Chesapeake Bay Program | Indicator Analysis and Methods Document
Chessie BIBI | Updated 10.16.2018

Indicator Title: Health of Freshwater Streams in the Chesapeake Bay Watershed

Relevant Outcome(s): Stream Health

Relevant Goal(s): Vital Habitats

Location within Framework (i.e., Influencing Factor, Output or Performance): Performance

A. Data Set and Source

(1) Describe the data set. What parameters are measured? What parameters are obtained by calculation? For what purpose(s) are the data used?

Datasets received from multiple federal, state, county, and volunteer monitoring programs are assembled in the "Chessie BIBI database" developed by ICPRB and CBPO Data Center staff. The database has gone through several iterations and now conforms to CEDR standards (Chesapeake Environmental Data Repository). It contains the taxonomic identifications and counts of stream benthic macroinvertebrate samples, and the habitat scores and water quality parameters (pH, conductivity, dissolved oxygen) measured at the time of sample collection. Most macroinvertebrate samples are collected with the kick-net method or variations thereof. Habitat scores are typically made with the Rapid Bioassessment Protocol, or variations thereof (EPA 1999). Some monitoring programs also collected water quality samples that are analyzed in the laboratory, and these data are also included in the database. The database is currently managed jointly by ICPRB and CBPO.

Earlier versions of the database were used to develop and refine the Chesapeake Basin-wide Index of Biotic Integrity for wadeable streams (Chessie BIBI). The Chessie BIBI development is described in Smith et al. 2017. The data are first "normalized" to reduce unintentional bias relating to programmatic differences. The condition of stream macroinvertebrate communities is then evaluated using a consistent approach applied to the data. The approach accommodates geographic differences and scores five or more macroinvertebrate metrics using thresholds derived from metric values in Reference populations. The numeric index scores can be classified into five ratings: Excellent, Good, Fair, Poor, and Very Poor. Ratings of individual sampling stations from across the Chesapeake basin are presently area-weighted (by the proportion of the HUC12 watershed area each station represents) and the proportions summed to obtain an overall proportion of each rating in all sufficiently sampled watersheds.

(2) List the source(s) of the data set, the custodian of the source data, and the relevant contact at the Chesapeake Bay Program.

- Source:

- Anne Arundel County (MD) Watershed, Ecosystem, and Restoration Service
- City of Baltimore (MD) Stream Monitoring Program
- Baltimore County (MD) Watershed Management and Monitoring
- District of Columbia (DC) Stream Monitoring Program
- Delaware (DE) Biological Monitoring Program
- Fairfax County (VA) Stream Quality Assessment Program
- Frederick County (MD) Watershed Management Program
- Howard County (MD) Bio-Monitoring and Assessment Program
- Loudoun County (VA) Stream Quality Assessment Program
- Montgomery County (MD) Stream Protection Program
- Maryland (MD) Biological Stream Survey
- New York (NY) Routine Statewide Monitoring Program
- Pennsylvania (PA) Other Water Quality Assessments
- Pennsylvania (PA) Surface Water Monitoring Programs
- Pennsylvania (PA) USGS
- Pennsylvania (PA) Unassessed Watersheds
- Prince Georges County (MD) Biological Assessment and Monitoring Program
- Susquehanna River Basin Commission (SRBC) Watershed Assessment and Protection - TMDL
- Susquehanna River Basin Commission (SRBC) Watershed Assessment Program
- EPA EMAP Wadeable Stream Assessment (EMAP)
- EPA Wadeable Stream Assessment (WSA)
- EPA National Rivers and Streams Assessment (NRSA)
- EPA Mid-Atlantic Highlands Assessment (MAHA)
- National Forest Service (NFS) Stream Assessment
- US Geological Survey (USGS) National Water Quality Assessment Program
- Virginia (VA) Benthic Monitoring Program
- Virginia (VA) Commonwealth University's Interactive Stream Assessment Resource Program (INSTAR)
- West Virginia (WV) Watershed Monitoring Program

- Custodian: Claire Buchanan (Interstate Commission on the Potomac River Basin) and Michael Mallonee (Interstate Commission on the Potomac River Basin @ CBPO)
- Chesapeake Bay Program Contact (name, email address, phone number): Jennifer Greiner, Jennifer_greiner@fws.gov

(3) Please provide a link to the location of the data set. Are metadata, data-dictionaries and embedded definitions included?
Complete dataset can be obtained from CBPO: http://www.chesapeakebay.net/data/downloads/watershed_wide_benthic_invertebrate_database

B. Temporal Considerations

(4) Data collection date(s): Most programs monitor benthic macroinvertebrates on a rotating cycle, with data from one year representing a two- to 10-year period. Data reflect samples taken between 2006-2011, although different states did this sampling in different years and not throughout the sample period.

(5) Planned update frequency (e.g., annual, biannual, etc.):
- Source Data: TBD
- Indicator: TBD

(6) Date (month and year) next data set is expected to be available for reporting: TBD. Participants at the Chessie BIBI workshop in April 2018 agreed to providing data on a regular basis but no formal process has been set up yet.

C. Spatial Considerations

(7) What is the ideal level of spatial aggregation (e.g., watershed-wide, river basin, state, county, hydrologic unit code)? Given the sampling program limitations, the HUC 12 watershed seems to work best.

(8) Is there geographic (GIS) data associated with this data set? If so, indicate its format (eg. Point, like, polygon). Polygon data (HUC12).

(9) Are there geographic areas that are missing data? If so, list the areas.
Many rural watersheds do not have sampling programs, but partners are working to collect samples in unsampled areas. Therefore , these areas with sparse data will change and hopefully shrink over time.

(10) Please submit any appropriate examples of how this information has been mapped or otherwise portrayed geographically in the past.
Information on state's 303(d) biological impairment assessments were mapped based on stream segment in the 2007 Bay Barometer. B-IBI data was also presented in the 2008 and 2009 Bay Barometer's using different methodology and portrayed as a map showing individual scores at each sampling location. In 2010 this information was similarly mapped and summarized by sub-watersheds. At the April 2018 workshop, various approaches for mapping the Chessie BIBI were illustrated and discussed. Workshop presentations with these examples are located here.

D. Communicating the Data

(11) What is the goal, target, threshold or expected outcome for this indicator? How was it established?

The 2014 Chesapeake Bay Agreement established the following goal and outcome:

> *Continually improve stream health and function throughout the watershed. Improve health and function of ten percent of stream miles above the 2008 baseline for the Chesapeake Bay watershed.*

A year later, the Management Strategy developed by the CBP Stream Health Workgroup identified the Chessie BIBI as the stream health indicator to be used to measure progress achieving this goal. At the "2008 Baseline" workshop convened in April 2018, participants selected as the stream health indicator the family-level version of the index adjusted to the Chesapeake's 12 bioregions and identified 2006 – 2011 as the baseline period against which to measure progress. They agreed to use a predictive tool developed by USGS to estimate and rate stream health in under-represented watersheds. Minutes of the workshop are available here.

In general, Chessie BIBI index scores greater than the 10th percentile of all scores in a bioregion's Reference population are considered stream conditions supportive of Chesapeake Bay TMDLs and other jurisdictional goals. Scores greater than the 10th percentile are rated Fair, Good, or Excellent.

(12) What is the current status in relation to the goal, target, threshold or expected outcome?

Healthy fresh water streams are intrinsically related to a healthy Bay. It can generally be said that a healthy Bay Watershed should have a large proportion of sites rated Fair, Good, or Excellent. The Chesapeake basin can be divided into 2,555 HUC12-bioregion units, or watersheds, totaling 167,000 km^2. Calculations derived from *sampled* HUC12-bioregion watersheds suggest about 55% of the Chesapeake Bay region had streams rated Fair, Good or Excellent during the 2006 – 2011 baseline period. However, over half of HUC12-bioregion watersheds had insufficient data, i.e., less than two data points (Figure 1). Therefore, when including all HUC12-bioregion watersheds to assess the status of the complete watershed, the percentage of streams in the Fair, Good or Excellent categories goes down to 25% due to the high percentage of HUC12-bioregion watersheds with insufficient data (54%). The USGS Leetown Science Center is currently using a 2008 land use layer to recalibrate its Random Forest predictive model of Chessie BIBI ratings for the 2006 – 2011 baseline period. When completed, model results will be combined with sample data to complete Figure 1. A manuscript describing the Random Forest model and its results for an earlier time period was recently accepted for publication by the Journal of Freshwater Science (Maloney et al. in press).

(13) Has a new goal, target, threshold or expected outcome been established since the last reporting period? Why?

Yes. The 2014 Chesapeake Bay Agreement established a new goal for stream health (see above). State data providers agreed at the April 2018 workshop to the period of 2006-2011 as the baseline because it provides the good areal coverage. The bioregion, family-level version of the Chessie BIBI index was selected as the stream health indicator.

(14) Has the methodology of data collection or analysis changed since the last reporting period? How? Why?

Yes. The Chessie BIBI database was updated and the index refined in 2017. Please refer to Smith et al. 2017.

(15) What is the long-term data trend (since the start of data collection)?

Although the database includes samples from the 1990's through 2015, only a few stations are sampled more than once because most monitoring programs apply a random, stratified sampling design. This design makes it difficult to confidently detect trends. Also, the state programs typically monitor benthic macroinvertebrates on a rotating cycle, with data from one year used to represent a 2- to 10-year period. The potential for identifying trends at the few repeatedly sampled sites (e.g., sentinel sites) will be explored in the future.

(16) What change(s) does the most recent data show compared to the last reporting period? To what do you attribute the change? Is this actual cause or educated speculation? No changes can be determined at this point.

(17) What is the key story told by this indicator?

Stream health in the Chesapeake basin as measured by the Chessie BIBI ranges from Very Poor to Excellent. The basin holds 2,555 unique "HUC12-bioregion" units, which represent small (~89 km^2, or 34.4 mi^2) watersheds with homogeneous natural features. A little more than half of these watersheds were sufficiently sampled in 2006 – 2011, the baseline period, to estimate stream conditions. In the near future, results of a USGS predictive model (Maloney et al. in press) will be added to the monitoring results and the combination will give us a picture of stream health for the entire Chesapeake basin. The monitoring data already give us a preview of what to expect. We will likely find roughly 19% of streams in Very Poor condition and 26% in Poor condition. Conversely, about 55% should be in Fair, Good, or Excellent.

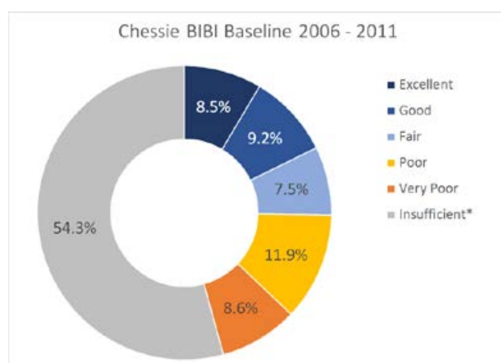

Figure 1. Stream ratings in Chesapeake basin HUC12-bioregion watersheds during the 2006 – 2011 baseline period. A Random Forest predictive model developed by USGS uses landscape variables to predict Chessie BIBI ratings in watersheds with few or no sampling stations. These model results are pending and

There is a clear link between the Chessie BIBI scores and land-based activities and land uses in the individual watersheds. The highest stream indexes tend to occur in heavily forested areas with low levels of pollution and stable in-stream and streamside habitats. The poorest stream indexes occur in highly urbanized watersheds such as those in the Baltimore-Washington D.C. metropolitan region. See Smith et al. 2017 for further discussion about this.

E. Adaptive Management

(18) What factors influence progress toward the goal, target, threshold or expected outcome?

- **Ecological stressors and factors**:
  - physical, chemical and biological factors that impair or limit stream health recovery
- **Within the stream channel and floodplain factors**
  - Excessive sediment and nutrients in-stream from unstable stream banks and legacy sediments in the floodplain
  - Limited nutrient and organic processing-instream
  - Alteration in channel form and function resulting in instability and disequilibrium affecting diversity and quality of habitat
  - Concentrated flows and reduction in baseflows
  - Piped and channelize streams
  - Removal/Loss of forested riparian areas and the benefits provided by shading
- **Watershed based factors**
  - Impervious cover and increases in stormwater runoff
  - Significant changes in watershed hydrology (time of concentration) related to overland flow impacted by road drainage, ag land drainage, driveways, stormwater collection systems, etc.
  - Flow alteration and flashy hydrology
  - Excessive nutrient loading to streams from excess untreated runoff (agricultural and urban) from the upland areas in the watershed and groundwater
  - Implementation of stormwater management controls (e.g. BMPs)
  - Leaky wastewater infrastructure
  - Toxicity of effluent from resource extraction activities (i.e., acid mine drainage, fracking)
  - Road de-icing practices (salt)
  - Thermal impacts
  - Invasive species
  - Endocrine disrupting chemicals
- **Policy and administrative factors**
  - Review and approval of stream restoration projects for WIP implementation

- Lack of common watershed, stressor and stream assessment and restoration guidelines
- Integration of water quality and living resource goals during WIP stream restoration
- MS4 permits focus on water quality
- Adequate financial resources to support local implementation efforts
- Adequate extension infrastructure to communicate newest research and technical guidance to jurisdictions
- In very urban area, the availability of land to retrofit and implement upland BMPs

- **Scientific knowledge and application of research**

- Stressor identification and prioritization procedures
- Functional metrics that correlate with priority stressors identified for measurement
  - Robust stream restoration monitoring to evaluate the potential functional lift or improvement in stream functions from BMP implementation
  - Possible lag times that affect the ability to evaluate the effect of upland BMP on stream health
  - Research needed to guide the selection of achievable reference conditions/design approaches based on watershed and stream functions to include an urban reference continuum
  - Insufficient data to develop Bay-wide fish-based indicator to complement macro-invertebrate indicator (Chessie BIBI)
  - Lengthy timeframe for adjusting BMP credit or recognizing new BMPs
  - Limitations of the applicability of the Chessie BIBI (and other similar ecological data) to streams where restoration work is being conducted on an annual basis.
  - Identify nutrient hotspot in stream valley where erodible geologic materials and soils contain excess nutrients
  - Additional research to refine nutrient credits for stream restoration projects as supported by the Expert Panel recommendations on Individual Stream Restoration Projects to include for example bioavailability of nutrients.

- **Partner coordination**
- **Funding**

(19) What are the current gaps in existing management efforts?

- **Information & Data**
  - Benthic macroinvertebrate data from enough streams with enough frequency to track progress over time. Chessie BIBI provides limited capacity for annual tracking, trend analysis less than 5-7 yrs.

- Bay-wide and stream metrics other than biological indices, such as the Chessie BIBI, to assess physical and chemical health and functions of streams
- Update or review of methods to define reference conditions or endpoints for streams
- Sufficiency of data to demonstrate effectiveness of stream restoration practices
- Sufficiency of data to demonstrate restoration of stream processes following installation of upland watershed BMPs. Sufficiency
- Cumulative effects and interactions between stressors
- Completion of stressor analysis for additional watersheds

- **Regulatory & Programmatic**
  - Project design process for stream restoration that can measure change in stream functions and/project success based on a project goals and objectives. Specific to the Bay TMDL, a design process for restoration projects to reduce nutrient and sediments loads delivered downstream while at the same time ensuring optimal habitat conditions are restored.
  - Information needs to support innovative, effective design approaches to identify restoration potential and success for different land uses, stream types, and current and future site constraints, causes of impairment/stressors (e.g. legacy sediment, contaminants in water and sediment, runoff volume and velocity).
  - Identification of local and watershed priority stressors that affect local steam health and management actions to results in associated function lift
  - Collaboration with the Healthy Watersheds GIT to identify marginal streams and various definitions for stream health (i.e., Chessie BIBI to individual state metrics).

- **Prioritization**
  - Targeting procedures for cost-effective restoration actions and design approaches that will achieve both water quality and biological functional improvement. WIPs provide a level of analysis on the type and mix of projects to meet load reductions and associated costs. The process to identify the projects varies by jurisdiction along with cost estimates.
  - Investments in research to improve the body of knowledge surrounding restoration techniques and net benefit to stream and watershed health.

(20) What are the current overlaps in existing management efforts?

- Development of 2008 baseline for the Chessie BIBI
- Pooled monitoring approach to stream restoration projects
- Ongoing monitoring efforts

- EPA National Rivers and Stream Assessment: The EPA NRSA sampled between 90 and 100 randomly selected sites in the Chesapeake Bay watershed. These sites have benthic invertebrate, fish, periphyton, water quality and habitat data. The EPA NRSA surveys are conducted every five years, including 2008/2009, 2013/2014, with the next one scheduled for 2018/2019.
- State 305b (Integrated Report) Reports (e.g. see http://www.mde.state.md.us/programs/Water/TMDL/Pages/Programs/WaterPrograms/tmdl/bsid_studies.aspx. Accessed Jan 15, 2015.)
- Tidal Network monitoring sites
- Non-Tidal Network monitoring sites
- National Park Service has five inventory and monitoring networks operating within the Chesapeake Bay (provided by Marian Norris);
- Maryland Biological Stream Survey (MBSS): Sampled 252 randomly selected sites during 2007- 2009 to characterize Maryland's ecological condition. Round Four is scheduled for 2014-2018.
- County monitoring programs

- MS4 permits that have incorporated Bay TMDL goals
- WIP implementation of BMPs
- Development of Chesapeake Bay BMP verification

(21) According to the management strategy written for the outcome associated with this indicator, how will we (a) assess our performance in making progress toward the goal, target, threshold or expected outcome, and (b) ensure the adaptive management of our work?

a) **Assessing Progress:**
The Chesapeake Bay Program annual progress reports on BMP implementation, specifically BMPs identified to impact critical steam functions (e.g. stream restoration, stream fencing and forest buffers) can be used to estimate the project nutrient and sediment load reductions expected from practice implementation. Assessing progress should also focus on remediation of principal stressors and stream reach functional lift based on stream restoration project goals and objectives. While projects undertaken for Bay TMDL purpose focus on nutrient and sediment reduction, information available from completed stressor identification analysis should be taken into consideration to improve stream health, as well as to instream and floodplain habitats. While we want to encourage the remediation of priority stressors to improve stream health, or maximize functional lift for all stream restoration projects, we cannot require it given site specific constraints and the ability to address watershed stressors affecting the health of the stream. It is important that a progress reporting process be developed that can be used to assess progress up through biology but allow for lower levels (i.e., stability) of report only.

b) **Adaptive Management**

A process that communicates the current state of the science on the influence of efforts to improve stream health now, with periodic updates, would help ensure the most successful practices are implemented and the most benefits possible for stream health are achieved.

The current Management Strategies include:

1. Identify an appropriate suite of metrics to measure the multiple facets of stream health to complement the baywide Chessie BIBI.
2. Provision of adequate funding and technical resources to support functional lift in stream restoration projects, in addition to nutrient and sediment reduction.
3. Ongoing coordination with stream restoration stakeholders (e.g., state and federal stream and wetland permitting authorities, natural resource agencies, local governments, non-profit organizations, stream restoration designers, researchers) needs to be improved to identify and remove barriers providing a clearly defined path to expedite the submittal and review of permit applications, whether the proposed activity is for marginal streams, impaired streams, or for credit in the Bay TMDL.
4. Develop and promote holistic stream restoration design guidelines that identifies the level of degradation and improvement of stream functions and key stressors/factors limiting potential uplift
5. Work with Chesapeake Bay Partners to include the Enhancing Partnering, Leadership and Management GIT, to enhance the capacity of local governments, organizations and landowners of beneficial stream restoration and maintenance practices

F. Analysis and Interpretation

Please provide appropriate references and location(s) of documentation if hard to find.

(22) What method is used to transform raw data into the information presented in this indicator? Please cite methods and/or modeling programs.

Benthic macroinvertebrates are collected in the field by multiple agencies and sorted, identified, and counted in various laboratories. In the future, the agencies collecting the data will submit data directly to CBPO, where it will be QA'ed and incorporated into the Chessie BIBI database. Community biometrics are calculated and scored with computer programs (developed by ICPRB), and the metric scores are averaged to obtain the Chessie BIBI index score. To enhance interpretation and usefulness of the results, index scores are grouped by the smallest feasible watershed size. At this time, the smallest feasible watershed size is a unit defined by unique combinations of the 1,978 USGS Hydrologic Unit Category (HUC) 12 and twelve bioregions in the Chesapeake basin which yield a total of 2,555 HUC12-bioregion. If two or more stations are located in a HUC12-bioregion watershed, the stations' ratings are weighted by equal proportions of that HUC12-bioregion watershed's area, and the area-weighted results for all watersheds are rolled up to obtain Chessie BIBI rating proportions for the entire Chesapeake basin (e.g. Figure 1). Details are available in Smith et al. 2017.

(23) Is the method used to transform raw data into the information presented in this indicator accepted as scientifically sound? If not, what are its limitations?

The Reference-based approach used to develop the index scored biological metrics of species habit, tolerance, richness, and composition according to the distributions of each metric's values in Reference populations. The Reference-based approach is a well-established method that has been used to develop numerous terrestrial, marine, and freshwater indices of biotic integrity. The Chessie BIBI is unusually in that it combines data from multiple sources. Another example of an index developed from multiple sources is the Potomac River Basin index (Astin 2006, 2007). The Chessie BIBI has undergone extensive technical and peer review by Chesapeake Bay Partnership members of two technical advisory groups of biologists as well as participants of the April 2018 "2008 Baseline" workshop. The indicator has also been reviewed by the Non-Tidal Water Quality Workgroup, the Stream Health Workgroup, and the Scientific and Technical Analysis and Reporting team.

(24) How well does the indicator represent the environmental condition being assessed?

The ability of the bioregion, family-level version of the Chessie BIBI index to correctly identify independently identified Reference and Degraded conditions varies depending on bioregion. Classification efficiency (CE) ranged from 70.4% to 90.0% (Smith et al. 2017). The Random Forest model that predicts Chessie BIBI ratings from landscape variables has similar CE's (Maloney et al. in press).

(25) Are there established reference points, thresholds, ranges or values for this indicator that unambiguously reflect the desired state of the environment?

Fair, Good, and Excellent ratings represent desirable Chessie BIBI scores and they encompass 90% of index scores found in undisturbed, Reference populations. Desirable Chessie BIBI index scores are roughly comparable to the desirable index scores for estuarine phytoplankton and benthic communities which are "greater than or equal to 3 on a scale of 1 – 5."

The Chesapeake Bay Program (CBP) has adopted the approach of presenting indicators as a "percent of restoration goal achieved." Currently, there is no percentage goal for streams with ratings of Fair, Good, or Excellent. The Strategy for Protecting and Restoring the Chesapeake Bay Watershed (2010) which was drafted in response to Executive Order 13508 had a goal to

> *"improve the health of streams so that 70 percent of sampled streams throughout the Chesapeake watershed rate fair, good or excellent, as measured by the Index of Biotic Integrity, by 2025."*

However, this goal was superseded by the ambiguous goal in the 2014 Bay Agreement which calls for *"continually improve stream health and function throughout the watershed. Improve health and function of ten percent of stream miles above the 2008 baseline for the Chesapeake Bay watershed."*

(26) How far can the data be extrapolated? Have appropriate statistical methods been used to generalize or portray data beyond the time or spatial locations where measurements were made (e.g., statistical survey inference, no generalization is possible)?

This has not been thoroughly investigated. Preliminary data analyses show that BIBI ratings for HUC12-bioregion watersheds with only two randomly located sampling stations closely resemble the ratings for watersheds with five or more randomly located stations. This suggests two sampling stations may adequately represent the HUC12-bioregion watershed.

G. Quality

Please provide appropriate references and location(s) of documentation if hard to find.

(27) Were the data collected and processed according to a U.S. Environmental Protection Agency-approved Quality Assurance Project Plan? If so, please provide a link to the QAPP and indicate when the plan was last reviewed and approved. **If not, please complete questions 29-31.**

FGDC compliant metadata for most of the programs that provided stream data circa 2010 are located at www.chesapeakebay.net. Many of the monitoring programs use CWA106 or other EPA funding and should thus be required by EPA to have QAPPs.

(28) *If applicable:* Are the sampling, analytical and data processing procedures accepted as scientifically and technically valid?

Yes. Most field sampling procedures are modifications of the Environmental Protection Agency's Rapid Bioassessment Protocols for Use in Stream and Wadeable Rivers (Plafkin et al. 1989, Barbour et al. 1999). Laboratory methods for identifying and counting benthic macroinvertebrates are similar – the largest difference being the number of individuals counted per sample. Data submitted by the data providers are incorporated into the database structure and format established by the CBP Chesapeake Environmental Data Repository (CEDR), which forces consistency and quality assures the data.

(29) *If applicable:* What documentation describes the sampling and analytical procedures used?

See metadata at:
https://www.chesapeakebay.net/what/downloads/watershed_wide_benthic_invertebrate_database

(30) *If applicable:* To what extent are procedures for quality assurance and quality control of the data documented and accessible?

The current document can be found here:
https://www.chesapeakebay.net/what/downloads/watershed_wide_benthic_invertebrate_database

When the Chessie BiBi update is finalized, a new QAQC document will be released.

(31) Are descriptions of the study design clear, complete and sufficient to enable the study to be reproduced?

Yes.

(32) Were the sampling, analytical and data processing procedures performed consistently throughout the data record?

No. There are slight differences in sampling methodology and laboratory analysis between sampling agencies. These are minimized to the greatest extent possible when the raw data are prepared for calculating the Chessie BIBI.

(33) If data sets from two or more sources have been merged, are the sampling designs, methods and results comparable? If not, what are the limitations?

Sampling designs are not 100% comparable, however calculation of the Chessie BIBI standardizes most of the data so it can be compared across jurisdictions (see Smith et al. 2017, Buchanan et al. 2011, Astin 2006, 2007 for discussions of this topic).

(34) Are levels of uncertainty available for the indicator and/or the underlying data set? If so, do the uncertainty and variability impact the conclusions drawn from the data or the utility of the indicator?

Classification efficiencies and jackknife validations of the BIBI error were done when the index was first developed (Buchanan et al. 2011) and when it was refined (Smith et al. 2017). Therefore, questions of uncertainty and variability in the index have been resolved to a large extent during indicator development. More analysis is needed to determine uncertainty in model predictions of the BIBI (Maloney et al. in press) and in trend analyses.

(35) For chemical data reporting: How are data below the MDL reported (i.e., reported as 0, censored, or as < MDL)? If parameter substitutions are made (e.g., using orthophosphate instead of total phosphorus), how are data normalized? How does this impact the indicator?

No water quality data are used to calculate the Chessie BIBI.

(36) Are there noteworthy limitations or gaps in the data record?

See Smith et al. 2017 and minutes of the “2008 Baseline” workshop.

H. Additional Information (*Optional)*

(37) Please provide any further information you believe is necessary to aid in communication and prevent any potential misrepresentation of this indicator.

References

Astin, L.E. 2006. Data synthesis and bioindicator development for nontidal streams in the interstate Potomac River basin, *USA.* Ecological Indicators 6: 664-685.

Astin, L. E. 2007. Developing biological indicators from diverse data: The Potomac Basin-wide Index of Benthic Integrity (B-IBI). Ecological Indicators 7: 895-908.

Barbour, M. T., et al. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. Second Edition. EPA 841-B-99-002.

Buchanan, C., K. Foreman, J. Johnson, and A. Griggs. 2011. Development of a Basin-wide Benthic Index of Biotic Integrity for Non-Tidal Streams and Wadeable Rivers in the Chesapeake Bay Watershed: Final Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. ICPRB Report 11-1. Prepared for the US Environmental Protection Agency, Chesapeake Bay Program.

Federal Leadership Committee for the Chesapeake Bay. Strategy for Protecting and Restoring the Chesapeake Bay Watershed. Executive Order 13508. May 12, 2010.

Foreman, K., Buchanan, C., Nagel, A., 2008. Development of ecosystem health indexes for non-tidal wadeable streams and rivers in the Chesapeake Bay basin. Report to the Chesapeake Bay Program Non-Tidal Water Quality Workgroup. December 5, 2008.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6:21-27.

Maloney, K. O., Z. M. Smith, C. Buchanan, A. Nagel, and J. A. Young. (in press) Predicting Biological Conditions for Small Headwater Streams in the Chesapeake Bay Watershed. Journal of Freshwater Science.

Plafkin, J. L. et al. 1989. Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish. EPA/440/4-98/001. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D. C.

Smith, Z., C. Buchanan, and A. Nagel. 2017. Refinement of the Basin-Wide Index of Biotic Integrity for Non-Tidal Streams and Wadeable Rivers in the Chesapeake Bay Watershed. Interstate Commission on the Potomac River Basin, ICPRB report 17-2. Available online at www.potomacriver.org.

US Environmental Protection Agency. 1997. Field and laboratory methods for macroinvertebrate and habitat assessment of low gradient nontidal streams. Mid-

Atlantic Coastal Streams Workgroup, Environmental Services Division, Region 3, Wheeling, WV; 23 pages with appendices.

Table 1. Benthic Macroinvertebrate sampling details for selected data sources in the Chesapeake Bay B-IB

Organization	Methodology	Sampling Details	Listing Criteria
Virginia	Rapid Bioassessment Protocol (EPA RBP II procedures)	Reference and target sampling 2 times/year, random once/year	Stream Condition Index (SCI biocriteria, SCI<=60 impaired
West Virginia	Rapid Bioassessment Protocol (EPA RBP II procedures) Wadeable riffles and runs	Random and targeted sampling on a 5 yr sampling rotation	Stream Condition Index (SCI biocriteria, SCI<60.6 impaire
Delaware	Mid-Atlantic Coastal Streams Workgroup Protocol, modified version of EPA RBP	Sampling in the fall, ~50 sites/year, presently targeted sampling driven by TMDL investigations, some targeted sites some random sites	Biological Index (BI) used fo <=66 impaired
Pennsylvania	Rapid Bioasseessment Protocol (EPA RBP III procedures)	Probabalistic is rotating watersheds. Targeted for TMDLs and threatened areas.	Index of Biotic Integrity (IBI)
Maryland	Maryland Biological Stream Survey (MBSS) protocol. 1st-4th order streams.	Random stratified sampling design. Probablistic sampling via 8 digit HUC sampling unit, 250 sites. Targeted sites sampled as well.	The IBI is scored on a 1-5 sc greater passing. Maryland es percentage of degraded stre each 8-digit watershed (equi All stations in each watershe reference conditions. If degra compose significantly greate the watershed is impaired.
SRBC	Project-dependent: RBPIII in PA and Md, mostly; RBPIII and NYS biomonitoring protocol in NY	Project-dependent: rotating basin surveys every 6 years; yearly sampling at selected sites; most sampling during baseflow conditions (summer/early fall)	<53% of reference is conside SRBC only lists streams as " insufficient data" - not as imp
Fairfax County	Mid-Atlantic Coastal Streams Workgroup Protocol, EPA RBP	Random stratified sampling design. Probablistic sampling of 40 randomly selected sites per year. Thirteen Reference sites are sampled each year, eleven piedmont sites and 2 coastal plain sites. Several trend sites are sampled on a rotating basis.	Sites are scored using an Ind Integrity (IBI) and rated on a broken down into Very Poor, and Excellent rating categori sites are combined to create Stream Quality Index (SQI).
Montgomery County	Modified version of Maryland Biological Stream Survey (MBSS) protocol	Both targeted and probability-based sampling, depending on management need. Sites selected in one of three ways using geographic and stream order stratification: 1) Reaches are randomly selected and sites are randomly chosen on the reach. 2) reaches are targeted and sites are randomly chosen on the reach, or 3) both reaches and sites are targeted. Baseline sites are revisited on a 5-year watershed rotational basis.	BIBI scored on a 8-40 scale. in channery silt loam ecoregi loam), 26-35 is Good, 17-25 Poor. All stations are scored reference stream conditions.
Prince George's County	Mid-Atlantic Coastal Streams Workgroup Protocol, modified version of EPA RBP (same as MBSS).	Random stratified sampling design. Probablistic sampling via the County's 41 designated watersheds. 5-year cycle (1999-2004), approximately 255 sites. Also, some targeted sites each year and special studies.	Use the MD DNR MBSS IBI. date of database the IBI is S or Southerland et al. 2005.
USFS	Rapid Bioassessment Protocol (EPA RBP II procedures)	Project-dependent. Reference, inventory and/or target sampling 1 time a year	MAIS (macroinvertebrate ag streams) developed by Vosh from 0-18, with 18 being the
New York*	NYS Biomonitoring Rapid Bioassessment Protocol	RIBS (Rotating Intensive Basin Studies) network monitoring is conducted in 2-4 watersheds/yr, 5 yr sampling rotation	Biological Assessment Profil combo of metrics and catego impaired on a scale of 0 to 1 severely impacted.

***Need updated information, however data not included in 2008 indicator**

Table 1 continue

Organization	Type of sampling	What time of year sample	Sampling equipment	Subsample size	Precisio method
Virginia	Reference, random, and targeted	Spring (April-May) and Fall (September-November)	D-Frame Net 500um	110	10% of s QA/QC'c
West Virginia	Random and targeted	April 15 thru Oct 1	500 um mesh, 0.5 m wide kick net	200 org	yes
Delaware	mixed: Targeted, random, reference	Fall, October - November	d-net, 20 1-meter long jabs (≈6 m^2)	100	No calcu estimate data is a
Pennsylvania	Random and Targeted	Freestone: Separate benchmarks for Nov to May and July to September, generally avoid October and June. Limestone and Pool/Glide: February to May only, try to avoid blackfly hatches in pool/glide	D-frame nets	200 organisms except limestone which is 300 due to lower diversity	All precis seasona to 100 s
Maryland	Random and Targeted	March and April	D-net, 540 microns	100	Duplicat sites sar
SRBC	Targeted	summer/fall	project dependent - d-frame or kick net	200	
Fairfax County	Reference, Random and Targeted	March/April	D-frame nets	200	10% of s
Montgomery County	Reference, random, and targeted	March 15 to April 30	D-net, 540 microns	100	Some du 10% of s per year
Prince George's County	Random and Targeted	Spring (Mar and Apr). Some early data in Feb.	D-net, 500-micron	100	Duplicat probabili
USFS	Reference, random, and targeted	Spring (March-May)	Kick Net 500um	200	10% of s QA/QC'c
New York*					

***Need updated information, however data not included in 2008 indicator**