Proof of concept for a novel procedure to standardize multispecies catch and effort data

Henning Winker[a,1] , Sven E. Kerwath[b,a] and Colin G. Attwood[a]

[a]*Marine Research Institute, Zoology Department, University of Cape Town, Private Bag Rondebosch 7701, South Africa.*

[b]*Department of Agriculture Forestry and Fisheries, Private Bag X2, Roggebaai 8012, South Africa*

---

[1] To whom all corresponding should be addressed:
e-mail: henning.winker@gmail.com
telephone: +27 (0) 21 6503459; fax: +27 (0) 21 6503301

ABSTRACT

To estimate reliable abundance indices for multispecies fisheries, the effect of variability in targeting needs to be removed from catch-per-unit-effort (CPUE) data. We test a novel Generalized Additive Model (GAM) that includes principal component scores (PCs) derived from the species composition in the catch, called the ‘Direct Principal Component’ (DPC) procedure, for its ability to remove the effect of variable targeting. A simple operating model is used to simulate exponential biomass trends with variable rates for two multi-species, multi-habitat fisheries scenarios: (i) four species distributed differentially across two habitats and (ii) ten species distributed differentially across four habitats. Tweedie distributed CPUE records are generated from the biomass trends for a fishery with constant targeting (control scenarios) and time-varying targeting (test scenarios). The DPC procedure is simulation-tested for its ability to estimate the underlying biomass trends for all species relative to the non-standardized CPUE index for the control and test scenarios. The DPC procedure proved to be more accurate and precise compared to the nominal CPUE trends in the test scenarios. Even in the control scenarios, the DPC method offers greater precision by removing substantial variation from the data, with a very low penalty on accuracy. The Kaiser-Guttman rule and Cattel’s scree-test proved to be suitable heuristics to select the optimal number of PCs to be included in the GAM. We expect that the DPC procedure is applicable in standardization procedures for a variety of multispecies fisheries, including hook and line, longline and trawl.

Keywords: CPUE standardization, simulation-testing, multispecies targeting, Tweedie distribution, GAM

1. Introduction

The standardization of catch-per-unit-effort (CPUE) is now widely regarded as a prerequisite for the use of CPUE as abundance index in stock assessment models (Maunder and Punt, 2004; Maunder et al., 2006). The nominal CPUE index, derived from yearly means of the raw CPUE data, can be severely biased due to non-random allocation of fishing effort over time (Harley et al., 2001; Maunder et al., 2006; Carruthers et al., 2010). The most commonly used standardization procedures entail the application of Generalized Linear Models (GLMs) or Generalized Additive Models (GAMs), which aim to isolate the temporal abundance trends from the total variation in the CPUE data by adjusting for confounding effects on the estimated abundance trends (Guisan et al., 2002; Maunder and Punt, 2004). Influences on the CPUE other than abundance are remarkably diverse and typically include time-variant changes in spatial and seasonal effort distribution, gear, fishing power and fishing behavior (Punt et al., 2000; Maunder and Punt, 2004; Carruthers et al., 2010). The problem of estimating reliable abundance indices is exacerbated in multispecies fisheries for which the available CPUE records may reflect a number of different fishing strategies, each associated with a particular choice of fishing-ground, habitat-type, and fishing-technique; even within the same fishing trip (Pelletier and Ferraris, 2000; Palmer et al., 2009; Winker et al., 2013).

An important consideration for the standardization of multispecies CPUE data is that the choice of fishing tactic allocates effort towards a particular target species or species complex and away from others, where the term ‘fishing tactic’ is defined here as a sequence of choices of fishing strategies made by the skipper during a fishing trip (Pelletier and Ferraris, 2000; Winker et al., 2013). Temporal variations in fishing tactics inevitably violate the fundamental assumption that CPUE is proportional to the product of abundance and a constant catchability (i.e. fraction of biomass/abundance caught per unit effort), because the latter will strongly depend on the choice of fishing tactic (Pelletier and Ferraris, 2000; Carvalho et al., 2010; Winker et al., 2013).

Conventional multispecies standardization models often include the catch rates of alternative target or bycatch species as covariates to correct for the effort directed away from the target species or species under consideration (Glazer and Butterworth, 2002; Maunder and Punt, 2004; Su et al., 2008). Importantly, the alternative species should not co-occur with the target species. For example, if two species were to co-occur in the catches and would be fished down simultaneously, the use of the catch rate of the one species as a negative predictor of the CPUE of the other may result in an erroneous removable of the underlying year-effect (Glazer and Butterworth, 2002; Maunder and Punt, 2004). An additional problem in situations where a large number of species are caught by the fishery is the objective selection of species-specific catch rates to be included as covariates in the standardization model.

An alternative approach is based on clustering fishing trips according to their similarity in catch composition (He et al., 1997; Pelletier and Ferraris, 2000; Carvalho et al., 2010). The identified clusters are assumed to be a representation of fishing tactics, which may be treated as categorical variables in the standardization model in order to adjust for differences in catchability associated with each cluster (Pelletier and Ferraris, 2000; Carvalho et al., 2010; Winker et al., 2013). This approach typically requires the implementation of a rather complex analytical framework based on a sequence of ordination and clustering techniques and involves several steps that can be associated with elements of subjectivity (Pelletier and Ferraris, 2000; Deporte et al., 2012; Winker et al., 2013).

A more direct method for the standardization of multispecies CPUE records was recently proposed by Winker et al. (2013). This ‘Direct Principal Component’ procedure (DPC) uses continuous principle component scores (PCs), derived from a Principal Component Analysis (PCA) of the catch composition data, as nonlinear predictor variables in a GAM framework to adjust for the effect of temporal variations in fishing tactics. The DPC procedure is based on the common assumption that information on the direction and extent of targeted effort can be found in the species composition of the catch (Pelletier and Ferraris, 2000; Carvalho et al., 2010). However, although the species composition does not hold direct information about the magnitude of the catch, it is arguably of concern that the information contained in the predictor variables derived from the catch composition is not entirely independent from the response CPUE and may have unpredictable impacts on the standardized CPUE trends. The standardization procedure would fail if variation in abundance of a particular species is falsely attributed to variation in targeting. Common model selection procedures, such as analysis of deviance, Akaike's information criterion (AIC) or cross-validation methods only evaluate the model based on how well it fits the data, but may fail to identify the model that provides least biased representation of the true abundance pattern (Hinton and Maunder, 2003; Carruthers et al., 2010).

The aim of this study was therefore to use simulation testing to evaluate if the DPC method is able to accurately track 'true' abundance trends. The specific objectives were: (i) to test the efficacy of the DPC method in eliminating the effect of time-varying trends in fishing tactics on the nominal CPUE record, (ii) to evaluate the risk associated with the DPC method to introduce potential bias and (iii) to evaluate alternative selection criteria for including the optimal number of PC covariates in the DPC standardization model.

2. Materials and methods

Simulation testing is a powerful tool to evaluate the performance of CPUE standardization models (Carruthers et al., 2010; Lynch et al., 2012; Thorson et al., 2012). The advantage of this approach is that the simulated abundance trends are known, so that the standardization model can be directly tested in terms of how well it predicts ‘true’ abundance trends. We developed a simple operating model to simulate non-standardized nominal CPUE records per trip for two scenarios: (i) a simple multispecies-fishery scenario, comprising four target species that are unevenly distributed across two different habitats; and (ii) a more complex multispecies-fishery scenario, comprising ten target species that are unevenly distributed across four different habitats. The chosen scenarios broadly resemble the habitat associations and catch rates of several common target species in the South African multispecies hand-line fishery, which represented the initial case study for the DPC procedure (Winker et al., 2013).

As is common practice, the use of CPUE as an index of abundance was based on the concept that catch is proportional to the product of catchability and biomass: CPUE = qB, where q is the catchability representing the fraction of biomass caught by expending one standard unit of effort and B is the biomass (Maunder and Punt, 2004). This relationship only holds if q is constant, which is almost certainly violated in multispecies-fisheries that employ a variety of fishing tactics. To simulate this effect, we assumed that the choice of targeting tactic is reflected by the choices of up to two target habitats during a fishing trip and that each habitat is associated with a species-specific catchability based on the conceptual consideration outlined in Stephens and MacCall (2004) and Winker et al. (2013). All simulations were conducted within the statistical environment R (R Development Core Team, 2011).

2.1. Basic dynamics

The ‘true’ underlying abundance trends were simulated in the form of biomass trajectories for each species i in year y over a period of 20 years as a function of:

$$B_{i,y} = B_{i,1}e^{(r_i(y-1))} \qquad y = 1, 2, \ldots, 20. \qquad (1)$$

where $B_{i,1}$ is the biomass of species i at start of the time-series and r_i is the rate of increase (or decrease) for species i. Nominal CPUE records for each species i and each trip t in year y were assumed to be Tweedie distributed and were expressed as:

$$CPUE_{i,t,y} = \sum_h q_{i,h} B_{i,y} f_h \qquad (2)$$

where $q_{i,h}$ is the catchability of species i in habitat h and f_h is the fraction of trip effort allocated to habitat h during a fishing trip. The Tweedie distribution belongs to the family of exponential dispersion models and is characterized by the two-parameter power mean-variance function of the form Var(Y)= $\phi\mu^p$,where ϕ is the dispersion parameter, μ is the mean (i.e. here CPUE$_{i,t}$) and p is the power parameter. Depending on the power parameter the Tweedie model includes the four well-known distributions: normal (p = 0), Poisson (p = 1), gamma (p = 2) and inverse-gamma (p= 3) (Dunn and Smyth, 2005). Here, we considered Tweedie distributions for the case $1 < p < 2$, which represents the class of Poisson mixtures of gamma distributions. These mixed distributions have the advantage that high frequencies of zeros in combination with right-skewed real numbers of positive catches can be handled in a natural way (Candy, 2004; Shono, 2008).

2.2. Standardization model

The simulated nominal CPUE$_{t,i,y}$ data were standardized by applying the DPC procedure (Winker et al., 2013). This method was developed on the premises that continuous principal component scores (PCs), derived from a PCA of the catch composition data, can be used as non-linear predictor variables for targeted effort within a Generalized Additive Model (GAM) framework (Winker et al., 2013). The performance of this method is tested by comparing standardized CPUE indices with corresponding nominal CPUE indices.

The first step of the DPC procedure entails applying a PCA to a multidimensional $\mathrm{CPUE}_{t,i,y}$ matrix. For this purpose, a data matrix only comprising $\mathrm{CPUE}_{t,i,y}$ records was extracted from the simulation dataset. The $\mathrm{CPUE}_{t,i,y}$ records were normalized into relative proportions by weight to eliminate the influence of catch volume and then either square-root (R2) or fourth-root (R4) transformed to further shift the source of information away from raw abundance (Winker et al., 2013). In the next step, the predicted principal component scores (PCs) were directly aligned with the CPUE records in the datasets for subsequent use as covariates in the GAM analysis.

First, the nominal target species CPUE indices were derived by applying a simple model of the form:

$$\log(\mathrm{CPUE}) = \alpha + Year \qquad (3)$$

where α is the intercept and *Year* denotes the year-effect treated as categorical variable. To estimate standardized abundance indices based on the DPC method, GAMs for each species were formulated as:

$$\log(\mathrm{CPUE}) = \alpha + Year + s(\mathrm{PC1}) + \ldots + s(\mathrm{PC}n) \qquad (4)$$

where PCn denotes the n^{th} principal component scores included in the model and $s()$ denotes a thin plate regression spline smoother function. The maximum number of knots was limited to $k \le 6$ in order to reduce the flexibility of the splines and to improve the model convergence probability and convergence time during simulation runs. The ’mgcv’ package in R was used to fit all models assuming a Tweedie distribution of the dependent variable with a log-link function.

2.3. Simulation scenarios

Simulation tests involving multi-species, multi-habitat operating models can quickly become mired in a vast number of permutations of scenarios and model formulations. In an attempt to simplify the study and facilitate comparisons we devised the following system of nomenclature. The successive terms in the following example string refer to (in order): Number of habitats (H*n*), ‘.’ the number of species (S*n*), ‘.’ the type of data transformation used (2nd 4th-root, abbreviated R2 and R4, respectively), ‘.’ the number of PC axes used in the GAM (PC*n*). The term “Nom” refers to the nominal CPUE for any given scenario.

First, a simple multispecies fishery is simulated, which targets four species in two different habitats (H2.S4). This scenario includes the species silver kob (KOB), geelbeek (GLBK), hake (HAKE) and panga (PANG) that are distributed across shallow- and deep water soft sediment habitats (Fig. 1). KOB and GLBK are abundant in shallow sediment habitats but are less common in deeper waters, whereas HAKE and PANG represent the dominant target species in deepwater soft sediment habitats. The species-specific distributions across the two habitats are determined by 4 species (i) × 2 habitat (h) matrixes of $q_{i,h}$ values, summarized in Table 1.

Two alternative effort scenarios were considered to simulate the distribution of fishing trips across habitats. The first effort scenario (E1) simulates time-invariant probabilities $\pi_{h,y}$ and acts as a ‘control’ (Fig. 2a), while the second effort scenario represents the ‘test’ scenario (E2), in which an abrupt switch in $\pi_{h,y}$ between the two habitats is simulated (Fig. 2b).

In the more complex scenario, the multispecies-fishery is extended to ten target species, which are distributed across four different habitats (H4.S10): shallow- and deep water soft sediment and shallow- and deep water reef habitats (Fig. 1). The species-assemblages of shallow- and deep water soft sediment habitats correspond to the first multispecies fishery scenario H2.S4. The shallow water reef assemblage is dominated by roman (ROMN), dageraad (DRGD), red stumpnose (RSTM) and santer (SNTR), while carpenter (CRPN) represents the dominant target species over deep water reefs (Fig. 1). There is some distributional overlap among reef associated species. In addition, we introduced the group 'sharks' (SHRK), for which small catches are occasionally made in all four habitats, but which are more common in both shallow water habitats . The species-specific distributions across habitats are determined by 10 species (i) × 4 habitat (h) matrixes of $q_{i,h}$ values, summarized in Table 1. As in scenario H2.S4, we considered two alternative effort scenarios for the distribution of $\pi_{h,y}$, with the first effort scenario (E1) acting as 'control' by simulating time-invariant trends in the probabilities $\pi_{h,y}$ (Fig 2c). The 'test' scenario (E2) simulates linear increases and decreases in $\pi_{h,y}$ for two habitats as well as an abrupt switch in $\pi_{h,y}$ between the other two habitats (Fig. 2d).

The idea of applying the PCA to the catch composition data is that nontrivial sources of variation are retained in the first few PC-axes, which are meant to represent meaningful separations of alternative fishing tactics associated with each target habitat. Conceptually, the number of Principal Components (PCs) required to correctly separating the species assemblages that are associated with each target habitat h is given by the total number habitats minus one (Fig. 3). For example, in the simple case of the two-habitat scenario the first PC is expected to separate the species assemblage into the dominant shallow water soft sediment species, KOB and GLBK, on the one side and the dominant deep water soft sediment species, HAKE and PANG, on the other side of the spectrum, as this split among species should account for most of the variation in the species composition data (Fig. 3a). Accordingly, the first three PCs are required to separate the species assemblage of the four-habitat scenario according to habitat preference (Figs. 3b-c). In the example given in Fig. 3b, the first PC separates the species assemblage into sediment and reef associated species, while the second PC is separating shallow and deep water sediment species. The third PC then isolates CRPN, being the dominant deep water reef species, from the species assemblage (Fig. 3c).

To confine the analysis to a reasonable number of scenarios, the two-habitat scenarios were only tested to include a maximum of two PCs as non-linear predictors. For the same reason, we considered the inclusion of the first two to four PCs for all four-habitat scenarios.

2.4. Data generation

A total of 200 simulation datasets were randomly generated for each simulation scenario. Each simulation dataset consisted of 500 trips per year and correspondingly a total of 10000 trip records over the 20 year period. The following randomization procedures were applied in order to generate the simulation datasets:

(1) Random biomass time series for each species i, $B_{i,y}$ (eq. 1), were generated by drawing random variants of r_i from uniform distribution with bounds at -0.1 and +0.1, U(-0.1, 0.1). Random biomass values at the start of the time series, $B_{i,1}$, were generated from a lognormal distribution as $B^*_{i,1} = 200^{(\varepsilon)}$ and $\varepsilon \sim N(0, 0.5^2)$ for the abundant species KOB, GLBK, HAKE and CRPN; and as $B^*_{i,1} = 50e^{(\varepsilon)}$ and $\varepsilon \sim N(0, 0.5^2)$ for the less abundant species PANG, SNTR, ROMN, DGRD, RSTM and SHRK (Fig. 4a-b).

(2) The distribution of effort across the alternative target habitats was determined as the probability $\pi_{h,y}$ that habitat h is targeted in year y such that $\sum_h \pi_{h,y} = 1$. The vectors $\pi_{h,y}$ that determine the probably for each habitat h being targeted in year y were randomly resampled without replacement, to vary the effort trends among habitats. Note that there were only two possible habitat × $\pi_{h,y}$ vector combinations for the two-habitat scenarios but 24 possible habitat × $\pi_{h,y}$ vector combinations for the four-habitat scenarios. For each simulation run, we introduced random noise, such that:

$$\pi^*_{1,y} = \pi_{1,y} e^{\varepsilon} \quad \text{and} \quad \pi^*_{2,y} = 1 - \pi^*_{1,y}, \qquad (5)$$

in the case of the two-habitat scenario and

$$\pi^*_{h,y} = \pi_{h,y} e^{\varepsilon} \quad \text{for } h = 1,2,3 \quad \text{and} \quad \pi^*_{h,y} = \begin{cases} 1 - \sum_{h=1}^{3} \pi^*_{h,y} & \text{if } \sum_{h=1}^{3} \pi^*_{h,y} \le 1 \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

in the case of the four-habitat scenario, where $\varepsilon = N(0,0.2^2)$ (Fig. 4c-d).

(3) The operating model permits the targeting of up to two habitats during each trip such that the catch composition can represent mixtures of two fishing strategies. This was realized by introducing a 25% probability that an additional target habitat is randomly drawn from the as a function $\pi_{h,y}$ during trip t. The fraction of the trip effort that is allocated to one of two habitats was generated from a sigmoid function of the form:

$$f_{h1} = \frac{1}{1 - \exp(x - 0.5)/0.5} \qquad (7)$$

where x is random uniform variable. The trip effort that is allocated to the second habitat is correspondingly given by $f_{h2} = 1 - f_{h1}$.

(4) Random CPUE$_{i,t}$ deviates (equation 2) were generated from Tweedie distribution using the random Tweedie number generator available in the ‘mgcv’ package for R. Specifically, we set $p = 1.3$ in combination with a fairly large dispersion parameter of $\phi = 10$. In combination with the values $B^{*}_{i,1}$ and $q_{i,h}$, this choice of parameterization produced realistic ranges of CPUE$_{i,y}$ records when compared to the hand-line fishery data. The resulting frequency distributions closely resembled the shape of the delta-lognormal distribution, which is commonly assumed for CPUE data (Lo et al., 1992; Maunder and Punt, 2004; Thorson et al., 2012). Examples of simulated frequency distributions of nominal CPUE records of four species are illustrated for the H2.S4.E2 scenario in Fig. 5.

2.5. Performance evaluation

The performance of the DPC standardization models was evaluated in terms of the ability to accurately estimate r_i in comparison to the nominal CPUE indices. Estimates of $\hat{r}_i$ were obtained from a simple linear regression of the form:

$$\log\left(\overline{\mathrm{CPUE}}_{i,y}\right) = \alpha + \hat{r}_i y \qquad y = 1, 2, \ldots, 20 \qquad (8)$$

where $\overline{\mathrm{CPUE}}_{i,y}$ is the estimated mean CPUE for species i in year y based on either the nominal or standardized CPUE, and a is the intercept term.

The scenarios H2.S4.E1 and H4.S10.E1 act as controls, for which $\hat{r}_i$ estimated from the nominal CPUE indices is expected to be unbiased. The Absolute Relative Error (ARE) was used to summarize the estimation performance of $\hat{r}_i$ relative to the 'true' values r_i that govern $B_{i,y}$ (Ono et al., 2012). The ARE quantifies the average model precision and therefore provides a relative estimate for the goodness-of-the-fit:

$$ARE_{i,k} = \left| \frac{\hat{r}_{i,k} - r_{i,k}}{r_{i,k}} \right| \qquad (9)$$

where $ARE_{i,k}$ is the Relative Error for species i from the k^{th} simulation and $\hat{r}_{i,k}$ is the estimate of the 'true' r_i generated in the k^{th} simulation. Smaller values of the $\text{ARE}_{i,k}$ mean that $\hat{r}_{i,k}$ was estimated closer to $r_{i,k}$.

To evaluate how precisely the nominal and standardized CPUE followed the simulated biomass $B_{i,y}$, the indices were rescaled through normalization by the geometric mean and compared based on the mean deviation between the simulated and estimated normalized abundance indices, such that:

$$MD_{i,k} = \sqrt{\frac{1}{20}\sum_{y=1}^{20}(\hat{I}_{i,y,k} - I_{i,y,j})^2} \qquad (10)$$

where $\text{MD}_{i,k}$ is the mean deviation for species i from the j^{th} simulation, $\hat{I}_{i,y,k}$ is the normalized abundance index for year y based on nominal or standardized CPUE and $I_{i,y,j}$ is the normalized abundance index based on the simulated biomass.

Comparisons of accuracy and precision among species were achieved by calculating the difference in ARE and MD medians between the nominal CPUE and best-performing DPC procedures. The purpose of the comparisons was to examine the influence of species distribution among habitat on the relative performance of each procedure.

2.6. Selection of Principal Components

Considering that the number of meaningful fishing tactics associated with each target habitat will often be unclear, it was an important objective of this study to evaluate alternative selection criteria that can potentially provide guidance for selecting the optimal number PCs in the standardization models. Representing widely used 'stopping rules' for the selection PCs, we first considered the Kaiser-Guttman rule (Guttman, 1954) and Cattel's scree-test (Cattell, 1966). The Kaiser-Guttman rule states that only PCs with Eigenvalues larger than one are retained. The reasoning behind this rule is that selected PCs summarize more information than any single original component (Jackson, 1993). The scree-test is a visual approach and is conducted by plotting the PCs (x-axis) against their Eigenvalues (y-axis). The smaller Eigenvalues of the higher order PCs tend to lie in along a straight line. The point where the first Eigenvalues markedly depart from this line is assumed to separate the trivial PCs, representing random noise, from the meaningful PCs that should be retained for further analysis (Jackson, 1993). In addition to the two stopping rules, we also noted the AIC and the Bayesian Information Criterion (BIC) for each model fit to evaluate whether or not these two standard model selection criteria are suitable for determining the optimal number of PCs to be included in the GAM framework.

3. Results

3.1 Simulation datasets

A total of 200 simulation datasets were generated for each control and test of the two- and four-habitat scenarios (H2.S4.E1, H2.S4.E2, H4.S10.E1 and H4.S10.E2), respectively. The simulation datasets covered wide ranges of proportions of zero catches (Table 2). Because fishing trips were distributed among more habitats in the four habitat datasets, the proportion of zeros in CPUE records in these scenarios was generally larger (Table 2). In the two-habitat scenario datasets, proportion of zeros ranged from 0.01 to 0.45 for the abundant species and from 0.19 to 0.74 for the less abundant PANG. In the four-habitat scenario dataset, proportion of zeros ranged from 0.28 to 0.75 and from 0.31 to 0.86 for abundant and less abundant species, respectively. Variation in CPUE among species was typically four-fold (Table 2). The positive catches were right-skewed (Fig. 5) with occasional very large maximum CPUE values in comparison to the mean (Table 2).

3.2 Standardization performance

Normalized annual abundance indices were calculated for each simulated dataset following the five approaches (nominal, PC1.R2, PC1.R4, PC2.R2, PC2.R4) for the two-habitat scenarios (S2.H4.E1 and S2.H4.E2) and the seven approaches (nominal, PC2.R2, PC2.R4, PC3.R2, PC3.R4, PC4.R2, PC4.R4) for the four-habitat scenarios (H4.S10.E1 and H4.S10.E2). The abundance indices over the 20 year simulation period were compared to the known biomass trends in terms of the accuracy metric ARE and the precision metric MD.

Across the 200 simulation runs, the nominal CPUE resulted in the lowest medians of the ARE in the control scenarios H2.S4.E1 and H4.S10.E1, but performed worst compared to all DPC models in the test scenarios H2.S4.E2 and H4S.10.E2, which included systematic trends in the annual proportions of fishing trips targeting the alternative habitats (Fig. 6). Whereas nominal

CPUE performance altered between control and test scenarios, medians and variation of the ARE and MD for the DPC models were comparable between control and test scenarios.

Overall, the best-performing DPC models in terms of both accuracy and precision is the PC1.R4 model in the two-habitat scenarios (Fig. 6a-d) and the PC3.R4 model in the four-habitat scenarios (Fig .6 e-h). An exception is that the DPC models with only the first two PCs produced marginally lower ARE medians then DPC models with three PCs in the control scenario H4.S10.E2 (Fig. 6e). The PC1.R4 and PC3.R4 models achieved notably lower MD medians than the nominal CPUE in both control scenarios H2.S4.E1 (Fig. 6b) and H4.S10.E1 (Fig. 6f), respectively, and largely improved the precision in the corresponding test scenarios H2.S4.E2 (Fig. 6d) and H4S10.E2 (Fig. 6h). These models remove a substantial amount of variation in the data while producing fairly accurate abundance trends, over the 20 years period. The results from the four-habitat scenarios indicate that including too many PCs in the DPC models carries a higher risk of introducing bias in the estimated abundance trends than using too few PCs (Fig. 6). In fact, the difference between H4.S10.PC2 and H4.S10.PC3 models was very small, whereas the H4.S10.PC4 models resulted in notable increases of the ARE and MD medians and variation. The choice of transformation had relatively small effects on the ARE and MD medians, but the 4th-root transformation consistently provided slightly better accuracy and precision estimates (Fig. 6).

In general, the optimal DPC models included the number of PCs that was equivalent to the number of simulated habitats minus one (i.e. one and three PCs for the H2.S4 and H4.S10 scenarios, respectively) and were based on a 4th-root transformation of the catch composition data. The performance of the DPC standardization model procedure is illustrated by the simulation runs for control- and test scenario H2.S4.E1 (Fig. 7) and H2.S4.E2 (Fig. 8). In the control scenario, the normalized nominal CPUE indices provide marginally more accurate but less precise estimates of the normalized biomass trends when compared to the PC1.R4 standardization model (Fig. 7). Whereas in the test scenario, the normalized nominal CPUE shows large departures from the normalized biomass trends, while the PC1.R4 model provide accurate and precise abundance trends by effectively adjusting for the time-varying shifts in targeting of the two habitats (Fig. 8).

3.3 Patterns among species

Model performance consistently differed among species. The hypothetical species assemblage varied in two important aspects, namely overall abundance and distribution among habitats. The more complex four-habitat scenario was used to demonstrate patterns in model performance with respect to species abundance and habitat-specific catchability. Relative to the nominal CPUE, the DPC procedure performed better in terms accuracy and precision for the abundant species KOB, GLBK, HAKE and CRPN than for the remaining, less abundant species (Table 3). The nominal CPUE model only outperformed the DPC model for the species SHRK in terms of both metrics, but this was due more to the better performance of the nominal CPUE rather than a failure of the DPC procedure. The species was unusual in that it was more uniformly distributed among the habitats than any other species (Table 1).

3.4 Selection of Principal Components

The Kaiser-Guttmann rule (Eigenvalue > 1) retained only the first PC for 88.5%, 88.5%, 96.0% and 97.5% of the PCA simulations for the two-habitat scenarios H2.S4.E1.R2, H2.S4.E2.R2, H2.S4.E1.R4 and H2.S4.E2.R4, respectively (Fig. 9) and first three PCs for 100% of the PCA simulations for all of the considered four-habitat scenarios (H4.S10.E1.R2, H4.S10.E2.R4, H4.S10.E1.R4 and H4.S10.E2.R4) (Fig. 9). The boxplots in Fig.9, representing the simulated PC Eigenvalues, indicate that the scree-test corroborates these results. For the H2.S4 scenarios, the median Eigenvalues for PC2 to PC4 tend to lie in a fairly straight line, whereas the median Eigenvalue of PC1 departs from this trend (Fig. 9). Similarly, the median Eigenvalues for PC4 to PC10 follow a linear trend for all H4.S10 scenarios, with first departures from this trend evident for the PC3 Eigenvalue (Fig. 9). By contrast, the AIC and BIC always selected the DPC models with the first two PCs for the H2.S4 scenarios and the first four PCs for H4.S10 scenarios, which therefore represented the models that included the highest available number of PCs in a given scenario.

4. Discussion

We have demonstrated the application of the DPC standardization process for simulated multispecies CPUE data sets of different complexity. The simulation results provide strong support for DPC procedure above nominal CPUE in the presence of time-varying trends in targeted fishing effort and habitat-dependent catchability. Even in cases where targeted effort is constant over a time series, the DPC method offers greater precision by removing substantial variation from the data, with a very low penalty on accuracy. By contrast, the use of nominal CPUE as an abundance index was only acceptable in two situations: (1) the catchablity of a species is similar among alternative target habitats (here in the case of SHRK) or (2) targeted effort is uniformly distributed over the time series (control scenarios).

In multispecies fisheries, however, these two situations are rarely the norm (Maunder et al., 2006; Wilberg et al., 2009). Fishes are not randomly distributed across the fishing grounds but tend to have distinct species-specific habitat preferences, which causes catch rates to vary according to the habitat that is targeted (Bigelow et al., 2002; Stephens and MacCall, 2004; Thorson et al., 2012). Occasional exceptions from this might be found, for example, in ubiquitous pelagic species that are picked up by demersal gear or could be artificially induced by improper classification of species into higher-order taxonomic groups (e.g. ‘sharks’ or ‘rays’) or polyphyletic artificial groups (e.g. ‘bycatch’). Systematic long-term changes in targeted effort are commonly indentified in fisheries-dependent catch and effort time series due to a large number of drivers (Wilberg et al., 2009). These include changes in market conditions, fuel price, gear and management regulations, the expansion of the fishery to new target species or fishing grounds, and advances in technology (Pelletier and Ferraris, 2000; Holley and Marchal, 2004; Maunder et al., 2006; Quirijns et al., 2008; Winker et al., 2013).

It remains implicit to always compare the DPC procedure with alternative models that do not adjust for targeting. Furthermore, we advise to inspect the catch composition data for temporal shifts in targeting before accepting abundance indices based on the DPC procedure. A simple indication for time-varying trends in targeted effort can often be inferred by evaluating the yearly catch compositions for systematic changes over the time series. Other, more comprehensive approaches that are useful for detecting systematic trends in targeting include the application of hierarchical agglomerative clustering (Pelletier and Ferraris, 2000; Holley and Marchal, 2004), non-hierarchical clustering methods (Castro et al., 2010; Punzón et al., 2010) and multivariate regression trees (Philippi et al., 1998). Our simulation results show that the nominal CPUE can exhibit considerable departures from the DPC-standardized CPUE (and the ‘true’ biomass) in the presence of shifts in targeted effort over time. Even for cases where the DPC-derived abundance index closely resembles the trends from alternative models, we expect that the DPC procedure will provide an attractive option for removing noise from the data. However, caution is recommended if the DPC-derived index diverges noticeably from alternative models despite no clear evidence of shifts in targeting.

PCA represents one of the most commonly used approaches to describe patterns of variations in multivariate datasets (Jackson, 1993). The idea is that meaningful sources of variation are retained in the first few PC-axes (nontrivial PCs). One of the difficulties will be to select the relevant number of nontrivial PCs, which are meant here to represent meaningful separations of fishing tactics. It was therefore an important objective to evaluate heuristics and potential model selection criteria for selecting the optimal number of PCs for DPC method, given that the number of fishing tactics will often be unclear. We found that the Kaiser-Guttman rule (Eigen values > 1) and the Cattel's scree-test proved to be suitable heuristics for selecting the optimal number of PCs for all our simulation scenarios. However, more comprehensive simulation testings, using a larger variety of simulated scenarios, are warranted. Another aspect that should be addressed in future research is the obvious failure of the common selection criteria AIC and BIC to select optimal number of PCs as covariates. For all simulated data sets, both criteria consistently lead to the selection of the most complex DPC model, which was often associated with a heavy penalty on both accuracy and precision. Similar failures to identify the most appropriate covariate-structure for CPUE standardization models, combined with the tendency to select unnecessary complex models, have also been noted in comparable simulation studies (Carruthers et al., 2010; Lynch et al., 2012). In the case of the DPC procedures, we suspect that the tendency to select overly complex models could be further increased because the information contained in the PC predictor variables might not be entirely independent from the response CPUE. In general, we suggest the Kaiser-Guttman rule (Eigen values > 1) and the Cattel's scree-test as most appropriate among the tested selection criteria. As the inclusion of too few PCs as covariates seem to produce less biased abundance indices than DPC model that include too many, we further recommend a conservative approach for the selection of PCs.

Another positive aspect of the DPC procedure is the adequate performance for CPUE data with high proportions of zero catches. This finding represents an important advancement compared to the initial case study for the DPC procedure (Winker et al., 2013), in which zero CPUE values were simply omitted from the standardization data sets prior to fitting the DPC models with assumed log-normal error distributions. In this specific case, Winker et al. (2013)

argued that the exclusion of zeros was unlikely to have notable impact on the standardized CPUE trends because the two species under assessment represented abundant target species that were likely to be caught when targeted so that only a minor fraction of zeros were caused by failed targeted effort. The risk of excluding zero catches is to produce positively biased, hyper-stable CPUE trends (Kimura, 1981). This is particularly eminent for less abundant and bycatch species for which the information in zero catches is often crucial to estimate accurate abundance trends (Maunder and Punt, 2004; Ortiz and Arocha, 2004; Minami et al., 2007).

Based on the concepts outlined in Martin et al. (2005), we would expect two principal origins of zeros in multi-species catch and effort data: (1) the species is targeted but no catch is made (‘true zeros’) and (2) zero catches are caused non-targeted effort (‘false zeros’), for example, by fishing in habitats that are unlikely to produce catches of the species under assessment. It follows that the proportion of true zeros should decrease with increasing abundance of a species in its prime habitat, where it is directly targeted. High proportions of false zeros will inevitably lead to zero-inflation if the effect of targeting is not accounted for. These mechanisms determining zero catches were effectively simulated by generating nominal CPUE deviates from the Tweedie distribution as function of the product of biomass and habitat-specific catchability. A favourable property of the Tweedie distribution is that it naturally handles zero catches together with right-skewed real numbers, which makes it an attractive error model for the analysis of CPUE data that are expressed in weight (Candy, 2004; Tascheri et al., 2010; Lorance et al., 2011). Whereas discrete count distributions (e.g. Poisson, negative binomial), which also naturally include zeros, could be considered for the DPC procedure when catches are recorded in numbers.

We have demonstrated the importance of accounting for shifts targeted effort for the standardization of multispecies CPUE data. Our operating model was based on the premises that fishing tactic is reflected by the choices of different target habitats and that each habitat is associated with a species-specific catchability (Stephens and MacCall, 2004). This simple concept is fairly general and should, in principal, not only apply to hand-line or rod and line fisheries (Stephens and MacCall, 2004; Winker et al., 2013), but to other multispecies fisheries as well. Many bottom-trawl fisheries, for example, show distinct patterns in the catch composition, which has been often linked with habitat-related factors such as bottom-type and depths strata (Fall et al., 2006; García-Rodríguez et al., 2006; Palmer et al., 2009). In fact, a trawl net will be less selective than hook and line gear since it will sweep up most fish in its path. We predict that the PCs derived from a typically large number of species in the bottom-trawl catch (García-Rodríguez et al., 2006; Palmer et al., 2009; Deporte et al., 2012) should clearly point towards a particular target habitat-type as origin of the catch. Further research could also investigate the applicability of the DPC procedure in long-line CPUE standardization procedures, for which shifts in targeted effort have been addressed by a number of studies (He et al., 1997; Bigelow et al., 2002; Maunder et al., 2006; Carruthers et al., 2010; Carvalho et al., 2010; Lynch et al., 2012)

In conclusion, the DPC standardization procedure provides a reliable method for removing the effects of targeting on multispecies CPUE, and should improve assessments of target and bycatch species alike. An obstacle to its immediate widespread application could be the onerous data requirement in the form of detailed records of catches for all or the majority of species at the trip or drag level. The integrity of bycatch records is often marred by the variable extent of discarding, and it may be necessary to include only those species which are not discarded. The method could provide additional impetus to improve bycatch reporting, as the more species that are included in the catch data, the greater the information content.

Acknowledgements

This work was partly financed under contract by the South African Department of Agriculture, Forestry and Fisheries. Additional funding and support was provided by the Ma-Re BASICS project of the University of Cape Town Vice-Chancellors Strategic Initiative and by the National Research Foundation. Any opinion, findings and conclusions or recommendations expressed in this material are those of the authors. We wish to thank Prof. Doug Butterworth and the members of the members of the review panel for the 2012 International Fisheries Stock Assessment Workshop in Cape Town for guidance and constructive comments on earlier drafts of this work.

5. References

Bigelow, K.A., Hampton, J., Miyabe, N., 2002. Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). Fish. Oceanogr. 11, 143-155.

Candy, S., 2004. Modelling catch and effort data using generalised linear models, the Tweedie distribution, and random vessel effects. CCAMLR Sci. 11, 59-80.

Carruthers, T.R., McAllister, M.K., Ahrens, R.N.M., 2010. Simulating spatial dynamics to evaluate methods of deriving abundance indices for tropical tunas. Can. J. Fish. Aquat. Sci. 67, 1409-1427.

Carvalho, F.C., Murie, D.J., Hazin, F.H.V., Hazin, H.G., Leite-Mourato, B., Travassos, P., Burgess, G.H., 2010. Catch rates and size composition of blue sharks (Prionace glauca) caught by the Brazilian pelagic longline fleet in the southwestern Atlantic Ocean. Aquat. Living Resour. 23, 373-385.

Castro, J., Punzón, A., Pierce, G.J., Marín, M., Abad, E., 2010. Identification of métiers of the Northern Spanish coastal bottom pair trawl fleet by using the partitioning method CLARA. Fish. Res. 102, 184-190.

Cattell, R.B., 1966. The scree test for the number of factors. Journal of Multivariate Behavioral Research 1, 245-276.

Deporte, N., Ulrich, C., Mahévas, S., Demanèche, S., Bastardie, F., 2012. Regional métiers definition: a comparative investigation of statistical methods using a workflow applied to international otter trawl fisheries in the North Sea. ICES J. Mar. Sci. 69, 331-342.

Dunn, P., Smyth, G., 2005. Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing 15, 267-280.

Fall, M., Samba, A., LaloÃ«, F., 2006. Fishing tactics and strategies in coastal demersal trawling fisheries in Senegal. Aquat. Living Resour. 19, 307-316.

García-Rodríguez, M., Fernández, Á.M., Esteban, A., 2006. Characterisation, analysis and catch rates of the small-scale fisheries of the Alicante Gulf (SE Spain) over a 10 years time series. Fish. Res. 77, 226-238.

Glazer, J.P., Butterworth, D.S., 2002. GLM-based standardization of the catch per unit effort series for the South Africna west coast hake, focusing on adjustment for targeting other species. Afr. J. Mar. Sci. 24, 323-339.

Guisan, A., Edwards, T.C., Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol. Model. 157, 89-100.

Guttman, L., 1954. Some necessary conditions for common factor analysis. Psychometrika 19, 149-161.

Harley, S.J., Myers, R.A., Dunn, A., 2001. Is catch-per-unit-effort proportional to abundance? Can. J. Fish. Aquat. Sci. 58, 1760-1772.

He, X., Bigelow, K.A., Boggs, C.H., 1997. Cluster analysis of longline sets and fishing strategies within the Hawaii-based fishery. Fish. Res. 31, 147-158.

Hinton, M.G., Maunder, M.N., 2003. Methods for standardizing CPUE and how to select among them. ICCAT SCRS/2003/034, 1-11.

Holley, J.-F., Marchal, P., 2004. Fishing strategy development under changing conditions: examples from the French offshore fleet fishing in the North Atlantic. ICES J. Mar. Sci. 61, 1410-1431.

Jackson, D.A., 1993. Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. Ecology 74, 2204-2214.

Kimura, D.K., 1981. Standardized measures of relative abundance based on modelling log (c.p.u.e.), and their application to Pacific ocean perch (Sebastes alutus). Journal du Conseil 39, 211-218.

Lo, N.C., Jacobson, L.D., Squire, J.L., 1992. Indices of relative abundance for fish spotter data based on delta-lognormal models. Can. J. Fish. Aquat. Sci. 49, 2515-2526.

Lorance, P., Pawlowski, L., Trenkel, V.M., 2011. Standardizing blue ling landings per unit effort from industry haul-by-haul data using generalized additive models. ICES J. Mar. Sci.

Lynch, P.D., Shertzer, K.W., Latour, R.J., 2012. Performance of methods used to estimate indices of abundance for highly migratory species. Fish. Res. 125-126, 27-39.

Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J., Possingham, H.P., 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. Ecol. Lett. 8, 1235-1246.

Maunder, M.N., Punt, A., 2004. Standardizing catch and effort data: a review of recent approaches. Fish. Res. 70, 171-195.

Maunder, M.N., Sibert, J.R., Fonteneau, A., Hampton, J., Kleiber, P., Harley, S.J., 2006. Interpreting catch per unit effort data to assess the status of individual stocks and communities. ICES J. Mar. Sci. 63, 1373-1385.

Minami, M., Lennert-Cody, C.E., Gao, W., Román-Verdesoto, M., 2007. Modeling shark bycatch: the zero-inflated negative binomial regression model with smoothing. Fish. Res. 84, 210-221.

Ono, K., Punt, A.E., Rivot, E., 2012. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fish. Res. 125, 173-183.

Ortiz, M., Arocha, F., 2004. Alternative error distribution models for standardization of catch rates of non-target species from a pelagic longline fishery: billfish species in the Venezuelan tuna longline fishery. Fish. Res. 70, 275-297.

Palmer, M., Quetglas, A., Guijarro, B., Moranta, J., Ordines, F., Massutí, E., 2009. Performance of artificial neural networks and discriminant analysis in predicting fishing tactics from multispecific fisheries. Can. J. Fish. Aquat. Sci. 66, 224-237.

Pelletier, D., Ferraris, J., 2000. A multivariate approach for defining fishing tactics from commercial catch and effort data. Can. J. Fish. Aquat. Sci. 57, 51-65.

Philippi, T.E., Dixon, P.M., Taylor, B.E., 1998. Detecting trends in species composition. Ecol. Appl. 8, 300-308.

Punt, A., Walker, T.I., Taylor, B.L., Pribac, F., 2000. Standardization of catch and effort data in a spatially-structured shark fishery. Fish. Res. 45, 129-145.

Punzón, A., Hernández, C., Abad, E., Castro, J., Pérez, N., Trujillo, V., 2010. Spanish otter trawl fisheries in the Cantabrian Sea. ICES J. Mar. Sci. 67, 1604-1616.

Quirijns, F.J., Poos, J.J., Rijnsdorp, A.D., 2008. Standardizing commercial CPUE data in monitoring stock dynamics: Accounting for targeting behaviour in mixed fisheries. Fish. Res. 89, 1-8.

R Development Core Team, 2011. *R*: A language and environment forstatistical computing. R Foundation for Statistical Computing, Vienna.

Shono, H., 2008. Application of the Tweedie distribution to zero-catch data in CPUE analysis. Fish. Res. 93, 154-162.

Stephens, A., MacCall, A., 2004. A multispecies approach to subsetting logbook data for purposes of estimating CPUE. Fish. Res. 70, 299-310.

Su, N.-J., Yeh, S.-Z., Sun, C.-L., Punt, A.E., Chen, Y., Wang, S.-P., 2008. Standardizing catch and effort data of the Taiwanese distant-water longline fishery in the western and central Pacific Ocean for bigeye tuna, *Thunnus obesus*. Fish. Res. 90, 235-246.

Tascheri, R., Saavedra-Nievas, J.C., Roa-Ureta, R., 2010. Statistical models to standardize catch rates in the multi-species trawl fishery for Patagonian grenadier (Macruronus magellanicus) off Southern Chile. Fish. Res. 105, 200-214.

Thorson, J.T., Stewart, I.J., Punt, A.E., 2012. Development and application of an agent-based model to evaluate methods for estimating relative abundance indices for shoaling fish such as Pacific rockfish (Sebastes spp.). ICES J. Mar. Sci. 69, 635-647.

Wilberg, M.J., Thorson, J.T., Linton, B.C., Berkson, J., 2009. Incorporating Time-Varying Catchability into Population Dynamic Stock Assessment Models. Rev. Fish. Sci. 18, 7-24.

Winker, H., Kerwath, S.E., Attwood, C.G., 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fish. Res. 139, 118-131.

Table 1

Specified values of catchability coefficients *q* for each species by habitat. The dashed box represents the reduced matrix of *q* values used for the two-habitat four species simulation scenarios.

	Target habitats			
Species	Shallow sediment	Deep sediment	Shallow reef	Deep reef
KOB	1.00	0.02	0.01	0.01
GLBK	0.90	0.10	0.01	0.01
HAKE	0.05	1.00	0.01	0.01
PANG	0.10	0.80	0.01	0.01
CRPN	0.01	0.01	0.10	1.00
SNTR	0.01	0.01	1.00	0.05
ROMN	0.01	0.01	0.80	0.20
DGRD	0.01	0.01	0.70	0.20
RSTM	0.01	0.01	0.90	0.30
SHRK	0.50	0.05	0.30	0.05

Table 2

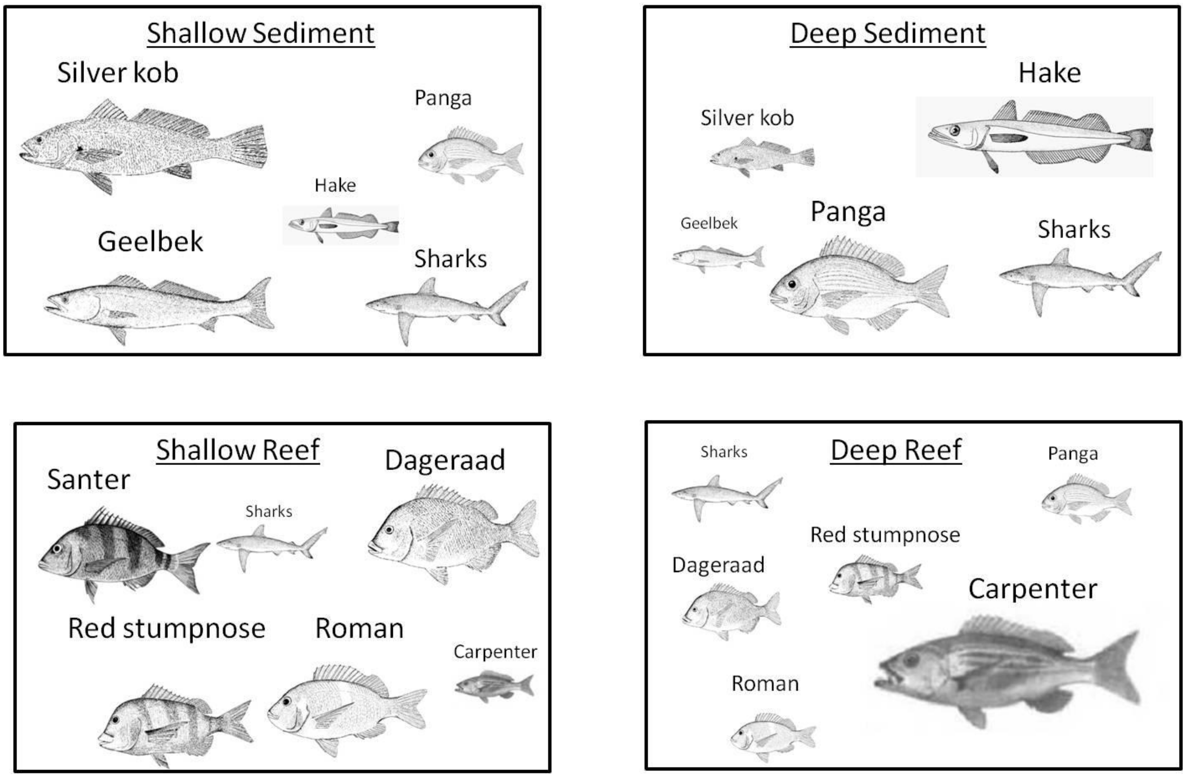
Statistics summarizing the 200 simulation data sets by species generated for four different test and control scenarios. The summary includes the proportion of zero records, averaged mean CPUE values, averaged maximum CPUE values for each species, with numbers in brackets denoting the 2.5th and 97.5th percentiles for the 200 simulation runs.

Scenario	Species	Proportion zeros	mean CPUE	Maximum CPUE
H2.S4.E1	KOB	0.20 [0.05, 0.45]	74.2 [14.7, 210.6]	582.93 [207.9, 1399.3]
	GLBK	0.10 [0.01, 0.37]	80.1 [17.2, 267.7]	585.95 [188.3, 1757.0]
	HAKE	0.14 [0.02, 0.35]	79.8 [17.8, 288.9]	485.26 [178.0, 1448.0]
	PANG	0.46 [0.21, 0.74]	15.1 [2.7, 41.5]	164.10 [71.3, 409.4]
H2.S4.E2	KOB	0.22 [0.03, 0.45]	75.9 [13.8, 249.8]	593.66 [198.3, 1620.7]
	GLBK	0.15 [0.01, 0.40]	73.5 [13.77, 238.0]	550.95 [206.1, 1532.6]
	HAKE	0.14 [0.02, 0.34]	84.6 [18.3, 248.5]	509.12 [187.8, 1457.2]
	PANG	0.46 [0.19, 0.72]	15.7 [3.0, 48.7]	168.20 [69.3, 435.8]
H4.S10.E1	KOB	0.59 [0.39, 0.75]	33.3 [6.8, 125.7]	543.5 [191.6, 1704.4]
	GLBK	0.49 [0.28, 0.69]	39.6 [7.4, 138.6]	573.5 [189.4, 1693.7]
.	HAKE	0.53 [0.33, 0.70]	39.5 [7.7, 120.8]	499.5 [152.9, 1421.0]
	PANG	0.68 [0.47, 0.85]	8.5 [1.5, 31.1]	162.6 [66.1, 450.4]
	CRPN	0.50 [0.30, 0.70]	40.2 [5.8, 140.2]	480.9 [157.6, 1283.4]
	SNTR	0.71 [0.54, 0.86]	8.5 [1.6, 26.7]	208.1 [84.0, 508.8]
	ROMN	0.65 [0.43, 0.85]	9.8 [1.5, 30.8]	205.2 [69.3, 481.4]
	DGRD	0.66 [0.43, 0.84]	8.4 [1.6, 27.9]	182.4 [78.0, 432.8]
	RSTM	0.62 [0.41, 0.82]	11.0 [1.9, 35.0]	219.2 [70.5, 527.2]
	SHRK	0.61 [0.31, 0.83]	7.9 [1.5, 32.8]	143.0 [61.4, 404.7]
H4.S10.E1	KOB	0.57 [0.40, 0.74]	35.6 [6.6, 97.8]	558.5 [195.8, 1486.5]
	GLBK	0.49 [0.30, 0.69]	35.9 [7.1, 95.7]	532.7 [194.4, 1486.4]
	HAKE	0.53 [0.38, 0.71]	36.2 [6.4, 98.0]	455.9 [158.7, 1248.7]
	PANG	0.68 [0.51, 0.85]	7.8 [1.6, 27.1]	152.7 [60.5, 397.7]
	CRPN	0.50 [0.33, 0.68]	38.0 [7.2, 116.0]	465.8 [161.8, 1250.6]
	SNTR	0.72 [0.55, 0.86]	8.8 [1.5, 26.2]	219.3 [77.2, 534.0]
	ROMN	0.66 [0.46, 0.84]	8.7 [1.7, 28.5]	192.7 [84.4, 506.3]
	DGRD	0.68 [0.46, 0.85]	7.6 [1.4, 27.7]	173.2 [65.9, 400.4]
	RSTM	0.62 [0.39, 0.84]	11.1 [1.7, 34.5]	218.9 [73.1, 516.9]
	SHRK	0.62 [0.32, 0.82]	7.5 [1.5, 25.7]	139.0 [58.0, 376.2]

Table 3

Comparisons of differences in ARE and MD medians by species between the nominal CPUE and the best-performing DPC model (PC3.R4) for control and test simulations of the four-habitat scenario (H4.S10). Positive values indicate a better performance of the DPC model.

Scenario	Species	Δ ARE (Nom - DPC)		Δ MD (Nom - DPC)	
		Control	Test	Control	Test
H4.S10	KOB	-0.10	0.66	0.61	1.53
	GLBK	-0.08	0.70	0.59	1.40
.	HAKE	-0.11	0.75	0.63	1.56
	CRPN	-0.30	0.41	0.43	1.15
	PANG	-0.16	0.48	0.44	1.38
	SNTR	-0.32	0.71	0.34	1.31
	ROMN	-0.17	0.45	0.35	0.98
	DGRD	-0.16	0.38	0.33	0.95
	RSTM	-0.12	0.42	0.37	0.93
	SHRK	-0.10	-0.03	0.10	0.30


Shallow Sediment

Deep Sediment

Shallow Reef

Deep Reef

Fig. 1. The size of the fish image illustrates the habitat preference of fish species considered for the two-habitat four species (H2.S10) and the four-habitat ten species (H4.S10) multispecies fisheries scenarios. The H2.S4 scenarios are based on four species that are distributed across shallow- and deepwater habitats, excluding sharks (upper panel). For the H4.S10 scenarios is extended to ten target species, which are distributed across the four different habitats: shallow- and deep water sediment and shallow- and deep water reef.

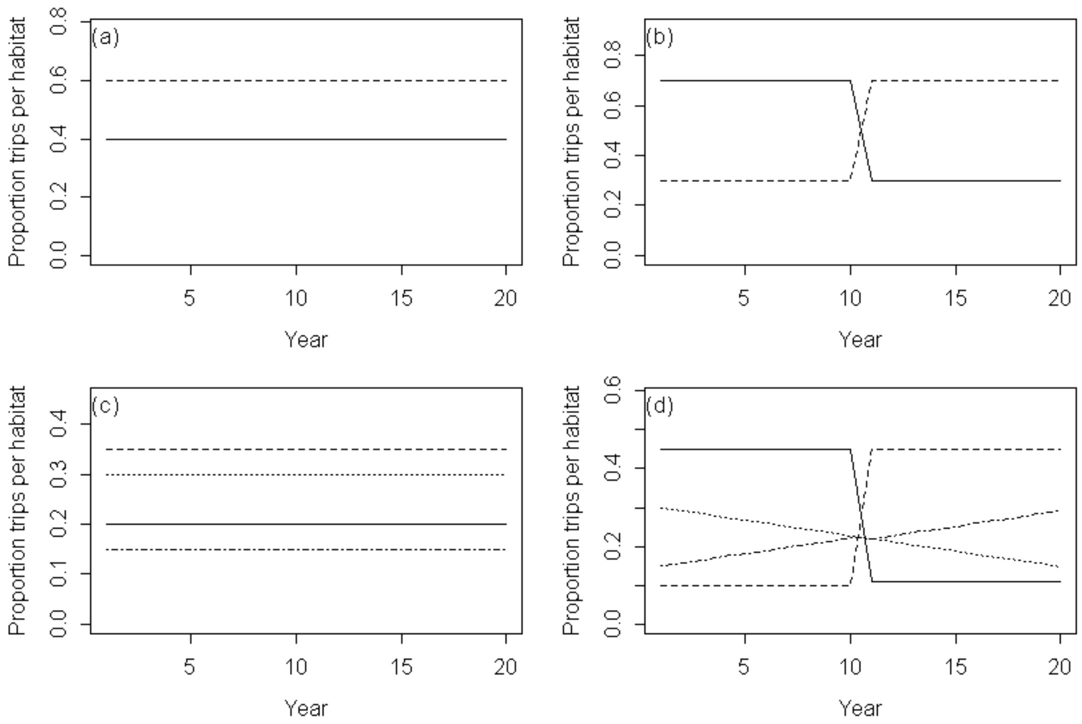


Fig. 2. Underlying trends of the yearly proportion fishing trips directed to two (a) – (b) and four (c) – (d) alternative target habitats. The control scenarios H2.S4.E1 (a) and H4.S10.E2 (c) are based on constant targeted-effort, whereas time-varying trends of targeted-effort were implemented for test scenarios H2.S4.E2 (b) and H4.S10.E2 (d).

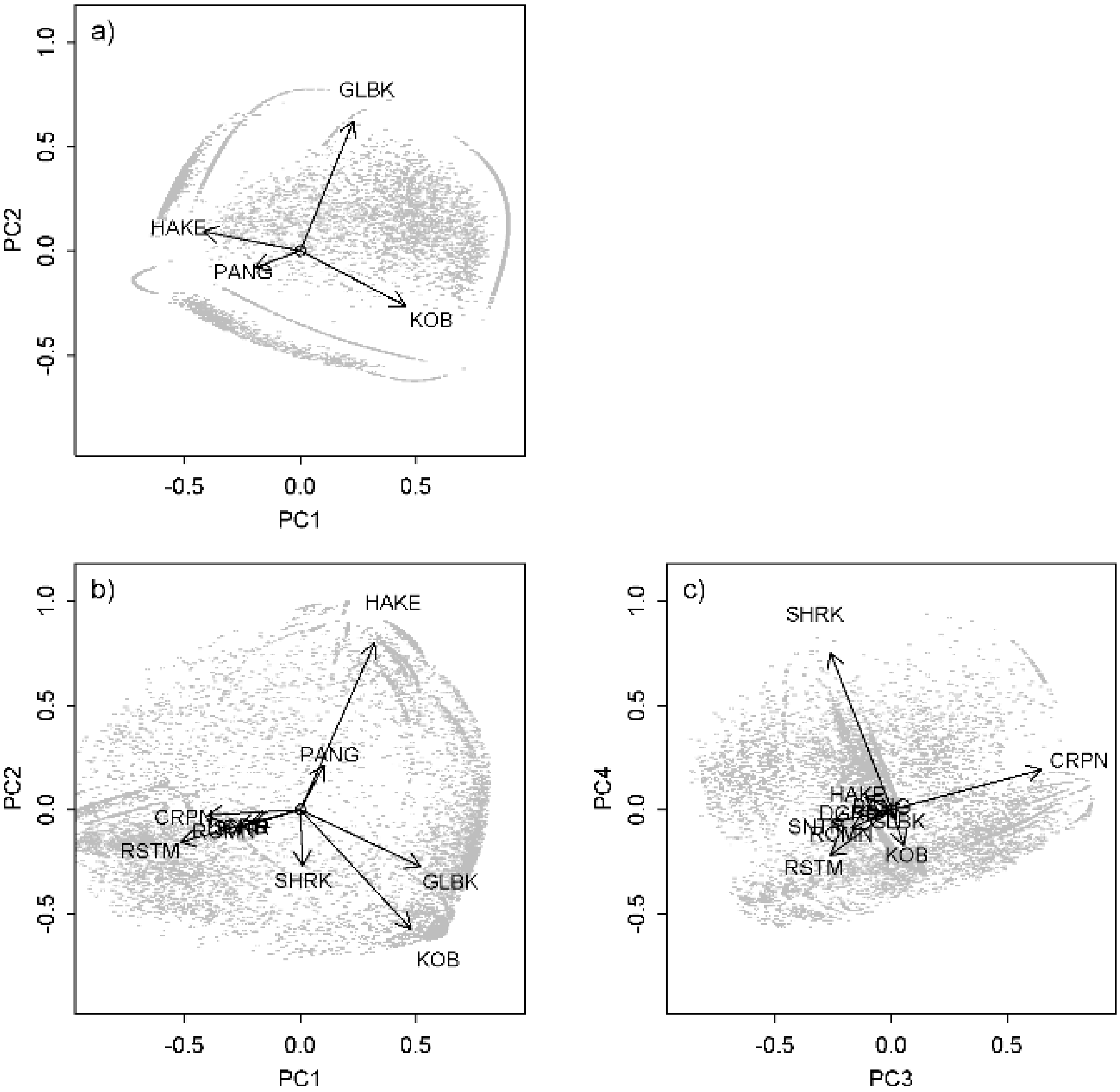


Fig. 3. Examples of biplots illustrating the results from two selected Principal Components Analyses (PCA) of simulated 4th-root transformed catch composition matrixes based on randomly chosen simulation runs for H2.S4.E2 scenario (a) and the H4.S10.E2 scenario (b) – (c), respectively.

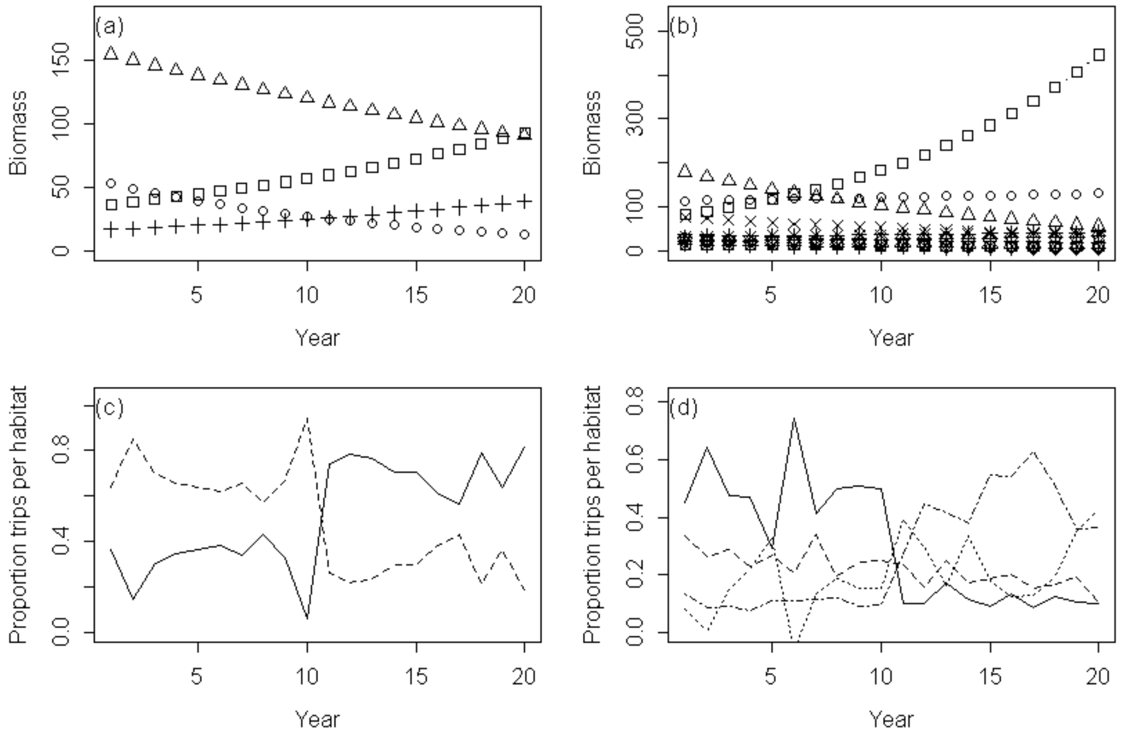


Fig. 4. Randomly selected simulation replicates of generated biomass trajectories of four species for the H2.S4 scenario (a) and of ten species for the H4.S10 scenario (b); and simulated random noise associated with the underlying trends of the yearly proportion fishing trips directed to two (c) and four (d) alternative target habitats for the test scenarios H2.S4.E2 and H4.S10.E2, respectively.

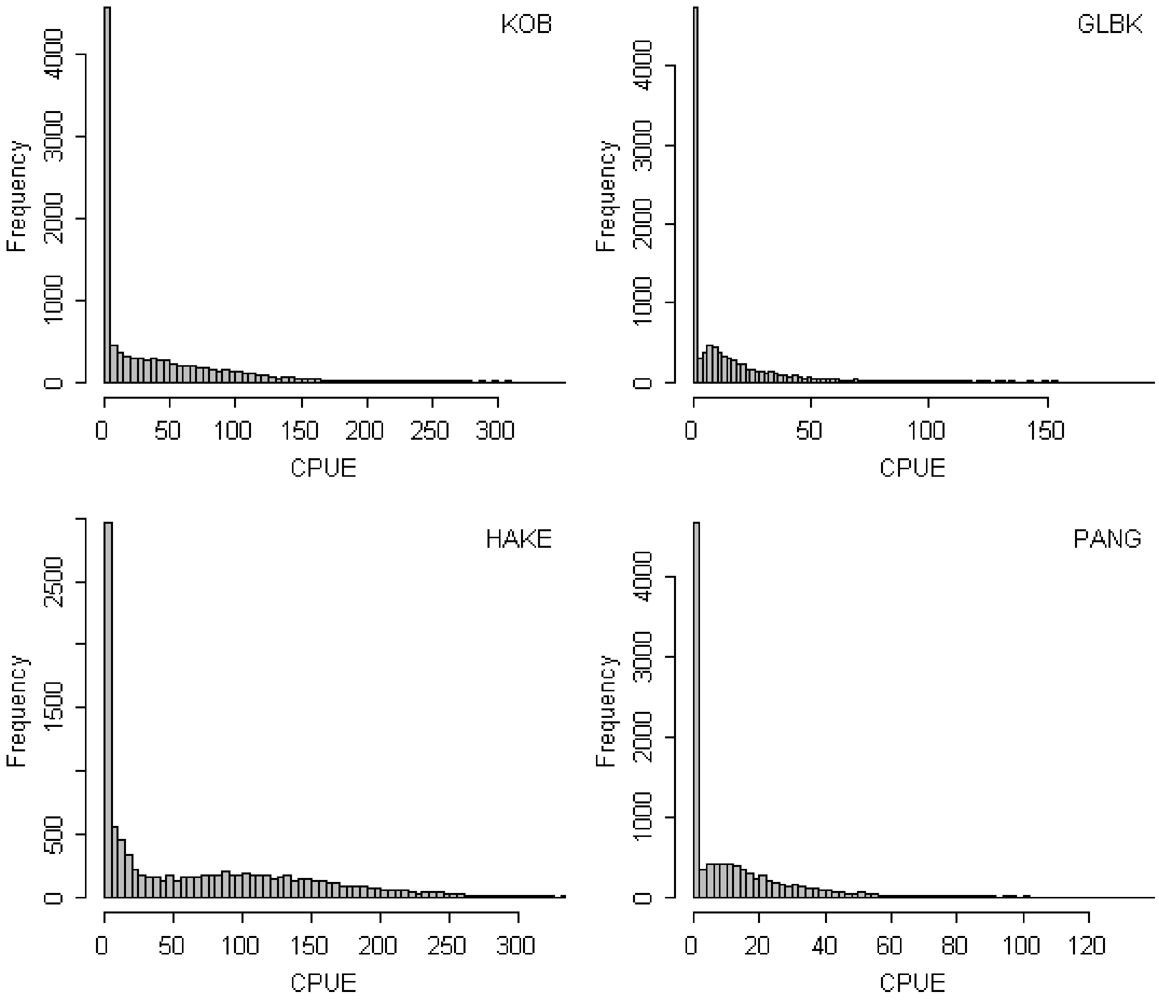


Fig. 5. Simulated frequency data of nominal CPUE records of for four species that are generated by randomly chosen simulation replicates for the test scenario H2.S4.E2.

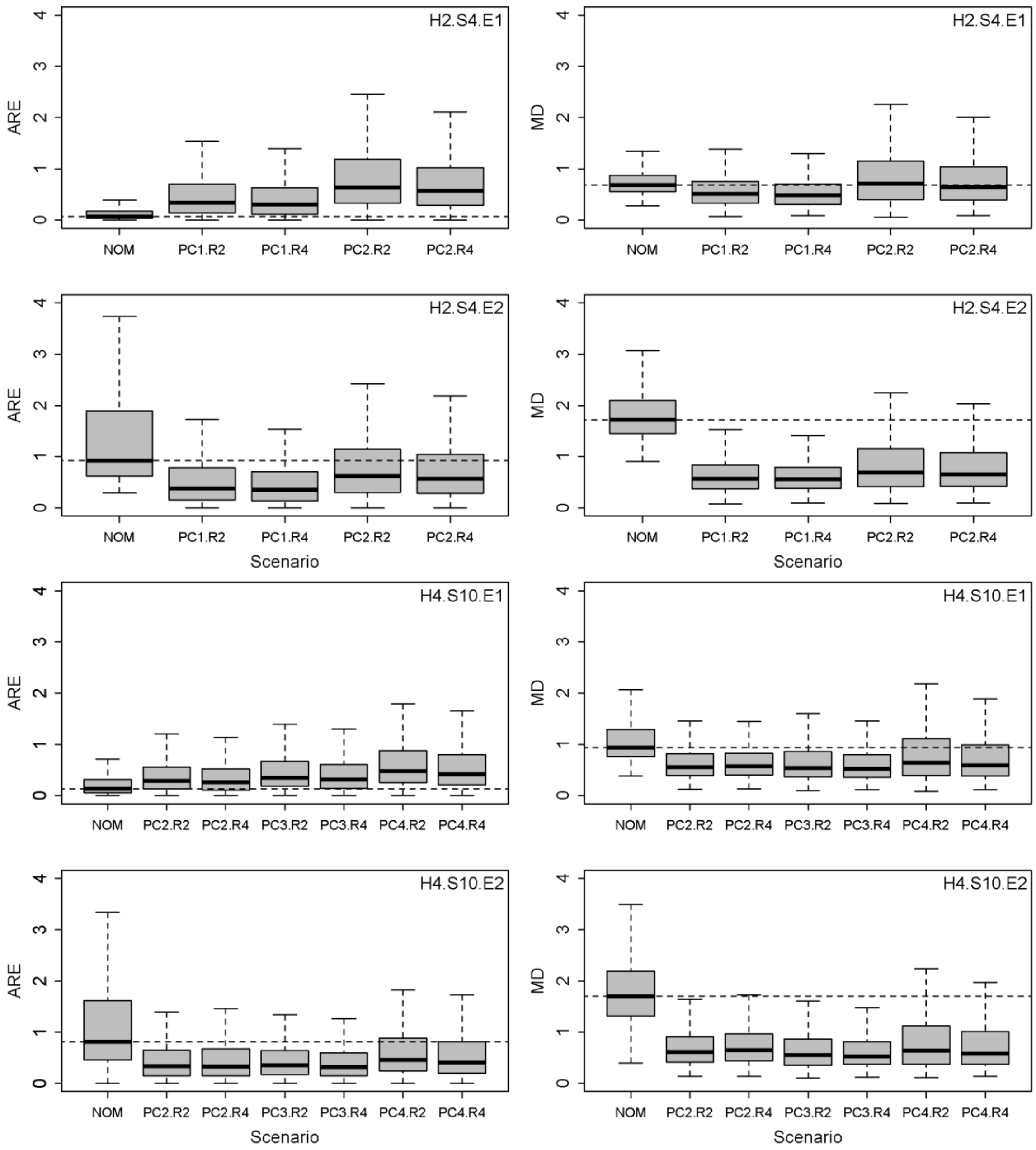


Fig. 6. Boxplots of Absolute Relative Errors (ARE) and mean deviations (MD), averaged over all species, for each scenario and standardization model, with NOM denoting the nominal CPUE. Dashed lines indicate the median ARE and MD estimated for the nominal CPUE (NOM).

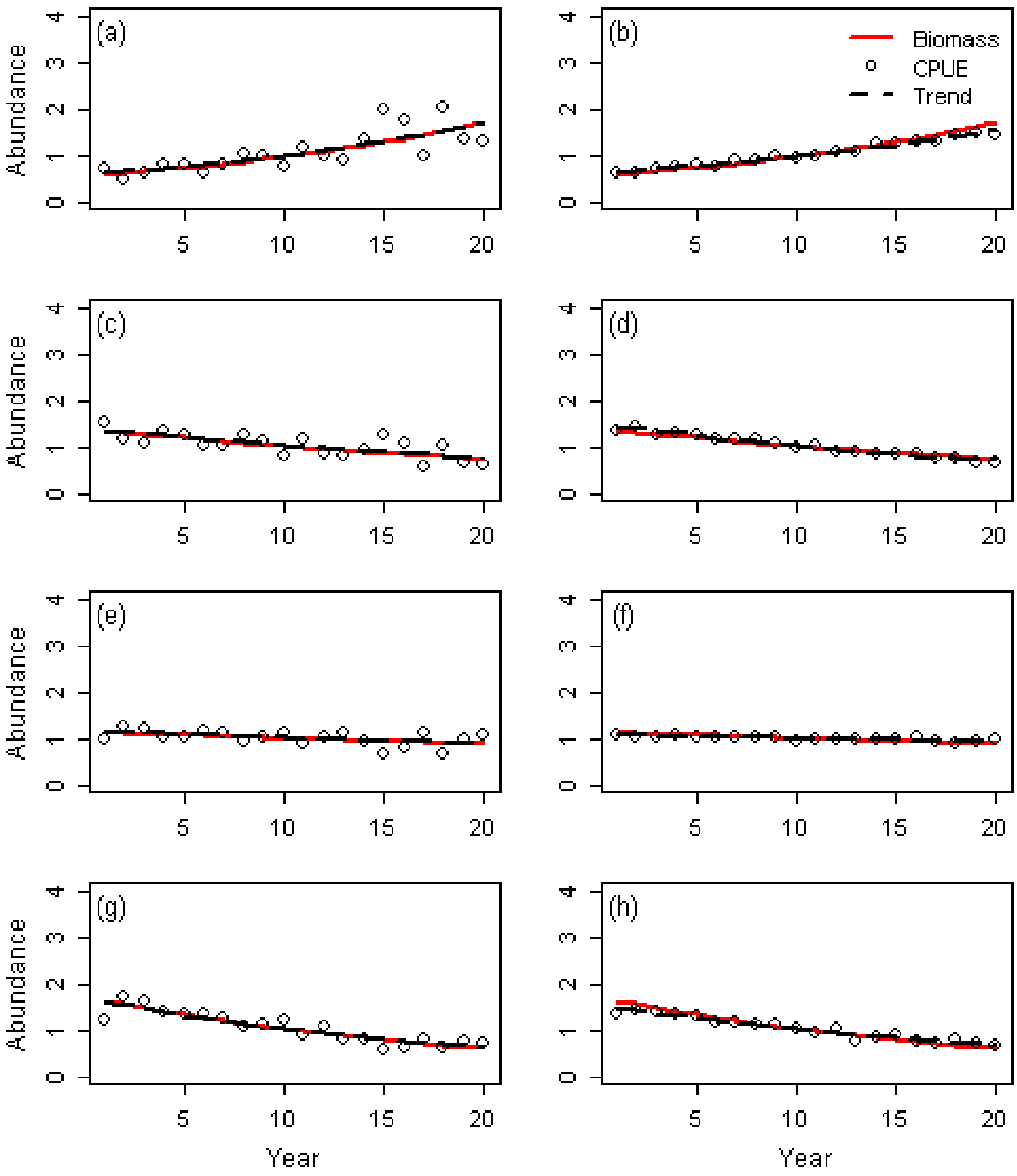


Fig. 7. Normalized Simulated 'true' biomass trajectories of KOB (a) – (b), GLBK (c) – (d), HAKE (e) – (f) and PANG (g) – (h) for the two-habitat four species control scenario (H2.S4.E1) in comparison to normalized nominal CPUE (left panel) and standardized CPUE (right panel) using PC1.R4 GAMs. The dashed lines show trends derived from simple log-linear regressions fitted to the CPUE estimates.

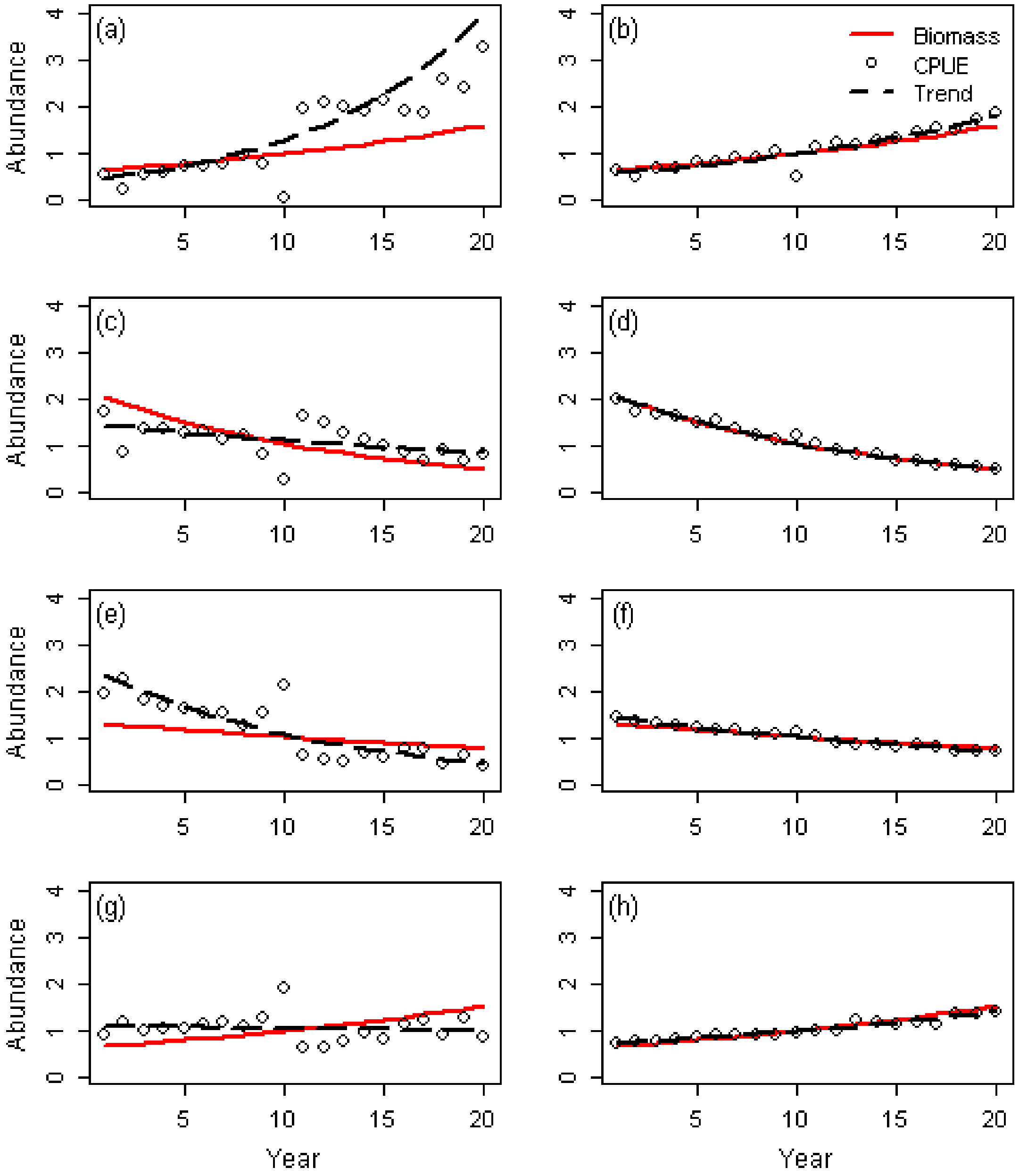


Fig. 8. Normalized Simulated ‘true‘ biomass trajectories of KOB (a) – (b), GLBK (c) – (d), HAKE (e) – (f) and PANG (g) – (h) for the two-habitat four species test scenario (H2.S4.E2) in comparison to normalized nominal CPUE (left panel) and standardized CPUE (right panel) using PC1.R4 GAMs. The dashed lines show trends derived from simple log-linear regressions fitted to the CPUE estimates.

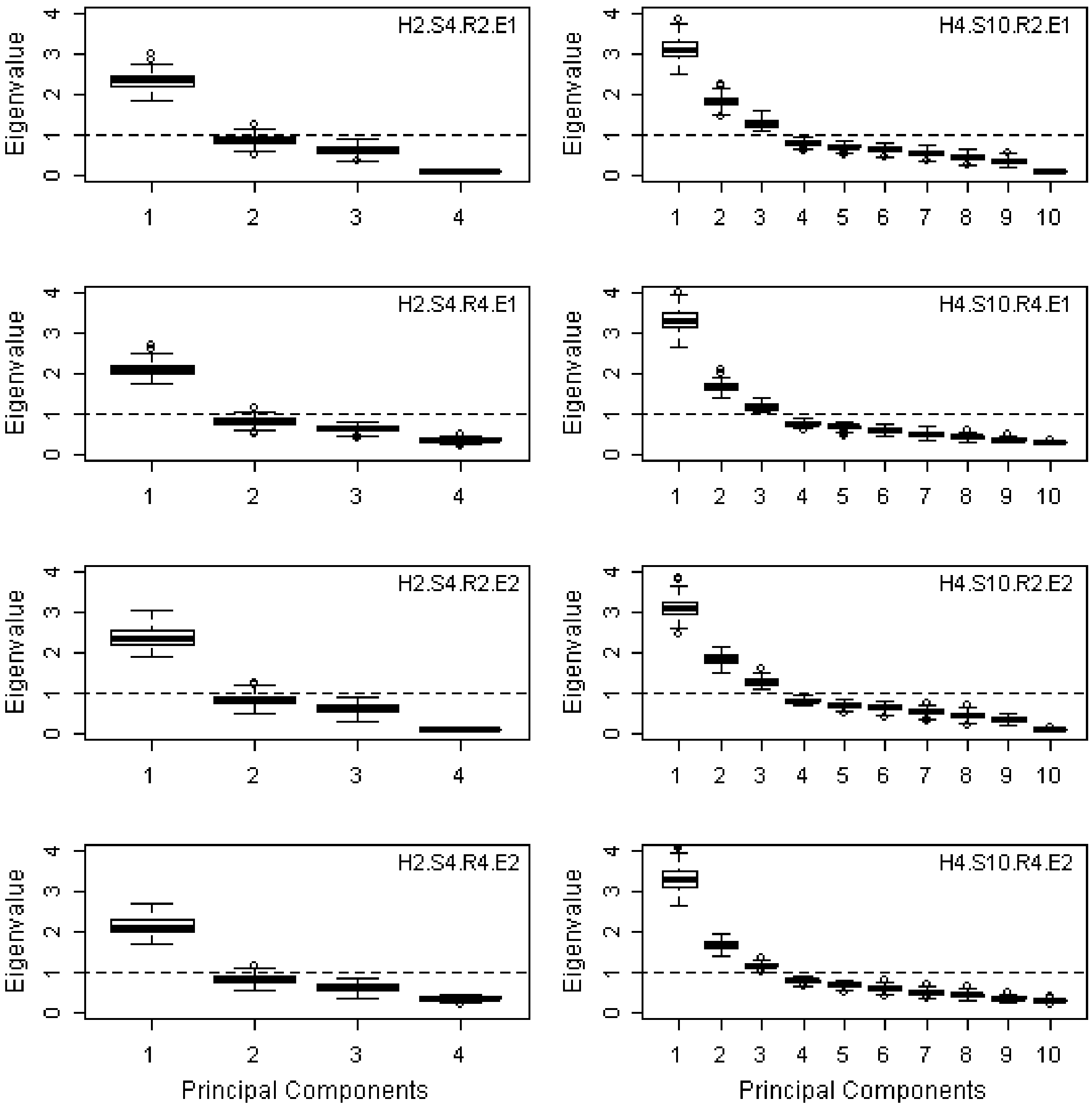


Fig.9. Boxplots of Eigenvalues for Principal Components (PCs) for all simulation scenarios.