

Review of 2021 ASMFC ARM revision

Kevin Shoemaker, Ph.D.,

Associate Professor, University of Nevada, Reno

September 2022

This is an expert review of the Adaptive Resource Management plan (ARM) proposed by ASMFC to guide management of the Horseshoe Crab fishery in Delaware Bay, performed by Kevin Shoemaker, Ph.D.

Dr. Shoemaker has a Ph.D. in Conservation Biology, a Master of Science degree in Conservation Biology, both from SUNY-ESF in Syracuse, NY, and a Bachelor of Science degree in Biology from Haverford College. He is a former Postdoctoral Fellow in the Department of Ecology and Evolution at Stony Brook University and a former Senior Scientist at Applied Biomathematics, an ecological research and development company located in Setauket, NY. Dr. Shoemaker is currently employed as an Associate Professor of Population Ecology at the University of Nevada, Reno. He has over 15 years of experience as a wildlife conservation scientist and has authored over 45 peer-reviewed scientific articles and book chapters on topics in wildlife ecology and conservation. He has expertise in Bayesian inference, machine learning, population ecology, and ecological modeling.

OVERVIEW

This report presents my review of the Adaptive Resource Management plan (ARM) proposed for use by the Atlantic States Marine Fisheries Commission (ASMFC) as a tool for guiding management of the horseshoe crab (HSC) fishery in Delaware Bay and protecting the Federally Threatened *Rufa* Red Knot (*Calidris canutus rufa*; REKN). Delaware Bay is a critical stopover site for REKN in their spring migration to breeding grounds in the high arctic from wintering grounds as far south as Tierra del Fuego (USFWS 2021). Specifically, HSC eggs deposited on coastal beaches provide a necessary high-calorie food resource for REKNs and other migrating shorebird

species as they replenish fat reserves depleted from their long migration and prepare for breeding. At the heart of the proposed ARM framework is an optimization model that provides harvest recommendations for female and male HSC, conditional on current estimates of HSC and REKN abundance. These recommendations are calibrated to maximize HSC harvest while causing minimal risk to the REKN population. The optimization model is based on a linked two-species simulation model (comprising a HSC and a REKN simulation model) that incorporates a one-way biotic interaction in which annual REKN survival and recruitment depend on female HSC abundance in Delaware Bay (among other covariates). While the stated objectives of the revised ARM are sensible, my review identified several concerns that suggest the revised ARM framework is not an appropriate tool for managing risk to HSC or REKN populations. Specifically, this report identifies six main areas of concern:

- (1) **The fitted relationship between HSC abundance and REKN vital rates (survival and fecundity) is of insufficient magnitude to forecast a decline in mean projected REKN population growth even under a total collapse of the HSC population.** The extremely weak REKN/HSC relationship used in the revised ARM is inconsistent with previous research documenting HSC eggs as a critical food resource for migrating REKN and with the documented decline of the REKN population over recent decades, which experts have linked to increases in HSC bait harvest during the 1990s (Niles et al. 2009; USFWS 2014). If the REKN population model is inconsistent with what has been observed in the recent past, it seems unlikely to yield robust forecasts of future risk to the REKN population (or recovery of this population) from which to base management decisions. The inclusion of a REKN population model within the ARM framework (both the initial and revised versions) presupposes that HSC harvest could put REKN populations at risk, at least under some scenarios. As it stands, the apparent inability of the revised ARM model to predict a decline of the REKN population even under a total collapse of the HSC population seems to violate this premise, and practically guarantees that the REKN population model will play an insignificant role in setting optimal HSC harvest rates.
- (2) **The HSC population simulation model fails to correctly propagate uncertainty about mean recruitment rates.** In specifying the bivariate normal distribution used to generate

annual male and female HSC recruitment rates (the most consequential empirically fitted parameters of the HSC simulation model), the proposed ARM framework treats uncertainty about annual recruitment rates as representative of temporal process variance (natural year-to-year fluctuations) rather than as a mixture of parameter uncertainty and process variance (Link and Nichols 1994; Regan et al. 2002; McGowan et al. 2011). This subtle but significant shortcoming will tend to manifest in simulation replicates that closely resemble one another, since key sources of uncertainty “regress to the mean” (good years cancel out bad years) instead of propagating over time. The importance of this distinction is magnified for long-lived iteroparous species like HSC, since these populations tend to be resilient to short-term fluctuations in reproduction or recruitment (Lovich et al. 2015). When this issue is corrected (using the same Bayesian approach used to treat process variation and uncertainty in the REKN simulation models in the revised ARM framework), preliminary simulation results suggest a highly uncertain outlook for the HSC population in Delaware Bay, especially when faced with harvest pressures. In sharp contrast to the ARM report and supplement, the population of HSCs in Delaware Bay appears to have a substantial (17.5%) probability of falling below the lowest previously estimated levels even in the absence of all direct anthropogenic sources of mortality (bait harvest, biomedical bleeding and discard mortality) over the next 50 years. Furthermore, a scenario in which HSCs are harvested annually at the current maximum allowable rates is accompanied by a severe risk of decline (33.45%) and disruption to the population age structure (lower multiparous/primiparous ratios than previously observed). Finally, an extreme harvest scenario in which two million male and female HSCs are harvested each year results in near-certain catastrophic population collapse over the 50-year time horizon, in contrast to the (original) ARM report, which suggests a relatively stable HSC population even under this extreme scenario (which greatly exceeds current maximum allowable rates).

- (3) **The Catch Multiple Survey Analysis (CMSA) exhibits poor fit to training and independent data, raising concerns about its use in projecting future HSC abundance.** Aside from being able to explain the apparent difference in mean HSC abundance before and after the “VT gap years” (see below; higher HSC abundance is both predicted and observed after the

period 2013-2016), the CMSA model explains very little, if any, of the observed variation in the primary data sources (three trawl surveys conducted in and around Delaware Bay). The CMSA results exhibit relatively good fit ($R^2 > 0.5$) to the recruitment data (primiparous abundance); however, this is unsurprising since there is only one source of data (VT swept area surveys) for estimating annual primiparous abundance versus three sources for estimating adult (multiparous) and total abundance. Given the overall lack of fit to training data, the HSC simulation model is unlikely to perform well for predicting independent validation data (data not used to fit the model). Indeed, when the CMSA results are challenged against the HSC spawning surveys – an independent estimate of HSC abundance for this region – there is no detectable relationship between these two independent estimates of HSC abundance. This lack of fit to both training and validation data raises concerns about the utility of the CMSA model, which informs all aspects of the proposed ARM, including the REKN IPM (where it represents the abundance of female HSC each year), the HSC projection model, and the annual harvest recommendation.

(4) **The “gap years” in the VT trawl survey data raise concerns about HSC recruitment estimates from the Catch Multiple Survey Analysis (CMSA).** As noted above, the CMSA is fundamental to all aspects of the proposed ARM framework. For the HSC population simulation models, the primary role of the CMSA is to parameterize HSC recruitment rates (which are the most consequential empirically derived inputs for the HSC simulation model). Unfortunately, of the three trawl surveys used to fit the CMSA models, the only survey that provides information for estimating recruitment – the Virginia Tech (VT) trawl surveys – was not conducted during a critical four-year period from 2013 to 2016 (hereafter referred to as the “VT gap”, during which no direct information was available for estimating annual HSC recruitment rates). The CMSA results suggest that the HSC population underwent a substantial state transition during the VT gap years in which the population was small but stable prior to the gap, and larger and more variable after the gap. More concerning, the CMSA predicts much higher average recruitment rates during the VT gap (for which no data are available for estimating recruitment) than at any single year before or after. The inflated average recruitment rates during the VT gap period are subsequently used for estimating

mean HSC recruitment rate for the HSC simulation models (thereby increasing estimated population resilience to harvest) – but unfortunately these high recruitment rates cannot be verified empirically. If average recruitment rates were computed from only those years in which recruitment could be verified empirically (i.e., excluding estimates from the VT gap years) the expected resilience of the HSC population to harvest would be substantially reduced.

(5) **The proposed ARM framework lacks ‘null model’ benchmarks and independent performance validation.** Null models are simplified representations of a system that lack many or all the proposed mechanisms that may help to explain the system dynamics; the typical null model in statistics assumes all observed variation is the result of a single random error process. By comparing complex models such as those used in the revised ARM with one or more null-model benchmark(s), researchers can determine whether the more complex models represent useful learned knowledge about a system (Koons et al. 2022). If a complex model fails to outperform a null model in terms of bias or precision (typically using independent validation data), the complex model is likely to be improperly specified or “overfitted” (whereby parameters are fitted to “noise” rather than true signal; Radosavljevic and Anderson 2014) and therefore not useful for prediction. The CMSA model fails to outperform even the simplest statistical null model (single intercept term with sampling error) for at least one data source (the VT swept-area estimate of female multiparous abundance). For the REKN component of the revised ARM, it would be informative to compare the performance of the REKN simulation model against a null model that omits any effect of female HSC abundance. It was recently demonstrated (Koons et al. 2022) that the ARM framework for guiding North American mallard harvest was unable to outperform a null model, and it would be instructive to pose a similar challenge to the REKN simulation model. If either model fails to outperform a null model, it should prompt managers to acknowledge that our current understanding of the effects of harvest on HSC populations remains insufficient for robust forecasting (Dietze 2017), and that a more precautionary approach may be warranted.

(6) **Lack of transparency.** The public still has no access to the data and code used for estimating

REKN population parameters, simulating REKN and HSC population dynamics, and running optimization routines (the CMSA code and data were made available). Without this data and code, it is difficult to fully assess the proposed ARM framework and to run scenario tests. If granted access to the code and data, there are a number of important null model tests (see above) and scenario tests that can be run, including (1) developing and testing the HSC and REKN models against a “null model” benchmark, (2) determining the ‘optimal’ female HSC harvest rates from the “canonical” versions of the HSC and REKN models in the absence of defined harvest limits, and (3) running the REKN simulation model under a scenario representing near-total collapse of the HSC population. The concerns identified above, which arise from analysis of the limited data and code made available to date, demonstrate, at a minimum, that such further testing is warranted. It seems prudent to delay implementation of the new ARM framework until the public and outside experts have had adequate time to scrutinize the statistical and simulation models that play such a central role in this proposed decision-making framework.

SUPPORTING EVIDENCE AND ANALYSES

The remainder of this report provides additional supporting details for the six major areas of concern identified above, including results and figures from re-analyses of the data presented in the ARM report.

1. The fitted relationship between HSC abundance and REKN vital rates (survival and fecundity) is of insufficient magnitude to forecast a decline in mean projected REKN population growth even under a total collapse of the HSC population

Including a model of REKN population dynamics as part of the previous and revised versions of the ARM framework implicitly acknowledges that reduction of the HSC population could, under some circumstances, have a negative impact on REKN populations. This assumption has a strong empirical basis, as multiple lines of evidence suggest that HSC eggs are an extremely important resource for migrating REKNs during their spring migration (e.g., Karpanty et al. 2006; Niles et al. 2009; USFWS 2014; USFWS 2021). Therefore, it is surprising that the fitted relationship between HSC abundance and REKN survival used in the revised ARM is very weak and appears to be

overwhelmed by random among-year variation (Fig. 47 from ARM Report; Fig. 9 from Supplemental Report; hereafter, I will use the notation 'ARM Fig. 47/9'). In fact, it appears from the ARM report that estimated REKN survival rates have generally decreased weakly over time despite an estimated increase in HSC abundance (ARM Fig. 44/7). Years with the lowest HSC abundance in the study period (at or near the lowest HSC abundances ever recorded in Delaware Bay) are coincident with the highest estimated REKN survival rates (ARM Fig. 47/9). Given this weak fitted relationship, simulated REKN abundance based on this model seems unlikely to be very sensitive to changes in HSC abundance. Indeed, a 'back of the envelope' calculation based on the REKN vital rates presented in the ARM report (and the slightly modified numbers presented in the Supplement) shows that the mean population growth rate (Λ) of the REKN population is likely to remain at or above replacement levels ($\Lambda \geq 1$) even at HSC population size equal to zero (Fig. 1). This calculation was produced by using the mean survival from Supplemental Table 8, mean recruitment estimated from Supplemental Fig. 7b, and the standardized logistic regression coefficients from Supplemental Table 9 (effect size = 0.37 for survival and -0.14 for recruitment) to model REKN survival and recruitment as a function of HSC abundance. As a brief aside, the regression coefficients presented in the ARM report (e.g., effect of HSC on survival) are standardized and are on the logit (log-odds) scale, making them difficult to interpret. A quick example may help to aid interpretation of the effect size of this relationship: given a coefficient of 0.37 (the mean regression coefficient for the relationship between HSC abundance and REKN survival from the ARM Supplement, Table 8), a loss of 1 million female horseshoe crabs from Delaware Bay would result in REKN survival rate declining by only 0.004 (from 0.93 to 0.926). This is consistent with visual inspection of ARM Fig. 47/9.

Although I did not have access to the code and data used to fit the relationships between HSC abundance and REKN survival and recruitment, the relationships I used to generate Fig. 1 closely match the relationships presented in ARM Fig. 46/8 (Fig. 2). Interestingly, the value for mean recruitment provided in Supplemental Table 8 ($\rho_{\text{mean}} = 0.063$) yields a declining REKN population ($\Lambda = 0.99$) even under average conditions from 2005 to 2017. Since this result is inconsistent with the reported Λ of 1.04 during that same period from ARM Table 25 (and the generally increasing population trajectories indicated in ARM Fig. 58/15), I chose to use the

mean annual recruitment estimated from Supplemental Fig. 7b, which I calculated to be 0.109 (or geometric mean of 0.099). Using these mean recruitment values resulted in a Lambda of 1.035 (for arithmetic mean) or 1.027 (for geometric mean), more closely resembling but still below the reported baseline Lambda of 1.04 from the ARM report; setting baseline Lambda to 1.04 would only make a stronger case that REKN populations would not be expected to decline under an HSC population collapse (Fig. 1). This simulation exercise makes it very clear that the REKN model used in the revised ARM would not be able to predict or explain the decline in the REKN population observed during the 1990s, which has been attributed to unregulated harvest of HSCs in Delaware Bay (Niles et al. 2009; USFWS 2014). If this framework is unable to explain the decline of the REKN population in the first place, it does not appear to be an appropriate tool for helping to reverse the decline and promoting the recovery of this threatened subspecies.

Note that the population vital rates used to generate Fig. 1 represent point estimates. Because there was uncertainty associated with the estimate of Lambda (CI from 1.00 to 1.06; ARM Table 25), and with the effect size of HSC abundance on survival rate (CI from 0.12 to 0.63; ARM supplemental Table 9), some simulation runs (i.e., those with small Lambda and larger effect size sampled randomly from the joint posterior distribution) are likely to indicate REKN population decline at low HSC abundances. It is likely that these (probably rare) simulations drive the shape of the REKN “harvest function” yielded by the approximate dynamic programming algorithm. However, without access to the IPM and simulation code, I am not able to formally test the behavior of the REKN simulation model under scenarios of HSC population decline or collapse.

Scenario: HSC population collapse

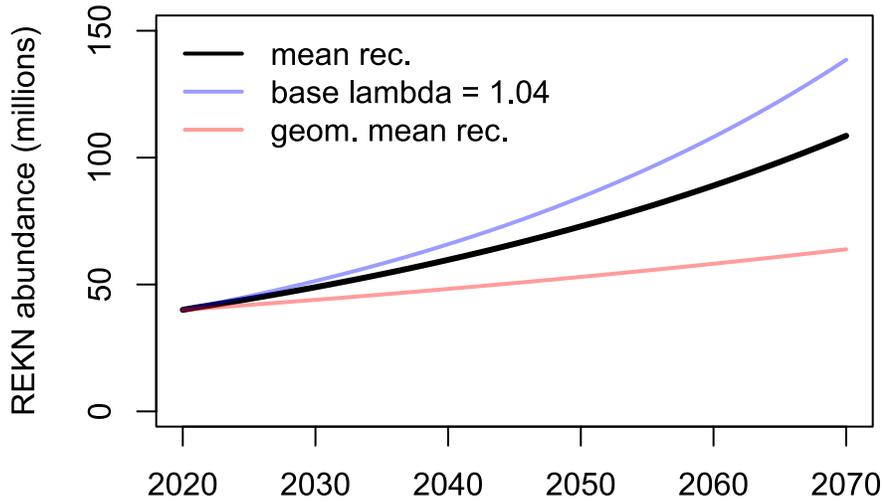


Figure 1. Results from a ‘back of the envelope’ calculation of REKN population growth under a scenario with depleted HSC population (female HSC abundance = 0 based on numbers presented in the ARM report. Mean recruitment rate was computed in three ways: arithmetic mean of values from ARM Supplemental Fig. 7b (“mean rec”), the geometric mean of these same values (“geom. mean rec.”), and a value fitted to ensure a population growth rate (Lambda) of 1.04, as indicated in the ARM report. Although somewhat simplistic, this figure illustrates that the reduction in REKN survival due to the collapse of HSCs in Delaware Bay appears to be insufficient to induce a meaningful REKN population decline. This figure is based on a simple age-structured population model and does not incorporate a density-dependence mechanism (the revised ARM includes a density ceiling that prevents the REKN population from growing above ~150k).

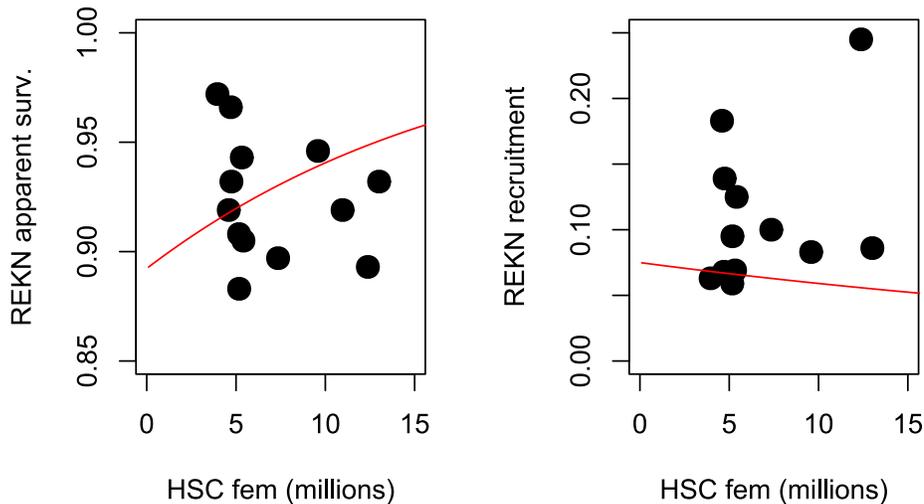


Figure 2. Relationships between female HSC abundance and REKN survival (left panel) and recruitment (right panel), recreated from information in the ARM supplemental report for the purpose of calculating the expected REKN population response to changes in the HSC population. Solid black dots represent annual vital rates estimated from ARM Supplement Fig. 9, and the red lines represent the fitted relationships presented in ARM Supplement Table 9.

Due to the weakness of the HSC/REKN relationship used in the revised ARM, and due to the complexity of the Integrated Population Model (IPM) framework used to represent the REKN population in the revised ARM, the relationship between HSC abundance and REKN population vital rates are likely to be unstable (sensitive to new data and alternative model specifications). Therefore, it is not implausible that the fitted relationship may disappear (become “non-significant”) – or even flip sign to become a negative relationship – when the IPM is fitted to additional observations. This outcome would pose an existential problem for the ARM framework, decoupling the two-species framework and rendering the REKN model unusable in the context of management. There does not appear to be a contingency plan for this outcome. More generally, the REKN IPM appears to have gone through several distinct versions before researchers settled on a final set of decisions to incorporate into the final model (there are several important differences between an earlier version of the IPM presented in Tucker [2019] and the ARM report). Ideally, the results from alternative representations of the REKN system should be considered in aggregate to better represent structural uncertainty about this system (Williams 2011).

The linked two-species modeling framework in the revised ARM assumes the relationship between REKN and HSC is independent of REKN densities (i.e., it assumes a prey-dependent functional response). Under this assumption, larger REKN populations do not require larger abundances of HSC females (i.e., more HSC eggs deposited) to support adequate per-capita weight gain; in other words, the ARM model assumes that a REKN population of 40k would experience the same per-capita survival and fecundity as a population of 400k for a given abundance of female HSC. Implicitly, this assumes a lack of interference among REKN individuals, and no decline in the mean quality or accessibility of HSC egg resources at elevated REKN abundances (Karpanty et al. 2011). Some researchers have argued convincingly that a ratio-dependent functional response – in which per-capita prey consumption depends on the ratio between prey and predator abundances – is likely to be more realistic for simulation models with discrete time steps that span the entire reproductive periods of predator and prey (Abrams and Ginzburg 2000), such as the linked two-species model used in the revised ARM.

The previous ARM framework used data gathered from multiple sources of data outside

Delaware Bay to parameterize the simulation models. The revised ARM attempts to use Delaware Bay data sources wherever possible – which is a significant advance in many ways, as the revised ARM is “fine-tuned” for the system and can be updated relatively easily as new data are collected. However, this modeling decision also limits the analyses to a small geographic area over a short period of time, potentially ignoring relevant evidence from other regions and/or time periods. Furthermore, the time frame over which data are available for fitting the population models used in the revised ARM represents a limited scope of historical variation during which populations of REKN and HSC were relatively small in comparison with earlier estimates. Using these models to forecast system dynamics under conditions outside the range of values used to fit the model (e.g., lower HSC abundances, higher REKN abundances) therefore requires extrapolation, which can be highly uncertain (and often inaccurate). Since both the HSC and REKN simulation models tend to produce forecasts that differ from current conditions (e.g., larger numbers of both species), and because the optimization routine relies on these simulated results, the management recommendations emerging from the revised ARM rely on highly uncertain extrapolations about HSC and REKN population dynamics and about how these two species may interact (analogous to extrapolations of species and community distributions under climate change; Araujo and Rahbek 2009). On one hand, the ARM framework is designed to be able to refine management policies as new data become available and as sources of uncertainty are reduced (Nichols et al. 2007). On the other hand, it does not seem prudent to implement management “experiments” that could potentially imperil a threatened or endangered species (TES), even under the rubric of adaptive management.

In summary, the relationship between HSC abundance and REKN survival appears to be too weak to induce a decline in REKN abundance (Fig. 1). If all HSCs in Delaware Bay disappeared today, the model would continue to predict a generally stable or increasing population of REKN over the next 50 years. Therefore, the revised ARM model would be unable to predict the decline of REKNs that was observed in recent decades, and which has been attributed in part to the decline in the HSC population (Niles et al. 2009; USFWS 2014). This lack of consistency between the revised ARM model and recent historical observations raises significant doubts about the ability of this model to accurately reflect future risks to the REKN population or to guide HSC

harvest decisions in a way that promotes REKN survival and recovery. Furthermore, the decision to include a REKN population model as part of the ARM framework (in both the original and revised versions) presupposes that HSC harvest could result in risk to the REKN population; the apparent inability of the ARM model to predict a decline in REKN abundance under a total HSC population collapse violates this premise and undermines the apparent purpose of the model.

2. The HSC population simulation model fails to propagate uncertainty about mean recruitment rates

The HSC recruitment process is the most consequential empirically fitted component of the HSC simulation model. Other elements of the HSC simulation model are not fitted to data – for example, natural mortality rate, the biomedical mortality rate, and bait harvest rates are fixed by the modelers. In the revised ARM, the recruitment process is fitted to data indirectly via the CMSA model; annual male and female recruitment estimates were used to fit a bivariate log-normal distribution (defined by a mean and standard deviation for each sex, along with a covariance between sexes – all on a logarithmic scale), which was then used to represent annual recruitment in the simulation model. The only other parameter fitted in the CMSA model – initial abundance – is not directly used in the simulation model. Recruitment is critical for any assessment of population resilience to harvest, since (in the absence of immigration, which is not included in the revised ARM), it is the only process that enables the population to overcome sources of mortality. Therefore, it is not surprising that the HSC simulation model is highly sensitive to changes in mean (log) fecundity (ARM Fig. 33; note that when I omit any reference to the supplemental report, I am referring to the primary ARM report). Given the high sensitivity of the HSC simulation model to the (log) mean HSC recruitment for males and females, it is critical that uncertainty about these parameters is properly represented in simulation models. However, the revised ARM framework incorrectly treats uncertainty about annual recruitment rates as representative of temporal process variance (natural year-to-year fluctuations) rather than as a mixture of parameter uncertainty and process variance (Link and Nichols 1994; Regan et al. 2002; McGowan et al. 2011). This is a subtle but consequential error, as sources of uncertainty will tend to “regress to the mean” (with good years cancelling bad years) instead of propagating over time.

To estimate the parameters for the log-normal recruitment process in the revised ARM, the following steps were taken: (1) log-normal distributions were separately fitted to each estimate of primiparous abundance (separately for each year and sex), based on estimates of parameter uncertainty (95% confidence intervals) derived from the CMSA results, (2) this collection of lognormal distributions (representing parameter uncertainty) was used to simulate annual male and female primiparous abundance for the years represented in the CMSA model (confusing parameter uncertainty with temporal process variation), and then (3) data from these simulations were used to fit a bivariate lognormal distribution (via maximum likelihood) for representing annual HSC recruitment in the ARM model. In general, parameter uncertainty should be represented in simulation models by drawing a single sample per replicate from a distribution of values representing parameter uncertainty (or by running replicates with “worst-case” and “best case” values for key parameters). However, the “canonical” version of the HSC projection model fails to address parameter uncertainty – most notably, uncertainty about the mean HSC recruitment rate, to which the HSC projection model is highly sensitive (ARM Fig. 33). Therefore, there is more uncertainty about the future of the HSC population in Delaware Bay than the revised ARM acknowledges. It is important to note that a sensitivity analysis was run in which expected recruitment was allowed to vary across simulation replicates within ca. 5% or 10% of the median recruitment value. This sensitivity test demonstrates an appropriate method for modeling parameter uncertainty; however, this test fails to represent the extent of uncertainty about the median HSC recruitment, which extends far beyond 10% of the mean estimated value (Fig. 3). Furthermore, this treatment of uncertainty was only run as a scenario test and was omitted from the ‘canonical’ version of the ARM that is proposed for use in managing the HSC harvest in Delaware Bay.

Interestingly, the REKN projection model in the revised ARM appears to represent parameter uncertainty appropriately. The key parameters of the REKN model were estimated using an Integrated Population Model (IPM), which were fitted in a Bayesian framework. In this framework, parameter uncertainty is represented by a joint posterior distribution that embodies the set of values that are consistent with the observed data. Furthermore, temporal process variation in the REKN population model is treated by explicitly modeling annual variability in key

vital rates (survival and recruitment) via annual random effects fitted with hyperparameters (Kery and Schaub 2011). This Bayesian hierarchical approach enables parameter uncertainty and process variation to be interpreted and modeled separately in a straightforward and intuitive manner. Specifically, parameter uncertainty is incorporated by running multiple replicates with different values drawn from the joint posterior distribution, and temporal process variation is included by sampling from the hyperparameters across years within each replicate (Goodman 2002).

To enable sensible propagation of parameter uncertainty in the HSC simulation model (analogous to the REKN model in the ARM), I constructed and fitted a hierarchical Bayesian version of the CMSA model. This model was fitted using the same data and model structure as the CMSA model included in the revised ARM. However, instead of estimating annual recruitment separately for each year and sex, the Bayesian CMSA model included an explicit representation of temporal process variance in recruitment (i.e., a “random effect” describing inter-annual variation in recruitment). This temporal process model was specified using a bivariate lognormal distribution exactly analogous to the HSC simulation model included in the ARM model, which included “hyperparameters” for male and female (log) mean recruitment, male and female (log) standard deviation, and a correlation term. By estimating temporal process variation directly, the Bayesian CMSA closely mirrors the HSC simulation model (analogous to the direct relationship between the IPM and the REKN simulation model), circumventing the multi-step process used in the ARM to generate the bivariate lognormal distribution from the CMSA results, and (most importantly) enabling the parameters of the bivariate lognormal distribution to be estimated directly from the data. To simulate HSC abundance over time, parameters for each replicate were drawn from the joint posterior distribution (representing parameter uncertainty), and temporal process variation within each replicate was simulated by sampling from the bivariate lognormal distribution. For the simulations, I incorporated the same restrictions in the stock-recruitment relationships indicated in the ARM report (driven by abundance and sex ratios for the years in which recruits were expected to have hatched).

Results from the Bayesian CMSA model indicate substantial uncertainty around mean HSC recruitment rates for both males and females (Fig. 3). Simulations (50 year time horizon) from

- Midwood, J. D., Cairns, N. A., Stoot, L. J., Cooke, S. J., & Blouin-Demers, G. (2015). Bycatch mortality can cause extirpation in four freshwater turtle species. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 25(1), 71-80.
- Nichols, J. D., Runge, M. C., Johnson, F. A., & Williams, B. K. (2007). Adaptive harvest management of North American waterfowl populations: a brief history and future prospects. *Journal of Ornithology*, 148(2), 343-349.
- Niles, L. J., Bart, J., Sitters, H. P., Dey, A. D., Clark, K. E., Atkinson, P. W., ... & Veitch, C. R. (2009). Effects of horseshoe crab harvest in Delaware Bay on red knots: are harvest restrictions working?. *BioScience*, 59(2), 153-164.
- Radosavljevic, A., & Anderson, R. P. (2014). Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of biogeography*, 41(4), 629-643.
- Regan, H. M., Colyvan, M., & Burgman, M. A. (2002). A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological applications*, 12(2), 618-628.
- Riecke, T. V., Williams, P. J., Behnke, T. L., Gibson, D., Leach, A. G., Sedinger, B. S., ... & Sedinger, J. S. (2019). Integrated population models: model assumptions and inference. *Methods in Ecology and Evolution*, 10(7), 1072-1082.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2016). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant.
- Tucker, A. (2019). Stopover ecology and population dynamics of migratory shorebirds. Dissertation. Auburn University.
- Tucker, A. M., McGowan, C. P., Lyons, J. E., DeRose-Wilson, A., & Clark, N. A. (2021). Species-specific demographic and behavioral responses to food availability during migratory stopover. *Population Ecology*, 64(1), 19-34.
- USFWS 2014. Endangered and Threatened Wildlife and Plants; Threatened Species Status for the Rufa Red Knot. *Federal Register / Vol. 79, No. 238*
- USFWS 2021. Endangered and Threatened Wildlife and Plants; Designation of Critical Habitat for Rufa Red Knot (*Calidris canutus rufa*). *Federal Register / Vol. 86, No. 133*
- Williams, B. K. (2011). Resolving structural uncertainty in natural resources management using POMDP approaches. *Ecological Modelling*, 222(5), 1092-1102.