Requested Review of Procedures
of the UMD/MAWP Best Management Practice Project
Year 2

STAC BMP Efficiencies Task Group

James Pease
Saied Mostaghimi
David Hansen
David Sample
Douglas Beegle
Steven Hodges

October 20, 2008

In June 2007, the University of Maryland/Mid-Atlantic Water Program (UMD/MAWP) requested that the Scientific and Technical Advisory Committee (STAC) review certain aspects of their Best Management Practice (BMP) Project. The two-year Project is charged with developing BMP reduction efficiencies for use in the Phase V Chesapeake Bay Model. At the end of year 1, the tasks requested of STAC were to: 1) review the relative rankings of BMP reduction effectiveness coefficients, and 2) review the process of developing BMP reduction effectiveness coefficients. The STAC Task Force directed to address the UMD/MAWP request considered the issues and returned the attached report, "Requested Review of Procedures for the UMD/MAWP Best Management Practice Project." The Task Force did not consider the relative rankings of BMP reduction efficiencies to be a scientific issue, but instead addressed the second task, that of reviewing the logic and process whereby MAWQ/UMD assessed recommendations made by experts and, in some cases, modified such recommendations. The Task Force considered justification of the direction and magnitude of such modifications to be the critical issue. Transparency and consistency are critical elements in this process.

At the end of the Project year 2, UMD/MAWP returned to STAC with a request for review of the process developed to produce Bay Model reduction effectiveness estimates for a second set of BMPs, which are listed in Table 1.

Table 1. UMD/MAWP Year 2 BMPs*	
Name	**Definition**
Ammonia Emission Reduction	Poultry litter acidifier treatment, biofilters and permeable plastic covers
Dairy Precision Feeding	Feed formulation so as to reduce N and P in manure
Dirt/gravel Road Erosion/Sediment Control	Driving Surface Aggregate Raising the Profile Grade Breaks Additional Drainage Outlets Berm Removal
Horse Pasture Management	>50% Cover Managed Species Traffic Management
Livestock/Poultry Mortality Composting	On-farm Composting vs. burying
Livestock Pasture Management	Rotational grazing Dairy Managed Intensive Grazing Beef and Other Livestock Managed Intensive Grazing
Infiltration/filtration	Bioretention Filters Open Channel Permeable Pavement and Pavers Infiltration Basins and Trenches
Nutrient Use Efficiency	Reduced application rate "Decision" agriculture
*Note: Table represents Task Force interpretation of UMD/MAWP documents.	

Acting on the basis of the Year 1 report from STAC, UMD/MAWP developed BMP definitions and effectiveness estimates for practices listed in Table 1, for which there are generally limited research results reflecting nutrient and sediment reductions. For all BMPs except ammonia emissions reductions, a panel of scientists with specific BMP expertise was convened and consulted to develop recommendations. For the ammonia emissions BMP, a single expert was consulted to develop the recommendation. In each case, the experts were directed to consider the following issues for the BMP in question:

- *Are natural characteristics (soil type, climate, flow paths, geology, vegetation, etc.) of the research site similar to conditions in the Chesapeake Bay watershed?*
- *Is the practice consistent with NRCS codes, jurisdictional stormwater design manuals? If not, how would effectiveness estimates be different?*
- *How critical is the duration of the experiment to the reported effectiveness results?*
- *Do results reflect changes in pollution reduction benefits over the lifetime of the BMP?*
- *Briefly explain the study method used?*
- *What parameters were sampled and monitored?*

- *Who conducted the research?*
- *How was the effectiveness estimate calculated?*
- *What was the scale of the study?*
- *What assumptions, outside of experimental results, were made in reaching the conclusions?* (Document: "Explanation of the application of 'Best Professional Judgment' in recommendation of BMP effectiveness estimates," Appendix C)

In addition, the following guidelines were used in selecting supporting literature for BMP effectiveness estimates:

- *Effectiveness estimates should reflect operational conditions, defined as the average watershed wide condition. Research scale effectiveness estimates should be adjusted to account for differences upon scaling up to operational conditions.*

- *Where studies with negative pollution reduction data (the BMP acted as a source, not a sink for pollution) are found, they should be included in the effectiveness development process as they reflect operational conditions.*

- *Peer reviewed literature has been subject to stringent evaluation and results from that literature are given more weight than literature that has not undergone the same review process by independent scientists. As such, peer reviewed literature should be given more weight than design standards and manuals. For this BMP, however, no peer reviewed literature was available and gray literature, or limited research scale type publications, and best professional judgment was used.*

- *Data from individual BMP project sites are to be utilized over median or average values calculated from multi-site analysis (meta-analysis). Single site studies evaluate individual BMP projects, while multi-site analyses are a collection of BMP projects.* (Document: "Explanation of the application of 'Best Professional Judgment' in recommendation of BMP effectiveness estimates," Appendix B)

The UMD/MAWP project developed a systematized process for adjusting literature-based or panel-developed BMP reduction effectiveness coefficients. Adjustments were made on the basis of BMP specification similarity to technical standards and to Chesapeake Bay soil and hydrologic conditions, and on the basis of the scientific support for results, results variability, and the number of studies supporting results. The process developed by UMD/MAWP to adjust effectiveness coefficients is presented in Table 2, with further explanation in Table 3.

Table 2. UMD/MAWP Decision Matrix/Adjustment Matrix*

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5
Effectiveness Estimates Assigned → **Research Attributes ↓**	**Average**	**Below average (between average and 1st quartile)**	**Low end of range (within 1st quartile)**	**Conservative estimate with maximum of 30%**
Applicability/ Specification of BMP	Within State TS definition and NRCS standards; matches Stormwater Manual Design Specifications	Generally representative	Somewhat representative	n/a
Study Location	Within Chesapeake Bay Watershed – representative soils and hydrology	Generally representative	Somewhat representative	n/a
Results Variability	Low variability	Medium variability	High variability	n/a
No. of studies	High	Medium	Low/limited	None
Scientific Support	Operational scale research (peer reviewed)	Research scale (peer reviewed)	Not peer reviewed ("gray" literature)	Best professional judgment, observation and/or extrapolation
*Note: Table represents Task Force interpretation of UMD/MAWP documents.				

Table 3. Description of Table 2 Items

Applicability/Specification of BMP

a) **Completely consistent** within jurisdiction and NRCS technical standards; or stormwater manual design standards: use average of the range of results

b) **Generally consistent**: use a value below the average and above the 25th percentile of research results. "Generally" is defined as representing 67% or greater of the standards and specifications within jurisdiction and NRCS technical standards or stormwater manual design standards.

c) **Somewhat consistent**: use a value no greater than the 25th percentile of research results. "Somewhat" is not defined by UMD/MAWP.

Study Location

Location is defined as the average soil conditions and hydrologic regime associated with typical land use.

a) **Completely Representative**: When all studies are representative of the conditions within the Chesapeake Bay watershed, the average of the range will be selected.

b) **Generally Representative**: When the natural conditions of the research area are generally representative of those in the Chesapeake Bay watershed, a value below the average and above the 25th percentile of research results will be selected. "Generally" is defined as the study being similar to, but not exactly the same as the soils and hydrology of the Chesapeake Bay watershed.

Somewhat Representative: When the location of the studies are somewhat representative of the soil and hydrologic conditions in the Chesapeake Bay watershed, a value no greater than the 25th percentile of research results will be selected. "Somewhat" is not defined by UMD/MAWP.

Results Variability

a) **Wide range** of results variability: a value no greater than the 25th percentile of research results will be selected. "Wide" is not defined by UMD/MAWP.

b) **Medium range** of results variability: use a value below the average and above the 25th percentile of research results.

c) **Low range** of results variability: use average of results range

Number of Studies

a) **High**: Greater than 6: use average within the results range

b) **Medium**: 4-6 studies: use value below the average of range reported

c) **Low/limited**: Less than or equal to 3 studies: use low end of results range or more conservative effectiveness estimate

d) **None**: No more than 30% effectiveness (See 'scientific support' below)

Scientific Support

a) **Peer reviewed studies that analyze practices in an operational setting on local watersheds** that are applicable to expected conditions throughout watershed: use average within the results range

d) **Research plot scale**: Studies that investigate practices on research plots on local watersheds that are applicable to expected conditions throughout watershed: use a value below the average and above the 25th percentile of research results

b) **Gray literature**: White paper, or limited research scale type publications, regardless of location: use a value no greater than the 25th percentile of research results

c) **Other**: Best professional judgment, observation, and extrapolation: conservative effectiveness estimate below 30% will be used. Rationale for selecting a 30% effectiveness estimate: most watershed studies show that when applying a suite of BMPs to a watershed, maximum reductions are about 30%. As such, no effectiveness estimate for a single practice recommendation based primarily on best professional judgment, extrapolation or observation should be more than 30%.

*Note: Table represents Task Force interpretation of UMD/MAWP documents.

The Task Force commends the efforts of the UMD/MAWP Project during year 2 for increasing transparency and consistency in the BMP adjustment process. Our interpretation of the Adjustment Matrix (called the Decision Matrix in the UMD/MAWP documents) application is thus: if (for example) a large number of operational scale peer-reviewed studies were determined to have been conducted in the Chesapeake Bay region with BMP specifications within state or Natural Resources Conservation Service (NRCS) technical standards, then the average value of reported results would be calculated to determine the BMP reduction effectiveness to be presented to the Chesapeake Bay Program.

The Task Force has questions about the Adjustment Matrix process, as well as its application for specific BMPs:

General questions concerning the efficacy of the Matrix:
   a) The Task Force considers that there is no scientific justification for assigning reduction effectiveness for BMPs with only 'gray' literature, Best Professional Judgment, observation, or extrapolation support. Standards of research quality and review form the basis of public trust in Chesapeake Bay Program efforts, and should not be compromised.
   b) BMP effectiveness over time is not considered in the Matrix. The UMD/MAWP documents do not provide information how BMP effectiveness may decline over time or be restored with maintenance activities.
   c) The number of studies supporting BMP effectiveness is, in some cases, very small. The Task Force recommends that no BMP effectiveness be recommended if not supported by at least 3 peer-reviewed studies, unless a panel of at least 3 experts recommends a BMP reduction effectiveness based on their judgment and agreement.

Questions about how the Matrix was used that should be clarified:
   a) BMP reduction effectiveness adjustments would be more transparent if panel/outside expert recommendations were listed, research attributes indicated by the Matrix as presented in UMD/MAWP documents were identified for each BMP and UMD/MAWP justification given for the magnitude of adjustments. This transparency would improve the credibility of the adjustment process.
   b) The Matrix is ambiguous in its prescribed adjustments for cases of research with mixed attributes. For example, how are two research results weighted when one is "completely consistent" with technical standards, and another is only "somewhat consistent." If supporting studies include some combination of peer-reviewed operational scale and peer-reviewed plot scale, how is research quality assessed and weighted in determining the BMP reduction effectiveness? It appears from the UMD/MAWP documents that each study result in weighted equally, without consideration of research attributes. Adjustment rules for research results with differing attributes should be clarified.

c) The Matrix states that the "average or median" of research results will be presented as the BMP reduction effectiveness if research attributes satisfy the standards in the matrix second column (e.g. equal to NRCS Technical Standards, representative of Chesapeake Bay soils and hydrology, high number of operational scale, peer-reviewed studies with low variability of results). Examination of UMD/MAWP recommendations for specific BMPs shows that the arithmetic mean of research results has been utilized, so the Task Force presumes that use of the word "median" in the Matrix is in error. If not, then conditions appropriate for selection of the mean or median should be specified.

d) Column 3 of the Matrix states that with less-than-optimal research attributes (e.g. BMP specification is "generally" representative of standards), the UMD/MAWP BMP effectiveness recommendation will be between the average value of research results and the 25th percentile of such results. If research results are ordered by magnitude, the 25th percentile is the value below which 25% of the ordered results lie. Since there is no unambiguous relationship between the arithmetic mean and the value at the 25th percentile (the 25th percentile may be higher or lower than the arithmetic mean), the matrix does not provide a clear rule for determining BMP effectiveness recommendations in these situations. Assumptions of symmetry or normality of research results should be clearly stated. In any case, the Task Force finds no application of this adjustment rule in the UMD/MAWP documents, and the BMP research addressed by UMD/MAWP may not have fallen into these adjustment categories.

e) An adjustment specified by column 4 of the Matrix (e.g. BMP specification is "somewhat" representative of standards) is ambiguous, stating that a BMP reduction effectiveness value will be selected that is no greater than the 25th percentile of research results. However, this would involve rank-ordering of research results, calculation of the percentiles of reported results, and linear interpolation between ranks to calculate the 25th percentile. In any case, the Task Force finds no application of this adjustment rule in the UMD/MAWP documents, and the BMP research addressed by UMD/MAWP may not have fallen into these adjustment categories.

f) An adjustment specified by column 5 of the Matrix, for which only Best Professional Judgment, observation, or extrapolation BMP effectiveness indicators are available, indicates that a recommended BMP effectiveness value will be no more than 30% from the non-BMP loss scenario. Justification is stated as: "*Rationale for selecting a 30% effectiveness estimate when best professional judgment is used is justified because most watershed studies show that when applying a suite of BMPs to a watershed, maximum reductions are about 30%. As such, no effectiveness estimate for a single practice recommendation based primarily on best professional judgment, extrapolation or observation should be more than 30%.*" The UMD/MAWP Project thus states its preference for a specific reduction effectiveness value for those BMPs for which there is no scientific support. However, the Task Force considers that the Project should provide an objective reference for its claim of "maximum reductions are about 30%," and justification for selection of a particular effectiveness magnitude between 0% and 30% for particular BMPs.

g) The "results variability" appears to be a judgment call by UMD/MAWP, and thus there are grounds for claims of subjectivity and bias. The Task Force recommends that a specific metric be developed to define "high", "medium", and "low" variability. Some problems in defining "results variability" will be solved by defining weights for research quality.

Questions concerning application of the Matrix to specific BMPs:
   a) In general, UMD/MAWP documents do not describe the prior-condition assumptions about practices prior to application of the BMP. The Task Force considers that BMP reduction effectiveness should clearly describe the condition and the nutrient/sediment losses from which reductions occur upon application of the BMP.
   b) UMD/MAWP documents indicate that it was impossible to empanel an expert group for the ammonia emissions BMP, so one expert was charged with developing the BMP reduction effectiveness factors. The Task Force considers the judgment of a single expert, no matter how qualified, to be insufficient to support BMP reduction effectiveness factors.
   c) Weighting Chesapeake Bay research more heavily than research in other locations is not justified for every BMP. For example, substantial research has been conducted in Arkansas, North Carolina and Georgia on alum applications to poultry litter. The Task Force asserts that conditions in a poultry house in Arkansas do not differ substantially from those in a similar house in the Chesapeake Bay watershed, and Arkansas research results should be considered equal to any obtained in the Bay watershed.
   d) For the Dirt Road BMP, since the research did not consider the first flush, and since there is only one study, the Task Force suggests that there is not enough scientific support to develop a BMP effectiveness reduction factor.
   e) For the Infiltration/Filtration BMP, there was some concern among the Task Force members that only positive reductions reported in the literature were considered in developing the BMP effectiveness reduction factor. A process needs to be developed to consider the negative values as well.
   f) For the Pasture Management BMPs, the Task Force asserts that there are runoff studies that provide results applicable to Bay watershed pasture management. The Task Force was asked to provide guidance on using RUSLE2 for estimating pasture phosphorus and sediment reductions. If this option is chosen, it is important that pre-BMP pasture conditions be carefully specified and simulated for conditions representative of the watershed, as well as post-BMP pasture conditions. RUSLE 2 is an index based method to estimate soil loss and does not attempt to explicitly model erosion processes (RUSLE 2 Users Guide – Draft version 2). RUSLE 2 also does not estimate erosion due to concentrated flow, which may well occur on poorly managed horse pastures. To generate any estimates of soil losses, detailed information on each hill slope would be needed. In addition, differences in canopy cover, ground cover, soil (surface) roughness, ridge height (although probably not applicable in this case), soil biomass, soil consolidation, and antecedent soil moisture of pre-BMP and post-BMP conditions are required, and estimates would apply only to the specific site. Although sediment losses could be estimated in this manner, RUSLE 2 would not assist in calculating nutrient losses. To truly model nutrient losses, a GLEAMS-based modeling approach (such as Answers or

SWAT) would be preferable, but even these models specify explicitly that they do not accurately model P movement in situations where soil P increases rapidly from repeated manure applications. Such would be the case when animals are seldom off the pasture and where supplemental feeding of animals and fertilization/manure deposition occur. In this case, there is no good substitute for well designed field research studies.

g) UMD/MAWP suggests use of the Chesapeake Bay Model to estimate nutrient and sediment reductions from Nutrient Use Efficiency (NUE). For the Reduced Application Rate BMP, a straightforward reduction in fertilizer applications in a model segment will result in decreased nutrient loadings. However, for the Precision Agriculture BMP, the Bay Model is not adequate to reflect a change in nutrient/sediment loadings, since such factors as variable application rates, precision application, and optimal application timing are not represented in the Model. Yet, these factors may substantially increase nutrient uptake efficiency, and thus reduce nutrient losses without reducing nutrient application rates. The Task Force suggests further review of the literature to provide guidance for establishing BMP effectiveness factors for Precision Agriculture. The Panel was confused by the term "decision agriculture", and found it unhelpful to define a new term that is not backed up by significant practical or conceptual science.

h) For the Dairy Precision Feeding BMP, the documents state that "The average (of research results) was calculated and rounded down to the nearest factor of five because there were 4 studies with direct data available for phosphorus reductions, in agreement with the decision matrix." As interpreted by the Task Force, use of 4 studies indicates that a value less than the average would be recommended, and the adjustment is ambiguous. In addition, UMD/MAWP has not indicated how the relevant research fared on the other Matrix research attributes.

i) For the Livestock/Poultry Mortality Composting BMP, two concerns are expressed by the Task Force. First, the baseline comparison of mortality composting is carcass burial. For permitted poultry operations in Virginia and Delaware, routine carcass burial is not allowed, so the baseline comparison with burial is invalid. In essence, a Virginia or Delaware poultry operation using mortality composting will be counted as reducing nutrient loadings, when in fact no such reduction occurred. A second concern is applicable to both Mortality Composting and other BMPs that have air quality as well as water quality impacts. It is recommended that BMP reduction effectiveness factors consider the net effect of nitrogen compound losses to the atmosphere and to water bodies. Nutrient loss reductions to water bodies should not be fully credited without consideration of offsetting re-deposition of atmospheric nitrogen.

Recommendations

The Task Force commends the efforts of the UMD/MAWP project to provide transparency and consistency to the process of estimating BMP reduction effectiveness factors. The Adjustment Matrix is a considerable step forward from the process reviewed for Year 1 BMPs. We suggest that Temporal Performance of BMPs should be an additional factor for determining reduction efficiencies based on research results. More formalized metrics for factors such as results variability should be developed and utilized to ensure transparency and consistency. We recommend that the Chesapeake Bay Program Office provide resources to review Year 1 BMPs using the adjustment principles developed for Year 2.

In addition, the Task Force recommends that the CBPO adopt formalized procedures such as initiated by the UMD/MAWP Project to evaluate BMP research results when establishing or reviewing BMP reduction efficiencies. We recommend that the CBPO undertake periodic literature reviews to assess the current state of scientific evidence supporting existing and new BMPs and update the reduction effectiveness factors on a regular basis.