Attachment 2

Comments of Rheinhardt are in blue italics, following each numbered issue.

DESCRIPTION OF SCIENTIFIC TOPICS TO BE ADDRESSED BY REVIEWERS

The statute mandate for external scientific peer review (Health and Safety Code Section 57004) states that the reviewer's responsibility is to determine whether the scientific portion of the proposed rule is based upon "*sound scientific knowledge, methods, and practices."*

We request that you make this determination for each of the following issues that constitute the scientific basis of CRAM. An explanatory statement is provided for each issue to focus the review.

For those work products which are not proposed rules, as with the subject of this review, reviewers must measure the quality of the product with respect to the same exacting standard as if it was subject to Health and Safety Code Section 57004 requirements.

This Attachment 2 consists of 32 issues which the reviewers are asked to address.

The issues are essentially statements of principles, assumptions, and conclusions. The issues are arranged in five groups according to the relevant chapters of the following basic references:

1. *California Rapid Assessment Method for Wetlands User's Manual version 5.0.2* (cited below as the CRAM Manual). (1. Collins, et al., 2008a)
2. CRAM Estuarine, Riverine and Depressional Wetlands Field Books (2,3,4, Collins, et al., 2008b,c,d).
3. *Using CRAM To Assess Wetland Projects As An Element of Regulatory And Management Programs Technical Bulletin- June 5, 2009* (Technical Bulletin) (5. CWMW, 2009).

In addition, each issue includes reference where appropriate. Literature references, are numbered in order of citation, and presented in Attachment 4.

GOAL AND USES OF CRAM

CRAM is designed to provide information about the condition of wetlands and to identify the sources of stress that may adversely affect them.

Goal of CRAM*: To provide rapid, scientifically defensible, standardized, cost-effective assessments of the status and trends in the condition of wetlands and the performance of related policies, programs and projects throughout California.*

The CRAM Manual (1) and the CWMW Technical Bulletin (5) list various appropriate and inappropriate uses of CRAM. These recommendations are presented in Table 1:

TABLE 1:

Recommended Appropriate Uses of CRAM

a. Ambient assessment of wetland condition.
b. Monitoring of ecological reserves, mitigation banks, wildlife refuges or similar management units. Evaluation of wetland Beneficial Uses.
c. Evaluation of pre-project conditions at potential impact sites.
d. Evaluation of impacts associated with unauthorized (enforcement) actions.
e. Evaluation of pre-project conditions at potential mitigation or restoration sites.
f. Assessment of performance/success of mitigation or restoration sites.
g. Assessment of compliance of mitigation sites with required targets or performance criteria.
h. Comparison of proposed alternatives for regulatory or restoration planning purposes.

Generally Inappropriate Uses of CRAM

a. Jurisdictional determinations.
b. Focused species or threatened and endangered species monitoring.
c. Evaluation of specific management questions that call for Level 3 monitoring.
d. Evaluation of compliance with water quality objectives.
e. Assessment of mechanisms or processes of wetland function (diagnostic evaluation of wetland function).
f. Assessment of wetland values. It has been well documented that wetlands provide a variety of values that are beneficial to people, such as floodflow attenuation, aesthetics, and contaminant sequestration. CRAM is designed to evaluate the ecological condition of a wetland in terms of its ability to support characteristic plants and animals. Human use values cannot be appropriately assessed using CRAM.
g. Use of CRAM metric descriptors as stand-alone project design templates. While the narrative descriptions of best attainable conditions for the CRAM metrics can be used as general guidelines for overall project designs, they do not account for site-specific constraints and opportunities or design objectives. Because CRAM has been calibrated against statewide conditions, it is not appropriate to design a specific project based on the descriptions contained within each metric.

ISSUES TO BE REVIEWED

The Basis of CRAM

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. (CRAM Executive Summary, Ch. 1.3; Technical Bulletin Sec. 3B- C).

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

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)

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.

> *Stein, E., M. Sutula, R. Clark, A. Wiskand, and J. Collins. 2007. Improving monitoring and assessment of wetland and riparian areas in California through implementation of a Level 1,2,3 Framework. Southern California Coastal Water Research Project Technical Report 555.*

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

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. 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. 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.

<u>GROUP 2 ISSUES: (related to CRAM Manual Chapters 2, 5, and Appendix III; Technical Bulletin Sections 2, 3, and 4)</u>

4. Sources of stress or <u>pressure</u> affecting the condition or <u>state</u> of wetlands are identified in CRAM. This design facilitates management <u>*responses*</u> 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).

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.

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,).

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 large-scale 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 real-world 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.

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)

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. 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.

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).

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 <u>within</u> a wetland too, e.g., forest clearcutting, filling, ditching.

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)

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.

9. Metrics are measurable components of the four wetland attributes recognized by CRAM: Buffer and landscape context, hydrology, physical structure, and biotic structure. 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).

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.

10. CRAM has "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).

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).

11. Calibration is a process whereby iterative adjustments to the CRAM wetland classification system and metrics are made through multiple field tests. Through repeated field evaluations, the descriptions of the range of potential conditions are adjusted in relation to each 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 begins 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).

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 non-native vegetation" vs. "substantial amounts of non-native 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.

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)

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 patch-types 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 non-native vegetation" (Rating B) from "substantial amounts of non-native 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).

Experts might be able to obtain the same ratings for conditions B-D, especially if they have a chance to spend time together in the field to calibrate with one another, but non-experts would be likely to disagree, due to lack of experience, inherent bias, or hidden agendas. If a rating could be reasonably be interpreted in another way, then it probably will be. For most of the metrics or attributes that require subjective interpretation (BPJ), the level of expertise needed is beyond that possessed by most delineators, particularly indicators of hydrologic condition.

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.

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.

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, Sec. 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).

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.

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 Sec. 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 criterion 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.

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

16. 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). 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)

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).

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.

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.

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.

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.

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.

GROUP 3 ISSUES: (related to CRAM Manual Chapters 1.3, and Technical Bulletin Sections 2, 3, & 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 Sec. 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; ,23).

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.

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.

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.

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).

"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.

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).

This assumption is probably <u>not true</u> 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.

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)

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

24. CRAM Quality Assurance/Quality Control (QA/QC) procedures are consistent with sound scientific data management practices. Methods specified are sufficient to assure consistency in the statewide collection of data over time. (CRAM Ch. 3.8.2)

The QA/QC goals and procedures 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).

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 reliance on BPJ in interpreting metric or attribute scores.

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 Assessment Areas (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).

<u>GROUP 4 ISSUES: (see CRAM Manual Chapters 2,3,4, Appendix I; Technical Bulletin Sections 2,3, 4, and 5)</u>

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).

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.

26. In regards to Step 4 above: AA's, 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 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 AA's are summarized in Table 3.5 in CRAM, Ch. 3.5: Special criteria for delineation of AA's for Riverine and Vernal Pool wetlands are also described.

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.

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 AA's pertain to both of these purposes for using CRAM. However, they may require different numbers of AA's.

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.

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).

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.

The Readiness of CRAM Modules for Specific Wetland Types

GROUP 5 ISSUES: (related to CRAM Estuarine and Riverine Wetland Field Books)

A. CRAM for Estuarine and Riverine Wetlands

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 (2,3), 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).

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.

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).

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

B. <u>Depressional Wetlands</u>

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 (4). 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.

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.

C. <u>Slope, Playa, Lacustrine, and Vernal Pool Wetlands</u>

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.

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.

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.

The Big Picture

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.

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.

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.

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. *<u>Landscape Connectivity</u>: 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.

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.

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).

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.

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).*

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.

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 non-urbanized 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 km^2, 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 likely… 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.