AA) as large as that proposed (1 ha), it would be very difficult to determine, for example, if cover is 11 % for a layer that is 6% of the total area. That would require being able to estimate a total cover of 0.66% over a 1-ha site.

p. 109, Glossary. The term "channelization" is not defined, nor was it used as an indicator of connectivity between channel and floodplain. This is a serious omission, since California has a large number of channelized streams (Warner and Hendrix 1984).

Richard E. Warner, R.E, and K.M. Hendrix (eds.) 1984. California Riparian Systems: Ecology, Conservation, and Productive Management. University of California Press, Berkley, CA, USA.

Estuarine Field Book

p. 17, under the seasonal estuarine type.
Paragraph two is confusing. If a wetland is treated as a riverine wetland when the tidal inlet is closed, then it is not clear when, if ever, the wetland would be treated as a seasonal estuarine wetland. Only two choices are provided: perennial estuarine or riverine (presumably non-confined). Or there was a typo and "riverine" should have been "seasonal estuarine."

p. 19, Hydrologic Connectivity.
The rating of alternative states for hydroperiod (Table 4.12) requires that the user know how much the tidal prism has been altered, which cannot be determined during a half-day (Level-2) field evaluation. To determine the rating for hydroperiod, CRAM directs the user to look for field indicators like algal blooms and encroachment of freshwater vegetation to determine whether there has been any alteration in tidal prism leading to a lack of tidal flushing. While algae and vascular plants could provide useful clues that there may be lack of tidal flow and the stressor "flow obstruction: culvert" could be checked off in Worksheet 5.10, the user still must rely on best professional judgment (BPJ) to determine the degree to which tidal prism has been altered (rating scale in Table 4.12).

The stressor checklist (Worksheet 5.1, p. 93) provides two levels of stress that can be used as a basis for rating: (a) "Present and likely to have negative effect on AA"" or (b) "Significant negative effect on AA." These choices require using BPJ and so are subject to a wide variety of interpretations. A better approach would be to use the culvert and other indicators together to produce a narrative that boxes the user into a choice that does not to allow for subjective interpretations about whether a stressor is "present and likelyP effect" or a "significant negative effect."

Riverine Field Book

p. 21, Table 4.10. column 1 under "Reduce Extent," third bullet.
The word "**into**" should be "**out of**" to be consistent with Table 4.10 in the main manual.

p. 25, under "Note, first sentence."
Delete "uncertainty about."

p. 33, Layer definitions. The layer definitions from medium vegetation and taller do not correspond to the definitions on p. 77 in Table 4.8.1.

p. 33, Very Tall Vegetation. Evaluating the "very tall vegetation" (>2.0-m or > 3.0-m tall), as a part of biotic structure may make CRAM amenable to evaluating vegetation success in wetlands naturally dominated by herbaceous or shrub vegetation, but for naturally forested wetland ecosystems, such as naturally forested riverine wetlands, assessing the composition of 3-m tall vegetation will not be sufficient to define Best Achievable condition. This is because best achievable in a forested wetlands usually means that it has a mature canopy. Defining very tall vegetation as being > 3.0 tall might make CRAM amenable judging whether a mitigation site has met success criteria, but since the appropriateness of vegetation (correct species) and the relative abundance of species are not criteria measured by CRAM, a true evaluation of potential mitigation success is not possible with CRAM protocol. Effectively evaluating restoration success of vegetation in naturally forested ecosystems is a problem with all assessment protocols currently in use due to the unavailability of interim vegetation standards based on reference (successional) wetlands.

Depressional Field Book

p. 5, under Section 3.2.2.3, first sentence.
Add "and nutrient limitations" behind "impervious substrate" and change "controls" to "control."

p. 5.
Kettle holes in moraines, referenced on p. 5, are usually controlled by groundwater. Are there kettle holes in CA and if so, are they groundwater driven? Are there other groundwater driven depressions in California? If so, then the flowchart (Fig. 3.2) should be revised to include groundwater driven depressions.

p. 9, Fig. 4.2.
A depressional system should be used as the example for calculating buffer in the depression field book, not a riverine example.

p. 16,
Table 4.10, under reduced extent of inundation. In the cell in the top left, "**into**" should be "**out of**."

p. 17.
Paragraph 1 references Table 4.11b, but there is no Table 4.11b in the Depressional Field Book.

S-RHEINHARDT-BP-h: These numerous specific comments have either been addressed in the previous analysis, or they can be addressed as the CRAM manual and other supporting materials are revised.

WARDROP-BP: No BP comments received from Wardrop

Part 2
Discussion of Primary Technical Concerns

The message that emerges from the four reviews is that CRAM is adequate to meet the needs of the State for rapid assessment of overall condition of estuarine and riverine wetlands, and that the process of continued development of CRAM is adequate to produce additional CRAM modules for other wetland classes if the procedures for QA/QC, phased implementation and adaptive management are maintained.

The reviews mainly focus on ways to improve the method, but a few technical concerns merit further consideration. These are addressed below.

Objectivity and Subjectivity

The need for greater objectivity was abundantly discussed by Rheinhardt. No other reviewer recognized this need. This issue is addressed here because of Rheinhardt's repeated reference to it. .

Subjective Metrics: Rheinhardt's concern about subjectivity centers on use of what he believes to be subjective metrics. He identifies these metrics as relying on visual inspection and pattern recognition rather than quantitative measurements or counts. The metrics identified as "subjective" by this reviewer are similar to practices that are widely used in ecological field studies for measuring, sampling and characterizing vegetation and ecological patterns in the landscape. Using these techniques for CRAM purposes will not be associated with any more subjectivity than is their use in other kinds of ecological studies, such as:

> 1. The USACE delineation methodology: The 1987 Manual and the regional supplements, all use estimates of plant coverage to identify "dominance" in determining the degree to which a site satisfies the vegetation criterion, and
>
> 2. VegCAMP, CWHR and many others: Nearly all standard vegetation description/classification protocols in use, including the procedure specified by both the California Native Plant Society and the Vegetation Classification and Mapping Program (VegCAMP) of the Department of Fish and Game, use similar methods. It should also be noted that the same methodology links directly into the California Wildlife Habitat Relationships System, and is therefore also a standard basis for identifying and classifying wildlife habitats in California, including the use of the CWHR by federal agencies.
>
> 3. Reference Texts: Estimating percent cover is a standard procedure for sampling vegetation, taught in almost every university/college curriculum in ecology. How to do this is covered in many textbooks and in the lab manuals prepared by most faculty for use in their courses.

If these practices, when used in CRAM, are subject to being criticized as "subjective," then it would have to be agreed that (at minimum) both wetland identification and vegetation classification as carried out in the United States are also "subjective."

The problem here is that the reviewer apparently expects an "objective" methodology to be one where something is "counted" in binary terms (e.g., is present or absent, black or white, so forth) or scalar terms (e.g., 1,2,3,4,5...) without being subject to a judgment on the part of the evaluator. This is an unrealistic view of how field science is carried out. Staff finds that it is inappropriate to criticize CRAM methodology for using the same standards of practice that are readily accepted for so many other practices.

Perceived Source of Subjectivity in CRAM Development: Rheinhardt's concern stems in part from his assertion that CRAM was developed before definition of the wetland types or classes to which CRAM will be applied (See S-Rhienhardt-3b, 3c and 10 above). He contends that this forced adoption of subjective metrics as the only way to apply a single set of metrics across the broad diversity California wetlands. He therefore applauds our effort to develop class-specific wetland modules, and recommends that these modules minimize the use of subjective metrics.

Staff shares Rheinhardt's concern about subjectivity, and we agree that the modular approach can help address this concern. Some background information is needed, however, to understand the evolution of this approach.

The consortium of agencies that initiated CRAM development requested that the PIs *try* to develop a single method that could track the overall performance of projects and changes in overall ambient condition for all wetlands in California. The rationale for this approach is simply that one method is far less costly to develop and implement than multiple methods. The corollary is that a "one-size–fits–all" method may be too general to meet the objectives. In recognition of this risk, the classification system and the CRAM framework were developed at the same time, using the classification system to identify class-specific and regional experts to help make sure the framework was as broadly applicable as possible, while still being able to differentiate among wetlands along gradients of condition.

A statewide, multi-agency steering committee guided the development efforts through multiple seasons of iterative field tests. Eventually, a standardized framework was developed with three initial modules that could be adjusted as needed to account for differences among regions and wetland classes. Modular development focused on estuarine, riverine, and depressional wetlands because most of the other kinds of wetlands, for which additional modules might be required, were considered to be subsets of these three major wetland classes.

Ongoing efforts to develop CRAM and the classification system are tightly coupled. A statewide stream and wetland status and trends project is currently leading an effort to standardize classifications, with broad participation from many interested agencies and oversight from CWMW (follow this link to: SCCWRP Status & Trends Project)[3] . Broad ambient surveys are used to test the applicability of CRAM across wetland classes. New class-specific modules are developed only if strong field-based evidence that they are needed is found.

For example, experience with the draft depressional module indicates that some of its metrics are biased against seasonal systems. The existing depressional module will be tested along a gradient of hydroperiod to delineate seasonal from perennial systems, based on CRAM performance. The established process of module development for seasonal and perennial depressional systems will be applied to discern an appropriate cut-off between the two CRAM sub-classes.

---

[3] Full URL for SCCWRP Status and Trends Project is: http://www.sccwrp.org/ResearchAreas/Wetlands/WetlandMonitoringAndAssessment/Landscape-ScaleAssessment/WetlandStatusAndTrends.aspx

The reviewer's concern about what he regards as subjective metrics is not that they aren't scientific or potentially useful, but that they can introduce observer bias and thus reduce the precision and repeatability of CRAM. He is especially concerned that these kinds of metrics reduce the ability of CRAM to discern between different wetlands in the middle range of condition. He suggests that the efficacy of these metrics depends on their use by highly trained experts (i.e., experts with more expertise than is usually needed for wetland delineation).

Staff shares the reviewer's concerns. It is essential to establish minimum acceptable levels of CRAM precision and repeatability that are linked to the intended uses of CRAM. At this time, the thresholds are set at 10 CRAM points for overall index scores and attribute scores, and one grade level for metrics.

Field experience so far suggests that these thresholds can be met by the given QAQC procedures:

- Provision of clear manuals and fieldbooks
- Intensive standardized training at reference sites
- Tracking of practitioners through the CRAM database
- Auditing practitioners

Staff agrees with L2 that these procedures are capable of achieving CRAM's established objectives for precision and repeatability. Staff also notes L2's clearly stated possible error rate associated with observer bias for CRAM (i.e., 10 points). This allows those who analyze CRAM data to understand the level of uncertainty associated with CRAM scores and to account for this uncertainty in their decisions to the extent that the decisions rely on CRAM data.

Staff, L2 and the reviewer agree that existing class-specific modules should be adjusted, and new modules should be developed as needed, to meet the objectives for precision and repeatability. This objective will of necessity need to be balanced with CRAM's other objectives for representativeness, rapidity, and cost-effectiveness.

There will always be some tension among these objectives. CWMW and L2 are in place to manage that tension.

Staff also agrees with Rheinhardt that the differentiation of moderate-condition wetlands should be a goal of any assessment method (See S-NIEMI-12 and S-RHEINHARDT-12, 22, and BPa above). In the context of practical applications, however, there is a need to decide how much difference in wetland condition makes a difference (or should make a difference) to managers and regulators.

As stated above, the thresholds for precision have been set at 10 CRAM points for overall index scores and attribute scores and one grade level for metric scores. This means that, if necessary, CRAM should be able to distinguish between two wetlands differing by only one letter grade for any one metric. To determine whether or not this degree of differentiation is warranted, surveys of wetland condition for each validated wetland module along single-stressor gradients, to the extent that such gradients can be discerned by independent measures other than CRAM, have been proposed. Tests such as this would help us determine if the precision objectives of CRAM exceed the need for precision, based on the intended uses of CRAM.

Validation

Niemi has indentified a need to revise the validation process and a need to further validate the estuarine and riverine CRAM modules. He also expressed skepticism that the required validation process would not be implemented due to funding constraints. No other reviewers raised these concerns. In fact, two other reviewers (Anderson and Wardrop) applauded the validation effort. Rheinhardt accepted it is as the most practical approach.

Niemi summarized his concerns about validation by stating that it should require the gathering of data unique to the CRAM process; that it should include a wide spectrum of wetland characteristics including benthos, birds and plants, but also amphibians and fish as well as water quality; and that it should include more data points (more sample sites) than have been available thus far.

He acknowledges that the use of existing Level 3 data is fine for some of the validation, but that the analysis is not extensive enough, and that the process is biased due to the limited amount and kinds of Level 3 data that are collected explicitly as part of CRAM validation.

He also expresses some concern that the development process does not include the diversity of personnel necessary to minimize these biases, that CRAM does not include training, and he suggests that professional statisticians should be consulted on the adequacy of the validation process and interpretation of the results. He concludes that neither the estuarine nor riverine modules should be implemented until further validation is accomplished.

Staff agrees that further validation can be helpful in many ways, but we find that Niemi's comments reflect some false assumptions, and for both practical and technical reasons we disagree with his summary comments and conclusion.

His false assumptions are that CRAM does not involve training or a diverse scientific community, and that professional statisticians have not guided the validation efforts. In fact, both CRAM development and implementation involve training. Draft modules are shared with scientists from many disciplines in multi-day field trials that involve intensive trainings. Validation involves training of assessment teams including inter-team calibration (an annual exercise, not a one-time event in the past). These teams, in combination, involve personnel from physical and biological branches of environmental science. A standardized curriculum is being used to train practitioners and trainers. A separate curriculum is used to brief agencies and scientific institutions.

The design of the validation procedure and the interpretation of results have involved senior sampling design statisticians from Oregon State University, USEPA's Office of Research and Development, and California State University at Monterey.

Stressor Gradients: The basic problem, Niemi suggests, is the lack of an independent human stressor gradient along which Level 3 and CRAM data can be co-collected.

Rheinhardt also recognizes this concern, but acknowledges that such gradients perhaps do not exist for all wetland classes or for all regions of the State. He suggests that the most practical approach to validation is to assume that the best achievable condition and the worst observed condition define the possible range in condition and that the associated range in metric scores reflects the continuum between least and most altered condition. This is essentially the approach taken thus far in the development of CRAM.

CRAM's P.I.s used the Level 3 data to define stressor gradients, chose Level 3 sample sites along those gradients, and assessed the same sites using CRAM. A CRAM survey of riverine condition along local single-stressor gradients is planned (e.g., distance downstream from flow diversion or regulation, or timber harvest intensity) to test CRAM sensitivity among mid-gradient wetlands, but not to validate the riverine module.

<u>Validation is not Calibration:</u> An important aspect of CRAM validation is that the correlations between CRAM scores and Level 3 data are not necessarily used to revise or adjust CRAM. In other words, validation is not calibration; CRAM modules are not calibrated or tuned to improve the correlation between their scores and any particular Level 3 dataset.

Instead, the Level 3 data are developed or selected to represent the range in kind and level of functions that the module should be able to assess. Building a comprehensive set of Level 3 data that represents all the kinds and levels of functions of any wetland class is not possible. However, it is possible to build a few Level 3 datasets that together represent a broad range of functions for one or more wetland classes.

<u>Validation by Comparison With Level 3 Data:</u> The concern about the limited kinds of Level 3 data used to validate CRAM leads to two fundamental questions: How many kinds are enough, and which kinds are most important? As explained in the supporting documentation presented to the reviewers (See Peer Review Request Attachment 4 – References), there is no universally accepted dataset for overall wetland condition that can be used to assess CRAM performance – no "gold standard." CRAM was developed with what was judged to be the only practical approach, an approach Rheinhardt points out: Compare CRAM scores to Level 3 data that define a gradient of condition.

But, what Level 3 data should be used? Experience has been that each kind of Level 3 data has a community of scientists that will rationally lobby for its use. CRAM's P.I.s therefore developed criteria for Level 3 data selection that addressed timeliness, QAQC, documentation, availability, access to data authors, geographic scope, food web representation, and range of condition represented.

The selected datasets were constrained by these criteria. The requirement for statewide distribution was the single biggest constraint on the Level 3 datasets. In the absence of any "gold standard," the comparisons between Level 3 data and CRAM scores informed a consensus opinion on the performance of each metric. This is the process that was recommended by the P.I. team statisticians.

Niemi raises some concerns about the particular parameters from the Level 3 datasets that were selected to compare to CRAM scores. For example, he expressed concern that overall avian species richness was an inappropriate parameter because of its generality. However, Staff concurs with CRAM's developers that since CRAM is designed to assess the overall conditions of wetland areas, meaning their ability to support high levels of all their expected functions including species support, Level 3 parameters such as species richness are appropriate for testing CRAM validity.

The concern about the limited amount of Level 3 data used to validate CRAM leads to another fundamental question: How much Level 3 data is enough? It is true that Level 3 data sets that are statistically appropriate for validation studies are not commonly available, and not as diverse as the universe of potential factors that could be used.

Niemi suggests that the size of the data set should be proportional to the abundance of the resource being assessed. CRAM's developers instead accepted that data were sufficient when they could adequately define a gradient of condition and populate the gradient with enough sample sites to test for correlation between the Level 3 data and CRAM scores. According to this rule, the existing validation efforts have adequate sample sizes. This difference in rules for data sets would be difficult to test, since data sets such as those proposed by Niemi are rare, and would be costly in time and funds to compile. Staff accepts the CRAM developer's methods as being the only practicable approach to the problem of data set size

Staff notes that L2 is considering regionalizing the validation process so that regional Level 3 datasets can be used. This approach raises concerns about extrapolation of validation results from one region to another, but it will increase the likelihood of affording new Level 3 data dedicated to CRAM validation, and it will support the "weight of evidence" approach to assessing the validity of CRAM.

Staff disagrees with Niemei's conclusion that the estuarine and riverine modules should not be used until they are further validated. Validation studies were conducted with statewide Level 3 datasets that define broad gradients in condition. The correlations between CRAM scores and the selected Level 3 data were at least adequate if not robust for both wetland classes. These studies employed multiple lines of evidence. In most cases multiple Level 3 variables converged on similar relationships with CRAM scores, providing increased confidence in the validity of the method. The results should be sufficient to encourage continued development of the method and prudent application of the method to the recommended uses according to an accepted implementation schedule.

However, Staff finds that ongoing validation studies on various levels should be a part of CRAM's continuing development (or any scientific process). Such studies should include replications of previous work, repetitions of previous work in different settings or conditions, and unique new investigations. Additional validation would be helpful to guide the interpretation of CRAM scores for functions not represented by the existing validation, and would help verify (or refute) previous conclusions.

Staff notes that processes are in place to guide further development of these modules while they are being used. The riverine and Estuarine modules are currently being used in conjunction with Level 3 data to assess riverine condition statewide (over 2000 sites for this joint effort of SWAMP and CDFG have been assessed to date for one study alone), and to assess riverine and estuarine project performance in selected regions. These efforts are being guided by statewide teams of wetland managers, regulators, and senior scientists. This implementation is likely to continue to yield Level 3 and CRAM results that can be used to further validate these modules.

Separation of Level 1 and Level 2 (CRAM) assessments

Each of the reviewers recognized the value of Level 1 assessments. That is, the reviewers identified reasons to assess wetland condition in the context of their surrounding landscapes. This was a primary interest of Wardrop, however. She, and the other reviewers to much lesser degrees, raised three concerns relating to the use of Level 1 data.

Value of Level 1 Assessments: One concern was that Level 1 assessments (i.e., landscape profiles of wetland condition) have value unto themselves, and that the State should consider developing standardized Level 1 assessment methodologies. Niemi noted that substantial amounts of Level 1 data are publicly available that could be used to conduct Level 1 assessments, including identification of stressor gradients. Wardrop noted that the USEPA "1-2-3 Framework" (which has been proposed as a State Water Board framework for wetland and riparian area monitoring and has been endorsed by the California Water Quality Monitoring Council) calls for independent Level 1, 2 and 3 assessments.

Staff concurs with the reviewers and with the Monitoring Council's guidance on this point. An approach to Level 1 assessments, apart from Level 2 (CRAM) should be and is being developed. The State, in cooperation with federal agencies, has already begun to develop wetland and riparian area mapping tools for a statewide aquatic resources "base map" that is consistent with federal standards and that can serve the varied interests of the State's aquatic resource management and regulatory needs.

Large pilot mapping efforts that employ the new tools in the San Francisco Bay Area and along the southern California coast are providing a basis for developing approaches to Level 1 assessments. The U.S. EPA has recently provided funding for development of a Level 1 Status and Trends program for wetlands which will help address these needs; that study is already underway.

One main purpose of the base map is to serve as the sample frame for probabilistic Level 2 surveys of ambient overall condition using CRAM.

Wardrop notes that assessing wetlands unto themselves or as components of landscape mosaics is likely to influence the assessment results. Staff agrees, and therefore recommends the current guidance provided by the CRAM PIs and L2; i.e., that the analysis of wetland mosaics should typically be an aspect of Level 1 assessments and not CRAM. The question is not fully closed and merits further consideration. Future development may allow for means to assess multi-class wetland mosaics with CRAM.

Classification Issues Across Level 1 and 2: Wardrop suggests that CRAM metrics might not reflect the distinguishing characteristics of different kinds of wetlands. In other words, Level 2 assessments may or may not be affected by lumping wetland classes together. She suggests that CRAM might be used to assess the validity of the classification system by comparing CRAM scores for wetland classes assessed separately and in combination.

In the CRAM PIs' approach to this issue, wetland classes are primarily defined by water source and other aspects of hydrology. These criteria are not arbitrary. The hydrological differences translate into meaningful differences in function and hence can drive differences in management objectives. This is one of the main reasons for developing separate CRAM modules for each of these classes.

Since new wetland classes are only recognized when the existing CRAM modules are deemed non-applicable without major revision, it is very likely that each module will yield different scores for different combinations of wetland classes, and that each module will not be completely applicable to multiple classes.

CRAM scores are intentionally positively correlated to wetland complexity, and since assessment areas that include multiple wetland classes are likely to be more complex than areas of a single wetland class, we expect that the areas of multiple classes will tend to have higher scores, regardless of their particular class composition.

For these reasons, it would be inappropriate for CRAM to be used to assess the validity of the classification system, or to determine when classes can be lumped together for the purpose of a Level 2 assessment.

Separation of Landscape and Buffer Condition from Within-wetland condition: The third main concern about Level 1 data raised by Wardrop and touched upon by other reviewers is the possible need to maintain a separation between the assessment of landscape and buffer condition from the assessment of condition within wetlands.

Wardrop points out that the conditions within a wetland are strongly influenced by surrounding land uses, a fact that is reflected in CRAM's PSR model and conceptual models of forces that control wetland form and structure.

Rheinhardt points out that landscape and buffer characteristics could provide a context for evaluating condition within a wetland, but perhaps are not a part of that condition. In other words, landscape and buffer condition might be assessed separately and used to stratify within-wetland condition assessments.

Niemi similarly suggests that Level 1 assessments based on sound geographic information systems (GIS) could provide landscape assessments that might be verified by CRAM.

Wardrop notes that combining Level 1 and Level 2 data deviates significantly from this concept of independent but functionally interconnected Level 1-3 assessments. She further suggests that, since CRAM assumes that condition in the landscape and in the buffer adjoining a wetland influence conditions within the wetland, combining these measures of condition into one score eliminates the ability to test the effects of landscape condition or buffer condition on wetlands condition. This this would mean that CRAM could not test the validity of its models for PSR and the forcing functions that control wetland form and structure.

Staff shares this concern. One of the reasons for being able to easily extract each metric score or attribute score from the overall index score is to be able to test for their correlations. Investigations which explore possible correlations between the landscape-buffer attribute scores and the other components CRAM scores would be informative.

However, while landscape condition and buffer condition influence within-wetland condition, the reverse also occurs. This is reflected in the definition of the overall condition of any given wetland area as its potential to provide high levels of all the functions that are expected of it, including its landscape- and buffer-level functions, such as flood water storage that leads to flood stage desynchronization, and wildlife support that leads to safe dispersal and migration.

Landscape and buffer metrics for CRAM were intended to relate to these kinds of functions. These functions are considered to be separate from other functions that emerge at the landscape scale, such as the cumulative effects of local flood water storage or sediment filtration on flood frequency and watershed yields of sediment, and the contributions of all wetlands in a landscape to its overall biological diversity. Therefore, staff finds that no conflict exists between using carefully selected Level 1 data to assess the condition of a wetland in its landscape and buffer contexts (i.e., adjacency to other aquatic areas, ecological corridor continuity, buffer condition), and using other Level 1 data (i.e., wetland extent, size distribution, and variety) to assess aspects of condition that emerge at the landscape scale.

Stressor Indices

CRAM uses a stressor checklist to identify possible causes for low scores for metrics of condition, and to post-stratify the condition assessments to facilitate a statistical test of correlations between condition and likely kinds of stress.

This use of a stressor checklist was acceptable to the reviewers. However, most of the reviewers touched on the subjects of moving from the stressor checklist to a more quantitative assessment of stress, and that stress and condition together might comprise two aspects of a Level 2 assessment.

For example, Niemi noted that the checklist is a reasonable start for a Level 2 stressor assessment, but that future improvements could be made in this regard.

Rheinhardt suggested that, although stress and condition were separated in CRAM, stressors could be used advantageously in more instances to help explicitly evaluate condition. He went further to suggest that, since the stressor checklist only provides two levels of stress for the user to identify, and both choices require using BPJ, the stressor assessment is more subjective than it could (or should) be.

Wardrop noted that stress and condition are very difficult to separate, since many factors can be stressors for some functions and be aspects of condition for other functions. She noted the complications presented by invasive species in this regard, given that they contribute to structure and provide some habitat value, but can also have serious detrimental impacts.

Staff concurs with L2 and the CRAM PIs that Level 2 assessment of stress should be further developed. This subject has been explored in the past. It was found that some basic aspects of stress, such as its intensity, magnitude, and frequency, are seldom possible to assess rapidly. Current effort is aimed at further development of the lists of possible stressors for each metric, and to record the number of possible stressors that are observed. This is analogous to some of the condition metrics, such as physical patch richness and the number of plant layers, and is consistent with the newly developed USA RAM, a level 2 assessment tool that will be used in 2011 to help assess wetland condition nationwide.

Part 3
Priorities for Future CRAM Development
Based on the State Water Board Peer Review

Different priorities for the future development of CRAM have been suggested by different sectors of the user and support communities. The peer review conducted by the Engineer Research and Development Center (ERDC) of the USACE places a premium on the development of performance curves to enable wetland planners and managers to estimate how the overall condition of wetland projects will improve over time. The Sacramento and Los Angeles Districts of the USACE are emphasizing validation of the vernal pool module and development of an arid riverine module, respectively. Outreach to the communities of wetland interests in the Central Valley and the Sierra Nevada stress the importance of validating the depressional wetland module and developing a montane wet meadow/slope wetland module. The California Wetland Monitoring Workgroup (CWMW) of the California Water Quality Monitoring Council (Council) has established a statewide Level 2 Committee that, among other tasks, will serve as a clearing house and coordinating body for future changes to CRAM. The Level 2 Committee will need to rectify the priorities suggested by the State Water Board peer review with the other priorities listed above.

Based on this peer review, the priorities for future development of CRAM seem straightforward. These priority actions are listed in the order in which Staff, in consultation with L2 and the CRAM PIs, think they can be accomplished.

For all actions, better documentation and publication of all CRAM developmental work would serve to answer many questions and provide a stronger foundation for improvement and management of the method.

Staff is cognizant that funding for any of these priority actions is a separate issue, and understands that development schedules can proceed only as funding becomes available.

This lack of funding should not delay initial uses of CRAM, but it can inhibit or delay implementation of more complex uses of CRAM that otherwise might be feasible in later phases of development.

1. **Test CRAM sensitivity along single stressor gradients.** Minor funding is available at this time to test the sensitivity of the riverine modules along one stressor gradient defined by flow modification in the Bay Area. Additional funding to do this for depressional wetlands has been awarded but delayed. This effort could be used as a pilot for developing and testing a sample design and analytical approaches. The focus would be on the sensitivity of each metric of each attribute of condition plus the index of overall condition in the mid range of the gradient. This step would entail work, already under consideration by L2, to convert the existing stressor checklist into a semi-quantitative stressor index, possibly based on the USA-RAM model.

2. **Develop a strategic plan for validation of all CRAM modules.** It is clear from this review that the process of validation, as already established, is critical to the scientific integrity and validity of CRAM. While the amount of validation and other details relating to its geographic scope and the necessary amounts and kinds of Level 3 data might vary among the wetland classes, such that the cost of validation will also vary among the classes, it is unlikely that the needed validation will occur in any case unless a basic plan

is in place – including funding strategies. This action could be initiated by the CWMW, completed by the L2 Committee, and implemented through the Monitoring Council. This should include development of statistically valid data set goals for the purpose setting benchmarks for the points at which CRAM is judged to reliably perform the uses for which it was intended as listed in CWMW Bulletin 1.

3. **Explore correlations among the landscape context and buffer metrics and the other metrics for within-wetland condition.** This is necessary to determine the need and opportunity to elaborate the landscape and buffer assessments into a stand-alone Level 2 module. This action can be based on the existing data from the riverine and estuarine validation efforts and from the statewide and regional Level 2 ambient surveys of estuarine and riverine condition. No new data are necessary to initiate and perhaps to complete this action. However, if the decision is to develop a stand-alone landscape-buffer module, a plan and funding for module development will be needed. In pursuing this step, the landscape metric is faced with two different but interrelated issues. The first concerns the extent of wetland elements in the landscapes and whether the existing metric appropriately captures the desired relationships. The second is whether there ought to be a "real" landscape ecological assessment that identifies the status of a "landscape," rather than just the importance of wetlands. These issues are already before the L2 Committee based on work on the vernal pool modules.

4. **Revise the CRAM manual and existing fieldbooks.** This peer review has generated a wealth of suggestions for improving the clarity and utility of the documents that support CRAM practitioners. Grant funded work toward this end is underway.

5. **Develop a Level 1 assessment methodology:** This peer review indicates a need for a standardized approach to a basic Level 1 assessment. Continued Level 1 development is also needed to improve our ability to assess net change in wetland extent and condition over time. The CWMW has also recognizes this need. Initial funding for this effort has been obtained and a Level 1 Committee is being formed with good interagency participation and cooperation.

6. **Develop capacity to conduct ambient assessments:** Develop capacity to conduct ambient assessments for all wetland classes in all ecoregions of California