StrathE2E2 Model parameter optimization, sensitivity analysis and Monte Carlo computation of credible intervals of outputs.

Michael R. Heath & Douglas C. Speirs

Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK.
E-mail: m.heath@strath.ac.uk

Date: April 2020

Computational scheme for fitting model parameters

Introduction

Many parameters in a model such as StrathE2E2 cannot be determined independently, since there are either no data available, or they refer to processes operating on coarse functional guilds where experimental data on individual species are unlikely to be directly applicable. Hence parameter selection has to involve some sort of optimization to fit the whole model to relevant observational data. This presents both a conceptual and an algorithmic challenge. The algorithmic challenge revolves around the number of parameters involved, the likely degree of interactions between parameters, and limitations on their values. Hand fitting is practically out of the question under these circumstances, so we need to seek computational solutions. Conceptual issues revolve around whether to fit a transient or stationary model to observed data.

Transient fitting attempts to optimize parameters so that model results match as closely as possible to the mean levels, time-trends and inter-annual variability of observed data on ecosystem components over as long a period as possible. In contrast, stationary fitting attempts to optimize parameters so that, with stationary annual cycles of driving data, then annual averages, integrated fluxes, or mean seasonal cycles of the stationary model match as closely as possible with long-term average observations on the state of the ecosystem.

Both the transient and stationary fitting approaches have benefits and problems. In the transient case, the results are extremely sensitive to the initial conditions at the start of the fitting time period and how close these are to being stationary with respect to the initial values of the driving data. Initial conditions therefore become parameters in their own right, and in most cases these cannot be well constrained by observations since the real-world includes variability which is highly unlikely to be fully explained by the model. It must also be assumed that all of the real-world drivers causing trends in the ecosystem are explicitly included in the model. Alternatively, models must include time-varying bias correction parameters which effectively absorb trends and variability in the observed data which are not capable of being explained by generic parameters of the model (e.g. recruitment variability of fish). However, the benefits of transient fitting are that, provided there is sufficient contrast in the observed data, the reactivity of the model can be accurately parameterised.

The main benefits of stationary state fitting are firstly that potentially a greater diversity of observational data may be available, and secondly that for a deterministic model the results are independent of initial conditions. However, the assumption that a long-term average of the real-world represents something close to a stationary state with respect to both model components and drivers is clearly a major issue. In addition, the stationary state fitting approach will not necessarily constrain the reactivity of the system to changes in mean levels of external drivers.

Parameter optimization for StrathE2E2

The parameter optimization scheme provided with StrathE2E2 is configured as a stationary state fitting method, though there is no intrinsic reason why is could not be adapted for time-series fitting. The methodology involves likelihood estimation and a simulated annealing scheme using the Metropolis-Hastings iterative algorithm (Bertsimas & Tsitsiklis, 1993; Cerny, 1985; Kavanagh *et al.*, 2004; Kirkpatrick *et al.*, 1983; Matear, 1995).

The Metropolis-Hastings algorithm implemented with StrathE2E2 automates the acceptance or rejection of proposed new randomly generated sets of model parameters (Figure 1). The objective is to maximize the likelihood of the set of observed (target) indices given a vector of model parameter values. Starting with an initial informed-guess parameter vector θ[0], the procedure iterates for k = 1, 2, 3..., and at each iteration the parameters are independently 'jiggled' by proposing θ* = θ[k-1] + δ[k], where δ[k] is a vector of random values from a gaussian distribution of mean 0 and standard deviation given by a fixed coefficient of variation applied to the current value of each parameter selected for fitting ($\theta_{current}$ = θ[k-1]). There are some special-case parameters in StrathE2E2 which are bounded by known limits, and the preference parameters which require to be renormalized to sum to 1.0 for each predator after the addition of δ[k].

After running the model to a stationary state with the θ* parameter vector, and extracting simulated versions of the observed ecosystem state indices (*i* = 1 to *I*) (where the observed indices each have a standard deviation σ_i), an overall error function $\chi^2_{\theta*}$ is calculated for the model as follows:

$$\chi^2_{\theta*} = \frac{\sum_{i=1}^{i=I}\left(\frac{\left(observed_i - simulated_{\theta*,i}\right)^2}{2\sigma_i^2}\right)}{I} \qquad \text{eqn 1}$$

The likelihood of the suite of observations given the parameter values is then estimated as:

$$P(observations|\theta^*) = \exp(-\chi^2_{\theta*}) \qquad \text{eqn 2}$$

The change in performance of the model due to the proposed jiggled parameter vector is measured by $\left(\chi^2_{\theta*} - \chi^2_{\theta_{current}}\right)$. A simple 'hill-climbing' scheme would accept the proposed vector θ* as a new version of $\theta_{current}$ for the next iteration only if the likelihood ratio (LR) $exp\left(\chi^2_{\theta*} - \chi^2_{\theta_{current}}\right) > 1$. Over many iterations $\theta_{current}$ should then migrate towards a maximum likelihood fit. However, such schemes are notorious for becoming trapped in local optima. Simulated annealing attempts to mitigate this risk by accepting a proportion of instances of θ* which produce worse results than $\theta_{current}$ (i.e. likelihood ratio < 1), thereby exploring a wider range of the parameter space. The process is an analogy of the way in which crystals form in a liquid metal as it cools. Taking Γ[k] to represent the 'temperature' of the system at iteration k, the probability of accepting θ* as the new $\theta_{current}$ is dependent on the modified likelihood ratio:

$$LR_m = \exp\left(\left(\chi^2_{\theta*} - \chi^2_{\theta_{current}}\right)/\Gamma[k]\right) \qquad \text{eqn 3}$$

The value of Γ[k] is allowed to 'cool' as the iterations progress (increasing k). As Γ[k] decreases towards 0 the system assumes a simple hill-climbing mode (*i.e.* only vectors θ* producing likelihood ratios > 1 are accepted). In the early stages, when Γ[k] >> 0, a proportion of θ* producing likelihood ratios < 1 are accepted. Practically, after each model run θ[k] = θ* only if $LR_m \geq \lambda$, where λ is drawn from a uniform random distribution between 0 and 1, otherwise θ[k] = θ[k-1] for the next iteration. At each iteration, both the currently

accepted and the proposed parameter vectors are saved, along with the resulting likelihood of the observations given the parameters.

The sequence of *P(observed|θ)[k])* should converge to a maximum value representing the best attainable fit between the model and the observations given the model structure and driving data. However, the rate at which convergence is achieved, and potentially the converged vector itself, are dependent on the cooling schedule and the magnitude of the jiggling coefficient of variation (cv) used to generate δ[k]. A geometrically decreasing temperature is applied (Γ[k] = r.Γ[k-1]), and after experimentation with the value of r and cv, the finally adopted values for fitting StrathE2E2 models to achieve consistency of convergence within a sensible number of iterations (<10000) are: r = 0.975, cv = 0.005. The best-fit value of θ[k] is regarded as being that remaining when no new parameter vectors are accepted within 200 iterations.
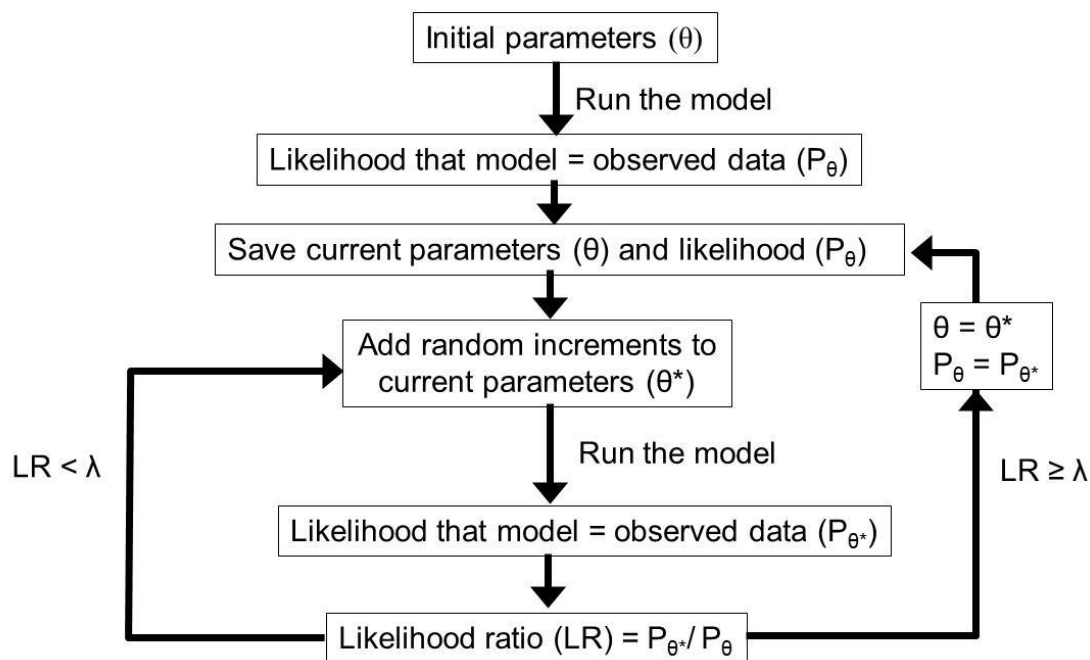
λ *is a random number (z ≤ λ ≤ 1); z increases from 0 towards 1 with iterations*

FIGURE 1 Flow diagram from the implementation of the simulated annealing parameter optimization scheme.

The target data to which the model is fitted are a set of indices of the real-world state of the ecosystem and their associated uncertainties (standard deviations). The indices include annual average biomasses or concentrations, annual integrated production rates of food web components, annual integrated dietary intakes, fishery landings and discards, denitrification rate, and other properties (Table 1). For each index, the code calculates an equivalent value for the final year of each model run as the basis for the likelihood estimation. Individual indices can be included or excluded from the likelihood calculation by means of a switch (value 1 or 0) in the input file containing the observational data.

TABLE 1 List of the observational indices of ecosystem state that are available for inclusion in the likelihood calculation as part of the model optimization process. The simulated annealing code derives equivalent values from the final year of model output for each iteration of the fitting process. Individual indices can be included or excluded from the process by parameter settings in the input file.

Description	Units
Whole domain annual total primary production	$mMN.m^{-2}.y^{-1}$
Whole domain annual new production from depth integrated nitrate drawdown	$mMN.m^{-2}.y^{-1}$
Annual within forest net production of macroalgae	$gC.m^{-2}.y^{-1}$
Whole domain annual omnivorous zooplankton gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual carnivorous zooplankton gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual planktivorous fish gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual demersal fish gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual filter/deposit feeding benthos gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual carnivorous/scavenge feeding benthos gross production	$mMN.m^{-2}.y^{-1}$
Whole domain annual net production of birds	$mMN.m^{-2}.y^{-1}$
Whole domain annual net production of pinnipeds	$mMN.m^{-2}.y^{-1}$
Whole domain annual net production of cetaceans	$mMN.m^{-2}.y^{-1}$
Whole domain annual maximum monthly averaged concentration of larvae of filter/deposit feeding benthos	$mMN.m^{-3}$
Whole domain annual maximum of monthly averaged concentration of larvae of carnivorous/scavenge feeding benthos	$mMN.m^{-3}$
Whole domain annual consumption of planktivorous fish by fish	$mMN.m^{-2}.y^{-1}$
Whole domain annual consumption of demersal fish by fish	$mMN.m^{-2}.y^{-1}$
Whole domain annual consumption of omnivorous zooplankton by fish and fish larvae	$mMN.m^{-2}.y^{-1}$
Whole domain annual consumption of omnivorous zooplankton by carnivorous zooplankton	$mMN.m^{-2}.y^{-1}$
Whole domain annual consumption of benthos by fish	$mMN.m^{-2}.y^{-1}$
Whole domain annual food consumption by birds	$mMN.m^{-2}.y^{-1}$
Whole domain proportion planktivorous fish in diet of birds	dimensionless
Whole domain proportion demersal fish in diet of birds	dimensionless
Whole domain proportion migratory fish in diet of birds	dimensionless
Whole domain proportion discards in diet of birds	dimensionless
Whole domain annual food consumption by pinnipeds	$mMN.m^{-2}.y^{-1}$
Whole domain proportion planktivorous fish in diet of pinnipeds	dimensionless
Whole domain proportion demersal fish in diet of pinnipeds	dimensionless
Whole domain proportion migratory fish in diet of pinnipeds	dimensionless
Whole domain annual food consumption by cetaceans	$mMN.m^{-2}.y^{-1}$
Whole domain proportion planktivorous fish in diet of cetaceans	dimensionless
Whole domain proportion demersal fish in diet of cetaceans	dimensionless
Whole domain proportion migratory fish in diet of cetaceans	dimensionless
Whole domain annual planktivorous fish landings (live weight)	$mMN.m^{-2}.y^{-1}$
Whole domain annual demersal fish landings (live weight)	$mMN.m^{-2}.y^{-1}$

Whole domain annual migratory fish landings (live weight)	mMN.m^{-2}.y^{-1}
Whole domain annual filter/deposit feeding benthos landings (live weight)	mMN.m^{-2}.y^{-1}
Whole domain annual carnivorous/scavenge feeding benthos landings (live weight)	mMN.m^{-2}.y^{-1}
Whole domain annual carnivorous zooplankton landings (live weight)	mMN.m^{-2}.y^{-1}
Whole domain annual macrophyte landings (live weight)	mMN.m^{-2}.y^{-1}
Whole domain annual carbon gross production:biomass ratio of macrophytes	y^{-1}
Whole domain annual gross production:biomass ratio of larvae of filter/deposit feeding benthos	y^{-1}
Whole domain annual gross production:biomass ratio of larvae of carnivorous/scavenge feeding benthos	y^{-1}
Whole domain annual gross production:biomass ratio of filter/deposit feeding benthos	y^{-1}
Whole domain annual gross production:biomass ratio of carnivorous/scavenge feeding benthos	y^{-1}
Whole domain annual gross production:biomass ratio omnivorous zooplankton	y^{-1}
Whole domain annual gross production:biomass ratio of carnivorous zooplankton	y^{-1}
Whole domain annual gross production:biomass ratio of larvae of planktivorous fish	y^{-1}
Whole domain annual gross production:biomass ratio of larvae of demersal fish	y^{-1}
Whole domain annual gross production:biomass ratio planktivorous fish	y^{-1}
Whole domain annual gross production:biomass ratio of demersal fish	y^{-1}
Whole domain annual gross production:biomass ratio of migratory fish	y^{-1}
Whole domain annual net production:biomass ratio of birds	y^{-1}
Whole domain annual net production:biomass ratio of pinnipeds	y^{-1}
Whole domain annual net production:biomass ratio of cetaceans	y^{-1}
Annual average proportion of kelp carbon uptake which is exuded	dimensionless
Whole domain annual average molar nitrogen:carbon ratio of kelp	dimensionless
Whole domain annual denitrification	mMN.m^{-2}.y^{-1}
Whole domain proportion of demersal fish catch discarded	dimensionless
Annual average ammonia concentration in inshore and offshore porewater of sand (grain size >0.15mm)	mMN.m3
Annual average ammonia concentration in inshore and offshore porewater of mud (grain size < 0.15mm)	mMN.m3
Annual average nitrate concentration in inshore and offshore porewater of sand (grain size >0.15mm)	mMN.m3
Annual average nitrate concentration in inshore and offshore porewater of mud (grain size <0.15mm)	mMN.m^{-3}
Annual average total (labile + refractory) organic N content in inshore and offshore of sand (grain size > 0.15mm)	%N (gN.(g dry sed)$^{-1}$)
Annual average total (labile + refractory) organic N content in inshore and offshore of mud (grain size < 0.15mm)	%N (gN.(g dry sed)$^{-1}$)
Whole domain average winter (Nov-Feb) nitrate concentration in	mMN.m^{-3}

the shallow layer	
Whole domain average summer (May-Aug) nitrate concentration in the shallow layer	$mMN.m^{-3}$
Whole domain average winter (Nov-Feb) nitrate concentration in the deep layer	$mMN.m^{-3}$
Whole domain average summer (May-Aug) nitrate concentration in the deep layer	$mMN.m^{-3}$
Whole domain average winter (Nov-Feb) ammonia concentration in the shallow layer	$mMN.m^{-3}$
Whole domain average summer (May-Aug) ammonia concentration in the shallow layer	$mMN.m^{-3}$
Whole domain average winter (Nov-Feb) ammonia concentration in the deep layer	$mMN.m^{-3}$
Whole domain average summer (May-Aug) ammonia concentration in the deep layer	$mMN.m^{-3}$
Inshore:offshore ratio of annual mean carnivorous zooplankton depth averaged concentration	dimensionless
Inshore:offshore ratio of annual mean omnivorous zooplankton depth averaged concentration	dimensionless
Inshore:offshore ratio of annual mean phytoplankton surface layer concentration	dimensionless
Inshore:offshore ratio of annual mean nitrate surface layer concentration	dimensionless
Inshore:offshore ratio of annual mean ammonia surface layer concentration	dimensionless
Inshore:offshore ratio of annual mean planktivorous fish density per m^{-2}	dimensionless
Inshore:offshore ratio of annual mean demersal fish density per m^{-2}	dimensionless
Whole domain annual by-catch of birds	$mMN.m^{-2}.y^{-1}$
Whole domain annual by-catch of pinnipeds	$mMN.m^{-2}.y^{-1}$
Whole domain annual by-catch of cetaceans	$mMN.m^{-2}.y^{-1}$
Proportion of kelp annual nitrogen uptake exported as beach-cast	dimensionless
Whole domain cetacean landings by whaling vessels	$mMN.m^{-2}.y^{-1}$

Versions of the simulated annealing scheme

Three groups of parameters in the combined ecology and fishing fleet model system might be considered as candidates for optimization, depending on the availability of external data:
- Ecology model parameters
- Scaling parameters linking effort to harvest ratio
- Fishing gear activity rates

It is clearly unrealistic to expect to be able to optimize all of these sets of parameters concurrently since there is a risk that some may be confounded. So, the R package provides three versions of the simulated annealing system, each of which optimizes one of these groups of parameters, given knowledge on the other two.

Ecology model parameter optimization

Optimizing the parameters of the ecology model requires that fishing gear activity rates and the parameters linking effort to harvest ratios are know.

Certain parameters of the ecology model are excluded from the optimization process because they are either confounded with other parameters in some way or there is a clear route to establishing values from external evidence (Table 2). All other parameters, defining the prey preferences of all the living components of the model, the uptake and mortality rate coefficients, and the rate coefficients for the biogeochemical processes, are eligible for optimization. A subset of the feeding rate parameters for the top-predators can be optionally excluded from the optimization process if reasonable values can be independently estimated. Separate coefficients of variation can be set for different classes of parameters (preference parameters, maximum uptake rates, half saturation coefficients, microbial rates, density dependent mortality rates) via a setup text file. These can be changed during a run, to enable exploration of the effects of parameter search patterns.

The running code is initialised with a proposed parameter set formed of three asci csv files (preference parameters, uptake and mortality parameters, microbiology and other parameters). At the end of a run the code exports a new set of these three files containing the accepted values for the maximum likelihood model.

TABLE 2 Background to the parameters of the ecological model which require to be fixed from external evidence and are excluded from the simulated annealing process.

Parameter type	Description
Assimilation efficiencies for each living guild in the model (Heath 2012)	Proportion of the ingested mass of food that contributes to new body tissue, after subtracting defecation and the metabolic costs of digestion and synthesis.
Biomass loss rates due to temperature-dependent metabolism for each living resource guild	Proportion of biomass lost to ammonia per day due to non-feeding related metabolism at a given reference temperature. Temperature dependency following a Q_{10} function.
Q_{10} values for temperature dependent processes, and the Q_{10} reference temperature	Separate Q_{10} values for autotrophic uptake of nutrient, heterotrophic feeding, and heterotrophic metabolism based on literature data.
Light intensity required to saturate autotrophic nutrient uptake	Light saturation intensity for nutrient uptake cannot be treated as a fitted value since it is confounded with other uptake parameters. Value estimated from survey of field and laboratory experiments.
Annual weight specific fecundities of planktivorous and demersal fish guilds and the two benthos guilds in the model (filter/deposit feeders and carnivore/scavenge feeders)	Guild-level values of the annual proportion of biomass which is shed as eggs
Spawning start and end dates for fish and benthos	Day-number (assuming a 360-day year) of the onset and end of spawning to be based on such empirical data as are available.
Recruitment start and	Day-number (assuming a 360-day year) of the onset and end of

end dates for fish and benthos	recruitment from the larval stage to the main guild biomass. Recruitment is assumed to occur at a uniform daily rate between the start and end dates.
Extra-domain stock biomass of migratory fish, and the proportion invading the domain each year. Start and end dates for the annual invasion, and start and end dates for the emigration. (see description in this document).	Biomass of the wider stock contributing to the proportion forming the transient population in the model domain extracted from fisheries stock assessments. Day-number (assuming a 360-day year) of the onset and end of immigration and emigration phases. Immigration parameterised as fixed mass flux per day between the onset and end dates. Emigration parameterised as a proportion of offshore biomass leaving per day between the onset and end dates. Values based on such empirical data from fisheries and stock assessments as are available.
Harvestable biomass density threshold for each resource guild.	The living resource guilds in the model represent a mixture of harvestable and non-harvestable species, especially the invertebrate guilds. The density threshold parameter sets a limit for the guild biomass below which the harvestable species are assumed to be exhausted. Values to be based on such empirical data from surveys, fisheries and stock assessments as are available.
Minimum inedible biomass of carnivorous zooplankton	The carnivorous zooplankton guild is a key component of the food web, predated on by all the fish and top-predators. However it represents an extremely diverse range of fauna many of which are not edible in significant quantities by the guild predators, e.g. scyphomedusae. A minimum edible threshold is set to ensure that the guild as a whole cannot be extirpated by predation.

Optimization of scaling parameters linking fishing effort to harvest ratio

Optimizing the scaling parameters linking fishing effort to harvest ratios requires that all of the ecology model parameters are known, and that the activity rates and distribution of each fishing gear is known. This is a situation which might apply when the coupled fishing and ecology model system has been comprehensively fitted for a data-rich period of time, and is then being applied to a different data-poor period where at least the gear activity rates are known.

Proposed values of the scaling parameters are contained in the configuration file for the fishing fleet model. In addition, the model setup includes a separate asci csv file (harvest_ratio_multiplier.csv) containing multipliers to be applied to each of these scaling parameters. These are useful when conducting scenario experiments, but would ordinarily be set to 1.0. The simulated annealing code for optimizing the scaling values iterates the multiplier values to seek the combination producing the best fit of the overall model to the set of target data. At the end of the process, the code saves the history of proposed and accepted multiplier values. If deemed reasonable, the final accepted version of the multipliers can then be manually applied to the corresponding rates in the fishing fleet configuration to produce an updated version. The effort-harvest ratio multipliers associated with birds, pinnipeds and cetaceans are likely to be poorly constrained by observational data and may need to be set independently. Nevertheless they are included in the process here.

Optimization of fishing activity rates

Usually, the fishing gear activity rates are a known input to the model. Sometimes, however, it may be necessary to use the model to hindcast the likely rates of activity given established values for the ecology model parameters and the scaling parameters linking effort to harvest ratios. This is a situation that might apply when the coupled fishing and ecology model system has been comprehensively fitted for a data-rich period of time, and is then being applied to a different data-poor period where the fishing gear activity rates are unknown or only approximately known.

Proposed values of the gear activity rates are contained in an input file for the fishing fleet model. In addition, the model setup includes a separate asci csv file (fishing_activity_multiplier.csv) containing multipliers to be applied to each of these rates. These are useful when conducting scenario experiments, but would ordinarily be set to 1.0. In the same way as for the effort-harvest ratio scaling value fitting, the simulated annealing code for optimizing the gear activity rates iterates the multiplier values to seek the combination producing the best fit of the overall model to the set of target data. At the end of the process, the code saves the history of proposed and accepted multiplier values. If deemed reasonable, the final accepted version of the multipliers can then be manually applied to the corresponding rates in the initial input file to produce an updated version.

There are two alternative functions for optimizing fishing activity rates. The first is conceptually equivalent to the simulated annealing functions for optimizing the ecology model parameters and the harvest ratio scaling parameters, in that the target for fitting is the same database of observational indices of ecosystem status. This uses the fully coupled fishing fleet and ecology model system and hence is costly to run. The second version is much faster and relies only on the fishing fleet model, but here the target for optimization is a set of known harvest ratios in each spatial zone. The circumstances under which the past harvest ratios for each guild are sufficiently well known to enable this fitting approach are likely to be rather restricted, but this method is probably the more robust. In both approaches the chief problem is that the selectivity patterns of the fishing gears are likely to be highly overlapping so that it will be difficult to identify a realistic maximum likelihood set of activity rates. For this reason the functions make use of a table of 'gear linkage' parameters which constrain the extent to which groups of gears can vary independently during the parameter searching process. This makes since management measures which may have restricted the activities of e.g. demersal trawlers, will also have been applicable to e.g. beam trawlers. However, the linkage table is populated partly by judgement rather than by hard data.

Sensitivity analysis

Sensitivity analysis helps to highlight the parameters which have the most influence in the simulated annealing optimization process. In addition the analysis can help to identify parameters for which constraint by external information would be most beneficial.

Simple one-at-a-time sensitivity analysis of parameters in a model is a straightforward process when there are few parameters and relatively little interaction between parameters. However, StrathE2E2 include many parameters and there is a strong possibility of interactions, so a more sophisticated scheme is required.

The R package includes a function to perform global analysis of parameter sensitivity using the Morris method (Morris, 1991). This involves a factorial sampling scheme to generate parameter values for replicate model runs and generate values for the 'elementary effect' (EE) of each parameter. The mean and standard deviation of EE values over many runs

allows ranking of the parameters in terms of the sensitivity of their effects on the model, and the strength of interactions with other parameters.

The parent parameter set for the sensitivity analysis (θ) incorporates all of the parameters and inputs to the model, i.e. the maximum-likelihood set of fitted ecology model parameters produced by the simulated annealing scheme, plus the fixed ecological parameters, the physical configuration parameters of the model (layer thicknesses, inshore/offshore areas), the parameters passed to the ecology model from the fishing fleet model, and the fishing activity rates and environmental driving data . The sensitivity to the environmental driving data is investigated by applying a scaling factor uniformly to the annual cycle of environmental data values.

From this parent set, a series of child-sets (θ^*_k where $1 \le k \le r$) is generated by applying a separate random increment to each parameter; $\theta^*_k = \theta + \delta[k]$ where $\delta[k]$ is a vector of random values from a gaussian distribution of mean 0 and standard deviation given by a fixed coefficient of variation applied to the parent-set value of each parameter. This process is equivalent to that for generating parameter proposals in the simulated annealing scheme, except that all of the parameters and model inputs are varied in the sensitivity analysis, rather than just the subset of parameters to be fitted.

For each of the child-sets of parameters, the likelihood of the suite of observed data indices $P(observations|\theta^*_{k,0})$ is calculated following runs of the model to stationary state, exactly as in simulated annealing scheme. We refer to these as trajectories.

Then, for each trajectory, the parameters ($1 \le i \le n$) are varied in turn, one at a time, by adding a fixed proportionality increment, the model re-run, and the likelihood computed ($P(observations|\theta^*_{k,i})$). The proportionality increment for a given trajectory is drawn at random from a set of four fixed levels in the range $\pm 10\%$ of the child parameter set (-10%, -5%, +5%, +10%, i.e. $\Delta = 0.9, 0.95, 1.05, 1.10$), so that:

$$\theta[i]^*_{k,i} = \theta[i]^*_{k,0} \cdot \Delta[z] \qquad\qquad \text{eqn 4}$$

where $\theta[i]^*$ is the ith element of θ^*, $\Delta[z]$ is the zth element of Δ, and z is drawn at random from the series (1, 2, 3, 4) for each value of k. Hence, for each trajectory the model runs are repeated n+1 times, where i=0 corresponds to the baseline run for each trajectory. The total number of nested runs to support the sensitivity analysis is thus r.(n+1).

From each level-run of the model (in which a single parameter is perturbed), and its corresponding likelihood, the elementary effect of the parameter is calculated as

$$EE_{k,i} = \frac{P(observations|\theta^*_{k,i}) - P(observations|\theta^*_{k,0})}{\sqrt{(\Delta[z]-1)^2}} \qquad\qquad \text{eqn 5}$$

On completion of the runs for all trajectories, the mean (μ_i) and standard deviation (S_i) of the r elementary effects for each parameter (*i*) are calculated. For the mean,

$$\mu_i = \frac{1}{r}\sum_{k=1}^{r} EE_{k,i} \qquad\qquad \text{eqn 6}$$

The magnitude of the mean represents the sensitivity of each parameter, and the corresponding standard deviation indicates the degree of non-linearity in the response or interaction with other parameters. The standard error of the mean for each parameter (SEM_i) is given by:

$$SEM_i = \frac{S_i}{\sqrt{r}} \qquad\qquad \text{eqn 7}$$

If $S_i > |\mu_i| \cdot \frac{\sqrt{r}}{2}$ then we can approximately conclude that μ_i is significantly greater than zero.

Monte Carlo computation of credible intervals for the model outputs

The Monte Carlo scheme provided with the StrathE2E2 package has two modes of operation. The first is referred to as 'baseline mode'. This involves generating a list of ecology model parameter sets and, for each set, determining the likelihood of a suite of observational data consistent with the environmental and fishery drivers. The model outputs generated with each set are then weighted by the likelihood before deriving quantiles of their distribution. The quantile ranges represent credible intervals of the model outputs.

The second mode of operation is referred to as the 'scenario mode'. In this case, the parameter sets and their associated likelihoods from a baseline mode simulation, are used to generate model outputs for scenarios of environmental or fishing drivers – e.g. increased activity by selected gears, or warmer sea temperatures. These scenario outputs are then weighted by the baseline mode likelihoods before derivation of quantiles and credible intervals.

Baseline mode operation.

The initial (parent) ecology model parameter set (θ) for a baseline mode Monte Carlo simulation should be the maximum-likelihood set produced by the simulated annealing scheme. From this parent set, a series of child-sets (θ^*_k, where $1 \leq k \leq r$ and r = e.g. 1000), is generated by applying a separate random increment to each of the parameters; $\theta^*_k = \theta + \delta[k]$ where δ[k] is a vector of random values from a uniform (rather than a gaussian) distribution of mean 0 and given range.

For each of the parameter sets, the likelihood of the suite of observed data indices P(observations|θ^*_k) is calculated following a run of the model to stationary state, and all the output from the final, stationary year of each run is saved (k = 0 to r, where k = 0 corresponds to the maximum likelihood (parent) parameter set).

To calculate the credible interval for any direct or derived model output variable (e.g. annual average mass density of a state variable, or the mass density on a given day in the final year), the values from the individual model runs (V_k) and the associated likelihoods (P_k) are assembled as a list of (r + 1; k = 0 to r) pairs ($a_k = (V_k, P_k)$). The list is then sorted by ascending values of V (a_j ; j = 1 to (r + 1) such that $V_j \geq V_{(j-1)}$). The vector of cumulative likelihoods is then calculated as follows:

$$C_j = (\ P_{j=1},\ \sum_{j=1}^{j=2} P_j\ ,\ \sum_{j=1}^{j=3} P_j\ ,\ \sum_{j=1}^{j=4} P_j\ \dots\ \sum_{j=1}^{j=(r+1)} P_j\ ) \qquad\qquad \text{eqn 8}$$

and the proportions of maximum cumulative likelihood as $Q_j = C_j/C_{(r+1)}$.

Finally, values of V corresponding to discrete values of Q = {0.005, 0.25, 0.5, 0.75, 0.995} are then extracted by interpolation. These values span the 0.5% and 99.5% credible intervals of the model output given the observed target data and uncertainty in the fitted ecology parameters. Note that uncertainties in the fixed ecology parameters, fishing fleet parameters, and environmental driving data are not reflected in these credible intervals.

Scenario mode operation

The aim of the scenario mode is to estimate credible intervals of model outputs for an experimental scenario state in which environmental and/or fishery driving data have been altered from the baseline to reflect, e.g. warming conditions or alternative fishery configurations. In such cases there cannot be any target data against which to measure the performance of the model. Instead, we use the parameter sets and associated likelihoods generated in a baseline mode simulation to generate outputs and credible intervals with the scenario drivers.

To calculate the credible interval for output variables in scenario mode, the values from the individual scenario model runs (VS_k) are paired up with the <u>baseline</u> mode likelihoods (P_k) associated with each parameter set. Then, as in the baseline mode, the values are assembled as a list of ($r + 1$; $k = 0$ to r) pairs ($a_k = (VS_k, P_k)$). The list is then sorted by ascending values of VS (a_j ; $j = 1$ to ($r + 1$) such that $VS_j \geq VS_{(j-1)}$). The vector of cumulative likelihoods is then calculated as in equation 8, leading to the scenario mode proportions of maximum cumulative likelihood $QS_j = C_j/C_{(r+1)}$.

References

Bertsimas, D. & Tsitsiklis, J. (1993). Simulated annealing. *Statistical Science*, **8**, 10-15.

Cerny, V. (1985). A thermodynamic approach to the travelling salesman problem: An efficient simulation. *Journal of Optimization Theory and Applications*, **54**, 41-51.

Heath, M.R. (2012). Ecosystem limits to food web fluxes and fishery yields in the North Sea simulated with an end-to-end food web model. *Progress in Oceanography*, **102**, 42-66.

Kavanagh, P., Newlands, N., Christensen, V. & Pauly, D. (2004). Automated parameter optimization for Ecopath ecosystem models. *Ecological Modelling*, **172**, 141-149.

Kirkpatrick, S., Gelett, C.D. & Vecchi, M.P. (1983). Optimisation by simulated annealing. *Science*, **220**, 621-630.

Matear, R.J. (1995). Parameter optimization and analysis of ecosystem models using simulated annealing: A case study at Station P. *Journal of Marine Research*, **53**, 571-607.

Morris, M.D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, **33**, 161-174.