SCIENTIFIC COMMITTEE
FIFTEENTH REGULAR SESSION
Pohnpei, Federated States of Micronesia
12–20 August 2019

Simulation analysis of pole and line CPUE standardization approaches for skipjack tuna in the WCPO

WCPFC-SC15-2019/SA-WP-04

N. Ducharme-Barth[1], M. Vincent[1], G. Pilling[1], J. Hampton[1]

[1]Oceanic Fisheries Programme, The Pacific Community

Contents

1 Executive summary

A simulation of the Japanese pole-and-line fishery for skipjack tuna was used as a case study to evaluate the effectiveness of CPUE standardization model performance in the case where spatial sampling coverage decreases over time. Key findings include:

- Geostatistical delta-GLMMs improve upon the performance of conventional delta-GLMs in simulations where shifts in spatiotemporal sampling occurred.
- Geostatistical delta-GLMMs have the flexibility to correctly estimate divergent regional trends, if present.
- The ability to estimate changes in catchability over time using the geostatistical delta-GLMM was influenced by the spatial distribution of the data. Shifts in spatial sampling were confounded with changes in catchability.
- Conventional delta-GLMs with additive spatial and temporal effects perform just as well as geostatistical delta-GLMMs models provided spatiotemporal shifts in sampling are not too extreme.
- Interpolating into unsampled areas is only valid if the assumption that biomass still exists there and that those areas are unsampled due to external barriers is met. Otherwise the model can be modified to not predict into unsampled areas.

We invite SC15 to:

- Discuss the findings in relation to the 2019 skipjack tuna stock assessment.
- Support further work done in this area, including simulation analysis of longline fisheries in preparation for the 2020 bigeye tuna and 2020 yellowfin tuna stock assessments.

2 Introduction

Skipjack tuna (*Katsuwonus pelamis*) represent the largest fishery by catch volume in the Western and Central Pacific Fisheries Commission (WCPFC) convention area (Williams and FFA, 2019). Standardized, fisheries-dependent catch-per-unit-of-effort (CPUE) indices remain an important input for the stock assessment process (McKechnie et al., 2016). In the case of skipjack tuna in the WCPFC convention area, catches are dominated by the purse seine sector of the fishery. Standardized CPUE indices for the purse seine sector are challenging to create due to potential changes in catchability over time (Vidal et al., 2019a). In the few cases where indices exist for the purse seine, they are limited in geographical scope (Vidal et al., 2019b). Given the historical distribution of the Japanese Pole-and-Line (JPPL) fishery (Ogura and Shono, 1999a), the standardized CPUE index from this fishery has been used to provide relative abundance inputs for the assessment across a broad geographic area (Ogura and Shono, 1999b; Langley et al., 2010; Kiyofuji et al., 2011; Kiyofuji, 2016). However, in recent decades, the JPPL fishery has shifted westwards across the Pacific and no longer samples skipjack tuna across the entire model region (Kiyofuji and Okamoto, 2013).

Traditionally, skipjack CPUE indices were created by using conventional delta-generalized linear models (delta-GLMs) fit to data within each assessment model region. Delta-GLMs are the combination of two underlying GLMs: a logistic model predicting the probability of a positive catch occurring and a second model predicting the magnitude of the positive catch rate (Lo et al.,

1992; Stefansson, 1996). More recently, geostatistical delta-generalized linear mixed models (delta-GLMMs) have been used for WCPFC stock assessments of bigeye tuna, yellowfin tuna, and South Pacifc albacore tuna (McKechnie et al., 2017; Tremblay-Boyer et al., 2017b,a, 2018).

The primary difference between these two approaches is how space and time are modeled. In a conventional delta-GLM this is usually done by specifying separate additive effects for spatial cell and year-quarter. This allows for the estimation of differential abundance in each spatial cell, however there is no interaction with time so the abundance in each spatial cell is assumed to have the same relative trend. This may not be a realistic assumption given that fish are not distributed homogeneously across ocean basins and could likely have different relative trends in non-adjacent portions of their range. An improvement to the additive model can be made by assuming an interactive effect between the spatial and temporal effects. This allows for independent temporal trends to be estimated in each spatial cell. However, models with interactive effects require observations in each combination of spatiotemporal strata which limits their application unless the observations from missing spatiotemporal strata are imputed or interpolated (Campbell, 2015). In a geostatistical delta-GLMM, the delta-GLM with interactive effects of space and time is extended in a mixed-effects framework to formally define the spatial and/or temporal correlation of observations, commonly defined using a Gaussian random field governed by a Matern covariance function (Thorson et al., 2016a; Macdonald et al., 2018). Using the estimated correlation structure of the data, geostatistical delta-GLMMs can simultaneously interpolate abundance of unobserved strata.

From an assessment standpoint, use of geostatistical models for CPUE standardization are appealing for a few different reasons. The explicit spatiotemporal correlation structure of the model allows for prediction into areas that are no longer fished. Additionally, the spatially explicit nature of the models allows for the simultaneous estimation of abundance indices across model regions using a single model and eliminates the need for post-hoc scaling of regional indices. Despite being shown to perform well with fisheries-independent data, even in cases where the underlying species distribution has shifted (Thorson et al., 2015, 2016b; Thorson and Barnett, 2017), these methods have had limited testing in their application to fisheries dependent catch rate data (Thorson, 2019).

While fisheries independent data come from statistically designed surveys that ensure the random distribution of samples across the spatial domain and temporal strata, the same assumption of appropriate spatiotemporal coverage cannot be made for fisheries dependent data. Holes in the spatiotemporal coverage from fisheries dependent data can arise from sampling preferentially with respect to abundance, changes in spatial targeting due to economic or management factors, as well as restricted access to fishing grounds due to regulatory or competitive forces. These anomalies in spatiotemporal sampling could lead to a disconnect between the underlying species abundance trend and the trend estimated from catch rate data. Beyond the fisheries dependent simulation testing already conducted (Grüss et al., 2019; Zhou et al., 2019), there exists a need to test these geostatistical methods in the case where fisheries spatial sampling coverage changes over time.

Given the apparent gap in the scientific literature, the spatial decrease in JPPL fishing effort and the importance of the JPPL index to the WCPFC skipjack tuna stock assessment; we evaluate the effectiveness of both the conventional and geostatistical modeling approaches in a case where spatial sampling coverage decreases over time. To do so, we developed a simulation model replicating the dynamics of the JPPL skipjack tuna fishery in the western and central Pacific Ocean (WCPO). This simulation study will allow us to identify issues in the estimated abundance indices that arise from 1) differences in the spatiotemporal sampling pattern and 2) differences in how space and/or time are handled in the model.

3 Methods

In order to explore these two different objectives, a simulation model was created using R v3.5.1 (R Core Team, 2018). This simulation model was broken up into 3 separate compartments: an operating model responsible for creating the abundance field of skipjack tuna in the WCPO, an observation model responsible for generating spatio-temporally referenced fisheries logbook observations of the simulated skipjack tuna abundance, and an estimation model component to create indices of abundance from the simulated logbook observations within the WCPFC assessment model regions.

Briefly, SEAPODYM output of adult skipjack biomass at a quarterly time step over a 32 year period (1979 - 2010) was used as the underlying operating model (Senina et al., 2016). Adult skipjack biomass was provided at a $1^\circ \times 1^\circ$ spatial resolution. All $1^\circ \times 1^\circ$ cells within the WCPO formed the spatial domain of the simulation (Figure 1). To set-up the observation models, a fleet of unique vessels were simulated to match the characteristics of the JPPL fleet. Next, for each vessel; the years active in the fishery, number of trips taken per year, and the gear configuration were simulated based on the JPPL fleet characteristics. Logbook records were simulated at the trip level according to agent based movement of individual vessels adapted from Ducharme-Barth et al. (2018). Having subdivided the model region into $1^\circ \times 1^\circ$ spatial cells, vessel movements among adjacent valid fishing cells within a trip were defined according to a simple "rook's case" random walk (diagonal movements were not permitted). Valid fishing cells were identified as those cells within the WCPO that were not on land.

Two different scenarios were used to determine the starting point and movement conditions for the random walk, resulting in two spatial sampling observation models (Figure 2): fisheries independent (*random*) and fisheries dependent (*JPPL*). In the random effort distribution scenario each cell had an equal probability across space and time of serving as the starting point for a non-directed random walk representing a vessel's fishing trip. The JPPL effort distribution scenario mimicked the preferential spatial sampling process of the JPPL fishing fleet. The starting cell was allocated preferentially with respect to skipjack biomass, and the probability of selecting a starting location closer to Japan increased over time. Within a trip, the probability of jumping to adjacent cells was also proportional to the underlying skipjack biomass. Additionally, in the case of the JPPL effort distribution, the number of vessels active in the fishery declined over time to match the pattern seen in reality. Once fishing locations were allocated to each vessel, catch was determined to be proportional to the biomass at each location.

Annual indices of relative abundance were estimated from the simulated fisheries logbook data using two estimation model structures: 1) a conventional delta-GLM and 2) a geostatistical delta-GLMM. More specifically, the delta-GLM was formulated with either additive or interactive spatiotemporal effects. The delta-GLMM structure was implemented using the R package VAST (Thorson et al., 2015; Thorson and Barnett, 2017; Thorson, 2019). Sensitivity to how this model was specified in terms of the configuration of the spatial random field, spatiotemporal correlation structure, and inclusion of an environmental covariate was also investigated. Indices were created for each of the 5 regions used in the 2016 skipjack tuna assessment and in each of the 8 regions proposed for use in the 2019 skipjack tuna assessment. In the case of the delta-GLM this meant fitting a separate model to the simulated data coming from each model region. Additionally, a nominal index was calculated in each region by taking the average of logbook catches (catch per sample) in each region within each year. Since the trend of the index and not the magnitude of the index was the quantity of interest, all indices were rescaled to a mean of 0 and a standard deviation

of 1 prior to comparing model performance. Model performance was assessed by calculating the root-mean-squared-deviance (RMSD) between the rescaled *true*, simulated abundance trend and the rescaled estimated index for each model region. Additional details of the simulation framework just described can be found in the subsequent Technical Annex.

This framework was developed in a nested structure in terms of the scope and uncertainty included in the simulation. To begin with, a *proof-of-concept* model was developed to test the effects of observation and estimation model type on the accuracy of the estimated index before moving on to a more detailed, and complex *case-study* model simulation of the JPPL skipjack tuna fishery. The case-study model simulated 32 years worth of data at a quarterly time-step, resulting in ∼750,000 observations. Movement in the case-study was defined according to the random walk patterns described above. In contrast, the proof-of-concept model simulated 15,000 observations over a ten year period and at an annual time scale. Additionally, the movement driving the observation patterns was simplified to the aggregate fleet level for the proof-of-concept simulation.

Within each of the proof-of-concept and case-study models, data uncertainty increased progressively from a *basic* case with zero error. Gradually uncertainty was introduced in the form of process error (*patchy* case) where 10% of the cells in the abundance field were replaced with zero skipjack abundance. The probability of each cell being replaced with zero abundance was inversely proportional to the skipjack abundance it initially contained. This had the effect of creating a patchier abundance field towards the periphery of the skipjack distribution. Building onto the patchy case, 10% lognormal observation error was introduced (*e10*) to vessel catches at each fishing location. Finally, catchability effects were added to vessel catches at each location on top of the lognormal observation error (*q10* case). The catchability effects were based on the vessel's gear configuration of class, number of poles fished, bird radar, bait tank, sonar, NOAA receiver, and a normally-distributed random vessel effect. Number of poles fished on a trip were determined by vessel class, as distant-water (PLDW) vessels are larger and typically fish more poles than offshore (PLOS) vessels. The remaining gear categories (bird radar, bait tank, sonar, and NOAA receiver) were allowed to be upgraded over time according to the empirical probabilities of having a specific gear configuration as seen in the JPPL fishery (Kiyofuji, 2013). This resulted in an increase in catchability over time for the two different vessel groups (Figure 3).

Using the above simulation framework, a full factorial experimental design of 50 simulated replicates from the proof-of-concept model was used to test for the effects of observation model type, estimation model type, and simulated uncertainty on the accuracy of the estimated abundance index. A single replicate of the case-study model was used to investigate the effects of including temporally changing catchability to the performance of the geostatistical model.

4 Results

Though indices were calculated for each model region used in the 2016 and 2019 WCPFC skipjack tuna assessment structures, we chose to focus our presentation of the results on a representative subset of the 8 region structure considered in the 2019 assessment: the entire WCPFC skipjack tuna assessment area, region 2, region 3, region 7 and region 8. As seen from (Figure 1) regions 7 and 8 are within the core of the distribution of skipjack biomass, while regions 2 and 3 are more to the periphery of the distribution. In terms of spatial sampling under the JPPL observation model (Figure 2), region 3 is the only region consistently sampled in all time periods. Since region 2 is on the periphery of skipjack abundance, sampling is patchy to begin with and coverage decreases spatially with time. Region 7 and half of region 8 are well sampled initially as they are in the core

of the skipjack distribution. However, by the end of the simulation region 8 is not sampled at all and only a portion of region 7 is still sampled.

Beginning with the simplest case of the proof-of-concept model, random effort distribution and no uncertainty (*basic_rand*), all estimation models are able to correctly estimate the true simulated trend in abundance across all 50 simulated replicates (Figure 4). Though an expected result, it serves as a baseline for comparison with other uncertainty and effort distribution scenarios. It is also worth noting that the inherent intraannual variation at the northern limit of skipjack tuna range (region 2) manifests itself as higher levels of variability across the 50 replicates for each index type. This variability would decrease with an increase in sampling intensity.

Moving to a more complex and realistic case with process error and JPPL effort distribution (*basic_jppl_patchy*), differences in model performance begin to occur (Figure 5). The most extreme example of this is in region 8 where the nominal and both delta-GLM indices exhibit large departures from the true trend as spatial sampling coverage decreases ultimately to zero by the end of the model period. The geostatistical indices, given their ability to interpolate abundance into unsampled areas, perform much better though it is interesting that none of the different formulations appear to make much of a difference in this scenario.

Summarizing the results from the proof-of-concept model simulation in terms of RMSD across all combinations of observation models, estimation models, and levels of data uncertainty (Figure 6); the ability to estimate the true trend is dependent on the observation model that generated the data. Holding the observation model constant, and looking across the scale of uncertainty (basic, patchy, and e10_patchy) model performance does not change very much. Conversely, if data uncertainty is held constant, model performance greatly changes depending on the observation model type that generated the data. As mentioned above, in model regions where a large shift in spatial sampling occured, the geostatistical models outperformed the conventional delta-GLM indices. Additionally, in this simulation, the formulation of the geostatistical model (knot structure, environmental covariate, or temporal correlation) did not appear to alter the estimated index.

The pattern in relative performance based on the observation models and levels of data uncertainty held true for the case-study. The inclusion of catchability effects into the data uncertainty (on top of the process and observation error) provided an interesting wrinkle to the case-study model. Looking specifically at region 8 (Figure 7), which exhibited the greatest shift in spatial sampling, the ability to estimate and correct for an increasing catchability trend over time appeared to be confounded with the perceived change in abundance caused by the shift in spatial sampling. This can be contrasted with the case where spatial cells were consistently sampled across time (*rand.q10*). In that scenario, the geostatistical model was able to properly estimate the catchability covariates and correct for the bias seen in the nominal index. However, the previous finding that consistent, random sampling could result in correct estimates of catchability covariates did not hold up across all regions. In region 2 (Figure 8), both types of observation model, random and JPPL, failed to estimate and account for the change in catchability. In fact, attempting to estimate the catchability covariates resulted in a worse index than the nominal. This is likely due to temporal confounding between the increasing trends in both catchability and abundance. There did not appear to be enough contrast in the data to resolve the relative difference in these two trends. When the catchability covariates were standardized out it resulted in an erroneous, decreasing trend.

5 Discussion

Spatial sampling pattern played a large role in the ability to correctly estimate the underlying trend. Both estimation model types were able to correctly estimate the trend under a random effort observation model. However, under the JPPL effort observation model, in regions where sampling coverage decreased, substantial differences emerged between the two estimation model types. In these cases geostatistical indices were able to achieve a much closer fit than delta-GLM indices in the proof of concept simulation. This simulation reinforces the importance of knowing what sampling framework was used when generating observations for standardizing fisheries catch rate information since the estimation models did not perform equally well for both sampling types. Ideally, a fisheries independent monitoring program would be used to collect samples. Then the choice of estimation model is less important. However, given the large spatial domain that tuna and tuna-like species occupy and the prohibitive cost associated with appropriately sampling across spatiotemporal strata, such a survey program is unrealistic. This limits the analysis to fisheries dependent catch rate data and thus care must be taken to select an appropriate estimation model.

The spatial shift observation model used in both the proof-of-concept and the case-study simulations where region 8 is completely unfished by JPPL vessels by the end of the simulation is an exaggeration of the spatial shift seen in the actual JPPL fishery. However, it serves to illustrate the point that the decision on how a model implicitly or explicitly interpolates the unfished cells in an important one (Walters, 2003; Carruthers et al., 2011). A conventional delta-GLM fit to data without interpolation into unfished cells makes the implicit assumption that they have the same abundance as the mean of the sampled cells. This can be an appropriate assumption to make if fishing effort does not shift relative to the underlying abundance distribution. However, as shown in region 8 of the simulation, if a persistent shift in spatial sampling effort exists and the cells being sampled are no longer representative of the abundance for the entire model region, the estimated index can depart from the true underlying trend.

Relatedly, one of the assumptions made with the geostatistical model is that abundance is interpolated into unfished areas based on the spatial correlation structure of the observations. Interpolating into unsampled areas is only valid if the assumption that biomass still exists there and that those areas are unsampled due to external barriers is met. In the case of skipjack, continued catches by the purse seine fishery in areas no longer fished by the JPPL fishery support the assumption of model interpolation into those areas. For a tropical species such as skipjack, it is important to ensure that biomass is not extrapolated outside of the biological and/or physiological range of the species. In these cases the geostatistical model should be modified to not interpolate biomass into these areas. For instance, the albacore assessment in 2018 used a modified geostatistical model to only predict biomass into cells that met a biologically realistic temperature threshold (Tremblay-Boyer et al., 2018).

Previous comparisons of conventional delta-GLMs and geostatistical models (Tremblay-Boyer et al., 2018) noted that the regional indices produced by a single geostatistical model, modeling the data together, were much more similar to each other than the regional indices estimated from conventional delta-GLMs, modeling each region independently. It was hypothesized that geostatistical models may not have sufficient flexibility to estimate divergent trends in model regions, and that modeling all of the data together would artificially minimize differences in abundance trends between regions. The simulation work presented here suggests that the geostatistical model does have sufficient flexibility to estimate divergent trends using a single model, if those divergent trends exist.

Another interesting result was the effect of allowing catchability to increase over time in the case-

study. When the effort distribution was random and catchability was accounted for, the geostatistical model was usually able to correctly estimate the index. However, when the effort distribution was generated by the JPPL observation model, confounding between the shift in spatial sampling and the change in catchability prevented the model from correctly estimating the underlying index. Furthermore, there was a notable exception to the above results. Even with a random effort distribution, the geostatistical model was unable, due to temporal confounding, to disentangle changes in catchability from changes in abundance if both changed in the same direction. In this case the estimated index was substantially worse than the nominal index.

This simulation is the first step in a larger analysis looking into the effects of different spatial sampling patterns and interpolation assumptions on the effects of estimated indices. One of the initial hypotheses prior to the study was that the different configurations of the geostatistical model (knot configuration, temporal covariance structure, and inclusion of an environmental covariate) would have an effect on the interpolated biomasses and thus on the estimated indices. This did not turn out to be the case, so future work will explore additional scenarios to try to identify where different configurations make more of a difference. Additionally, the case-study model only considered a single simulation replicate and did not compare the performance of the geostatistical model to the conventional delta-GLM. Future work will fully compare the two estimation model types across a larger number of replicates of the case-study similar to the proof-of-concept simulation.

In conclusion, the proof-of-concept simulation showed that geostatistical models appear to be an improvement to the conventional delta-GLM in this particular case where effort has shifted spatially over time. However, it is important to consider the factors that may have led to the effort shift and evaluate if the interpolation assumptions made in the geostatistical model remain appropriate. It was also shown that despite modeling all data and regions together, geostatistical models maintain the flexibility to correctly estimate divergent regional indices if they exist. This research can serve as a building block for additional analyses exploring the effects of spatial sampling patterns on estimated abundance indices, and can help guide decisions made in estimating abundance indices for use in WCPFC stock assessments.

We invite SC15 to discuss the findings in relation to the 2019 skipjack tuna stock assessment. We also invite the SC to support further work done in this area, including simulation analysis of longline fisheries in preparation for the 2020 bigeye tuna and 2020 yellowfin tuna stock assessments.

6 Acknowledgments

We are grateful for our colleagues H. Kiyofuji and Y. Aoki at the National Research Institute of Far Seas Fisheries, Fisheries Research Agency Shimizu Japan for stimulating discussions that led to this simulation project, and facilitating further collaboration on the geostatistical analysis of Japanese pole-and-line CPUE data. Thank you to S. Hare for his thoughtful comments on an earlier draft of this report.

References

Campbell, R. A. (2015). Constructing stock abundance indices from catch and effort data: Some nuts and bolts. *Fisheries Research*, 161:109–130.

Carruthers, T. R., Ahrens, R. N. M., McAllister, M. K., and Walters, C. J. (2011). Integrating

imputation and standardization of catch rate data in the calculation of relative abundance indices. *Fisheries Research*, 109(1):157–167.

Ducharme-Barth, N. D., Shertzer, K. W., and Ahrens, R. N. (2018). Indices of abundance in the Gulf of Mexico reef fi sh complex: A comparative approach using spatial data from vessel monitoring systems. *Fisheries Research*, 198:13.

Enea, M. (2017). speedglm: Fitting Linear and Generalized Linear Models to Large Data Sets.

Grüss, A., Walter, J. F., Babcock, E. A., Forrestal, F. C., Thorson, J. T., Lauretta, M. V., and Schirripa, M. J. (2019). Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. *Fisheries Research*, 213:75–93.

Kiyofuji, H. (2013). Reconsideration of CPUE for albacore caught by the Japanese pole and line fishery in the northwestern North Pacific Ocean. Technical Report ISC/13/ALBWG-01/11.

Kiyofuji, H. (2016). Skipjack catch per unit effort (CPUE) in the WCPO from the Japanese pole-and-line fisheries. Technical Report WCPFC-SC12-2016/SA-WP-05, Bali, Indonesia, 3–11 August 2016.

Kiyofuji, H. and Ochi, D. (2016). Proposal of alternative spatial structure for skipjack stock assessment in the WCPO. Technical Report WCPFC-SC12-2016/SA-IP-09, Bali, Indonesia, 3–11 August 2016.

Kiyofuji, H. and Okamoto, H. (2013). Decadal and spatial analysis of Japanese pole and line fisheries for improving catch per unit effort of skipjack tuna in the Western and Central Pacific Ocean (WCPO). Technical Report WCPFC-SC9-2013/ SA-WP-14, Pohnpei, Federated States of Micronesia.

Kiyofuji, H., Uosaki, K., and Hoyle, S. (2011). Up-to-date CPUE for skipjack caught by Japanese distant and offshore pole and line in the western central Pacific Ocean. Technical Report WCPFC-SC7-2011/SA-IP-13, Pohnpei, Federated States of Micronesia.

Langley, A., Uosaki, K., Hoyle, S., Shono, H., and Ogura, M. (2010). A standardized CPUE analysis of the Japanese distant-water skipjack pole-and-line fishery in the western and central Pacific Ocean (WCPO), 1972-2009. Technical Report WCPFC-SC6-2010/SA-WP-08.

Lo, N. C., Jacobson, L. D., and Squire, J. L. (1992). Indices of Relative Abundance from Fish Spotter Data based on Delta-Lognormal Models. *Canadian Journal of Fisheries and Aquatic Sciences*, 49:2515–2526.

Macdonald, J. I., Logemann, K., Krainski, E. T., Sigurosson, P., Beale, C. M., Huse, G., Hjollo, S. S., and Marteinsdottir, G. (2018). Can collective memories shape fish distributions? A test, linking space-time occurrence models and population demographics. *Ecography*, 41(6):938–957.

McKechnie, S., Hampton, J., Pilling, G. M., and Davies, N. (2016). Stock assessment of skipjack tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC12-2016/SA-WP-04, Bali, Indonesia, 3–11 August 2016.

McKechnie, S., Pilling, G., and Hampton, J. (2017). Stock assessment of bigeye tuna in the western and central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-05, Rarotonga, Cook Islands, 9–17 August 2017.

Ogura, M. and Shono, H. (1999a). Factors affecting the fishing effort of the Japanese distant water pole and line vessel and the standardization of that skipjack CPUE - Part A Description of the fishery and data. Technical Report SCTB12.

Ogura, M. and Shono, H. (1999b). Factors affecting the fishing effort of the Japanese distant water pole and line vessel and the standardization of that skipjack CPUE. Part B - Calculation of CPUE standardization. Technical Report SCTB12.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Senina, I., Lehodey, P., Calmettes, B., Nichol, S. J., Caillot, S., J, H., and P, W. (2016). Predicting skipjack tuna dynamics and effects of climate change using SEAPODYM with fishing and tagging data. Technical Report WCPFC-SC12-2016/EB WP01.

Smith, T. and Reynolds, R. (1981). NOAA Smith and Reynolds Extended Reconstructed Sea Surface Temperature (ERSST) Level 4 Monthly Version 5 Dataset in netCDF.

Stefansson, G. (1996). Analysis of groundfish survey abundance data: Combining the GLM and delta approaches. *Ices Journal of Marine Science*, 53(3):577–588.

Thorson, J. T. (2019). Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, 210:143–161.

Thorson, J. T. and Barnett, L. A. K. (2017). Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science*, page 11.

Thorson, J. T., Fonner, R., Haltuch, M. A., Ono, K., and Winker, H. (2016a). Accounting for spatiotemporal variation and fisher targeting when estimating abundance from multispecies fishery data. *Canadian Journal of Fisheries and Aquatic Science*, page 14.

Thorson, J. T., Pinsky, M. L., and Ward, E. J. (2016b). Model-based inference for estimating shifts in species distribution, area occupied and centre of gravity. *Methods in Ecology and Evolution*, 7(8):990–1002.

Thorson, J. T., Shelton, A. O., Ward, E. J., and Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *Ices Journal of Marine Science*, 72(5):1297–1310.

Tremblay-Boyer, L., McKechnie, S., and Pilling, G. (2018). Background Analysis for the 2018 stock assessment of South Pacific albacore tuna. Technical Report WCPFC-SC14-2018/ SA-IP-07, Busan, South Korea, 8-16 August 2018.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017a). Geostatistical anal- yses of operational longline CPUE data. Technical Report WCPFC-SC13-2017/SA-WP-03, Rarotonga, Cook Islands, 9-17 August 2017.

Tremblay-Boyer, L., McKechnie, S., Pilling, G., and Hampton, J. (2017b). Stock assessment of yellowfin tuna in the Western and Central Pacific Ocean. Technical Report WCPFC-SC13-2017/SA-WP-06, Rarotonga, Cook Islands, 9-17 August 2017.

Vidal, T., Muller, B., and Pilling, G. (2019a). Tropical WCPO purse seine effort creep indicators. Technical Report WCPFC-SC15-2019/MI-IP-05, Pohnpei, Federated States of Micronesia.

Vidal, T., Pilling, G., Tremblay-Boyer, L., and Usu, T. (2019b). Standardized CPUE for skipjack tuna Katsuwonus pelamis from the Papua New Guinea archipelagic purse seine fishery. Technical Report WCPFC-SC15-2019/SA-IP-05, Pohnpei, Federated States of Micronesia.

Walters, C. (2003). Folly and fantasy in the analysis of spatial catch rate data. *Canadian Journal of Fisheries and Aquatic Sciences*, 60(12):1433–1436.

Williams, P. and FFA (2019). Overview of tuna fisheries in the WCPO, including economic conditions – 2018. Technical Report WCPFC-SC15-2019/GN-IP-01.

Zhou, S., Campbell, R. A., and Hoyle, S. D. (2019). Catch per unit effort standardization using spatio-temporal models for Australia's Eastern Tuna and Billfish Fishery. *ICES Journal of Marine Science*.

7 Figures

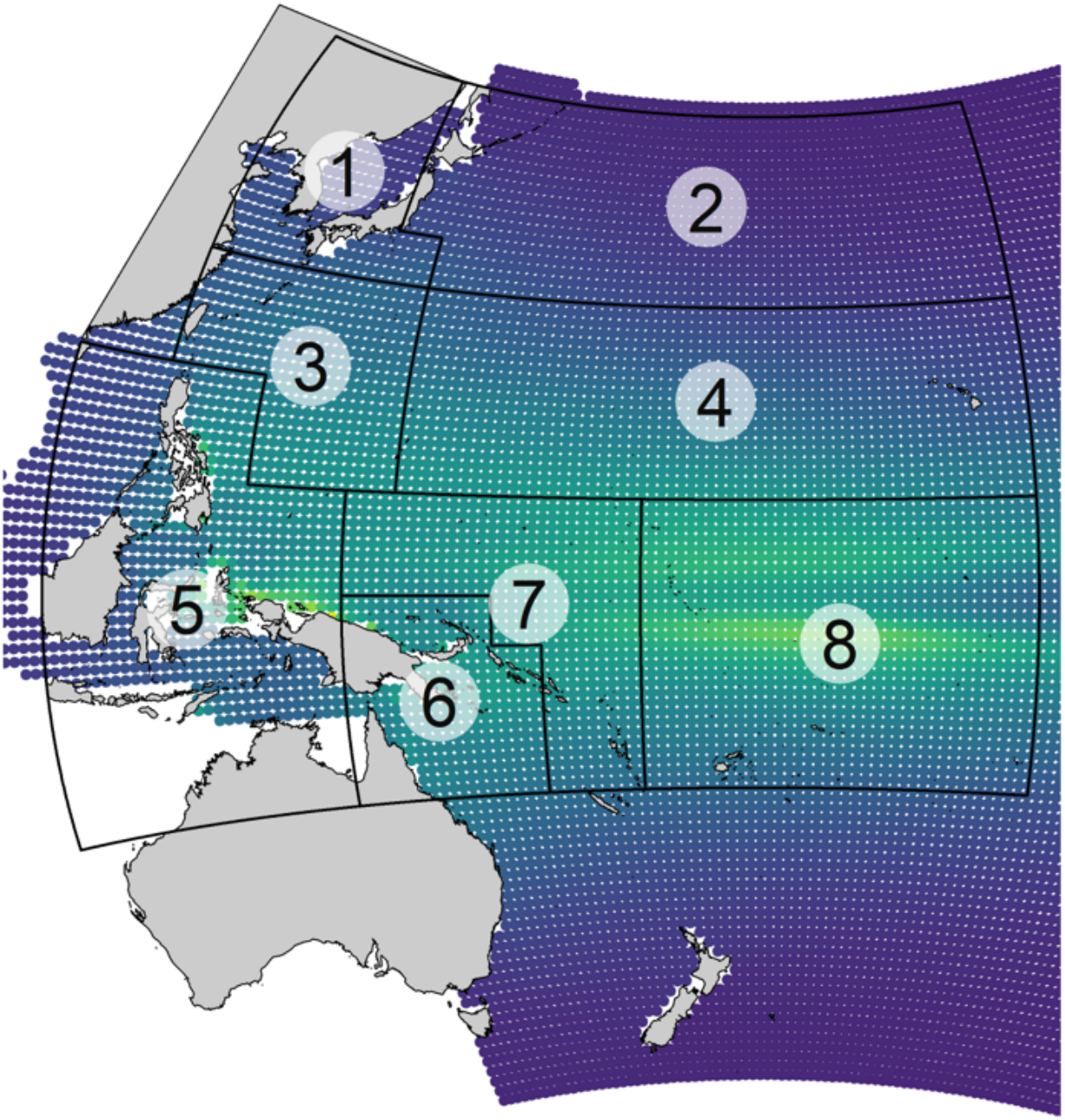

Figure 1: Average spatial distribution of adult skipjack tuna biomass from SEAPODYM used as the operating model in the simulation. Lighter colors indicate areas of higher abundance, and darker colors indicate areas of lower skipjack abundance. The numbers correspond to the 8 regions proposed for use in the 2019 skipjack tuna assessment.

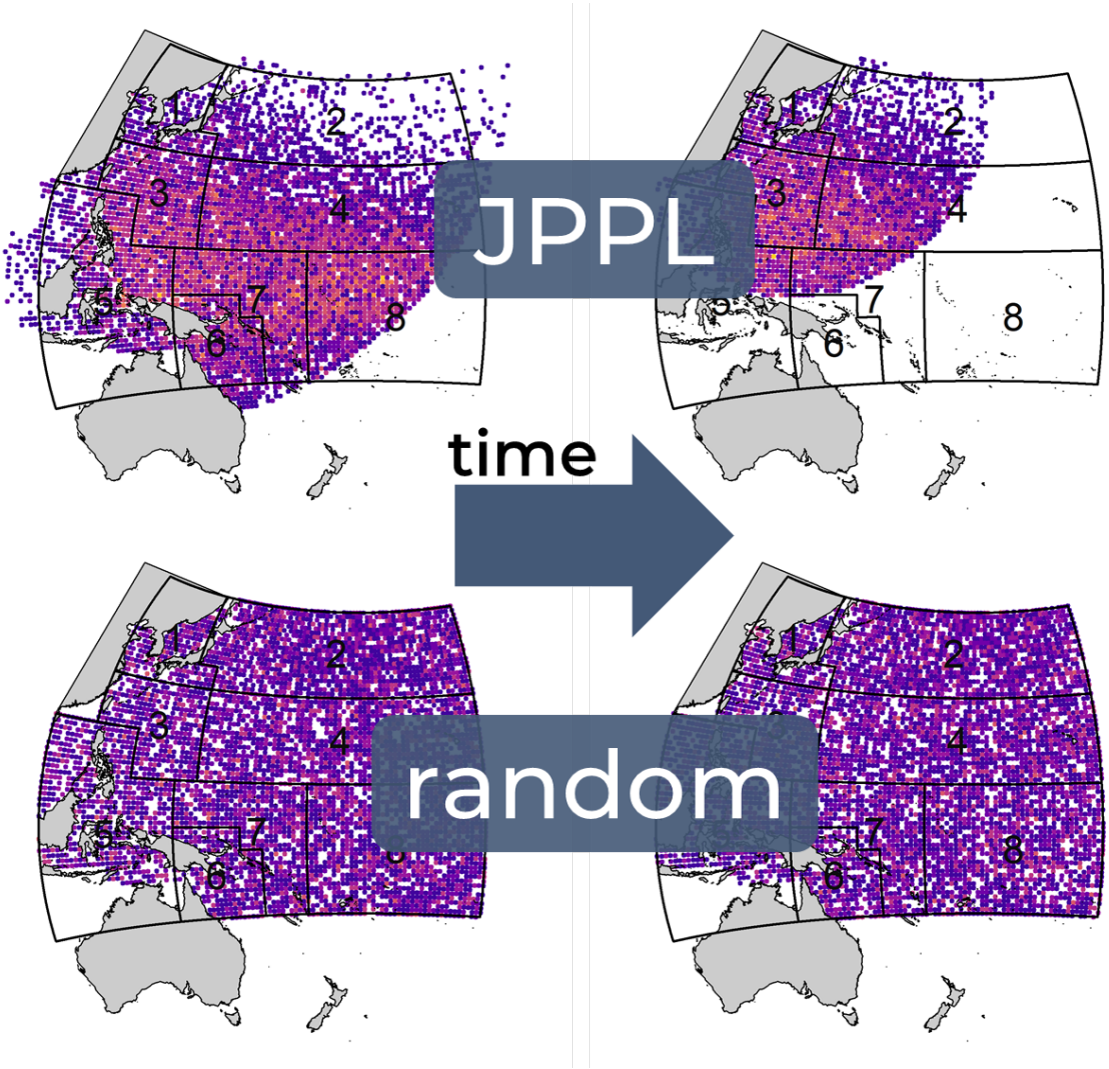

Figure 2: The effort distribution patterns for the two different observation models (*random* and *JPPL*) at the start and end of the simulation. Lighter colors indicate a greater density of effort, and darker colors indicate lower levels of effort. Areas in white are unsampled in that time period.

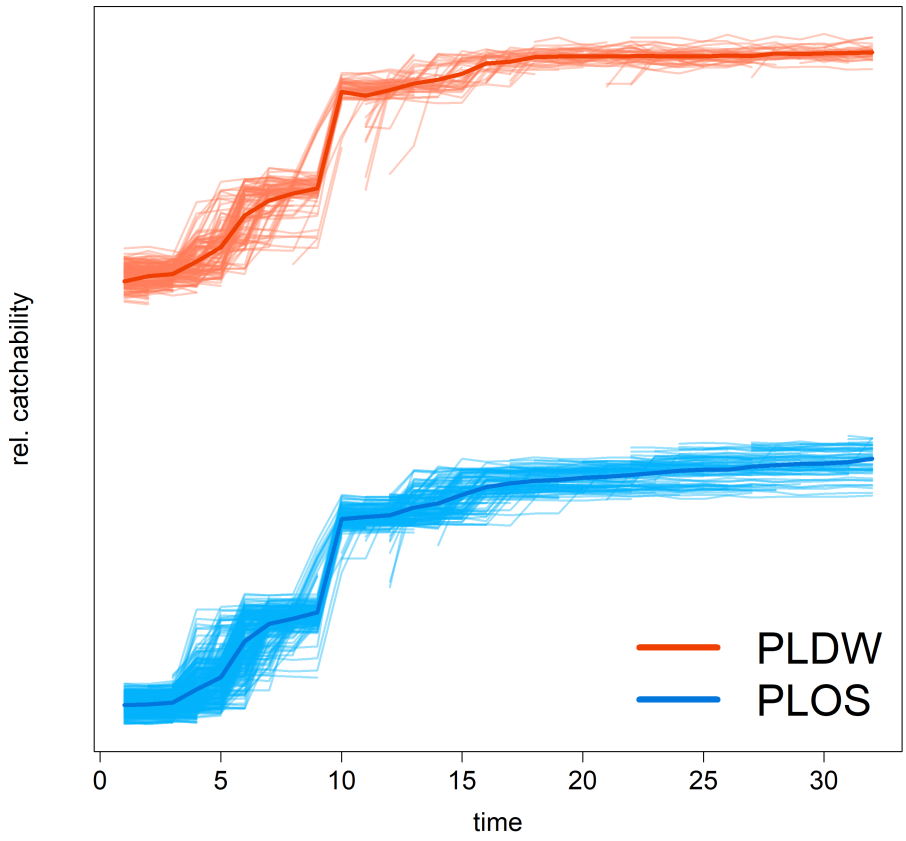

Figure 3: Cumulative increase of catchability over time for the simulated DW and OS fleets in the case-study. Each line represents an individual vessel. The thicker line denotes the median across all vessels.

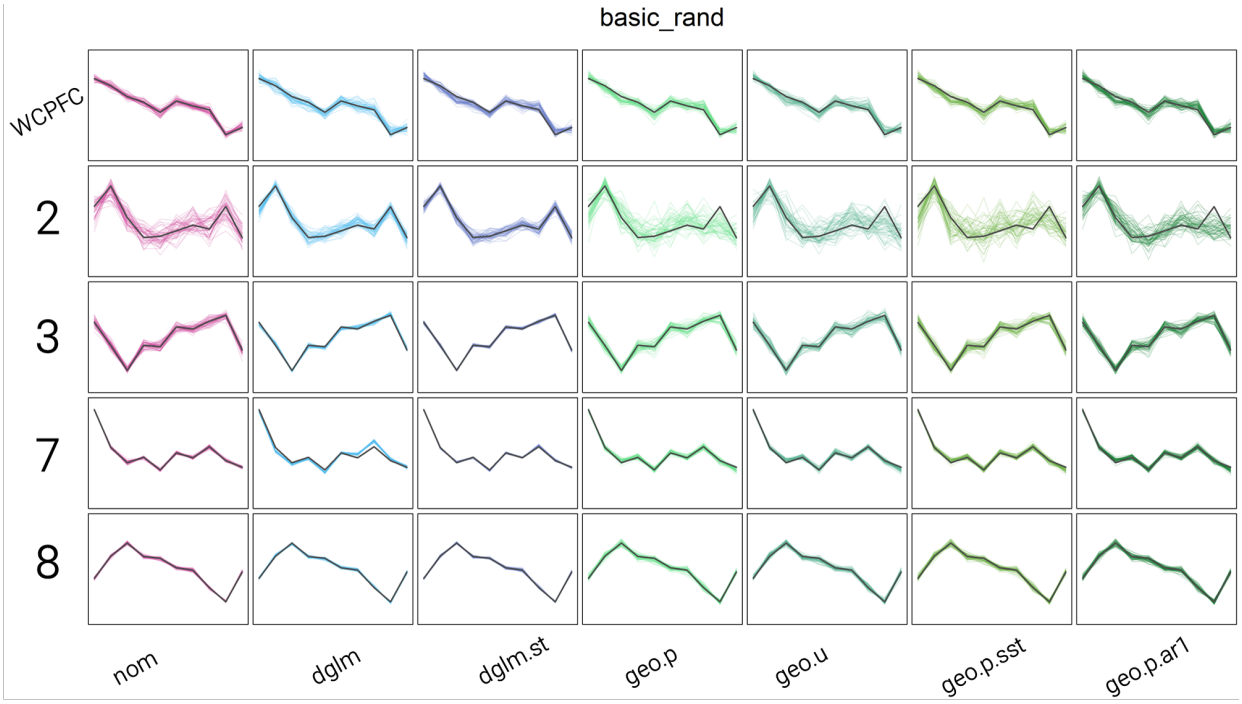

Figure 4: Summary of the results of the proof-of-concept model run without uncertainty and a random effort distribution. The true simulated trend is shown in black and the colored lines show the estimated trends across the 50 simulated replicates. Moving across the columns the results are shown for the nominal indices (*nom*), delta-GLM with additive spatiotemporal effects (*dglm*), delta-GLM with interactive spatiotemporal effects (*dglm.st*), geostatistical model with preferential knots (*geo.p*), geostatistical model with uniform knots (*geo.u*), geostatistical model with preferential knots and sea-surface temperature covariate (*geo.p.sst*), and geostatistical model with first order temporal autocorrelation (*geo.p.ar1*). The rows correspond to a subset of the 8 regions proposed for use in the 2019 skipjack tuna assessment. The *WCPFC* row is the index estimated across regions 1-8 combined.

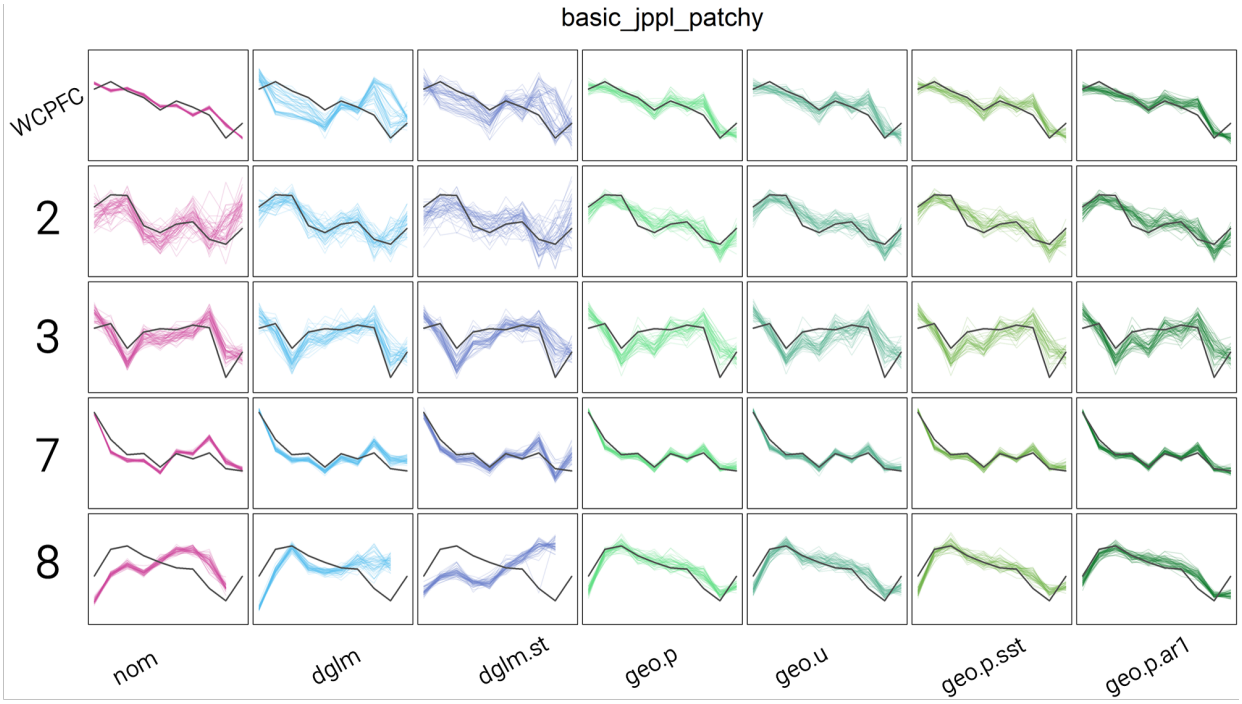

Figure 5: Summary of the results of the proof-of-concept model run with process error and JPPL effort distribution. The true simulated trend is shown in black and the colored lines show the estimated trends across the 50 simulated replicates. Moving across the columns the results are shown for the nominal indices (*nom*), delta-GLM with additive spatiotemporal effects (*dglm*), delta-GLM with interactive spatiotemporal effects (*dglm.st*), geostatistical model with preferential knots (*geo.p*), geostatistical model with uniform knots (*geo.u*), geostatistical model with preferential knots and sea-surface temperature covariate (*geo.p.sst*), and geostatistical model with first order temporal autocorrelation (*geo.p.ar1*). The rows correspond to a subset of the 8 regions proposed for use in the 2019 skipjack tuna assessment. The *WCPFC* row is the index estimated across regions 1-8 combined.

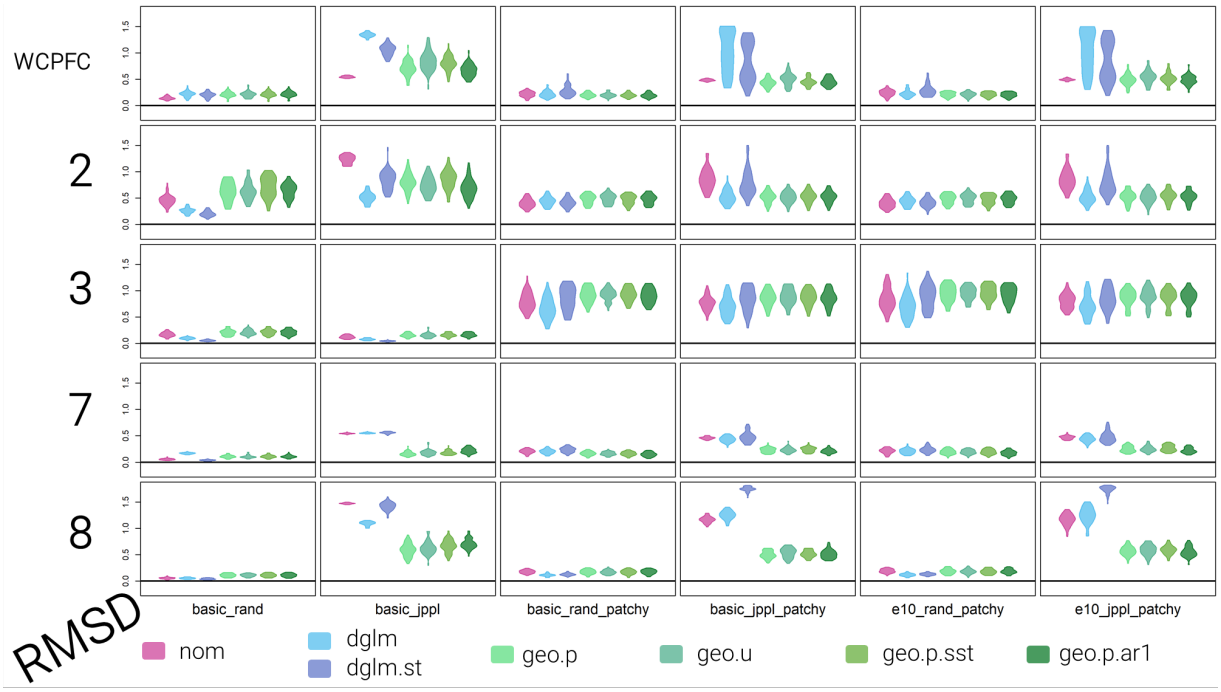

Figure 6: Summary of the root-mean-square-deviance (RMSD) across all proof-of-concept simulation runs. Each violin represents the distribution of RMSD across all 50 simulations. The closer the violin is to the x-axis in each panel, the greater accuracy of the estimation model used. The more vertically compressed the violin is, the greater precision in the estimate. Within each panel the results are shown for the nominal indices (*nom*), delta-GLM with additive spatiotemporal effects (*dglm*), delta-GLM with interactive spatiotemporal effects (*dglm.st*), geostatistical model with preferential knots (*geo.p*), geostatistical model with uniform knots (*geo.u*), geostatistical model with preferential knots and sea-surface temperature covariate (*geo.p.sst*), and geostatistical model with first order temporal autocorrelation (*geo.p.ar1*). The rows correspond to a subset of the 8 regions proposed for use in the 2019 skipjack tuna assessment. The *WCPFC* row is the index estimated across regions 1-8 combined.

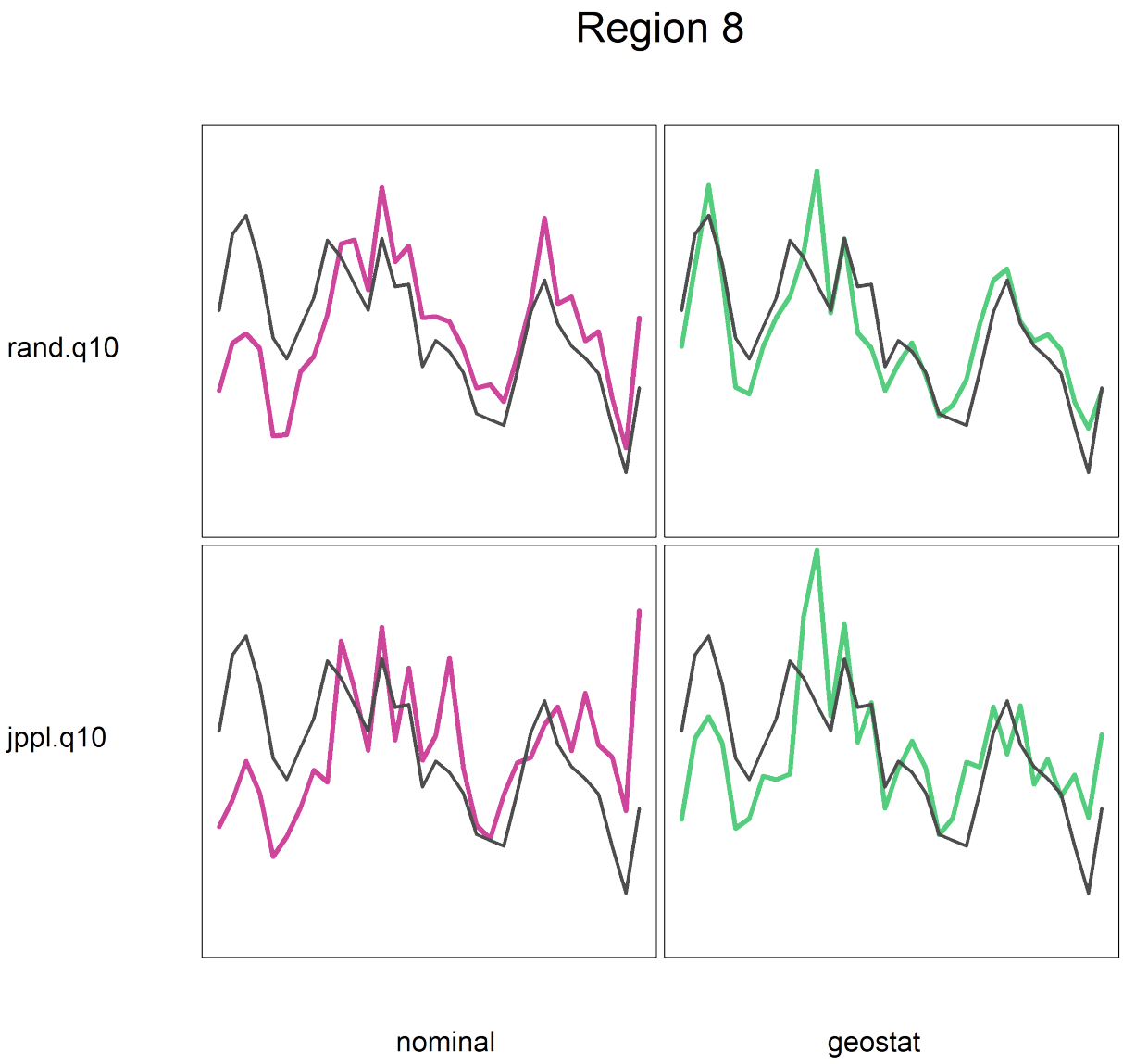

Figure 7: Effect of incorporating catchability effects into the observation models on the estimated indices. The left two panels show the nominal index under the *rand.q10* and *jppl.q10* estimation models, and the right two panels show the estimated index from a geostatistical model including terms for the catchability effects.

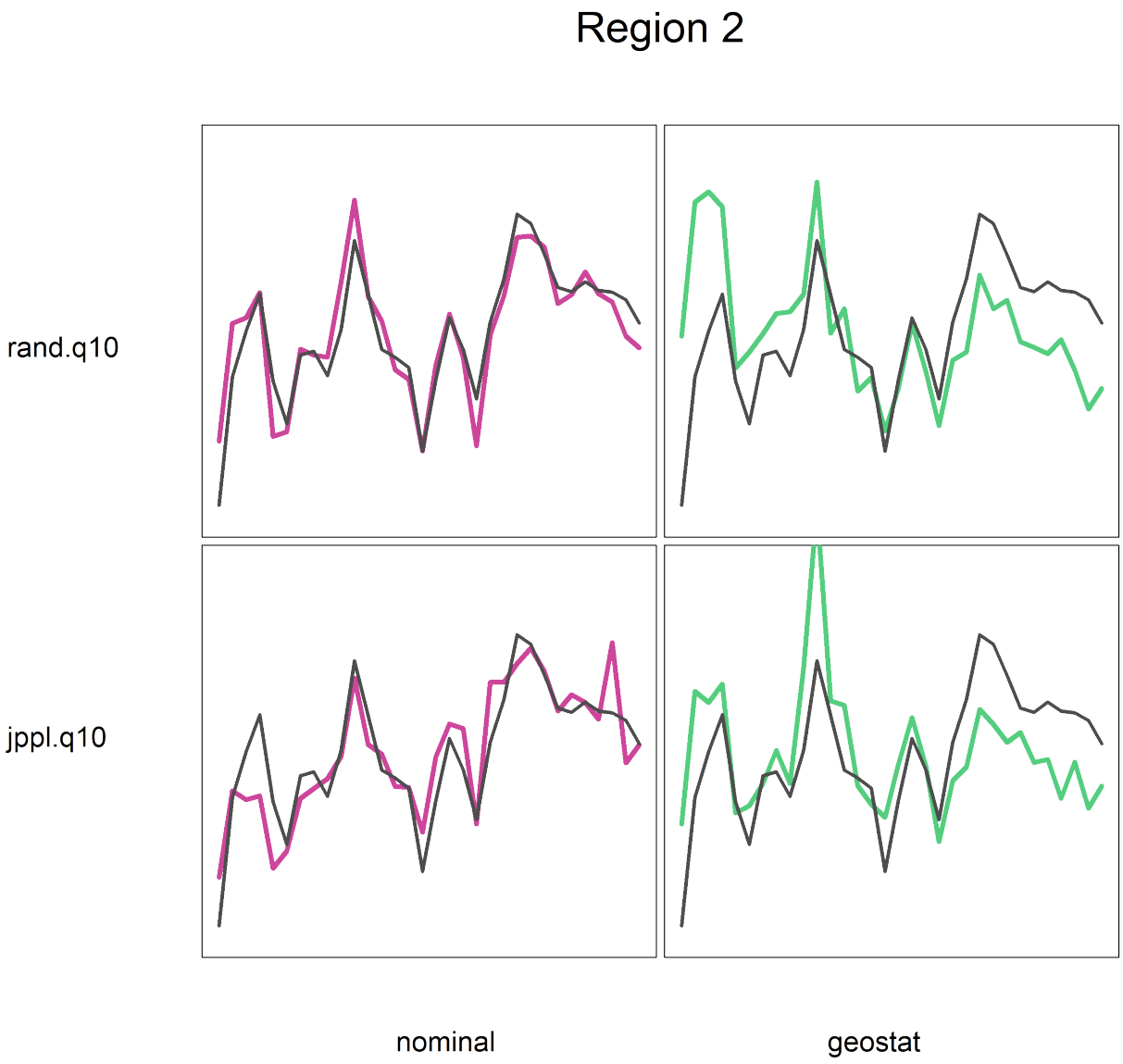

Figure 8: Effect of incorporating catchability effects into the observation models on the estimated indices. The left two panels show the nominal index under the *rand.q10* and *jppl.q10* estimation models, and the right two panels show the estimated index from a geostatistical model including terms for the catchability effects.

8 Technical Annex

The structures of the proof-of-concept and case-study simulations are very similar, although the case-study has some added complexity in order to simulate observations at the individual vessel level to more closely replicate the dynamics of the JPPL fishery.

8.1 Proof-of-concept simulation

8.1.1 Effort distributions

The samples generated in each year of the simulation were defined as a random draw from a $Multinomial(n_s; p_1, ..., p_i, ..., p_T)$ where n_s is the total number of samples across all years and p_1, p_i, and p_T are the event probabilities for the first year, i^{th}, and terminal years of the simulation. Under a random effort distribution all p_i were equal but under the JPPL effort distribution the p_i followed a temporal trend matching the pattern of unique vessels active in the fishery described in Kiyofuji and Okamoto (2013).

Next, for each sample in each year, cells were selected to be sampled from a pool of valid fishing cells. At a minimum, fishing cells were defined as those cells not on land and that were within the WCPO. Under the random effort distribution, all valid fishing cells had an equal probability of being selected across space and time. Under the JPPL effort distribution scenario, the pool of valid fishing cells diminished each year as cells further from Japan were sequentially removed. Within the pool of valid fishing cells for a given year, the probability of selecting a given cell i was defined as $p_i = \frac{a_i}{\sum_{i=1}^{N} a_i}$ where a_i is the abundance in cell i. Using these probabilities, the fishing cells were drawn from a $Multinomial$ distribution.

8.1.2 Data uncertainty

In the *basic* uncertainty case, there is no uncertainty so the simulated logbook catch record for a fished cell i was defined as $c_i = A_o \times a_i$ where A_o is a constant applied to prevent numerical overflow.

In the *patchy* uncertainty case, the simulated logbook catch was defined as above. As mentioned in the Methods section above, process error was introduced by replacing 10% of a_i with 0 where the probability of replacing a cell i's abundance was defined as $p_i \propto \frac{\sum_{i=1}^{N} a_i}{a_i}$. A $Multinomial$ distribution was then used to select the cells for replacement using these probabilities. In this way, zero catches were introduced into the simulated logbook.

In the $e10$ uncertainty case, 10% lognormal observation error was introduced on top of the existing *patchy* uncertainty. The simulated logbook catch record for a fished cell was defined as

$$ln(c_i) \sim Normal(\mu = A_o a_i, \sigma = 0.1 A_o a_i)$$

.

8.2 Case-study simulation

8.2.1 Effort distributions

In order to simulate logbook records at the trip level, individual vessels and their trips had to be generated. For the JPPL effort distribution case, the number of unique vessels (per vessel class:

PLDW or PLOS) fishing in a given year was determined using the numbers reported in Kiyofuji and Okamoto (2013). In the random effort distribution case, the number of unique vessels fishing in each year for each class was not assumed to decline but was held fixed at the mean value from Kiyofuji and Okamoto (2013).

Based on the numbers of unique vessels active in a given year per class, the entry and exit dates into and out of the fishery were simulated in an iterative process beginning in the first year. In year 1, all vessels active in that year were assumed to have entered in that year. Each of these vessels were then assigned a longevity based on a random draw from a *Negative Binomial* distribution with $\mu = 11.78$ and $size = 3.31$. These parameters were estimated from the time distribution of unique vessels in Kiyofuji and Ochi (2016). In each subsequent year, vessels were allowed to either retire or enter the fishery. Retirements were handled first and occurred if the vessel reached the longevity assigned to it or if there were too many unique vessels of that class in the fishery as determined by the Kiyofuji and Okamoto (2013) data. If there were too many active vessels, vessels of the appropriate class were randomly selected (with equal probability of selection) for retirement until the number of unique vessels matched the numbers from Kiyofuji and Okamoto (2013). If after accounting for vessel retirements there were too few unique vessels in the fishery in a given year for a given class, vessels were allowed to enter as described for year 1 of the simulation.

Fishing trips per vessel were simulated at an annual level, so for each vessel the number of trips it took in a given year was simulated based on vessel class. The number of trips per year taken by a PLDW vessel was drawn from a $Conway - Maxwell - Poisson$ (CMP) distribution with $\mu = 5$ and $\nu = 10$, and the number of trips per year by an PLOS vessel given by $CMP(\mu = 20, \nu = 10)$ where the μ's were informed by Kiyofuji (2019, pers. comm.). Trips were then assigned to take place within a quarter with equal probability. Now at the trip level, length of trip was randomly drawn from a CMP distribution with $\mu = 40$ for the PLDW vessels and $\mu = 10$ for the PLOS vessels (Kiyofuji and Okamoto, 2013). In both cases, ν was assumed to be 5. Within a trip, movement across valid fishing cells for each observation model is described by the two random walk scenarios in the Methods section above and depicted for a trip length of 6 in Figure 9. Logbook catch records were simulated for each cell visited on a trip, where the number of cells visited is equal to the length of the trip.

8.2.2 Data uncertainty

Data uncertainty for the *basic*, *patchy*, and $e10$ levels were simulated in the same way as the proof-of-concept simulation. For the case-study, the $q10$ data uncertainty level built off of $e10$. When each vessel entered the fishery, the starting gear configuration of bird radar (BR), bait tank (BT), NOAA receiver (NR), and sonar (S) were defined by the empirical probability of having that configuration as defined in Kiyofuji (2013) and shown in Figure (10). A normally-distributed vessel random effect (VRE) was also assigned upon entering the fishery, however the random effects were pre-drawn and sorted so that vessels entering later in the fishery received a higher vessel random effect. Additionally, in each subsequent year, the gear configuration could be upgraded to a higher level (e.g. BR1 to BR2) based on the empirical probabilities from Kiyofuji (2013). Gear configurations were assumed to never downgrade, and higher levels of a gear configuration were always assumed to be better. Lastly, at the trip level, the number of poles fished by a vessel was determined by a draw from a CMP distribution parameterized for each class($Poles_{PLDW} \sim CMP(\mu = 22, \nu = 10)$ and $Poles_{PLOS} \sim CMP(\mu = 14, \nu = 10)$). Therefore the catchability effect for vessel v on trip t in year y is given by:

$$q_{v,y,t} = Class_v + BR_{v,y} + BT_{v,y} + NR_{v,y} + S_{v,y} + Poles_{v,t} + VRE_v$$

Under the $q10$ uncertainty level the simulated logbook catch record was defined as

$$ln(c_i) \sim Normal(\mu = qA_o a_i, \sigma = 0.1 qA_o a_i)$$

given the appropriate q for that vessel and trip.

8.3 Estimation models

8.3.1 Conventional delta-GLM

The convetional delta-GLM model was based off the structure described in Campbell (2015), and implemented using the speedglm package in R (Enea, 2017). Additive effect of time and space where β_t is the temporal effect at time t and δ_r is the spatial effect of area r:

- Binomial component: $logit(p_i) \sim \alpha + \beta_t + \delta_r + ... + \epsilon; \epsilon \sim Binomial$
- Positive component $(c_i > 0)$: $ln(c_i) \sim \alpha + \beta_t + \delta_r + ... + \epsilon; \epsilon \sim Normal$

In the case of this simulation β_t is the year effect, and δ_r is the effect for each spatial cell r.

Interactive effect of time and space:

- Binomial component: $logit(p_i) \sim \alpha + \beta_t \times \gamma_a + \delta_r + ... + \epsilon; \epsilon \sim Binomial$
- Positive component $(c_i > 0)$: $ln(c_i) \sim \alpha + \beta_t \times \gamma_a + \delta_r + ... + \epsilon; \epsilon \sim Normal$

For the interactive effect model, a sub-region term was added γ_a. Model regions were sub-divided into 4 roughly equal area sub-regions. With the interactive effect, this allowed 4 separate trends to be estimated in each model region. Spatial cells were assigned to each sub-region based on a k-means clustering of their spatial coordinates. This is the same process that is used within the VAST package (Thorson, 2019) to allocate observations to spatial knots.

8.3.2 Geostatistical delta-GLMM

The geostatistical model implemented by the VAST package v5.4.0 (Thorson et al., 2015; Thorson, 2019) is a generalized extension of a conventional delta-GLM with interactive spatiotemporal effects. The basic form is described in the following equations:

- Binomial component: $p_i \sim \beta_1(t_i) + \omega_1(x_i) + \phi_1(x_i, t_i) + ... + \epsilon$
- Positive component $(c_i > 0)$: $c_i \sim \beta_2(t_i) + \omega_2(x_i) + \phi_2(x_i, t_i) + ... + \epsilon$

$\beta(t_i)$: intercept at time t

$\omega(x_i)$: spatial random effect at location or knot x

$\phi(x_i, t_i)$: spatiotemporal random effect at time t and location x

$\omega \sim MVN(0, R)$ where R is a Matern correlation function

$$\phi \sim \begin{cases} MVN(0, R) & \text{if t=1} \\ MVN(\rho\phi_{t-1}, R) & \text{if t>1, default} \rho = 0 \end{cases}$$

Given this framework, the user has considerable flexibility to specify the error structures of the two model components, the number of x_i, how the x_i are distributed spatially, whether to allow for the estimation of ρ, or the inclusion of an environmental covariate to aide in the spatial prediction and interpolation of abundance.

This simulation did not specifically evaluate the effect of choice for number of x_i. However, Thorson (2019) suggests using the most knots that is computationally feasible. This study used 100 knots in the proof-of-concept simulation and 200 knots in the case-study simulation.

The baseline assumption in VAST is to spatially distribute the x_i in proportion to the location of the observations via k-means clustering of their spatial coordinates. Other research has noted that this assumption of preferential knot distribution should be evaluated when modeling fisheries-dependent data (Tremblay-Boyer et al., 2018; Thorson, 2019). Two separate methods were used to distribute knots spatially for the simulations: the VAST default of preferential knots (*geo.p*) and uniform knots (*geo.u*). In contrast to the *geo.p*, the *geo.u* spatially distributes the x_i in proportion to the unique $1^\circ \times 1^\circ$ cells that are sampled. The difference between these can be seen in Figure 11 and Figure 12.

Another baseline assumption made in the VAST package is that the spatiotemporal random effects (ϕ) are independent across time ($\rho = 0$). However, if a strong temporal correlation exists in the spatial observations, allowing for a non-zero ρ could improve the models' ability to predict biomass into unfished temporal strata and thus potentially improve the estimated indices. Allowing for the estimation of a non-zero ρ was used to evaluate this effect.

An additional functionality of VAST is the ability to include environmental covariates in the model to aide in the prediction of abundance, i.e. their effect is not standardized out along with vessel or catchability effects. This is done by adding in the following term, $\xi \times \psi(x_i, t_i)$, into both components of the model where $\psi(x_i, t_i)$ is the value of the covariate defined at each spatial knot and each time step, and ξ is the linear effect of that covariate. The environmental covariate used in the simulation is the monthly NOAA Smith & Reynolds Extended Reconstructed Sea Surface Temperature (Smith and Reynolds, 1981) corresponding to the dates of the SEAPODYM biomass extract (Figure 13). This data is available at a $2^\circ \times 2^\circ$ spatial resolution and was disaggregated to match the $1^\circ \times 1^\circ$ spatial resolution of the SEAPODYM output. Only a single linear effect was considered in these simulations but this can be modified to allow for more complex relationships; the inclusion of spline functions for example. It is worth noting that very recent versions of VAST have the functionality to include a temporal index as a covariate, and allow for the environmental covariates to be defined at the observation rather than the knot level.

To model catchability in the case-study simulation covariate effects were added as additive categorical effects (vessel class, bird radar, bait tank, and NOAA receiver) or additive linear effects (number of poles fished). The vessel effect was modeled as a normally distributed random effect.

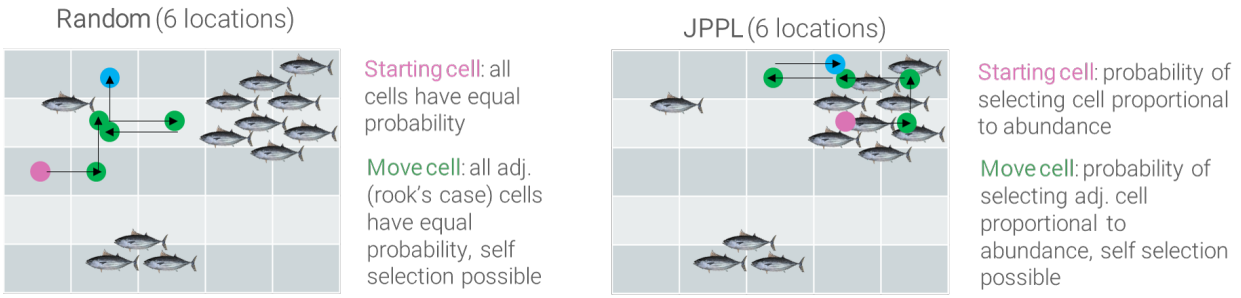

Figure 9: Depiction of the cells visited along a random walk defining the movement of the two observation models (random and JPPL) for a trip length of 6.

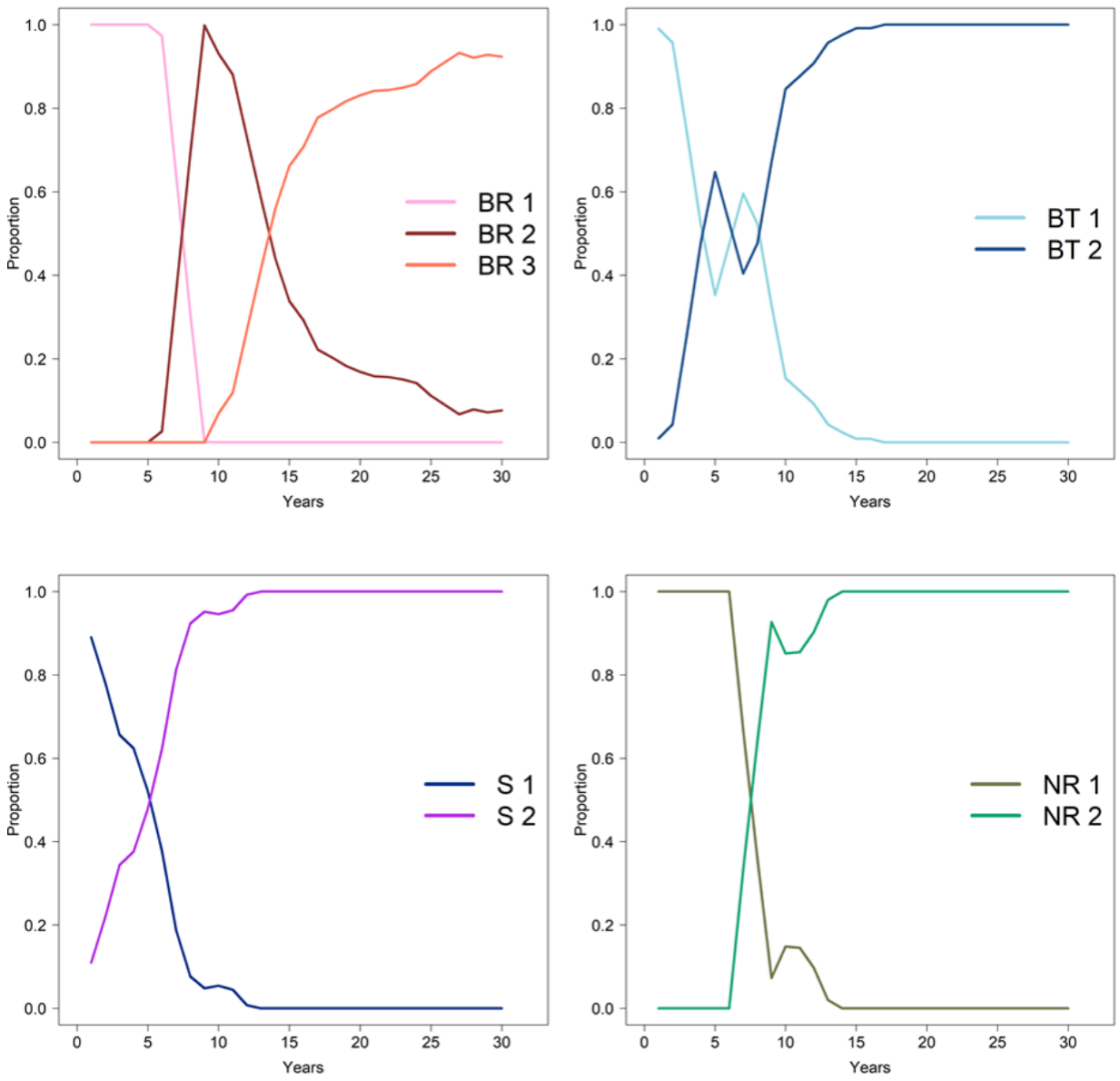

Figure 10: Depiction of the empirical probabilities of having a particular gear configuration in a given year of the simulation for bird radar (BR), bait tank (BT), NOAA receiver (NR) and sonar (S) as reported in Kiyofuji (Kiyofuji, 2013).

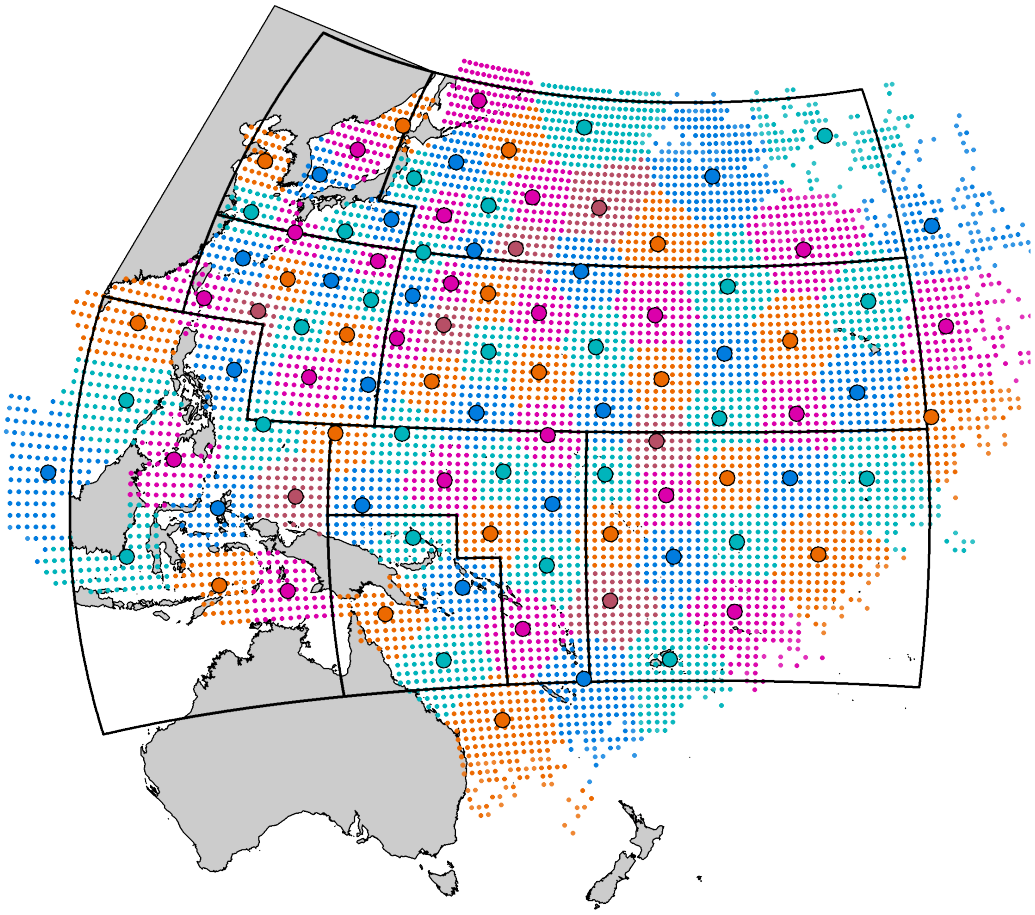

Figure 11: Preferential placement of 100 knots given the observations generated by the JPPL observation model. The knots are shown by the large colored dots, the corresponding dots of matching color are the $1^\circ \times 1^\circ$ grid cells with observations that are associated with each knot. Note that knots are clustered closer together adjacent to Japan where effort density is highest in the simulation.

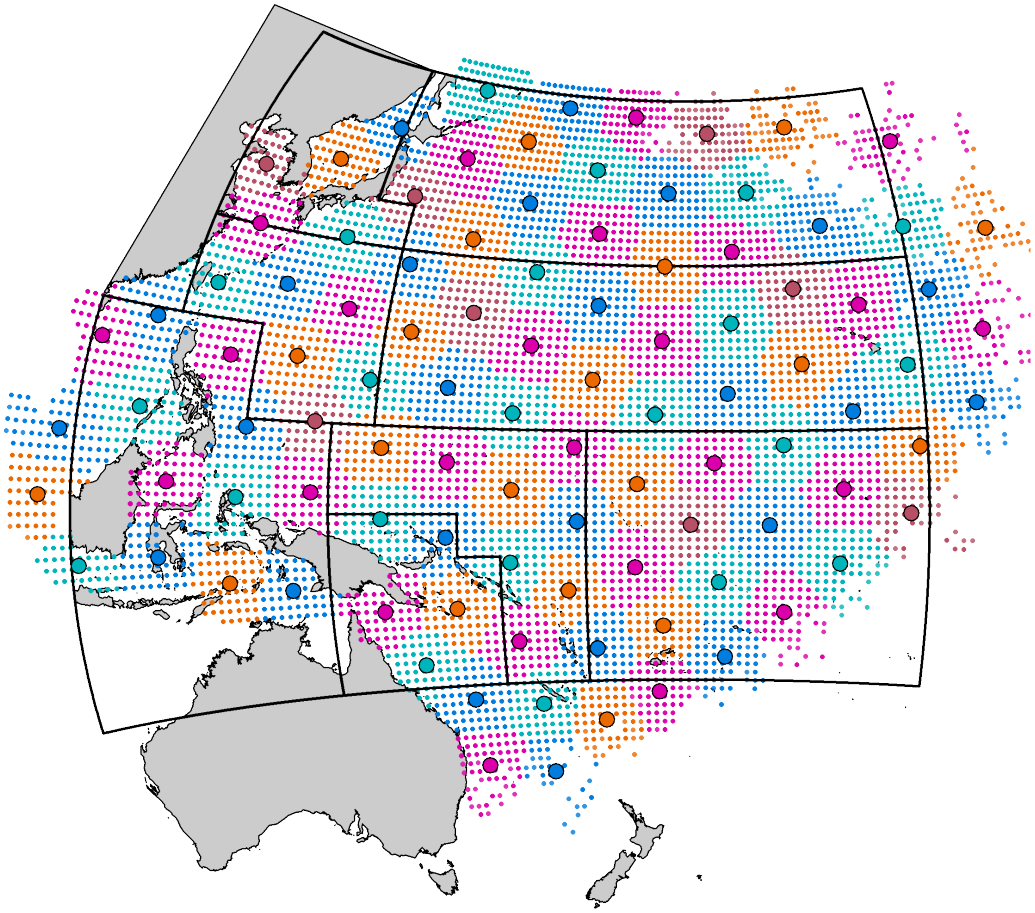

Figure 12: Uniform placement of 100 knots given the observations generated by the JPPL observation model. The knots are shown by the large colored dots, the corresponding dots of matching color are the $1^\circ \times 1^\circ$ grid cells with observations that are associated with each knot. Note that knots are evenly spaced across the $1^\circ \times 1^\circ$ grid cells with observations.

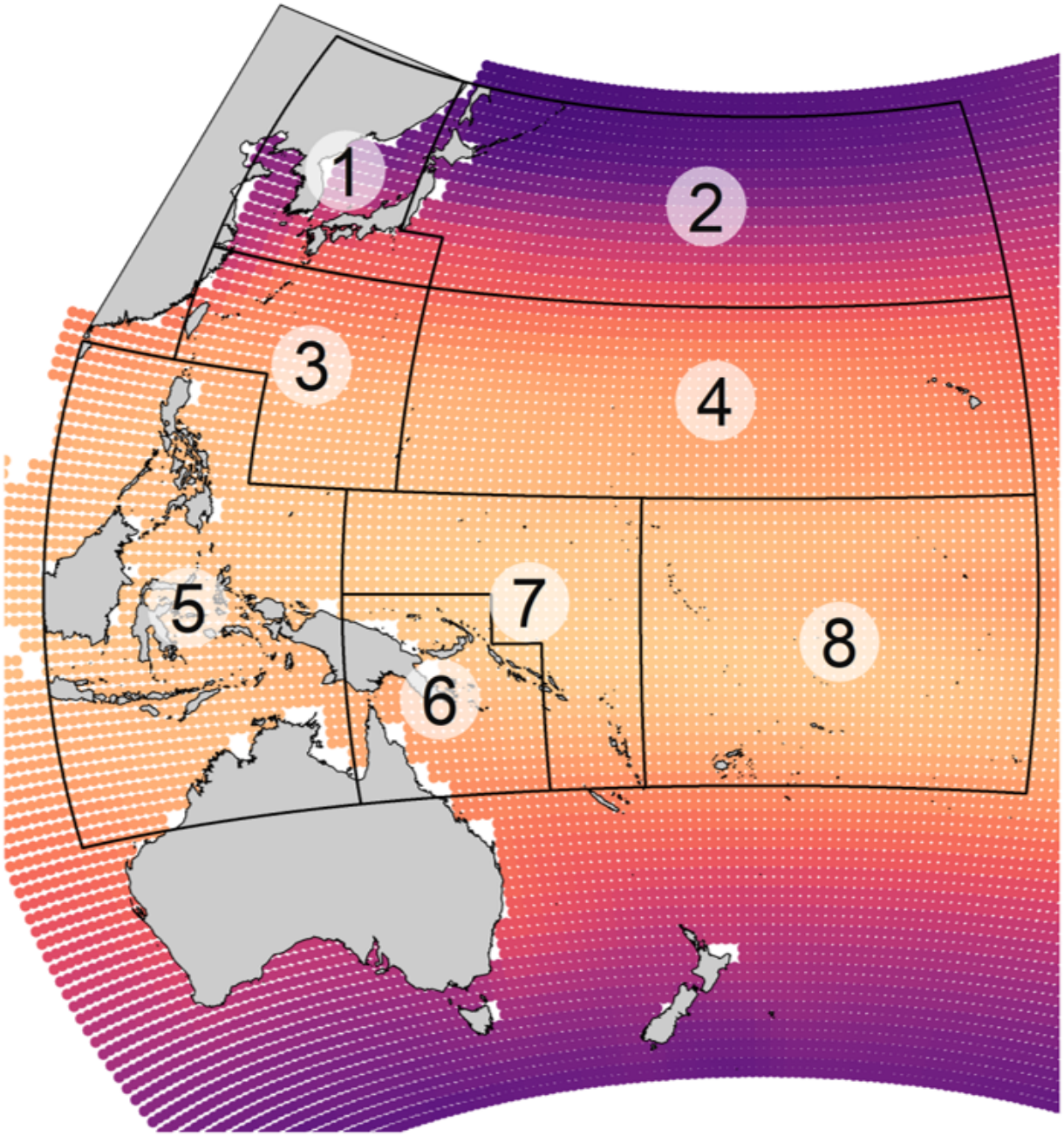

Figure 13: Average sea surface temperature (SST) distribution from the NOAA Smith & Reynolds model. Lighter colors indicate warmer temperature, and darker colors indicate cooler temperatures. The numbers correspond to the 8 regions proposed for use in the 2019 skipjack tuna assessment.