Impacts of different types of data integration on the predictions of spatio-temporal models: a fishery application and simulation experiment

Arnaud Grüss[1, a*], Richard L. O'Driscoll[1, b], James T. Thorson[2, c], Jeremy R. McKenzie[3, d], Sira L. Ballara[1, e], Anthony R. Charsley[1, f]

[1] NIWA (National Institute of Water and Atmospheric Research Ltd), 301 Evans Bay Parade, Greta Point, Wellington 6021, New Zealand.

[2] Resource Ecology and Fisheries Management, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way N.E., Seattle, WA 98115, USA.

[3] NIWA (National Institute of Water and Atmospheric Research Ltd) Auckland, Private Bag 99940, Auckland, 1149, New Zealand.

Author email addresses

[a] Arnaud.Gruss@niwa.co.nz
[b] richard.odriscoll@niwa.co.nz
[c] james.thorson@noaa.gov
[d] Jeremy.McKenzie@niwa.co.nz
[e] Sira.Ballara@niwa.co.nz
[f] Anthony.Charsley@niwa.co.nz

Keywords: Integrated modelling; spatio-temporal models; spatially varying catchability; southern hake; simulation experiment.

Funding: This work was supported by NIWA Strategic Science Investment Funding and NIWA Structural Internal Project Funding. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Corresponding author

Dr. Arnaud Grüss
National Institute of Water and Atmospheric Research
301 Evans Bay Parade
Greta Point, Wellington 6021, New Zealand
Telephone: +64 4 386 0580
Email: Arnaud.Gruss@niwa.co.nz

ABSTRACT

Integrated spatio-temporal models, which enable the sharing of information across locations, time and data sources, are gaining traction for their potential to generate more precise and more accurate estimations compared to models fitted to single data sources. Standard integrated spatio-temporal models combine multiple data sources via a catchability factor. Recently, spatially varying catchability (SVC) integrated spatio-temporal models were developed to implement data integration via the estimation of an SVC term for the least reliable data sources. Expanded-domain integrated spatio-temporal models are models integrating data from different spatial areas. Spatio-temporal models can combine standard or SVC integrated modelling with expanded-domain integrated modelling. The above-mentioned types of data integration have never been evaluated through a comparative analysis. Here, we investigate the impacts of these different types of data integration on the predictions of spatio-temporal models, via an application to the southern hake (*Merluccius australis*) HAK4 stock, where the bottom trawl data collected within the New Zealand observer programme are integrated with the data from five different bottom trawl research surveys, and a simulation experiment. In total, six models were compared in the present study, where the three last models constitute expanded-domain integrated models: (Model 1) a model fitted to observer-only data for HAK4; (Model 2) a standard integrated model fitted to both observer and survey data for HAK4; (Model 3) an SVC integrated model fitted to both observer and survey data for HAK4; (Model 4) a model fitted to observer-only data for HAK4 and the other New Zealand hake stocks (HAK1 and HAK7); (Model 5) a standard integrated model fitted to both observer and survey data for HAK4, HAK1 and HAK7; and (Model 6) an SVC integrated model fitted to both observer and survey data for HAK4, HAK1 and HAK7. For the simulation experiment, we produced simulated data from Model 5, fitted the six models to the simulated data, and evaluated the performance of the models by comparing their estimations

65 to the simulated data. Overall, the indices obtained with the different types of integrated
66 models outperformed the indices obtained with models using observer-only data: indices from
67 the integrated models were more precise, better matched the traditional stratified random
68 index and had less bias and a smaller root-mean-squared-error, yet characterised uncertainty
69 less well. Moreover, expanded-domain integrated models outperformed other models
70 regarding habitat assessments: (1) they provided insights into spatial density patterns for
71 much larger regions and predicted these patterns more precisely for the area common to all
72 models; and (2) models combining expanded-domain integrated modelling with standard or
73 SVC integrated modelling predicted patterns of distribution shifts and range
74 expansion/contraction more precisely than the expanded-domain integrated model employing
75 observer-only data. Finally, expanded-domain integrated models did not outperform other
76 integrated models regarding supporting stock assessments. The indices obtained with
77 expanded-domain integrated models were more precise and had less bias and a smaller root-
78 mean-squared-error, but agreed less with the traditional stratified random index and
79 characterised uncertainty less well. This result suggests that standard or SVC integrated
80 models should be preferred to produce indices for local areas.

81

1. Introduction

Resource models, including population, stock assessment and species distribution models, are central tools to support decision-making in marine and terrestrial systems, including fisheries management (Robinson et al., 2020; Schaub et al., 2024). Population and stock assessment models assist conservation efforts by evaluating species' conservation status and extinction risk in response to population dynamics drivers and management actions (Ellner and Fieberg, 2003; Saunders et al., 2018), and harvest management by determining sustainable levels of removals (Hilborn and Walters, 1992). Species distribution models estimate relationships between species' probability of encounter or abundance and covariates such as abiotic environmental variables; these relationships can then be used to inform resource management about spatial patterns of species' probability of encounter or abundance in the past or the future (Elith and Leathwick, 2009; Sofaer et al., 2019). For example, species distribution models can provide insights into the potential impacts of climate change on living resources and fisheries (e.g., Brodie et al., 2022; Liu et al., 2023) and contribute information to marine spatial planning (e.g., Grüss et al., 2019a; Paradinas et al., 2022).

Resource models should be fitted to the best available data to support resource management. Therefore, structured data, i.e., data collected using well-designed sampling protocols such as the data derived from well-designed research surveys, are a preferred choice for resource modelling (Isaac et al., 2020). However, obtaining structured data across the entire spatial habitat of a population is generally cost prohibitive; the collection of structured data tends to be restricted to particular locations and times of the year (Whittaker et al., 2005; Bourdaud et al., 2017; Webster et al., 2020; Rufener et al., 2021). Consequently, resource models relying solely on a given source of structured data (e.g., the data collected by a given well-designed research survey) often provide statistical inferences that are too limited to be useful to resource management (Schaub and Abadi, 2011; Robinson et al., 2020).

Unstructured data (e.g., data collected by observers placed on commercial fishing boats) are usually obtained at much larger spatial and temporal scales and in larger quantity, but are of lesser quality because they are associated with less formal or no sampling design and protocols (Fithian et al., 2015; Isaac et al., 2020; Grüss et al., 2023d).

Instead of choosing between individual data sources and addressing trade-offs between data quality and data quantity, multiple data sources can be leveraged in the same resource model. Employing multiple data sources in the same population, stock assessment or species distribution model is referred to as "data integration", "data fusion" or "integrated analysis" (Pacifici et al., 2017; Miller et al., 2019; Isaac et al., 2020; Dovers et al., 2024). In recent decades, we have seen the increasing development of integrated population models (IPMs), integrated stock assessment models (ISAMs) and integrated species distribution models (ISDMs), which implement the model-based integration of multiple different datasets where each dataset provides some information about the population of interest.

IPMs and ISAMs, which have been employed in numerous studies to estimate population abundance and demographic rates, are very similar approaches, with the former typically applied in terrestrial studies and the latter usually used in fisheries research (Schaub et al., 2024). IPMs consist of the joint analysis of population count data and demographic data, with the goal of evaluating the demographic drivers of change in a population and trends in demographic rates and population abundance (Schaub and Abadi, 2011; Schaub and Kéry, 2022). ISAMs result from the complexity of modern stock assessments, which require a diversity of data types collected via different sampling techniques that contain information about absolute stock abundance and fishing mortality (Maunder and Punt, 2013; Punt et al., 2013). Nowadays, fisheries scientists very often employ ISAMs to conduct stock assessments, where a population dynamics model is fitted to catch data, age or length composition data from fisheries and/or research surveys, and indices of relative abundance, among, potentially,

other datasets (Methot, 2009; Methot and Wetzel, 2013; Doonan et al., 2016; Punt et al., 2020). Inferences with IPMs and ISAMs are based on the multiplication of the likelihoods from the different sources of information, i.e., the joint likelihood (or the sum of the negative log-likelihoods from the different sources of information, i.e., the joint negative log-likelihood) (Maunder, 2004; Schaub and Abadi, 2011). The calculation of the joint likelihood is possible when several components of that likelihood have one or several parameters in common and when two key assumptions are met: (1) the individuals from the different data sources have the same demography; and (2) the different data sources are independent (unless analysts explicitly model dependency among data sources) (Schaub and Kéry, 2022; Schaub et al., 2024). IPMs and ISAMs can provide numerous benefits. They can deliver a larger number of demographic quantities and more precise estimates for the demographic quantities delivered (e.g., Besbeas et al., 2002; Brooks et al., 2004; Abadi et al., 2010, 2012). Moreover, according to simulation experiments, they can estimate population abundance and demographic rates more accurately (Dorazio, 2014; Zipkin et al., 2017).

ISDMs consist of the integration of multiple different monitoring data sources including, typically, research surveys, programmes placing observers onboard fishing vessels (observer programmes), and/or datasets gathering miscellaneous occurrence observations (presence-only data) (Fithian et al., 2015; Grüss et al., 2017, 2023a; Dovers et al., 2024). Other sources of data such as predators' stomach contents ("predators as samplers") can also be integrated (e.g., Grüss et al., 2023d). In some applications, the different data sources integrated in the same species distribution model also correspond to different data types, e.g., encounters/non-encounters, counts and biomass-sampling data in Grüss and Thorson (2019). Like IPMs and ISAMs, ISDMs rely on the joint likelihood for inferences and must meet the assumptions that the individuals from the different datasets have the same demographic characteristics and that there is independence between the different data sources (Fithian et

al., 2015; Pacifici et al., 2017; Grüss and Thorson, 2019; Miller et al., 2019; Grüss et al., 2023d). Like IPMs and ISAMs, ISDMs come with many potential advantages, including: (1) reduced uncertainty in model estimates (Grüss and Thorson, 2019; Thompson et al., 2022; Dovers et al., 2024); (2) more accurate model estimates, according to cross-validation procedures and simulation experiments (Fithian et al., 2015; Fletcher et al., 2019; Grüss and Thorson, 2019; Robinson et al., 2020; Thompson et al., 2022); (3) a better coverage of different environmental conditions allowing for enhanced predictions of species' whole ranges and population dynamics across space and time, which is central to efficient conservation efforts (Guisan et al., 2013; Pagel et al., 2014; Jetz et al., 2019; Pinto et al., 2019; Mäkinen et al., 2024); (4) the production of less uncertain indices of relative abundance for a longer time period, which is highly beneficial to stock assessments and fisheries management (O'Leary et al., 2020; Monnahan et al., 2021; Rufener et al., 2021; Grüss et al., 2023d); and (5) enhanced spatial biodiversity and species richness predictions to assist conservation planning and ecosystem-based management, particularly for under-sampled taxa and regions (Isaac et al., 2020; Mäkinen et al., 2024).

Many of the ISDMs used in fisheries science are spatio-temporal models, which estimate unmeasured (latent) spatial variation that is constant over time (henceforth "spatial variation") and unmeasured spatial variation that changes over time (henceforth "spatio-temporal variation") (e.g., Monnahan et al., 2021; Grüss and Thorson, 2019; Gonzalez et al., 2021; Rufener et al., 2021; Thompson et al., 2022; Grüss et al., 2023a, 2023d; Nephin et al., 2023). Spatio-temporal models, even when they do not integrate multiple data sources, allow for the sharing of information across locations and time periods, thereby usually generating less uncertain (Shelton et al., 2014; Thorson et al., 2015; Thorson and Haltuch, 2019; Hodgdon et al., 2020) and more accurate estimates (Grüss et al., 2019b; Zhou et al., 2019; Brodie et al., 2020; Hsu et al., 2022) compared to models that do not account for spatial or

182 spatio-temporal variation (i.e., non-geostatistical models). While integrated analysis with
183 spatio-temporal models can be implemented in different ways, a terminology of integrated
184 spatio-temporal models is lacking. In the present study, we provide such a terminology.
185 There are three types and five sub-types (categories) of integrated spatio-temporal
186 models (Table 1). The three types of integrated spatio-temporal models include standard
187 integrated models, spatially varying catchability (SVC) models, and expanded-domain
188 models. Standard and SVC integrated spatio-temporal models combine multiple data sources
189 via, respectively, a non-spatially varying catchability term and an SVC term for the least
190 reliable data sources, while expanded-domain integrated spatio-temporal models are models
191 that integrate data sources for different spatial areas. Because spatio-temporal models can
192 combine standard or SVC integrated modelling with expanded-domain integrated modelling,
193 we can then distinguish between five sub-types (categories) of integrated spatio-temporal
194 models: (1) standard models that integrate data from different sources but from the same
195 spatial area, via a non-spatially varying catchability term; (2) SVC models that integrate data
196 from different sources but from the same spatial area, via an SVC term; (3) expanded-domain
197 models that integrate data from the same source but from different spatial areas (e.g., from
198 different quota management areas (QMAs)); (4) expanded-domain models that integrate data
199 from different sources and from different spatial areas, via a non-spatially varying catchability
200 term; and (5) expanded-domain models that integrate data from different sources and from
201 different spatial areas, via an SVC term (Table 1).
202 Standard integrated spatio-temporal models are the most common type of integrated
203 spatio-temporal models. Compared to spatio-temporal models fitted to one single data source,
204 standard integrated spatio-temporal models include a catchability factor accounting for the
205 fact that catchability differs among the different data sources that are being integrated. This
206 catchability factor, which is modelled either as a random effect (e.g., Grüss et al., 2018a,

2018b, 2020; Thompson et al., 2022; Nephin et al., 2023) or as a fixed effect (e.g., Grüss and Thorson, 2019; Gonzalez et al., 2021; Rufener et al., 2021; Alglave et al., 2022; Grüss et al., 2023a), enables the estimation of fishing-power ratio of each data source relative to a reference data source (deemed to be the most reliable data source) (Grüss and Thorson, 2019; Monnahan et al., 2021). The estimation of the fishing-power ratio of each data source relative to the reference data source ensures that predictions are intercalibrated to be in the units of the reference data source (Grüss et al., 2023a). Treating the catchability factor as a fixed effect rather than as a random effect may be preferable to avoid implicitly assuming that the fishing-power ratio follows some shared distribution; treating the catchability factor as a random effect leads to implicitly assuming that the different data sources have similar fishing power on average (Grüss and Thorson, 2019; Grüss et al., 2023a).

SVC integrated spatio-temporal models are a very recent development in the field of spatio-temporal modelling (Bolser et al., 2023; Grüss et al., 2023d; Olmos et al., 2023; Thorson et al., 2023). SVC integrated spatio-temporal models consist of substituting the catchability factor employed in standard integrated spatio-temporal models with spatially varying coefficients representing SVC for the non-reference data sources (Grüss et al., 2023d). In this setting, the reference dataset is assumed to have constant catchability (like in most analyses involving a single data source; Thorson et al., 2013), and an SVC is estimated for each non-reference data source which has the effect of substantially downweighting the influence of the non-reference data sources on model predictions (see Figures 4 and 8 in Grüss et al. (2023d)). SVC spatio-temporals models enable catchability coefficients for the non-reference data sources to vary over space, i.e., allow the response curve for the catchability of the non-reference data sources to change smoothly across the modelled region (Thorson et al., 2023). By expressing the fishing-power ratio of non-reference data sources as a function of unmeasured (latent) spatial variables and processes, SVCs enable spatio-

temporal models to attribute some of the residual variance of the spatio-temporal models to these latent spatial variables and processes; consequently, SVCs can enhance the fit of spatio-temporal models to data and strengthen statistical inferences (Finley, 2011; Rollinson et al., 2021; Thorson et al., 2023). We can then posit that SVC integrated spatio-temporal models have a better predictive performance than standard integrated spatio-temporal models, although this assumption has never been evaluated. Among existing SVC integrated spatio-temporal models, that of Olmos et al. (2023), which uses SVC to integrate monitoring data collected in different seasons, is worth highlighting. Olmos et al. (2023) used SVC to represent seasonality, thus addressing the problem posed by the spatial imbalance of many monitoring and fisheries datasets across seasons to successfully estimate a seasonal spatio-temporal model (Thorson et al., 2020a).

Expanded-domain integrated spatio-temporal models constitute the third type of integrated spatio-temporal models. These models involve the integration of datasets for different areas, which is particularly useful for tracking spatio-temporal shifts in marine species across countries' exclusive economic zones or across management areas within the waters of a given country (Ono et al., 2018; Maureaud et al., 2021; O'Leary et al., 2022). Existing expanded-domain integrated spatio-temporal models have achieved data integration through the estimation of a non-spatial catchability term for the non-reference data sources (e.g., a catchability factor treated as a random effect in Maureaud et al. (2021)). However, it is also feasible to carry out data integration in an expanded-domain integrated spatio-temporal model by estimating an SVC for the non-reference data sources. Hence, as mentioned earlier, we can distinguish between three different sub-types (categories) of expanded-domain integrated spatio-temporal models: (1) a first category where the data that are integrated come from the same source but were collected within different areas; (2) a second category consisting of the combination of expanded-domain integrated spatio-temporal modelling

(because data come from multiple areas) and standard integrated spatio-temporal modelling (because data integration is achieved through the estimation of a non-spatial catchability term for the non-reference data sources), like in Maureaud et al. (2021); and (3) a third category consisting of the combination of expanded-domain integrated spatio-temporal modelling and SVC integrated spatio-temporal modelling, i.e., like what was achieved in Maureaud et al. (2021) but with the non-spatial catchability term for the non-reference data sources substituted with an SVC term (Table 1).

While integrated spatio-temporal modelling is increasingly being considered and employed worldwide, the above-mentioned types of data integration, i.e., standard, SVC and expended-domain integrated modelling, have never been evaluated through a comparative analysis. Here, we address this research gap through an investigation of the impacts of these different types of data integration on the predictions of spatio-temporal models, via: (1) an application to the southern hake (*Merluccius australis*) HAK4 stock, where the bottom trawl data collected within the New Zealand observer programme are integrated with the data from five different bottom trawl research surveys; and (2) a simulation experiment. Our modelling work involves six alternative spatio-temporal models, including a non-integrated model and the five sub-types of integrated models described earlier and in Table 1: (1) a spatio-temporal model fitted to one single data source (observer-only data in our case); (2) a standard integrated spatio-temporal model; (3) an SVC integrated spatio-temporal model; and (4) the three categories of expanded-domain integrated spatio-temporal models, including (i) an expanded-domain integrated spatio-temporal model fitted to data from the same source (bottom trawl data collated within the New Zealand observer programme in our case) but collected within different areas, (ii) a spatio-temporal model combining expanded-domain and standard integrated modelling, and (iii) a spatio-temporal model combining expanded-domain and SVC integrated modelling.

2. Material and methods

2.1. Spatio-temporal modelling

In fisheries science, spatio-temporal models tend to be fitted to catch rate data, with the primary goal of producing indices of relative abundance to support stock assessments (Hodgdon et al., 2020; O'Leary et al., 2020; Thorson et al., 2020b; Adams et al., 2021). When spatio-temporal models are fitted to data coming from one single data source, they typically express catch rate d at each site *s* and in each time period *t* (assumed to be a given year henceforth) in the link space as follows (Thorson, 2022)[1]:

$$g\big(d(s_i, t_i)\big) = \beta(t_i) + \omega(s_i) + \varepsilon(s_i, t_i) + \sum_{k=1}^{n_k} \lambda(k) Q(i, k) \quad (1)$$

where *i* indexes samples; s_i is the site where sample *i* was collected; t_i is the year in which sample *i* was collected; *g*() is a link function; $\beta(t_i)$ is the intercept in year t_i; $\omega(s_i)$ is spatial variation at site s_i; $\varepsilon(s_i, t_i)$ is spatio-temporal variation at site s_i in year t_i; k indexes catchability covariates, i.e., covariates affecting the process of obtaining measurements; $Q(i, k)$ is an element of the matrix $\mathbf{Q}$ composed of n_k covariates explaining variation in catchability; and $\lambda(k)$ is the estimated impact of catchability covariate *k*. This formulation does not include any density covariates (e.g., environmental covariates influencing density such as sea surface temperature or bottom depth); however, future implementations of our spatio-temporal models could make use of density covariates (see Section 4). The year intercepts β tend to be treated as fixed effects in spatio-temporal models (Thorson, 2019), which is what we also assume here, whereas the spatial variation term ω and the spatio-temporal variation term ε are both treated as random effects.

[1] In the present study we use mathematical notation similar to the C++ code used to define VAST models in the Template Model Builder (TMB). The notation we employ is close to common recommendations (e.g., from Edwards and Auger-Méthé (2019)) and is typically adopted in peer-reviewed publications that report VAST modelling results.

The spatial and spatio-temporal variation terms are the core terms of spatio-temporal models. These terms account for, respectively, multiple latent static variables and multiple latent dynamic variables that have an influence on densities (Shelton et al., 2014; Thorson et al., 2015; Grüss et al., 2023a; Dovers et al., 2024). Both the spatial variation term ω and the spatio-temporal variation term ε are modelled as Gaussian Markov random fields following a multivariate normal (MVN) distribution. In addition, because the spatial distribution of sampling can change substantially across years, the spatio-temporal variation term ε is often modelled as a first-order autoregressive (AR1) process in spatio-temporal models (e.g., Charsley et al., 2023; Grüss et al., 2023a, 2023c, 2024; Kawauchi et al., 2023):

$$\boldsymbol{\omega} \sim MVN\left(\mathbf{0}, \sigma_{\omega}^{2}\mathbf{R}(\kappa)\right) \tag{2}$$

$$\boldsymbol{\varepsilon}(t) \sim \begin{cases} MVN\left(\mathbf{0}, \sigma_{\varepsilon}^{2}\mathbf{R}(\kappa)\right) \; if \; t = t_{min} \\ MVN\left(\rho_{\varepsilon}\boldsymbol{\varepsilon}(t-1), \sigma_{\varepsilon}^{2}\mathbf{R}(\kappa)\right) \; if \; t > t_{min} \end{cases}$$

where $\boldsymbol{\omega}$ and $\boldsymbol{\varepsilon}(t)$ are vectors of length n_s, where n_s is the number of locations; $\mathbf{R}(\kappa)$ is the correlation among locations given an estimated decorrelation distance κ (Thorson et al., 2015); ρ_{ε} is first-order temporal autocorrelation, which analysts can set to 0 if it is not warranted to model an AR1 process; σ_{ω}^{2} is the estimated pointwise variance of spatial variation; σ_{ε}^{2} is the estimated pointwise variance of spatio-temporal variation; and t_{min} is the first year in the time series.

Catchability covariates (also called "detectability covariates") are particularly relevant to include in spatio-temporal models when working with fisheries-dependent monitoring data such as observer data (Wilberg et al., 2010; Grüss et al., 2019b, 2023b; Ducharme-Barth et al., 2022). However, the catchability of fisheries-dependent monitoring programs tends to be the result of numerous complex, and often interacting, causes, because fisheries-dependent monitoring programs rely on the strategies of fishing boats, which target specific places, times and species (Maunder and Punt, 2004; Ye and Dennis, 2009) and because fishers regularly

adapt their behaviours and gears to management changes, technological developments, and other constraints (Hilborn and Walters, 1992; Quinn and Deriso, 1999). We encapsulate catchability effects in a vessel effect term, $\eta(v_i)$, where v_i is the vessel with which sample i was collected. This vessel effect term represents multiple latent catchability variables not explicitly modelled (Thorson and Ward, 2014). Because differences in catchability among fishing boats are various, complex and typically poorly understood, many spatio-temporal modelling studies have employed the vessel effect term $\eta(v_i)$ in lieu of several explicit catchability covariates (e.g., Xu et al., 2019; Rufener et al., 2021; Grüss et al., 2023b, 2023c, 2023d). Under this premise, Equation 1 becomes:

$$g\big(d(s_i, t_i)\big) = \beta(t_i) + \omega(s_i) + \varepsilon(s_i, t_i) + \eta(v_i) \quad (3)$$

where vessel effects are modelled as random effects following a normal distribution with a mean of zero and a standard deviation estimated in the model (Bell et al., 2021; Ducharme-Barth et al., 2022; Grüss et al., 2023b).

The above assumed the use of one single data source. Let us now imagine that the analysts wish to employ two or more data sources in the same spatio-temporal model, e.g., both observer and research survey datasets. In this situation, relying on Equation 3 to produce statistical inferences would result in "data pooling" rather than data integration, where observations from different sources are used in the same model without any acknowledgment of the data sources (Fletcher et al., 2019; Isaac et al., 2020). In the present study, we do not consider the data pooling approach, although one sub-type (category) of expanded-domain integrated spatio-temporal models described in the following could also be considered as a form of data pooling as it does not involve the estimation of data source impacts (the first category of expanded-domain integrated models; see below). We distinguish between three types of integrated analysis in the present study: (1) standard integrated spatio-temporal

modelling; (2) SVC integrated spatio-temporal modelling; and (3) expanded-domain integrated spatio-temporal modelling.

The first and most common type of integrated spatio-temporal models are standard integrated spatio-temporal models (e.g., Grüss et al., 2018a, 2018b, 2020b; Grüss and Thorson, 2019; Gonzalez et al., 2021; Rufener et al., 2021; Alglave et al., 2022; Thompson et al., 2022; Nephin et al., 2023), which account for the differing catchabilities of the different data sources as follows:

$$g\big(d(s_i,t_i)\big) = \beta(t_i) + \omega(s_i) + \varepsilon(s_i,t_i) + \eta(v_i) + \sum_{m=1}^{n_m} \delta(m)M(i,m) \quad (4)$$

where m indexes data sources; n_m is the number of data sources considered in the model; $\sum_{m=1}^{n_m} \delta(m)M(i,m)$ is the effect of data sources on expected catch rates; the design matrix $M(i,m)$ is 1 for the data source associated with sample *i* and 0 otherwise; and the data source impact $\delta(m)$ is set up so that $\delta(m) = 0$ for the reference data source, which is a necessary constraint for the identifiability of the year intercepts (Grüss and Thorson, 2019). When the data sources employed by the analyst include one single research survey, that research survey is identified as the reference data source. On the other hand, when the spatio-temporal model integrates several research surveys (among, potentially, other data sources), the research survey with the largest sample size is designated as the reference data source, unless the sampling design and/or protocols of that particularly survey have changed markedly over time compared to the other surveys available (Grüss et al., 2023a).

With standard integrated spatio-temporal models, a non-spatial catchability term $\delta(m)$ is estimated for each non-reference data source but not for the reference data source, while the vessel effect term $\eta(v_i)$ serves as an additional catchability term encompassing numerous latent variables not explicitly modelled that have an influence on catchability (Grüss et al., 2023a). As the reference data source impact is set to zero, the specified impact of data sources

allows for the estimation of fishing-power ratios relative to the reference data source for the non-reference data sources (Monnahan et al., 2021). As indicated earlier, the data source catchability factor in Equation 4 can be treated as a fixed or a random effect, but we assume a fixed effect to avoid the shrinking of model predictions towards a mean (this shrinking towards a mean is what is done with the the vessel effect term $\eta(v_i)$; Grüss et al., 2023a). In doing so, we do not implicitly assume that the different data sources have similar fishing power on average.

The second type of integrated spatio-temporal models are SVC integrated spatio-temporal models (e.g., Bolser et al., 2023; Grüss et al., 2023d; Olmos et al., 2023), which can be roughly described as the model presented in Equation 4 where the non-spatial data source catchability term for each non-reference data source is replaced with a spatially varying catchability (SVC) term. Thus, with SVC integrated spatio-temporal models, Equation 4 becomes (Grüss et al., 2023d):

$$g\big(d(s_i,t_i)\big) = \beta(t_i) + \omega(s_i) + \varepsilon(s_i,t_i) + \eta(v_i) + \sum_{m=1}^{n_m} \xi(s_i,m)M(i,m) \quad (5)$$

where $\xi(s_i,m)$ is the additive, zero-centred, spatially varying impact of data source m at location s_i, which is set to 0 for the reference data source and is estimated for the non-reference sources. The $\xi(s_i,m)$ term for the non-reference data sources is estimated as a random effect following an MVN distribution:

$$\boldsymbol{\xi}(m) \sim MVN\left(\mathbf{0}, \sigma_\xi^2(m)\mathbf{R}(\kappa)\right) \quad (6)$$

where $\sigma_\xi^2(m)$ is the estimated pointwise variance of the spatially varying response to non-reference data source m. Comparing Equations 6 and 2, we can see that the specification of the random spatially varying impacts of individual non-reference data sources is similar to the specification of the random spatial variation term. This specification results in the estimation of totally separate spatially varying impacts for each non-reference data source, thereby

avoiding the shrinking of model predictions towards a mean common to all non-reference data sources.

The third and last type of integrated spatio-temporal models are expanded-domain integrated spatio-temporal models (e.g., Ono et al., 2018; Maureaud et al., 2021; O'Leary et al., 2022). Their specificity is that they enable the integrated analysis of datasets collected within spatially distinct areas, either within distinct management areas in a country's exclusive economic zone (e.g., the eastern Bering Sea, Aleutian Islands and Gulf of Alaska in Ono et al. (2018)) or within the waters of different countries (e.g., the northern and eastern Bering Sea off Alaska and the Russian Pacific in O'Leary et al. (2022)). Data integration in expanded-domain integrated spatio-temporal models can involve the estimation of a non-spatial catchability term (Equation 4) or an SVC term (Equation 5) for the non-reference data sources. In these situations, spatio-temporal models combine, respectively, expanded-domain and standard integrated modelling and expanded-domain and SVC integrated modelling. In a third situation, data integration in expanded-domain integrated spatio-temporal models resembles data pooling. In this third situation, the datasets that are being integrated can be considered to be similar as they come from the same monitoring programme but from spatially distinct areas (e.g., observer data collected within the same programme but within distinct management areas of a country's economic exclusive zone).

In the present study, we provide the first-ever comparison of three above-mentioned types of integrated spatio-temporal models, i.e., standard, SVC and expanded-domain integrated spatio-temporal models. Our comparison involves an application to the southern hake HAK4 stock (described in Subsection 2.2) and a simulation experiment (described in Subsection 2.3).

2.2. Fishery application

418 *2.2.1. Southern hake and the HAK4 stock*

419 Southern hake (HAK) are large fish that are commonly encountered in deep waters of
420 New Zealand, almost entirely south of 40°S (Fig. 1; Anderson et al., 1998; O'Driscoll et al.,
421 2003). They support economically important commercial trawl fisheries and are caught both
422 as target species and as valuable bycatch by vessels targeting hoki (*Macruronus*
423 *novaezelandiae*), southern blue whiting (*Micromesistius australis*) and other species; hake are
424 virtually all caught by trawlers, approximately equally by mid-water trawl and bottom trawl
425 (Devine, 2009; Ballara, 2013, 2015, 2018).

426 Hake undertake ontogenetic habitat shifts (Horn, 2015). Hake young-of-the-years
427 inhabit inshore areas (<250-m depth) primarily off the central and northwest coasts of South
428 Island, while juvenile hake (1-2 years old) can be found at considerably deeper depths (up to
429 1000-m depth) generally north of 45°S (Horn, 2015; McGregor, 2021). Immature and adult
430 hake (2-7 years old) occur at depths greater than 250-350 m and less than 1400 m, but are
431 abundant at 800 m; compared to juvenile hake, immature and adult hake extend into deeper
432 waters and south into the waters south of 46°S (i.e., into the Sub-Antarctic), but also
433 (although in very small numbers) along the North Island southeast coast (Hurst et al., 2000;
434 Horn, 2015).

435 Since the introduction of the quota management system in New Zealand on 1st October
436 1986 (Lock and Leslie, 2007), hake has been managed through total allowable catches
437 allocated to four QMAs: (1) the HAK4 QMA, which covers the Chatham Rise and the East
438 Coast South Island (ECSI); (2) the HAK7 QMA, which encompasses the Challenger Plateau
439 and the fraction of the West Coast South Island (WCSI) located north of 46°S; (3) the HAK10
440 QMA, in the Kermadec area, where hake has never been recorded in catches and which will
441 be disregarded henceforth; and (4) the HAK1 QMA, which includes all of New Zealand's
442 exclusive economic zone not covered by the three other hake QMAs (Fig. 1a; Horn, 2015;

443 Dunn et al., 2021). The hake QMAs do not correspond with speculated hake biological stock
444 boundaries (Horn, 2015). The degree of mixing between the HAK4, HAK1 and HAK7 stocks,
445 as well as seasonal migration patterns in hake, are very poorly understood (Horn, 2015) and
446 cannot inform the present study.

447 In the present study, we focus on the HAK4 stock. However, our study region is wider
448 than the HAK4 QMA and covers the HAK4 and HAK7 QMAs and the fraction of the HAK1
449 QMA located south of 42.17°S, i.e., only the Sub-Antarctic part of the HAK1 QMA (Fig. 1b).
450 We consider this wider region, because some of the spatio-temporal models that we
451 developed for hake are expanded-domain integrated models fitted to data for the HAK4 QMA
452 combined with data for the HAK7 and HAK1 QMAs (see Subsection 2.2.3). Our study region
453 includes only the Sub-Antarctic part of the HAK1 QMA, as was the case in other hake
454 modelling studies (e.g., Dunn et al., 2021), because, within the HAK1 QMA, hake has rarely
455 been encountered outside of the Sub-Antarctic (Anderson et al., 1998; O'Driscoll et al.,
456 2003).

457

458 *2.2.2. Data for hake*

459 In the present study, we relied on bottom trawl biomass catch rate data for hake
460 (HAK) for the period 1991–2022, coming from: (1) the Fisheries New Zealand (FNZ)
461 database *cod*, which contains data collected by observers placed onboard commercial trawlers
462 (Sanders and Fisher, 2020); and (2) the FNZ database *trawl*, which compiles the data collated
463 within the research surveys taking place in New Zealand's exclusive economic zone (Mackay,
464 2020). We used only the *cod* data collected onboard bottom trawlers, hereafter referred to as
465 "COD BT data," in modelling analyses. Regarding research surveys, we employed data
466 collected within the following five bottom trawl surveys: (1) the Chatham Rise middle depth
467 trawl survey (CHAT MD survey); (2) the Sub-Antarctic summer middle depth trawl survey

(SUBA SUM survey); (3) the Sub-Antarctic autumn middle depth trawl survey (SUBA AUT survey); (4) the Southland middle depth survey (SOUTH MD survey); and (5) the WCSI Tangaroa middle depth survey (WCSI MD survey) (Table 1 and Fig. 1a). Three inshore research surveys are also conducted in our study region, namely the inshore trawl survey of the WCSI and Tasman/Golden Bays (WCSI TBGB survey) and the ECSI summer and winter inshore trawl surveys (ECSI SUM and ECSI WIN surveys). However, we disregarded the WCSI TBGB, ECSI SUM and ECSI WIN surveys in the present study, because they do not sample older hakes (2+ year-old individuals) and mainly catch 0-1 year old individuals (Horn, 2015). By contrast, the five above-mentioned middle depth surveys sample older hake like the COD BT monitoring programme (yet the five middle depth surveys seem to miss some of the younger ages, namely 3-5 year-old individuals, which may be found deeper) (Horn, 2015). Excluding the WCSI TBGB, ECSI SUM and ECSI WIN surveys from the present study was done to meet the assumption that the individuals from the different datasets integrated in a single model have the same demographic characteristics (Fithian et al., 2015; Pacifici et al., 2017; Grüss and Thorson, 2019; Miller et al., 2019; Grüss et al., 2023d).

The *trawl* database gathers data collated using various protocols and designs and in different areas of New Zealand, sometimes for very distinct purposes (Mackay, 2020). Moreover, not all the observations compiled in the *trawl* database are valid for biomass estimation, because, for example, of poor gear performance. Therefore, it is necessary to carefully select the data from the *trawl* database that go into modelling analyses. To address this issue, we employed the enhanced version of the *trawl* database presented in Grüss et al. (2023a). We extracted the biomass catch rate data ($kg \cdot km^{-2}$) collected in five above-mentioned surveys (the CHAT MD, SUBA SUM, SUBA AUT, SOUTH MD and WCSI MD surveys), which we cleaned via procedures that are standard in New Zealand (Morrison et al., 2013; Grüss et al., 2023a, 2023d; Ballara and O'Driscoll, 2024). For the period 1991–2022,

493 our cleaned survey dataset for the HAK4 QMA (including only CHAT MD survey data)
494 counted a total of 3117 records, while our cleaned survey dataset for HAK encompassing the
495 HAK4, HAK7 and HAK1 QMAs (including CHAT MD, SUBA SUM, SUBA AUT, SOUTH
496 MD and WCSI MD survey data) counted a total of 6052 records (Fig. A1). The high-quality
497 CHAT MD survey was the survey with the largest sample size and was, therefore, nominated
498 as the reference data source in all the integrated spatio-temporal models that we developed for
499 hake.

500 The *cod* database consists of data that have been collected onboard commercial
501 trawlers by observers since 1986 (Sanders and Fisher, 2020). Although the *cod* database also
502 contains age, length and biological information obtained by observers, we used only the catch
503 and effort (confidential) data collated onboard bottom trawlers that are compiled in the *cod*
504 database. We cleaned the data that we extracted from the *cod* database through procedures
505 that are standard in New Zealand (Edwards and Mormede, 2023; Grüss et al., 2023a, 2023d).
506 We calculated an area swept (km^2) as the product of doorspread (km) and distance trawled
507 (km) and, then, biomass catch rate ($kg \cdot km^{-2}$) by dividing the catch (kg) by the area swept.
508 For the period 1991–2022, our cleaned observer dataset for the HAK4 QMA counted a total
509 of 53 783 records, whereas our cleaned observer dataset for HAK encompassing the HAK4,
510 HAK7 and HAK1 QMAs counted a total of 134 177 records (Fig. A1).

511

512 *2.2.3. Spatio-temporal modelling for hake*

513 Prior to fitting spatio-temporal models, we constructed a prediction grid, following the
514 methodology of Grüss et al. (2018c, 2023a). First, we produced a 10 km × 10 km spatial grid
515 for New Zealand's exclusive economic zone. Then, using all the encounter data for hake
516 (HAK) from the original, larger *trawl* and *cod* databases (data collected with a bottom trawl
517 or other gears), we computed hake longitudinal, latitudinal and depth ranges. The depth data

518 for this step came from Mitchell et al. (2012). Finally, based on the computed longitudinal,
519 latitudinal and depth ranges of hake, we subsetted the 10 km × 10 km spatial grid for New
520 Zealand's exclusive economic zone, to come up with a 10 km × 10 km prediction grid for the
521 present study (Fig. 1b). The prediction grid was employed to determine the spatial distribution
522 of "knots" in the spatio-temporal models and generate spatial log-density maps from the
523 model outputs, and it was further subsetted to construct indices for the HAK4 stock (see
524 below).

525 We developed six different spatio-temporal models and compared the predictions of
526 the six models. Fisheries-dependent data are the major source of information for the great
527 majority of stocks in New Zealand including hake stocks (Dunn et al., 2021; Grüss et al.,
528 2023b; Ballara and O'Driscoll, 2024) and our primary interest was to understand whether and
529 how the integration of fisheries-dependent data with survey data could benefit New Zealand
530 stock and habitat assessments, using HAK4 as a case study. As such, the first of the six
531 models that we estimated was a spatio-temporal model fitted to observer-only data for the
532 HAK4 QMA (Model 1). After the observer-only data model for HAK4 was fitted, we
533 developed five other spatio-temporal models, including: (Model 2) a standard integrated
534 model fitted to both observer and survey data for HAK4 (where the survey data came only
535 from the CHAT MD survey); (Model 3) an SVC integrated model fitted to both observer and
536 survey data for HAK4 (where, again, the survey data came only from the CHAT MD survey);
537 and (Models 4–6) three different expanded-domain integrated models fitted to data for HAK,
538 i.e., data collected within not only the HAK4 QMA but also within the HAK7 and HAK1
539 QMAs (where the survey data came from the CHAT MD, SUBA SUM, SUBA AUT,
540 SOUTH MD and WCSI MD surveys). The first of the three expanded-domain integrated
541 models (Model 4) was fitted only to observer data, but these observer data were not only for
542 the HAK4 QMA but also for the HAK7 and HAK1 QMAs. The two other expanded-domain

543 integrated spatio-temporal models consisted of fitting a standard integrated model (Model 5)
544 or an SVC integrated model (Model 6) to both observer and survey data that were not only for
545 the HAK4 QMA but also for the HAK7 and HAK1 QMAs. In summary, we developed: (1)
546 three models for HAK4, including an observer-only data model (Model 1), a standard
547 integrated model (Model 2), and an SVC integrated model (Model 3); and (2) three expanded-
548 domain integrated models for the HAK population (i.e., encompassing the HAK4, HAK7 and
549 HAK1 stocks), including an observer-only data model (Model 4), a standard integrated model
550 (Model 5), and an SVC integrated model (Model 6).

551 We implemented the six spatio-temporal models with version 3.10.0 of the vector
552 autoregressive spatio-temporal (VAST) modelling platform (Thorson, 2019), whose R code
553 and documentation are available publicly at https://github.com/James-Thorson-NOAA/VAST.
554 Given that all the datasets fed into the models included many zeros, all six models were two-
555 stage delta models combining together the probabilities of encounter estimated by a first
556 linear predictor and the positive catch rates estimated by a second linear predictor (Lo et al.,
557 1992; Stefánsson, 1996). Specifically, all six models were "Poisson-link delta models" as
558 described in detail in Thorson (2018) and Grüss et al. (2023a). In all standard and SVC
559 models, the catchability factor (Equation 4) or the SVC term (Equation 5) was included only
560 in the first linear predictor (Grüss and Thorson, 2019; Grüss et al., 2023a, 2023d). In the
561 models fitted to both observer and survey data, the vessel effect η was not modelled (i.e., was
562 "turned off") for the survey data. In the six models, we modelled spatio-temporal variation as
563 an AR1 process and utilised the SPDE-Barrier model described in Grüss et al. (2023a) to
564 estimate covariances among sites, because of the presence of convoluted coastlines and
565 islands in our study region. Moreover, in all models, we distributed $n_x = 200$ knots
566 uniformly over the 10 km × 10 km prediction grid, and densities were predicted across 2000
567 grid cells covering the prediction grid (Grüss et al., 2020a). We confirmed that the parameter

568 estimates and predictions of the model were qualitatively similar when we specified more
569 than 200 knots. After the different spatio-temporal models were fitted, we ascertained that
570 they had successfully converged by checking that the gradient of the marginal log-likelihood
571 was less than 0.0001 for all fixed effects and that the Hessian matrix of secondary derivatives
572 of the negative-likelihood was positive definite. Then, we evaluated the models by employing
573 a procedure relying on R package *DHARMa* (Hartig, 2020) which is standard for spatio-
574 temporal models implemented with the VAST modelling platform. More precisely, we
575 considered simulation-based probability-integral-transform (PIT) residuals (Smith, 1985;
576 Warton et al., 2017) via Q-Q plots; we utilised R package *DHARMa* to simulate datasets from
577 the predictive distribution of the data conditioned on estimated fixed and random effects, and
578 we then calculated the PIT residuals from the observed and simulated values to create
579 diagnostic objects for the simulation-based quantile residuals from a delta model. We
580 confirmed that all models passed the evaluation tests (Appendix A2).

581 We compared the capabilities of the six spatio-temporal models to support habitat and
582 stock assessments. Regarding habitat assessments, we first analysed the following outputs of
583 the six models: (1) mean spatial patterns of log-density for hake; and (2) the coefficients of
584 variation (CVs) of mean spatial log-densities, which were computed from the standard errors
585 (SEs) associated with the mean spatial patterns of log-density. To obtain SEs, we: (1) drew
586 1000 samples from the predictive distributions by sampling from the joint precision of the
587 fixed and random effects of the spatio-temporal models; and (2) calculated SEs for log-
588 density estimates from the 1000 samples (Goodman et al., 2022; Grüss et al., 2023c). For the
589 three expanded-domain integrated models, we also examined: (1) predicted annual eastward
590 and northward centres of gravity (COGs) for the HAK population (encompassing the HAK4,
591 HAK1 and HAK7 stocks), which shed light on density shift patterns (Thorson et al., 2016a,
592 2016b); (2) the CVs of predicted eastward and northward COGs; (3) the annual effective

areas occupied estimated for the HAK4 stock and its adjacent HAK1 stock by the three expanded-domain integrated models, where estimated changes in effective area occupied are informative about patterns of range expansion/contraction in populations (Thorson et al., 2016a; Han et al., 2021); and (4) the CVs of these effective areas occupied. We applied the Mann-Kendall test for trend (Mann, 1945; Kendall, 1975; Gilbert, 1987) to the eastward and northward COG and effective area occupied estimates, and checked whether the *p*-value associated with the Mann-Kendall test was statistically significant at the $\alpha = 0.05$ level.

To compare the capabilities of the six spatio-temporal models to support stock assessments, we looked at: (1) the indices of relative biomass (henceforth simply "indices") estimated for the area encompassed by the CHAT MD survey, i.e., the Chatham Rise middle depth area; (2) the CVs of these indices; and (3) the interannual variability of these indices. We configured the spatio-temporal models in VAST so as to be able to retrieve index values for each site *s* in each year *t* and their associated SEs (Thorson, 2022). To obtain the indices for the Chatham Rise middle depth area, we subsetted the prediction grid to retain only the index values and their associated SEs that corresponded to the Chatham Rise middle depth area (Grüss et al., 2023a). To calculate the SEs around indices (and around COGs and effective areas occupied), we relied on the generalised delta method (Kass and Steffey, 1989) implemented in the Template Model Builder (TMB; Kristensen et al., 2016) that is leveraged by the VAST modelling platform (see Appendix A2). The interannual variability of the index, $V(t)$, was calculated as:

$$V(t) = \frac{1}{(n_t - 1)} \sum_{t}^{n_t - 1} \left[\hat{I}(t+1) - \hat{I}(t)\right] \quad (7)$$

where n_t is the number of years in the study period (32); and $\hat{I}(t)$ is the estimated index in year t.

We compared the indices estimated by the six spatio-temporal models fitted in the present study to the traditional stratified random index computed from the CHAT MD survey data with the *SurvCalc* software, henceforth referred to as the "SurvCalc index" (Francis, 2009). The SurvCalc index that we employed in the present study came from Stevens et al. (2023). The CHAT MD survey sampling design changed in 2010 and then again in 2016, with deeper 800-1300 m strata added to the sampling design. For this reason, we used SurvCalc index for the core 200-800 m strata, as these strata have been surveyed consistently in all years of our study period (1991–2022) (Stevens et al., 2023). Comparing the indices estimated by catch rate standardisation models to the SurvCalc indices is standard practice in New Zealand (Grüss et al., 2023a, 2023d). The SurvCalc index for the CHAT MD survey for the core 200-800 m strata is considered to be a good index to gauge the indices estimated by the spatio-temporal models, because this SurvCalc index is produced from fisheries-independent-only data collected by well-designed and consistent research monitoring within the core 200-800 m strata (Grüss et al., 2023d).

2.3. Simulation experiment

In addition to the application for hake, we conducted a simulation experiment where the operating model (OM) was the standard integrated spatio-temporal model for the HAK population (encompassing the HAK4, HAK1 and HAK7 stocks), i.e., the spatio-temporal model combining standard and expanded-domain integration. The goal of the simulation experiment was to further assess the capability of the six forms of spatio-temporal modelling described in Subsection 2.2.3 to support stock assessments. The simulation experiment consisted of: (1) utilising a simulator, which produces new values of random effects conditional on the maximum likelihood estimates (MLEs) for the fixed effects estimated by the OM and then generates simulated data conditional upon fixed and random effect values;

640 (2) considering that the indices derived from simulated data represent "true indices;" (3)
641 fitting alternative estimation models (EMs) to the simulated data, resulting in alternative
642 estimated indices; and (4) comparing the estimated indices (obtained from the EMs) to the
643 true indices (obtained from the simulated data) using performance metrics to evaluate the
644 bias, error and confidence interval coverage of the indices estimated by the alternative EMs.
645 The alternative EMs were the six spatio-temporal models from Subsection 2.2.3, namely:
646 (Model 1) the model fitted to observer-only data for HAK4; (Model 2) the standard integrated
647 model fitted to both observer and survey data for HAK4; (Model 3) the SVC integrated model
648 fitted to both observer and survey data for HAK4; (Model 4) the expanded-domain integrated
649 model fitted to observer-only data for HAK (encompassing the HAK4, HAK1 and HAK7
650 stocks); (Model 5) the standard integrated model fitted to both observer and survey data for
651 HAK, i.e., the model combining expanded-domain and standard integrated modelling; and
652 (Model 6) the SVC integrated model fitted to both observer and survey data for HAK, i.e., the
653 model combining expanded-domain and SVC integrated modelling.
654 We carried out the simulation experiment by employing the "self-test simulator" that
655 is included in R package *VAST* (Thorson, 2019), which involved six alternative scenarios
656 corresponding to the six spatio-temporal models described in the previous paragraph and 100
657 simulations for each of the six scenarios. The estimated indices (obtained from the EMs) were
658 for the HAK4 stock (i.e., corresponded to the HAK4 QMA) and were compared to the true
659 indices (obtained from the OM) for that stock. Because of the high computation costs of our
660 simulation experiment, we implemented it using the New Zealand eScience Infrastructure
661 (NeSI) high-performance computing facilities (http://www.nesi.org.nz/).
662 The following metrics (detailed, e.g., in the appendices of Grüss and Thorson (2019))
663 were computed to gauge the performance of the six alternative EMs relative to one another:
664 (1) a bias metric that indicates if changes in the true index are estimated accurately (the closer

665 to 1 the better; Thorson et al., 2015); (2) root mean squared error (the lower the better; Stow
666 et al., 2009; Norberg et al., 2019); and (3) coverage for a 50% confidence interval (henceforth
667 usually "coverage"; Agresti and Coull, 1998; Newcombe, 1998; Brown et al., 2001).
668 Coverage is defined as the percentage of time periods (years in our case) that the 50%
669 confidence interval of an index estimated by a model contains the "true" index; if the
670 coverage values are smaller (greater) than 50%, then confidence intervals are too narrow (too
671 wide) (Bolker, 2008; Johnson et al., 2019). The bias metric we considered was the coefficient
672 b of the following linear model (Thorson et al., 2015):

$$\begin{aligned} \hat{I}(t) &= a + b \times I(t) \\ E(t) &\sim Normal(0, \sigma_E^2) \end{aligned} \qquad (8)$$

673 where a is an intercept; $I(t)$ is the "true" index in year *t*; $E(t)$ is the "estimation error" in
674 $\hat{I}(t)$; and σ_E^2 is the variance of E. A b of 1 is indicative that changes in the true index are
675 reflected accurately by the estimated index, while a b greater than 1 (lower than 1) indicates
676 that the estimated index $\hat{I}(t)$ underestimates (overestimates) changes in the true index $I(t)$
677 (Wilberg et al., 2010; Thorson et al., 2015).
678
679 **3. Results**
680 The six spatio-temporal models (Models 1–6) that we estimated for hake predicted
681 similar density patterns in the HAK4 QMA, particularly higher densities on the northern
682 Chatham Rise (Fig. 2). The expanded-domain integration models (Models 4–6) also allowed
683 for insights into the spatial density patterns of hake beyond the HAK4 QMA, i.e., in the
684 HAK1 and HAK7 QMAs as well. All expanded-domain integration models predicted that the
685 highest hake densities are located off the WCSI, primarily between 42°S and 43°S (Fig. 2).
686 Looking at the CVs associated with estimated spatial density patterns suggests that: (1)
687 the integrated models (Models 2–6) predicted more precise densities in the HAK4 QMA
688 compared to the model fitted to observer-only data for the HAK4 QMA only (Model 1); (2)

689 the SVC integrated models (Models 3 and 6) predicted more precise densities in the HAK4
690 QMA than the standard integrated models (Models 2 and 5); and (3) the expanded-domain
691 integrated models (Models 4–6) predicted more precise densities in the HAK4 QMA
692 compared to the other models (Models 1–3) (Fig. 3a). These results become clearer when
693 comparing the mean CV of density in the HAK4 QMA across all years for the six different
694 models (Fig. 3b). We note that the CVs of estimated densities are highest in the southernmost
695 areas of the HAK4 QMA (Fig. 3a), because there are limited monitoring data for hake south
696 of 45°S and east of 172°E (Figs. 1a and b); therefore, densities are inferred for these areas
697 primarily through interpolation and/or extrapolation.

698 All expanded-domain integrated models (Models 4–6) provided similar insights into
699 annual changes in eastward and northward COGs, i.e., similar insights into density shifts for
700 hake (Fig. 4). Specifically, all expanded-domain integrated models predicted a statistically
701 significant northwestern shift in the COG of the HAK population over time. Looking at the
702 95% confidence intervals of the annual COG estimates suggests that: (1) the standard and
703 SVC integrated models (Models 5 and 6, which were fitted to both observer and survey data)
704 yielded more precise insights into patterns of density shifts than the model fitted to observer-
705 only data (Model 4); and (2) the predicted changes in eastward COGs were less uncertain with
706 the standard integrated model (Model 5) than with the SVC integrated model (Model 6), while
707 the opposite was true for the predicted changes in northward COGs (Fig. 4). These results
708 become clearer when comparing the mean CVs of the eastward and northward COGs across
709 all years for three different expanded-domain integrated models (Figs. 5a and b).

710 All expanded-domain integrated models (Models 4–6) provided similar insights into
711 changes in effective area occupied for HAK4 and HAK1, i.e., similar insights into patterns of
712 range expansion/contraction for the two stocks (Fig. 6). Specifically, all expanded-domain
713 integrated models predicted virtually no changes in the effective area occupied of HAK4

between 1991 and 2022, but a statistically significant decrease in the effective area occupied of HAK1 over time (Fig. 6). Looking at the 95% confidence intervals of the annual effective area occupied estimates suggests that: (1) the standard and SVC integrated spatio-temporal models (Models 5 and 6, which are fitted to both observer and survey data) yielded more precise insights into patterns of range expansion/contraction than the model fitted to observer-only data (Model 4); and (2) the predicted changes in the effective area occupied of the HAK4 stock were less uncertain with the standard integrated model (Model 5) than with the SVC integrated model (Model 6), while the opposite was true for the predicted changes in the effective area occupied of the HAK1 stock (Fig. 6). These results become clearer when comparing the mean CVs of the effective area occupied estimates across all years for three different expanded-domain integrated models (Figs. 5c and d).

All six models predicted similar trends in relative biomass, with an overall decrease in the index between 1992 and 2020 and increase in 2021–2022 (Figs. 7 and 8). Results indicated that: (1) the indices obtained with integrated models (Models 2–6) were more precise than the index obtained with the model fitted to observer-only data from the HAK4 QMA (Model 1); (2) the indices estimated by the standard integrated models (Models 2 and 5) and the SVC integrated models (Models 3 and 6) had the same level of uncertainty; and (3) the indices obtained with the expanded-domain integrated models (Models 4–6) were more precise than those obtained with the models fitted to data from the HAK4 QMA only (Models 1–3) (Figs. 7 and 9a). In terms of interannual variability in relative biomass: (1) the index estimated by the model fitted to data for the HAK population (encompassing the HAK4, HAK7 and HAK1 stocks; Model 4) had, by far, the largest interannual variability; (2) the SVC integrated models (Models 3 and 6) produced indices with less interannual variability than the standard integrated models (Models 2 and 5); and (3) the indices obtained with the expanded-domain integrated models (Models 4–6) exhibited more interannual variability than

739 the indices obtained with the models fitted to data from the HAK4 QMA only (Models 1–3)
740 (Figs. 7 and 9b).

741 We also compared the indices estimated by the six models with the SurvCalc index,
742 which revealed that: (1) the indices obtained with the six models were all strongly (Pearson's
743 correlation coefficients: 0.5–0.7) to very strongly correlated (Pearson's correlation
744 coefficients: ≥0.7) with the SurvCalc index; (2) the indices produced by the SVC integrated
745 models (Models 3 and 6) were more strongly correlated with the SurvCalc index than the
746 indices produced by the standard integrated models (Models 2 and 5); and (3) the correlation
747 with the SurvCalc index was stronger for the models fitted to data from the HAK4 QMA only
748 (Models 1–3) than for the expanded-domain integrated models (Models 4–6) (Fig. 8).

749 The simulation experiment revealed that: (1) the integrated models (Models 2–6) had
750 less bias and less error than the model fitted to observer-only data from the HAK4 QMA
751 (Model 1), but smaller coverage (i.e., worse coverage); (2) the standard integrated models
752 (Models 2 and 5) and the SVC integrated models (Models 3 and 6) had similar performance in
753 terms of bias, error and coverage; and (3) expanded-domain integrated models (Models 4–6)
754 had less error and less bias, but smaller coverage (i.e., worse coverage) than the models fitted
755 to data from the HAK4 QMA only (Models 1–3) (Fig. 10).

756

757 **4. Discussion**

758 Integrated modelling, using IPMs, ISAMs and ISDMs, is gaining increasing traction
759 worldwide (Grüss et al., 2023a; Schaub et al., 2024; Punt et al., 2020). In this context, we
760 provided in the present study a terminology of spatio-temporal ISDMs and the first-ever
761 comparison of standard, SVC and expanded-domain integrated models implemented with the
762 widely employed VAST modelling platform (used in around 130 refereed publications as of
763 April 2024; https://github.com/James-Thorson-NOAA/VAST).

764 Our results indicated that standard, SVC and expanded-domain integrated spatio-
765 temporal models outperformed a model fitted to one single data source in pretty much all
766 respects, consistent with previous research (e.g., Pagel et al., 2014; Fithian et al., 2015; Grüss
767 and Thorson, 2019; O'Leary et al., 2020; Pinto et al., 2019; Robinson et al., 2020; Monnahan
768 et al., 2021; Rufener et al., 2021; Thompson et al., 2022; Grüss et al., 2023d; Dovers et al.,
769 2024; Mäkinen et al., 2024). Specifically, compared to the observer-only data model, our
770 integrated models that combined observer and survey data delivered: (1) more precise spatial
771 density predictions; (2) more precise insights into patterns of density shifts and range
772 expansion/contraction; and (3) indices that were more precise, better matched the traditional
773 stratified random index, and had less bias and less error. However, our integrated models
774 characterised uncertainty around the estimated indices less well than the observer-only data
775 model. Similarly, Grüss et al. (2023a) found that their standard integrated spatio-temporal
776 models had worse coverage for a 50% confidence interval than models fitted to survey-only
777 data. These findings highlight that integrated models are not without potential issues and
778 should, therefore, be compared to models fitted to single data sources in individual
779 applications (Isaac et al., 2020; Simmonds et al., 2020).

780 Our findings also suggested that, overall, SVC integrated spatio-temporal models
781 performed better than standard integrated spatio-temporal models in support of habitat and
782 stock assessments. First, our results indicated that SVC spatio-temporal models predicted
783 more precise spatial densities than standard integrated spatio-temporal models. Second, we
784 found that the indices obtained with SVC spatio-temporal models were more strongly
785 correlated with the traditional stratified random index. Otherwise, the SVC and standard
786 integrated spatio-temporal models had similar performance in terms of uncertainty in
787 predicted patterns of density shifts and range expansion/contraction and uncertainty around
788 the estimated indices, as well as in terms of bias, error and coverage. By expressing the

789 fishing-power ratio of non-reference data sources as a function of multiple latent spatial
790 variables and processes, spatially varying catchabilities enable spatio-temporal models to
791 attribute some of the residual variance to these latent spatial variables and processes, thereby
792 potentially improving model predictive performance (Finley, 2011; Rollinson et al., 2021).
793 The present study confirmed that modelling spatially varying catchabilities do result in
794 enhanced spatio-temporal modelling predictive performance so that SVC integrated spatio-
795 temporal models outperform standard integrated spatio-temporal models in several respects.

796 Whether expanded-domain integrated models should be chosen among the different
797 types of integrated spatio-temporal models appeared to depend on whether model outputs aim
798 to support habitat or stock assessments. Our findings suggested that spatio-temporal models
799 combining standard and expanded-domain integrated modelling or SVC and expanded-
800 domain integrated modelling should be preferred to assist habitat assessments, because: (1)
801 expanded-domain integrated models provided insights into spatial density patterns for much
802 larger regions and predicted these patterns more precisely for the area common to all the
803 models fitted in the present study; and (2) spatio-temporal models combining standard and
804 expanded-domain integrated modelling or SVC and expanded-domain integrated modelling
805 predicted patterns of distribution shifts and range expansion/contraction more precisely than
806 the expanded-domain integrated model fitted to observer-only data. The latter finding echoes
807 the results reported in Mäkinen et al. (2024) that models better predicted species' whole
808 ranges when employing data from multiple sources rather than from one single source.
809 Expanded-domain integrated spatio-temporal models predict spatial density patterns more
810 precisely because they rely on more data, coming from adjacent regions that are not
811 considered in standard or SVC integrated spatio-temporal models. Therefore, expanded-
812 domain integrated spatio-temporal models leverage more data when performing interpolations
813 and/or extrapolations in spatial areas where limited or no data are available.

By contrast, we found that expanded-domain integrated models did not outperform the other types of integrated models when the goal of the modelling analysis is to support stock assessments. Compared to the other types of integrated models, expanded-domain integrated models generated indices for a local area that were more precise and had less bias and less error, but that agreed less with the traditional stratified random index and characterised uncertainty less well (had a worse coverage for a 50% confidence interval). The lesser agreement with the traditional stratified random index is due to the fact that expanded-domain integrated models use data for the region of interest but also other regions to estimate indices, while the traditional stratified random index and the indices from the other types of integrated models are produced from data for the region of interest only. These results lead us to recommend indices for local areas be produced with spatio-temporal models using standard or SVC integrated modelling without combining it with expanded-domain integrated modelling. Moreover, we suspect that, in some applications, indices for local areas may be biased if data are borrowed from adjacent regions where rates of population change over time are very different. Therefore, we encourage future research to compare expanded-domain integrated modelling to standard and SVC integrated modelling under different scenarios, to determine the situations where standard and SVC integrated modelling should be strictly utilised to generate indices for local areas (as opposed to being merely preferred over expanded-domain integrated modelling).

Important to ISDM studies are the two assumptions that are worth repeating that: (1) the individuals from the different datasets have the same demographic characteristics; and (2) there is independence between the different data sources (Fithian et al., 2015; Pacifici et al., 2017; Grüss and Thorson, 2019; Miller et al., 2019; Grüss et al., 2023d). In our hake application, the exclusion of inshore survey datasets from the integrated models ensured that the first assumption was met. The second assumption was also met in the hake application as

all data sources (the observer programme and the five middle depth bottom trawl surveys) were independent from one another, being associated with distinct sites and/or times of the year. Regardless, it is interesting to note that Dovers et al. (2024) argued that, because Gaussian random fields in species distribution models (such as the ones representing spatial and spatio-temporal variation in our models) can be seen as capturing unmodelled covariates, they can also be seen as capturing dependence across the data sources caused by unmodelled covariates.

Because the focus of the present study was on different types of data integration, the spatio-temporal models developed here did not include any density covariates or any catchability covariates beyond the monitoring programme effects. Future studies could examine how the different types of spatio-temporal models perform when variables affecting density or catchability are also considered. It would be particularly valuable to explore how the inclusion of environmental covariates affecting density may benefit the different types of spatio-temporal models by improving their predictive performance outside of the areas sampled by any monitoring programme (Brodie et al., 2022; Nephin et al., 2023). In addition, given the increasing calls for investigations of climate change impacts on fishes and fisheries worldwide including in New Zealand (Mace, 2014; The Aotearoa Circle, 2020; Gerrard, 2021), developing integrated spatio-temporal models factoring in the effects of measured environmental variables is certainly a key avenue for future research. By explaining more of the variability in the data in the past through the sharing of information among sites, time and data sources, integrated spatio-temporal models may likely project the potential impacts of climate change with more certainty and more accuracy.

An obvious next step of integrated species distribution modelling research would be the comparison of different types of integrated spatio-temporal joint species distribution models (JSDMs). Spatio-temporal JSDMs are powerful tools that estimate spatial and spatio-

temporal variation but also associations among multiple species (Thorson et al., 2016a; Thorson and Barnett, 2017; Gao et al., 2020; Guan et al., 2020; Arimitsu et al., 2023). By sharing information across sites, time and data sources, but also among species, integrated spatio-temporal JSDMs may result in even more precise and more accurate estimates and, possibly, in a better characterisation of uncertainty around these estimates as well. We, therefore, encourage, future studies to produce a comparative analysis similar to the one reported in the present study but using spatio-temporal JSDMs in lieu of single-species integrated spatio-temporal models.

The spatio-temporal models reported in the present study did not account for preferential sampling, an issue often associated with unstructured data such as observer data (Pennino et al., 2019; Rufener et al., 2021; Alglave et al., 2022). Preferential sampling is the likely correlation between sampling locations and the response variable of species distribution models (Diggle et al., 2010; Conn et al., 2017). However, in preliminary analyses for the HAK7, we developed bivariate spatio-temporal models which analysed jointly catch rate data with gridded fishing intensity, where the correlation between hake density and fishing intensity was estimated, specifically the log-log effect of an increase of fishing intensity on an increased prediction of hake density (unpublished results). The estimated correlation between hake density and fishing intensity was very weak, which was expected because analyses of the data themselves suggested an absence of relationship between fishing intensity and hake catch rate (unpublished results). Consequently, we did not consider the issue of preferential sampling in the present study, but this issue was out of scope regardless. All this being said, it will be good for future spatio-temporal modelling studies using unstructured data potentially affected by sampling bias to carry out preliminary analyses to check for the existence of absence of a relationship between sampling intensity and catch rate (Olmos et al., 2023). We

888 also note that preferential sampling can be caused by factors other than fish density such as
889 distance to shore.

890 Although the intent of the present study was not to generate or verify knowledge for
891 New Zealand hake, our density maps and indices were in accordance with the results of
892 previous studies of the HAK population or the HAK4 stock. Thus, we found that, in New
893 Zealand, hake densities are highest on the WCSI, particularly between 42°S and 43°S, which
894 is in agreement with the spatial density patterns reported in Horn (2015). Moreover, our
895 predicted indices for the HAK4 stock matched the traditional stratified random index
896 produced from the CHAT MD survey data very well (Fig. 8; Stevens et al., 2023), as well as
897 the standardised catch-per-unit-effort indices for the period 1991–2019 from McGregor
898 (2021) and the biomass trends in 1991–2020 estimated in the latest HAK4 stock assessment
899 (Holmes, 2021). We encourage future studies to leverage and expand the spatio-temporal
900 models developed in the present study to re-evaluate hake biological stock boundaries and,
901 potentially recommend changes in the delineation of hake QMAs. It is unclear whether we
902 can consider that there is more than one hake stock in New Zealand waters, as analyses of
903 data on enzyme polymorphism and morphometric data were inconclusive (Smith et al., 1979;
904 J.A. Colman, NIWA, unpublished data), while some other analyses indicated there are either
905 three distinct biological stocks (analyses of age/length frequencies and year class strengths;
906 Horn, 1998, 1997, 2015; Dunn et al., 2021) or one stock (analyses of growth rates and ages at
907 maturity; Horn, 2015). To re-examine hake stock structure, not only spatio-temporal models
908 fitted to catch rate data will be needed, but spatio-temporal models fitted to catch-rate-at age
909 and/or catch-rate-at length data and length or weight data will be desirable as well (Lindegren
910 et al., 2022; Grüss et al., 2023b).

911 In conclusion, the present study concurs with previous research that all types of
912 integrated spatio-temporal models (standard, SVC and expanded-domain integrated models)

outperform models fitted to one single data source in pretty much all respects. The present study also suggests that expanded-domain integrated spatio-temporal models outperform standard and SVC integrated spatio-temporal models if the goal of modelling efforts is to support habitat assessments, but not if the goal is to assist stock assessments. We recommend standard or SVC integrated spatio-temporal models be preferred over other models to produce indices for local areas. The spatio-temporal modelling efforts reported in the present study could be continued to compare the performance of the different types of integrated spatio-temporal models when the goal of modelling efforts is to support climate-vulnerability assessments (Brodie et al., 2022; Liu et al., 2023), risk assessments (Grüss et al., 2024), or assessments of alternative sampling designs for particular monitoring programmes (Oyafuso et al., 2021, 2023).

Acknowledgments

Foremost, we express our gratitude to the scientists and the scientific observers that collected the data used in the present study and to Fisheries New Zealand (FNZ) for letting us employ these data, as well as to Pamela Mace (FNZ) and New Zealand's Statistics Assessments Methods Working Group (SAMWG) and Jennifer Devine and Matt Dunn (NIWA) for constructive discussions and feedback on our research. We also thank Brent Wood, Jeremy Yeoman, and Shaun Carswell (NIWA) very much for assisting with the FNZ research survey and observer databases, as well as the New Zealand eScience Infrastructure (NeSI) for the use of their high-performance computing resources. We are very grateful to two internal reviewers, Vidette McGregor (NIWA) and Andrew Allyn (Gulf of Maine Research Institute), as well as two anonymous journal reviewers, for their insightful comments on our manuscript. Financial support for this study was provided by NIWA Strategic Science Investment Funding and NIWA Structural Internal Project Funding.

Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the Department of Commerce.

DATA ACCESSIBILITY

The research survey data employed in the present study are the property of Fisheries New Zealand and can be obtained via an email to rdm.sharedrdm@mpi.govt.nz. The observer data used in the present study are confidential fisheries-dependent data that cannot be shared. The R codes associated with the present study can be found in the following GitHub repository: https://github.com/agruss2/Integrated-VAST-modelling.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version of the manuscript.

References

Abadi, F., Gimenez, O., Arlettaz, R., Schaub, M., 2010. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. Ecology 91, 7–14.

Abadi, F., Gimenez, O., Jakober, H., Stauber, W., Arlettaz, R., Schaub, M., 2012. Estimating the strength of density dependence in the presence of observation errors using integrated population models. Ecological Modelling 242, 1–9.

Adams, C.F., Brooks, E.N., Legault, C.M., Barrett, M.A., Chevrier, D.F., 2021. Quota allocation for stocks that span multiple management zones: analysis with a vector autoregressive spatiotemporal model. Fisheries Management and Ecology 28, 417–427.

Agresti, A., Coull, B.A., 1998. Approximate is better than "exact" for interval estimation of binomial proportions. The American Statistician 52, 119–126.

Alglave, B., Rivot, E., Etienne, M.-P., Woillez, M., Thorson, J.T., Vermard, Y., 2022. Combining scientific survey and commercial catch data to map fish distribution. ICES Journal of Marine Science 79, 1133–1149.

Anderson, O.F., Bagley, N.W., Hurst, R.J., Francis, M.P., Clark, M.R., McMillan, P.J., 1998. Atlas of New Zealand fish and squid distributions from research bottom trawls. NIWA Technical Report 42. National Institute of Water and Atmospheric Research, Wellington. 303 p.

Arimitsu, M.L., Piatt, J.F., Thorson, J.T., Kuletz, K.J., Drew, G.S., Schoen, S.K., Cushing, D.A., Kroeger, C., Sydeman, W.J., 2023. Joint spatiotemporal models to predict seabird densities at sea. Frontiers in Marine Science 10, 1078042.

Ballara, S.L., 2018. Descriptive analysis of the fishery for hake (*Merluccius australis*) in HAK 1, 4 and 7 from 1989–90 to 2014–15, and a catch-per-unit-effort (CPUE) analysis for Chatham Rise and WCSI hake. New Zealand Fisheries Assessment Report 2018/55. 57 p.

Ballara, S.L., 2015. Descriptive analysis of the fishery for hake (*Merluccius australis*) in HAK 1, 4 and 7 from 1989–90 to 2012–13, and a catch-per-unit-effort (CPUE) analysis for Sub-Antarctic hake. New Zealand Fisheries Assessment Report 2015/12. 47 p.

Ballara, S.L., 2013. Descriptive analysis of the fishery for hake (*Merluccius australis*) in HAK 1, 4 and 7 from 1989-90 to 2010-11, and a catch-per-unit-effort (CPUE) analysis for Chatham Rise and WCSI hake. New Zealand Fisheries Assessment Report 2013/45. 82 p.

Ballara, S.L., O'Driscoll, R.L., 2024. Catches and size and age structure of the 2021–22 hoki fishery and a summary of input data used for the 2023 stock assessment. New Zealand Fisheries Assessment Report 2024/16. 184 p.

Bell, R.J., McManus, M.C., McNamee, J., Gartland, J., Galuardi, B., McGuire, C., 2021. Perspectives from the water: Utilizing fisher's observations to inform SNE/MA windowpane science and management. Fisheries Research 243, 106090.

Besbeas, P., Freeman, S.N., Morgan, B.J.T., Catchpole, E.A., 2002. Integrating Mark-Recapture-Recovery and Census Data to Estimate Animal Abundance and Demographic Parameters. Biometrics 58, 540–547.

Bolker, B.M., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, NJ.

Bolser, D.G., Berger, A.M., Chu, D., de Blois, S., Pohl, J., Thomas, R.E., Wallace, J., Hastie, J., Clemons, J., Ciannelli, L., 2023. Using age compositions derived from spatio-temporal models and acoustic data collected by uncrewed surface vessels to estimate Pacific hake (*Merluccius productus*) biomass-at-age. Frontiers in Marine Science 10, 1214798.

Bourdaud, P., Travers-Trolet, M., Vermard, Y., Cormon, X., Marchal, P., 2017. Inferring the annual, seasonal, and spatial distributions of marine species from complementary research and commercial vessels' catch rates. ICES Journal of Marine Science 74, 2415–2426.

Brodie, S., Smith, J.A., Muhling, B.A., Barnett, L.A., Carroll, G., Fiedler, P., Bograd, S.J., Hazen, E.L., Jacox, M.G., Andrews, K.S., 2022. Recommendations for quantifying and reducing uncertainty in climate projections of species distributions. Global Change Biology 28, 6586–6601.

Brodie, S.J., Thorson, J.T., Carroll, G., Hazen, E.L., Bograd, S., Haltuch, M.A., Holsman, K.K., Kotwicki, S., Samhouri, J.F., Willis-Norton, E., 2020. Trade-offs in covariate selection for species distribution models: a methodological comparison. Ecography 43, 11–24.

Brooks, S.P., King, R., Morgan, B.J.T., 2004. A Bayesian approach to combining animal abundance and demographic data. Animal Biodiversity and Conservation 27, 515–529.

Brown, L.D., Cai, T.T., DasGupta, A., 2001. Interval estimation for a binomial proportion. Statistical Science 16, 101–133.
Charsley, A.R., Grüss, A., Thorson, J.T., Rudd, M.B., Crow, S.K., David, B., Williams, E.K., Hoyle, S.D., 2023. Catchment-scale stream network spatio-temporal models, applied to the freshwater stages of a diadromous fish species, longfin eel (*Anguilla dieffenbachii*). Fisheries Research 259, 106583.
Conn, P.B., Thorson, J.T., Johnson, D.S., 2017. Confronting preferential sampling when analysing population distributions: diagnosis and model-based triage. Methods in Ecology and Evolution 8, 1535–1546.
Devine, J.A., 2009. Descriptive analysis of the commercial catch and effort data for New Zealand hake (*Merluccius australis*) for the 1989-90 to 2005-06 fishing years. New Zealand Fisheries Assessment Report 2009/21. 74 p.
Diggle, P.J., Menezes, R., Su, T., 2010. Geostatistical inference under preferential sampling. Journal of the Royal Statistical Society Series C: Applied Statistics 59, 191–232.
Doonan, I., Large, K., Dunn, A., Rasmussen, S., Marsh, C., Mormede, S., 2016. Casal2: New Zealand's integrated population modelling tool. Fisheries Research 183, 498–505.
Dorazio, R.M., 2014. Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. Global Ecology and Biogeography 23, 1472–1484.
Dovers, E., Popovic, G.C., Warton, D.I., 2024. A fast method for fitting integrated species distribution models. Methods in Ecology and Evolution 15, 191–203.
Ducharme-Barth, N.D., Grüss, A., Vincent, M.T., Kiyofuji, H., Aoki, Y., Pilling, G., Hampton, J., Thorson, J.T., 2022. Impacts of fisheries-dependent spatial sampling patterns on catch-per-unit-effort standardization: A simulation study and fishery application. Fisheries Research 246, 106169.
Dunn, A., Mormede, S., Webber, D.N., 2021. Descriptive analysis and stock assessment model inputs of hake (*Merluccius australis*) in the Sub-Antarctic (HAK 1) for the 2020-21 fishing year. New Zealand Fisheries Assessment Report 2021/74. 52 p.
Edwards, A.M., Auger-Méthé, M., 2019. Some guidance on using mathematical notation in ecology. Methods Ecol Evol 10, 92–99. https://doi.org/10.1111/2041-210X.13105
Edwards, C.T.T., Mormede, S., 2023. Temporal and spatial distribution of non-target catch and non-target catch species in deepwater fisheries. New Zealand Aquatic Environment and Biodiversity Report No. 303. 81 pp.
Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40, 677–697.
Ellner, S.P., Fieberg, J., 2003. Using PVA for management despite uncertainty: Effects of habitat, hatcheries, and harvest of salmon. Ecology 84, 1359–1369.
Finley, A.O., 2011. Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence. Methods in Ecology and Evolution 2, 143–154.
Fithian, W., Elith, J., Hastie, T., Keith, D.A., 2015. Bias correction in species distribution models: pooling survey and collection data for multiple species. Methods in Ecology and Evolution 6, 424–438.
Fletcher, R.J., Hefley, T.J., Robertson, E.P., Zuckerberg, B., McCleery, R.A., Dorazio, R.M., 2019. A practical guide for combining data to model species distributions. Ecology 100, e02710.
Francis, R.I.I.C., 2009. SurvCalc User Manual. NIWA, Wellington, New Zealand. 39 p.
Gao, J., Thorson, J.T., Szuwalski, C., Wang, H.-Y., 2020. Historical dynamics of the demersal fish community in the East and South China Seas. Marine and Freshwater Research 71, 1073–1085.

Gerrard, J., 2021. The Future of Commercial Fishing in Aotearoa New Zealand. A report from the Office of the Prime Minister's Chief Science Advisor, Kaitohutohu Mātanga Pūtaiao Matua ki te Pirimia. Wellington, New Zeland.
Gilbert, R.O., 1987. Statistical methods for environmental pollution monitoring. John Wiley & Sons, New York, NY.
Gonzalez, G.M., Wiff, R., Marshall, C.T., Cornulier, T., 2021. Estimating spatio-temporal distribution of fish and gear selectivity functions from pooled scientific survey and commercial fishing data. Fisheries Research 243, 106054.
Goodman, M.C., Carroll, G., Brodie, S., Grüss, A., Thorson, J.T., Kotwicki, S., Holsman, K., Selden, R.L., Hazen, E.L., De Leo, G.A., 2022. Shifting fish distributions impact predation intensity in a sub-Arctic ecosystem. Ecography 2022, e06084.
Grüss, A., Biggs, C., Heyman, W.D., Erisman, B., 2018a. Prioritizing monitoring and conservation efforts for fish spawning aggregations in the US Gulf of Mexico. Scientific reports 8, 1–10.
Grüss, A., Biggs, C.R., Heyman, W.D., Erisman, B., 2019a. Protecting juveniles, spawners or both: a practical statistical modelling approach for the design of marine protected areas. Journal of Applied Ecology 56, 2328–2339.
Grüss, A., Charsley, A.R., Thorson, J.T., Anderson, O.F., O'Driscoll, R.L., Wood, B., Breivik, O.N., O'Leary, C.A., 2023a. Integrating survey and observer data improves the predictions of New Zealand spatio-temporal models. ICES Journal of Marine Science 80, 1991–2007.
Grüss, A., Drexler, M.D., Ainsworth, C.H., Babcock, E.A., Tarnecki, J.H., Love, M.S., 2018b. Producing Distribution Maps for a Spatially-Explicit Ecosystem Model Using Large Monitoring and Environmental Databases and a Combination of Interpolation and Extrapolation. Frontiers in Marine Science 5, 16.
Grüss, A., Gao, J., Thorson, J.T., Rooper, C.N., Thompson, G., Boldt, J.L., Lauth, R., 2020a. Estimating synchronous changes in condition and density in eastern Bering Sea fishes. Marine Ecology Progress Series 635, 169–185.
Grüss, A., McKenzie, J.R., Lindegren, M., Bian, R., Hoyle, S.D., Devine, J.A., 2023b. Supporting a stock assessment with spatio-temporal models fitted to fisheries-dependent data. Fisheries Research 262, 106649.
Grüss, A., Moore, B.R., Pinkerton, M.H., Devine, J.A., 2023c. Understanding the spatio-temporal abundance patterns of the major bycatch species groups in the Ross Sea region Antarctic toothfish (*Dissostichus mawsoni*) fishery. Fisheries Research 262, 106647.
Grüss, A., Perryman, H.A., Babcock, E.A., Sagarese, S.R., Thorson, J.T., Ainsworth, C.H., Anderson, E.J., Brennan, K., Campbell, M.D., Christman, M.C., et al., 2018c. Monitoring programs of the US Gulf of Mexico: inventory, development and use of a large monitoring database to map fish and invertebrate spatial distributions. Reviews in Fish Biology and Fisheries 28, 667–691.
Grüss, A., Rose, K.A., Justić, D., Wang, L., 2020b. Making the most of available monitoring data: A grid-summarization method to allow for the combined use of monitoring data collected at random and fixed sampling stations. Fisheries Research 229, 105623.
Grüss, A., Thorson, J.T., 2019. Developing spatio-temporal models using multiple data types for evaluating population trends and habitat usage. ICES Journal of Marine Science 76, 1748–1761.
Grüss, A., Thorson, J.T., Anderson, O.F., O'Driscoll, R., Heller-Shipley, M., Goodman, S., 2023d. Spatially varying catchability for integrating research survey data with other data sources: case studies involving observer samples, industry-cooperative surveys,

and predators-as-samplers. Canadian Journal of Fisheries and Aquatic Sciences 80, 1595–1615.
Grüss, A., Thorson, J.T., Sagarese, S.R., Babcock, E.A., Karnauskas, M., Walter III, J.F., Drexler, M., 2017. Ontogenetic spatial distributions of red grouper (*Epinephelus morio*) and gag grouper (*Mycteroperca microlepis*) in the US Gulf of Mexico. Fisheries Research 193, 129–142.
Grüss, A., Walter III, J.F., Babcock, E.A., Forrestal, F.C., Thorson, J.T., Lauretta, M.V., Schirripa, M.J., 2019b. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fisheries Research 213, 75–93.
Grüss, A., Winker, H., Thorson, J.T., Walker, N.D., Maureaud, A., Pacoureau, N., 2024. Coupling state-of-the-art modelling tools for better informed Red List assessments of marine fishes. Journal of Applied Ecology 61, 647–657.
Guan, L., Shan, X., Jin, X., Gorfine, H., Yang, T., Li, Z., 2020. Evaluating spatio-temporal dynamics of multiple fisheries-targeted populations simultaneously: A case study of the Bohai Sea ecosystem in China. Ecological Modelling 422, 108987.
Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P., Buckley, Y.M., 2013. Predicting species distributions for conservation decisions. Ecology Letters 16, 1424–1435.
Han, Q., Grüss, A., Shan, X., Jin, X., Thorson, J.T., 2021. Understanding patterns of distribution shifts and range expansion/contraction for small yellow croaker (*Larimichthys polyactis*) in the Yellow Sea. Fisheries Oceanography 30, 69–84.
Hartig, F., 2020. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.3 3.
Hilborn, R., Walters, C.J., 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Chapman and Hall, New York, NY.
Hodgdon, C.T., Tanaka, K.R., Runnebaum, J., Cao, J., Chen, Y., 2020. A framework to incorporate environmental effects into stock assessments informed by fishery-independent surveys: a case study with American lobster (*Homarus americanus*). Canadian Journal of Fisheries and Aquatic Sciences 77, 1700–1710.
Holmes, S.J., 2021. Stock Assessment of Hake (*Merluccius Australis*) on Chatham Rise for the 2019-20 Fishing Year. New Zealand Fisheries Assessment Report 2021/22. 55 p.
Horn, P.L., 2015. Southern hake (*Merluccius australis*) in New Zealand: biology, fisheries and stock assessment. In: H. Arancibia (Ed.) Hakes: Biology and exploitation, 1st edn, pp. 101-125. Wiley & Sons, Chichester, UK.
Horn, P.L., 1998. The stock affinity of hake (*Merluccius australis*) from Puysegur Bank, and catch-at-age data and revised productivity parameters for hake stocks HAK 1, 4, and 7 (New Zealand Fisheries Assessment Research Document No. 98/34). Ministry of Fisheries, Wellington, New Zealand.
Horn, P.L., 1997. An ageing methodology, growth parameters, and estimates of mortality for hake (*Merluccius australis*) from around the South Island, New Zealand. Marine and Freshwater Research 48, 201–209.
Hsu, J., Chang, Y.-J., Ducharme-Barth, N.D., 2022. Evaluation of the influence of spatial treatments on catch-per-unit-effort standardization: A fishery application and simulation study of Pacific saury in the Northwestern Pacific Ocean. Fisheries Research 255, 106440.

Hurst, R.J., Bagley, N.W., Anderson, O.F., Francis, M.P., Griggs, L.H., Clark, M.R., Paul, L.J., Taylor, P.R., 2000. Atlas of juvenile and adult fish and squid distributions from bottom and midwater trawls and tuna longlines in New Zealand waters. NIWA Technical Report 84. National Institute of Water and Atmospheric Research, Wellington. 162 p.

Isaac, N.J., Jarzyna, M.A., Keil, P., Dambly, L.I., Boersch-Supan, P.H., Browning, E., Freeman, S.N., Golding, N., Guillera-Arroita, G., Henrys, P.A., 2020. Data integration for large-scale models of species distributions. Trends in Ecology & Evolution 35, 56–67.

Jetz, W., McGeoch, M.A., Guralnick, R., Ferrier, S., Beck, J., Costello, M.J., Fernandez, M., Geller, G.N., Keil, P., Merow, C., Meyer, C., Muller-Karger, F.E., Pereira, H.M., Regan, E.C., Schmeller, D.S., Turak, E., 2019. Essential biodiversity variables for mapping and monitoring species populations. Nature Ecology & Evolution 3, 539–551.

Johnson, K.F., Thorson, J.T., Punt, A.E., 2019. Investigating the value of including depth during spatiotemporal index standardization. Fisheries Research 216, 126–137.

Kass, R.E., Steffey, D., 1989. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). Journal of the American Statistical Association 84, 717–726.

Kawauchi, Y., Yagi, Y., Yano, T., Fujiwara, K., 2023. Multi-decadal distribution changes of commercially important demersal species in the central-western Sea of Japan based on a multi-species spatiotemporal model. Regional Studies in Marine Science 61, 102899.

Kendall, M., 1975. Rank correlation methods (4th edn.). Griffin, London, UK.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B., 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70, 1–20.

Lindegren, M., van Deurs, M., Maureaud, A., Thorson, J.T., Bekkevold, D., 2022. A spatial statistical approach for identifying population structuring of marine fish species: European sprat as a case study. ICES Journal of Marine Science 79, 423–434.

Liu, O.R., Ward, E.J., Anderson, S.C., Andrews, K.S., Barnett, L.A., Brodie, S., Carroll, G., Fiechter, J., Haltuch, M.A., Harvey, C.J., 2023. Species redistribution creates unequal outcomes for multispecies fisheries under projected climate change. Science Advances 9, eadg5468.

Lo, N.C., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance from fish spotter data based on delta-lognornial models. Canadian Journal of Fisheries and Aquatic Sciences 49, 2515–2526.

Lock, K., Leslie, S., 2007. New Zealand's quota management system: a history of the first 20 years. Motu Working Paper 07-02. Available from: http://www.ecologyandsociety.org/vol15/iss3/art36/.

Mace, P., 2014. The Challenges Still Ahead. In: P. Mace & Vignaux, M. (Eds) A Celebration of 30+ Years of Fisheries Science. Ministry for Primary Industries (2014) Fisheries Assessment Plenary– Supplement. Ministry for Primary Industries, Wellington, New Zealand. 87 p.

Mackay, K.A., 2020. Database documentation for the Ministry for Primary Industries Fisheries research trawl survey database trawl. NIWA Fisheries Data Management Database Documentation Series. 82 p.

Mäkinen, J., Merow, C., Jetz, W., 2024. Integrated species distribution models to account for sampling biases and improve range-wide occurrence predictions. Global Ecology and Biogeography 33, 356–370.

Mann, H.B., 1945. Nonparametric tests against trend. Econometrica: Journal of the Econometric Society 245–259.

Maunder, M.N., 2004. Population viability analysis based on combining Bayesian, integrated, and hierarchical analyses. Acta Oecologica 26, 85–94.
Maunder, M.N., Punt, A.E., 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research 142, 61–74.
Maunder, M.N., Punt, A.E., 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries Research 70, 141–159.
Maureaud, A., Frelat, R., Pécuchet, L., Shackell, N., Mérigot, B., Pinsky, M.L., Amador, K., Anderson, S.C., Arkhipkin, A., Auber, A., et al., 2021. Are we ready to track climate-driven shifts in marine species across international boundaries? A global survey of scientific bottom trawl data. Global Change Biology 27, 220–236.
McGregor, V.L., 2021. Chatham Rise hake (Merluccius australis) 1989–90 to 2018–19: descriptive analysis of commercial catch effort data and standardised catch and effort data analysis (CPUE). New Zealand Fisheries Assessment Report 2021/34. 38 p.
Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142, 86–99.
Methot, R.D., 2009. Stock assessment: operational models in support of fisheries management. In: Beamish, R.J., Rothschild, B.J. (Eds.), The Future of Fisheries Science in North America. Springer Netherlands, Dordrecht, pp. 137–165.
Miller, D.A., Pacifici, K., Sanderlin, J.S., Reich, B.J., 2019. The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 10, 22–37.
Mitchell, J.S., Mackay, K.A., Neil, H.L., Mackay, E.J., Pallentin, A., Notman, P., 2012. Undersea New Zealand, 1:5,000,000. NIWA Chart, Miscellaneous Series No. 92.
Monnahan, C.C., Thorson, J.T., Kotwicki, S., Lauffenburger, N., Ianelli, J.N., Punt, A.E., 2021. Incorporating vertical distribution in index standardization accounts for spatiotemporal availability to acoustic and bottom trawl gear for semi-pelagic species. ICES Journal of Marine Science 78, 1826–1839.
Morrison, M.A., Bian, R., Jones, E.G., McKenzie, J., Parsons, D., Shankar, U., 2013. The relative abundance of red gurnard, John dory, tarakihi and other inshore finfish species around northern New Zealand–the potential for a new "northern hotspot" trawl survey series. New Zealand Fisheries Assessment Report 8, 130.
Nephin, J., Thompson, P.L., Anderson, S.C., Park, A.E., Rooper, C.N., Aulthouse, B., Watson, J., 2023. Integrating disparate survey data in species distribution models demonstrate the need for robust model evaluation. Canadian Journal of Fishereies and Aquatic Sciences 80, 1869–1889.
Newcombe, R.G., 1998. Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in Medicine 17, 857–872.
Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., Foster, S.D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N.A., Holt, R.D., Hui, F.K.C., Husby, M., Kålås, J.A., Lehikoinen, A., Luoto, M., Mod, H.K., Newell, G., Renner, I., Roslin, T., Soininen, J., Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N.E., Gravel, D., Ovaskainen, O., 2019. A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. Ecological Monographs 89, e01370.
O'Driscoll, R.L., Booth, J.D., Bagley, N.W., Anderson, O.F., Griggs, L.H., Stevenson, M.L., Francis, M.P., 2003. Areas of importance for spawning, pupping or egg-laying, and juveniles of New Zealand deepwater fish, pelagic fish, and invertebrates. NIWA Technical Report 119. National Institute of Water and Atmospheric Research, Wellington. 377 p.

O'Leary, C.A., DeFilippo, L.B., Thorson, J.T., Kotwicki, S., Hoff, G.R., Kulik, V.V., Ianelli, J.N., Punt, A.E., 2022. Understanding transboundary stocks' availability by combining multiple fisheries-independent surveys and oceanographic conditions in spatiotemporal models. ICES Journal of Marine Science 79, 1063–1074.

O'Leary, C.A., Thorson, J.T., Ianelli, J.N., Kotwicki, S., 2020. Adapting to climate-driven distribution shifts using model-based indices and age composition from multiple surveys in the walleye pollock (*Gadus chalcogrammus*) stock assessment. Fisheries Oceanography 29, 541–557.

Olmos, M., Ianelli, J., Ciannelli, L., Spies, I., McGilliard, C.R., Thorson, J.T., 2023. Estimating climate-driven phenology shifts and survey availability using fishery-dependent data. Progress in Oceanography 215, 103035.

Ono, K., Ianelli, J.N., McGilliard, C.R., Punt, A.E., 2018. Integrating data from multiple surveys and accounting for spatio-temporal correlation to index the abundance of juvenile Pacific halibut in Alaska. ICES Journal of Marine Science 75, 572–584.

Oyafuso, Z.S., Barnett, L.A., Kotwicki, S., 2021. Incorporating spatiotemporal variability in multispecies survey design optimization addresses trade-offs in uncertainty. ICES Journal of Marine Science 78, 1288–1300.

Oyafuso, Z.S., Barnett, L.A., Siple, M.C., Cooper, D.W., Kotwicki, S., 2023. Evaluating potential changes to the US Chukchi Sea bottom trawl survey design via simulation testing. Frontiers in Marine Science 10, 1214526.

Pacifici, K., Reich, B.J., Miller, D.A., Gardner, B., Stauffer, G., Singh, S., McKerrow, A., Collazo, J.A., 2017. Integrating multiple data sources in species distribution modeling: a framework for data fusion. Ecology 98, 840–850.

Pagel, J., Anderson, B.J., O'Hara, R.B., Cramer, W., Fox, R., Jeltsch, F., Roy, D.B., Thomas, C.D., Schurr, F.M., 2014. Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. Methods Ecol Evol 5, 751–760. https://doi.org/10.1111/2041-210X.12221

Paradinas, I., Giménez, J., Conesa, D., López-Quílez, A., Pennino, M.G., 2022. Evidence for spatiotemporal shift in demersal fishery management priority areas in the western Mediterranean. Canadian Journal of Fisheries and Aquatic Sciences 79, 1641–1654.

Pennino, M.G., Paradinas, I., Illian, J.B., Muñoz, F., Bellido, J.M., López-Quílez, A., Conesa, D., 2019. Accounting for preferential sampling in species distribution models. Ecology and Evolution 9, 653–663.

Pinto, C., Travers-Trolet, M., Macdonald, J.I., Rivot, E., Vermard, Y., 2019. Combining multiple data sets to unravel the spatiotemporal dynamics of a data-limited fish stock. Canadian Journal of Fisheries and Aquatic Sciences 76, 1338–1349.

Punt, A.E., Dunn, A., Elvarsson, B.Þ., Hampton, J., Hoyle, S.D., Maunder, M.N., Methot, R.D., Nielsen, A., 2020. Essential features of the next-generation integrated fisheries stock assessment package: a perspective. Fisheries Research 229, 105617.

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

Quinn, T.J., Deriso, R.B., 1999. Quantitative fish dynamics. Oxford University Press, New York, NY.

Robinson, O.J., Ruiz-Gutierrez, Viviana., Reynolds, M.D., Golet, G.H., Strimas-Mackey, M., Fink, D., 2020. Integrating citizen science data with expert surveys increases accuracy and spatial extent of species distribution models. Diversity and Distributions 26, 976–986.

Rollinson, C.R., Finley, A.O., Alexander, M.R., Banerjee, S., Dixon Hamil, K.-A., Koenig, L.E., Locke, D.H., DeMarche, M.L., Tingley, M.W., Wheeler, K., Youngflesh, C.,

Zipkin, E.F., 2021. Working across space and time: nonstationarity in ecological research and application. Frontiers in Ecology and the Environment 19, 66–72.
Rufener, M.-C., Kristensen, K., Nielsen, J.R., Bastardie, F., 2021. Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. Ecological Applications 31, e02453.
Sanders, B., Fisher, D., 2020. Database documentation for the Ministry for Primary Industries Centralised Observer Database: Cod. NIWA Fisheries Data Management Database Documentation Series. 628 p.
Saunders, S.P., Cuthbert, F.J., Zipkin, E.F., 2018. Evaluating population viability and efficacy of conservation management using integrated population models. Journal of Applied Ecology 55, 1380–1392.
Schaub, M., Abadi, F., 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152, 227–237.
Schaub, M., Kéry, M., 2021. Integrated population models: Theory and ecological applications with R and JAGS. Academic Press an imprint of Elsevier, London, UK.
Schaub, M., Maunder, M.N., Kéry, M., Thorson, J.T., Jacobson, E.K., Punt, A.E., 2024. Lessons to be learned by comparing integrated fisheries stock assessment models (SAMs) with integrated population models (IPMs). Fisheries Research 272, 106925.
Shelton, A.O., Thorson, J.T., Ward, E.J., Feist, B.E., 2014. Spatial semiparametric models improve estimates of species abundance and distribution. Canadian Journal of Fisheries and Aquatic Sciences 71, 1655–1666.
Simmonds, E.G., Jarvis, S.G., Henrys, P.A., Isaac, N.J., O'Hara, R.B., 2020. Is more data always better? A simulation study of benefits and limitations of integrated distribution models. Ecography 43, 1413–1422.
Smith, J.Q., 1985. Diagnostic checks of non-standard time series models. Journal of Forecasting 4, 283–291.
Smith, P.J., Patchell, G.J., Benson, P.G., 1979. Glucosephosphate isomerase and isocltrate dehydrogenase polymorphisms in the hake, *Merluccius australis*. New Zealand Journal of Marine and Freshwater Research 13, 545–547.
Sofaer, H.R., Jarnevich, C.S., Pearse, I.S., Smyth, R.L., Auer, S., Cook, G.L., Edwards Jr, T.C., Guala, G.F., Howard, T.G., Morisette, J.T., 2019. Development and delivery of species distribution models to inform decision-making. BioScience 69, 544–557.
Stefánsson, G., 1996. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science 53, 577–588.
Stevens, D.W., Ballara, S.L., Escobar-Flores, P.C., O'Driscoll, R.L., 2023. Trawl survey of hoki and middle depth species on the Chatham Rise, January 2022 (TAN2201). New Zealand Fisheries Assessment Report 2023/24. 122 p.
Stow, C.A., Jolliff, J., McGillicuddy Jr, D.J., Doney, S.C., Allen, J.I., Friedrichs, M.A., Rose, K.A., Wallhead, P., 2009. Skill assessment for coupled biological/physical models of marine systems. Journal of Marine Systems 76, 4–15.
The Aotearoa Circle, 2020. Seafood Sector Adaptation Strategy. Climate Adaptation Strategy 2021-2030. The Aotearoa Circle, Wellington, New Zealand.
Thompson, P.L., Anderson, S.C., Nephin, J., Robb, C.K., Proudfoot, B., Park, A.E., Haggarty, D.R., Rubidge, E., 2022. Integrating trawl and longline surveys across British Columbia improves groundfish distribution predictions. Canadian Journal of Fisheries and Aquatic Sciences 80, 195–210.
Thorson, J.T., 2022. VAST model structure and user interface. https: //github.com/James-Thorson-NOAA/VAST.

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., 2018. Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. Canadian Journal of Fisheries and Aquatic Sciences 75(9), 1369–1382.
Thorson, J.T., Adams, C.F., Brooks, E.N., Eisner, L.B., Kimmel, D.G., Legault, C.M., Rogers, L.A., Yasumiishi, E.M., 2020a. Seasonal and interannual variation in spatio-temporal models for index standardization and phenology studies. ICES Journal of Marine Science 77, 1879–1892.
Thorson, J.T., Barnes, C.L., Friedman, S.T., Morano, J.L., Siple, M.C., 2023. Spatially varying coefficients can improve parsimony and descriptive power for species distribution models. Ecography 2023, e06510.
Thorson, J.T., 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 74, 1311–1321.
Thorson, J.T., Clarke, M.E., Stewart, I.J., Punt, A.E., 2013. The implications of spatially varying catchability on bottom trawl surveys of fish abundance: a proposed solution involving underwater vehicles. Canadian Journal of Fisheries and Aquatic Sciences 70, 294–306.
Thorson, J.T., Haltuch, M.A., 2019. Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. Canadian Journal of Fisheries and Aquatic Sciences 76, 401–414.
Thorson, J.T., Ianelli, J.N., Larsen, E.A., Ries, L., Scheuerell, M.D., Szuwalski, C., Zipkin, E.F., 2016a. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Global Ecology and Biogeography 25, 1144–1158.
Thorson, J.T., Maunder, M.N., Punt, E., 2020b. The development of spatio-temporal models of fishery catch-per-unit-effort data to derive indices of relative abundance. Fisheries Research 230, 105611.
Thorson, J.T., Pinsky, M.L., 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, 990–1002.
Thorson, J.T., Rindorf, A., Gao, J., Hanselman, D.H., Winker, H., 2016c. Density-dependent changes in effective area occupied for sea-bottom-associated marine fishes. Proceedings of the Royal Society B: Biological Sciences 283, 20161853.
Thorson, J.T., Shelton, A.O., Ward, E.J., 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, 1297–1310.
Thorson, J.T., Ward, E.J., 2014. Accounting for vessel effects when standardizing catch rates from cooperative surveys. Fisheries Research 155, 168–176.
Warton, D.I., Thibaut, L., Wang, Y.A., 2017. The PIT-trap—A "model-free" bootstrap procedure for inference about regression models with discrete, multivariate responses. PloS one 12, e0181790.
Webster, R.A., Soderlund, E., Dykstra, C.L., Stewart, I.J., 2020. Monitoring change in a dynamic environment: spatiotemporal modelling of calibrated data from different types of fisheries surveys of Pacific halibut. Canadian Journal of Fisheries and Aquatic Sciences 77, 1421–1432.

Whittaker, R.J., Araújo, M.B., Jepson, P., Ladle, R.J., Watson, J.E.M., Willis, K.J., 2005. Conservation Biogeography: assessment and prospect. Diversity and Distributions 11, 3–23.

Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2010. Incorporating time-varying catchability into population dynamic stock assessment models. Reviews in Fisheries Science 18, 7–24.

Xu, H., Lennert-Cody, C.E., Maunder, M.N., Minte-Vera, C.V., 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Fisheries Research 213, 121–131.

Ye, Y., Dennis, D., 2009. How reliable are the abundance indices derived from commercial catch–effort standardization? Canadian Journal of Fisheries and Aquatic Sciences 66, 1169–1178.

Zhou, S., Campbell, R.A., 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 76, 1489–1504.

Zipkin, E.F., Rossman, S., Yackulic, C.B., Wiens, J.D., Thorson, J.T., Davis, R.J., Grant, E.H.C., 2017. Integrating count and detection–nondetection data to model population dynamics. Ecology 98, 1640–1650.

https://github.com/James-Thorson-NOAA/VAST

http://www.nesi.org.nz/

Figure legends

Fig. 1. (a) Spatial distribution of the bottom trawl data collected within the New Zealand observer programme (COD BT data) used in the present study, where each dot corresponds to an individual observer record made between 1991 and 2022. **(b)** Spatial distribution of the bottom trawl data collected within five New Zealand research surveys (Table 2) used in the present study, where each dot corresponds to an individual survey record made between 1991 and 2022. **(c)** Bottom depth (m) in the 10 km × 10 km prediction grid constructed for southern hake (HAK; *Merluccius australis*) in the present study. Also displayed in all panels are the New Zealand exclusive economic zone (delineated by the outer black polygon) and the HAK4, HAK1 and HAK7 quota management areas (QMAs). In the first panel, the following features are labelled: NI = North Island; SI = South Island; CR = Chatham Rise; SA = Sub-Antarctic; and CP = Challenger Plateau.

Fig. 2. Mean spatial patterns of log-density over the period 1991–2022 (log- kg · km^{-2}) predicted by the vector autoregressive spatio-temporal (VAST) models for the southern hake HAK4 stock (top panels) and the VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks) (bottom panels).

Fig. 3. (a) Spatial patterns of the mean coefficient of variation (CV) of density over the period 1991–2022 predicted by the VAST models for the southern hake HAK4 stock (Models 1–3; top panels) and the VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks) (Models 4–6; bottom panels). **(b)** Mean CV of density in the HAK4 quota management area (across all years and locations) for Models 1-6, expressed relative to Model 1.

Fig. 4. Changes in eastward and northward centres of gravity (COGs) predicted by the three VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks): a model fitted to observer-only data (expanded-domain integrated modelling; left panels); a standard integrated model (standard + expanded-domain integrated modelling; middle panels); and a spatially varying catchability (SVC) integrated model (SVC + expanded-domain integrated modelling; right panels). In all panels, the shaded areas represent 95% confidence intervals around VAST model predictions.

Fig. 5. Uncertainty in the habitat assessment for the three VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks): a model fitted to observer-only data ("Obs-only"; expanded-domain integrated modelling); a standard integrated model ("Std Int"; standard + expanded-domain integrated modelling); and a spatially varying catchability (SVC) integrated model ("SVC Int"; SVC + expanded-domain integrated modelling). **(a)** Mean coefficient of variation of eastward COG (across all years) relative to the Obs-only model; **(b)** Mean coefficient of variation of northward COG (across all years) relative to the Obs-only model. **(c)** Mean coefficient of variation of log-effective area occupied (across all years) in the HAK4 quota management area relative to the Obs-only model; **(d)** Mean coefficient of variation of log-effective area occupied (across all years) in the HAK1 quota management area relative to the Obs-only model.

Fig. 6. Changes in log-effective area occupied (log- km^{-2}) in the HAK4 and HAK1 quota management areas predicted by the three VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks): a model fitted to observer-only data (expanded-domain integrated modelling; left panels); a standard integrated model (standard + expanded-domain integrated modelling; middle panels); and a spatially varying catchability (SVC) integrated

model (SVC + expanded-domain integrated modelling; right panels). In all panels, the shaded areas represent 95% confidence intervals around VAST model predictions.

Fig. 7. Indices of relative biomass for the Chatham Rise middle depth area predicted by the VAST models for the southern hake HAK4 stock (Models 1–3; top panels) and the VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks) (Models 4-6; bottom panels). In all panels, the shaded areas represent 95% confidence intervals around VAST model predictions.

Fig. 8. As Fig. 7 but also displaying the stratified random (SurvCalc) index of relative biomass for the Chatham Rise middle depth area obtained from survey data (for years in which surveys were carried out) and the 95% confidence intervals of these SurvCalc index (vertical bars). Also shown are the Pearson's correlation coefficients between the VAST and the SurvCalc index and their *p*-values.

Fig. 9. Uncertainty and interannual variability in the index of relative biomass for the Chatham Rise middle depth area estimated by the VAST models for the southern hake HAK4 stock (Models 1–3) and the VAST models for the New Zealand HAK population (HAK4 + HAK1 + HAK7 stocks) (Models 4–6). **(a)** Mean coefficient of variation of the index (across all years) relative to Model 1. **(b)** Mean interannual variability in the index (across all years) relative to Model 1.

Fig. 10. Bias (the closer to 1 the better), root mean squared error (the lower the better), and coverage (in %; the closer to 50% the better) of the indices of relative biomass for the

1508 Chatham Rise middle depth area estimated within the simulation experiment, by the VAST
1509 models for the southern hake HAK4 stock (Models 1–3) and the VAST models for the New
1510 Zealand HAK population (HAK4 + HAK1 + HAK7 stocks) (Models 4–6).

1511 **Tables**

1512 **Table 1.** Types and sub-types (categories) of integrated spatio-temporal models.

Type	Sub-type (category)	Comments
Standard integrated spatio-temporal model		● Accounting for differences in catchability among data sources via a non-spatially varying catchability term for each non-reference data source
Spatially varying catchability (SVC) spatio-temporal model		● Accounting for differences in catchability among data sources via an SVC term for each non-reference data source
Expanded-domain integrated spatio-temporal model	Expanded-domain data integration only	● Integrating data from the same source but collected within different areas
	Combining expanded-domain and standard data integration	● Integrating data from different sources and collected within different areas ● Accounting for differences in catchability among data sources via a non-spatially varying catchability term for each non-reference data source
	Combining expanded-domain and SVC data integration	● Integrating data from different sources and collected within different areas ● Accounting for differences in catchability among data sources via an SVC term for each non-reference data source

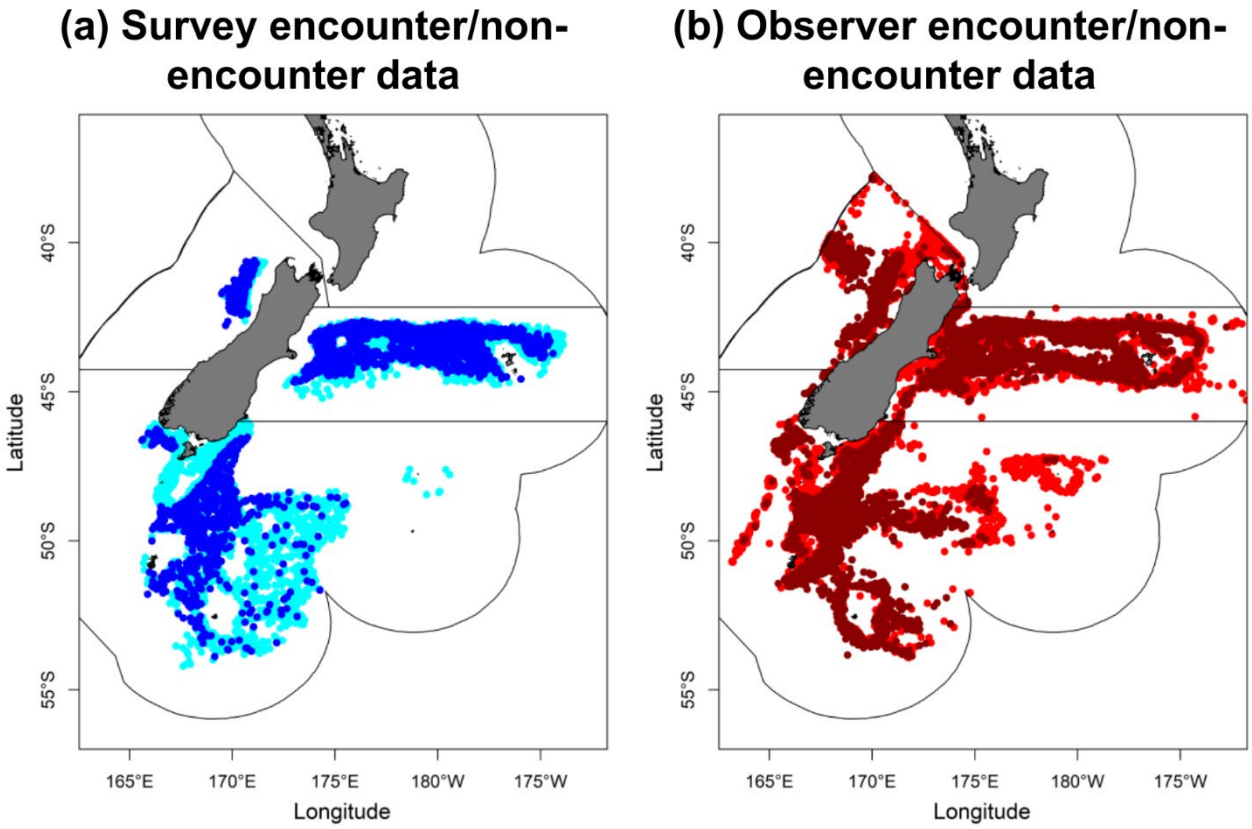
1513

1514 **Table 2.** New Zealand bottom trawl monitoring programmes considered in the present study.

Monitoring programme short name	Monitoring programme long name
COD BT	Fisheries New Zealand's observer programme – Bottom trawl component
CHAT MD	Chatham Rise middle depth trawl survey
SUBA SUM	Sub-Antarctic summer middle depth trawl survey
SUBA AUT	Sub-Antarctic autumn middle depth trawl survey
SOUTH MD	Southland middle depth survey
WCSI MD	West Coast South Island Tangaroa middle depth survey

Appendix A. Supplementary data

Fig. A1. Spatial distribution of the survey and observer encounter/non-encounter data for southern hake (*Merluccius australis*) over the period 1991–2022: (a) encounter (dark blue) and non-encounter (light blue) survey data; and (b) encounter (dark red) and non-encounter (light red) observer data.



Appendix A2. Estimation and evaluation of the spatio-temporal models.

Estimation of the spatio-temporal models

The fixed effects of the vector autoregressive spatio-temporal (VAST) models were estimated by identifying parameter values that maximised the marginal log-likelihood of parameters. The VAST modelling platform uses the Laplace approximation implemented by Template Model Builder (TMB) (Kristensen et al., 2016) to calculate the marginal likelihood of parameters by approximating the integration across all random effects during model fitting. TMB computes the model Hessian (employed by the Laplace approximation) in an efficient manner, as well as the gradient of the log-likelihood included in the Laplace approximation (which is employed during fixed effect maximisation). All the random effects are predicted by TMB by maximising the marginal log-likelihood of parameters given the maximum likelihood estimates of fixed effects. We used the bias-correction estimator from Thorson and Kristensen (2016) to correct for "retransformation bias" when predicting any derived quantity involving a non-linear transformation of random effects. Moreover, we relied on the generalised delta method implemented in TMB to calculate the standard errors of all fixed and random effects, as well as the standard errors of derived quantities (Kass and Steffey, 1989).

Evaluation of the spatio-temporal models

After model fitting, we confirmed model convergence by checking that the gradient of the marginal log-likelihood was less than 0.0001 for all fixed effects and that the Hessian (i.e., the matrix of second-order derivatives describing the curvature of the log-likelihood) was positive definite (the matrix-equivalent of scalar positiveness).

Next, we evaluated model fits using simulation-based probability-integral-transform (PIT) residuals (Smith, 1985; Warton et al., 2017). From the predictive distribution of the biomass catch rate data conditioned on the estimated fixed and random effects, we simulated datasets and then calculated PIT residuals from the observed and simulated values, using R package *DHARMa* (Hartig, 2020). Under the assumptions that the regression model is correct and true parameter values are known, PIT residuals represent exactly an iid sample from the standard uniform distribution. As such, PIT residuals are employed as diagnostic tools, where a PIT residual of 0 indicates that all simulated values were greater than the observed value, whereas a PIT residual of 0.5 indicates that half of the simulated values are greater than the observed value. Thus, spatio-temporal models that display an even spread of PIT residuals between zero and one are considered to have a good fit to the data (Hartig, 2020). We examined histograms and QQ-plots of PIT residuals to be able to gauge how the residuals distribute between 0 and 1. We also examined standardized DHARMa residuals versus model predictions.

Results of the evaluation of the models

The six spatio-temporal models fitted in the present study passed the convergence and evaluation tests (Figs. A2.1 and A2.2). For the six models, the histogram of DHARMa residuals showed a relatively even distribution of DHARMa residuals between 0 and 1, and the Q-Q plot of DHARMa residuals also had DHARMa residuals along the one-to-one line (Figs. A2.1 and A2.2). The QQ-plots of DHARMa residuals illustrated that DHARMa residuals approximately followed the expected distribution and the plots of standardized DHARMa residuals versus model predictions suggested that there was no evidence for

underdispersion or overdispersion, indicating that there were roughly as many residuals as expected in the tails of the distributions from the fitted models (Figs. A2.3 to A2.8).

References of Appendix A2

Hartig, F., 2020. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. R package version 0.3 3.

Kass, R.E., Steffey, D., 1989. Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). Journal of the American Statistical Association 84, 717–726.

Kristensen, K., Nielsen, A., Berg, C.W., Skaug, H., Bell, B., 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70, 1–20.

Smith, J.Q., 1985. Diagnostic checks of non-standard time series models. Journal of Forecasting 4, 283–291.

Thorson, J.T., Kristensen, K., 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fisheries Research 175, 66–74.

Warton, D.I., Thibaut, L., Wang, Y.A., 2017. The PIT-trap—A "model-free" bootstrap procedure for inference about regression models with discrete, multivariate responses. PloS one 12, e0181790.

Figures of Appendix A2

Figure A2.1. Histogram of DHARMa residuals (left panels) and QQ-plot of DHARMa residuals (right panels) for three of the spatio-temporal models developed in the present study: (Model 1) the model fitted to observer-only data for HAK4; (Model 2) the standard integrated model fitted to both observer and survey data for HAK4; and (Model 3) the spatially varying catchability (SVC) integrated model fitted to both observer and survey data for HAK4. In the left panels, the red bar at one shows the frequency of DHARMa residuals with a value of one.

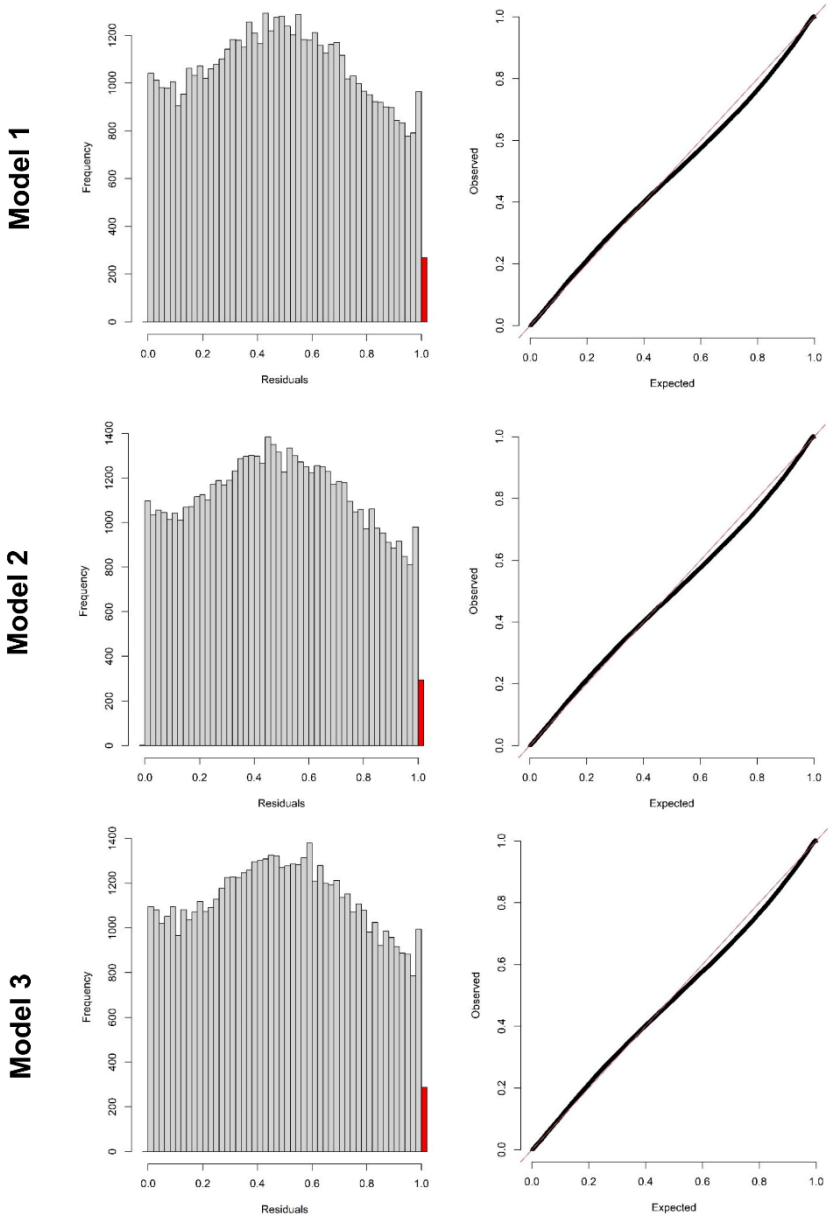


Figure A2.2. Histogram of DHARMa residuals (left panels) and QQ-plot of DHARMa residuals (right panels) for three of the spatio-temporal models developed in the present study: (Model 4) the expanded-domain integrated model fitted to observer-only data for the HAK population (encompassing the HAK4, HAK1 and HAK7 stocks); (Model 5) the standard integrated model fitted to both observer and survey data for HAK, i.e., the model combining expanded-domain and standard integrated modelling; and (Model 6) the spatially varying catchability (SVC) integrated model fitted to both observer and survey data for HAK, i.e., the model combining expanded-domain and SVC integrated modelling. In the left panels, the red bar at one shows the frequency of DHARMa residuals with a value of one.

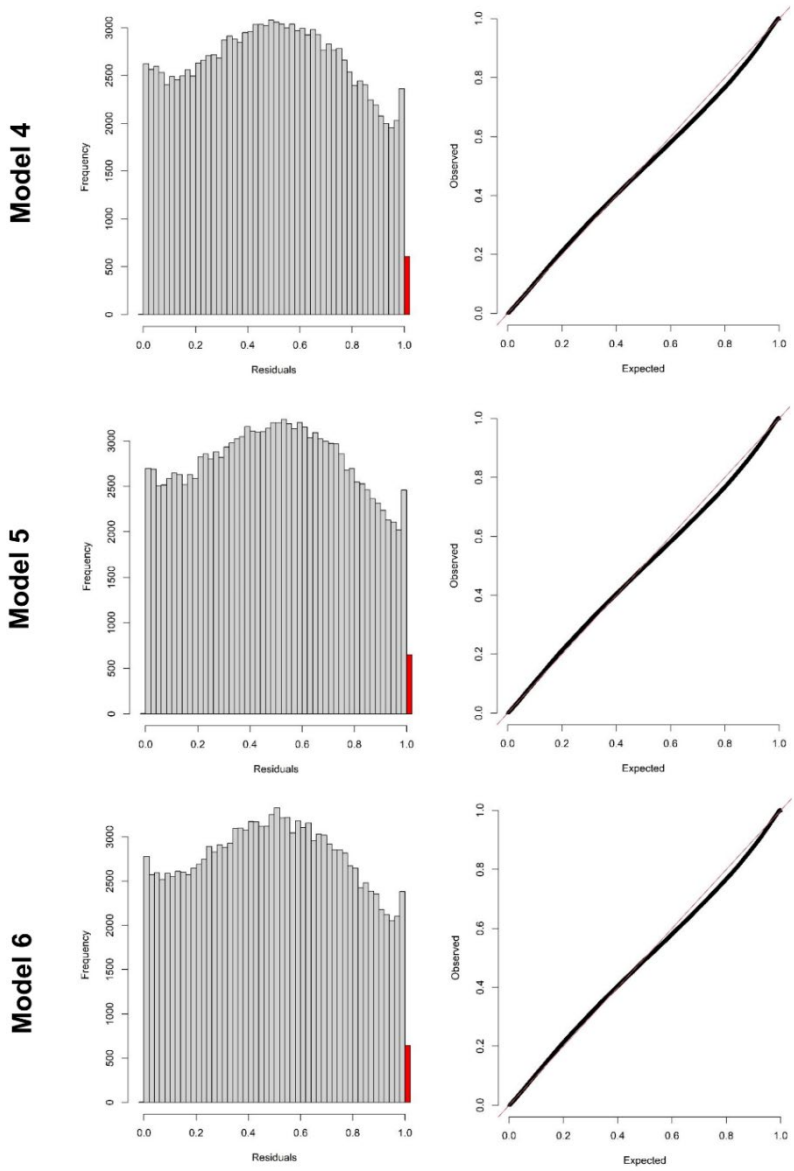


Figure A2.3. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the model fitted to observer-only data for HAK4 (Model 1).

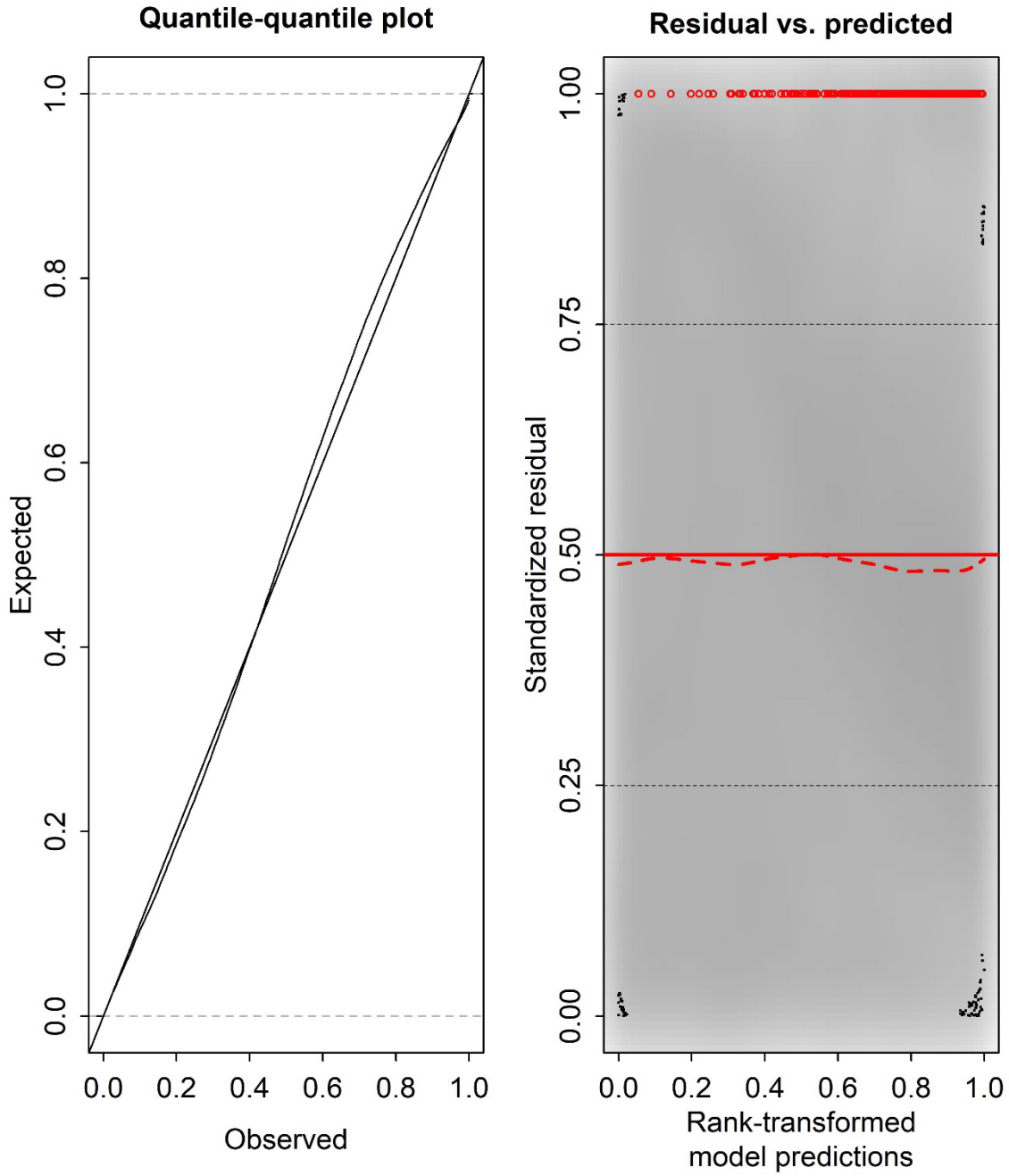


Figure A2.4. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the standard integrated model fitted to both observer and survey data for HAK4 (Model 2).

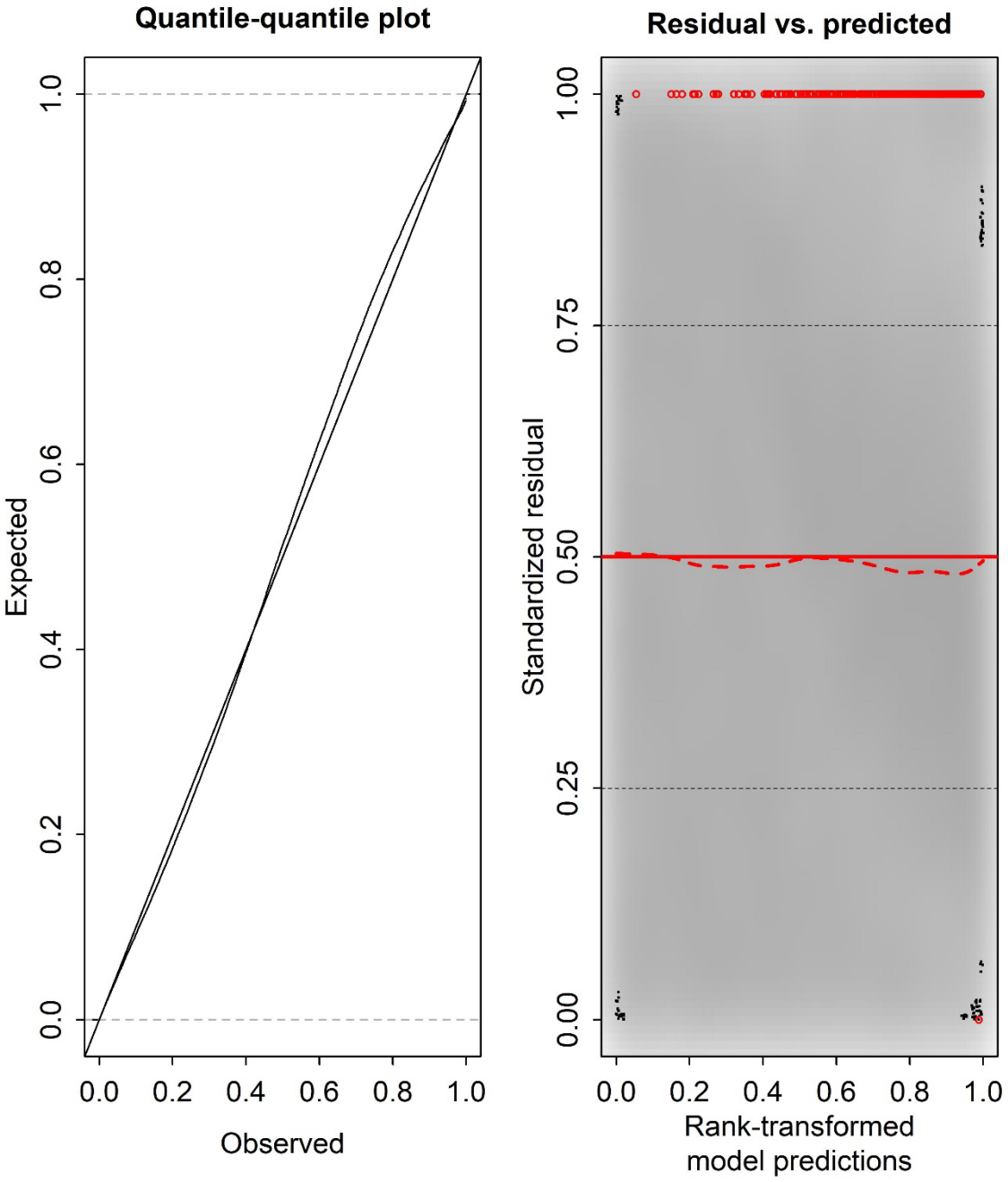


Figure A2.5. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the spatially varying catchability (SVC) integrated model fitted to both observer and survey data for HAK4 (Model 3).

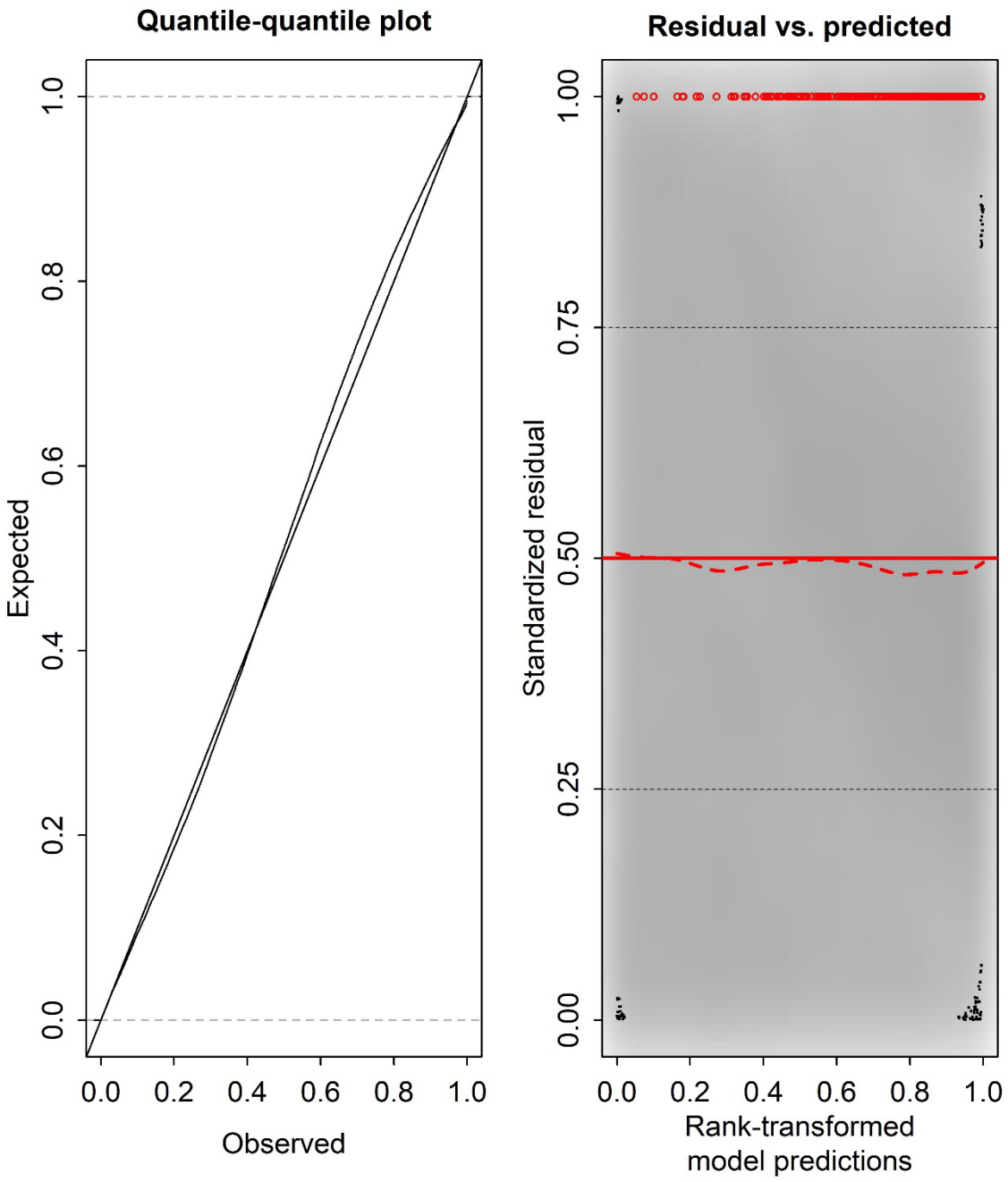


Figure A2.6. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the expanded-domain integrated model fitted to observer-only data for the HAK population (encompassing the HAK4, HAK1 and HAK7 stocks) (Model 4).

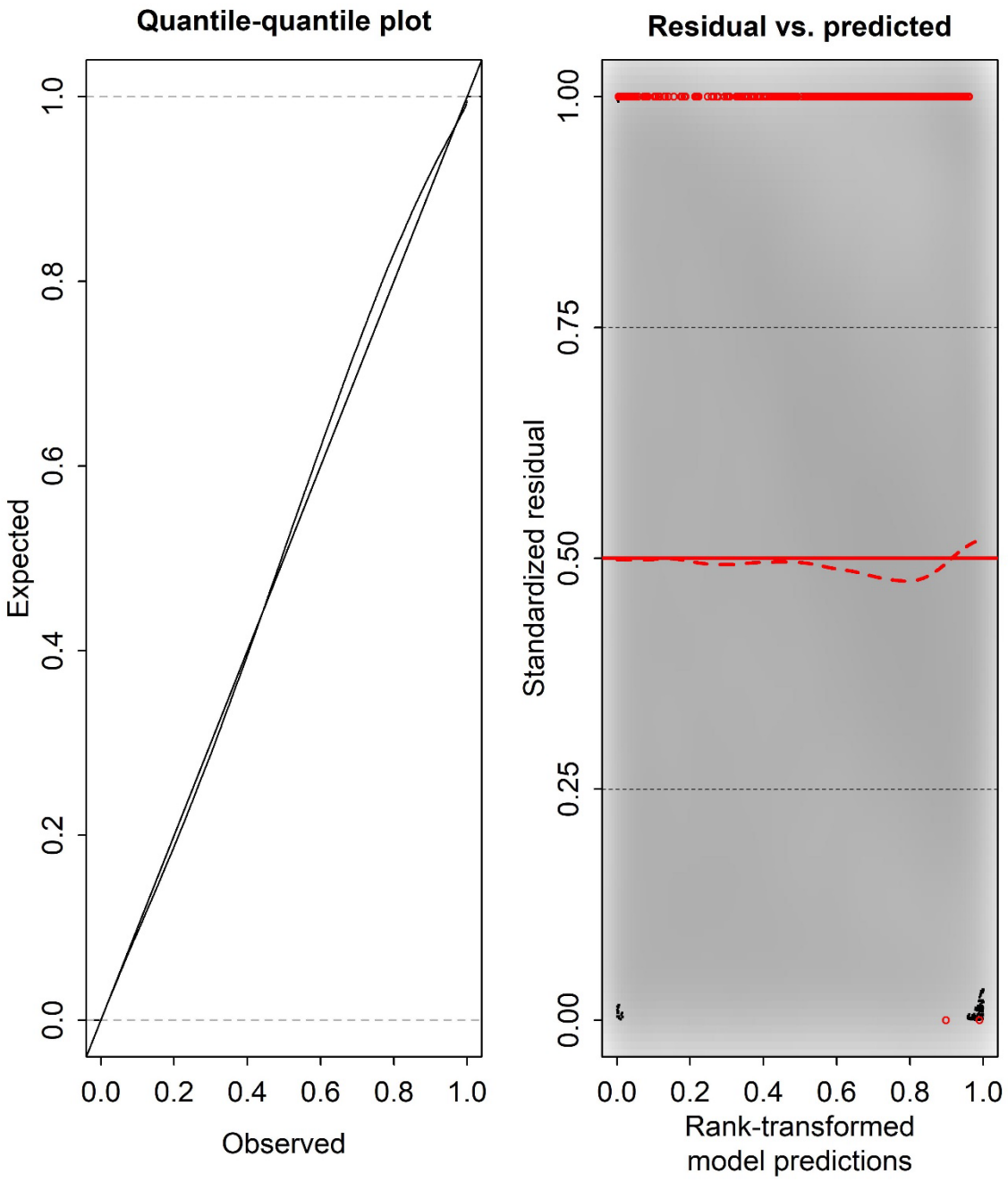


Figure A2.7. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the standard integrated model fitted to both observer and survey data for the HAK population (encompassing the HAK4, HAK1 and HAK7 stocks), i.e., the model combining expanded-domain and standard integrated modelling (Model 5).

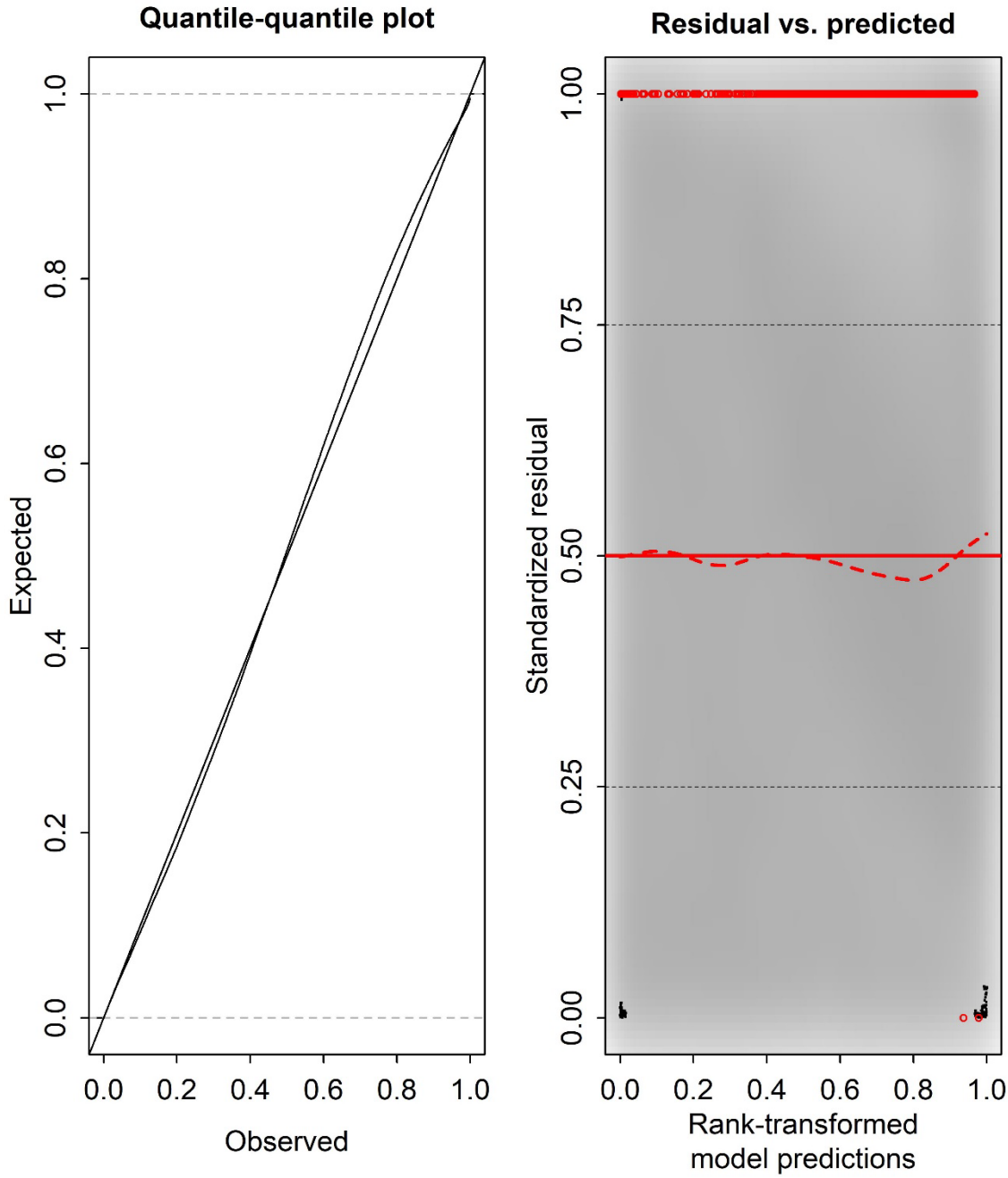


Figure A2.8. Q-Q plot of DHARMa residuals and standardized DHARMa residuals versus model predictions for the spatially varying catchability (SVC) integrated model fitted to both observer and survey data for the HAK population (encompassing the HAK4, HAK1 and HAK7 stocks), i.e., the model combining expanded-domain and SVC integrated modelling (Model 6).