

## APPLICATION

# EcoEnsemble: A general framework for combining ecosystem models in R

Michael A. Spence  | James A. Martindale  | Michael J. Thomson 

Centre for Environment, Fisheries and Aquaculture Science, Lowestoft, UK

## Correspondence

Michael A. Spence

Email: [michael.spence@cefas.gov.uk](mailto:michael.spence@cefas.gov.uk)

## Funding information

Centre for the Environment, Fisheries and Aquaculture Science Seedcorn, Grant/Award Number: DP436

Handling Editor: Edward Codling

## Abstract

1. Often there are several complex ecosystem models available to address a specific question. However, structural differences, systematic discrepancies and uncertainties mean that they typically produce different outputs. Rather than selecting a single 'best' model, it is desirable to combine them to give a coherent answer to the question at hand.
2. Many methods of combining ecosystem models assume that one of the models is exactly correct, which is unlikely to be the case. Furthermore, models may not be fitted to the same data, have the same outputs, nor be run for the same time period, making many common methods difficult to implement. In this paper, we use a statistical model to describe the relationship between the ecosystem models, prior beliefs and observations to make coherent predictions of the true state of the ecosystem with robust quantification of uncertainty.
3. We introduce EcoEnsemble, an R package that takes advantage of the statistical model's structure to efficiently fit the ensemble model, either sampling from the posterior distribution or maximising the posterior density.
4. We demonstrate EcoEnsemble by investigating what would happen to four fish species in the North Sea under future management scenarios. Although developed for applications in ecology, EcoEnsemble can be used to combine any group of mechanistic models, for example in climate modelling, epidemiology or biology.

## KEYWORDS

Bayesian statistics, ecosystem management, ecosystem models, ensemble modelling, R, uncertainty analysis

## 1 | INTRODUCTION

Ecosystem models are being increasingly used to support policy decisions (Hyder et al., 2015; Nielsen et al., 2018). Often, there are several such models available, each providing different outputs with potentially different implications for management actions. While

some models are better at capturing some aspects of the ecosystem than others, in general no model is uniformly better than the others (Chandler, 2013; Nilsen et al., 2022), and often the average of models outperforms all of the models individually (Rougier, 2016). Management decisions on the basis of a single model are therefore sensitive to the choice of model (Collie et al., 2016; Essington

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

& Plagányi, 2013; Fulton et al., 2003; Gaichas, 2008; Hart & Fay, 2020). Furthermore, by ignoring other available models, the amount of information utilised is limited, leading to an increase in uncertainty (Spence et al., 2018), often to a greater extent than formally recognised.

Instead of choosing one model, it is possible to combine them using an ensemble model. However, many such methods, such as weighting schemes (e.g. Bayesian model averaging, Banner & Higgs, 2017; Dormann et al., 2018), explicitly assume that one of the models correctly describe the truth (Chandler, 2013). Not only is this a strong assumption, but treating a suite of models as containing the true model tends to underestimate the uncertainty because alternative, possibly as yet undeveloped, models could give predictions outside of the current range (Chandler, 2013). Under this scheme, adding a new model to the suite could increase the uncertainty of the ensemble despite the additional information contained in the new model (Dormann et al., 2018).

Applying the above methods to complex ecosystem models is not straightforward. Often these models are fitted to different data (Ianelli et al., 2016), describe different processes and produce outputs that are not directly comparable (Spence et al., 2018). At the same time, complex ecosystem models often share many similarities. They can have similar or even identical processes and forcing inputs (e.g. Tittensor et al., 2018), or can be fitted using similar data, suggesting that the discrepancies in the models are not independent (Christiansen, 2021; Knutti, 2010; Rougier et al., 2013).

Spence et al. (2018) developed an ensemble model adopting a Bayesian approach that treats the individual models as exchangeable and drawn from a population of possible models. They separated the discrepancy between the models and the true variables of interest (Vol) into discrepancy shared between all of the models and discrepancies that are specific to each model.

By modelling in this way, the ensemble model exploits the strengths whilst discounting the weaknesses of each individual simulator (Chandler, 2013) while rigorously quantifying the uncertainty allowing for better management decisions (Spence, Griffiths, et al., 2021).

In this paper, we introduce EcoEnsemble, an R package that implements the ensemble model of Spence et al. (2018) for time-series outputs of complex models, where the models output linear combinations of the Vol. We introduce EcoEnsemble and the ensemble model in Section 2 and then demonstrate it on case study of fishing in the North Sea in Section 3.

## 2 | USING ECOENSEMBLE

EcoEnsemble is an R package for fitting and sampling from the ensemble model described in Spence et al. (2018). It is available on CRAN (<https://cran.r-project.org/web/packages/EcoEnsemble/index.html>) and GitHub (<https://github.com/CefasRepRes/EcoEnsemble>). To avoid confusion, we henceforth refer to complex ecosystem models as 'simulators'.

## 2.1 | Ensemble model overview

We are interested in the value of  $d$  Vol,  $\mathbf{y}^{(t)} = (y_1^{(t)}, \dots, y_d^{(t)})'$  for  $t = 1, \dots, T$ . In the absence of any other information, the Vol evolves according to a random walk,

$$\mathbf{y}^{(t)} \sim N(\mathbf{y}^{(t-1)}, \Lambda_y). \quad (1)$$

We have  $m$  simulators that describe the Vol. Not all of the simulators output all of the Vol over the full time period. To accommodate these differences, Spence et al. (2018) introduced a latent variable, known as the 'best guess',  $\mathbf{x}_k^{(t)}$ , which represents simulator  $k$ 's output if it described all  $d$  Vol at time  $t$  with no parameter uncertainty.

For each simulator, the 'best guess' is the Vol plus a discrepancy term  $\zeta_k^{(t)}$  (Kennedy & O'Hagan, 2001),

$$\mathbf{x}_k^{(t)} = \mathbf{y}^{(t)} + \zeta_k^{(t)}. \quad (2)$$

The discrepancy term is split between discrepancies that are shared by all of the simulators and discrepancies that are specific to the  $k$ th simulator. The shared discrepancy is split between the long-term shared discrepancy,  $\delta$ , and the short-term shared discrepancy,  $\boldsymbol{\eta}^{(t)}$ . Simulator  $k$ 's individual discrepancy is split between long-term individual discrepancy,  $\boldsymbol{\gamma}_k$ , and short-term individual discrepancy,  $\mathbf{z}_k^{(t)}$ , that is

$$\zeta_k^{(t)} = \delta + \boldsymbol{\eta}^{(t)} + \boldsymbol{\gamma}_k + \mathbf{z}_k^{(t)}. \quad (3)$$

The long-term individual discrepancies are normally distributed so that the long-term individual discrepancy of the  $k$ th simulator is

$$\boldsymbol{\gamma}_k \sim N(\mathbf{0}, C_\gamma). \quad (4)$$

The short-term discrepancy terms,  $\boldsymbol{\eta}^{(t)}$  and  $\mathbf{z}_k^{(t)}$ , follow stationary autoregressive models of order one, with autoregressive parameters  $R_\eta$  and  $R_k$ , and covariance parameters  $\Lambda_\eta$  and  $\Lambda_k$ , respectively.

A summary of the ensemble model is found in Table 1. For more details on the model, see Spence et al. (2018) or the 'EcoEnsemble' vignette in EcoEnsemble.

## 2.2 | Data requirements

In EcoEnsemble, the user requires:

- observations of the Vol ( $\hat{\mathbf{y}}^{(t)}$ ) and a covariance matrix describing observational uncertainty ( $\Sigma_y$ ).
- outputs from each of the simulators ( $\hat{\mathbf{x}}_k^{(t)}$ ) and covariance matrices describing the parameter uncertainty of each simulator ( $\Sigma_k$ ).
- prior beliefs about the discrepancy terms.

### 2.2.1 | Observations

Observations of the Vol at some times are included as an input. The ensemble model says noisy observations of the Vol are

**TABLE 1** A summary of the variables in the ensemble model. Values of  $n_k$  and  $M_k$  are simulator specific.

Variable	Dimensions	Description	Relationship
$\mathbf{y}^{(t)}$	$d$	Variables of interest	$\mathbf{y}^{(t)} \sim N(\mathbf{y}^{(t-1)}, \Lambda_y)$
$\hat{\mathbf{y}}^{(t)}$	$d$	Noisy observation of $\mathbf{y}^{(t)}$	$\hat{\mathbf{y}}^{(t)} \sim N(\mathbf{y}^{(t)}, \Sigma_y)$
$\delta$	$d$	Long-term shared discrepancy	
$\boldsymbol{\eta}^{(t)}$	$d$	Short-term shared discrepancy	$\boldsymbol{\eta}^{(t)} \sim N(R_\eta \boldsymbol{\eta}^{(t-1)}, \Lambda_\eta)$
$\boldsymbol{\mu}^{(t)}$	$d$	Simulator consensus	$\boldsymbol{\mu}^{(t)} = \mathbf{y}^{(t)} + \delta + \boldsymbol{\eta}^{(t)}$
$\gamma_k$	$d$	Simulator $k$ 's long-term individual discrepancy	$\gamma_k \sim N(\mathbf{0}, C_\gamma)$
$\mathbf{z}_k^{(t)}$	$d$	Simulator $k$ 's short-term individual discrepancy	$\mathbf{z}_k^{(t)} \sim N(R_k \mathbf{z}_k^{(t-1)}, \Lambda_k)$
$\mathbf{x}_k^{(t)}$	$d$	Simulator $k$ 's best guess	$\mathbf{x}_k^{(t)} = \boldsymbol{\mu}^{(t)} + \gamma_k + \mathbf{z}_k^{(t)}$
$\hat{\mathbf{x}}_k^{(t)}$	$n_k$	The expectation of simulator $k$ 's output $\mathbf{x}_k^{(t)}$	$\hat{\mathbf{x}}_k^{(t)} \sim N(M_k \mathbf{x}_k^{(t)}, \Sigma_k)$

$$\hat{\mathbf{y}}^{(t)} \sim N(\mathbf{y}^{(t)}, \Sigma_y), \quad (5)$$

whenever there are observations at time  $t$ .

The observations could be direct, noisy observations of the Vol, with some measure of uncertainty (e.g. survey data; Spence et al., 2018), or outputs from a statistical model that estimates the Vol (e.g. a state-space model, Auger-Méthé et al., 2021), with uncertainty (e.g. Spence et al., 2022; Spence, Griffiths, et al., 2021).

## 2.2.2 | Simulator outputs

Simulators should output linear combinations of the Vol. They do not need to produce outputs for the whole period, but a given simulator should output the same variables for each time that it does. The simulator outputs are noisy linear combinations of the 'best guess'. If the  $k$ th simulator is evaluated at time  $t$ , its output is

$$\hat{\mathbf{x}}_k^{(t)} \sim N(M_k \mathbf{x}_k^{(t)}, \Sigma_k), \quad (6)$$

where  $\Sigma_k$  reflects the parameter uncertainty of the  $k$ th simulator and  $M_k$  is an  $n_k \times d$  matrix relating the 'best guess' to the simulator output. If, for all simulators, Vol are either present or not, then the `fit_ensemble_model()` function infers  $M_k$  for  $k = 1, \dots, m$ . If other transformations of the Vol are outputted from the simulators, then the  $M_k$  matrices should be specified by the user.

Parameter uncertainty for simulator outputs should also be quantified in the covariance matrices  $\Sigma_k$ . This can be calculated, for example, by calibrating the simulator to empirical data (e.g. Spence, Thorpe, et al., 2021), or by projecting a distribution of elicited parameter values (e.g. Ducharme-Barth & Vincent, 2022; Thorpe et al., 2015).

## 2.2.3 | Prior beliefs

To include prior beliefs in EcoEnsemble, the user should place priors on  $m + 3$  different covariance matrices: the covariance of the random walk of the Vol  $\Sigma_y$ ; the covariance of the auto-regressive process on the shared short-term discrepancy,  $\Lambda_\eta$ ; the covariance of the long-term individual discrepancies,  $C_\gamma$ ; and the covariances of the auto-regressive processes on the individual short-term discrepancies,  $\Lambda_1, \dots, \Lambda_m$ . The prior on the long-term shared discrepancy,  $\delta$ , also needs to be specified.

For  $\Sigma_y$ , the only choice for prior distribution is an inverse-Wishart distribution.

In EcoEnsemble, we follow the approach of Spence et al. (2018) and decompose each of the discrepancy covariance matrices  $C \in \{\Lambda_\eta, \Lambda_1, \dots, \Lambda_m, C_\gamma\}$ , in the form

$$C = \Pi P \Pi, \quad (7)$$

where  $\Pi = \text{diag}(\pi_1, \dots, \pi_d)$  is a diagonal matrix of standard deviations, and  $P$  is a correlation matrix.

In EcoEnsemble, for  $C_\gamma$  and  $\Lambda_\eta$ , there are three available prior distributions for covariance matrices. In each case, the variance terms are parameterised by a gamma distribution, while the correlation matrices can be parameterised either by an LKJ distribution (Lewandowski et al., 2009), an inverse-Wishart distribution, or by beta distributions, using the method of concordance (Gokhale & Press, 1982; Zondervan-Zwijenburg et al., 2017). For  $\Lambda_k$ , for  $k = 1, \dots, m$ , there is an additional option of a hierarchical prior.

The prior on the long-term shared discrepancy is parameterised by a zero-centred normal distribution,

$$\delta \sim N(\mathbf{0}, \sigma_\delta^2 I_d), \quad (8)$$

so the user should specify the standard deviation  $\sigma_\delta$ . The user can also specify priors on the autoregressive parameters  $R_k$  for  $k = \eta, 1, \dots, m$ ,

$$\frac{R_k + 1}{2} \sim \text{beta}(\alpha_k, \beta_k). \quad (9)$$

In EcoEnsemble, priors are encoded by an EnsemblePrior object. For example:

```
priors <- EnsemblePrior(4,
  ind_st_params = IndSTPrior("hierarchical", list(-3, 1, 8, 4),
    list(0.1, 0.1, 0.1, 0.1), AR_params=c(2, 2)),
  ind_lt_params = IndLTPrior("lkj", list(1, 1), 1),
  sha_st_params = ShaSTPrior("lkj", list(1, 10), 1, AR_params=c(2, 2)),
  sha_lt_params = 5)
```

Further details of the available priors are given in Section S1 of the Supporting Information. For more information on eliciting priors, see O'Hagan et al. (2006).

## 2.3 | Fitting the model

The model described in Table 1 can be written as a dynamical linear model (see Supporting Information), meaning that the latent variables,  $\theta^{(t)} = (\mathbf{y}^{(t)}, \boldsymbol{\eta}^{(t)}, \mathbf{z}_1^{(t)}, \dots, \mathbf{z}_m^{(t)})'$  for all  $t$ , can be integrated out of the likelihood. EcoEnsemble uses the sequential Kalman filter (Chui & Chen, 2009; Kalman, 1960) to fit the ensemble model (see the Supporting Information for more details). The highest posterior density can be found or a full sample can be generated using the No U-turn Hamiltonian Monte Carlo (Hoffman & Gelman, 2011) in the package rstan (Stan Development Team, 2020).

The model is fit using the `fit_ensemble_model()` function, which creates an EnsembleFit object containing samples from Markov Chain Monte Carlo or point estimates of all of the parameters. For example:

```
fit <- fit_ensemble_model(
  observations = list(SSB_obs, Sigma_obs),
  simulators = list(list(SSB_ewe, Sigma_ewe, "EwE"),
                    list(SSB_lm, Sigma_lm, "LeMans"),
                    list(SSB_miz, Sigma_miz, "mizer"),
                    list(SSB_fs, Sigma_fs, "FishSUMS")),
  priors = priors)
```

This code uses datasets included in EcoEnsemble. See Section 3 for more details.

## 2.4 | Model outputs

EcoEnsemble includes functionality for sampling  $\theta^{(t)}$ , as well as the most likely  $\theta^{(t)}$ , from an EnsembleFit object. For each set of parameters from a fitted model, we calculate the most likely latent variables and sample them using the algorithm developed by Strickland et al. (2009), based on Durbin and Koopman (2002).

`generate_sample()` creates an EnsembleSample object, which contains samples of  $\theta^{(t)}$ . Calling the `plot()` function with an EnsembleSample object plots the Vol,  $\mathbf{y}^{(t)}$ , the simulators outputs,  $\hat{\mathbf{x}}_k^{(t)}$ , and the observations,  $\hat{\mathbf{y}}^{(t)}$ , for all  $t$  and  $k$  (see Figure 2).

```
samples <- generate_sample(fit)
plot(samples)
```

## 3 | CASE STUDY

We demonstrate EcoEnsemble using an example from fisheries. We investigated the change in spawning stock biomass (SSB) of four species in the North Sea, Norway pout *Trisopterus esmarkii*, Atlantic herring *Clupea harengus*, Atlantic cod *Gadus morhua* and common sole *Solea solea* if fished at maximum sustainable yield (MSY), a management strategy that maximises the long-term yield for each species,

from 2017 (Spence, Griffiths, et al., 2021). We used single-species stock assessments and four multispecies simulators to predict the SSB from 1984 to 2050 under this scenario. The data for this study are available in EcoEnsemble.

### 3.1 | Observations

Noisy observations of the SSB were taken from single-species stock assessments (ICES, 2020a, 2020b). Uncertainty provided in stock assessments formed the diagonal elements of  $\Sigma_y$ , with the off-diagonal elements being zero.

### 3.2 | Simulators

SSB was also predicted using four simulators:

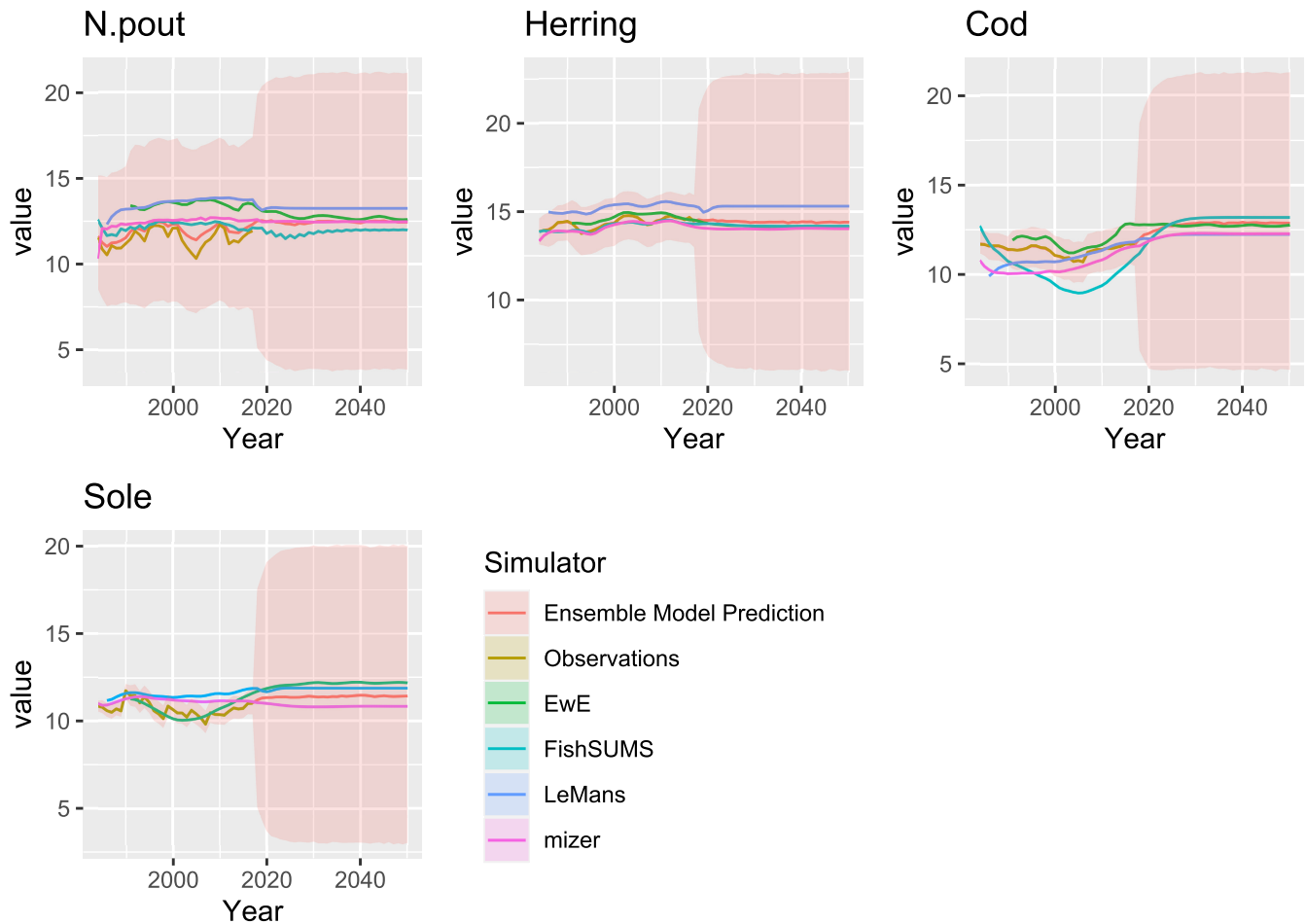
1. **EcoPath with EcoSim** (EwE) is a simulator with 60 functional groups for the North Sea (ICES, 2016). It outputs all four species from 1991 until 2050, meaning that  $n_1 = 4$  and  $M_1 = I_4$ . The covariance of EwE's output was calculated in Mackinson et al. (2018).
2. **LeMans** is a discrete time length-based model that describes growth and predation (Thorpe et al., 2015). It outputs all four species from 1986 until 2050, meaning that  $n_2 = 4$  and  $M_2 = I_4$ . Its covariance was calculated in Thorpe et al. (2015).
3. **mizer** is a size-based model that describes ontogenetic feeding and growth, mortality and reproduction driven by size-dependent predation and maturation processes (Blanchard et al., 2014). It outputs all four species from 1984 until 2050, meaning that  $n_3 = 4$  and  $M_3 = I_4$ , and its covariance was calculated in Spence et al. (2016).
4. **FishSUMS** is a discrete time length-based model that describes growth, density-dependent mortality and losses due to fishing and predation by explicitly modelled species, and seasonal reproduction (Speirs et al., 2016). It outputs Norway pout, Atlantic herring and Atlantic cod, that is it does not output sole, from 1984 until 2050, meaning that  $n_4 = 3$  and that

$$M_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (10)$$

Its covariance was calculated in Spence et al. (2018).

### 3.3 | Prior information

We used the default priors in EcoEnsemble to fit the model. See Supporting Information and Exploring priors for more information. The result of the prior predictive of the ensemble model is shown in Figure 1 using the `prior_ensemble_model()` and `sample_prior_mode()` functions.



**FIGURE 1** A plot of the natural logarithm of the prior predictive spawning-stock biomass of each species considered in the study, as predicted by each simulator, single-species assessments and the ensemble model. The edges of the shaded area represent the 5% and 95% quantiles of the prior predictive SSB.

### 3.4 | Results and discussion

The ensemble model prediction, simulator outputs and stock assessment estimates are shown in Figure 2. The probability that the SSB of each species increases is given in Table 2.

The simulators appear to capture the general trend correctly for most species but often struggle to capture short-term dynamics. For example, in the single-species model, low recruitment for Norway pout in 2003/2004 led to lower SSB in 2005 (ICES, 2020b), something that none of the simulators were able to capture. The ensemble model is able to account for similar unexpected changes by having large uncertainty in predictions of the future. However, the overall trend is for Norway pout to remain at a similar level as before, as suggested by all of the simulators.

For Herring, all of the simulators predicted a decrease in the biomass after 2017, and the ensemble model agreed. The ensemble model predicts a continued recovery of cod into the future under fishing at MSY (Table 2). EwE was able to capture the trends in the observed SSB and therefore contributes to the ensemble model predictions. For sole, none of the models predicted a large variation in the SSB, and so there was a large uncertainty in the ensemble model predictions.

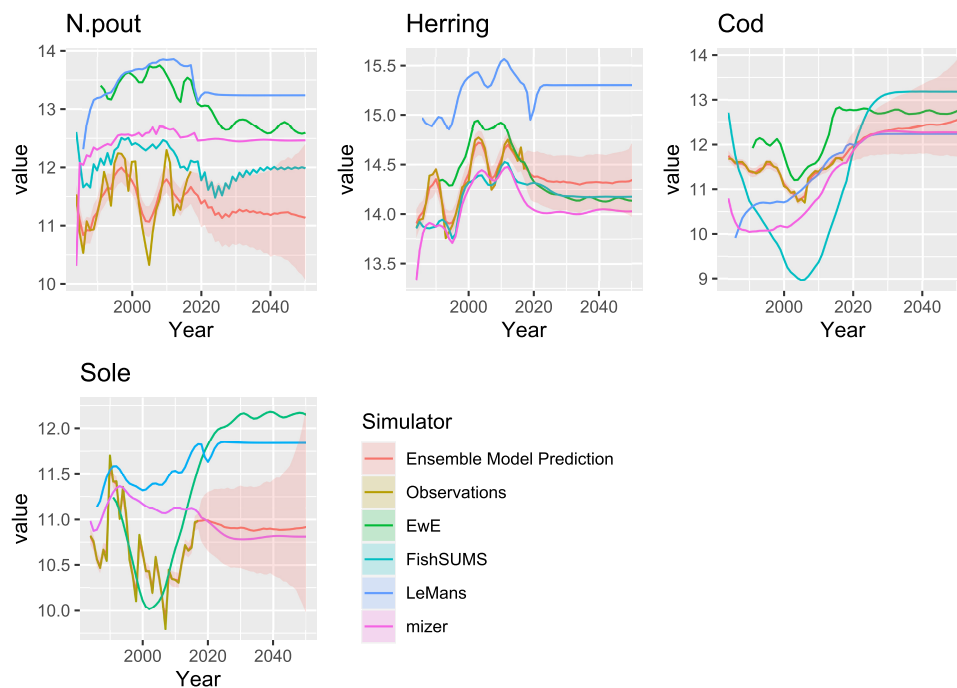
**TABLE 2** The ensemble model probability that the spawning stock biomass for species in 2050 is larger than the year in the first column.

Year	Norway pout	Herring	Cod	Sole
2000	0.158	0.378	0.983	0.726
2005	0.537	0.120	0.991	0.858
2010	0.156	0.177	0.975	0.873
2017	0.202	0.261	0.951	0.449

This study was completed as a demonstration of the EcoEnsemble package. See Spence et al. (2020) or <https://rconnect.cefas.co.uk/content/22> for the results of a more extensive study.

## 4 | CONCLUSIONS

Complex models are rapidly emerging as an invaluable tool for managing natural resources (Hyder et al., 2015). However, when models



**FIGURE 2** A plot of the posterior predictive natural logarithm of the spawning-stock biomass of each species considered in the study, as predicted by each simulator, single-species assessments and the ensemble model. The edges of the shaded area represent the 5% and 95% quantiles of the posterior predictive SSB.

produce different, and sometimes incomparable, outputs to one another, a single, clear interpretation for the purposes of decision-making can be difficult to find.

Spence et al. (2018) developed a statistical framework to combine outputs from multiple complex models by describing how these outputs relate to the Vol. EcoEnsemble is an R package that gives user-friendly way to configure priors, fit and sample from this ensemble model. We hope that this leads to an increase in the use of this model whenever multiple complex models are available.

While this framework was originally designed to answer questions in fisheries science, the tools presented here are general purpose and can be used in other disciplines, such as climate science (Spence et al., 2022) and epidemiology.

#### AUTHOR CONTRIBUTIONS

Micahel A. Spence and James A. Martindale conceived the ideas behind EcoEnsemble, designed the methodology and led the writing of the manuscript; James A. Martindale and Michael A. Spence developed EcoEnsemble, with contributions from Michael J. Thompson. All authors contributed critically to the drafts and gave final approval for publication.

#### ACKNOWLEDGEMENTS

This work was supported by the Centre for the Environment, Fisheries and Aquaculture Science Seedcorn project DP436 'Operationalising ensemble modelling for multispecies stock assessments'. The authors would also like to thank Robert Thorpe for comments on earlier versions of the paper.

#### CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

#### PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14148>.

#### DATA AVAILABILITY STATEMENT

EcoEnsemble can be downloaded from GitHub (<https://github.com/CefasRepRes/EcoEnsemble>) and the Comprehensive R Archive Network CRAN (<https://cran.r-project.org/web/packages/EcoEnsemble/index.html>) (Spence et al., 2023).

#### ORCID

Michael A. Spence  <https://orcid.org/0000-0002-3445-7979>

James A. Martindale  <https://orcid.org/0000-0002-1913-5592>

Michael J. Thomson  <https://orcid.org/0000-0003-0284-0129>

#### REFERENCES

- Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., & Thomas, L. (2021). A guide to state-space modeling of ecological time series. *Ecological Monographs*, 91, e01470.
- Banner, K. M., & Higgs, M. D. (2017). Considerations for assessing model averaging of regression coefficients. *Ecological Applications*, 27, 78–93.
- Blanchard, J. L., Andersen, K. H., Scott, F., Hintzen, N. T., Piet, G., & Jennings, S. (2014). Evaluating targets and trade-offs among fisheries and conservation objectives using a multispecies size spectrum model. *Journal of Applied Ecology*, 51, 612–622.



- Chandler, R. E. (2013). Exploiting strength, discounting weakness: Combining information from multiple climate simulators. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371, 20120388.
- Christiansen, B. (2021). Understanding the distribution of multimodel ensembles. *Journal of Climate*, 33, 9447–9465.
- Chui, C. K., & Chen, G. (2009). *Kalman filtering with real-time applications* (4th ed.). Springer.
- Collie, J. S., Botsford, L. W., Hastings, A., Kaplan, I. C., Largier, J. L., Livingston, P. A., Plagányi, E., Rose, K. A., Wells, B. K., & Werner, F. E. (2016). Ecosystem models for fisheries management: Finding the sweet spot. *Fish and Fisheries*, 17, 101–125.
- Dormann, C. F., Calabrese, J. M., Guillaera-Arroita, G., Matechou, E., Bahn, V., Barton, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J. J., Pollock, L. J., Reineking, B., Roberts, D. R., Schröder, B., Thuiller, W., Warton, D. I., ... Hartig, F. (2018). Model averaging in ecology: A review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological Monographs*, 88, 485–504.
- Ducharme-Barth, N. D., & Vincent, M. T. (2022). Focusing on the front end: A framework for incorporating uncertainty in biological parameters in model ensembles of integrated stock assessments. *Fisheries Research*, 255, 106452.
- Durbin, J., & Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89, 603–615.
- Essington, T. E., & Plagányi, E. E. (2013). Pitfalls and guidelines for “recycling” models for ecosystem-based fisheries management: Evaluating model suitability for forage fish fisheries. *ICES Journal of Marine Science*, 71, 118–127.
- Fulton, E., Smith, A., & Johnson, C. (2003). Effect of complexity of marine ecosystem models. *Marine Ecology Progress Series*, 253, 1–16.
- Gaichas, S. K. (2008). A context for ecosystem-based fishery management: Developing concepts of ecosystems and sustainability. *Marine Policy*, 32, 393–401.
- Gokhale, D. V., & Press, S. J. (1982). Assessment of a prior distribution for the correlation coefficient in a bivariate normal distribution. *Journal of the Royal Statistical Society: Series A (General)*, 145, 237–249.
- Hart, A. R., & Fay, G. (2020). Applying tree analysis to assess combinations of ecosystem-based fisheries management actions in management strategy evaluation. *Fisheries Research*, 225, 105466.
- Hoffman, M., & Gelman, A. (2011). The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1593–1623.
- Hyder, K., Rossberg, A. G., Allen, J. I., Austen, M. C., Barciela, R. M., Bannister, H. J., Blackwell, P. G., Blanchard, J. L., Burrows, M. T., Defriez, E., Dorrington, T., Edwards, K. P., Garcia-Carreras, B., Heath, M. R., Hembury, D. J., Heymans, J. J., Holt, J., Houle, J. E., Jennings, S., ... Paterson, D. M. (2015). Making modelling count—Increasing the contribution of shelf-seas community and ecosystem models to policy development and management. *Marine Policy*, 61, 291–302.
- Ianelli, J., Holsman, K. K., Punt, A. E., & Aydin, K. (2016). Multi-model inference for incorporating trophic and climate uncertainty into stock assessments. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 379–389. Understanding Ecosystem Processes in the Eastern Bering Sea IV.
- ICES. (2016). *Working Group on Multispecies Assessment Methods (WGSAM)*. Technical report. International Council for Exploration of the Seas.
- ICES. (2020a). *Herring Assessment Working Group for the Area South of 62°N (HAWG)*. Technical report, ICES scientific reports. 2:60. 1151 pp. ICES.
- ICES. (2020b). *Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak*. Technical report, ICES scientific reports. 3:66. 1281 pp. ICES.
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82, 35–45.
- Kennedy, M. C., & O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 425–464.
- Knutti, R. (2010). The end of model democracy? *Climatic Change*, 102, 395–404.
- Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100, 1989–2001.
- Mackinson, S., Platts, M., Garcia, C., & Lynam, C. (2018). Evaluating the fishery and ecological consequences of the proposed North Sea multi-annual plan. *PLoS ONE*, 13, 1–23.
- Nielsen, J. R., Thunberg, E., Holland, D. S., Schmidt, J. O., Fulton, E. A., Bastardie, F., Punt, A. E., Allen, I., Bartelings, H., Bertignac, M., Bethke, E., Bossier, S., Buckworth, R., Carpenter, G., Christensen, A., Christensen, V., Da-Rocha, J. M., Deng, R., Dichmont, C., ... Waldo, S. (2018). Integrated ecological-economic fisheries models—evaluation, review and challenges for implementation. *Fish and Fisheries*, 19, 1–29.
- Nilsen, I., Hansen, C., Kaplan, I., Holmes, E., & Langangen, O. (2022). Exploring the role of Northeast Atlantic cod in the barents sea food web using a multi-model approach. *Fish and Fisheries*, 23, 1083–1098.
- O'Hagan, A., Buck, C. E., Daneshkhan, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., & Rakow, T. (2006). *Uncertain judgments: Eliciting experts' probabilities*. Wiley.
- Rougier, J. (2016). Ensemble averaging and mean squared error. *Journal of Climate*, 29, 8865–8870.
- Rougier, J., Goldstein, M., & House, L. (2013). Second-order exchangeability analysis for multimodel ensembles. *Journal of the American Statistical Association*, 108, 852–863.
- Speirs, D., Greenstreet, S., & Heath, M. (2016). Modelling the effects of fishing on the North Sea fish community size composition. *Ecological Modelling*, 321, 35–45.
- Spence, M. A., Alliji, K., Bannister, H. J., Walker, N. D., & Muench, A. (2020). Fish should not be in isolation: Calculating maximum sustainable yield using an ensemble model.
- Spence, M. A., Blackwell, P. G., & Blanchard, J. L. (2016). Parameter uncertainty of a dynamic multispecies size spectrum model. *Canadian Journal of Fisheries and Aquatic Sciences*, 73, 589–597.
- Spence, M. A., Blanchard, J. L., Rossberg, A. G., Heath, M. R., Heymans, J. J., Mackinson, S., Serpetti, N., Speirs, D. C., Thorpe, R. B., & Blackwell, P. G. (2018). A general framework for combining ecosystem models. *Fish and Fisheries*, 19, 1031–1042.
- Spence, M. A., Griffiths, C. A., Waggitt, J. J., Bannister, H. J., Thorpe, R. B., Rossberg, A. G., & Lynam, C. P. (2021). Sustainable fishing can lead to improvements in marine ecosystem status: An ensemble-model forecast of the north sea ecosystem. *Marine Ecology Progress Series*, 680, 207–221.
- Spence, M. A., Lynam, C. P., Thorpe, R. B., Heneghan, R. F., & Dolder, P. J. (2022). Synthesizing empirical and modelling studies to predict past and future primary production in the North Sea. *Frontiers in Marine Science*, 9, 828623.
- Spence, M. A., Martindale, J. A., & Thompson, M. J. (2023). Cefasrepres/ecoensemble: Initial release. <https://doi.org/10.5281/zenodo.7950258>
- Spence, M. A., Thorpe, R. B., Blackwell, P. G., Scott, F., Southwell, R., & Blanchard, J. L. (2021). Quantifying uncertainty and dynamical changes in multi-species fishing mortality rates, catches and biomass by combining state-space and size-based multi-species models. *Fish and Fisheries*, 22, 667–681.
- Stan Development Team. (2020). RStan: The R interface to Stan. R package version 2.19.3.
- Strickland, C. M., Turner, I. W., Denham, R., & Mengersen, K. L. (2009). Efficient Bayesian estimation of multivariate state space models. *Computational Statistics & Data Analysis*, 53, 4116–4125.

- Thorpe, R. B., Le Quesne, W. J. F., Luxford, F., Collie, J. S., & Jennings, S. (2015). Evaluation and management implications of uncertainty in a multispecies size-structured model of population and community responses to fishing. *Methods in Ecology and Evolution*, 6, 49–58.
- Tittensor, D. P., Eddy, T. D., Lotze, H. K., Galbraith, E. D., Cheung, W., Barange, M., Blanchard, J. L., Bopp, L., Bryndum-Buchholz, A., Büchner, M., Bulman, C., Carozza, D. A., Christensen, V., Coll, M., Dunne, J. P., Fernandes, J. A., Fulton, E. A., Hobday, A. J., Huber, V., ... Walker, N. D. (2018). A protocol for the intercomparison of marine fishery and ecosystem models: Fish-MIP v1.0. *Geoscientific Model Development*, 11, 1421–1442.
- Zondervan-Zwijenburg, M., van de Schoot-Hubeek, W., Lek, K., Hoijtink, H., & van de Schoot, R. (2017). Application and evaluation of an expert judgment elicitation procedure for correlations. *Frontiers in Psychology*, 8, 90.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**Supporting Information S1.** Exploring priors.

**Supporting Information S2.** Case Study: Synthetic data.

**How to cite this article:** Spence, M. A., Martindale, J. A., & Thomson, M. J. (2023). EcoEnsemble: A general framework for combining ecosystem models in R. *Methods in Ecology and Evolution*, 14, 2011–2018. <https://doi.org/10.1111/2041-210X.14148>