SEQUENTIAL IMPORTANCE RESAMPLING FILTERING IN ECOSYSTEM MODELLING

S. Losa, G. Kivman, J. Schröter and M. Wenzel

sloza@awi-bremerhaven.de

Alfred Wegener Institute for Polar and Marine Research

Abstract

Sequential Important Resampling filter (SIRF) is applied for assimilating time-series data into an ecosystem model. Advantage of this Monte-Carlo based data assimilation approach for combined state and parameter estimation in ecosystem modelling has been already demonstrated in previous studies (Losa et al., 2003). Some aspects of the SIRF implementation for the highly non-linear system, however, still remain to be worked out. The filter is known to suffer from degeneration of the ensemble if either the system noise does not provide sufficient spreading of states which are resampled several times or the ensemble badly approximates the true prior distribution (the distance between the best member and the true state is too big). This problem is even more pronounced in the case of simultaneous state-parameter estimation where regenerating the number of samples in the parameter space is needed. In this study, we are focusing on the model noise optimization. Investigating the system noise would, probably, allow us to explain the notable seasonality obtained for some of the optimized parameters in our previous study (Losa et al., 2003).

1 Method

Initialization. An ensemble of initial conditions and model parameters $\psi_j^k, k = 1, ...N$ is drawn from a prior distribution.

Prediction. Each ensemble member evolves in accordance to model equations which are perturbed by a model noise till the next analysis step $j+1$ when data d_{j+1} become available.

Filtering. Assign to each $\psi_j^k, k = 1, ...N$ a weight ω_{j+1}^k according to the distance between the state described by ψ_0^k and the data d_{j+1}.

Resampling. Assign these weights ω_{j+1}^k as probabilities for each ensemble member to be resampled independently N times with replacement.

Prediction. The resampled ensemble $\psi_{j+1}^k, k = 1, ...N$ with some noise added to the parameters evolves till the next analysis step.

time step=j

time step=j+1
(data are available)

Time

Previously we estimated a proper level of the model noise simply by "trial- and errors".
Now we suggest considering levels (E) of the system noise ϵ as additional parameters to be optimized at every analysis step.

Model noise generating (and estimating) and jittering model parameters are two the very crucial factors affecting the filter performance and preventing the "ensemble collapse"

The parameters can be jittered by simply generating a new parameter set.
For example, if the parameter values are resampled many times, at an analysis step, a new parameter ensemble can be generated by redrawing from the uniform distribution whithin the interval [**p** - nearest smaller value, **p** + nearest higher value].

2 Experiment

We have implemented the SIRF for a highly non-linear nine-compartment ecosystem model, which possesses 15 poorly-known model parameters.

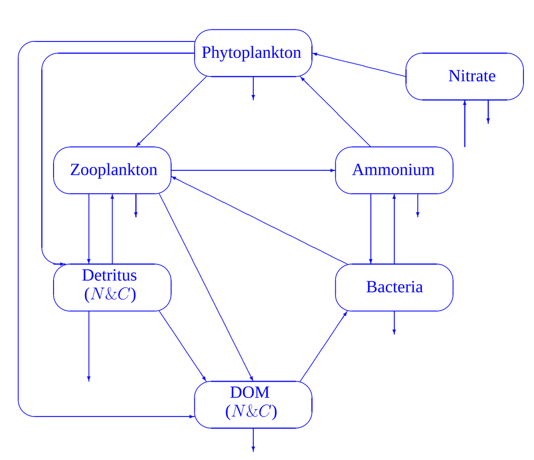
2.1 Ensemble Initialization

An ensemble ψ of $N = 1000$ members is generated for state variables X and model parameters P, including the levels of the system noise $E \in P$, from an exponential distribution

$$\rho(\psi) = \overline{\psi} \exp\left(-\frac{\psi}{\overline{\psi}}\right),$$

mean of the distribution $\overline{\psi}$ here assumed to equal to a first guess.

2.2 Ecosystem model



The ensamble particles evolves in accordance to the ecosystem model (see a schematic diagram, to the left) equations. The ecosystem model used is built on the seven-compartment nitrogen-based model of Fasham *et al* (1990) (FDM-model).

The schematic diagram shows the compartments and inter-compartmental flows of the upper mixed layer ecosystem.

Acknowledgments
We thank Dr. H. Drange for providing us with the model code.

2.5 Data and weighting

The model was constrained by data of the Bermuda Atlantic Time-series Study, particularly, by measurements (X_{obs}^k) of nitrate, chlorophyll, dissolved organic nitrogen and carbon concentrations for the period December 1988 to January 1994. All the data were averaged over the upper mixed layer (UML). The UML thickness were estimated by means of an analysis of BATS temperature profiles for the same period. The UML depth is determined as the depth at which the temperature is 0.5^0C less than that at the surface.
The relative weights w_j^k are calculated under the assumption of Gaussian data errors .

$$w^k = C \exp\left(-0.5\, \sigma_{X_{obs}}^{-2} (X^k - X_{obs}^k)\right),$$

$\sigma_{X_{obs}}$ is the error levels of the observations.

2.4 Optimized model parameters

- μ_1 – phytoplankton maximum specific mortality rate;
- k_1 – half-saturation constant for nitrate uptake;
- k_2 – half-saturation constant for ammonium uptake;
- k_5 – phytoplankton mortality half-saturation const.;
- ψ – nitrate uptake ammonium inhibition parameter;
- α – initial slope of the $P - I$ curve;
- μ_2 – zooplankton maximum loss rate;
- g – zooplankton maximum ingestion;
- k_3 – zooplankton ingestion half-saturation constant;
- k_6 – zooplankton loss rate half-saturation constant;
- μ_3 – bacterial excretion rate;
- V_b – bacterial maximum uptake rate;
- k_4 – bacterial half-saturation const. for uptake;
- μ_4 – detrital breakdown rate;
- w – detrital sinking rate;

3 Results

Model noise

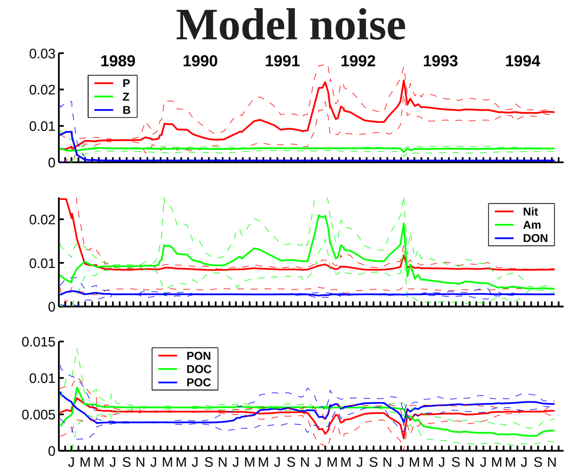


Figure 1: The evolution of the model noise variance for the period 1989 - 1994.

State variables

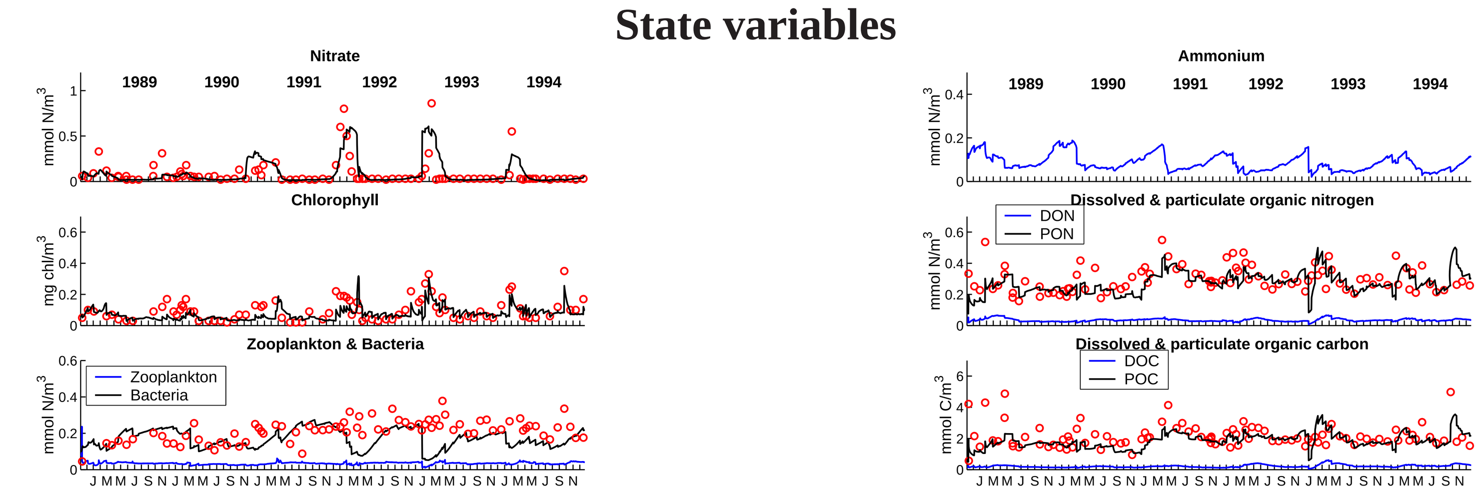


Figure 2: The time evolution of the ecosystem components at the Bermuda station for the period 1989-1994. The solid curve is a result of the sequential weak constraint parameter estimation. Red circles are data for nitrate, chlorophyll, particular organic nitrogen and carbon. Bacteria data are not assimilated.

Previous parameter estimates

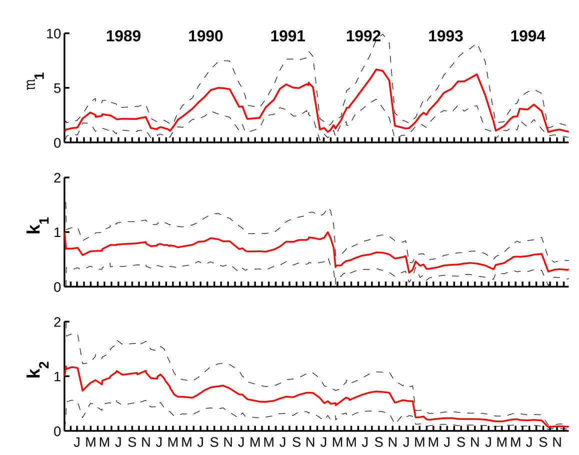


Figure 3: The evolution of the estimates for the model parameters obtained with a constant level of the model noise (Losa et al., 2003). Values are normalized with respect to model parameter initial values.

Model parameters

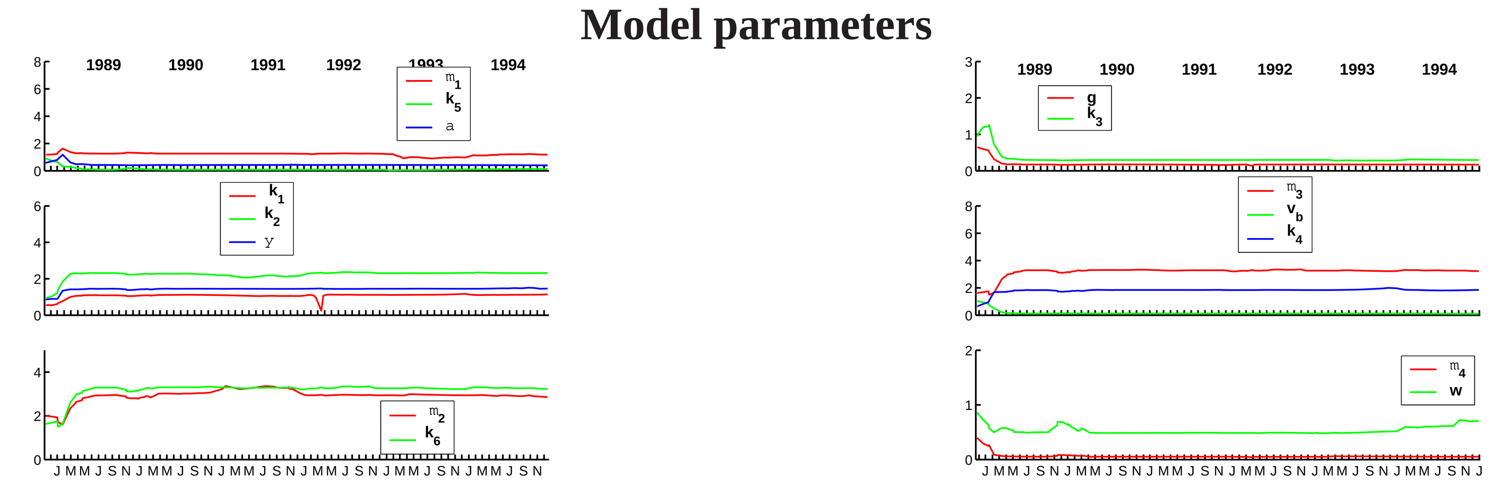


Figure 4: The evolution of the parameters tuned to BATS data for the period 1989 - 1994. Values of the model parameters are normalized with respect to their initial values.

4 Conclusions

The obtained results, to some extent, prove the idea that the notable seasonality obtained for some of the biological parameters can be linked to errors existing in forcing data and/or to uncertainties in the used biological parameterizations.

References

[1] Fasham M.J.R., Ducklow H.W., McKelvie S.M., 1990. A nitrogen-based model of plankton dynamics in the oceanic mixed layer. *J.Mar.Res.*, 48(3), pp. 591-639.

[2] Losa S.N., Kivman, G.A., Schröter, J. and Wenzel, M., 2003. Sequential weak constraint parameter estimation in an ecosystem model. Jornal of Marine Systems, 42, 31–49.

[3] Rubin D.B., 1988. Using the SIR algorithm to simulate posterior distribution, in Bayesian Statistics 3 (Eds. J.M. Bernardo, M.H. Degroof, D.V. Lindleyand, A.F.M. Smith). *Oxford Univ. Press.*,395-402.