Report of the External Independent Review under the Western Pacific Stock Assessment Review process

Level 1 & 2 Essential Fish Habitat Models for Main Hawaiian Islands Uku (*Aprion virescens*)

July 12-14, 2022
Council Office Conference Room
Honolulu, Hawaii

1. Executive summary

Level 1 (presence-absence) and Level 2 (density dependent) modeling approaches to improve the delineation of Uku (*Aprion virescens*) essential fish habitat (EFH) within the main Hawaiian Islands (MHI) were subjected to external review in relation to pre-defined Terms of Reference (TORs, see Annex I). Three external panelists received presentations reviewing the TORs noting that each modeling approach was to be evaluated independently according to appropriate Level and not compared against the other. Additional presentations were provided covering EFH definitions, Uku commercial and recreational catch, survey data, life history, biology, stock assessment and sources of fishery-independent abundance data. The information provided represents what is known from published literature upon which the existing definition of EFH of MHI Uku is based. The two researchers presented on each of their modeling approaches with respect to data sources, methods, results with relevant discussion and fielded questions from the panelists.

The reviews were conducted within the framework of the Western Pacific Stock Assessment Review (WPSAR) process. Franklin (2021) provided a Level 1 estimate of Uku relative abundance based on two presence-absence models using boosted regression trees. The models utilized shallow water diver observations (<30 m) and baited remote underwater video (BRUV) data (30 – 300 m) which in combination covered the full vertical habitat range of the species. Tanaka et al. (2022) developed a Level 2 generalized additive mixed-effects model using the diver data exclusively to estimate Uku population density for shallow-water areas (<30 m). Neither study examined egg or larval abundance, concentrating on EFH for large juvenile, sub/adult and adult stages. A variety of static and dynamic covariates were examined in the Level 1 and Level 2 modeling approaches.

Both approaches represent a great improvement over the existing literature-based description of Uku EFH (WPRFMC 2020). However, both approaches have a number of recommendations for development and improvement as detailed in this report. It should also be noted that the fishery independent data sources available for the modeling approaches have limitations as described in this report and generally represent low encounter rates of uku relative to other species. . Available fishery dependent data sources do not provide fine-scale resolution needed to model EFH.

The Level 2 analysis contains a scientifically sound approach for estimating density but the applied data set does not represent the primary depth distribution of the species in the MHI necessary for delineation of EFH boundaries. The model can be tested and applied to

shallow water reef species that remain mostly in shallow waters (<30 m) but should not be used for the delineation of MHI Uku EFH. The Level 2 modeling approach should be tested with deeper data to explore its utility for estimating Uku density throughout the MHI.

The Level 1 approach examined both shallow diver (<30 m) and the deeper data from baited remote underwater video surveys (BRUV) (30-300 m) that provided observational data for Uku throughout its known depth range. This is a large improvement in defining Uku EFH and the results can be combined with existing published information to improve the delineation of Uku EFH suitable for management purposes. However, there is a potential problem of spatial discontinuity as two separate models were used separately to estimate Uku occurrence in shallow and deep waters. The utility of combining the shallow and deep fisheries independent data sources into a single model should be explored.

2. WPSAR review panelists

a. David Itano - Chair
   Member WPRFMC Scientific and Statistical Committee (SSC), Fisheries consultant, Hawaii, USA

b. Szymon Smolinski, PhD
   National Marine Fisheries Research Institute, Department of Fisheries Resources, Poland

c. Francisco Javier Murillo Perez, PhD
   Ocean and Ecosystem Sciences Division, Department of Fisheries and Oceans, Bedford Institute of Oceanography, Nova Scotia, Canada

3. Researchers presenting models for the delineation of Essential Fish Habitat (EFH) for Uku (*Aprion virescens*) in the Main Hawaiian Islands (MHI)

a. Kisei Tanaka, PhD
   NOAA, NMFS, Pacific Islands Fishery Science Center
   Ecosystems Division, Stock Assessment, Honolulu, Hawaii

b. Eric C, Franklin, PhD
   Hawaii Institute of Marine Biology
   School of Ocean and Earth Science and Technology, Honolulu, Hawaii

4. Summary of review activities and background information on Main Hawaiian Island (MHI) Uku

a. Mark Fitchett, Council Staff opened the three-day meeting held in a hybrid mode with panelists, modelers, representatives from NMFS science center and reginal office meeting in person with Council Staff. Presenters and members of the public could attend virtually via online conferencing. Following introductions, the following background was provided:

b. David Itano, WPSAR Chair provided an overview of the main objectives and Terms of Reference (TORs) for the WPSAR EFH review process. It was noted the review should follow the TORs closely to determine if the two models (products) can be considered best scientific information available (BSIA) with respect to delineating EFH for MHI Uku.

   It was further noted that this would be the first ever model-based definitions of EFH for the region with two disparate approaches that should be reviewed independently and not compared. It was proposed that Dr. Franklin's model would be reviewed at Level 1 (presence/absence data) while Dr. Tanaka's product would take a Level 2 approach to examine Uku habitat related to Uku density for shallow areas.

   The products were to be evaluated according to the TORs as listed in Annex I. The panelists agreed that each model would need to be evaluated independently against each TORs and are presented in that format within this report. Recommendations for model improvement and future research priorities are provided.

c. David Delaney (PIRO Habitat Conservation Division) provided a detailed overview of EFH with respect to: definitions, designations, adverse effects, habitat types, habitat areas of particular concern (HAPC), the history of the current EFH definition for Uku, and a brief overview of the consultation process. It was clarified that there are currently no quantitative definitions for EFH in the region and that the existing definition of EFH for MHI Uku was very broad and general in nature with little information on life stages, particularly for early life stages. The current EFH definition for western Pacific Uku is provided in WPRFMC (2005).

d. Jason Helyer (Hawaii Division of Aquatic Resources, DAR) presented on commercial and non-commercial fisheries that catch and interact with Uku in the Main Hawaiian Islands. Helyer provided information on catch trends, spatial patterns, and discussed the resolution of fishery-dependent data for Uku (spatially and temporally). Recreational/non-commercial take of Uku in the MHI is estimated to be roughly equal to commercial landings and represents an important source of high quality bottomfish to the local fishing community that is easily accessible (relatively shallow) compared to the deepwater eteline snappers and are also taken at the surface with simple troll gear. Uku landings in the MHI are the second highest of any single bottomfish species. Annual catches peak in the summer period apparently coinciding with the spawning season when commercial fishers can more efficiently target Uku.

   Statewide landings are dominated by catches on the Penguin Banks, a shallow shelf extending south southeast from west Molokai at a favorable depth for Uku and from the of Maui Nui area (shelf area of Maui, Kaho'olawe, Lanai and Moloka'i). Fisheries dependent data resides in a long term (since 1946) fisher reported catch and effort database that is a reporting requirement for commercial license holders in the state. Reporting grids are relatively coarse and do not provide depth specific data so are less useful for delineation of EFH. For example the entire Penguin Banks is contained within a single reporting grid. However, gross data on relative abundance is available by grid area. Recreational data is available through the Hawaii Marine Recreational Fishing Survey (HMRFS) program using a combination of angler intercept surveys and mail in effort surveys. Shark depredation when fishing on Uku aggregations was noted as a potential influence on CPUE that if known could improve data quality for assessment.

e. Joseph O'Malley (PIFSC Life History Program lead) presented on the current state of knowledge of life history (growth, maturity, size and age ranges, natural mortality) and biology (life stage-specific habitats, diet, and movement) of Uku related to EFH. Uku are harvested by recreational fishers in shallow waters and by commercial bottom fishermen, typically in depths of 50 – 120 m but are known to aggregate in mid waters, presumably on spawning aggregations. Uku mature relatively early at around 3 years at an L50 of 449 mm (Everson et al 1989) but can become quite old with otolith ageing out to 30+ years. Uku spawning season, inferred by monthly mean gonosomatic index (GSI) peaks in May – July when it aggregates in shallow waters to spawn.

   Egg and larval stages are pelagic but details of depth, distribution and larval duration are unknown, but cross-channel larval dispersion is assumed. Juvenile Uku settle in shallow soft-bottom habitats different from typical adult habitats. Adults have been characterized as habitat generalists, occupying both hard and soft bottom areas but are typically targeted over hard bottom, high relief areas and often on the tops or edges of shallow banks to 180 m. The lower limit of adult Uku depth range was noted at 194 m (WPRFMC 2017). Seasonal movements from winter (October – April) core areas to summer (May – September) areas are hypothesized to be associated with spring/summer spawning areas. Diel movements are linked to tidal flow and possibly feeding movements (Myers et al. 2007)

f. Marc Nadon (PIFSC, Stock Assessment) presented on the most recent 2020 benchmark stock assessment for MHI Uku and current stock status. Parameters of interest include spatial patterns of catch, data gaps, and patterns and trends of fishery exploitation. Stock Synthesis 3 was used with the commercial marine catch data standardized in a number of ways over the long history of the database. Recreational catch data was modeled from survey data which was adjusted in the Stock Synthesis 3 modeling framework. Stock status was assessed (2020) as not overfished and not experience overfishing with total biomass and female spawner biomass steadily increasing from 2004 onward to 2018. Recruitment estimates have also increased in recent decades. CPUE is increasing across all gear types but it is not clear if this is due to changes in catchability, efficiency of the fishery, environmental variables or issues within the DAR dataset.

5. Fishery Independent Data Sources for MHI Uku

a. Audrey Rollo (NOAA, PIFSC) provided an overview of camera systems utilized in the Level 1 modeling approach. The Modular Optical Underwater Survey System (MOUSS), BotCam, and baited remote underwater videos (BRUVs) were used in survey effort. Video surveys were deployed during daytime hours on soft and hard bottom in a range of ~75 – 400 m. All video surveys were baited to draw subjects but survey times were limited to reduce the likelihood of attraction from a larger area.

b. Tye Kindinger (NOAA, PIFSC) presented on paired-diver stationary point count (SPC) survey data through the NOAA National Coral Reef Monitoring Program (NCRMP) that were used in both the Level 1 and Level 2 analyses. Replicate counts are made on SCUBA gear within two adjacent 15 m diameter cylindrical areas, enumerating fish from the bottom to as high in the water column as feasible. Species that tend to be elusive or avoid divers are noted early in an attempt to include them in the survey frame. However, protocols may have changed over time. Survey divers estimate fish numbers and length. These surveys were designed to target shallow-water coral reef associated fish species so are concentrated on hard bottom substrate to a maximum depth of 30 m. It was noted that Uku are not attracted to divers or diver activity and in fact tend to avoid divers, especially in areas commonly dove or swept by spearfishermen. Observations of Uku on these surveys were relatively uncommon and normally consisted of a single individual moving across the bottom.

6. WPSAR Presentations from Authors on Level 1 and Level 2 EFH Analyses.

a. Uku abundance was assessed by Dr. Eric Franklin at Level 1 using a combination of diver and stereo video camera derived observations and at Level 2 by Dr. Kisei Tanaka using diver survey data to assess shallow water distribution of MHI Uku. Figure 1 lists the four data levels that can be applied to designate EFH. Level 3 and Level 4 data useful for the delineation of EFH for MHI Uku does not currently exist so evaluations were restricted to Level 1 and 2. Information on the egg and larval stage of MHI Uku is limited or unknown so EFH determinations were restricted to larger juvenile and adult stages for both studies.

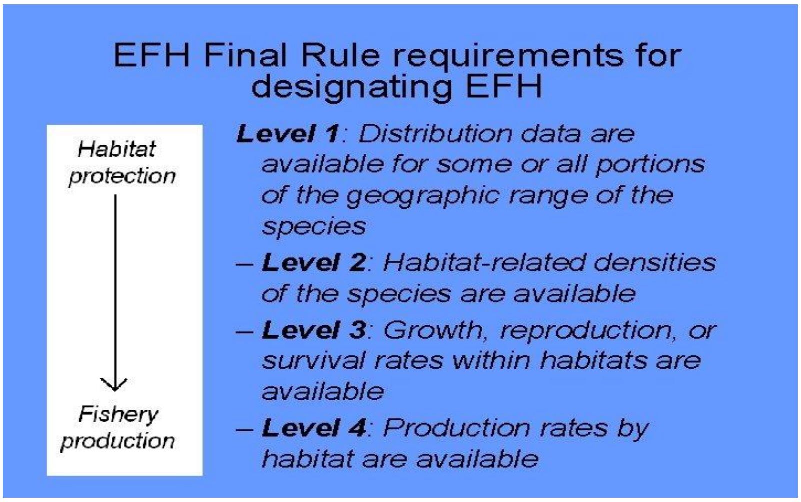

Figure 1. Data Levels 1 - 4 for designating EFH

7. Level 1 EFH model: Distribution data for Uku presence-absence. Franklin, E.C. 2021. Model-based essential fish habitat definitions for the Uku *Aprion virescens* in the Main Hawaiian Islands. (Overview)

Dr. Erik Franklin presented on a Level 1 modeling approach to improve on the definition of EFH for MHI Uku. The presentation was provided in respect to data sources, methods, results and relevant discussion. The approach used observational presence-absence data from shallow diver and deeper underwater video observations. The combination of the two datasets provided data that encompasses the known vertical habitat distribution of Uku. Species Distribution Models (SDM) previously used in Alaska for defining EFH were adapted to examine the shallow diver (<30 m) and deeper (30 – 330 m) video datasets separately (Laman et al. 2018). SDMs can predict occurrence or abundance of a target organism across the study domain. The presence-absence of combined juvenile and adult Uku was assessed to determine habitat suitability (probability of occurrence). Influence of static variables (bathymetry, slope, aspect, rugosity, % sand) and dynamic variables (shallow data only) of mean and maximum significant wave height were evaluated. Fishery dependent data was not used in the analysis. However, the modeled encounter probabilities using fishery independent data did correspond to the Hawaii Division of Aquatic Resources (HDAR) Fishery Reporting System grids of highest encounter rates from fishery dependent data.

8. Level 1 EFH evaluation against TORs: distribution data (presence-absence)

a. ***Level 1/TOR 1*: Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?**

YES. The author conducted a thorough search of scientific literature and fishery-independent and fishery-dependent information relevant to the delineation of MHI Uku. Information from relevant information sources were incorporated into the paper and included among the cited references.

Two sources of visual survey data were selected: shallow (0-30 m) NOAA Fisheries field observations from stationary point count (SPC) diver fish surveys and deeper (30-300m) observations from three separate NOAA baited stereo video survey programs. Based on the information provided during the WPSAR meeting (presentations by Audrey Rollo and Tye Kindinger on 12 July 2022, and further discussions), these data sets are considered the best available quantitative information on the Uku occurrence in the MHI.

Large datasets from diver survey data of similar depth range as the NOAA diver surveys from the Hawaii Division of Aquatic Resources (HDAR) and The Nature Conservancy were excluded from the study due to differences in data collection and sampling methodologies. However, the gear and methodology from the three NOAA video survey programs (MOUSS, BotCAM, and NOAA PIFSC BRUVs surveys) were considered similar enough to pool these data sources to inform spatial presence-absence and depth information for MHI Uku. Only surveys deeper than 30 m and less than 300m with fish positively identified as Uku were included in the deep model analysis. Sampling station data with clearly erroneous geographical locations (on land, unrealistically deep) were also excluded from the analysis.

In general, the author did excellent work to use the shallow diver and combined deep video surveys to examine for the presence of Uku across an appropriate depth range. However, there is a need for a more complete presentation of the inclusion/exclusion criteria applied. More complete information on the other, available but unused, datasets for MHI should be presented. A comprehensive list of all available datasets found by the author would be useful. The explanation of the problems and technical challenges faced that prevented the usage of particular data sets in this modelling exercise would be of high value for future studies on the fish distribution in the MHI aimed at the delineation of EFH. Such information might also help in the identification of future work priorities and improve existing monitoring surveys.

i) **Recommend adding clarifying description in the report on the inclusion/exclusion criteria for data and explanations of why certain data sets were not included in the analysis (short term, high priority)**

b. ***Level 1/TOR 2*: Are the data properly applied and appropriate for this species and habitat?**

YES. The data sets used in this study to model relationships between the probability of occurrence of the species and environmental factors are properly applied and appropriate. The combination of shallow and deep water survey data encompassed the depth range of the species in the MHI (in line with the previous EFH definition based on the literature review presented in WPRFMC, 2005, and in line with more current primary scientific sources, e.g. Asher *et al.*, 2017) allowing the study to provide a more complete estimate of Uku EFH in the MHI. The usage of two data types and specific data sets is well explained and technical details of survey methodology are presented in the report.

One exception has to do with the characterization of Uku as a habitat generalist found on both hard and soft bottom habitats. Diver surveys were designed to capture more coral reef species and targeted hard and coral reef substrates. However, deeper water video surveys were distributed over both bottom types. Another issue may be related to diver avoidance suggesting that Uku may have been undersampled in the shallow diver surveys.

c. ***Level 1/TOR 3*: Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?**

YES. Applied models are appropriate to delineate EFH for Uku. Boosted regression trees are powerful machine learning algorithms that were used for similar tasks both in the terrestrial and marine systems (Elith *et al.*, 2008a). The author provided detailed information on the optimization, training, and validation processes in the report. The study applied Species Distribution Modeling (SDM) to model Uku presence-absence or degree of occurrence. The optimal models for both shallow and deep habitats had "good" model fits (>0.80) based on the Area Under the Receiving Operating Curve (AUC) calculation and the True Skill Statistic (TSS). This level of accuracy is considered excellent according to the broadly accepted rules of thumb (Hosmer and Lemeshow, 2000, p. 164). See Smolinksi (2022), Annex II for further supporting language.

The existing Level 1 analysis combines two separate models for diver (0-30 m) and video derived (30-300 m) data. Considering the discontinuity of shallow and deep-water model outputs, the possibility of a single model that combines dive and drop camera data should be investigated.

i. **Recommend combined modeling and analysis of shallow and deep data. This would include an evaluation of the combined model and adoption in the final report if deemed an improvement over the existing two-model approach. (short term, high priority)**

d. ***Level 1/TOR 4*: Are decision points and input parameters reasonably chosen?**

YES. The selected environmental covariates are reasonable for this spatial modeling exercise and appropriate for the study. A range of static variables were examined e.g. bathymetry, slope, aspect, rugosity, % sand) that have been shown to have high predictive power to the distribution of coral reef-associated fish (Knudby et al., 2010). Dynamic variables (shallow data only) of mean and maximum significant wave height were also evaluated for shallow water data (<30 m).

i. **Recommend providing an expanded description of the rationale for why certain variables were included and a summary of their value (or not) to the density predictions. (short term, high priority)**

ii) **Recommend the testing of additional variables that may be informative to Uku EFH (i.e. salinity, water temperature, water quality) – especially variables which can be helpful in the improvement of model performance in the areas of high false positive EFH classifications. (mid term, medium priority)**

e. ***Level 1/TOR 5*: Are primary sources of uncertainty documented and presented?**

NO. The model output should be better described, including typical summary statistics, like standard errors of the parameter estimates, which give a general overview of the uncertainty in the model. See Murillo 2022 (Annex III) for guidelines on how to address uncertainty in the modelling effort.

i. **Recommendation: Uncertainty of the models need to be investigated in detail and documented.**
   - **(1) calculate AUC values and other performance metrics presented in the report, as well as the coefficient of variation of the ensemble predictions separately by area, i.e. island or windward/leeward subareas. (see Annex II) (short term, high priority).**
   - **(2) Provide a map of uncertainty in the predictions. (see Annex III) (short term, high priority).**
   - **(3) Provide more information on the uncertainty of species identification from video surveys (short term, medium priority)**

f. ***Level 1/TOR 6*: Are model assumptions reasonably satisfied?**

YES. Model assumptions are reasonably satisfied. Boosted regression trees are among highly flexible data-driven methods (Elith *et al.*, 2008a) without the traditional statistical parameters (e.g. as in the traditional regression-based techniques). Since these methods are non-parametric, a minimal level of assumptions is to be made by the modeler. Tuning of model hyperparameters (tree complexity, learning rate, number of trees, and the bag-fraction) and selection of the optimal model separately for shallow and deep waters follow a well-established methodology (e.g. Knudby *et al.*, 2010a, 2010b). The reviewers agree that all statistical requirements are satisfied for both shallow and deep models.

g. ***Level 1/TOR 7***: **Are the final results scientifically sound, particularly delineating EFH boundaries?**

YES. The shallow and deep presence-absence data is considered the best scientific information available (BSIA) for MHI Uku and the combination of the two datasets provide observations that encompasses the known depth distribution of the species. The datasets have issues but the model results are a significant improvement over the existing literature-based EFH definition (see WPRFMC 2005).

i. Recommend correction of the categorical maps (revision of the calculation of quantiles following Laman et al. (2018) (short term, high priority)

ii. Recommend replacing descriptive and subjective terminology (e.g. "hot spots", "core areas") with terms that describe the relative probability of occurrence (i.e. top 25%, top 50%, top 95% = EFH). (short term, high priority)

9. Can the Level 1 results be used to address management goals stated in the relevant FEP or other documents provided to the review panel?

The short term, high priority Recommendations should first be addressed and evaluated, after which the Level 1 results can be used to delineate Uku EFH boundaries suitable for management purposes and can be considered current BSIA. This would include an evaluation of the combined model and adoption if deemed an improvement over the existing two-model approach.

10. Minor suggestions and comments to address for Level 1 model improvement and research priorities (short term, high priority)

a. Add 30 and 300-m bathymetry lines to Figures 8 to 11.
b. Add one figure with the presence-absence records overlaying the final probability map.
c. Add supporting information from Friedman and Meulman (2003) and Pyle et al. (2016) to the paper and references section.
d. Provide a figure in the Appendix or supplements with the correlation values between all the environmental variables used.

++

11. **Level 2 EFH model: Tanaka, K. R., Schmidt, A. L., Kindinger, T. L., Whitney, J. L., and Samson, J. C. 2022. DRAFT. Spatiotemporal assessment of Aprion virescens density in shallow Main Hawaiian Islands Waters, 2010-2019. (Overview)**

Dr. Kisei Tanaka presented a Level 2 modeling approach to predict Uku density using NOAA dive survey data in shallow MHI habitats to improve the definition of EFH for MHI Uku. The study was described with respect to data sources, methods, results, and relevant discussion. Dr. Tanaka developed a statistical EFH level 2 modeling framework employing a combination of *in-situ A. virescens* density data enhanced by various gridded satellite products to estimate the species' abundance in shallow MHI waters. This is the first study to use a large, fishery-independent database as a data source to provide model-based dynamic density estimates of a bottom-fish management unit species (MUS) across the MHI shallow waters.

The approach used observational fishery-independent density data from shallow (<30m) stationary point count (SPC) SCUBA diver observations from 2010 to 2019. The fishery-independent SPC survey protocol was based on a stratified random sampling design using the paired-diver SPC method. The divers deploy a marked transect line across the main axis of the cylinders and use this line as a frame of reference to estimate the extent of the cylinders. These surveys provide site-level density and biomass records across a range of fish species and trophic groups. Relevant environmental variables that are associated with the density distribution of *A. virescens* were included in the analysis: depth, sea surface temperature (SST), surface chlorophyll-a concentration (Chl α), and surface wind speed (SWS). Temporally corresponding SST, Chl α, and SWS values at each surveyed location were obtained for every time-stamped and georeferenced SPC survey record (n=2968). For every dated SPC survey record, each environmental variable's mean, standard deviation, 5th quantile, and 95th quantile values over the past one month were calculated from the corresponding survey date.

A spatiotemporal modeling technique was used to predict changes in the species' localized density (spatially resolved number of individual estimates per 100 m2) in relation to dynamic environmental variables. The statistical mixed-modeling approach (generalized additive mixed-effect model; GAMM) accounts for spatial autocorrelation between spatially referenced observations and effects of environmental drivers and incorporates a spatially explicit temporal trend (i.e., local trend) alongside spatial (temporally constant) and spatiotemporal (time-varying) components, thereby imposing correlation across space and time in the estimates of target response variables. Using this approach, size-aggregated *A. virescens* density is modeled as a function of 'fixed' effects resulting from explicit habitat variables and random effects as a product of unobserved or 'latent' spatiotemporal effects. Overall, the model indicated low but steady Uku densities in shallow MHI waters (0-30 m).

12. **Level 2 EFH evaluation against TORs: MHI Uku (density)**

    a. ***Level 2/TOR 1*: Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?**

<u>YES with caveats</u>. The Level 2 model to predict MHI Uku density used the single data set in relatively shallow waters (<30 m) from the NCRMP dive surveys. These data provided estimates of Uku densities sampled in a randomized fashion and were appropriate and justified for this purpose. However, other data sets were potentially available but not utilized, particularly for deeper observational data. A summary of available datasets and criteria for inclusion/exclusion from the analysis would be informative and help in the identification of future work priorities and improvements to monitoring surveys useful for the delineation of EFH.

i. **Recommend the report provide details on the data inclusion/exclusion criteria and discussion as to why certain data sets were not included in the analysis. (short term, high priority)**

b. ***Level 2/TOR 2*: Are the data properly applied and appropriate for this species and habitat?**

<u>YES with caveats</u>. The data are properly applied but in total are not appropriate to delineate Uku EFH as applied data are limited to shallow waters. Uku is characterized in the literature and this study as a habitat generalist found in both hard and softbottom habitats. The shallow diver surveys tended to favor hard bottom and coral reef substrates so may have missed out on Uku frequency of occurrence on soft bottom habitats. The main issue of concern has to do with the exclusive use of diver survey data that is limited to daytime depths to a maximum of 30 m. The species ranges down to at least 180 m with depth preferences greater than 30 m noted for sub-adult and adult stages. The choice of this single shallow water data source limited the utility of the exercise by excluding the bulk of the sub-adult/adult biomass of Uku from the Level 2 density analysis. Inclusion of other data sources or newly available data that encompasses the full depth range of MHI Uku is recommended.

i. Recommend that the Level 2 density model be applied to sources of fishery-independent data that extend below 30m and encompass the depth range of Uku found in the MHI. Specifically, use of baited stereo-video survey data should be used (medium term, high priority).

c. ***Level 2/TOR 3*: Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?**

<u>YES.</u> The statistical approach developed in the Level 2 analysis is correct and constitutes the state-of-the-art framework for the analysis of Uku density data.

i. Recommend that some aspects of the model development procedures are improved including model selection, presentation of parameter estimates, and estimates of variance components. (short term, medium priority)

d. ***Level 2/TOR 4*: Are decision points and input parameters reasonably chosen?**

YES. The environmental variables selected for this modelling are appropriate and have potential biological meaning.

i. **Recommend to include in the paper a description of the rationale for why the selected environmental variables were included and an assessment of how influential each proved to be. (short term, high priority)**

e. ***Level 2/TOR 5:*** **Are primary sources of uncertainty documented and presented?**

NO. Uncertainty in the model should be more thoroughly documented and presented in the descriptive document. The model output should be better described, including typical summary statistics, like standard errors of the parameter estimates, which give a general overview of the uncertainty in the model.

i. **Recommendation: Uncertainty of the models need to be investigated in detail and documented. For example, the uncertainty of the predictions using bootstrapping. (short term, high priority)**

f. ***Level 2/TOR 6:*** **Are model assumptions reasonably satisfied?**

YES. Model assumptions appear to be satisfied. However, some discrepancies were observed in the quantile-quantile plots presented during the WPSAR meeting. Also, the scatterplot of the predicted versus observed values suggests that the model overpredicts the density when no fish are observed, while underpredicting positive values.

i. **Recommend additional validation and diagnostics tests (zero inflation and dispersion in the data) and how well the model is in the predictions of zero by the calculation of AUC (assuming zero and non-zero groups) to test model performance. (short term, high priority)**

g. ***Level 2/TOR 7:*** **Are the final results scientifically sound, particularly delineating EFH boundaries?**

YES with caveats. The Level 2 modeling approach produces scientifically sound results within the model domain. However, this domain is limited to data and output for coastal areas less than 30 m in depth around the MHI. Various life stages of Uku are found at these depths but the majority of the biomass and adult biomass occupies deeper waters. This limitation in data input and predictive capacity critically limits the use of the model in this configuration for the delineation of Uku EFH. However, this model using current data sources would be well suited and appropriate to the delineation of EFH for shallow-water reef associated species with vertical habitat mainly shallower than 30 m.

In summary, the Level 2 analysis contains a scientifically sound approach but the applied data set does not represent the main distribution of the species in the MHI needed for delineation of EFH boundaries.

i. **Recommend the preparation of procedures to transform density predictions into categorial maps (medium-term medium-priority)**

13. **Can the Level 2 results be used to address management goals stated in the relevant FEP or other documents provided to the review panel?**

In summary, the analysis contains a scientifically sound Level 2 approach but the applied data set does not represent the main vertical distribution of the species in the MHI needed for delineation of EFH boundaries and cannot be considered BSIA for this purpose.

14. **Technical suggestions and comments to address for Level 2 model improvement and research priorities (short term, high priority)**

a. **Expand description of available Uku fisheries-dependent data.**
b. **Items specific to Tanaka et al. 2021**
   1) Lines 167-168: please be more precise (as there was at least one previous study: Franklin, 2021).
   2) Fig. 2b: indicate what is presented with error bars
   3) Fig. 2c: it is not clear why discrete scales are presented in this figure. It needs more explanation in the caption
   4) Line 193: please correct the number of observations reported
   5) Table 1: temporal range or temporal resolution?
   6) Line 244: reconsider if these are continuous values
   7) Line 298: it seems that part of the sentence is missing. Personally, I do not agree that the model performance is close to ideal (see comments in sections 3.3. and 3.6.). Also, the second part of the sentence "variability in model accuracy increased at higher density" is imprecise. I suggest rewriting the whole sentence.
   8) Figure 4a: it is not clear what is presented in the plot as ranges of color gradients do not match maximal values reported above the panels. Please clarify.
   9) Figure 5: please describe beta and p in the caption. Confidence intervals are described in the caption but not present. As mentioned in the previous part of the review report, the note on the models used to describe temporal trends should be included in the methods section.
   10) Line 205: please provide the reasoning for using one-month time windows
   11) Lines 316-318: This sentence is not clear to me.

12) The full bibliographic data for Tanaka and Oliver 2021 is not in the list of references
13) Figure S2: red dots show land masses, not green. What then is the distinction between green and blue dots?

15. Recommendations for future research priorities

a. Explore additional scientific approaches commonly used in SDMs to discriminate thresholds for EFH delineation. For example, EFH can be defined by environmental covariates or by the physiology of the species in question.

b. Investigate the relationships between probability of occurrence and density estimates of focal species.

c. Conduct video data processing to obtain additional length data and length-based SDMs for separate ontogenetic stages.

d. Conduct additional MOUSS (baited underwater stereo video camera) surveys in mid-depth hard and soft-bottom habitats poorly sampled by existing diver and video survey effort.

e. Develop a focused research effort to identify the location and timing of Uku spawning aggregations throughout the MHI and the geographic extent of the population that contributes to different aggregations. Acoustic and conventional tagging techniques can be utilized.

f. Investigation into the diurnal and intra- and inter-island movements and habitat utilization of Uku. Acoustic and conventional tagging techniques can be utilized.

g. Designation of EFH according to their ecological function: spawning habitats, feeding habitats, nursery grounds, etc.

h. Explore the concept of ensemble approaches and comparing different model outputs in order to provide additional information on uncertainty and relationships between species distribution and environmental conditions.

i. Additional SPC dive surveys on soft bottom areas and seasons not covered by the NCRMP program.

j. Revisit existing data and literature on Uku behaviour in areas of heavy diving and spearfishing with potential bias to underwater visual census (Gray et al. 2016).

16. Comments on Terms of Reference

- The TORs were adopted from the original WPSAR process to review stock assessment modeling approaches and retain terminology and phrases like “input parameters” and “model assumptions reasonably satisfied” which could be better adapted to terms and the evaluation of spatial modeling and EFH delineation.

- Some of the TORs combine multiple questions such as “Are the data properly applied and appropriate for this species and habitat?” which should be split into separate TORs. Otherwise it is difficult to assign a Yes or No response, even with caveats.

- For TOR 4, it would be appropriate to replace “input parameters” with “environmental covariates”.

17. Public comment

A member of the public asked if there has been further consideration of other life stages of Uku not considered by the models under review (e.g. egg and larval stages). It was noted that there is a general lack of data for early life stages of Uku but that a recent or ongoing larval survey will add to available data sources.

18. References

Asher, J., Williams, I. D., and Harvey, E. S. 2017. An Assessment of Mobile Predator Populations along Shallow and Mesophotic Depth Gradients in the Hawaiian Archipelago. Scientific Reports, 7: 1–18. Springer US.

Elith, J., Leathwick, J. R., Hastie, T., and R. Leathwick, J. 2008a. Elith, Leathwick & Hastie A working guide to boosted regression trees - Online Appendices Page 1. Journal of Animal Ecology, 77: 802–13. http://www.ncbi.nlm.nih.gov/pubmed/18397250.

Everson, A. R., H. A. Williams, and B. M. Ito. 1989. Maturation and reproduction in two Hawaiian eteline snappers, Uku, *Aprion virescens* and Onaga, *Etelis coruscans*. Fishery Bulletin 87:877-888.

Franklin, E.C. 2021. Model-based essential fish habitat definitions for the Uku *Aprion virescens* in the Main Hawaiian Islands. Report to Western Pacific Regional Fishery Management Council. 34 pp.

Gray, A.E., Williams, I.D., Stamoulis, K.A., Boland, R.C., Lino, K.C., Hauk, B.B., Leonard, J.C., Rooney, J.J., Asher, J.M., Lopes Jr, K.H. and Kosaki, R.K., 2016. Comparison of reef fish survey data gathered by open and closed circuit SCUBA divers reveals differences in areas with higher fishing pressure. *PLoS One*, *11*(12), p. e0167724.

Knudby, A., LeDrew, E., and Brenning, A. 2010a. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. Remote Sensing of Environment, 114: 1230–1241.

Knudby, A., Brenning, A., and LeDrew, E. 2010b. New approaches to modelling fish–habitat relationships. Ecological Modelling, 221: 503–511.

Laman, E. A., Rooper, C. N., Turner, K., Rooney, S., Cooper, D. W., & Zimmermann, M. (2018). Using species distribution models to describe essential fish habitat in Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, *75*(8), 1230-1255.

Meyer, C. G., Y. P. Papastamatiou, and K. N. Holland. 2007. Seasonal, diel, and tidal movements of green jobfish (*Aprion virescens*, Lutjanidae) at remote Hawaiian atolls: implications for marine protected area design. Marine Biology 151:2133–2143.

Murillo, J. 2022. External Independent Peer Review under the Western Pacific Stock Assessment Review framework. Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*). 11 pp.

Smolinski, S. 2022. External Independent Peer Review under the Western Pacific Stock Assessment Review framework. Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*). 32 pp.

Tanaka K. R., A. L. Schmidt, T. L. Kindinger, J. L. Whitney and J. C. Samson. 2022. DRAFT: Spatiotemporal assessment of *Aprion virescens* density in shallow MHI Waters, 2010-2019. NOAA NMFS-PIFSC draft NOAA Technical Memorandum.

WPRFMC. 2005. Essential Fish Habitat Descriptions for Western Pacific Archipelagic and Remote Island Areas Fishery Ecosystem Plan Management Unit Species (Crustacean, Bottomfish, Precious Coral and Coral Reef Ecosystem Species). Western Pacific Regional Fishery Management Council, Honolulu, Hawaii.

WPRFMC 2017, Amendment 4 to the Fishery Ecosystem Plan for the Hawaii Archipelago. Revised Descriptions and Identification of Essential Fish Habitat and Habitat Areas of Particular Concern for Bottomfish and Seamount Groundfish of the Hawaiian Archipelago. Honolulu, Hawaii.

Annex I. Terms of Reference for the Peer Review

Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*)

External Independent Peer Review under the Western Pacific Stock Assessment Review framework: Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*).

For questions 1-7 and their subcomponents, reviewers shall provide a “yes” or “no” answer and will not provide an answer of “maybe”. Only if necessary, caveats may be provided to these yes or no answers, but when provided they must be as specific as possible to provide direction and clarification to NMFS.

1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?

2. Are the data properly applied and appropriate for this species and habitat?

3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?

4. Are decision points and input parameters reasonably chosen?

5. Are primary sources of uncertainty documented and presented?

6. Are model assumptions reasonably satisfied?

7. Are the final results scientifically sound, particularly delineating EFH boundaries?

8. Can the results be used to address management goals stated in the relevant FEP or other documents provided to the review panel? If any results of these models should not be applied for management purposes with or without minor short-term further analyses (in other words, if any responses to any parts of questions 1-7 are “no”), indicate:

    Which results should not be applied and describe why, and

    Which alternative set of existing EFH definitions should be used to inform Uku EFH instead and describe why.

9. As needed, suggest recommendations for future improvements and research priorities. Indicate whether each recommendation should be addressed in the short/immediate term (2 months), mid-term (3-5 years) and long-term (5-10 years). Also indicate whether each recommendation is high priority (likely most affecting results and/or interpretation), mid priority, or low priority.

10. Draft a report (individual reports from each of the panel members and an additional Summary Report from Chair) addressing the above TOR questions.

Annex II. Smolinski, S. 2022. External Independent Peer Review under the Western Pacific Stock Assessment Review framework. Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*).

External Independent Peer Review

under the Western Pacific Stock Assessment Review framework

Level 1 and Level 2 Essential Fish Habitat Models

for the Main Hawaiian Islands Uku (*Aprion virescens*)

Szymon Smoliński

National Marine Fisheries Research Institute

Gdynia, Poland

ssmolinski@mir.gdynia.pl

St. John’s, July 2022

Table of contents

1. Executive summary 20

2. EFH Level 1: distribution data (presence-absence) 22

2.1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented? 22

2.2. Are the data properly applied and appropriate for this species and habitat? 24

2.3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data? 25

2.4. Are decision points and input parameters reasonably chosen? 26

2.5. Are primary sources of uncertainty documented and presented? 28

2.6. Are model assumptions reasonably satisfied? 30

2.7. Are the final results scientifically sound, particularly delineating EFH boundaries? 31

2.8. Additional technical comments to the report 34

3. EFH Level 2: habitat-specific densities 35

3.1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented? 35

3.2. Are the data properly applied and appropriate for this species and habitat? 37

3.3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data? 38

3.4. Are decision points and input parameters reasonably chosen? 40

3.5. Are primary sources of uncertainty documented and presented? 41

3.6. Are model assumptions reasonably satisfied? 42

3.7. Are the final results scientifically sound, particularly delineating EFH boundaries? 43

3.8. Additional technical comments to the report 45

4. General comment on the research priorities 46

5. Notes to the ToRs 47

6. List of references 48

1. Executive summary

I independently evaluated Level 1 and Level 2 Essential Fish Habitat (EFH) models for the Main Hawaiian Islands (MHI) Uku (*Aprion virescens*) described in the provided reports (Franklin, 2021; Tanaka *et al.*, 2022) with respect to a set of pre-defined Terms of Reference. Evaluated reports present background information, methodology, and results of the spatial distribution modeling study conducted for MHI Uku with an adequate discussion. The results of these exercises constitute a great improvement over the existing EFH descriptions that are based on the literature review (WPRFMC, 2005). The fishery-independent survey data and spatial distribution models were involved to obtain a more accurate designation of the habitats essential for the species and provide adequate maps. Presented studies provide additional information on the spatial distribution of the species and its relationships with environmental conditions.

In both cases, a more in-depth presentation of inclusion/exclusion criteria of available data sources on Uku occurrence or density is necessary. The statistical approaches applied (boosted regression trees and generalized additive mixed-effect models with Tweedy distribution) are well-established methods that have been previously used for the spatial distribution modeling of marine living resources in different parts of the global ocean. The modeling procedures applied to derive both Level 1 and Level 2 products are correct and follow current statistical standards.

Obtained results allow for the reliable designation of EFH at Level 1. There are, however, certain methodological aspects of Level 1 product that need further deliberation or additional work. Since two separate models were used to model Uku occurrence in shallow and deep waters, there is a potential problem of spatial discontinuity in models' predictions and a lack of comparability between both models' outputs. Thus, I advise detailed inspection of these issues and consideration of the combined model. Another aspect that needs correction is the procedure used for the transformation of original predicted maps into categorical maps.

The statistical framework presented for the Level 2 constitutes a good baseline for the modeling of fish species distribution in MHI. Despite the fact, that the presented statistical framework for the Level 2 data is correct and comprehensive, the limitation of data available for the development of the Level 2 model is considered a serious threat to the reliability of the prediction of the model and its validity and in consequence for the determination of the EFH. Therefore, it is suggested that the exercise for Uku is repeated after the inclusion of additionally available data sets or when more data become available.

There are also important areas in which both Level 1 and Level 2 products can be improved. Particularly noteworthy is the appropriate description of prediction uncertainty, which should be presented in spatial dimension and constitute a complementary element to the maps of the occurrence or density of the MHI Uku population. In both cases, biological and oceanographical interpretation of the detected relationships should be provided, as it might serve as a basis for further studies on Uku biology and distribution.

2. EFH Level 1: distribution data (presence-absence)

2.1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?

Two separate data types were used in this modeling of Essential Fish Habitat for Uku at Level 1. National Oceanic and Atmospheric Administration (NOAA) fish diver surveys were used for shallow (0-30 m) depths and NOAA and University of Hawaii baited stereo-video camera arrays were used for deeper (30-300 m) areas. Based on the information provided during the WPSAR meeting (presentations by Audrey Rollo and Tye L. Kindinger on 12 July 2022, and further discussions), these data sets are considered the best available quantitative information on the Uku occurrence in the MHI. Similar types of data are commonly used for the development of spatial distribution models and the provisioning of fish species distribution maps in other marine areas (e.g., Knudby *et al.*, 2010a; Langlois *et al.*, 2020). In general, the author did excellent work in combining selected databases (MOUSS, BotCAM, and NOAA PIFSC BRUVs surveys, together with Stationary Diver Fish Surveys). However, there is a need for a more clear presentation of the inclusion/exclusion criteria applied. More complete information on the other, available but unused, datasets for MHI should be presented. A comprehensive list of all available datasets found by the author would be useful. The explanation of the problems and technical challenges faced that prevented the usage of particular data set in this modeling exercise would be of high value for future studies on the fish distribution in MHI aimed at the delineation of EFH. Such information might also help in the identification of future work priorities and improve existing monitoring surveys.

In summary, more details on the inclusion/exclusion criteria and explanations of why certain data sets were not included are expected. This information was provided during the discussions in the WPSAR meeting but is not present in the report. Without the proper documentation of these decisions, the reader is not able to unequivocally assess if the exclusion was justified.

However, I consider the inclusion of four selected data sets fully appropriate (see section 2.2). This is a short-term recommendation of high priority.

2.2. Are the data properly applied and appropriate for this species and habitat?

Yes, the data sets used in this study to model relationships between the probability of occurrence of the species and environmental factors are properly applied and appropriate. These types of data constitute one of the typical sources of georeferenced observations of living marine resources. Diver surveys are high-quality data products provided using a standardized methodology that covers most of the shallow waters in the studied area of MHI. Comprehensive camera-based surveys are the second type of data used for the deeper parts. Importantly, the monitoring data used in this modeling includes depth from ~3 m to ~314 m and fully capture the range of depths of the Uku distribution in MHI (in line with the previous EFH definition based on the literature review presented in WPRFMC, 2005, as well as in line with more current primary scientific sources, e.g. Asher *et al.*, 2017). The usage of two data types and specific data sets is well explained and technical details of these surveys are presented in the report.

2.3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?

Yes, applied models are appropriate to delineate EFH for Uku. Boosted regression trees are powerful machine learning algorithms that were used for similar tasks both in the terrestrial and marine systems (Elith *et al.*, 2008a). The author is experienced in the application of this statistical method and used this approach, e.g. for the prediction of coral distribution and abundance in the Hawaiian Islands (Franklin *et al.*, 2013). The author provided detailed information on the optimization, training, and validation processes in the report. The search for the optimal hyperparameters of the model seems correct and the procedure follows a previous published peer-reviewed study conducted by the author (Franklin *et al.*, 2013). Similarly, the training and validation phases follow established standards and are correct. Calculated AUC (area under the ROC curve) and other performance metrics (specificity, sensitivity) suggest the good performance of the model. For example, the AUC value was above 0.8 both for the shallow water model (AUC = 0.82) and deep water model (AUC = 0.86), and this level of accuracy is considered excellent according to the broadly accepted rules of thumb (Hosmer and Lemeshow, 2000, p. 164). Therefore, I conclude that the models can properly represent relationships between Uku occurrences and environmental variables, and their accuracy is high.

I suggest providing short information in the report on True Skill Statistic (TSS), Specificity, and Sensitivity that can be used in the interpretation of the results as “good” or “bad” - similarly to the description provided for the AUC metric. In the current form, only ranges of TSS (from -1 to 1) are indicated. Lobo *et al.*, 2008 and Mouton *et al.*, 2010 are potentially useful references.

2.4. Are decision points and input parameters reasonably chosen?

Yes, selected environmental variables are considered reasonable for this spatial modeling task. Selected environmental variables refer to the static features of the environment, such as depth, and relief-related parameters (rugosity, slope, etc.). It has been shown previously that these types of variables have high predictive power in the model of distribution of coral reef-associated fish (e.g. Knudby *et al.*, 2010a, but their importance in the prediction of coral distribution has been shown also by the author's previous studies, such as Franklin *et al.*, 2013). Moreover, dynamic environmental data were included in the model. These dynamic variables included maximum significant wave height and mean significant wave height for the climatology from 2010-2019, as well as the 8-day composite sea surface temperature (SST) and net primary productivity (NPP), obtained from the Aqua MODIS sensors. Comprehensive tests including, static, dynamic, or both static and dynamic factors have been conducted in the study on Uku Level 1 EFH, allowing for the selection of the best predictive model. This part of the model's development is very well presented in the report, easy to follow, and correct in my opinion.

My suggestion for this part, as a short-term recommendation of medium priority, is to provide a description of the rationale for why certain variables were included. A form of working hypothesis for each variable, together with the description of the expected shape and direction of the relationships could be useful in the identification of spurious correlations between response (probability of occurrence) and explanatory environmental variables. Such a summary can be useful for future modeling of fish distribution in MHI.

I suggest considering the inclusion of additional variables – especially variables which can be helpful in the improvement of model performance in the areas of high false positive classifications (which potentially can be the case in the human-impacted shallow bays, like Pearl Harbour - see e.g. Coles *et al.*, 1997; Coles, 2006). In order to fine-tune the model for these areas, additional data might be needed, as these areas are underrepresented in the data set. This is a mid-term recommendation of medium priority.

Another aspect of the modeling worth consideration is the application of different spatial and time “windows” used to aggregate the environmental data. Since the dynamic variables used in the modeling are represented on a temporal scale, an additional selection of a time window for each predictor is necessary. Typically, a priori selected time windows are considered, but there are potentially many plausible competing environmental signal hypotheses (van de Pol *et al.*, 2016; Smoliński, 2019; Smoliński *et al.*, 2021). Uku distribution can be driven by more direct (immediate) impacts of the environmental conditions showing short time lags. Alternatively, mean environmental conditions can shape the overall character of the habitat driving Uku distribution, thus longer time “windows” (e.g. mean values from the 6 months) can be better predictors of fish occurrence. Similarly, different spatial ranges can be considered for the aggregation of some selected environmental variables taking into account the information on the Uku movements (Meyer *et al.*, 2007). Currently available statistical approaches allow for the comparison of regression models supporting different hypotheses and the selection of the optimal prediction in a statistically rigorous manner (Bailey and van de Pol, 2016). However, an adaptation of the approach to the boosted regression trees requires manual programming as currently there is no ready-to-use software available. During the WPSAR meeting, these aspects were raised by the reviewers and discussed with the authors. Since the current version of the model shows good performance, this suggestion is considered a long-term research opportunity focused on improving the existing models, rather than a priority.

2.5. Are primary sources of uncertainty documented and presented?

No, this part of the study should be better addressed and reported.

Typically, SDMs are developed on spatially and temporally biased samples and then applied over much larger spatial scales. It is advisable to test the assumption that the statistical relationships between species observations and environmental predictors are applicable to other locations and times (Haulsee *et al.*, 2020). The information on the primary uncertainties in the data used for modeling can be unavailable and these types of uncertainties are in general hard to quantify. Thus, it is expected to appropriately acknowledge at least statistical uncertainties in the model predictions. Boosted regression trees are modeling approaches based on the independent bootstrapped iterations of the model fits, which compose an ensemble of trees in the final model (Elith *et al.*, 2008b). For this reason distribution of predictions obtained for each grid cell from single trees in the ensemble can be utilized to characterize uncertainty in the model predictions. While this approach does not address the uncertainties in the baseline data, it can provide a relative comparison of the prediction certainty in space, e.g. help to identify areas with higher uncertainty associated with the underrepresentation of the area in the data set, differences in the relationships between species occurrences and environmental variables or other environmental factors not incorporated in the modeling that contribute to the higher uncertainty in a particular area. The coefficient of variation of the ensemble predictions in each grid cell can be calculated to provide a visualization of spatial regions with higher uncertainty levels. For that purpose, the ModelMap package might be helpful, as it provides the features to read large GIS data in sections and maintain a reasonable usage of computer memory (Freeman *et al.*, 2010).

To provide a basic description of model uncertainty, my short-term recommendation of high priority for the author is to calculate AUC values and other performance metrics presented in the report, as well as the coefficient of variation of the ensemble predictions separately by area, i.e. island or windward/leeward subareas. Such a simple test can provide a general view of the area- or subarea-specific inaccuracies of the model and help to quickly identify potential regions of unreliable predictions (Smoliński and Radtke, 2017).

Although the importance of bias-reduced spatial resampling methods for performance estimation has been emphasized repeatedly in recent years, many studies in recent years neglect this problem (Brenning, 2012; Schratz *et al.*, 2019). This is the case also for the Uku Level 1 product. Therefore, a suggestion for the mid-term goal is to conduct area-stratified cross-validation (in spatial blocks) and a more in-depth evaluation of the model transferability (Wenger and Olden, 2012). Model validation with the consideration of the spatial context can provide valuable information on the biological and ecological variability in the region. Also, it can help to assess inaccuracies of the predictions caused by the differences in the spatial coverage of the data.

Model results show that the "*aspect*" variable was among the top variables in terms of predictive power. While the "*aspect"* per se has probably no biological meaning as a predictor, it may represent actual differences in the oceanographic properties between windward/leeward areas (as indicated in the materials and methods in the section *Study area*). For this reason, the exercise of the spatially-stratified cross-validation is also advised with the split of the areas (islands) into windward and leeward parts. These additional tests can shed a light on the possible differences in the species occurrence-environment relationships, but might also provide additional information that can be used in the designation of the Uku monitoring and selection of survey locations. In the process of EFH designation, confidence in the predictions is crucial, and an appropriate presentation of uncertainty is important for end-users to understand data reliability (Caldow *et al.*, 2015).

I suggest consideration of the potential biases in the dataset caused by the exclusion of records with "uncertain" categories: "Lutjanid/ae", "Perciformes", "Teleost", or "too dark to annotate". I don't think that they have any major impact on the final results, but they might have unequal distribution in the gradient of depth and somehow affect final predictions. At least information on the percentage of 'uncertain" observations and comments on this issue in the report would be helpful.

2.6. Are model assumptions reasonably satisfied?

Yes, model assumptions are reasonably satisfied. Boosted regression trees are among highly flexible data-driven methods (Elith *et al.*, 2008a) without the traditional statistical parameters (e.g. as in the traditional regression-based techniques). Since these methods are non-parametric, a minimal level of assumptions is to be made by the modeler. Tuning of model hyperparameters (tree complexity, learning rate, number of trees, and the bag-fraction) and selection of the optimal model separately for shallow and deep waters follow a well-established methodology (e.g. Knudby *et al.*, 2010a, 2010b). I think that all statistical requirements are satisfied for both shallow and deep models.

2.7. Are the final results scientifically sound, particularly delineating EFH boundaries?

Yes, obtained results can be used to provide new EFH definitions for Uku and are based on the best scientific information available.

However, considering the discontinuity of shallow and deep water model outputs, this aspect should be further inspected with the possibility of the usage of a combined model, which incorporates data sources (diver surveys or baited camera surveys) as an additional factor. Such a construction of the model may improve the model performance gained from the larger number of observations (two data sets combined). While the data on the occurrence of Uku along the depth gradient (top predictor in shallow and second most important in deep water submodel) should provide essentially the same information to the system, it may potentially help in the evaluation of the interactions between different variables and their nonlinearities which are not appreciated when using two models and separate shallow and deep data sets. In my opinion, the advantage of such an approach is that the effect of different survey methods and their interactions with other variables are assessed internally in the model and no inter-calibration factor is needed to correct raw values. Moreover, it is advised to consider all available data and take advantage of the observations from diver survey and baited camera survey in the overlapping range of depths – there are n=40 observations in the range 4-30m of baited camera surveys that were excluded in the current version of the study. With the combined model it is possible to provide maps of the mean probability of occurrence calculated from the predicted probability for each data source or method (factor incorporated in the model). It would be helpful to contrast predictions from the separate sub-models with the predictions of the combined model by e.g. calculation of the correlation coefficients and scatterplots of predicted values from the submodel and combined model separately for each depth category (shallow and deep). This additional exercise is recommended as a short-term task of high priority, followed by a careful evaluation of the model output. Depending on the result of this test, based on the model performance measures and reliability of the output, the author can consider the usage of such a combined model to delineate EFH for Uku.

Alternatively, the current sub-models for the shallow and deep waters can be used with a clear appreciation of the split into two sub-models and the possible spatial discontinuity of the predicted values. When interpreting the delineated EFH from both sub-models and applying this information in the management process, caution should be taken especially in the transition zone around 30m depth, where discontinuity might occur and results can be largely affected by the survey methods. If two separate models are to be used in the management process, it should be clearly stressed that the predictions and assigned categories of importance (e.g. top 25%, top 50%, top 95 %) shouldn't be compared between depth categories (between shallow and deep).

In the process of EFH delineation I suggest adopting the approach used in Alaska by the North Pacific Fishery Management Council (Laman *et al.*, 2018; NPFMC, 2020), which defined EFH based on the quantitative description:

"*EFH is defined in the Magnuson-Stevens Fishery Conservation and Management Act (Magnuson-Stevens Act) as "those waters and substrate necessary to fish for spawning, breeding, feeding, or growth to maturity." EFH for groundfish species is determined to be the general distribution of a species described by life stage. General distribution is a subset of a species' total population distribution, and is identified as the distribution of 95 percent of the species population, for a particular life stage, if life history data are available for the species. Where information is insufficient and a suitable proxy cannot be inferred, EFH is not described. General distribution is used to describe EFH for all stock conditions whether or not higher levels of information exist, because the available higher level data are not sufficiently comprehensive to account for changes in stock distribution (and thus habitat use) over time"* (NPFMC, 2020).

It is also advised to avoid descriptive terminology ("hot spots", "core habitats"), and instead use only basic EFH definitions and describe other areas using the respective percentile applied (i.e. "top 25%", "top 50%", "top 95%"="EFH").

In summary, the presented model at Level 1 constitutes a major improvement of the EFH descriptions over the previously adopted literature-based designation and provides important knowledge on the distribution of Uku in MHI. Obtained results can be used to delineate EFH boundaries. However, some additional changes and tests are recommended before the product is finalized.

2.8. Additional technical comments to the report

- Page 10, reconsider this sentence: “a lack of geographic and depth overlap between diver surveys and stereo-video camera surveys” because there are samples overlapping in the depth gradient.
- There is a lack of description of fisheries-dependent data.
- If the data are available, usage of catch per unit effort (CPUE, instead of total catch) to compare with model predictions would be advised. A simple regression test of mean model prediction for the probability of occurrence in the fishing area with the CPUE in the area would be useful.
- Provide information on the number and percentage of records with “Lutjanid/ae”, “Perciformes”, “Teleost”, or “too dark to annotate” that were not included in the analysis

3. EFH Level 2: habitat-specific densities

3.1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?

In the presented model for Level 2 Uku in MHI, one data set (the National Coral Reef Monitoring Program - NCRMP) was used for the modeling. The authors carefully selected this data set in order to meet all of the statistical assumptions of the model applied. Specifically, they were looking for the estimates of the species densities sampled in a randomized fashion in order to meet the criteria of the spatial independence of the observations. After the selection, the NCRMP data sampled using the paired-diver stationary point count method was selected. Other available data on the distribution of the Uku do not provide standardized information on the densities of the species or do cause other technical difficulties. While these aspects were raised during the WPSAR meeting, they are not presented in the report. For this reason, there is a need for a more clear presentation of the data sets inclusion/exclusion criteria applied. More complete information on the other, available but unused, datasets for MHI should be presented. Similarly, to the Level 1 product, it would be very useful to provide a comprehensive list of all available datasets found by the authors. The explanation of the problems and technical challenges faced that prevented the usage of particular data set in this modeling exercise would be of high value for future studies on the fish distribution in MHI aimed at the delineation of EFH. It might also help in the identification of future work priorities and improvements in the monitoring surveys.

In summary, more details on the inclusion/exclusion criteria and an explanation of why certain data sets were not included are expected. I recommend providing this information as a short-term high-priority task. Without the proper documentation of these decisions, the reader is not able to unequivocally assess if the exclusion was justified. However, I consider the data set included in the study appropriate for this modeling exercise (see section 3.2.).

3.2. Are the data properly applied and appropriate for this species and habitat?

Yes, data are properly applied, but they are not fully appropriate for this species and habitat as they are limited to shallow waters. It is known from the literature that Uku occurs in a broad range of depths from shallow areas to areas with >150 m depth (WPRFMC, 2005). Similarly, the description of the species' habitat provided in the Level 1 product (Franklin, 2021) shows a high probability of occurrence of Uku in the waters up to ~130 m. Therefore, current data limitation to the depths of 0-30 m prevents direct application of this product for the delineation of EFH (see section 3.3.).

The authors are to be congratulated for the transparency of the work and the ease of accessibility of the data used in the study. Since all R-code developed in this work and the full dataset are available through GitHub, a more in-depth inspection of the data and code is possible, helping in a more reliable review. The project file includes well-organized R scripts and datasets and is described so that consecutive steps of the analysis are easy to follow. It has been noticed that number of observations n=1484 was limited to n=1271 due to the i) exclusion of observations without environmental data and ii) exclusion of samples obtained above 30 m and below 0 m depth. The relatively low number of observations in the data available for the development of the Level 2 model is considered a severe additional limitation to the reliability of the prediction of the model and its validity and in consequence for the determination of the EFH.

I recommend as a medium-term with a high priority goal that the Level 2 density model be applied to sources of fishery-independent data that extend below 30 m. Inclusion of other data sources or newly available data to encompass the full depth range of the Uku species is needed.

3.3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?

Yes, chosen statistical approach is correct and constitutes the state-of-the-art framework for the analysis of the specie's density data. The model applied in the development of Level 1 product, implements geostatistical spatial and spatiotemporal random structures and is appropriate for these data (Anderson *et al.*, 2022). However, some aspects of the model development procedure need further improvement. While the ability of models based on the Tweedie distribution to handle zero-inflated data has been shown, other models, such as zero-inflated Poisson or quasi-Poisson, or negative binomial distribution should be considered. A formal model selection procedure based on the information criterion is advised in order to compare fits of models with different distributions assumed. In fact, the NCRMP survey data were collected using the paired-diver stationary point count method (McCoy *et al.*, 2019). Thus, primary Uku counts given per standard area (353 m2) can be modeled even with the Poisson distribution as they can be presented as whole numbers (zero and positive integers).

In the current form, the model seems to overpredict zeros and systematically underpredict positive values. These aspects of the model performance can be potentially improved with other distributions, but it needs more detailed investigation and should be tested statistically. Akaike Information Criterion can be used to compare different models in a similar way as fixed effects have been selected. I recommend a comparison of different models as a short-term goal of medium priority.

The selection of the fixed structures of the model based on the AIC is correct and follows current statistical standards. It allowed for the identification of the optimal fixed parameters for the prediction of Uku densities in the studied area. However, the same procedure should be applied based on the models fitted with a restricted maximum likelihood approach (REML, Zuur *et al.*, 2009) for the selection of the random effect structures. Spatial and spatiotemporal random effect structures should be tested in order to provide statistical support for the inclusion of these parameters (Bates *et al.*, 2015). While for example Barnett et al., 2021 found support for the inclusion of local trends in the majority of studied cases, it should be tested case-by-case. I recommend conducting such tests as a short-term high-priority task.

A more in-depth description of the developed model is needed, including the presentation of the model's parameter estimates. Among others, the estimated variance associated with each random effect (and possibly also Intraclass Correlation Coefficient) should be presented. Moreover, marginal R^2 (which provides the information on the variance explained only by fixed effects) and conditional R^2 (which provides the variance explained by the entire model, i.e., both fixed effects and random effects) should be presented (Nakagawa and Schielzeth, 2013). These metrics can be used to assess the leading component of variance and evaluate the part of variance that cannot be explained by the covariates used. Following Thorson (2015) and Grüss *et al.* (2020), covariates can be rescaled to have a mean of 0 and a standard deviation of 1, which facilitates comparison between covariates of different units and allows for comparison between the estimated coefficients and the marginal standard deviation of spatial and spatiotemporal variation (Lindmark *et al.*, 2022). This information is important for the application of the modeling results in the delineation of EFH and management. As an example, in an extreme case, variance observed in the density of Uku can be dominated by the local trends, which call for serious caution at the stage of EFH designation. Thus, provisioning of the complete model description is recommended as a short-term high-priority goal.

As a minor note, there is a lack of description of linear models used for the assessment of the directionality of the temporal trends presented in Figure 5. I assume that this test is based on the output of the first model and only point estimates are used as an input into the separate linear models to test the significance of slopes in four areas. While this test provides a reasonable overview, using point estimates in the second test has some caveats (Hadfield *et al.*, 2010; Houslay and Wilson, 2017). In future studies, I advise considering the inclusion of continuous Year effects (optionally in the interaction with factor area) in the fixed part of the model and testing the presence of temporal trend internally in the model.

3.4. Are decision points and input parameters reasonably chosen?

Yes, the environmental variables selected for this modeling are appropriate and have potential biological meaning. Utilized sources of environmental data provide the best available scientific information. My suggestion for this part, as a short-term recommendation of medium priority, is to provide a description of the rationale for why certain variables were included. A form of working hypothesis for each variable, together with the description of the expected shape and direction of the relationships could be useful for future modeling of fish distribution in MHI.

The framework of the Environmental Data Summary (EDS) and the tools developed by the authors represent a major improvement in data aggregation techniques. The inclusion of environmental data statistics other than the mean value (standard deviation, upper and lower percentiles) is noteworthy, as these aspects are often overlooked in ecology. Also, in this case, full transparency and availability of the R code constitute the great added value of the study.

Similar to the Level 1 product, different spatial and time "windows" used to aggregate the environmental data can be considered (see section 2.4.). For example, according to Meyer *et al.* (2007), the tagged Uku exhibited diel and tidal habitat shifts, with the latter resulting in round trips of up to 24 km in 24 h. Summary statistics provided for different areas and periods can be stored as a matrix and further utilized as alternative explanatory variables. Methods for testing different spatial or time "windows" in a systematic way using statistical comparison of models can also be potentially incorporated into the procedures. Since fitting multiple models using sdmTMB can be computationally heavy, alternative packages based on the Integrated Nested Laplace Approximation (INLA) can be considered (Rue *et al.*, 2009). These aspects were raised and discussed with the first author during the WPSAR meeting. Similar to the Level 1 product, this suggestion is considered a long-term research opportunity focused on improving the existing methods, rather than a recommendation of priority.

3.5. Are primary sources of uncertainty documented and presented?

No, this part of the study should be better addressed and reported.

The model output should be better described, including typical summary statistics, like standard errors of the parameter estimates, which give a general overview of the uncertainty in the model. Moreover, summary statistics of the accuracy of the prediction by area or subarea (windward/leeward) can be provided as a first step. The description of uncertainties should aim at the full quantitative evaluation of the model's performance. The information on the prediction uncertainty should be presented in spatial dimension and constitute a complementary element to the maps of the density of the MHI Uku population. In the case of the Level 2 product, uncertainty in the spatial predictions could be monitored using bootstrap approaches. I refer to section 2.5, which describes the methods of validation (spatial blocks) and presentation of uncertainty in the results of modeling that can be applied in the Level 2 product.

3.6. Are model assumptions reasonably satisfied?

Yes, in general, model assumptions seem to be satisfied. However, some discrepancies were observed in the quantile-quantile plots presented during the WPSAR meeting. Also, the scatterplot of the predicted versus observed values suggests that the model overpredicts the density when no fish are observed, while underpredicting positive values. Careful tests for overdispersion and zero-inflation are advised to check if the model assumptions are fully satisfied (Zuur and Ieno, 2016). Simulation-based approaches can be useful here (Hartig, 2020). I recommend these additional tests as short-term high priority tasks.

It is also recommended to test how good the model is in the predictions of zero by the calculation of AUC (assuming zero and non-zero groups). These aspects of the model performance are critical for the delineation of EFH since false positive predictions can extend putative habitats into areas of minor importance for the species (the prediction may indicate positive densities while the species is absent or the probability of occurrence is negligible). Systematic underprediction of positive values is not so impactful, because some kind of rescaling is done at the very late stage when presenting maps, but zeros are critical as they can overvalue habitats actually not preferred by the species. This simple test should be included in the previous recommendation and considered a short-term high priority task.

3.7. Are the final results scientifically sound, particularly delineating EFH boundaries?

Yes, a scientifically sound statistical approach was presented in the Level 2 report. This framework can be used in future studies on other species in the MHI region. However, currently, the data limitation prevents robust usage of the model for the delineation of EFH boundaries for MHI Uku. Most importantly, the data is limited to the most shallow waters from 0 to 30 m. While Uku occurs in this range of depths, the species prefer deeper waters and therefore shallow areas are not the essential habitats for it (Meyer *et al.*, 2007; Franklin, 2021). Since the data from the deeper areas were not included in the modeling, I think that the sentence in the executive summary (line 116: "*The results indicate shallow-water habitats in the MHIs are likely essential for A. virescens*" is not well supported).

There are also numerical problems in the data set. Among the 1271 observations used for the modeling, ~89% were zeros and only n=143 (~11%) observations showed positive values. Such distribution of the survey data causes serious challenges in the modeling. In my opinion, such a limited data set calls for caution and do not provide sufficient baseline data for the delineation of EFH. The problem of data limitation was also stressed by the authors (line 155 of the report). This problem is not uncommon for many marine species as preferred habitat and species distribution data are often sparse (Moore *et al.*, 2016).

The developed framework for Level 2 provides predictions of the Uku densities. If the product is to be used for the delineation of EFH in the future, the adoption of EFH definition and preparation of the categorical maps, similar to the Level 1 model (see section 2.7.), would be advised (Laman *et al.*, 2018; NPFMC, 2020). However, in my opinion, currently, the data limitation prevents robust usage of the model for the delineation of EFH boundaries, thus further development of the categorical maps at this stage is not necessary from a short-term perspective. I recommend the preparation of the procedures to transform density predictions into categorial maps as a mid-term medium-priority task.

When more reliable information on the density of the species is available, it is recommended to investigate the relationships between density and the probability of occurrence of the species. Models of the relationships between density and probability of occurrence can potentially offer a cost-effective alternative to direct measurements.

In summary, data limitation prevents reliable usage of this model at the current stage. Therefore, more data on the Uku densities should be collected to provide a solid background for the spatial distribution model of Uku in MHI.

3.8. Additional technical comments to the report

- Lines 167-168: please be more precise (as there was at least one previous study: Franklin, 2021).
- Fig. 2b: indicate what is presented with error bars
- Fig. 2c: it is not clear why discrete scales are presented in this figure. It needs more explanation in the caption
- Line 193: please correct the number of observations reported
- Table 1: temporal range or temporal resolution?
- Line 244: reconsider if these are continuous values
- Line 298: it seems that part of the sentence is missing. Personally, I do not agree that the model performance is close to ideal (see comments in sections 3.3. and 3.6.). Also, the second part of the sentence "variability in model accuracy increased at higher density" is imprecise. I suggest rewriting the whole sentence.
- Figure 4a: it is not clear what is presented in the plot as ranges of color gradients do not match maximal values reported above the panels. Please clarify.
- Figure 5: please describe beta and p in the caption. Confidence intervals are described in the caption but not present. As mentioned in the previous part of the review report, the note on the models used to describe temporal trends should be included in the methods section.
- Line 205: please provide the reasoning for using one-month time windows
- Lines 316-318: This sentence is not clear to me.
- The full bibliographic data for Tanaka and Oliver 2021 is not in the list of references
- Figure S2: red dots show land masses, not green. What then is the distinction between green and blue dots?

4. General comment on the research priorities

- In the future process of EFH designation, different roles of habitats (e.g., spawning habitats, feeding habitats, nursery grounds) should be considered. These aspects can be partly addressed by the development of size-stratified or stage-specific spatial distribution models and the definition of separate EFH. For this reason, estimation of the fish lengths from the available video recordings and utilization of other available data is advised.
- There is a considerable lack of knowledge on the Uku seasonal migrations and distribution of spawning aggregations. Further research on the seasonal patterns in the species distribution is recommended.
- There is a need for additional data on the distribution of Uku in the mid-depth hard habitats.
- Despite the different purposes of Level 1 and Level 2 products, density or biomass variability can be potentially approximated as a function of the probability of occurrence. Therefore, further investigation of occurrence-density relationships is advised. While direct measurements of densities and predictions at the Level 2 can be difficult to obtain, the relationships between occurrence and density (or biomass) can be used to obtain proxies. This kind of approximation can be a useful and cost-effective basis for the delineation of the EFH.
- The concepts of ensemble approaches can be explored in future studies to compare the output of different models. It can provide additional information on method-specific uncertainty. Such exercises help also in the comparison of species-environment relationships identified by different models and allow for detecting statistical errors (e.g. spurious correlations, unreliable relationships). Further, the application of multiple model techniques in ensemble predictions can reduce the error associated with any given model and technique so that higher predictive accuracy is achieved.

5. Notes to the ToRs

- Since terms of reference (ToRs) were adopted from the previous WPSAR process focused on the stock assessment procedures, there are certain phrases, like “model assumptions reasonably satisfied” or “input parameters”, which could be fine-tuned to better reflect the components of the works on spatial modeling and EFH determination. Therefore, I suggest a reconsideration of the wording in ToRs to better address the specificity of the tasks evaluated.
- Some of the ToRs, like “Are the final results scientifically sound, particularly delineating EFH boundaries?” combine multiple aspects that can be split into separate ToRs. Since reviewers are obliged to answer “yes” or “no”, splitting currently formulated questions into separate ToRs could help in the easier description of the strengths and weaknesses of the evaluated works in the review process.

6. List of references

Anderson, S. C., Ward, E. J., English, P. A., and Barnett, L. A. K. 2022. sdmTMB: an R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. bioRxiv: 1–17.

Asher, J., Williams, I. D., and Harvey, E. S. 2017. An Assessment of Mobile Predator Populations along Shallow and Mesophotic Depth Gradients in the Hawaiian Archipelago. Scientific Reports, 7: 1–18. Springer US.

Bailey, L. D., and van de Pol, M. 2016. climwin: An R Toolbox for Climate Window Analysis. PLoS ONE, 11: 1–27.

Bates, D., Kliegl, R., Vasishth, S., and Baayen, R. H. 2015. Parsimonious Mixed Models. arXiv. https://arxiv.org/abs/1506.04967.

Brenning, A. 2012. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. International Geoscience and Remote Sensing Symposium (IGARSS): 5372–5375.

Caldow, C., Monaco, M. E., Pittman, S. J., Kendall, M. S., Goedeke, T. L., Menza, C., Kinlan, B. P., *et al.* 2015. Biogeographic assessments: A framework for information synthesis in marine spatial planning. Marine Policy, 51: 423–432.

Coles, S. L., Carlton, R. C., G., D. L., and T., E. J. 1997. Biodiversity of marine communities in Pearl Harbor, Oahu, Hawaii with observations on introduced exotic species. http://hbs.bishopmuseum.org/pdf/PHReport.pdf.

Coles, S. L. 2006. Marine communities and introduced species in Pearl Harbor, O’ahu, Hawai’i. The Environment in Asia Pacific Harbours: 207–228.

Elith, J., Leathwick, J. R., Hastie, T., and R. Leathwick, J. 2008a. Elith, Leathwick & Hastie A working guide to boosted regression trees - Online Appendices Page 1. Journal of Animal Ecology, 77: 802–13. http://www.ncbi.nlm.nih.gov/pubmed/18397250.

Elith, J., Leathwick, J. R., and Hastie, T. 2008b. A working guide to boosted regression trees. The Journal of Animal Ecology, 77: 802–813.

Franklin, E. C., Jokiel, P. L., and Donahue, M. J. 2013. Predictive modeling of coral distribution and abundance in the Hawaiian Islands. Marine Ecology Progress Series, 481: 121–132.

Franklin, E. C. 2021. Model-based Essential Fish Habitat Definitions for the Uku Aprion virescens in the Main Hawaiian Islands by.

Freeman, E. a, Frescino, T. S., and Moisen, G. G. 2010. ModelMap : an R Package for Model Creation and Map Production.

Grüss, A., Gao, J., Thorson, J. T., Rooper, C. N., Thompson, G., Boldt, J. L., and Lauth, R. 2020. Estimating synchronous changes in condition and density in Eastern Bering Sea fishes. Marine Ecology Progress Series, 635: 169–185.

Hadfield, J. D., Wilson, A. J., Garant, D., Sheldon, B. C., and Kruuk, L. E. B. 2010. The misuse of BLUP in ecology and evolution. American Naturalist, 175: 116–125.

Hartig, F. 2020. Residual diagnostics for hierachical (multi-level/mixed) regression models. https://cran.r-project.org/web/packages/DHARMa/DHARMa.pdf.

Haulsee, D. E., Breece, M. W., Fox, D. A., and Oliver, M. J. 2020. Simple is sometimes better: a test of the transferability of species distribution models. ICES Journal of Marine Science, 77: 1752–1761.

Hosmer, D., and Lemeshow, S. 2000. Applied Logistic Regression. John Wiley & Sons.

Houslay, T. M., and Wilson, A. J. 2017. Avoiding the misuse of BLUP in behavioural ecology. Behavioral Ecology, 28: 948–952.

Knudby, A., LeDrew, E., and Brenning, A. 2010a. Predictive mapping of reef fish species richness, diversity and biomass in Zanzibar using IKONOS imagery and machine-learning techniques. Remote Sensing of Environment, 114: 1230–1241.

Knudby, A., Brenning, A., and LeDrew, E. 2010b. New approaches to modelling fish–habitat relationships. Ecological Modelling, 221: 503–511.

Laman, E. A., Rooper, C. N., Turner, K., Rooney, S., Cooper, D. W., and Zimmermann, M. 2018. Using species distribution models to describe essential fish habitat in Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1230–1255.

Langlois, T., Goetze, J., Bond, T., Monk, J., Abesamis, R. A., Asher, J., Barrett, N., *et al.* 2020. A field and video annotation guide for baited remote underwater stereo-video surveys of demersal fish assemblages. Methods in Ecology and Evolution, 11: 1401–1409.

Lindmark, M., Anderson, S. C., Gogina, M., Casini, M., Station, P. B., Canada, O., and Lindmark, M. 2022. Evaluating drivers of spatiotemporal individual condition of a bottom-associated marine fish. bioRxiv, 46: 1–27. https://doi.org/10.1101/2022.04.19.488709.

Lobo, J. M., Jiménez-valverde, A., and Real, R. 2008. AUC: A misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography, 17: 145–151.

McCoy, K., Asher, J., Ayotte, P., Gray, A., Kindinger, T., and Williams, I. 2019. Pacific Reef Assessment and Monitoring Program Data Report - Ecological monitoring 2019 - Reef fishes and benthic habitats of the main Hawaiian Islands: PIFSC data report DR-19-039: 1–2.

Meyer, C. G., Papastamatiou, Y. P., and Holland, K. N. 2007. Seasonal, diel, and tidal movements of green jobfish (Aprion virescens, Lutjanidae) at remote Hawaiian atolls: Implications for marine protected area design. Marine Biology, 151: 2133–2143.

Moore, C., Drazen, J. C., Radford, B. T., Kelley, C., and Newman, S. J. 2016. Improving essential fish habitat designation to support sustainable ecosystem-based fisheries management. Marine Policy, 69: 32–41. Elsevier. http://dx.doi.org/10.1016/j.marpol.2016.03.021.

Mouton, A. M., De Baets, B., and Goethals, P. L. M. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling, 221: 1995–2002. Elsevier B.V. http://dx.doi.org/10.1016/j.ecolmodel.2010.04.017.

Nakagawa, S., and Schielzeth, H. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution, 4: 133–142.

NPFMC. 2020. Fishery management plan for groundfish of the Gulf of Alaska 2015. https://www.fisheries.noaa.gov/management-plan/groundfish-gulf-alaska-management-plan.

Rue, H., Martino, S., and Chopin, N. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 71: 319–392.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., and Brenning, A. 2019. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. Ecological Modelling, 406: 109–120. Elsevier. https://doi.org/10.1016/j.ecolmodel.2019.06.002.

Smoliński, S., and Radtke, K. 2017. Spatial prediction of demersal fish diversity in the Baltic Sea: comparison of machine learning and regression-based techniques. ICES Journal of Marine Science, 74: 102–111.

Smoliński, S. 2019. Incorporation of optimal environmental signals in the prediction of fish recruitment using random forest algorithms. Canadian Journal of Fisheries and Aquatic Sciences, 76.

Smoliński, S., Langowska, A., and Glazaczow, A. 2021. Raised seasonal temperatures reinforce autumn Varroa destructor infestation in honey bee colonies. Scientific Reports, 11: 22256. Nature Publishing Group UK. https://doi.org/10.1038/s41598-021-01369-1.

Sundblad, G., Bergström, U., Sandtrom, A., and Eklöv, P. 2014. Nursery habitat availability limits adult stock sizes of predatory coastal fis. ICES Journal of Marine Science, 71: 672–680.

Tanaka, K. R., Schmidt, A. L., Kindinger, T. L., Whitney, J. L., and Samson, J. C. 2022. DRAFT : Spatiotemporal assessment of Aprion virescens density in shallow Main Hawaiian Islands waters, 2010-2019.

Thorson, J. T. 2015. Spatio-temporal variation in fish condition is not consistently explained by density, temperature, or season for California Current groundfishes. Marine Ecology Progress Series, 526: 101–112.

van de Pol, M., Bailey, L. D., McLean, N., Rijsdijk, L., Lawson, C. R., Brouwer, L., and Gimenez, O. 2016. Identifying the best climatic predictors in ecology and evolution. Methods in Ecology and Evolution, 7: 1246–1257.

Wenger, S. J., and Olden, J. D. 2012. Assessing transferability of ecological models: An underappreciated aspect of statistical validation. Methods in Ecology and Evolution, 3: 260–267.

WPRFMC. 2005. Essential Fish Habitat Descriptions for Western Pacific Archipelagic and Remote Island Areas Fishery Ecosystem Plan Management Unit Species (Crustacean, Bottomfish, Precious Coral and Coral Reef Ecosystem Species).

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., and Smith, G. M. 2009. Mixed effects models and extensions in ecology with R. Springer, New York. 580 pp.

Zuur, A. F., and Ieno, E. N. 2016. A protocol for conducting and presenting results of regression-type analyses. Methods in Ecology and Evolution, 7: 636–645.

Annex III. Murillo, J. 2022. External Independent Peer Review under the Western Pacific Stock Assessment Review framework. Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*)

External Independent Peer Review under the Western Pacific Stock Assessment Review framework

Level 1 and Level 2 Essential Fish Habitat Models for the Main Hawaiian Islands Uku (*Aprion virescens*)

Javier Murillo
Department of Fisheries and Oceans Canada
Dartmouth, Nova Scotia, Canada
Javier.murillo-perez@dfo-mpo.gc.ca

Dartmouth, NS, July 2022

Executive summary

Two documents were prepared for review under the WPSAR process (Franklin 2021 and Tanaka et al. 2022). Both documents try to delineate essential fish habitat (EFH) through the use of species distributions modelling for the Main Hawaiian Islands Uku (*Aprion virescens*). Franklin (2021) provides a level 1 EFH based on two presence-absence models using boosted regression trees created for the full extent of the habitat of the species. Whereas, Tanaka et al. (2022) provides a level 2 EFH based on density data using generalized additive mixed-effect model for only the shallow portions of the habitat. Both documents use the same data sources for the shallow part but different environmental covariates and modelling approaches. They answer two separate questions on varying levels of EFH and therefore two independent reviews have been provided.

1.- Franklin, E.C. 2021. Model-based essential fish habitat definitions for the Uku Aprion virescens *in the Main Hawaiian Islands. Report to Western Pacific Regional Fishery Management Council. 34 pp.*

Overview:

This document provides Level 1 EFH based on two presence-absence models created at different depths. The author provided a thoughtful review of the data available and through the use of species distribution modelling studied the relationship of Uku (*Aprion virescens*) with different habitat covariates. The document is well written and structured and the results constitute a great improvement over the existing EFH descriptions. Additionally, the author provided a fantastic presentation of the results and addressed all the comments during the 3-day WPSAR process. For this review I let myself be guided by the suggested Terms of Reference, and have provided responses related to each one below.

1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?

Yes with caveats. The author did a good review of the literature and data availability and documented properly the response data used, both the shallow diver surveys and deep baited stereo-video cameras. It seems that there are some diver survey data from the Hawaii Division of Aquatic Resources and Nature Conservancy that were not included in the shallow model likely due to the lack of QA/QC procedures and other reasons. However, although this data may not be appropriate for inclusion in the model, they still represent an independent data set that could be used for model validation of the shallow part of the model.
The selection of habitat covariates was appropriated and justified. However, there are more dynamic variables such as temperature (only included in the deep model) and salinity that could be important in explaining the distribution patterns of Uku.
Suggested recommendation: Evaluate the performance of the model (sensitivity and specificity) in the shallow waters using the survey data from Hawaii Division of Aquatic Resources and Nature Conservancy. (*short/immediate term and mid priority*).
Evaluate the use of other dynamic variables, such as temperature, salinity and others in the model. (*short/immediate term and mid priority*).

2. Are the data properly applied and appropriate for this species and habitat?

Yes. The author used fishery-independent data not biased spatially and habitat covariates that are expected to have an influence in the distribution of this species.

3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?

Yes with caveats. The models were properly applied, adequate, and appropriate for the species and available data. However, by providing two models for the shallow and deep waters separately, the shallow model predicts maximum probability of occurrence in the deepest areas (30 m; Figure 12), whereas the deep model predicts the maximum probability around the 60-80

m depth (Figure 13).
This discrepancy may create some discontinuity or spurious effect around the 20-30 m depth.

Suggested recommendation: Provide a model combining the shallow and deep waters data. (*short/immediate term and high priority*). It can be argued that the two different data sets should not be combined. However, it is common in the literature to combine data from different sources for presence-absence models (e.g. Morato et al. 2020; Beazley et al. 2021), and more in poor data areas. Additionally, the prevalence from both data sets do not differ substantially (14.9% for shallow data and 14.2% for deep data which ranges from 5.2% and 26.2% depending of the video survey used). In fact, there is more variation of prevalence between the difference video surveys than between diver surveys (14.9%) and the average of video surveys (14.2%±10.8). This could be for differences between the systems or related to the spatio-temporal differences in the distribution of the species.

4. Are decision points and input parameters reasonably chosen?

Yes with caveats. The input parameters were reasonably chosen. However, some decision points in the model approach should be improved. Threshold used to calculate sensitivity and specificity is not discussed. Going from probability to classification maps is not done in the same way than Laman et al (2018) and when this is done prevalence is very important and this does seem to have been considered here.

Suggested recommendation: Use a prevalence threshold or optimal threshold in model validation for the calculation of sensitivity and specificity. (*short/immediate term and high priority*).

Both Uku shallow and deep data are characterized by a higher number of absences relative to presences (i.e. unbalanced species prevalence). Classification accuracy in decision trees is prone to bias towards the majority class when the categorical response variable is highly imbalanced (Chen et al., 2004). This is due to over-representation of the majority class in the bootstrap sample leading to a higher frequency in which the majority class is drawn, therefore skewing predictions in that favor. Several different approaches have been used to address imbalanced data: 1) assign a high cost to misclassification of the minority class, 2) down-sample the majority class, and 3) up-sample the minority class (Evans et al., 2011). Although several studies suggest a balanced modelling prevalence of 0.5 (McPherson et al., 2004; Liu et al., 2005), this approach may result in a loss of information particularly for rare species, and may not be necessary when the model training data is reliable and not biased spatially and/or environmentally (Jiménez-Valverde and Lobo, 2006). Another widely-used approach is to adjust the threshold used to divide the probabilistic predictions of occurrence into discrete predictions of presence or absence, to match modelling prevalence (Liu et al., 2005). The latter approach has shown to produce constant error rates and optimal model accuracy measures compared to balancing modelling prevalence (Liu et al., 2005; Hanberry and He, 2013). Other approaches use a threshold that maximize sensitivity and specificity (e.g. Morato et al. 2020, Beazley et al. 2021). The function 'optimal.thresholds' from the PresenceAbsence R package (Freeman and Moisen, 2008; Freeman 2022) provides different options.

5. Are primary sources of uncertainty documented and presented?

No. No uncertainty documented for the response data (uncertainty in the species identification), predictors or results.

Suggested recommendation: Show extrapolated areas. (*short/immediate term and high priority*). The reliability of predictions from tree-based models decreases in areas of extrapolation outside of the domain of the environmental predictors (Liu et al., 2020). Areas of extrapolation can be shown to evaluate the reliability of the predictions and show the habitat covariates causing this extrapolation (e.g Wang et al. 2022). High probability in non-extrapolated areas would have more certainty that the same probability associated with extrapolated areas. The Extrapolation Detection (ExDet) tool from the 'dsmextra' R package (Bouchet et al., 2020) could be used.
Provide a map of uncertainty in the predictions. (*short/immediate term and high priority*). There are several ways of doing that. Some authors have used different models and averaged the results in an ensemble model with the associated variation (e.g. Georgian et al. 2019). Other option is to create some uncertainty from the different predictions in each tree or to provide a cross-validation based on spatial blocks which reduces the effects of spatial autocorrelation on model prediction error (Roberts et al. 2017). However, for this work, if a single model combining the shallow and deep water is provided, different models can be done with the different data sets, leaving one out every time. There is one data set of diver surveys (D, diver) and 3 from underwater videos (V, videos). A full model combining all data could be done to create the final probability surface. However, for validation, different data sets could be created using 3 data sources each time (DV_1V_2, DV_1V_3, DV_2V_3, $V_1V_2V_3$). This would create 4 models for validation that could be averaged and a surface with the standard deviation of these 4 models created. Areas with higher standard deviation would be associated with higher uncertainty.
Provide the original resolution of the habitat covariates before being resampled to the 50 x 50 m raster grids. (*short/immediate term and high priority*).
Provide more information in the uncertainty of the species identification (*short/immediate term and mid priority*). During one of the presentations (A. Rollo – Pacific Islands Fisheries Science Center Bottomfish Data Collection Methods) it was mentioned some classes as possible uku or too dark to annotate. This adds uncertainty to the response data used and this information (e.g. % of observations) should be commented in the methods.

6. Are model assumptions reasonably satisfied?

Yes. Boosted regression tree models is a non-parametric technique with little assumptions which is one of the strength of the method.

7. Are the final results scientifically sound, particularly delineating EFH boundaries?

Yes with caveats. By doing two separate models, it is possible that the deep waters in the shallow model (~30 m) have higher probability of occurrence than the shallower areas (~40-50 m) in the deep model. This may show a discontinuity in the probability of occurrence and be an artifact. Looking at Figure 12 is seems that maximum probability in relation to depth increases monotonically and reaches a maximum around 60-80 m depth.

Following Laman et al. (2018) the categorical maps should be done based on the quantiles from the raster cell distribution rather than from the probability. However, it is recommended not using these classes and creating just two classes, one delineating EFH boundaries and another indicative of non-EFH.

Suggested recommendation: Modify the approach used to delineate EFH boundaries. (*short/immediate term and high priority*). Once that a single model combining the shallow and deep data is created, the threshold (prevalence, MSS, etc.) used in model validation for the calculation of sensitivity and specificity (ToR 4) should also be used to classify the probabilistic map in a binary map where one of the classes (above the threshold) would define the suitable habitat for Uku that would be equivalent to the EFH, whereas below the threshold would define non-suitable habitat and therefore non-EFH (e.g. Morato et al. 2020; Beazley et al. 2021; Wang et al. 2022).
Areas of higher probability inside the EFH supported by presence records can help managers to visualize more important areas inside the EFH or in the future prediction of abundance/density could help to create a EFH level 2 from this EFH level 1. (*mid-term and mid priority*).

8. Can the results be used to address management goals stated in the relevant FEP or other documents provided to the review panel? If any results of these models should not be applied for management purposes with or without minor short-term further analyses (in other words, if any responses to any parts of questions 1-7 are "no"), indicate:

Which results should not be applied and describe why, and

Based on the previous ToRs, the model can be used to address management goals for EFH Level 1 with suggested recommendations.

Which alternative set of existing EFH definitions should be used to inform uku EFH instead and describe why.

Level 0 EFH should be used until the suggested recommendations (*short/immediate term and high priority*) are provided.

Other minor comments;
- Add 30 and 300-m bathymetry lines to Figures 8 to 11.
- Add one figure with the presence-absence records overlaying the final probability map.
- Add Friedman and Meulman (2003) and Pyle et al. (2016) to the References section.
- Provide a figure in the Appendix or supplements with the correlation values between all the environmental variables used.

Research priorities:
- Understanding habitat use of other life stages (<9cm) and improve separation between sub/adult and adult classes from the deep video surveys. (*mid-term and mid priority*).

- Exploring other environmental variables (e.g. current speed, fishing effort, mixed layer depth, etc.) that can help to explain the observed distribution patterns of Uku in the Main Hawaiian Islands. (*mid-term and mid priority*).
- Validation of the models in predicted EFH where no data are available through future sampling surveys (*mid-term and mid priority*).
- Exploring other data sources such as recreational fisheries, iNaturalist, citizen science, etc., that could provide additional records for model validation. https://www.inaturalist.org/guide_taxa/928917#:~:text=Summary&text=The%20green%20jobfish%2C%20Aprion%20virescens,180%20ft%20(55%20m). (*long-term and low priority*).

References

Beazley, L., Kenchington, E., Murillo, F.J., Brickman, D., Wang, Z., Davies, A.J., Martyn Roberts, E., and Rapp, H.T. 2021. Climate change winner in the deep sea? Predicting the impacts of climate change on the distribution of the glass sponge *Vazella pourtalesii*. Marine Ecology Progress Series, 657: 1-23.

Bouchet, P. J., Miller, D. L., Roberts, J. J., Mannocci, L., Harris, C. M. and Thomas, L. 2020. Dsmextra: Extrapolation Assessment Tools for Density Surface Models. Methods Ecol. Evol. 11, 1464–1469.

Chen, C., Liaw, A., and Breiman, L. 2004. Using random forest to learn imbalanced data. University of California, Berkeley. 12 p.

Evans J.S., Murphy, M.A., Holden, Z.A., and Cushman, S.A. 2011. Modeling Species Distribution and change Using Random Forests. *In* Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications. Edited by C.A Drew, Y.F. Wiersma, and F. Huettmann. Springer, New York. pp. 139–159.

Freeman, E.A., Moisen, G. 2008. “PresenceAbsence: An R Package for Presence Absence Analysis.” Journal of Statistical Software, 23(11), 1–31.

Freeman, E. 2022. Package ‘PresenceAbsence’ version 1.1.10. https://cran.r-project.org/web/packages/PresenceAbsence/PresenceAbsence.pdf

Georgian, S.E., Anderson, O.F., Rowden, A.A. 2019. Ensemble habitat suitability modeling of vulnerable marine ecosystem indicator taxa to inform deep-sea fisheries management in the South Pacific Ocean. Fisheries Research, 211: 256-274.

Hanberry, B.B., and He, H.S. 2013. Prevalence, statistical thresholds and accuracy assessment for species distribution models. Web Ecology, 13: 13–19.

Jiménez-Valverde, A. and Lobo, J. M. 2006. The ghost of unbalanced species distribution data in geographical model predictions. Divers. Distrib., 12: 521–524.

Laman, E.A., Rooper, C.N., Turner, K., Rooney, S., Cooper, D.W., Zimmermann, M. 2017. Using species distribution models to describe essential fish habitat in Alaska. Canadian Journal of Fisheries and Aquatic Sciences, 75: 1230-1255

Liu, C., Wolter, C., Xian, W. and Jeschke, J. M. 2020. Species Distribution Models Have Limited Spatial Transferability for Invasive Species. Ecol. Lett. 23, 1682–1692.

McPherson, J.M., Jetz, W., and Rogers, D.J. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artifact? J. Appl. Ecol., 41: 811–823.

Morato, T., González-Irusta, J.-M., Dominguez-Carrió, C., Wei, C.-H., Davies, A., Sweetman, A.K., et al. 2020. Climate-induced changes in the suitable habitat of cold-water corals and commercially important deep-sea fishes in the North Atlantic. Global Change Biology, 26: 2181-2202.

Liu, C., Berry, P.M., Dawson, T.P., and Pearson, R.G. 2005. Selecting thresholds of occurrence in prediction of species distribution. Ecography 28: 385–393.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S. and others. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40: 913−929.

Wang, S., Murillo, F.J. and Kenchington, E. 2022. Climate-Change Refugia for the Bubblegum Coral *Paragorgia arborea* in the Northwest Atlantic. Frontiers in Marine Science. doi: 10.3389/fmars.2022.863693.

2.- Tanaka, K.R., Schmidt, A.L., Kindinger, T.L., Whitney, J.L., and Samson, J.C. 2022. Spatiotemporal assessment of Aprion virescens *density in shallow Main Hawaiian Islands waters, 2010-2019. U.S. Dept. of Commerce, NOAA Technical Memorandum NMFS-PIFSC-##, 29 p*

Overview:

This document provides Level 2 EFH for shallow coastal waters of Uku (*Aprion virescens*). The authors studied the relationship between density data and different environmental variables considering as well the spatio-temporal effect through the use of a generalized additive mixed-effect model. The document is well written and structured and the results improve the existing EFH description for the shallow part of the habitat of Uku. Additionally, the first author provided a fantastic presentation of the results and addressed all the comments during the 3-day WPSAR process. For this review I let myself be guided by the suggested Terms of Reference, and have provided responses related to each one below.

1. Of the data considered for inclusion in the EFH model, were final decisions on inclusion/exclusion of particular data appropriate, justified, and well-documented?

Yes with caveats. The main objective of this document is to provide a Level 2 EFH model for the shallow waters. However, during the presentation of the Fishery Independent Data Sources for MHI Uku by PIFSC Staff it was clear that some abundance data could be obtained from the baited stereo-video survey data for the deep areas but probably there was not consistency in the recording of this data type because uku is not a priority species. More detailed information about the different data sources and why not a density model was created for the deep waters should be included in the document.

Suggested recommendation: Provide more information about the reason of not creating a density model with the baited stereo-video survey data *(short/immediate term and high priority*). This could be for lack of data, lack of calibration between sampling devices, lack of random stratified design? Knowing more about the reason for that would help to identify caveats or limitations that could be improved with future research.

2. Are the data properly applied and appropriate for this species and habitat?

Yes with caveats. The modelling approach used has been commonly used in the literature for fish species and is appropriated when spatial data from different years are available. And the Environmental Data Summary provides a great tool to obtain temporally corresponding environmental data when there is a lack of collection of *in situ* environmental data. However, *A. virescens* is often seen at depths of 0-180 m and one of the considered species' core habitats (Penguin Bank) is deeper than the depth range considered in this study. Additionally, most of the data records include single observations. Therefore, the shallow waters do not represent well the main habitat of this species.

Suggested recommendation: Exploring the use of baited stereo-video survey data to provide a similar density model in deeper waters. (*mid-term and mid priority*).

3. Are the models used reliable, properly applied, adequate, and appropriate for the species, habitat, and available data?

Yes with caveats. The model does not seem to predict well the 0 values and underpredict the high density values. This could be related to the chosen of the Tweedie distribution. Perhaps a hurdle model could work better. It is important to have good predictions of the absences too in order to separate the EFH from non-EFH.

Suggested recommendation: Evaluate other model approaches using a hurdle approach where first the probability of occurrence is modelled based on presence-absence data, and secondly, only the positive values are modelled following a Poisson or Gamma distribution. This approach would allow to obtain both EFH Levels (1 and 2) with the same modelling approach. *(mid-term and mid priority*).

4. Are decision points and input parameters reasonably chosen?

Yes with caveats. There is not comparison of models without using the spatio-temporal effect to see if by adding this and complexity to the model, the results are improved.

Suggested recommendation: Include the variance explained by the “fixed” effect and by the full model comparing models with and without the spatio and temporal effect. *(short/immediate term and high priority*)

5. Are primary sources of uncertainty documented and presented?

No. Not uncertainty documented for the results. There is not data used for validation. I understand that the data are limiting and it is difficult to do a partition to train the model and another to test it. However, although all data can be used to build the model, two metrics can be provided. One would be the explanatory power using all the data, and another could be the predictive power partitioning the data or through other bootstraping technique that would allow to create a surface of variability in the predictions.

Suggested recommendation: Provide both the explanatory and predictive power and a surface of variability in the predictions. *(short/immediate term and high priority*)

6. Are model assumptions reasonably satisfied?

Yes. The authors have carefully designed the modelling approach and input data to satisfy all the assumptions of the models. However, as mentioned in ToR 3, the residual plots show some discrepancy between the observed and predicted values that could be improved using a hurdle model.

7. Are the final results scientifically sound, particularly delineating EFH boundaries?

Yes with caveats. The model is restricted to shallow waters and based on density. However, the species is more common in deep waters and most of the observations are single individuals. Applying management decisions to an incomplete distribution model can lead to misleading decisions. However, this approach is scientifically sound and can be used to other species in the future.

Suggested recommendation: Exploring the use of baited stereo-video survey data to provide a similar density model in deeper waters. (*mid-term and mid priority*).

8. Can the results be used to address management goals stated in the relevant FEP or other documents provided to the review panel? If any results of these models should not be applied for management purposes with or without minor short-term further analyses (in other words, if any responses to any parts of questions 1-7 are “no”), indicate:

Which results should not be applied and describe why, and

Based on the previous ToRs, the model can be used to address management goals for EFH Level 2 only for the shallow waters with suggested recommendations.

Which alternative set of existing EFH definitions should be used to inform uku EFH instead and describe why.

EFH level 1 can be used for the full distribution range of uku.

Other minor comments;

- Add Tanaka and Oliver (2021) and Cressie and Wikle (2015) to the References section.

- Figure 3. The relationship with deep looks more exponential than linear.
- Add the model equation with the parameters and error associated to the results.

Research priorities:

- Improve the understanding of habitat use of early life stages (<9cm) (*mid-term and mid priority*).
- Improve the abundance data obtained from baited stereo-video surveys. (*mid-term and mid priority*).
- Calibrate differences in abundance data between baited stereo-video surveys. (*mid-term and mid priority*).
- Investigate the relationships between probability of occurrence and density estimates. (*mid-term and mid priority*).
- Consider other environmental variables (e.g. topographic variables, current speed, fishing effort, mixed layer depth, etc.) that can help to explain the observed distribution patterns of Uku in the Main Hawaiian Islands. (*mid-term and mid priority*).
- Improve the sampling in soft bottom habitat. (*mid-term and mid priority*).
- Validation of the models in predicted EFH where no data are available through future sampling surveys (*mid-term and mid priority*).