ENCLOSURE 2
EPA'S LIST DEVELOPMENT PROCESS

Clean Water Act (CWA) Section 303(d), 33 U.S.C. § 1313(d), (Section 303(d)) requires states to identify those waters within their jurisdiction for which effluent limitations required by CWA Section 301(b)(1)(A) and (B), 33 U.S.C. § 1311(b)(1)(A) and (B), are not stringent enough to implement any applicable water quality standard, to establish a priority ranking for such waters, and to submit a listing of such waters to the U.S. Environmental Protection Agency (EPA) (Section 303(d) list).

On December 21, 2012, EPA received from the West Virginia Department of Environmental Protection (WVDEP) West Virginia's 2012 Section 303(d) list of water quality limited segments (WQLSs) (West Virginia's Section 2012 303(d) list), as part of the Integrated Report submitted by WVDEP (submission) to meet the requirements of CWA Sections 303(d), 305(b), and 314; 33 U.S.C. § §1313(d), 1315(b), and 1324. As described in Enclosure 1, EPA has partially disapproved West Virginia's 2012 Section 303(d) list submission because WVDEP did not evaluate all existing and readily available water quality-related data and information, specifically, information related to whether certain waters are achieving West Virginia's narrative water quality criteria as applied to the aquatic life uses (W. Va. CSR § 47-2-3.2(e) & (i)), when it developed West Virginia's 2012 Section 303(d) list. See 40 CFR 130.7(b)(5). As required by 40 CFR 130.7(d)(2), EPA developed a list of waters that are not achieving West Virginia's water quality standards by evaluating this existing and readily available water quality related data and information utilizing a methodology previously used by WVDEP in connection with previous Section 303(d) lists.

In so doing, EPA utilized as its starting point WVDEP's most recent methodology (the West Virginia Stream Condition Index (WVSCI) for assessing compliance with narrative water quality criteria as applied to achieving and maintaining aquatic life uses. Below is a description of the methodology and sources of data that EPA used to (1) assess the specific water quality data and information; and (2) determine whether/which waters were not in compliance with West Virginia's narrative water quality criteria as applied to the aquatic life uses. Using that methodology and considering this existing and readily available data, EPA proposes to add 255 water quality limited segments to West Virginia's Section 303(d) list. A list of the waters that EPA proposes to add to West Virginia's Section 303(d) list is in Enclosure 3. EPA will issue a notice in the Federal Register of our proposed action within 30 days of this disapproval. There will be a 30 day public comment period. Upon completion of the public comment period, EPA will review all comments and make changes to the proposed list accordingly.

West Virginia's Narrative Water Quality Criteria

West Virginia's narrative water quality criteria (W. Va. CSR § 47-2-3.2(e) & (i)) provide:

3.2. No sewage, industrial wastes or other wastes present in any of the waters of the state shall cause therein or materially contribute to any of the following conditions thereof:

* * *

3.2.e. Materials in concentrations which are harmful, hazardous or toxic to man, animal or aquatic life;

* * *

[and] 3.2.i. Any other condition, including radiological exposure, which adversely alters the integrity of the waters of the State including wetlands; no significant adverse impact to the chemical, physical, hydrologic, or biological components of aquatic ecosystems shall be allowed.

WVDEP's Methodology for Assessing Its Narrative Water Quality Criteria on Past Section 303(d) Lists

EPA used as its starting point the methodology that WVDEP has historically applied for assessing waters against West Virginia's narrative water quality criteria. Starting with its 1998 Section 303(d) list through its 2010 Section 303(d) list, WVDEP has considered the health of the macroinvertebrate community as its primary means of directly measuring whether the narrative water quality criteria as applied to the aquatic life uses are being satisfied.[1] Beginning with its 2002 Section 303(d) list and continuing through its 2010 Section 303(d) list, WVDEP utilized the WVSCI as its methodology for assessing whether streams are achieving West Virginia's narrative criteria as applied to the aquatic life use. WVSCI consists of six benthic community metrics combined into a single multimetric index. It was developed by Tetra Tech, Inc. on behalf of WVDEP and in coordination with EPA in 2000 using WVDEP and EPA data collected from riffle habitats in wadeable streams. Generally, all metric values were converted to a standard 0 (worst) to 100 (best) point scale. The six standardized metric scores were then averaged for each benthic sample site to come up with a final index score ranging from 0.0 to 100.0. Using the distribution of scores from all sites that are considered reference sites, a threshold score of 68.0, representing the 5th percentile of reference sites, was identified by WVDEP as the lowest WVSCI score that was considered as fully supportive of the narrative criteria as applied to the aquatic life uses.[2] This means that 95% of all reference sites had a higher score. Setting a threshold as a percentile of the reference population corresponds to setting the acceptable significance of a hypothesis test (α), or the acceptable type 1 error rate (false positive), as the reference percentile.

In its 2002 through 2010 Section 303(d) lists, WVDEP described how it used WVSCI to assess waters against West Virginia's narrative water quality criteria for achievement of the aquatic life use. Generally, a score greater than 68.0 was considered unimpaired. Waters with a score of less than 60.6 were considered impaired and placed on the Section 303(d) list. Waters with scores of 60.6-68 (gray zone) were considered neither impaired nor unimpaired.

---

[1] West Virginia also assesses ambient levels of various parameters against numeric water quality criteria established to protect the aquatic life use.

[2] As a general matter, the reference sites will have experienced some alteration and thus represent some degree of departure from truly natural conditions. To account for this, many states (Virginia for example) use 10th percentile of reference, or even the 25th percentile of reference. EPA agreed with WVDEP's use of the 5th percentile of reference because of the high quality and general confidence in West Virginia's reference samples as representative of something closer to natural conditions.

EPA's List Development Process

EPA started by reviewing the biological data assembled by WVDEP for the CWA Section 305(b) portion of its Integrated Report. This included the WVDEP "Decision Database" that was provided with the submission of WVDEP's final 2012 IR. The Decision Database is an Access database that contains relevant water quality monitoring data including but not limited to biological assessment data. The database also includes the various lists of waters that comprise WVDEP's Integrated Report.

Because the basis of EPA's partial disapproval of the 2012 Section 303(d) list is WVDEP's failure to evaluate certain information against West Virginia's narrative water quality criteria, EPA limited itself to data that was readily available to and assembled by WVDEP, but had not been evaluated previously in connection with the 2010 or other past approved Section 303(d) lists. Phrased differently, EPA did not consider samples that were collected and assembled prior to July 1, 2009 because those samples were evaluated by WVDEP as part of West Virginia's 2010 Section 303(d) list, which was approved by EPA on February 8, 2011. It is not EPA's purpose to re-visit evaluations that form part of past approved Section 303(d) lists or to re-visit EPA's approval of those past lists.

EPA then applied the same methodology (WVSCI) used by WVDEP in connection with the 2010 Section 303(d) list to the assembled existing and readily available data described above. WVDEP had retained on the 2012 Section 303(d) list waters previously identified as impaired by the WVSCI methodology on the basis that the WVSCI methodology was "valid" when those waters were identified as impaired. Given the language of SB 562 instructing that WVDEP "may not establish measurements for biologic components of West Virginia's narrative water quality standards that would establish standards less protective than requirements that exist at the time [of SB 562's enactment]" and WVDEP's representation that SB 562 does not change West Virginia's water quality standards, EPA concludes that it can appropriately apply WVSCI to evaluate the existing and readily available data assembled by WVDEP while not pre-empting the methodology development process currently being undertaken by WVDEP pursuant to SB 562.

Using WVDEP's methodology from the 2010 Section 303(d) list, EPA considered waters with a WVSCI score greater than 68 as unimpaired. Waters with a WVSCI score below 60.6 were placed on the list of waters that EPA proposes to add to the Section 303(d) list. With respect to assessing specific data for waters with WVSCI scores from 60.6-68, EPA departed from WVDEP's methodology. As described briefly above, WVDEP historically has considered waters with WVSCI scores in the range 60.6-68 to be in a "gray zone." WVDEP has provided the following explanation in the 2010 Section 303(d) list:

> To address the potential variability associated with a number of factors (collector, microhabitat, subsampling, etc.) a precision estimate was determined by analysis of duplicate biomonitoring data. The precision estimate (7.4 WVSCI points) was subtracted from the impairment threshold to define a "gray zone" of WVSCI scores between 60.6 and 68.0 for which adverse impact to biological integrity is less than certain.

The effective use of limited TMDL development and implementation resources requires the avoidance of impairment misclassifications. Although the true WVSCI impairment threshold is 68.0, DEP identified biological impairment in the Section 303(d) listing process only in response to WVSCI scores less than 60.6, so as to allow the highest degree of confidence in the validity of the listed biological impairments,

In past comments on West Virginia's 2010 Section 303(d) list, EPA had informed WVDEP that its use of a precision estimate to establish the "gray zone" is not statistically supportable. This is because the potential variability for which the gray zone is purported to account already is accounted for by variability in the reference sites.

To determine impairment of a test site compared to a regional reference condition, a single test site score is compared to the reference distribution of index scores (normally composed of single samples from numerous reference sites that are considered natural or near natural). This approach asks whether a single test site (represented by a single observation) is a member of a population of sites (represented by single samples from numerous reference sites). In most cases, the population of reference sites consists of single scores at reference sites and is not an "error free distribution." In other words, the distribution of reference site scores already includes measurement error (due to field and lab methods) and sampling error (the reference site range is only a sample of the reference sites, not the entire population of reference sites). In this case, the procedure is simply to use the best available estimate for the site (single observation, mean, or median), and if that is below the threshold, then the site is rated impaired.

Restating this in statistical terms, the null hypothesis is that a site is a member of the unimpaired (reference) population. Setting a threshold as a percentile of the reference population sets the acceptable significance of the test (α), or the acceptable type 1 error rate (false positive), as the reference percentile. For example, if the criterion is set at the 5th percentile of the reference distribution (as in WV), then any site below that threshold will be rated "impaired". Note that, on average, 5% of reference sites in the distribution of reference sites will also be rated impaired – these are the false positives, or type I error. Setting $\alpha = 0.05$ means that this error rate is acceptable. As noted in footnote 2, the 5th percentile is quite low and assumes that most of the reference sites are in a natural state.

WVDEP's gray zone purports to adjust for the possibility of the effects of measurement error by subtracting confidence intervals (based on the standard deviation of within-site variance calculated from replicate samples collected at several sites) from their initial threshold determination. This approach, however, is only appropriate if the reference range is composed of an "error free" distribution of scores. Estimating the "error-free" distribution of reference site values can be fairly straightforward if multiple measurements are available for each reference site. However, this analysis must be performed *before* thresholds are determined. Any percentile estimated from a raw distribution of single reference site values will include the effects of sampling variability and measurement error, and further adjustment for sampling variability would account for sampling variability twice in the threshold determination (i.e., sampling variability would be double-counted).

Because, as described above, EPA has determined that WVDEP's use of the "gray zone" is statistically unsupported, EPA finds that WVSCI scores from 60.6-68 indicate that waters do not achieve the West Virginia narrative criteria as applied to the aquatic life uses. EPA proposes to add to West Virginia's Section 303(d) list previously unevaluated waters for which there is readily available stream specific data where the WVSCI score is below 68.

Using the foregoing methodology, EPA proposes to add 255 water quality limited segments to West Virginia's Section 303(d) list. A list of the waters that EPA proposes to add to West Virginia's Section 303(d) list is in Enclosure 3. EPA will open a public comment period on these proposed additions to West Virginia's Section 303(d) list and will, if appropriate, revise the list of added waters and pollutants following consideration of any comments received.

EPA is aware that, in its comments on West Virginia's 2010 Section 303(d) list, EPA both had raised concerns regarding how WVDEP was utilizing WVSCI and also had recommended that WVDEP move toward use of a more rigorous genus-level metric called the Genus-Level Index of Most Probable Stream Status (GLIMPSS) as its assessment methodology for the narrative criteria.

GLIMPSS has undergone external peer review and publication in *Environmental Monitoring and Assessment* Volume 185 Number 2 (2012). EPA's view that GLIMPSS is a more rigorous assessment tool that is more consistent with the state of the science remains unchanged. EPA also has commented in the past that, to the extent WVSCI is used, its application should consider the greater number of reference sites that have become available since 2000. That view also remains unchanged.[3] Nevertheless, EPA is mindful that the West Virginia legislature has instructed WVDEP to develop and submit for rulemaking a new assessment methodology for West Virginia's narrative water quality criterion, and that WVDEP is in the process of complying with that instruction. Given that context and WVDEP's representation that WVSCI was a "valid" assessment methodology at the time that prior impairments were identified, EPA has elected for purposes of the 2012 Section 303(d) list to apply WVSCI, WVDEP's past methodology, rather than utilize GLIMPSS, a new methodology not previously utilized by WVDEP. EPA's partial disapproval is based not upon WVDEP's selection of an assessment methodology, but rather upon WVDEP's failure to evaluate certain existing and readily available biological data using any assessment methodology. EPA's decision to utilize WVSCI, a methodology that WVDEP acknowledges was valid in the past, is consistent with the basis of its partial disapproval while avoiding introduction of a new methodology while WVDEP undertakes methodology development pursuant to SB562.[4] If a new methodology is not in place in time for the 2014 Section 303(d) list, EPA will reconsider the range of existing and readily available information, including available assessment

[3] In addition, a review of the WVSCI technical paper reveals that the developers of WVSCI recommended reconsideration of several of the original reference sites to ensure those sites properly represented reference conditions. Several original reference sites apparently were reconsidered and ultimately removed from the set of reference sites; however, the 5th percentile score was not re-calculated to account for this change in the reference set.

[4] It is unclear whether WVDEP's methodology will be ready in time for use in connection with West Virginia's 2014 Section 303(d) list. Today's action should not be construed as a preview of how EPA would view West Virginia's 2014 Section 303(d) list if WVDEP has not completed its assessment methodology or if a future assessment methodology does not include analysis of the benthic macroinvertebrate community using the state of the science.

methodologies, at that time. As WVDEP moves forward, EPA recommends that WVDEP incorporate GLIMPSS into its assessment methodology. EPA stands ready to work with WVDEP as it develops a scientifically rigorous assessment methodology.