3.5 Results of space-time modelling of IPHC fishery-independent setline survey WPUE and NPUE data

Raymond A. Webster

Abstract

Space-time modelling of survey data was used to produce regulatory-area estimates of mean weight-per-unit-effort (WPUE) and numbers-per-unit-effort (NPUE) for Pacific halibut from 1998-2016, with corresponding estimates of uncertainty. While the primary data source was the International Pacific Halibut Commission fishery-independent setline survey (setline survey), this was augmented with data from three other sources: the National Marine Fisheries Service (NMFS) annual sablefish longline survey for deep waters in the Gulf of Alaska; the NMFS annual trawl survey of the Bering Sea flats; and the Alaska Department of Fish and Game triennial trawl survey of Norton Sound. Approaches to calculating adjustments for hook competition and the timing of the setline survey relative to the commercial Pacific halibut harvest were modified in 2016 with input from the Science Review Board. Overall, trends in model-estimated mean WPUE and (NPUE) by area are consistent with previous empirical estimates, with some modest changes in these indices among regulatory areas. Revised apportionment estimates maintain our current understanding of the biomass distribution among areas, although with some small, but locally important, changes from previous estimates.

Introduction

The International Pacific Halibut Commission (IPHC)'s fishery-independent setline survey (setline survey) provides data used to compute indices of Pacific halibut density for use in monitoring stock trends, apportioning biomass, and as an important input in the stock assessment. Biomass apportionment among regulatory areas is based on the annual mean weight-per-unit-effort (WPUE) for each area, computed as the average of WPUE of O32 (greater than or equal to 32" or 81.3cm in length) halibut caught at each station in an area (Webster and Stewart 2017). Mean numbers-per-unit-effort (NPUE) is used to index the trend in halibut density in the stock assessment models (Stewart and Hicks 2017). Until this year, both indices were computed as the simple arithmetic mean of station-level WPUE or NPUE for each area, with various adjustment scalars applied to the means to account for incomplete setline survey coverage in some regions in some years. Further adjustments to the WPUE index were made to account for different levels of competition with halibut for baits in each regulatory area, and the variability in the timing of the setline survey relative to the total harvest. In this work, we call this direct data-based approach to estimating density indices the "empirical method".

Webster (2016a) outlined an improved approach to estimating density indices which makes use of additional information within the setline survey catch data, along with auxiliary information collected on the survey (such as station depth). Specifically, improvements in estimation are made by fitting models to the data that account for spatial and temporal dependence, which make use of the degree to which the halibut distribution is patchy (has regions of high and low density), and that those patches tend to persist with time. For example, if WPUE is high at a particular location, it is more likely to be high at nearby locations, and at the same location in previous and subsequent years. Therefore, we not only have information about density at a location and time from a direct observation, but from other data recorded nearby in space and time. Similarly, such an approach also allows estimation of a density index at a location with no data (e.g., a location between stations, a station with an ineffective set, or a region not surveyed annually). Further improvements can come from careful selection of a covariate model for density.

This report presents the results of the first full year of using space-time modelling to estimate WPUE and NPUE indices. We begin with a short review of the statistical methods, and discuss how adjustment factors and data from other, complementary, surveys are incorporated into the modelling. In presenting the results, new estimates of the indices are compared with estimates based on methods used previously. The special case of Area 4CDE is discussed separately due to the relatively complex data compilation process it requires. Finally, revised biomass apportionment estimates are compared with estimates from the previous empirical method.

Methods

As discussed in Webster (2016b), our approach is to fit space-time models using INLA (Integrated Nested Laplace Approximation), implemented in the statistical software R through the R-INLA package (Lindgren and Rue 2015). This provides an efficient way to undertake Bayesian modelling of a process observed on a continuous spatial domain (a Gaussian random field, or GF, which in our case is a region of the ocean).

In our work, separate models were fitted to data from each regulatory area for the period 1998-2016, which spans the period in which the current setline survey grid design has been in use. Due to spatial dependence, data from nearby stations in neighbouring regulatory areas also inform the modelling for the area of interest, and we included stations within approximately a degree of latitude and longitude outside the boundaries of the regulatory area of interest.

Details of how the models are set up (definition of boundary regions and meshes) are described in Webster (2016b), but we begin by reviewing the statistical models also presented in that report.

Review of statistical models

In our modelling this year, we used semi-continuous models for both WPUE and NPUE data. The count data from which NPUE is computed could also be modelled using some variation of the Poisson distribution, but that approach was more challenging when data from other surveys were used (particularly in Area 4CDE), as count and effort data are not separately calibrated. The large number of unique values non-zero NPUE data typically take in an area means they are well approximated by a continuous distribution. Therefore, for both WPUE and NPUE modelling, we used the semi-continuous model described in the R-INLA tutorial document (Krainski et al. 2015).

For data collected in a given year, let w_i be the WPUE or NPUE value for station *i*. The value is either non-zero, if some halibut were caught, or zero if no halibut were caught. Two new variables are defined, one for presence of halibut in the catch (x_i) and the other for the WPUE or NPUE value when halibut are present (y_i):

$$x_i = \begin{cases} 1 & w_i > 0 \\ 0 & w_i = 0 \end{cases}$$

$$y_i = \begin{cases} w_i & w_i > 0 \\ NA & w_i = 0 \end{cases}$$

Here the *NA* means that y_i is a random variable that can only take non-zero values, and is therefore undefined when $w_i = 0$. The x_i have a Bernoulli distribution, $x_i \sim \text{Bern}(p_i)$, while we use a gamma distribution for the y_i, $y_i \sim \text{Gamma}(a_i, b_i)$, which has mean $\mu_i = a_i / b_i$. Now let ε_i be a GF (a Gaussian process in two or more dimensions) which is shared by both component random variables in the following way:

$$\text{logit}(p_i) = \alpha_x + \varepsilon_i$$

$$\log(\mu_i) = \alpha_y + \beta_\varepsilon \varepsilon_i$$

where α_x and α_y are intercept terms and β_ε is a scaling parameter on the shared random effect.

Spatial and temporal dependence are modelled through the spatial random field (SRF), ε_i. This is assumed to be a GF with mean zero and covariance matrix Σ, that is, $\varepsilon_i \sim \text{GF}(0, \Sigma)$. A stationary Matérn model (Cressie 1993) is assumed for the spatial covariance model, which specifies how the dependence between observations at two locations decreases with increasing distance between them. A simple autoregressive model of order 1 (AR(1)) is used to model temporal dependence, using the method described in Lindgren and Rue (2015), with the temporal correlation parameter denoted by ρ.

Covariate models

R-INLA provides a flexible framework for including covariates in the model. We included depth in the models of both the zero and non-zero processes through a random walk model of order 1. Year, also modelled as a random walk, was included in models for most areas through the non-zero process only. The reason for including this covariate was to improve prediction at unsurveyed locations distant in space and time from observed data: without this covariate, estimates of WPUE or NPUE at such locations approach the regulatory area average over the entire 1998-2016 data range, and this covariate instead ensured they followed the general trend of an area's index. Following a suggestion by the IPHC's Science Review Board (SRB), distance from the Area 4CDE shelf edge (actually the 400 fathom, or 732 m, contour) was used as a covariate for that area. WPUE and NPUE have historically been greatest along the 4CDE shelf edge, with much lower densities to the north and east along the Bering Sea flats, and the goal here was to improve prediction in the large unsurveyed regions in that part of Area 4CDE.

Other data sources

Bering Sea trawl surveys

The IPHC annual setline survey does not include stations on the eastern Bering Sea flats, except for those around St Matthew Island and the Pribilof Islands. Instead, we rely on data from annual National Marine Fisheries Service (NMFS) trawl surveys (Lauth and Nichol 2013) to construct density indices through a calibration with data from the IPHC 2006 and 2015 setline surveys in the eastern Bering Sea (Webster et al. 2016). An index for the northern Bering Sea was constructed using a combination of data from a 2010 NMFS trawl survey, and the relationship of Norton Sound trawl survey density index estimated from the approximately triennial Alaska Department of Fish and Game (ADFG) surveys (Soong and Hamazaki 2012) in that region with southern Bering Sea trawl density index from annual NMFS surveys (see Webster 2014 for details). In order to include data from these surveys in the space-time modelling, the calibration and scalings must be applied to data from individual survey stations. This is relatively straightforward for the NMFS trawl survey data: the region-wide calibration curve and scaling are applied to station-level trawl data to produce WPUE and NPUE values at each NMFS station on the same scale as the IPHC setline survey indices.

For the ADFG Norton Sound trawl survey, we have station-level catch-per-unit-effort (CPUE) data spanning the full 1998-2016 period being modelled here. In order to convert these to WPUE and NPUE indices on the same scale as the setline survey, we made use of the overlap of this survey and the 2010 NMFS northern Bering Sea trawl survey. From the latter, we have station-level WPUE and NPUE indices using the calibration and scaling discussed above. Taking the mean of these indices within the area covered by the standard ADFG Norton Sound trawl survey grid and dividing by the mean CPUE index from the same stations (computed in the same way as ADFG CPUE values) gives us multipliers to convert Norton Sound station-level CPUE values to WPUE and NPUE.

NMFS Alaska sablefish longline survey

The sablefish longline survey is conducted annually in the Gulf of Alaska, and biennially alternating in the Bering Sea and around the Aleutian Islands (Sigler and Lunsford, Unpub[1]). Since 2014, data from the NMFS sablefish longline survey have been used to index deep water (>275 fm, 503 m) in Areas 2C, 3A, 3B, and the Area 4D edge, with the Area 2C deep-water index also being applied to Area 2B (Webster et al. 2015). Sablefish survey NPUE, (i.e., numbers of halibut per 45-hook skate) indices were computed for both deep (275-400 fm, 503-732 m) and standard depth (0-275 fm, 0-503 m) waters on each set that spanned both depth ranges. The ratio of deep to standard depth NPUE was converted to a WPUE index by multiplying by the setline survey index from stations in 20-275 fm (37-503 m, the standard survey depth range excluding expansion stations) for the regulatory area in which the set was located. These set-level data were included in the modelling with locations being the mean latitude and longitude of the 275-400 fm (503-732 m) portion of the set. Inclusion of these data was important for Areas 2C, 3A, and 3B, which otherwise lacked any deep-water setline survey data, and preliminary modelling showed estimates for future deep-water stations would otherwise have been heavily influenced by higher catch rates at nearby existing stations at shallower depths.

[1]Survey protocol for the Alaska sablefish longline survey. http://www.afsc.noaa.gov/ABL/MESA/pdf/LSprotocols.pdf. Retrieved December 2016.

NMFS West Coast trawl survey

The West Coast trawl survey (Bradburn et al. 2011) catches few halibut, and Webster (2016a) showed that there was no relationship between mean setline survey WPUE and trawl WPUE where the two surveys overlapped. Because the trawl survey appears to provide a poor index of local halibut density, we did not consider using station-level data from that survey to supplement the setline survey data in California, where the latter has had no coverage south of 39°N, and coverage elsewhere in California in 2013 and 2014 only. Instead, we only estimated WPUE and NPUE down to 39°N, and for the purposes of apportionment, scaled the mean index for Area 2A using the scale factor of 0.916 computed in 2015 to account for lower densities south of there (see Webster and Stewart, 2017).

Standardisations for competition and survey timing

Prior to 2016, the WPUE series used for apportionment included two standardisations to account for competition for baits among Pacific halibut and other species, and for the timing of the setline survey relative to the total halibut harvest. Both were applied to each regulatory area's WPUE mean value, and not to the WPUE of individual stations. The standardisations were not previously applied to NPUE series used in the stock assessment, but for the space-time modelling, both indices had the standardisations applied for consistency.

For the hook competition standardisation, the relevant data (proportion of hooks returned with baits) are recorded on each set. For the space-time modelling, a station's raw WPUE was standardised for the effect of competition prior to using the data in the model. Importantly, this approach means that variability in the standardisation among stations is accounted for in the models, whereas the previous approach ignored this source of uncertainty by aggregating data across each regulatory area.

Some care was required in handling sets with no baits returned, as the usual standardisation formulae would lead to infinite WPUE in such cases. In computing the standardisation at each station, the quantity Z is required:

$$Z = \log\left(\frac{h}{b}\right)$$

where h is the number of hooks on a set and b is the number of baits returned. To avoid division by zero when no baits are returned, we add a small quantity δ to both h and b:

$$Z = \log\left(\frac{h+\delta}{b+\delta}\right)$$

where $\delta = h/100$. This choice means δ is proportional to the number of hooks set, ensuring that if no baits are returned, Z will be the same for sets of different lengths (e.g., 5 skates vs 6 skates). The adjustment factor for the standardisation is given by

$$f_H = \frac{Z}{1-e^{-Z}}$$

The standardised WPUE is then found by multiplying raw WPUE by f_H. The largest possible value for f_H with $\delta = h/100$ is 4.66, i.e., with no baits returned, WPUE is scaled up by a factor of 4.66.

For average numbers of returned baits, the standardisation leads to WPUE indices that are over twice the raw values, implying that in the absence of competition, the halibut catch would have been twice what was actually caught. In order to keep the index on a scale familiar to stakeholders, standardised WPUE is divided by the mean coastwide value of standardised WPUE in 1998 (the first year of the modern survey design). Prior to 2016, hook competition adjustment factors for each regulatory area were such that there was no adjustment made to coastwide WPUE. This had the effect of removing any effects of changes in competition over time at the coastwide level. An advantage of the new method is that now changes in WPUE due to changes in the degree of coastwide competition for baits are accounted for in the standardisation.

The second standardisation we make is to account for the effect of halibut harvest taken prior to the summer setline survey, and how this varies among areas. The input data for computing this standardisation are aggregated at the regulatory area level, and so we are unable to calculate standardisation factors for individual stations. The calculations for this standardisation have also been somewhat circular, as they depend on estimates of realised harvest rates, which depend on an estimate of exploitable biomass from the stock assessment, which in turn depends on the standardised index. To avoid this, the Scientific Review Board recommended use of the target harvest rates instead of realised harvest rates (Cox et al. 2017). This has the added advantage that harvests in excess of the target in an area are not being "rewarded" with increases in the standardisation factor the following year. A disadvantage from the perspective of stakeholders in an area with a high proportion of the harvest occurring prior to the setline survey (historically only Area 2A) is that the upward scaling of WPUE will be less in years when the harvest rate exceeds the target than it would have been using the previous approach.

While we have discussed the effect of these standardisations on WPUE, we also use exactly the same standardisations for NPUE. For the space-time modelling, the input WPUE and NPUE data include both the station-level hook competition standardisation, and the area-level timing standardisation applied to station-level WPUE and NPUE. For data from other sources described above, the mean standardisations for an area or region are applied. Thus, the sablefish survey WPUE and NPUE values for a regulatory area and year are scaled by the average adjustment factors calculated for that area and year. For trawl survey WPUE and NPUE data in Area 4CDE, the mean hook adjustment factor from the IPHC's setline surveys in 2006 and 2015 are used in those years, while the average of 2006 and 2015 factors are used in other years. All Area 4CDE input data use the same annual timing standardisation.

Prediction of WPUE and NPUE

Models for each area were refitted with the observed data augmented with a full grid of official station locations in each year with missing values instead of WPUE or NPUE data. The model output then includes predictions of the zero and non-zero processes at each of the locations and years with missing data, which are multiplied to give WPUE or NPUE. In order to estimate uncertainty (standard deviations and credible intervals), we generated 2,000 samples from the posterior distributions for each area, and computed station-level WPUE or NPUE for each sample. Mean WPUE or NPUE for a regulatory area was then computed for each of the 2,000 samples by averaging over the station values, and the final posterior mean index value was the average of these 2,000 means. Posterior standard deviations and quantiles (e.g., 95% intervals) were computed in the same way.

For areas which have already had a setline survey expansion (2A, 4A and 4CDE), prediction was made at the official location of each station in each year, whereas the modelling itself used the observed station locations which can depart slightly from the official locations. Prediction was only made at stations with mean depth across all years in the range 10-400 fm (19-732 m), which is the range of the setline survey including expansion stations. A few expansion stations had depths outside of this range (see, for example, Webster and Soderlund 2017), and although these data were included in the modelling, we did not predict WPUE or NPUE at these stations. For areas with planned future expansions (2B, 2C, 3A, 3B and 4B), prediction was done at all existing and future potential expansion stations. The prediction at stations with no data is one of the ways in which the space-time modelling provides a major improvement on the previous approach. Until now, stations or regions with missing data were either assumed to have the same mean WPUE or NPUE as the regulatory area as a whole, or (for areas which have had setline survey expansions) adjustment scalars were calculated from years in which a full expanded survey occurred and used to scale the indices in years without complete coverage. The latter method assumes that the relationship between indices in areas with annual coverage and those surveyed once (or more in some cases) remains fixed with time. The assumptions inherent in the previous approach were unlikely to hold. A further disadvantage of the empirical method is that we were essentially predicting at unsurveyed locations without accounting for any of the uncertainty in prediction.

For Area 4CDE, prediction on a 10 nmi (18.5 km) grid throughout the area was not computationally feasible, and instead we predicted at official setline survey station locations in the 75-400 fm (137-732 m) range on the IPHC 10 nmi (18.5 km) grid along the Area 4CDE edge, at the official setline station positions around St Matthew and the Pribilof Islands, and at the station locations on the standard NMFS trawl survey 20 nmi (37.0 km) grid (including 2010 northern Bering Sea stations) elsewhere. Mean WPUE and NPUE for Area 4CDE were computed by first weighting stations inversely proportional to density.

Apportionment

The 2,000 posterior samples obtained as output from each regulatory area's model are also used to compute biomass apportionment estimates. We begin by calculating apportionment estimates for each of the 2,000 samples in the usual way, by calculating a biomass index for each area as the mean predicted WPUE multiplied by bottom area, and dividing these by the sum of the biomass indices over all areas. The second and final step is to average the 2,000 apportionment estimates for each area to get the posterior means, used as the final apportionment estimates. Calculation of posterior standard deviations and quantiles is similar. Note that because we have fitted models to each area separately, the posterior samples from each area are assumed to be independent. This assumption is unlikely to be true (WPUE trends in several areas are similar), and has some effect on the estimates of uncertainty (i.e., on standard deviation and quantiles). Nevertheless, this is the first time uncertainty in the apportionment estimates has been quantified.

Results

Modelling summary

There were strong relationships of WPUE and NPUE with depth in all models, and the year variable was important for all regulatory areas except Area 3B (the reasons for this are not clear, but year was omitted from the model used for predictions for Area 3B). A useful parameter for understanding the strength of spatial dependence is the practical range, which is the distance beyond which spatial autocorrelation can be considered close to zero. The posterior mean of this parameter varied with area, and for WPUE went from 0.0033 radians (11 nmi or 20 km) for Area 2C to 0.02 radians (69 nmi or 128 km) for Area 3B. The temporal correlation parameter had posterior means from 0.89 to 0.96 for WPUE, showing strong temporal dependence in all areas (the parameter takes the value 1 with perfect correlation between observations in successive years). Parameter values were similar for total NPUE models.

WPUE and NPUE indices

Figures 1 and 2 show the mean O32 WPUE and mean total NPUE time series from the space-time modelling, compared with the time series obtained using previous, empirical methods. The empirical method uses hook standardisations applied at the regulatory area-level, but for fairness of comparison, the same timing adjustment was used (as the change in its calculation recommended by the SRB, discussed below, was not dependent on the use of space-time models). Note that due to the different ways the hook standardisation was scaled for the two methods (see above), we may expect differences between the two time series in terms of the absolute value of WPUE or NPUE. Of greatest interest here is how the trends in the time series compare, and in general the trends follow each other very closely. Even though the empirical method uses a 75:20:5 weighting on the three most recent years (greatest weighting on current year), the space-time modelling's accounting for temporal dependence often leads to a smoother time series. This is most clear in Area 2, and in Area 2A in particular. For Area 2A, the large spikes in the empirical time series were due to years with very high values of the hook competition adjustment factor; the change to station-level adjustments has helped smooth out these otherwise anomalous spikes in the series (the reason for this is discussed below).

The space-time modelling also paints a more realistic picture of uncertainty in the WPUE and NPUE time series by accounting for spatial and temporal dependence in the data. The confidence intervals for the empirical method assume observations at all stations within an area are completely independent, which the space-time modelling demonstrates is not true: when spatial dependence is present, the effective sample size will be less than the number of survey stations, leading to the standard deviation of the mean to be overestimated. Also unlike the empirical method, the space-time modelling incorporates variability among stations in the hook standardisation. Finally, by predicting WPUE or NPUE at unobserved locations, a further source of variability is accounted for in the space-time modelling, one that is ignored by the empirical method.

It is no surprise, therefore, that in most cases, the credible intervals of the space-time estimates are wider than the approximate confidence intervals from the empirical method. A clear exception to this is Area 2A, where the new space-time modelling makes use of the strong underlying spatial and temporal dependence to estimate indices that appear much more precise than those previously estimated.

Area 4CDE

Due to the relatively complex compilation of data sources required for use in the space-time modelling, and the large gaps in coverage in most years, the results for Area 4CDE are of particular interest. Figure 3 breaks down the WPUE into four components: the region covered by the annual setline survey along the continental shelf edge, the two clusters of setline survey stations around the Bering Sea islands of St Matthew and the Pribilof Islands (Area 4C islands) and the large shallow region not covered by the IPHC setline survey. The IPHC setline survey along the Area 4CDE continental edge began in 2000, surveying waters in the 75-275 fm (137-503 m) depth range, and the only information prior to that came from nearby trawl stations, which are only set as deep as 200 m (109 fm). The low catch rates at most of these trawl stations has led to low estimates of WPUE along the edge (the region in Area 4CDE covering waters from 75-400 fm, 137-732 m) in 1998 and 1999, and has likely influenced WPUE along the edge in subsequent years (due to spatial and temporal dependence). The far north of the edge was surveyed for the first time in 2016, and low catch rates in most years in nearby trawl stations could also have had an influence on estimates of WPUE in this region. This helps explain why the space-time estimates of the WPUE times series is lower in the early years relative to recent years compared with empirical estimates (Fig. 1) – in recent years, there is very close agreement in the trends of both estimates. The stations around St Matthew and the Pribilof Islands were surveyed by the IPHC for the first time in 2006, so estimates prior to that are largely informed by data from nearby trawl stations. In Webster (2014) it was noted that while there was some agreement in the data from the setline and trawl surveys around the Area 4C islands, the NMFS stations near St Matthew were deeper than the IPHC stations and did not cover the higher-density shallow waters in which the IPHC stations are set. This helps explains why the trends in the St Matthew and Area 4C Islands' time series diverge prior to 2006, as the influence of the IPHC data fades as the number of years prior to 2006 increases. Note also that the low number of stations in these small island regions means that uncertainty is high in these WPUE estimates.

Figure 4 shows a map of predicted WPUE for 2016 at each station location used in compiling the overall index (together with 15 stations shallower than 10 fm, or 18 m, mainly in Norton Sound, not used in the index), and the corresponding standard deviations. With large coverage gaps in the northern Bering Sea, we are interested in how well the model is predicting WPUE in this region in the absence of direct observations. The NMFS 2010 trawl survey caught almost no halibut in much of this area, with the higher catch rates occurring at the stations closest to shore in the east, and in Norton Sound. Indeed, the 2016 estimates are very consistent with that. In the absence of covariates, locations far in time and space from observed data tend towards the mean across the entire region being modelled. For this reason, we included distance from the shelf edge (the 400 fm, 732 m contour) as a covariate, in order to improve prediction in areas most distant from observed data. This appears to have achieved its goal, and even six years after the only full northern Bering Sea survey, WPUE is estimated to be close to zero where the 2010 survey caught few halibut. Note that halibut captured in the ADFG Norton Sound surveys (most recently in 2014) continue to result in higher estimates of WPUE in that region in 2016.

Apportionment

Estimated biomass proportions for the start of 2017 calculated from the 2,000 posterior samples from the space-time modelling are shown in Figure 5, alongside estimates from the empirical method. There is generally good agreement between estimates from the two methods. The largest absolute differences between the two methods occur in Areas 2B, 2C and 4B. For Area 2A, the difference is also relatively large (2.22% of coastwide biomass for space-time method, vs 2.62% for empirical method), although small in absolute terms.

In Area 2B, prediction of WPUE in the shallow waters of Dogfish Bank and the unsurveyed Strait of Georgia (influenced by low WPUE in the adjacent waters of the Salish Sea in Area 2A) contributed to lower WPUE estimates relative to other areas compared to estimates obtained by the empirical method, which had assumed WPUE in those regions was the same as WPUE measured elsewhere in Area 2B.

In Areas 2A, 2C and 4B, the change to a station-based hook adjustment, instead of one computed at the regulatory area level, was a major reason for the apportionment changes. In computing the overall hook standardisation factor for a regulatory area, the previous method gave greater weight to stations with lower levels of competition (more baits returned). The revised approach leads to different weightings for the stations that depend on their level of competition and WPUE. In Area 2A, stations with lower than average observed WPUE had more competition on average than stations with greater WPUE: mean standardisation factors were 1.54 and 1.40 for below and above average WPUE respectively. To get a standardised mean WPUE outside of the modelling, we multiply station-level adjustment factors by the station WPUE. The lower adjustments for higher WPUE stations means they have less weight in computing a mean WPUE (the mean of the adjusted station WPUEs) than the lower WPUE stations. This in turn leads to a lower standardised WPUE for Area 2A compared with areas where competition was more evenly distributed among high and low-WPUE stations, or areas where the greatest competition was at high-WPUE stations. Area 4B is an example of the latter, with mean adjustment factors of 0.79 and 1.05 for below and above-average WPUE stations respectively, while Area 2C falls in between Areas 2A and 4B. The change in the hook competition standardisation may have also affected Area 2B's apportionment estimate, with above-average WPUE stations having a mean adjustment factor of 1.33 compared with 1.43 for below-average WPUE stations.

Discussion

The space-time modelling approach to estimating mean WPUE and NPUE time series has several important advantages over the previous empirical method. More information in the data is used, specifically information on spatial and temporal dependence, and auxiliary information from covariates can be incorporated into the modelling to further improve estimates. Estimates for regions without survey coverage can be included without making possibly unreasonable assumptions about the relationship of density in such regions with density elsewhere in an area. Estimation of uncertainty is greatly improved, by estimating uncertainty in regions without data, allowing for variation among stations in the adjustment factors calculated for the hook competition standardisation, and by not assuming that stations are sampled independently (an important, but incorrect, assumption in previous estimates of the standard deviation). On a practical level, the space-time modelling approach simplifies calculations by dispensing with the complex and ad hoc set of adjustments for incomplete spatial coverage previously employed. The one disadvantage of the new approach is the large amount of computing effort required, which has placed some constraints on the scope of the modelling we are able to attempt (for example, fitting coastwide models are not yet possible), and on our ability to produce high-resolution maps of predicted

WPUE and NPUE. With improved computing power becoming available, we are hopeful some of these limitations will be overcome in the near future.

There are still improvements that can be made to the modelling itself. This includes consideration of additional covariates, such as the oceanographic variables measured using water column profilers deployed on setline survey sets in recent years, e.g., pH, bottom temperature, and dissolved oxygen (Sadorus and Walker 2015). Future work will also involve extending the time series backwards, to include the less comprehensive surveys undertaken in the 1990s and possibly earlier (Soderlund et al. 2012). Finding a way to incorporate uncertainty in the calibration and scaling of the trawl survey data in Area 4CDE is also a future goal. More work needs to be done examining parameter sensitivity, which appears to be an issue in some areas. For example, when modelling Area 4CDE data, a modest change to the model (addition of a covariate) led to large changes in the posterior distributions of the parameters of the spatial and temporal components of the model, but the reason for this was not obvious.

Improvements in the data are expected in coming years. Setline survey expansions in Areas 4B (2017), 2B and 2C (2018) and 3A and 3B (2019) are planned, and NMFS has proposed a repeat of the 2010 northern Bering Sea trawl survey in 2017. The ADFG Norton Sound trawl survey is also due to take place in 2017, which should allow us to improve the calibration of this survey with the NMFS trawl survey stations in Norton Sound. As the planned setline survey expansion nears completion, it is important for us to evaluate the need for repeats of the expansions in the future. The space-timing modelling is the perfect tool for this, as through modelling we can study the effect of new data on overall WPUE estimates and on estimates of their precision.

References

Bradburn, M. J., Keller, A. A., and Horness, B. H. 2011. The 2003 to 2008 U.S. West Coast bottom trawl surveys of groundfish resources off Washington, Oregon, and California: Estimates of distribution, abundance, length, and age composition. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFSC-114.

Cox, S. P., Ianelli, J. and Mangel, M. 2017. Reports of the IPHC Scientific Review Board. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC-2016-RARA-26-R: 443-452.

Cressie, N. 1993. Statistics for spatial data. Wiley, New York.

Lindgren, F., and Rue, H. 2015. Bayesian spatial modelling with R-INLA. Journal of Statistical Software. 63: 1-25.

Krainski, E. T., Lindgren, F., Simpson, D., and Rue, H. 2015. The R-INLA tutorial on SPDE models. Retrieved 30 November 2015. http://www.math.ntnu.no/inla/r-inla.org/tutorials/spde/spde-tutorial.pdf

Lauth, R. R., and D. G. Nichol. 2013. Results of the 2012 eastern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate resources. U.S. Dep. Commer., NOAA Tech. Memo. NMFS-AFSC-256.

Soong, J. and Hamazaki, T. 2012. Analysis of Red King Crab Data from the 2011 Alaska Department of Fish and Game Trawl Survey of Norton Sound. Alaska Department of Fish and Game Fishery Data Series No. 12-06.

Sadorus, L. L., and Walker, J. 2015. IPHC oceanographic monitoring program 2014. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2014: 351-356.

Soderlund, E., Randolph, D. L. and Dykstra, C. 2012. IPHC Setline Charters 1963 through 2003. IPHC Tech. Rep. No. 58.

Stewart, I. J. and Hicks, A. C. 2017. Assessment of the Pacific halibut stock at the end of 2016. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC-2016-RARA-26-R: 365-394.

Webster, R. A. 2014. Construction of a density index for Area 4CDE. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2013: 261-288.

Webster, R. A. 2016a. Indexing density in southern Area 2A using West Coast trawl survey data. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2015: 544-551.

Webster, R. A. 2016b. Space-time modelling of setline survey data using INLA. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2015: 552-568.

Webster, R. A., Dysktra, C. L., and Henry, A. M. 2016. Eastern Bering Sea setline survey expansion and trawl calibration. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2015: 530-543.

Webster, R. A., Dysktra, C. L., Henry, E., Soderlund, E. and Kong, T. 2015. Setline survey expansions in 2014 and use of sablefish longline survey data for a deep-water density index. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2014: 603-617.

Webster, R.A., and Soderlund, E. 2017. Area 4CDE edge IPHC survey expansion. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC-2016-RARA-26-R: 216-219.

Webster, R. A. and Stewart, I. J. 2017. Setline survey-based apportionment estimates. Int. Pac. Halibut Comm. Report of Assessment and Research Activities 2016. IPHC-2016-RARA-26-R: 395-402.

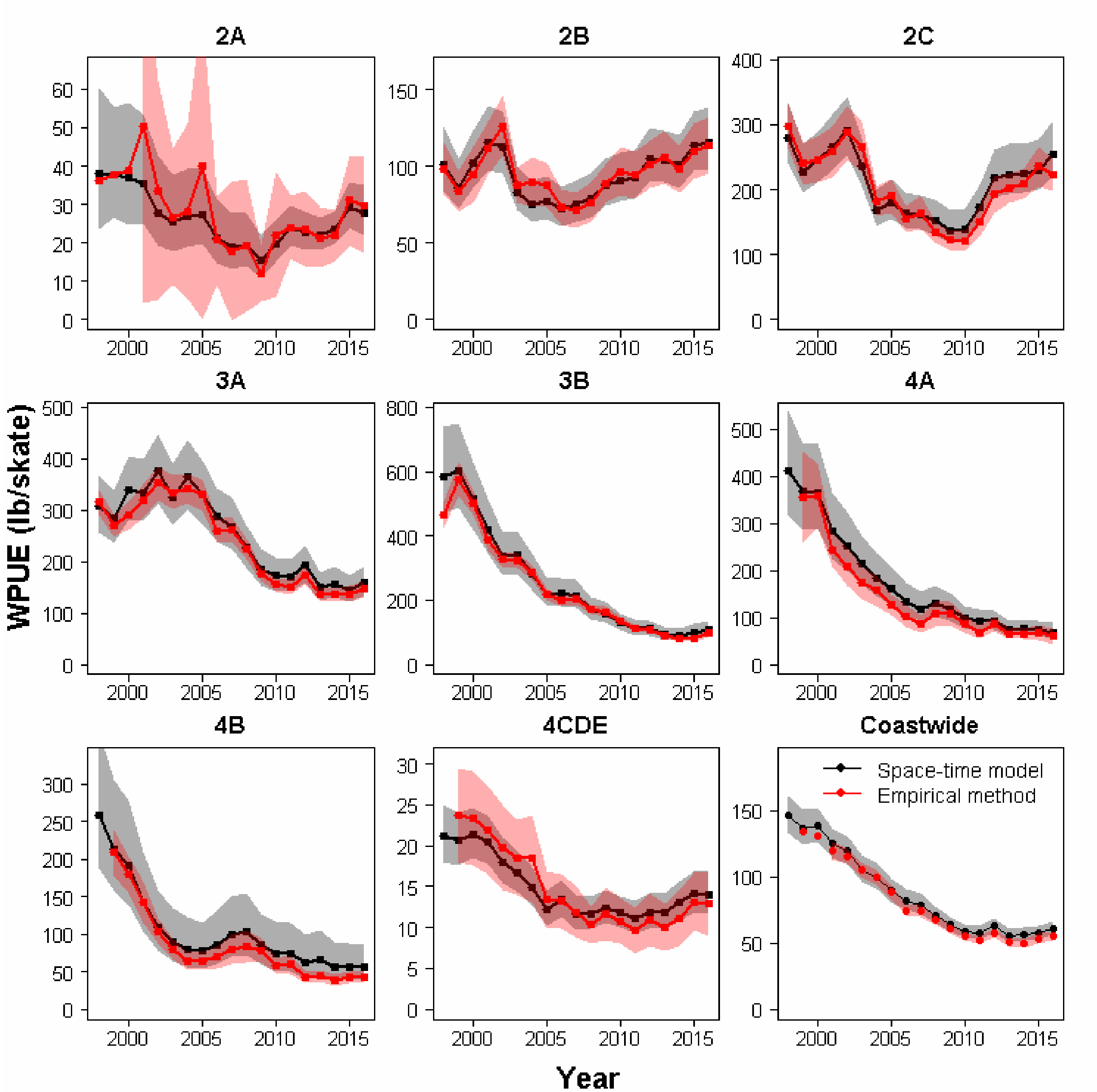

Figure 1. Comparison of estimated WPUE time series with the series calculated from space-time model predictions (“Space-time model”) and from observed data with adjustment factors applied (“Empirical method”). For space-time model results, the gray shaded region represents the 95% posterior credible interval, while for the empirical method, the red shaded region shows an approximate 95% confidence interval.

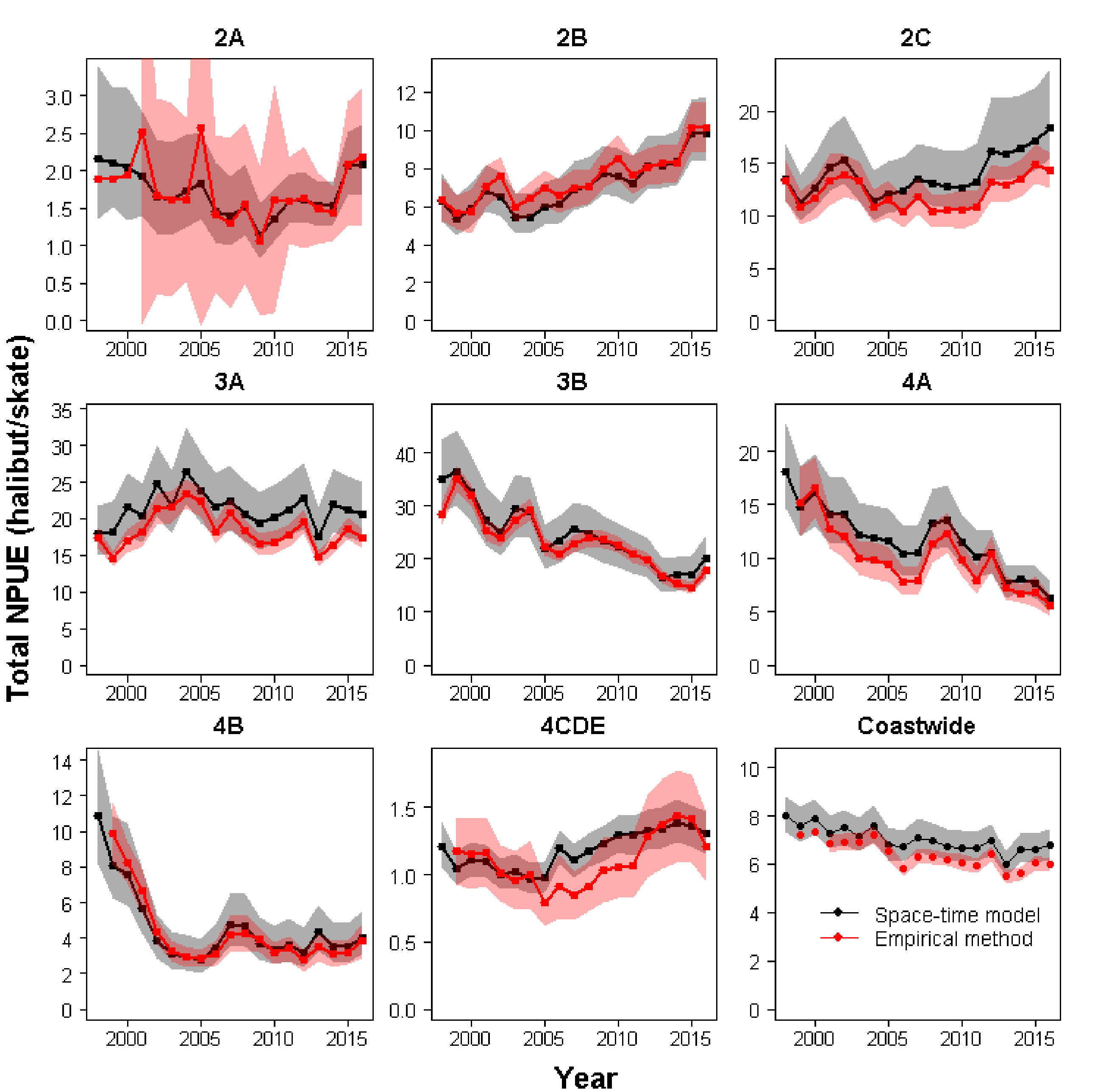

Figure 2. Comparison of estimated total NPUE time series with the series calculated from space-time model predictions ("Space-time model") and from observed data with adjustment factors applied ("Empirical method"). For space-time model results, the gray shaded region represents the 95% posterior credible interval, while for the empirical method, the red shaded region shows an approximate 95% confidence interval.

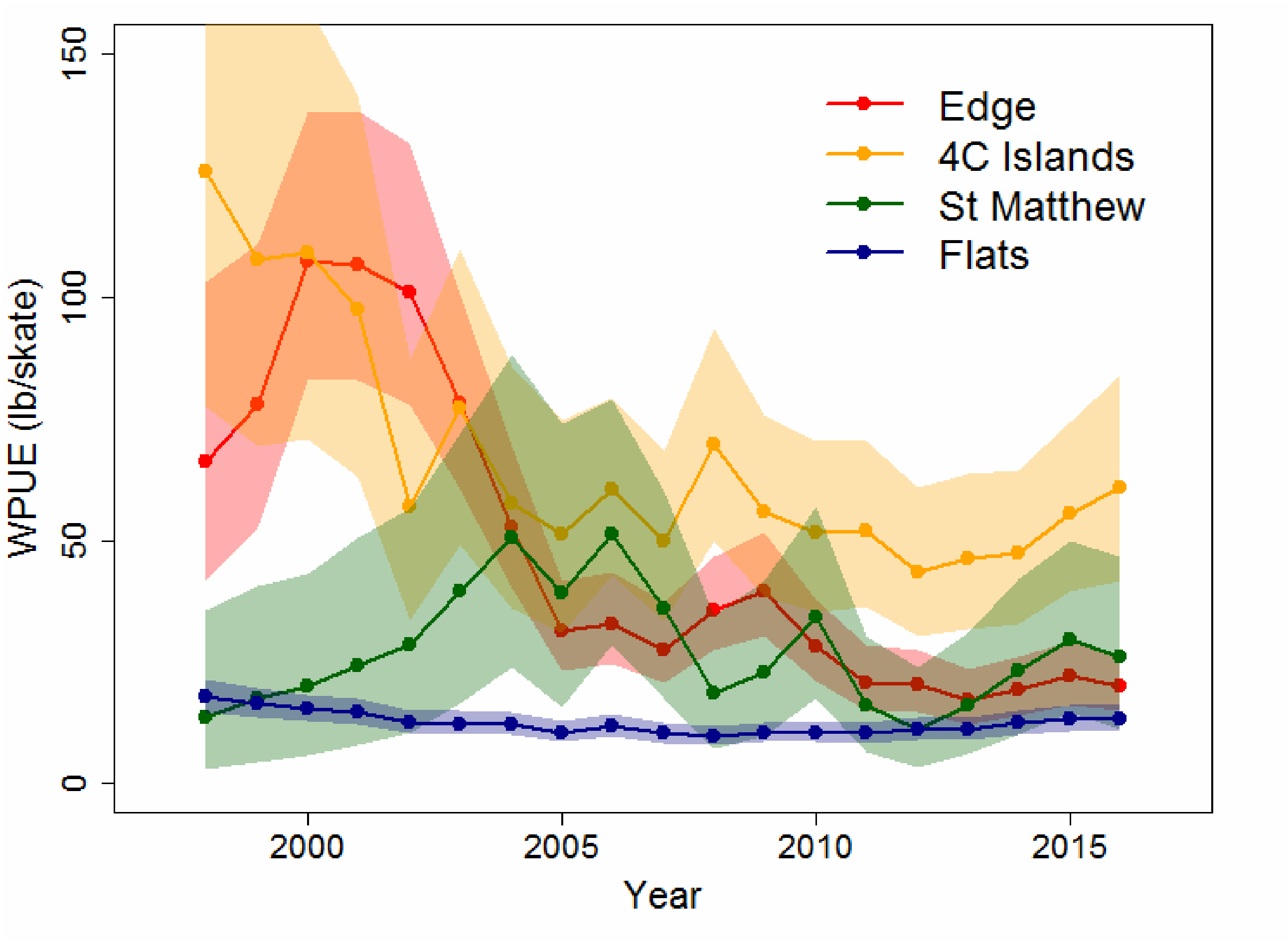

Figure 3. Comparison of estimated WPUE time series calculated from space-time model predictions for components of Area 4CDE. The shaded regions represent 95% posterior credible intervals.

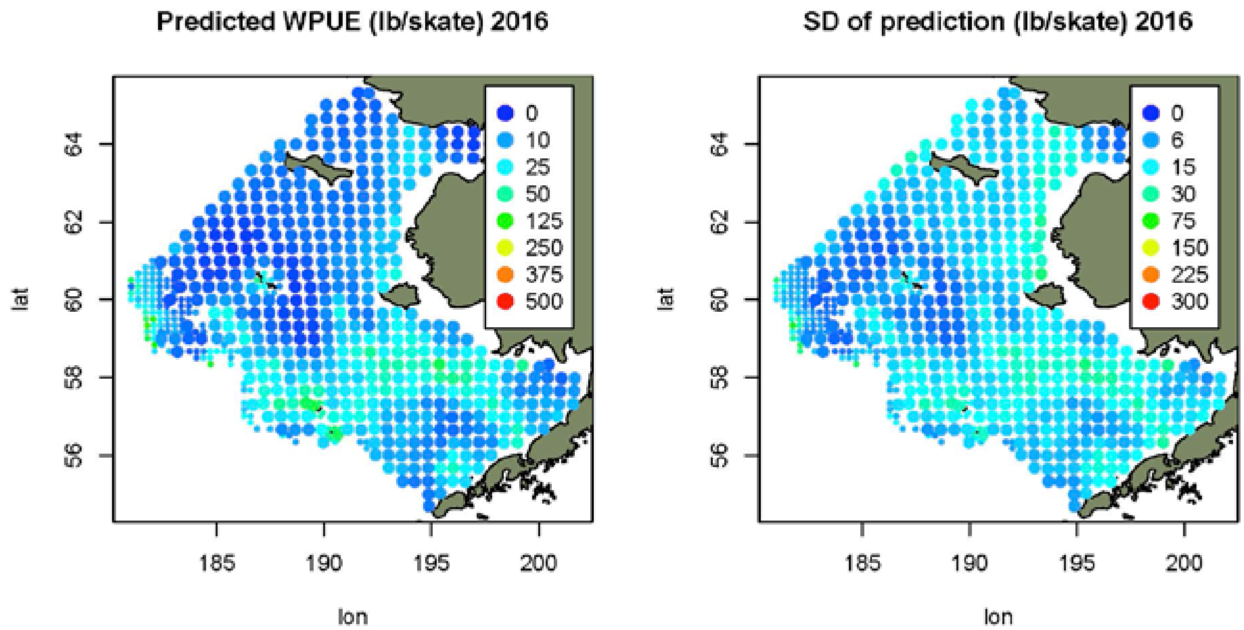

Figure 4. Predicted station WPUE in Area 4CDE in 2016, and corresponding standard deviations. Small circles represent the locations of IPHC fishery-independent setline survey stations, while large circles are NMFS trawl survey stations on the standard 20 nmi (37 km) grid.

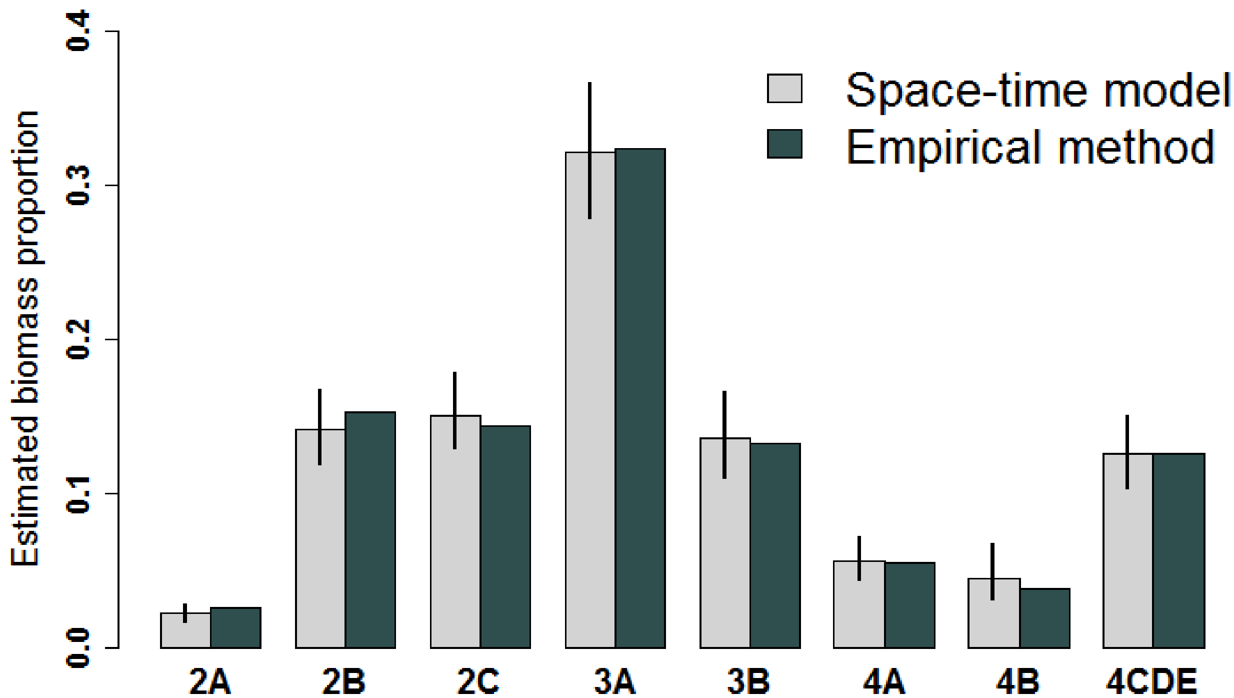

Figure 5. Posterior means of apportionment estimates (estimated biomass proportions) for the start of 2017 by area from the space-time modelling, compared with estimates from the empirical method used previously. The vertical bars represent 95% posterior credible intervals for the space-time model estimates.