

May 2026

***Independent Peer Review of the Benchmark Quantitative Fishery
Stock Assessment for Blue Crab in the Chesapeake Bay (2023-
2025)***

Annapolis, Maryland

April 13-16, 2026

Dr Simon de Lestang

**Department of Fisheries (Western Australia)
Western Australian Fisheries and Marine Research Laboratories
PO Box 20, North Beach, WA 6920, Australia**

Representing the Center of Independent Experts

Executive Summary

The NOAA Chesapeake Bay Office requested an independent peer review for the 2025 “Benchmark Quantitative Fishery Stock Assessment for Blue Crab in the Chesapeake Bay” (blue crab stock assessment). This updated stock assessment will provide a critical science product at a time of partnership-wide strategic planning for the Chesapeake Bay Program and to inform science-based management of the Chesapeake Bay blue crab fishery.

The term review is applied broadly, as the review panel (RP) was able to request additional analyses, error corrections and additional runs of the assessment models provided by the assessment panel.

The Center for Independent Experts (CIE) organized three reviewers to conduct the peer review. On the 24 March 2026, the Ecosystem Science and Synthesis Manager, NOAA Chesapeake Bay Office made available a summary draft stock assessment document, 13 additional working papers and background information via a google drive folder. Most of the background information related directly to the data collection, manipulation and modelling for the stock assessment. The background information was extensive and included all material required to conduct the review (see Appendix 1 for complete list). The CIE RP participated in a hybrid workshop held in Annapolis, Maryland. The meeting was held in Conference Rooms 1 and 2, 200 Harry S. Truman Parkway, Annapolis, MD, 21401, and streamed live on Microsoft Teams. One member of the RP attended the meetings in person while the two others attended remotely. The meeting began at 12 pm on Monday, April 13th and ended at 4:30 pm on Wednesday, April 15th. Live streaming of the workshop was recorded.

The assessment team presented key aspects of their research on the first day, with presentations structured around their terms of reference (TORs) research topics. Copies of the presentations were provided to the review panel (RP) via a google documents folder (https://drive.google.com/drive/folders/1UkSin2UYwSKB1pQCYk_GzZeSwesa4d5ca). The RP were able to discuss and query any aspects of the presentation as it was being delivered. The assessment team continued to present their research on the second day in addition to providing some analytical feedback from questions posed by the RP on day one. The last day of the review consisted of discussions concerning the assessment model, assessment outcomes and the results of additional analysis requested by the RP during days one and two.

The RP found that the process conducted by the assessment team to be extremely thorough and expansive, conducted in a very structured and outcome-driven format in keeping with best practice. By building on initial workshops structured around the TOR, and including experts from various stakeholder groups, the assessment team ensured the development of a comprehensive and valuable library of working papers on which the stock assessment could be based. A clear and concise draft stock assessment report, containing key aspects of the working papers, a description of the modelling framework, and assessment outcomes, was then compiled.

The data collated for the assessment was examined and refined to produce commercial landings and fishery independent survey data sets that could be truncated by size class bin (1 cm) on a monthly time scale encompassing the entirety of Chesapeake Bay. An environmental dataset was developed for water temperature. The types of robust datasets identified enabled the assessment team to use a data-rich size and sex specific integrated assessment (Sex and Length Structured Assessment Model: SLAM). The assessment team

utilised multiple modelling techniques (e.g., VAST and GAMLSS) to develop spatially and temporally standardized indices of abundance for SLAM. The RP considered these choices were appropriate for use in this assessment.

The assessment team conducted a thorough review of blue crab biology and life history, producing several valuable new analyses including sex-specific length-weight relationships, female maturation growth, maturation ogives for both sexes, revised fecundity estimates, and year-specific overwinter mortality linked to winter severity. However, the extent and scale of interannual variation in life history parameters was not fully addressed. While overwinter mortality and length-weight relationships were examined for temporal change, growth, non-winter natural mortality, size-at-recruitment, timing of recruitment, size-at-maturity, and fecundity were all assumed to be stationary across the 30-year assessment period without formal testing. This omission is surprising given the report's own documentation of temperature-dependent growth and changing environmental conditions over the assessment time frame.

The treatment of fishery-independent data is a substantial methodological advance, with eight surveys systematically evaluated and indices developed using appropriate spatiotemporal (VAST) and GAM-based frameworks at a monthly resolution. However, the assessment did not examine the spatial coherence of the survey data sources, despite surveys covering markedly different portions of the Bay. Size composition data should be compared across surveys to determine whether they are sampling similar population segments, and consideration given to breaking spatially extensive surveys like the WDS into sub-regions. A mismatch also exists between the careful standardisation applied to abundance indices and the unstandardised size composition data entering the model. Environmental and abiotic factors were not considered during index standardisation, which was specifically identified in the associated TOR. The report would benefit from showing the incremental impact of each standardisation factor on the resulting indices. The recruit censoring approach (size cutoffs of 35 mm and 65 mm) lacked formal sensitivity testing.

The assessment team's treatment of fishery-dependent data is thorough, with detailed descriptions of commercial reporting across all three jurisdictions, a well-justified fleet structure in the model, and a sound decision to exclude commercial CPUE given its unreliability and the availability of multiple fishery-independent surveys. However, the conversion of catch from numbers to weight uses fixed conversion factors assumed constant over the entire time series, which implicitly assumes stationary size structure in the commercial catch despite evidence of long-term changes in mean crab size. The treatment of discards accounts for crabs observed dead in commercial gear and fishing-induced mortality is accounted for in the trawl fishery. There was, however, no handling mortality included within the modelling for sublegal crabs that are caught in pots and returned.

The SLAM model represents a major advance, moving to a monthly, sex- and length-structured framework implemented in RTMB with clear, well-structured code. However, several structural and methodological concerns warrant attention. The estimation of separate fishing mortality for every month \times sex \times fleet combination (contributing to 2,639 total parameters) risks absorbing sex ratio discrepancies into F rather than into selectivity or population structure; a hybrid-F approach could achieve the same outcome with far fewer parameters. Survey selectivity splines for the MD and VIMS trawl surveys

show similar dome-shaped patterns in spring despite the surveys operating in different parts of the Bay. If these reflect migration-driven crab availability, the patterns should be opposing, not parallel, suggesting the selectivity splines may be compensating for structural model misspecification. The model lacks process error, assumes constant recruit size distributions, does not incorporate density-dependent mortality (cannibalism), and uses quasi-equilibrium initial conditions that produce large early residuals. The monthly time step may be unnecessarily complex during winter months when no growth, reproduction, or fishing occurs. Spatial heterogeneity in the Bay is substantial. WDS size-frequency data separated by state reveals markedly different population structures, yet the model is not spatially aggregated. Treating spatial sub-regions as separate fleets would better represent demographic differences. Data weighting relies on ad hoc effective sample sizes and fixed log-scale standard deviations rather than iterative methods such as Francis weighting, and bias ramping on recruitment deviations was not considered. Uncertainty was characterised through asymptotic standard errors, sensitivity analyses of fixed parameters, and jitter analyses. The jitter analysis demonstrated reasonable convergence stability, but the most significant finding from the sensitivity analysis is the influence of the growth parameters. This is particularly concerning since growth was calibrated internally through iterative manual adjustment rather than derived from an independent data source. If growth has changed over time, as the biological evidence suggests, the impact on assessment results could be considerable. A retrospective analysis with Mohn's rho was presented during the review meetings but is absent from the report and should be included as a standard diagnostic. The sensitivity analysis was limited to one-at-a-time perturbations, which does not capture interactions among assumptions. Additional uncertainty arising from spatial variation in size composition across the Bay, the quasi-equilibrium initial conditions, and the assumed data weighting structure remain underexplored.

The comparison between the previous 2011 model (SSCMSA) and the new SLAM model is logically structured and reveals strong correlations between design-based and model-based survey indices, providing reassurance that the shift in index standardisation methods has not fundamentally altered the input signals. However, SLAM estimates abundance approximately twice as high as the previous model, primarily due to different assumptions about WDS selectivity and the visibility of age-0 recruits. This scale difference is consequential and would benefit from independent ground-truthing using external abundance estimates (e.g., tag-recapture studies, surplus production models [JABBA/SPiCT], or depletion-based approaches). The progression from ADMB to STAN to RTMB across the comparison is an unusual frequentist–Bayesian–frequentist path, and the influence of prior specifications in STAN relative to penalty structures in the other platforms should be acknowledged. Despite the difference in absolute scale, both models suggest broadly similar stock sustainability relative to reference points, but if fishing mortality appears sustainable while recruitment is declining, the management framework risks maintaining pressure during periods of low productivity, underscoring the value of incorporating a recruitment index into management advice.

The adoption of SPR-based reference points is sound and avoids the estimation difficulties with stock-recruitment relationships that plagued the 2011 assessment. The choice of F40% as the threshold is conservative for crustacean fisheries but is justified given unexplained recruitment declines and recent collapses of other crab stocks under

environmental change; however, formal evaluation through MSE would strengthen this justification. Although the assessment did not fit a stock-recruitment relationship, MSY-based reference points could still be estimated through an iterative equilibrium yield approach using the full SLAM model dynamics with a representative constant recruitment. This would complement the SPR-based reference points and directly address the management question of maximum sustainable yield. Stock status results show fishing mortality well below target levels for both sexes and female abundance near the target, but male abundance has been persistently below the threshold and three-year average recruitment fell below the 25th percentile threshold in 2022 and 2023 for the first time. The recruitment reference point should be formally incorporated into management decision-making to enable proactive responses. Presenting stock status in terms of exploitation rate rather than instantaneous F , calculated using mid-season biomass, would improve communication with managers and stakeholders.

The assessment conducted several valuable ecosystem analyses including blue catfish predation effects, spatiotemporal temperature modelling, and overwinter mortality estimation, but these were largely standalone investigations external to the assessment framework. Environmental and climate drivers could and should have been examined within the VAST and GAMLSS index standardisation models, where covariates such as temperature, salinity, and dissolved oxygen can be fitted directly as catchability or density covariates. The inability to fit temperature-dependent growth could be addressed pragmatically by treating the current fixed monthly Gompertz b parameters as priors and allowing them to vary between years as random effects, providing a de facto mechanism for capturing interannual growth variation informed by the data. Important ecosystem interactions including cannibalism, habitat change, and predation by expanding species such as red drum were not examined or could not be completed within the assessment timeframe. The dominant drivers of the recent recruitment decline remain unidentified, which is the single most important unresolved question given that fishing mortality is low, female abundance is adequate, yet recruitment has fallen below the threshold. The RP agreed that the assessment team at least partially achieved their TOR and that this assessment was a significant improvement on the previous 2011 benchmark assessment. The current assessment represents the best science available and is appropriate for developing management advice.

Background

The NOAA Fisheries is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NOAA Fisheries science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards¹.

The NOAA Chesapeake Bay Office requests an independent peer review for the 2025 "Benchmark Quantitative Fishery Stock Assessment for Blue Crab in the Chesapeake Bay" (blue crab stock assessment). This updated stock assessment will provide a critical science product at a time of partnership-wide strategic planning for the Chesapeake Bay Program and to inform science-based management of the Chesapeake Bay blue crab fishery.

In the Chesapeake Bay, the blue crab (*Callinectes sapidus*) supports economically and culturally important commercial fisheries. Diverse gears are used to harvest blue crabs regionally and seasonally, and many of these were developed in response to spatiotemporal variability in the presence of specific life stages. The previous benchmark blue crab stock assessment for the Chesapeake Bay was completed in 2011 and developed female biological reference points based on maximum sustainable yield (MSY). Following the 2011 assessment, management measures were implemented to reach the reference points and target exploitation rate for adult female crabs. Despite management successfully achieving the reference points and target exploitation rates, the population has not responded as anticipated in recent years. Specifically, recruitment has been lower than expected in the most recent decade. Improved understanding of continued low recruitment has been increasingly called for, and a new benchmark stock assessment was requested to reevaluate reference points for sustainable management. Potential causes of low recruitment include increased predation of juvenile blue crabs by finfish predators, changes in the ocean habitat and currents for larval blue crabs, reduced

¹ https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

numbers of female spawners due to spring fishing for female blue crabs, and sperm limitation caused by too few adult male blue crabs.

NOAA Chesapeake Bay Office then requested the Center for International Experts (CIE) to conduct an independent review of this entire benchmarking process from data identification through to model development and assessment of stock status.

Three CIE reviewers conducted the peer review in accordance with the Terms of Reference (TORs) (see Appendix 2). Approximately three weeks before the peer review, NOAA Chesapeake Bay Office made available via both email and a google drive (<https://drive.google.com/drive/folders/1UkSin2UYwSKB1pQCYkGzZeSwesa4d5ca>) all necessary background information and reports for the peer review. The CIE reviewers participated in a hybrid panel review held in Annapolis, Maryland. The meeting was held in Conference Rooms 1 and 2, 200 Harry S. Truman Parkway, Annapolis, MD, 21401, and streamed live on Microsoft Teams. One member of the RP attended the meetings in person while the two others attended remotely. The meeting began at 12 pm on Monday, April 13th and ended at 4:30 pm on Wednesday, April 15th. Live streaming of the workshop was recorded.

The CIE reviewers were joined by Daniel Hennen from NOAA who acted as chair for the meetings. The RP met with key assessment scientists (AS) Michael Wilberg, Maya Drzewicki, Gina Ralph, Dong Liang, Madison Sholes, Rober Aguilar, Amanda Bevans, Ingrid Braun-Ricks, Mandy Bromilow, Glenn Davis, Alexa Galvan, Thomas Ihde, Robert Latour, Rom Lipcius, Brooke Lowman, Marvin Mace III, Thomas Miller, Matthew Ogburn, Gabrielle Saluta, Alexandra Schneider, Alexei Sharov, and Troy Tuckey. Observers were also present through most of the review workshop, both in person and online (Appendix 3). The assessment team presented the key aspects of their research on the first day according to the agenda in Appendix 4. Copies of the presentations were provided to the reviewers. The RP were able to discuss and query aspects of the presentation as it was being delivered. The assessment team continued to present key aspects of their research on the second day in addition to providing some analytical feedback from questions posed by the RP on day one. The third day of the review consisted of discussions concerning the assessment and feedback of requests being posed to the assessment team during days one and two. Additionally on day three each member of the RP discussed with the chair their overarching views on the completion of each of the assessment TORs) to provide initial feedback to the assessment team during a plenary presentation at the completion of day three. A fourth day, as outlined in the draft agenda, was not required and the RP used this opportunity to continue to develop their review reports.

The bench-mark stock assessment report generated by the assessment team was based around their TORs (which differed slightly from the TORs outlined for the CIE review and for which this review report is based upon):

Assessment Scientists TORs.

The 2026 Chesapeake Bay blue crab stock assessment was funded by grants from the Maryland Department of Natural Resources and the Virginia Marine Resources Commission through the Atlantic States Marine Fisheries Commission and by support from the NOAA Chesapeake Bay Office and the Potomac River Fisheries Commission. The three management agencies jointly developed the following 11 specific terms of

reference (TORs) that were approved by the Chesapeake Bay Stock Assessment Committee (CBSAC).

TOR 1: Critically review and estimate life history parameters and vital rates of blue crab in the Chesapeake Bay that are relevant to the stock assessment. In particular, the assessment should evaluate the extent and scale of interannual variation in life history parameters and vital rates of blue crab in the Chesapeake Bay.

TOR 2: Describe and quantify patterns in fishery-independent surveys to develop indices of abundance and characterize the size composition of the population. Analyses should include: (1) A comprehensive evaluation of the utility of fishery-independent surveys to inform the stock assessment; (2) Consideration of index standardization which may include effects of environmental and abiotic factors on survey catches; and (3) Characterization of uncertainty in indices of abundance.

TOR 3: Describe and quantify patterns in catch, effort, and CPUE. Analyses should include: (1) Estimation of catch and effort for each jurisdiction; (2) Evaluation of the utility of a commercial CPUE index in the assessment; (3) Examination of the impacts of reporting changes and trends in CPUE; (4) Evaluation and quantification of bycatch and/or discard mortality, and recreational harvest using available data from the jurisdictions; and (5) Characterization of uncertainty in the data.

TOR 4: Evaluate the feasibility of, and if possible, implement blue crab stock assessment models that operate on sub-annual time steps and/or at spatial resolutions lower than that of the entire Chesapeake Bay to better represent population dynamics.

TOR 5: Characterize uncertainty in assessment estimates (mortality and abundance).

TOR 6: Update the sex-specific catch survey models used in the 2011 benchmark stock assessment with relevant new data. Characterize major changes in assumptions between the 2011 assessment model and the 2023 model.

TOR 7: Based on assessment model results recommend appropriate biological reference points for management. To extent possible, evaluate the appropriateness and utility of (1) Aggregate bay-wide reference points; (2) Sex specific reference points; and (3) Recruitment reference points.

TOR 8: Evaluate stock status relative to recommended reference points. 2026 Chesapeake Bay blue crab stock assessment.

TOR 9: Identify relevant ecosystem and climate influences (such as habitat, environmental drivers, prey availability, and predation/cannibalism) on the population dynamics and fisheries and, to the extent possible, explore other analyses that support the assessment.

TOR 10: Identify existing data sources and gaps, and, to the extent possible, characterize the uncertainty in the relevant sources of data.

TOR 11: Report on the status of research recommendations from the most recent benchmark assessment. Identify and prioritize research recommendations for future work.

CIE TORs.

TOR 1: Does the assessment critically review and estimate life history parameters and vital rates of blue crab in the Chesapeake Bay that are relevant to the stock assessment?

In particular, does the assessment evaluate the extent and scale of interannual variation in life history parameters and vital rates of blue crab in the Chesapeake Bay?

TOR 2: *Does the assessment describe and quantify patterns in fishery-independent surveys to develop indices of abundance and characterize the size composition of the population?*

TOR 3: *Does the assessment describe and quantify patterns in catch, effort, and CPUE?*

TOR 4: *Does the assessment evaluate the feasibility of, and if possible, implement blue crab stock assessment models that operate on sub-annual time steps and/or at spatial resolutions lower than that of the entire Chesapeake Bay to better represent population dynamics?*

TOR 5: *Does the assessment characterize uncertainty in assessment estimates (mortality and abundance)?*

TOR 6: *Does the assessment update the sex-specific catch survey models used in the 2011 benchmark stock assessment with relevant new data? Does the assessment characterize major changes in assumptions between the 2011 assessment model and the 2023 model?*

TOR 7: *Does the assessment recommend appropriate biological reference points for management? Does the assessment evaluate stock status relative to recommended reference points?*

TOR 8: *Does the assessment identify relevant ecosystem and climate influences (such as habitat, environmental drivers, prey availability, and predation/cannibalism) on the population dynamics and fisheries and explore other analyses that support the assessment?*

TOR 9: *Does the assessment report on the status of research recommendations from the most recent benchmark assessment? Does the assessment identify and prioritize research recommendations for future work?*

TOR 10: *Based on your full review, is the 2025 Blue Crab Stock Assessment adequate for informing management decisions?*

Summary of Findings

The findings of the review have been reported following the terms of reference set for the CIE review workshop.

1. Does the assessment critically review and estimate life history parameters and vital rates of blue crab in the Chesapeake Bay that are relevant to the stock assessment? In particular, does the assessment evaluate the extent and scale of interannual variation in life history parameters and vital rates of blue crab in the Chesapeake Bay?

The assessment team conducted a comprehensive review of blue crab biology and life history in Chesapeake Bay, substantially updating the knowledge base from the 2011 benchmark. Several new analyses were undertaken specifically for this assessment that represent genuine advances. The length-weight relationships (Liang et al. 2026) were rigorously estimated with an explicit test for temporal change. The female maturation growth analysis (Ogburn et al. 2026) quantified the growth increment at the terminal moult and provided a functional relationship that was incorporated directly into the SLAM model. Maturation ogives were developed for both sexes using contemporary survey data (Ogburn et al. 2026; Ralph et al. 2026b), and the overwinter mortality analysis (Liang et al. 2026b) produced year, sex, and length-specific estimates that revealed substantial interannual variation linked to cumulative winter water temperatures. The reproductive biology sections were particularly strong, drawing on an impressive body of recent work including revised fecundity estimates (Schneider et al. 2024a), brood production rates (Schneider et al. 2024b), and the novel nemertean biomarker for spawning history (Schneider et al. 2023a,b).

However, TOR 1 explicitly requested that the assessment "evaluate the extent and scale of interannual variation in life history parameters and vital rates." This analysis was not consistently applied across all life history parameters examined. The overwinter mortality and length-weight analyses both tested for and characterised temporal variation. But for growth rate, size-at-recruitment, timing-of-recruitment, size-at-maturity, timing of maturity, non-winter natural mortality, and fecundity, stationarity was assumed rather than tested. This is a significant gap given the 30-year span of the assessment and the environmental changes that occurred over that period. At minimum, one would expect formal tests for temporal trends or regime shifts in these key parameters, even if the conclusion were that no significant change was detected.

What makes this omission surprising is that evidence for non-stationarity in growth-related processes is presented within the report itself. The assessment cites work demonstrating temperature-dependent growth in juvenile blue crabs (Glandon and Miller 2017; Glandon et al. 2019b), documents long-term changes in mean carapace width from the PEARL pot survey and presents environmental data showing trends in Bay water temperatures over the assessment period (Figure 1.1). As such, the assessment's own evidence base suggests that key vital rates have not been stationary, yet the modelling framework assumes they have been. Addressing temporal variation in growth and natural mortality, whether within the assessment model or through auxiliary analyses, should be a high priority for future work.

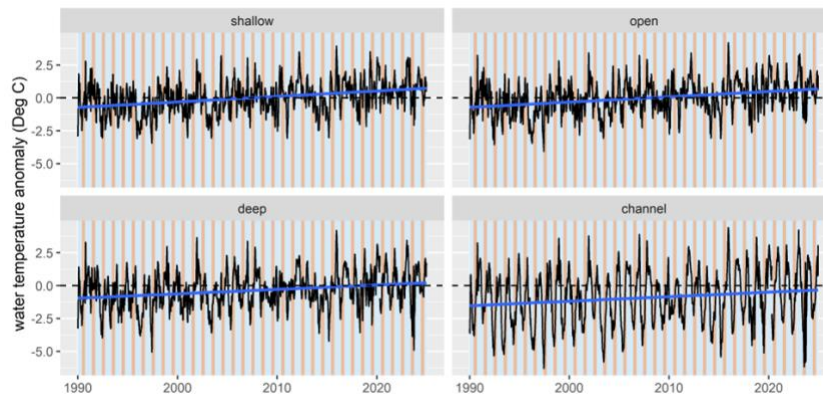


Figure 1.1. Standardised water temperature anomalies for Chesapeake Bay across the temporal scale of the SLAM model.

2. Does the assessment describe and quantify patterns in fishery-independent surveys to develop indices of abundance and characterize the size composition of the population?

The assessment team treatment of fishery-independent data represents a substantial advance over the 2011 benchmark. Eight long-term surveys were evaluated for their utility, with clear rationale for their subsequent inclusion/exclusion from the assessment. The adoption of VAST spatiotemporal models for random-design surveys and GAMLSS for fixed-site surveys is appropriate and well-implemented. Vessel calibration work for both the VIMS Trawl Survey and ChesMMAAP was thorough, supported by years of paired-tow experiments. The development of monthly indices to match the SLAM model's monthly time step is a clear improvement over the annual indices used previously, and the coherence analysis using Kendall rank correlations provides useful transparency about agreement among data sources.

There are no examinations of the spatial coherence of the survey data even though the surveys differ markedly in spatial coverage. The WDS spans the entire Chesapeake Bay, while the VIMS Trawl Survey covers only the area south of the Rappahannock River, ChesMMAAP samples the mainstem, the Maryland Trawl Surveys are restricted to eastern shore tributaries, and the PEARL Pot Survey monitors just a few sites near Calvert Cliffs. The relatively low among-survey coherence (Kendall τ ranging from -0.04 to 0.46) is most likely a consequence of these surveys sampling different portions of the Bay that support different components of the population. Before attempting to reconcile these indices within a single spatially-aggregated model, the assessment team should examine whether the size composition data and population demographics from each survey are consistent with one another (considering survey selectivity). Where size compositions/demographics are similar, the surveys are likely sampling similar segments of the population and can reasonably be treated as observing the same modelled quantity. Where they diverge, combining them under a single selectivity function introduces a

mismatch between what the survey observes and what the model predicts. Breaking the WDS into spatial sub-regions and treating them as separate fleets in the model, each with its own selectivity, would likely improve model fit and better represent the known spatial heterogeneity in blue crab population structure across the Bay, without requiring a more complex model structure (i.e., would not require multiple areas or movement).

A related concern is the inconsistency between how abundance indices and size composition data were treated. The indices were carefully standardised using spatiotemporal or GAM-based approaches that account for spatial, temporal, and vessel effects, yet the size composition data fed into the model were not standardised in the same way. This creates a potential mismatch in the signals being provided to the assessment model. The index says one thing about trends in abundance after removing confounding factors, but the raw size composition data may carry spatial or temporal biases that have not been similarly addressed. There are two approaches to resolving this: either standardise the length-frequency data using comparable model-based methods (which can be computer intensive), or subset the spatial coverage of each survey to identify regions where size compositions are consistent, thereby ensuring that the selectivity assumed in the model is appropriate for the spatial domain actually being indexed. The latter approach has the additional advantage of clarifying the appropriate spatial footprint over which each survey should be truncated to inform the model.

TOR 2 explicitly asks for consideration of environmental and abiotic factors in index standardisation. While the VAST framework provides terms for density and catchability covariates, in practice only vessel effects and area-swept offsets were employed. Temperature and salinity data are collected at every survey station and are known to affect blue crab availability. Productivity of the system, based on temporal and spatial proxies of nutrient load, are also available. This is particularly important for the WDS, where water temperature directly determines whether crabs are torpid and available to the dredge gear. The spatiotemporal random fields in VAST will capture some environmental variation implicitly, but an explicit evaluation of whether including environmental covariates improved index standardisation was not reported. Additionally, the report would benefit from showing the incremental impact of each standardisation component on the resulting index. Presenting the index sequentially, raw then with spatial effects, then with vessel effects, then with spatiotemporal effects, allows the reader to assess how much each factor contributes and whether any single adjustment is driving the overall trend. Some plots associated with the standardisation were provided upon request by the AS, but their inclusion into the report would aid transparency.

Finally, the censoring of recruits from summer–winter surveys using size cutoffs (< 35 mm in August, < 65 mm September–January) impacts the index magnitude and appears to have been determined by visual inspection of length-frequency distributions rather than through formal analysis. Since these cutoffs directly affect both the indices and the size compositions entering the model, and since the model's recruitment estimates depend on what is excluded, a sensitivity analysis on the cutoff values would strengthen confidence in the approach. It is also likely that the delineation between cohorts varies temporally (between years) with changes in recruitment time and growth rates. More representative delineation could be determined through the fitting of year-specific mixture models.

3. Does the assessment describe and quantify patterns in catch, effort, and CPUE?

The assessment team produced a thorough examination of fishery-dependent data from the three management jurisdictions. Descriptions of fishery regulations, reporting requirements, and data collection procedures are detailed and provide a clear understanding of the data. The separation of catch into six fleets (Maryland pot, Maryland trotline, Virginia pot, Virginia dredge, PRFC, and peeler) is well-justified based on differences in selectivity, seasonality, and spatial operation. The treatment of sex-specific catch composition using the Maryland Sentinel Fishery Monitoring Program data is methodical, with appropriate filtering to match legal retention rules across jurisdictions and time periods. The updated approaches to dead discard estimation, drawing on the Walters et al. (2025) and Lipcius et al. (2013) studies, represent improvements over the 2011 assessment. The recreational harvest estimation, while still relying substantially on older studies, was validated by the consistency between the Ashford et al. (2001) and Semmler et al. (2021) estimates for Maryland male recreational harvest (11% vs. 11.2% of commercial male harvest).

The decision not to pursue standardised commercial CPUE as an index of abundance is well-reasoned and clearly articulated. The assessment team correctly identifies fundamental problems with effort measurement in the blue crab fisheries as inconsistent interpretation of "hours fished," ambiguity in whether pot numbers represent pots checked, pots deployed, or maximum licence allowance, and the non-trivial challenge of defining a comparable unit of effort across diverse gear types. Given the availability of multiple high-quality fishery-independent surveys covering the Bay, commercial CPUE would add little information and could introduce misleading signals.

One aspect of the commercial catch information that requires further examination is the conversion of catch from reported units (bushels, numbers) to weight in pounds. This relies on fixed conversion factors that are assumed constant over the entire time series. For example, Maryland uses 0.42 lbs per crab for male hard crabs and 0.31 lbs per crab for females throughout the assessment period, while Virginia updated some conversions in 2018 but applied fixed values before and after that point. If average crab size has changed over time, and there is evidence within this assessment that it has, then applying a constant weight per crab implicitly assumes a stationary size structure in the commercial catch. This could bias the catch time series, potentially underestimating catches in earlier years when crabs were larger and overestimating them in more recent years. A time-varying conversion informed by the available length composition data from the Sentinel Monitoring Program and buyer sampling should be investigated.

The treatment of discards also warrants further clarification and development. The assessment accounts for dead crabs observed in commercial gear, animals that are already dead when the gear is retrieved. This addresses only one component of fishing-induced mortality on non-retained catch. What is not accounted for is the handling mortality imposed on sublegal crabs that are caught in pots, handled, and returned alive but subsequently experience a higher rate of mortality. Every pot haul exposes undersized crabs to air, temperature stress, and physical handling, and a proportion of these animals will subsequently die even though they are returned. This source of mortality is routinely estimated and included in trawl fishery assessments, and the Walters et al. (2025) study

cited in the report provides directly relevant mortality estimates for pot-caught blue crabs (ranging from 2% to 14% depending on salinity and season). Incorporating this handling mortality for returned sublegal crabs, in addition to the observed dead discards, would provide a more complete accounting of total fishing-induced mortality and is likely to be non-trivial given the large number of sublegal encounters in the pot fishery.

The characterisation of uncertainty in the catch data is clear and transparent, identifying specific periods and sources of concern including the early years of mandatory reporting in Virginia (1994–1996), the over-reporting of female hard crabs in Maryland following the introduction of bushel limits in 2008, and the lack of independent audit capability in the PRFC. One area where the treatment could be strengthened is the description and presentation of catch data prior to 1994, the start year of the assessment model. While mandatory reporting across all jurisdictions only begins in 1994, commercial landings data extend considerably further back. The trajectory of these catches through the 1980s and early 1990s would provide valuable context for evaluating the model's initial conditions in 1994 and how intensive the burn in period for the model should be. If catches were declining, stable, or at historically high levels entering the assessment period, this has implications for whether the assumption that the population was near equilibrium at the start of the model is reasonable, and for interpreting estimated abundance trends in the early years of the time series.

4. Does the assessment evaluate the feasibility of, and if possible, implement blue crab stock assessment models that operate on sub-annual time steps and/or at spatial resolutions lower than that of the entire Chesapeake Bay to better represent population dynamics?

The development of the Sex and Length Structured Assessment Model (SLAM) is a substantial improvement for the Chesapeake Bay blue crab stock assessment, moving from the age-based catch survey model used in 2011 to a monthly, length and sex structured framework implemented in RTMB. The model code is clear and well structured, and the process of tracking males, immature females, and mature females through 1-cm length bins, applying length-based natural mortality, overwinter mortality, and fleet-specific fishery selectivity is appropriate for this species. The decision to use RTMB provides automatic differentiation and access to random effects. The model converged with a small maximum gradient and an invertible Hessian, and the jitter analyses, while revealing some instability associated with PEARL catchability and ChesMMA selectivity, supported the robustness of the chosen model structure. However, the report's description of the model development process was incomplete and did not fully reflect the sequence of decisions and iterations that led to the final model structure, as described during the review process.

A concern with the current parameterisation is the estimation of separate fishing mortality intensities for every month \times sex \times fleet combination. While sex-specific fishing mortality is biologically plausible given different regulations and spatial distributions, estimating fully independent F values by sex means that any discrepancy between observed and predicted sex ratios in the catch can be absorbed into F rather than being attributed to selectivity or population structure. This can mask real conflicts in the data. If

male and female F estimates are highly correlated across time within a fleet, then splitting by sex is adding parameters without adding information. The assessment team should examine whether this is the case. An alternative approach such as a hybrid-F method, which solves for sex-aggregated fleet-specific harvest rates iteratively given observed total catches and predicted vulnerable biomass, would achieve the same outcome with far fewer estimated parameters and would force the model to reconcile sex ratios through population dynamics and selectivity rather than through unconstrained F estimates. Under this framework, observed catch sex ratios could be fitted as a separate likelihood component with uncertainty that reflects the actual sampling coverage (e.g., Sentinel program sample sizes, buyer sampling availability, and mixed-bushel allocation rules), rather than being implicitly treated as known through tight fits to sex-specific catches. More broadly, the model estimates 2,639 parameters, and the ratio of estimated fishing mortality parameters to the information content of the data should be critically evaluated. The current model with its large number of parameters to be estimated may carry a risk of overfitting.

The fishery selectivity raises several questions. Selectivity was fixed based on minimum size regulations rather than estimated, which is a defensible simplification, but the iterative addition of constants to the inflection point parameter to improve fits introduces a degree of ad hoc tuning that is difficult to evaluate. Was this required due to the large size bins (1 cm) and the assumption of an even distribution of crabs across a length bin? For the surveys, the use of spline-based selectivity for the Maryland and VIMS trawl surveys produces flexible curves, but several aspects of the results require closer examination and justification. First, the report presents estimated selectivity splines for only three months for each survey (May–July for the MD Trawl, April–June for the VIMS Trawl), with selectivity fixed at 1 for all remaining months. The rationale for this choice “that censoring of recruits precludes estimation of the ascending limb in later months” is stated but is not fully convincing, and selectivity for all months should be shown and the decision justified more transparently. Second, and more concerning, both surveys show strikingly similar dome-shaped selectivity patterns in the spring months, with reduced selectivity for large crabs. If these splines are capturing spatial availability driven by migration (mature females moving south toward the spawning grounds) then the patterns should be opposing, not parallel. Large crabs migrating from the upper Bay (Maryland) to the lower Bay (Virginia) should become less available to the MD Trawl and more available to the VIMS Trawl during the same months, not less available to both simultaneously. The fact that both surveys show the same declining selectivity for large crabs in spring suggests the splines may be compensating for a structural model-data conflict rather than representing genuine availability. This warrants investigation: are the splines absorbing misspecification in growth, natural mortality, or spatial population structure? Without a clear biological mechanism that explains why large crabs would be simultaneously unavailable to trawl surveys in both the upper and lower Bay, the estimated selectivity shapes should be treated with caution. More fundamentally, the model does not distinguish between encounter probability (does the gear contact the animal?) and retention (given contact, is the animal caught?). For the commercial fisheries, the assumption that recreational selectivity is identical to commercial selectivity for each jurisdiction is untested and likely incorrect given the different gears and fishing behaviours involved.

The model operates on a monthly time step throughout the year, but this level of temporal resolution may be unnecessarily complex during periods when little is happening biologically. During the winter months (December–March), there is no growth, no reproduction, and no commercial fishing other than the now-closed Virginia dredge fishery. Pooling these months into a single time step would reduce the number of estimated parameters without sacrificing biological realism and would simplify the model.

Several structural features that could improve model performance were not explored. First, the model does not include process error. (Was this because of the large number of fixed parameters slowed down model runs, making the inclusion of process error prohibitively slow?) All variability is attributed to observation error, which means the model is forced to track every signal in the data rather than allowing for unmodelled stochasticity in the population dynamics. Adding process error on abundance or recruitment would provide a more realistic characterisation of uncertainty and could improve the residual patterns observed in the early years of the time series. Second, the model's initial conditions assume quasi-equilibrium in 1994, but the large, standardised residuals in the mid-1990s suggest this assumption may be inappropriate. Exploring alternative burn-in approaches or extending the model backward using available pre-1994 catch data could help resolve this. Third, the size distribution of recruits is assumed constant over the entire time series, yet changes in the timing of settlement, environmental conditions, or nursery habitat quality could shift both the mean size and variability of recruits entering the population. Allowing non-stationarity in recruitment size parameters would provide flexibility to accommodate such changes and could act as a proxy for temporal variation in both the size and timing of recruitment, given the simplification of modelling it as a February pulse. Fourth, the model does not include any density-dependent natural mortality (cannibalism), despite cannibalism being well-documented in blue crab ecology and potentially important for understanding the relationship between recruitment and subsequent juvenile survival.

The model is spatially aggregated, treating the entire Chesapeake Bay as a single population unit despite substantial spatial heterogeneity in blue crab demographics. The WDS length-frequency data, when separated by state, reveal strikingly different population structures. Maryland stations show a population dominated by small crabs with a secondary mode (~130 mm) of medium-sized animals, while Virginia stations show a bimodal distribution with a much more prominent mode of large crabs (~145 mm), bigger than those in Maryland (Figure 3.1). These differences reflect the known spatial segregation of life stages, with juveniles concentrated in upper Bay nursery habitats and mature females migrating to the lower Bay for spawning. Fitting a single selectivity and catchability to the Bay-wide WDS index forces the model to average across what are essentially different demographic components. A more informative approach, as discussed under TOR 2, would be to treat spatial sub-regions of the WDS as separate fleets in the model, each with its own selectivity. This allows the model to fit the distinct size compositions observed in different parts of the Bay and better represent spatial population structure without requiring a fully spatial model.

The size-varying natural mortality applied to crabs ≤ 5 cm adds complexity in a size range where the data may be most limited due to the sampling limitations of the surveys (minimum depth of 1.5 m for WDS, gear selectivity for trawl surveys). The model could

potentially be simplified by recruiting at a larger initial size (e.g., 5 cm, still age-0+), which would reduce the need for the size-dependent M scaling and place the model's dynamics in a size range better supported by the available data.

Finally, the data weighting approach deserves scrutiny. Effective sample sizes for composition data were determined using the number of sampling trips or fixed values, with a weighting term of 0.25 applied to non-WDS surveys chosen such that standardised residuals were manageable. This approach, while practical, lacks the iterative self-consistency of established methods such as Francis weighting, which re-weights composition data based on the fit of the model to the mean length or age. The assumed log-scale standard deviations for catch (0.1) and survey indices (0.25 for WDS, 0.4 for others) were fixed rather than estimated or iteratively tuned, and the report does not discuss whether these values appropriately balance the information content of the different data sources. Bias adjustment or ramping on recruitment deviations, which is standard practice in models like Stock Synthesis to avoid biased estimates in the early and late parts of the time series, does not appear to have been considered. Given the sensitivity of the model to growth parameters documented in the sensitivity analysis, and the large number of fixed assumptions in the model, a more systematic approach to data weighting and model diagnostics would strengthen confidence in the assessment results. The peeler fleet, for which catch can be reported directly in numbers of crabs, could be fitted in numbers rather than weight, avoiding the additional uncertainty introduced by weight conversion factors.

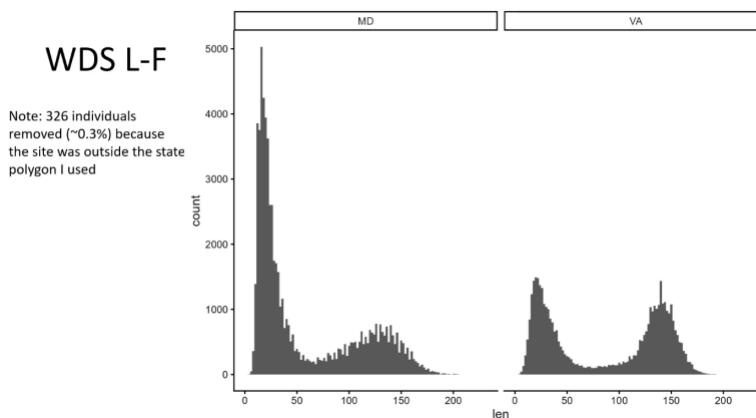


Figure 3.1. Size composition of blue crab sampled from the winter dredge survey pooled over a number of years and separated into surveys conducted in Maryland and Virginia waters. Figure was provided upon request during the review meeting.

5. Does the assessment characterize uncertainty in assessment estimates (mortality and abundance)?

The assessment characterises uncertainty through three approaches: asymptotic standard errors from the estimated variance-covariance matrix, sensitivity analyses of fixed parameters, and jitter analyses to test convergence stability. The jitter analysis demonstrated that the base model converged on a consistent solution, providing

reasonable confidence in the optimisation process. The sensitivity analysis systematically tested alternative values for growth, maturation, and natural mortality parameters, giving insight into which assumptions most influence model outputs. The asymptotic standard errors, while standard practice, may underestimate true uncertainty for a model with this many fixed structural assumptions.

The most significant finding from the sensitivity analysis was the influence of pre-determined growth parameters. Alternative growth parameterisations produced markedly different abundance and fishing mortality estimates, with faster-growth scenarios estimating lower abundance and some generating implausibly high fishing mortality in the first year. This sensitivity is particularly concerning because the growth parameters were not estimated from an independent data source but were instead constructed through iterative manual calibration within the SLAM model. This circularity means the growth parameters are conditioned on the model's other assumptions (natural mortality, selectivity, recruitment), and any errors in those assumptions will propagate into the growth specification and then back into abundance estimates. If growth has changed over time, as the biological evidence reviewed under TOR 1 suggests is plausible, the impact on assessment results could be substantial. Deriving growth parameters from an external data source, or at minimum developing an independent validation of the current parameterisation, should be a priority.

Retrospective analysis with associated Mohn's rho values was presented to the RP during the meetings and should be incorporated into the report as it provides critical information about the stability and reliability of the model's terminal year estimates, which are the estimates most directly relevant to management advice.

Several additional dimensions of uncertainty remain unknown. The surveys used in the model cover different spatial domains across the Bay, yet the model assumes they are all observing the same population with the same size composition, differing only through selectivity. In reality, size composition almost certainly varies spatially, with distance from the ocean, water depth, and habitat type, as well as along the mainstem axis of the Bay. Understanding these spatial patterns in size structure would help clarify whether the unusual survey selectivity patterns estimated by the model (particularly the dome-shaped splines for the MD and VIMS trawl surveys discussed under TOR 4) reflect genuine gear selectivity, spatial availability, or model misspecification. More broadly, the sensitivity analysis focused on one-at-a-time perturbations of fixed parameters, which does not capture interactions among assumptions. The large residuals in the early years of the model, related to the quasi-equilibrium initial conditions discussed under TOR 4, represent another source of structural uncertainty that has not been formally evaluated. Exploring alternative burn-in periods and initial condition assumptions would help quantify how much the early-period estimates, and by extension the baseline against which trends are measured, depend on this structural choice.

6. Does the assessment update the sex-specific catch survey models used in the 2011 benchmark stock assessment with relevant new data? Does the assessment characterize major changes in assumptions between the 2011 assessment model and the 2023 model?

The assessment team undertook a commendable set of comparisons between the 2011 model (SSCMSA) and the new SLAM model, providing important context for interpreting the transition between assessment frameworks. The comparison was well structured, progressing from removals data through fishery-independent indices to continuity model runs, allowing the reader to trace which changes in data treatment and model assumptions drove differences in results. The strong correlations between design-based and model-based survey indices, particularly for the WDS and VIMS Trawl Survey, provide reassurance that the shift to spatiotemporal index standardisation has not fundamentally altered the signals entering the model. The comparison of ADMB and STAN implementations of the SSCMSA is a useful exercise that identified numerical stability issues in the ADMB version for the shortened 1994–2023 time series and confirmed broadly similar population trajectories between the two platforms. However, the progression from ADMB (frequentist) to STAN (Bayesian) and then to RTMB (frequentist) for the final SLAM model is an unusual path for comparing model structures. The priors used in STAN will influence parameter estimates differently from the penalty functions in the frequentist implementations, and this should be acknowledged when interpreting differences between the platforms. A clearer discussion of how prior specifications in STAN were chosen and how they compare to the penalty structures in ADMB and RTMB would help the reader assess whether differences between implementations reflect genuine model behaviour or artefacts of the estimation framework.

The most significant finding from the model comparison is that SLAM estimates abundance approximately twice as high as the updated 2011 model. The assessment team attributes this primarily to different assumptions about the size distribution of recruits, WDS selectivity, and length-specific natural mortality. Specifically, SLAM assumes the WDS observes less than 10% of age-0 blue crabs in winter, whereas the 2011 model estimated approximately 60%. While this explanation is plausible, a factor-of-two difference in estimated abundance is significant and warrants further investigation to determine which estimate is closer to reality. Independent, external estimates of absolute biomass or abundance would be valuable for ground-truthing. Several potential sources exist: the Lipcius et al. (2026a,b) tag-recapture studies could inform Brownie-type abundance estimates given assumptions about exploitation rate; surplus production models such as JABBA or SPiCT applied to the available index and catch data could provide model-independent biomass trajectories; and Leslie-type depletion estimates, if feasible from within-season catch and effort data, could offer ballpark abundance estimates for comparison (generally within a year/fishery/area the relative bias in reporting of commercial effort is consistent between months). Even approximate external estimates would help bracket the plausible range and inform which set of assumptions about WDS catchability and recruit visibility is more defensible.

Despite the difference in absolute scale, both models suggest broadly similar conclusions about stock sustainability relative to threshold and target reference points, which provides some reassurance for management. However, if both models agree that fishing mortality

is currently at or below target levels, yet recruitment has been declining since 2010 and is at historically low levels, there is a risk that management could maintain or increase fishing pressure in the face of declining productivity. The current framework, which relies on exploitation rate and female abundance reference points, does not provide a mechanism to respond proactively to poor recruitment before it manifests as reduced harvestable abundance. Incorporating a recruitment index into the management framework, as discussed under TOR 7, would provide an early warning system that triggers precautionary action when recruitment falls below critical levels, even if current F appears sustainable.

7. Does the assessment recommend appropriate biological reference points for management? Does the assessment evaluate stock status relative to recommended reference points?

The adoption of spawning potential ratio (SPR) reference points is a sound decision that avoids the well-documented difficulties in estimating stock-recruitment relationship parameters that occurred in the 2011 assessment. The SPR approach is widely accepted in fisheries management. The development of separate SPR models for males and females is appropriate given the sex-specific differences in growth, maturation, selectivity, and fishing mortality. For females, the use of egg production (batch fecundity \times abundance of mature females \times probability of spawning) as the reproductive metric is well-justified and incorporates the revised fecundity estimates from Schneider et al. (2024a). For males, the reproductive potential calculation uses weight-at-length and an assumed mating frequency of twice per month during April–November, which is a reasonable approximation but should be acknowledged as uncertain. The recruitment reference point (three-year running average compared to the 25th percentile of estimated recruitment) is a pragmatic approach that recognises that sustained low recruitment is more concerning than individual poor years. The target fishing mortality rates, set at 75% of the threshold F40% values, provide a sensible buffer.

The choice of 40% SPR as the threshold is on the conservative end for crustacean fisheries, where reference points of 20–35% SPR are more common (e.g., 20% for many crab stocks on the US west coast). The assessment team justifies this conservatism on three grounds: robustness to potential climate effects on population dynamics (Shertzer et al. 2024), the recent unexplained decline in recruitment, and recent collapses of other crab stocks under environmental change (Szuwalski et al. 2023). While this precautionary approach is understandable given the current trajectory of the stock, the justification would benefit from a more formal evaluation (e.g., through management strategy evaluation (MSE)), to determine whether F40% achieves the stated management objectives under plausible future scenarios. It would also be worth examining whether the SPR calculations account for potential non-stationarity in size at maturity; if maturity ogives have shifted over time, using a fixed maturity schedule in the reference point calculations could bias the SPR estimates.

Although the assessment team chose not to fit a stock-recruitment relationship, this does not preclude estimating MSY-based reference points. An iterative equilibrium yield approach can be implemented directly using the full SLAM model dynamics: for each

candidate fishing mortality level, the population is projected forward through the complete monthly growth, natural mortality, overwinter mortality, maturation, and selectivity processes assuming a fixed representative recruitment (e.g., the median estimated recruitment over most recent ten years), allowed to reach equilibrium, and the resulting yield is calculated. Scanning across a range of F values produces an equilibrium yield curve whose peak defines an MSY proxy and the corresponding F_{msy} proxy. This approach exploits the full complexity of the model rather than relying on simplified per-recruit equations or an assumed SPR percentage as a proxy for sustainability. The assumption of constant recruitment means this approach will overestimate true MSY if recruitment declines at low stock sizes, but given that the current evidence suggests recruitment variability is driven primarily by environmental factors rather than spawning stock size, a constant-recruitment MSY estimate may be more defensible than forcing an ill-fitting parametric stock-recruit relationship. This approach could sit alongside the SPR-based reference points to provide complementary management information and would directly answer the question most relevant to managers: what is the maximum yield the stock can sustain?

The stock status results paint a mixed picture. Fishing mortality rates are currently well below both threshold and target levels for both sexes, and mature female abundance was slightly above the target in 2023, suggesting that overfishing is not occurring and the female spawning stock is not overfished by the recommended criteria. However, male abundance has consistently been below the N40% threshold for most of the time series, and the three-year average recruitment fell below the 25th percentile threshold in both 2022 and 2023. The juxtaposition of low fishing mortality, adequate female abundance, and declining recruitment strongly suggests that the driver of recent poor recruitment is environmental rather than fishing-related, which has important implications for management. If recruitment continues to decline despite low F , reducing fishing pressure further will not address the underlying problem. This reinforces the value of the recruitment reference point as an early warning indicator and supports the case for incorporating it formally into management decision-making.

From a communication and practical management perspective, the use of instantaneous fishing mortality rate (F) as the primary metric for comparing current removals to reference points, while technically correct, is not intuitive for managers, industry, or recreational fishers. Exploitation rate (the proportion of the population removed by fishing) is a more readily understood metric that translates directly into the fraction of the stock being harvested. It is also typically more stable across years than F -based estimates, which can fluctuate substantially. If exploitation rate is adopted as a management metric, it should be calculated using mid-season biomass (i.e., biomass adjusted for half a year's natural mortality) rather than simply using the abundance of recruits for the year, to provide a more accurate representation of the stock available to the fishery during the fishing season. This would be straightforward to derive from the existing model outputs and would improve the accessibility of assessment results to stakeholders.

8. Does the assessment identify relevant ecosystem and climate influences (such as habitat, environmental drivers, prey availability, and predation/cannibalism) on the population dynamics and fisheries and explore other analyses that support the assessment?

The assessment team investigated several important ecosystem and climate influences on blue crab population dynamics, covering blue catfish predation, spatiotemporal temperature dynamics, overwinter mortality, temperature-dependent growth, and hypoxia effects. This represents a broader scope of ecosystem investigation than was attempted in the 2011 assessment and a move forward to better understanding the drivers of the recent recruitment decline. The blue catfish analysis (Sholes et al. 2026) is particularly interesting given the rapid expansion of this invasive species throughout Chesapeake Bay tributaries, and the finding that juvenile blue crab abundance in 2023 may have been approximately 8% lower than it would have been in the absence of blue catfish provides a useful, if modest, quantification of this threat. The 4-dimensional temperature interpolation approach (Liang et al. 2026c) is an innovative method to produce daily, spatially resolved temperature fields across the Bay, a valuable data source that could have been used more within this assessment.

The investigation of ecosystem and climate influences was largely conducted as standalone analyses external to the assessment framework, when in many cases these drivers could and should have been examined within the VAST and GAMLSS index standardisation models. Both frameworks readily accommodate environmental covariates (temperature, salinity, dissolved oxygen, habitat type, productivity) as either catchability or density covariates, allowing their effects on survey catches to be quantified statistically and accounted for within the derived indices. This is a far more natural and statistically rigorous approach than attempting to add environmental drivers directly into a complex process model like SLAM. By fitting environmental covariates within the index standardisation stage, the indices entering the assessment model would already be corrected for environmental effects on availability and catchability, and the relative importance of different drivers could be assessed through standard model selection. For example, temperature effects on blue crab availability to the WDS (where torpor directly affects catchability), the relationship between hypoxia and survey encounter rates in the ChesMMAP, and the influence of salinity regime on juvenile crab distribution could all be evaluated within the existing VAST/GAMLSS frameworks using the environmental data that are already collected at every survey station. This approach would address the TOR's request to consider environmental and abiotic factors while keeping the assessment model itself tractable.

The failure to incorporate temporal variability in growth remains a significant limitation, but the solution need not require a fully mechanistic temperature-growth relationship. The current monthly Gompertz b parameters were fixed through manual calibration and represent average seasonal growth patterns across the entire time series. An alternative would be to treat these fixed values as priors or mean values and allow b to vary between years by a small amount — effectively estimating year-specific b parameters as random effects with the current values as the mean and a constrained variance. This would provide a mechanism for incorporating interannual variation in growth, informed by the model's ability to track changes in survey length compositions from year to year, without

requiring identification of the specific environmental driver. If growth genuinely varies across years (due to temperature, food availability, or other factors) this approach would allow the model to accommodate that variation, and the estimated year effects could subsequently be examined for correlations with environmental covariates as a post hoc diagnostic. This is preferable to the current approach, which assumes growth has been identical in every year from 1994 to 2023 and forces all interannual variation into recruitment magnitude and fishing mortality estimates. Given the documented sensitivity of the model to growth parameters, even modest year-to-year variation in growth could have meaningful effects on estimated abundance and fishing mortality trajectories. Several potentially important ecosystem interactions were not examined, or were identified as gaps that could not be addressed within the assessment timeframe. Cannibalism by adult blue crabs on juveniles is well-documented (Hines et al. 2026) and could represent a density-dependent mechanism linking adult abundance to recruitment survival, but it was not investigated as a driver of recruitment variability. The potential role of other predators, including the expanding red drum population, was acknowledged but not analysed due to data limitations. Changes in juvenile nursery habitat over time (seagrass coverage, marsh extent, and shoreline hardening) were intended to be investigated but were not completed. Oceanographic drivers of recruitment variability were deferred to another study. While no single assessment can address all possible ecosystem interactions, the collective effect of these gaps means that the dominant drivers of the recent recruitment decline remain unidentified. Given that the stock status determination shows low fishing mortality, adequate female spawning stock, but declining recruitment, understanding the environmental and ecological drivers of recruitment is arguably the single most important question for the future management of this stock.

9. Does the assessment report on the status of research recommendations from the most recent benchmark assessment? Does the assessment identify and prioritize research recommendations for future work?

The assessment team provided a thorough accounting of the status of research recommendations from the 2011 benchmark and the 2017 update, systematically addressing each recommendation with a clear description of what was done, what was partially addressed, and what remains outstanding. This is well-organised and transparent. Substantial progress has been made on several key recommendations: the incorporation of additional surveys (ChesMMAP, PEARL Pot), the development of sex-specific growth and monthly time steps, the evaluation of gear efficiency differences through VAST, and new biological research on reproduction, natural mortality, and maturation. The honest acknowledgement that some recommendations could not be addressed (including spatial modelling, recreational harvest surveys, and juvenile habitat change) due to time and resource constraints is appropriate, and the note that the US government shutdown prevented analysis of the SERC tagging data provides important context. The recurring theme across both the 2011 and 2017 recommendations is the difficulty of resolving spatial dynamics, growth, natural mortality, and recreational harvest, all of which remain priority data gaps in the current assessment.

The new research recommendations developed by the 2026 assessment team are comprehensive, covering 27 items organised by TOR. The priorities are well-chosen and reflect the key limitations identified throughout this review: improving understanding of growth through field methods (recommendation 1), better characterising crabs in shallow habitats not sampled by the WDS (2), improving natural mortality estimates (3), understanding spatial recruitment patterns (5), reconsidering the recruit censoring approach (11), and developing a spatial model (20). The recommendation to develop and test harvest control rules (21) is particularly important given the current reliance on reference points without a formal management strategy evaluation. However, the recommendations would benefit from explicit prioritisation, distinguishing between those that are critical for the next assessment update versus longer-term research goals. Currently, all 27 items are presented with roughly equal weight, which does not help managers and funders allocate limited resources. For example, improving growth estimates (1), developing a spatial model (20), and better understanding recreational harvest (16) are arguably foundational for the next assessment, whereas some other recommendations, while valuable, could be deferred without compromising assessment quality.

Several additional recommendations arising from this review should be considered.

1. Incorporating environmental covariates into the VAST and GAMLSS index standardisation models, an approach that would address multiple ecosystem-related questions without adding complexity to SLAM (as discussed under TOR 9/ecosystem).
2. Implementing time-varying growth through random effects on the Gompertz b parameters, which would address the non-stationarity concerns raised under TOR 1 and TOR 5 and is feasible within the existing RTMB framework.
3. Exploring a hybrid-F parameterisation to reduce the number of estimated parameters while maintaining the sex- and fleet-specific structure.
4. Examining spatial patterns in size composition across the Bay to inform whether splitting the WDS into sub-regional indices would improve model performance.
5. Evaluating the feasibility of external abundance estimates (tag-recapture, surplus production models) to help resolve the factor-of-two scale difference with the 2011 assessment.
6. Conducting a formal retrospective analysis and including it in the assessment report.

These recommendations would complement those already identified by the assessment team and address several structural issues identified during this review.

10. Based on your full review, is the 2025 Blue Crab Stock Assessment adequate for informing management decisions?

The 2026 Chesapeake Bay Blue Crab Stock Assessment represents a substantial advance over the 2011 benchmark in nearly every aspect. The move from an age-based catch survey model to a sex- and length-structured model (SLAM) operating on a monthly time step is well-suited to a species that cannot be aged reliably, and the implementation in RTMB

provides a modern, transparent, and extensible modelling platform. The breadth of new biological research incorporated into the assessment, including revised fecundity estimates, maturation ogives for both sexes, length-specific overwinter mortality, and the female maturation growth relationship, reflects a considerable investment of effort by the assessment team. The evaluation of eight fishery-independent surveys using state-of-the-art spatiotemporal and GAM-based standardisation methods is a major methodological improvement, and the comprehensive treatment of fishery-dependent data across three management jurisdictions is thorough and well-documented. The adoption of SPR-based reference points avoids the estimation difficulties that undermined confidence in the stock-recruitment-based reference points from the 2011 assessment, and the addition of a recruitment reference point provides a valuable early warning mechanism. On balance, the assessment is adequate for informing management decisions and represents the best scientific information currently available for the Chesapeake Bay blue crab stock.

References (including reference material not cited in the text):

- Abbe, G.R. 2002. Decline in size of male blue crabs (*Callinectes sapidus*) from 1968 to 2000 near Calvert Cliffs, Maryland. *Estuaries* 25(1):105–114.
- Ashford, J.R. and C.M. Jones. 2001. Survey of the blue crab recreational fishery in the Chesapeake Bay 2001. A completion report to Maryland Department of Natural Resources. Old Dominion University, Norfolk, VA.
- Ashford, J.R. and C.M. Jones. 2002. Survey of the blue crab recreational fishery in the Chesapeake Bay 2002. A completion report to Maryland Department of Natural Resources. Old Dominion University, Norfolk, VA.
- Ashford, J.R., Jones, C.M., and Fegley, L. 2009. A license registry improves sampling efficiency for a marine recreational survey. *Transactions of the American Fisheries Society* 138:984–989.
- Ashford, J., Jones, C.M., and Fegley, L. 2010a. Private waterfront households catch less per trip than other fishers: Results of a marine recreational survey. *Transactions of the American Fisheries Society* 139:1083–1090.
- Ashford, J., Jones, C., Fegley, L., and O'Reilly, R. 2010b. Catch data reported by telephone avoid public access bias in a marine recreational survey. *Transactions of the American Fisheries Society* 139(6):1751–1757.
- Ashford, J.R., Jones, C.M., and Fegley, L. 2013a. Independent estimates of catch by private and public access fishers avoid between-group sources of error in a recreational fishing survey. *Transactions of the American Fisheries Society* 142:422–429.
- Ashford, J.R., Jones, C.M., and Fegley, L. 2013b. Within-day variability in catch taken by public access fishers during a recreational fishing survey. *Transactions of the American Fisheries Society* 142:974–978.
- Bauer, L.J. and T.J. Miller. 2010a. Temperature-, salinity-, and size-dependent winter mortality of juvenile blue crabs (*Callinectes sapidus*). *Estuaries and Coasts* 33:668–677.
- Bauer, L.J. and T.J. Miller. 2010b. Spatial and interannual variability in winter mortality of the blue crab (*Callinectes sapidus*) in the Chesapeake Bay. *Estuaries and Coasts* 33:678–687.
- Bromilow, A.M. and R.N. Lipcius. 2017. Mechanisms governing ontogenetic habitat shifts: Role of trade-offs, predation, and cannibalism for the blue crab. *Marine Ecology Progress Series* 584:145–159.
- Brylawski, B.J. and T.J. Miller. 2006. Temperature-dependent growth of the blue crab (*Callinectes sapidus*): a molt process approach. *Canadian Journal of Fisheries and Aquatic Sciences* 63:1298–1308.
- Bunnell, D.B. and T.J. Miller. 2005. An individual-based modeling approach to per-recruit models: blue crab *Callinectes sapidus* in the Chesapeake Bay. *Canadian Journal of Fisheries and Aquatic Sciences* 62:2560–2572.
- Cao, J., Y. Chen, and A. Richards. 2017. Improving assessment of *Pandalus* stocks using a seasonal, size-structured assessment model with environmental variables: Part I: Model description and application. *Canadian Journal of Fisheries and Aquatic Sciences* 74(3):349–362.
- Clark, W.G. 1991. Groundfish exploitation rates based on life history parameters. *Canadian Journal of Fisheries and Aquatic Sciences* 48:734–750.
- Colton, A.R., M.J. Wilberg, V. Coles, and T.J. Miller. 2014. An evaluation of the synchronization in the dynamics of blue crab populations in the western Atlantic. *Fisheries Oceanography* 23:132–146.
- Crone, P.R. and D.B. Sampson. 1997. Evaluation of assumed error structure in stock assessment models that use sample estimates of age composition. *International Symposium on Fishery Stock Assessment Models for the 21st Century*, Anchorage, Alaska.
- Darnell, Z.M., D. Rittschof, K.M. Darnell, and R.E. McDowell. 2009. Lifetime reproductive potential of female blue crabs *Callinectes sapidus* in North Carolina, USA. *Marine Ecology Progress Series* 394:153–163.
- Davis, J.L.D., A.C. Young-Williams, A.H. Hines, and Y. Zohar. 2005. Assessing the potential for stock enhancement in the case of the Chesapeake Bay blue crab (*Callinectes sapidus*). *Canadian Journal of Fisheries and Aquatic Sciences* 62:109–122.
- Dickinson, G.H., D. Rittschof, and C. Latanich. 2006. Spawning biology of the blue crab, *Callinectes sapidus*, in North Carolina. *Bulletin of Marine Science* 79:273–285.
- Doering, K.L., M.J. Wilberg, D. Liang, and M. Tarnowski. 2021. Patterns in oyster natural mortality in Chesapeake Bay, Maryland using a Bayesian model. *Fisheries Research* 236:105838.
- Drzewicki, M.R. and M.J. Wilberg. 2026. Additional plots for the 2026 Chesapeake Bay blue crab assessment. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 13.
- Efron, B. and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*, 1st edition. Chapman and Hall/CRC, New York.
- Fabrizio, M.C. and T.D. Tuckey. 2016. Calibration of VIMS research vessel catch data to ensure continuity of recruitment indices for the Chesapeake Bay region. Final Report to NMFS, NOAA Chesapeake Bay Office.
- Fabrizio, M.C., Tuckey, T.D., Buchanan, J.R., and Fisher, R.A. 2025. Predation impacts of invasive blue catfish on blue crabs in estuarine environments. *Marine and Coastal Fisheries: Dynamics, Management and Ecosystem Science* 17(4).
- Fegley, L.W., Davis, B.K., Davis, G.R., Webb, K., Walstrum, J.C., and Brown, H.R. 2006. Harvest and effort characterization of the Maryland Chesapeake Bay blue crab fishery through a cooperative data collection program. Final report from MDNR to NOAA Chesapeake Bay Office, Annapolis, Maryland.
- Fournier, D.A., J.R. Sibert, J. Majkowski, and J. Hampton. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (*Thunnus maccoyii*). *Canadian Journal of Fisheries and Aquatic Sciences* 47:301–317.

Gabriel, W.L., M.P. Sissenwine, and W.J. Overholtz. 1989. Analysis of spawning stock biomass per recruit: an example for Georges Bank haddock. *North American Journal of Fisheries Management* 9:383–391.

Gartland, J. and R.J. Latour. 2024. Disentangling bottom-up and top-down controls on fish consumption of key prey in the Northeast US Shelf ecosystem. *ICES Journal of Marine Science* 81:1669–1684.

Glandon, H.L. and T.J. Miller. 2017. No effect of high pCO₂ on juvenile blue crab, *Callinectes sapidus*, growth and consumption despite positive responses to concurrent warming. *ICES Journal of Marine Science* 74(4):1201–1209.

Glandon, H.L., Kilbourne, K.H., Schijf, J., and Miller, T.J. 2018. Counteractive effects of increased temperature and pCO₂ on the thickness and chemistry of the carapace of juvenile blue crab, *Callinectes sapidus*, from the Patuxent River, Chesapeake Bay. *Journal of Experimental Marine Biology and Ecology* 498:39–45.

Glandon, H.L., Kilbourne, K.H., and Miller, T.J. 2019b. Winter is (not) coming: Warming temperatures will affect the overwinter behavior and survival of blue crab. *PLoS ONE* 14(7):e0219555.

Graham, D.J., Perry, H., Biesiot, P., and Fulford, R. 2012. Fecundity and egg diameter of primiparous and multiparous blue crab *Callinectes sapidus* (Brachyura: Portunidae) in Mississippi waters. *Journal of Crustacean Biology* 32:49–56.

Haefner, P.A. and C.N. Shuster. 1964. Length increments during terminal molt of the female blue crab, *Callinectes sapidus*, in different salinity environments. *Chesapeake Science* 5(3):114–118.

Hamel, O.S. and J.M. Cope. 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. *Fisheries Research* 256:106477.

Hewitt, D.A., D.M. Lambert, J.M. Hoenig, R.N. Lipcius, D.B. Bunnell, and T.J. Miller. 2007. Direct and indirect estimates of natural mortality for blue crab. *Transactions of the American Fisheries Society* 136:1030–1040.

Hilling, C.D., Schmitt, J.D., Jiao, Y., and Orth, D.J. 2023. Predatory impacts of invasive blue catfish in an Atlantic coast estuary. *Marine and Coastal Fisheries* 15:e210261.

Hines, A.H. 2007. Ecology of juvenile and adult blue crabs. Pages 565–654 in V.S. Kennedy and L.E. Cronin, editors. *The Blue Crab Callinectes sapidus*. Maryland Sea Grant College, College Park, MD.

Hines, A.H., A.M. Haddon, and L.A. Wiechert. 1990. Guild structure and foraging impact of blue crabs and epibenthic fish in a subestuary of Chesapeake Bay. *Marine Ecology Progress Series* 67:105–126.

Hines, A.H., P.R. Jivoff, P.J. Bushmann, J. van Montfrans, S.A. Reed, D.L. Wolcott, and T.G. Wolcott. 2003. Evidence for sperm limitation in the blue crab, *Callinectes sapidus*. *Bulletin of Marine Science* 72:287–310.

Hines, A.H., E.G. Johnson, A.C. Young, R. Aguilar, M.A. Kramer, M. Goodison, O. Zmora, and Y. Zohar. 2008. Release strategies for estuarine species with complex migratory life cycles: Stock enhancement of Chesapeake blue crabs (*Callinectes sapidus*). *Reviews in Fisheries Science* 16:175–185.

Hines, A.H., M.A. Kramer, K.D. Richie, L.D. Smith, R. Aguilar, M.B. Ogburn, E.G. Johnson, and G.M. Ruiz. 2026. Cannibalism, predation, and refuges of juvenile blue crabs from a 37-year experiment. *Proceedings of the National Academy of Sciences* 123:e2533461123.

Hyman, A.C., Chiu, G.S., Fabrizio, M.C., and Lipcius, R.N. 2022. Spatiotemporal modeling of nursery habitat using Bayesian inference: environmental drivers of juvenile blue crab abundance. *Frontiers in Marine Science* 9:834990.

Hyman, A.C., Chiu, G.S., Seebo, M.S., Smith, A., Saluta, G.G., and Lipcius, R.N. 2024. Ontogenetic patterns in juvenile blue crab density: effects of habitat and turbidity in a Chesapeake Bay tributary. *Marine Ecology Progress Series* 729:135–150.

Johnson, D.F. and K.W. Hess. 1990. Numerical simulations of blue crab larval dispersal and recruitment. *Bulletin of Marine Science* 46:195–213.

Johnson, E.G., Hines, A.H., Kramer, M.A., and Young, A.C. 2008. Importance of season and size of release to stocking success for the blue crab in Chesapeake Bay. *Reviews in Fisheries Science* 16:243–253.

Johnson, E.G., Young, A.C., Hines, A.H., Kramer, M.A., Bademan, M., Goodison, M.R., and Aguilar, R. 2011. Field comparison of survival and growth of hatchery-reared versus wild blue crabs, *Callinectes sapidus* Rathbun. *Journal of Experimental Marine Biology and Ecology* 402:35–42.

Karp, M.A., Kuriyama, P., Blackhart, K., Brodziak, J., Carvalho, F., Curti, K., Dick, E.J., Hanselman, D., Ianelli, J., Sagarese, S., Shertzer, K., and Taylor, I. 2022. Common model diagnostics for fish stock assessments in the United States. NOAA Technical Memorandum NMFS-F/SPO-240A.

Kristensen, K. 2025. RTMB: R Bindings for TMB. R package version 1.7. <https://CRAN.R-project.org/package=RTMB>.

Latour, R.J., Brush, M.J., and Bonzek, C.F. 2003. Toward ecosystem-based fisheries management. *Fisheries* 28:10–22.

Latour, R.J., Gartland, J., and Bonzek, C.F. 2023. Design and redesign of a bottom trawl survey in Chesapeake Bay, USA. *Frontiers in Marine Science* 10:1217792.

Liang, D., Nessler, G.M., Wilberg, M.J., and Miller, T.J. 2021. Ranking ecosystem impacts on Chesapeake Bay blue crab (*Callinectes sapidus*) using empirical Gaussian Graphical Models. *Canadian Journal of Fisheries and Aquatic Sciences* 78(3):245–254.

Liang, D., Bevans, A., Peng, E., Davis, G., Ihde, T., Lipcius, R., Saluta, G., Sholes, M., Tuckey, T., and Wilberg, M.J. 2026a. Length weight relationships of blue crab from the Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 2.

Liang, D., Davis, G., Drzewicki, M., Ralph, G., Sholes, M., Lipcius, R.N., Tuckey, T., and Wilberg, M.J. 2026b. Overwinter mortality of blue crabs in Chesapeake Bay using a Bayesian model. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 3.

Liang, D. 2026c. Development of daily temperature time series for the Bay, a 4D interpolation approach. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 4.

Lipcius, R.N. and M.S. Seebo. 2026. Blue crab spine width increases disproportionately at maturity, with consequences for demographic models. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 11.

Lipcius, R.N. and W.T. Stockhausen. 2002. Concurrent decline of the spawning stock, recruitment, larval abundance, and size of the blue crab *Callinectes sapidus* in Chesapeake Bay. *Marine Ecology Progress Series* 226:45–61.

Lipcius, R.N., Eggleston, D.B., Heck, K.L. Jr, Seitz, R.D., and van Montfrans, J. 2007. Ecology of postlarval and young juvenile blue crabs. Pages 535–564 in V.S. Kennedy and L.E. Cronin, editors. *The Blue Crab, Callinectes sapidus*. Maryland Sea Grant, College Park, MD.

Lipcius, R.N., McCulloch, D., and Seebo, M.S. 2013. Blue Crab Dredge Mortality Project Final Report. Virginia Institute of Marine Science, College of William and Mary.

Lipcius, R.N., Hoenig, J.M., Seebo, M.S., and Smith, A. 2026a. Tag-recapture estimates of survival of mature female blue crab *Callinectes sapidus*. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 10.

Lipcius, R.N., Hoenig, J.M., Lambert, D.M., Seebo, M.S., and Smith, A. 2026b. Longevity-based, tag-recapture estimation of natural mortality in mature female blue crabs (*Callinectes sapidus*) in Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 12.

Macedo, D., Caballero, I., Mateos, A., Leblois, R., McCay, S., and Hurtado, L.A. 2019. Population genetics and historical demographic inferences of the blue crab *Callinectes sapidus* in the US based on microsatellites. *PeerJ* 7:e7780.

McMillen-Jackson, A.L., Bert, T.M., and Steele, P. 1994. Population genetics of the blue crab *Callinectes sapidus*: modest population structuring in a background of high gene flow. *Marine Biology* 118:53–65.

McMillen-Jackson, A.L. and T.M. Bert. 2004. Mitochondrial DNA variation and population genetic structure of the blue crab *Callinectes sapidus* in the eastern United States. *Marine Biology* 145:769–777.

Miller, T.J. 2001a. Matrix-based modeling of blue crab population dynamics with applications to the Chesapeake Bay. *Estuaries* 24:535–544.

Miller, T.J. 2001b. The precautionary approach to managing blue crab in Chesapeake Bay: establishing limits and targets. UMCES Technical Series TS-340-01-CBL. University of Maryland Center for Environmental Science, Solomons, MD.

Miller, T.J. and E.D. Houde. 1999. Blue crab target setting. U.S. EPA – Chesapeake Bay Program, Annapolis, MD.

Miller, T.J., Martell, S.J.D., Bunnell, D.B., Davis, G., Fegley, L.A., Sharov, A.F., Bonzek, C.F., Hewitt, D.A., Hoenig, J.M., and Lipcius, R.N. 2005. Stock assessment of blue crab in Chesapeake Bay: 2005. UMCES Tech. Ser. No. TS-487-05-CBL. University of Maryland Center for Environmental Science, Solomons, MD.

Miller, T.J., Wilberg, M.J., Colton, A.R., Davis, G.R., Sharov, A., Lipcius, R.N., Ralph, G.M., Johnson, E.G., and Kaufman, A.G. 2011. Stock assessment of blue crab in Chesapeake Bay, 2011. University of Maryland Center for Environmental Science, Solomons, MD. 192 pp.

Miller, T.J., Davis, G., Liang, D., Drzewicki, M., Ralph, G., and Wilberg, M.J. 2026. Characterizing major changes in assumptions between the 2011 assessment model and the 2023 model. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 9.

Nelson, G.A. 2014. Cluster sampling: A pervasive, yet little recognized survey design in fisheries research. *Transactions of the American Fisheries Society* 143:926–938.

Ogburn, M.B., Richie, K.D., Jones, M.A., and Hines, A.H. 2019. Sperm acquisition and storage dynamics facilitate sperm limitation in the selectively harvested blue crab *Callinectes sapidus*. *Marine Ecology Progress Series* 629:87–101.

Ogburn, M.B., Roberts, P.M., Richie, K.D., Johnson, E.G., and Hines, A.H. 2014. Temporal and spatial variation in sperm stores in mature female blue crabs *Callinectes sapidus* and potential effects on brood production in Chesapeake Bay. *Marine Ecology Progress Series* 507:249–262.

Ogburn, M.B., Aguilar, R., Galvan Lam, Z., Hines, A.H., and Wilberg, M.J. 2026. Growth, maturation and operational sex ratio of female blue crabs in the Rhode River, Maryland, USA. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 7.

Prager, M.H., McConaughy, J.R., Jones, C.M., and Geer, P.J. 1990. Fecundity of blue crab, *Callinectes sapidus*, in Chesapeake Bay: Biological, statistical and management considerations. *Bulletin of Marine Science* 46:170–179.

Punt, A.E., Deng, R.A., Dichmont, C.M., Kompas, T., Venables, W.N., Zhou, S., Pascoe, S., Hutton, T., Kenyon, R., van der Velde, T., and Kienzle, M. 2010. Integrating size-structured assessment and bioeconomic management advice in Australia's northern prawn fishery. *ICES Journal of Marine Science* 67:1785–1801.

Punt, A.E., Huang, T., and Maunder, M.N. 2013. Review of integrated size-structured models for stock assessment of hard-to-age crustacean and mollusc species. *ICES Journal of Marine Science* 70:16–33.

Rains, S.M., Wilberg, M.J., and Miller, T.J. 2016. Sex ratios and average sperm per female blue crab in six tributaries of Chesapeake Bay. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science* 8:492–501.

Rains, S.M., Wilberg, M.J., and Miller, T.J. 2018. Evaluation of fishery-induced sperm limitation in Chesapeake Bay blue crab using an individual based model. *Marine Ecology Progress Series* 596:127–142.

Ralph, G.M. and R.N. Lipcius. 2014. Critical habitats and stock assessment: age-specific bias in the Chesapeake Bay blue crab population survey. *Transactions of the American Fisheries Society* 143:889–898.

- Ralph, G.M., Seitz, R.D., Orth, R.J., Knick, K.E., and Lipcius, R.N. 2013. Broad-scale association between seagrass cover and juvenile blue crab density in Chesapeake Bay. *Marine Ecology Progress Series* 488:51–63.
- Ralph, G.M., Gartland, J., and Latour, R.J. 2024. Evaluation of the Chesapeake Bay blue crab sanctuary through habitat suitability. *Marine Ecology Progress Series* 739:111–128.
- Ralph, G., Drzewicki, M.R., and Wilberg, M.J. 2026a. Estimation of monthly proportions of female blue crabs with eggs in Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 6.
- Ralph, G., Latour, R., and Wilberg, M.J. 2026b. Estimating size at maturity for male blue crabs in Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 5.
- Restrepo, V.R., Thompson, G.G., Mace, P.M., Gabriel, W.L., Low, L.L., MacCall, A.D., Methot, R.D., Powers, J.E., Taylor, B.L., Wade, P.R., and Witzig, J.F. 1998. Technical guidance on the use of precautionary approaches to implementing National Standard 1 of the Magnuson-Stevens Fishery Conservation and Management Act. NOAA Technical Memorandum NMFS-F/SPO-31. 54 pp.
- Rigby, R.A. and D.M. Stasinopoulos. 2005. Generalized additive models for location, scale and shape. *Applied Statistics* 54:507–554.
- Rome, M.S., Young-Williams, A.C., Davis, G.R., and Hines, A.H. 2005. Linking temperature and salinity tolerance to winter mortality of Chesapeake Bay blue crabs (*Callinectes sapidus*). *Journal of Experimental Marine Biology and Ecology* 319:129–145.
- Rugolo, L.J., Knotts, K., Lange, A., Crecco, V., Terceiro, M., Bonzek, C.F., Stagg, C., O'Reilly, R., and Vaughan, D.S. 1997. Stock assessment of Chesapeake Bay blue crab (*Callinectes sapidus*). National Oceanic and Atmospheric Administration. 267 pp.
- Schneider, A.K., Pomroy, A.K., and Shields, J.D. 2023a. Prevalence, intensity of infestation, and biomarker potential of the nemertean worm, *Carcinonemertes carcinophila*, on ovigerous blue crabs, *Callinectes sapidus*, in Chesapeake Bay, USA. *Journal of Crustacean Biology* 43:1–12.
- Schneider, A.K., Fabrizio, M.C., and Lipcius, R.N. 2023b. Reproductive potential of the blue crab spawning stock in Chesapeake Bay across eras and exploitation rates using nemertean worms as biomarkers. *Marine Ecology Progress Series* 716:77–91.
- Schneider, A.K., Shields, J.D., Fabrizio, M.C., and Lipcius, R.N. 2024a. Spawning history, fecundity, and potential sperm limitation of female blue crabs in Chesapeake Bay. *Fisheries Research* 278:107094.
- Schneider, A.K., Fabrizio, M.C., and Lipcius, R.N. 2024b. Reproductive phenology of the Chesapeake Bay blue crab population in a changing climate. *Frontiers in Ecology and Evolution* 11:1–15.
- Seitz, R.D., Lipcius, R.N., and Seebo, M.S. 2005. Food availability and growth of the blue crab in seagrass and unvegetated nurseries of Chesapeake Bay. *Journal of Experimental Marine Biology and Ecology* 319:57–68.
- Semmler, R.F., Ogburn, M.B., Aguilar, R., North, E.W., Reaka, M.L., and Hines, A.H. 2021. The influence of blue crab movement on mark-recapture estimates of recreational harvest and exploitation. *Canadian Journal of Fisheries and Aquatic Sciences* 78:371–385.
- Sharov, A.F., Volstad, J.H., Davis, G., Davis, B.K., Lipcius, R.N., and Montane, M.M. 2003. Abundance and exploitation rate of the blue crab (*Callinectes sapidus*) in Chesapeake Bay. *Bulletin of Marine Science* 72:543–565.
- Shertzer, K.W., Damiano, M.D., and Williams, E.H. 2024. Spawning potential ratio can provide reference points for fishery management that are robust to environmental variability. *Fishes* 9:497.
- Sholes, M., Fabrizio, M.C., Tuckey, T.D., Davis, G., Liang, D., and Wilberg, M.J. 2026. Effects of blue catfish on blue crab abundance in Chesapeake Bay tributaries during 1990–2023. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 1.
- Smith, S.G. 1997. Models of crustacean growth dynamics. Ph.D. Dissertation. University of Maryland, College Park.
- Smith, S.G. and E.S. Chang. 2007. Molting and growth. Pages 197–254 in V.S. Kennedy and L.E. Cronin, editors. *The Blue Crab Callinectes sapidus*. Maryland Sea Grant, College Park, MD.
- Szuwalski, C.S., Aydin, K., Fedewa, E.J., Garber-Yonts, B., and Litzow, M.A. 2023. The collapse of eastern Bering Sea snow crab. *Science* 382:306–310.
- Tagatz, M.E. 1968. Biology of the blue crab, *Callinectes sapidus* Rathbun, in the St. Johns River, Florida. *Fishery Bulletin* 67:17–33.
- Thorson, J.T. 2019. Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research* 210:143–161.
- Thorson, J.T. and L.A.K. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* 74:1311–1321.
- Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science* 72(5):1297–1310.
- Truesdell, S.B., Hart, D.R., and Chen, Y. 2017. Effects of unequal capture probability on stock assessment abundance and mortality estimates: an example using the US Atlantic sea scallop fishery. *Canadian Journal of Fisheries and Aquatic Sciences* 74:1904–1917.
- Van Engel, W.A. 1958. The blue crab and its fishery in Chesapeake Bay. Part 1 – reproduction, early development, growth and migration. *Commercial Fisheries Review* 20:6–17.

- Van Engel, W.A. 1990. Development of the reproductively functional form in the male blue crab, *Callinectes sapidus*. *Bulletin of Marine Science* 46:13–22.
- Vølstad, J.H., Sharov, A.F., Davis, G., and Davis, B. 2000. A method for estimating dredge catching efficiency for blue crabs, *Callinectes sapidus*, in Chesapeake Bay. *Fishery Bulletin* 98:410–420.
- Walters, E.A., Crowley-McIntye, C.E., Gandy, R.L., Martin, C.W., Shea, C.P., and Behringer, D.C. 2025. Estimating discard mortality in Florida's blue crab *Callinectes sapidus* fishery using reflex impairment, injuries, and salinity. *Marine and Coastal Fisheries* 17. <https://doi.org/10.1093/mcfafs/vtae001>.
- Wilberg, M.J. and M.R. Drzewicki. 2026. Estimation of the size distribution of age-0 blue crabs in Chesapeake Bay and Winter Dredge Survey selectivity. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 8.
- Wilberg, M.J., Robinson, J.M., Rains, S.A.M., Humphrey, J.L., and Lipcius, R.N. 2013. Effects of location errors on estimates of dredge catchability from depletion based methods. *Fisheries Research* 148:1–8.
- Wilberg, M.J., Woodward, R.T., Huang, P., and Tomberlin, D. 2019. Developing precautionary reference points for fishery management using robust control theory: application to the Chesapeake Bay blue crab fishery. *Marine and Coastal Fisheries* 11:177–188.
- Wood, S.N. 2017. *Generalized Additive Models: An Introduction with R* (2nd edition). Chapman and Hall/CRC Press.
- Wood, M.A. and R.N. Lipcius. 2022. Non-native red alga *Gracilaria vermiculophylla* compensates for seagrass loss as blue crab nursery habitat in the emerging Chesapeake Bay ecosystem. *PLoS ONE* 17:e0267880.
- Zhao, M., Plough, L.V., Behringer, D.C., Bojko, J., Kough, A.S., Alper, N.W., Xu, L., and Schott, E.J. 2023. Cross-hemispheric genetic diversity and spatial genetic structure of *Callinectes sapidus* Reovirus 1 (CsRV1). *Viruses* 15(2):563.
- Zmora, O., Findiesen, A., Stubblefield, J., Frenkel, V., and Zohar, Y. 2005. Large-scale juvenile production of the blue crab *Callinectes sapidus*. *Aquaculture* 244:129–139.
- Zmora, N., Trant, J., Zohar, Y., and Chung, J.S. 2009. Molt-inhibiting hormone stimulates vitellogenesis at advanced ovarian developmental stages in the female blue crab, *Callinectes sapidus* 1: an ovarian stage dependent involvement. *Saline Systems* 5:7.

Appendix 1: Bibliography of materials provided for review

Materials provided for the 2025 Chesapeake Bay blue crab stock assessment review

The 2025 Chesapeake Bay blue crab stock assessment report

Wilberg, M. J., Maya Drzewicki, Gina M. Ralph, Dong Liang, Madison Sholes, Rober Aguilar, Amanda Bevans, Ingrid Braun-Ricks, Mandy Bromilow, Glenn R. Davis, Alexa Galvan, Thomas Ihde, Robert Latour, Romuald. N. Lipcius, Brooke Lowman, Marvin M. Mace III, Thomas J. Miller, Matthew B. Ogburn, Gabrielle Saluta, Alexandra K. Schneider, Alexei Sharov, and Troy D. Tuckey. Stock Assessment of Blue Crab in Chesapeake Bay, Draft Assessment Report. March 11, 2026

The 2011 Chesapeake Bay blue crab stock assessment report

Miller, T.J., Wilberg, M.J., Colton, A.R., Davis, G.R., Sharov, A., Lipcius, R.N., Ralph, G.M., Johnson, E.G., Kaufman, A.G., 2011. Stock assessment of blue crab in Chesapeake Bay, 2011. University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD, p. 192.

Working papers

Sholes, M., M.C. Fabrizio, T.D. Tuckey, G. Davis, D. Liang., M.J. Wilberg. 2026. Effects of Blue Catfish on Blue Crab Abundance in Chesapeake Bay Tributaries during 1990-2023. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 1.

Liang, D., A. Bevans, E. Peng, G. Davis, T. Ihde, R. Lipcius, G. Saluta, M. Sholes, T. Tuckey, M.J. Wilberg. 2026a. Length Weight Relationships of Blue Crab from the Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 2.

Liang, D., G. Davis, M. Drzewicki, G. Ralph, M. Sholes, R.N. Lipcius, T. Tuckey, M.J. Wilberg. 2026b. Overwinter Mortality of Blue Crabs in Chesapeake Bay Using a Bayesian Model. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 3.

Liang, D. 2026. Development of Daily Temperature Time Series for the Bay, a 4D Interpolation Approach. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 4.

Ralph, G., R. Latour, M.J. Wilberg. 2026b. Estimating size at maturity for male blue crabs in Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 5.

Ralph, G., M.R. Drzewicki, M.J. Wilberg. 2026a. Estimation of monthly proportions of female blue crabs with eggs in Chesapeake Bay. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 6.

Ogburn, M.B., R. Aguilar, Z. Galvan Lam, A.H. Hines, M.J. Wilberg. 2026. Growth, Maturation and Operational Sex Ratio of Female Blue Crabs in the Rhode River, Maryland, USA. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 7.

Wilberg, M.J., M.R. Drzewicki. 2026. Estimation of the size distribution of age-0 blue crabs in Chesapeake Bay and Winter Dredge Survey Selectivity. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 8.

Miller, T.J., G. Davis, D. Liang, M. Drzewicki, G. Ralph, M.J. Wilberg. 2026. Characterizing major changes in assumptions between the 2011 assessment model and the 2023 model. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 9.

Lipcius, R.N., J.M. Hoenig, M.S. Seebo, A. Smith. 2026a. Tag-recapture estimates of survival of mature female blue crab *Callinectes sapidus*. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 10.

Lipcius, R.N., M.S. Seebo. 2026. Blue crab spine width increases disproportionately at maturity, with consequences for demographic models. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 11.

Lipcius, R.N., J.M. Hoenig, D.M. Lambert, M.S. Seebo, A. Smith. 2026b. Longevity-based, tag-recapture estimation of natural mortality in mature female blue crabs (*Callinectes sapidus*) in Chesapeake Bay. Working Paper. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 12.

Drzewicki, M.R., M.J. Wilberg. 2026. Additional plots for the 2026 Chesapeake Bay blue crab assessment. 2026 Chesapeake Bay Blue Crab Stock Assessment Working Paper 13.

Appendix 2: Performance Work Statement

**Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
NOAA Fisheries
Office of Science & Technology
Center for Independent Experts (CIE)
External Independent Peer Review
Benchmark Quantitative Fishery Stock Assessment for Blue Crab in the
Chesapeake Bay (2023-2025)**

Background

The NOAA Fisheries is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation’s marine living resources based upon the best scientific information available (BSIA). NOAA Fisheries science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards².

Scope

The NOAA Chesapeake Bay Office requests an independent peer review for the 2025 “Benchmark Quantitative Fishery Stock Assessment for Blue Crab in the Chesapeake Bay” (blue crab stock assessment). This updated stock assessment will provide a critical science product at a time of partnership-wide strategic planning for the Chesapeake Bay Program and to inform science-based management of the Chesapeake Bay blue crab fishery.

In the Chesapeake Bay, the blue crab (*Callinectes sapidus*) supports economically and culturally important commercial fisheries. Diverse gears are used to harvest blue crabs

² https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

regionally and seasonally, and many of these were developed in response to spatiotemporal variability in the presence of specific life stages. The previous benchmark blue crab stock assessment for the Chesapeake Bay was completed in 2011, and developed female biological reference points based on maximum sustainable yield (MSY). Following the 2011 assessment, management measures were implemented to reach the reference points and target exploitation rate for adult female crabs. Despite management successfully achieving the reference points and target exploitation rates, the population has not responded as anticipated in recent years. Specifically, recruitment has been lower than expected in the most recent decade. Improved understanding of continued low recruitment has been increasingly called for, and a new benchmark stock assessment was requested to reevaluate reference points for sustainable management. Potential causes of low recruitment include increased predation of juvenile blue crabs by finfish predators, changes in the ocean habitat and currents for larval blue crabs, reduced numbers of female spawners due to spring fishing for female blue crabs, and sperm limitation caused by too few adult male blue crabs.

Requirements

NOAA Fisheries requires **three (3)** reviewers to conduct an impartial and independent peer review in accordance with the Performance Work Statement (PWS), OMB guidelines, and the CIE Terms of Reference (ToRs) for reviewers listed below. The CIE reviewers shall have excellent communication skills in addition to working knowledge in fisheries stock assessment and modeling, statistics, fisheries science, and marine biology sufficient to complete the primary task of providing peer-review advice in compliance with the ToRs for the blue crab stock assessment. Preferably, reviewers will have specific expertise in crustacean fisheries. Additionally, the Chair, who is in addition to two CIE reviewers, will also be participating in this review and will be provided by the NOAA Chesapeake Bay Office. The Chair's participation (e.g., labor and travel) is not covered under this contract.

Terms of Reference for CIE Reviewers:

Each CIE reviewer shall complete the following tasks in accordance with the PWS and Schedule of Milestones and Deliverables herein.

- 1. Pre-review Background Documents:** A minimum of two weeks before the peer review, the NOAA Fisheries Project Contact will send by electronic mail or make available at an FTP site to the CIE reviewers all necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NOAA Fisheries Project Contract will consult with the Contractor on where to send the documents. The CIE reviewers shall read all documents in preparation for the peer review
- 2. Virtual Conferencing Test:** Additionally, one or two weeks prior to the peer review, the CIE reviewers will participate in a test to confirm that they have the necessary technical (hardware, software, etc.) capabilities to participate in the virtual panel in advance of the review meeting. This review's NOAA Fisheries Project Contact will provide the information for the arrangements for this test.

The NOAA Fisheries Project Contact is responsible for facility arrangement (e.g., video or teleconference arrangements). This virtual conferencing test will also serve as a check in on the panel review agenda, ensure reviewers have all necessary materials and address clarifying questions.

- 3. Attend and Participate in the Virtual Panel Review Meeting:** Each CIE reviewer shall conduct the independent peer review in accordance with the PWS and CIE ToRs for reviewers, and shall not serve in any other role unless, specified herein. Modifications to the PWS and the CIE ToRs for reviewers cannot be made during the peer review, and any PWS or CIE ToRs modifications prior to the peer review shall be approved by the Contracting Officer’s Representative (COR) and the CIE contractor. Each reviewer shall actively participate in a professional and respectful manner as a member of the review panel’s meeting. And their peer review tasks shall be focused on the CIE ToRs as specified herein.

- 4. Contract Deliverables – Independent CIE Peer Review Reports:** Reviewers are not required to reach a consensus. Each CIE reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each CIE reviewer shall complete the independent peer review addressing each of the questions in **Annex 2**. A draft agenda is provided in **Annex 3**. Each CIE reviewer will deliver their reports according to the specified milestone dates. Each reviewer shall assist the Chair of the meeting with contributions to the summary report.

Place of Performance

The place of performance shall be online.

Period of Performance

Each reviewer’s duties shall not exceed 14 days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule:

Within two weeks of award	Contractor selects and confirms reviewers
Two weeks prior to review	Contractor provides the pre-review documents to the reviewers
April 13-16, 2026	Panel review meeting
Within two weeks after review	Reviewers submit draft peer-review reports to the contractor for quality assurance and review
Within three weeks of receiving draft reports	Contractor submits final reports to the Government

*The Chair’s Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards

The acceptance of the contract deliverables shall be based on three performance standards:

- (1) The reports shall be completed in accordance with the required formatting and content
- (2) The reports shall address each of the questions under the CIE ToR for reviewers as specified
- (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Confidentiality and Data Privacy

This contract may require that services contractors have access to Privacy Information. Service contractors are responsible for maintaining the confidentiality of all subjects and materials and may be required to sign and adhere to a Non-disclosure Agreement (NDA).

Travel

Travel is not expected or authorized for this review.

NOAA Fisheries Project Contact:

Bruce Vogt

Ecosystem Science and Synthesis Manager

NOAA Chesapeake Bay Office

200 Harry S Truman Parkway, Annapolis, MD 21401

bruce.vogt@noaa.gov

Annex 1: Peer Review Report Requirements

1. The report must be prefaced with an Executive Summary providing a concise summary of the findings and recommendations and specify whether the stock assessment science reviewed is adequate for use in management.

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, and a summary of findings for each question under the ToRs (*strengths and weaknesses, conclusions and recommendations for changes or improvements*).

a. Reviewers must describe in their own words the review activities completed during the panel review meeting, including a brief summary of findings, of the science, conclusions, and recommendations.

b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, but especially where there were divergent views.

c. Reviewers should elaborate on any points raised in the summary report they believe might require further clarification.

d. Reviewers shall provide a critique of the NOAA Fisheries review process, including suggestions for improvements of both process and products.

e. The report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The report shall represent the peer review of each question under the ToR section, and shall not simply repeat the contents of the summary report.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review

Appendix 2: A copy of this Performance Work Statement

Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Benchmark Quantitative Fishery Stock Assessment for Blue Crab in the Chesapeake Bay (2023-2025)

CIE reviewers are contracted to complete their independent peer review based on the ToRs. Therefore, the CIE-NOAA Fisheries review and approval process is based on whether the CIE independent reports addressed each of the ToRs. The specific responsibilities of the CIE review are to provide detailed written responses to the questions below as part of their review. Responses to each of these questions should include strengths and weaknesses, conclusions and recommendations for changes or improvements.

- 1. Does the assessment critically review and estimate life history parameters and vital rates of blue crab in the Chesapeake Bay that are relevant to the stock assessment? In particular, does the assessment evaluate the extent and scale of interannual variation in life history parameters and vital rates of blue crab in the Chesapeake Bay?*
- 2. Does the assessment describe and quantify patterns in fishery-independent surveys to develop indices of abundance and characterize the size composition of the population?*
- 3. Does the assessment describe and quantify patterns in catch, effort, and CPUE?*
- 4. Does the assessment evaluate the feasibility of, and if possible, implement blue crab stock assessment models that operate on sub-annual time steps and/or at spatial resolutions lower than that of the entire Chesapeake Bay to better represent population dynamics?*
- 5. Does the assessment characterize uncertainty in assessment estimates (mortality and abundance)?*
- 6. Does the assessment update the sex-specific catch survey models used in the 2011 benchmark stock assessment with relevant new data? Does the assessment characterize major changes in assumptions between the 2011 assessment model and the 2023 model?*
- 7. Does the assessment recommend appropriate biological reference points for management? Does the assessment evaluate stock status relative to recommended reference points?*
- 8. Does the assessment identify relevant ecosystem and climate influences (such as habitat, environmental drivers, prey availability, and predation/cannibalism) on the population dynamics and fisheries and explore other analyses that support the assessment?*
- 9. Does the assessment report on the status of research recommendations from the most recent benchmark assessment? Does the assessment identify and prioritize research recommendations for future work?*
- 10. Based on your full review, is the 2025 Blue Crab Stock Assessment adequate for informing management decisions?*

Annex 3: Tentative Agenda
Chesapeake Bay Blue Crab Stock Assessment Review
Annapolis, MD

2026 Chesapeake Bay Blue Crab Stock Assessment Review
April 13 - 16, 2026
Meeting Link: Join the meeting via Microsoft Teams
Meeting ID: 211 216 057 105 70 Passcode: wB2JP7eM
Call:+1 202-991-0477 Conference ID: 486625540#

Visit the meeting webpage for meeting materials and additional information.
This meeting may be recorded for internal use only to assure the accuracy of meeting notes. To turn on closed captioning, click on the three ellipses (More actions), then click on “Turn on live captions” (preview). To request accommodations, please contact Christina Garvey (christina.garvey@noaa.gov).

Please read the following information carefully, as our meeting policies have changed:

- All meeting attendees' cameras and microphones will be muted at the start of the meeting.
- To request access to the microphone and camera, all meeting participants will be required to use the raised hand feature on Teams. Once access has been granted by the meeting organizer, you will then be allowed to unmute your mic and turn on your camera. Unless instructed otherwise, once a participant has microphone or camera access, they will have this permission for the remainder of the meeting.
- Access to chat will be provided as well. Should it be necessary, the Q&A feature on Teams will be utilized to field participant questions.

Compromised Meeting Plan: If the meeting's privacy is compromised, the meeting staffer and coordinator will send an email to all Members, alternates, staffers, coordinators, and interested parties. Within the email, you will find a new meeting link, instructions on sharing this information with external partners, and any necessary adjustments to the meeting schedule. Please do NOT share this information publicly or post it to the Chesapeakebay.net webpage.

Purpose: This meeting will convene modeling experts, state jurisdictions, and independent reviewers to evaluate the draft 2026 Chesapeake Bay Blue Crab Stock Assessment.

Agenda

Day 1: April 13, 2026

I. Welcome & Introductions 12:00pm – 12:15pm

Brief introductions of the Chesapeake Bay Blue Crab Stock Assessment Committee & the review panel

Requested Action: Non-decisional

Lead: Dan Hennen (NOAA)

Materials: N/A

II. Presentation of the 2026 Blue Crab Stock Assessment 12:15pm – 3:30pm

Review of TORs 4, 5, 7 and 8

Requested Action: Non-decisional

Lead: Michael Wilberg (UMCES)

Materials: Draft 2026 Blue Crab Stock Assessment Report

III. General / Open Question Period 3:30pm - 5:00pm

Opportunity for public comment & questions from the review panel

Requested Action: Non-decisional

Lead: Dan Hennen (NOAA)

Materials: Draft 2026 Blue Crab Stock Assessment Report
IV. Adjourn 5:00pm

Day 2: April 14, 2026

I. CIE Term of Reference Review and Discussion 9:00am – 1:00pm

Review of TORs 2 & 3
Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

II. Lunch 1:00pm – 2:00pm

III. CIE Term of Reference Review and Discussion (continued) 2:00pm - 5:00pm

Review of TORs 1, 6, 9
Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

IV. Adjourn 5:00pm

Day 3: April 15, 2026

I. CIE Term of Reference Review & Discussion 9:00am - 10:00am

Review of TORs 10 & 11
Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

II. Review Session [closed-door] 10:00am - 12:00pm

Opportunity for the review panel to discuss assessment methodologies and develop individual opinions. The review panel will initiate development of the panel summary report.

Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

III. Lunch 12:00pm - 1:15pm

IV. Review Session (continued) [closed-door] 1:15pm - 3:00pm

Opportunity for the review panel to discuss assessment methodologies and develop individual opinions. The review panel will initiate development of the panel summary report.

Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

V. Plenary Session 3:00pm - 4:30pm Review panel discloses decision on TORs.

Requested Action: Decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report

VI. Adjourn 4:30pm

Day 4: April 16, 2026

I. Review Session [closed-door] 9:00am - 12:00pm

Opportunity for the review panel to continue to discuss assessment methodologies and develop individual opinions. The review panel will continue the development of summary documents and present the summary report. Next steps and timeline for finalizing the summary report will be identified.

Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report
II. Lunch 12:00pm - 1:15pm

III. Review Session (continued) [closed-door] 1:15pm - 4:30pm

Opportunity for the review panel to continue to discuss assessment methodologies and develop individual opinions. The review panel will continue the development of summary documents and present the summary report. Next steps and timeline for finalizing the summary report will be identified.

Requested Action: Non-decisional
Lead: Dan Hennen (NOAA)
Materials: Draft 2026 Blue Crab Stock Assessment Report
IV. Adjourn 4:30pm

Appendix 3: CIE review panel, assessment team and review attendees

Review Panel:

- Dan Hennen (NOAA)
- Yong Chen (Stony Brook University/CFE)
- Nick Caputi (Western Australian Fisheries and Marine Research Laboratories, DPIRD)
- Simon Delestang (Department of Primary Industries and Regional Development (DPIRD), Australia)

Assessment Team:

- Michael Wilberg (UMCES CBL)
- Maya Drzewicki (UMCES CBL)
- Madison Sholes (UMCES CBL)
- Dong Liang (UMCES CBL)
- Gina Ralph (VIMS)
- Gabrielle Salula (VIMS)
- Rom Lipcius (VIMS)
- Ingrid Braun-Ricks (PRFC)
- Alexei Sharov (MDNR)
- Trey Mace (MDNR)
- Mandy Bromilow (MDNR)
- Amanda Bevans (MSU PEARL)
- Tom Ihde (MSU PEARL)
- Rob Latour (VIMS)
- Alexa Galvan (VMRC)
- Brooke Lowman (VMRC)
- Glenn Davis (MDNR)
- Matt Ogburn (SERC)
- Rob Aguilar (SERC)
- Tom Miller (UMCES)
- Troy Tuckey (VIMS)

Attendees:

Day 1:

Dan Hennen (NOAA), Michael Wilberg (UMCES CBL), Bruce Vogt (NOAA), Christina Garvey (CRC), Nick Staten (CRC), Madison Sholes (UMCES CBL), Dong Liang (UMCES CBL), Gina Ralph (VIMS), Gabrielle Salula (VIMS), Rom Lipcius (VIMS), Carrie Kennedy (MDNR), Maya Drzewicki (UMCES CBL), Yong Chen (Stony Brook University/CFE), Ingrid Braun-Ricks (PRFC), Alexei Sharov (MDNR), Trey Mace (MDNR), Mandy Bromilow (MDNR), Amanda Bevans (MSU PEARL), Tom Ihde (MSU PEARL), Rob Latour (VIMS), Alexa Galvan (VMRC), Brooke Lowman (VMRC), Matt Ogburn (SERC), Rob Aguilar (SERC), Tom Miller (UMCES), Troy Tuckey (VIMS), Nick Caputi (DPIRD), Simon Delestang (DPIRD), William Stoker (MDNR), Amy Schueller (NOAA), Sharon McBreen (PCT), Danielle Kozkinski (MBA), Matthew Zink (NCDEQ), Holly Funkhouser (SC DNR), CJ Schick (SC DNR), Jeff Brunson (SC DNR), Graham Wagner (SC DNR), Ron Owens (PRFC)

Day 2:

Dan Hennen (NOAA), Michael Wilberg (UMCES CBL), Bruce Vogt (NOAA), Christina Garvey (CRC), Nick Staten (CRC), Madison Sholes (UMCES CBL), Dong Liang (UMCES CBL), Gina Ralph (VIMS), Gabrielle Salula (VIMS), Rom Lipcius (VIMS), Carrie Kennedy (MDNR), Maya Drzewicki (UMCES CBL), Yong Chen (Stony Brook University/CFE), Ingrid Braun-Ricks (PRFC), Alexei Sharov (MDNR), Trey Mace (MDNR), Mandy Bromilow (MDNR), Amanda Bevans (MSU PEARL), Tom Ihde (MSU PEARL), Rob Latour (VIMS), Alexa Galvan (VMRC), Brooke Lowman (VMRC), Matt Ogburn (SERC), Troy Tuckey (VIMS), Nick Caputi (DPIRD), Simon Delestang (DPIRD), William Stoker (MDNR),

Matthew Zink (NCDEQ), CJ Schick (SC DNR), Ron Owens (PRFC), Adriana Celtruda (UMD), Matthew Damiano (NC DMF)

Day 3:

Dan Hennen (NOAA), Michael Wilberg (UMCES CBL), Bruce Vogt (NOAA), Christina Garvey (CRC), Nick Staten (CRC), Madison Sholes (UMCES CBL), Dong Liang (UMCES CBL), Gina Ralph (VIMS), Gabrielle Salula (VIMS), Rom Lipcius (VIMS), Carrie Kennedy (MDNR), Maya Drzewicki (UMCES CBL), Yong Chen (Stony Brook University/CFE), Ingrid Braun-Ricks (PRFC), Trey Mace (MDNR), Mandy Bromilow (MDNR), Amanda Bevans (MSU PEARL), Tom Ihde (MSU PEARL), Rob Latour (VIMS), Alexa Galvan (VMRC), Brooke Lowman (VMRC), Matt Ogburn (SERC), Troy Tuckey (VIMS), Nick Caputi (DPIRD), Simon Delestang (DPIRD), Heather Hayden (MDNR), William Stoker (MDNR), Tom Miller (UMCES)