Validating and calibrating MERSEA ecosystem models

S. Losa, J. Schröter, M. Wenzel

sloza@awi-bremerhaven.de

Alfred Wegener Institute for Polar and Marine Research

Abstract

A new 1-dimensional ecosystem model, developed within the MERSEA project is calibrated with and validated against time series data collected at 6 different locations of the World Ocean. The new model describes the dynamics of phytoplankton N, zooplankton Z, nutrients N and detritus D under different environmental conditions. The quality of the model performance strongly depends on the model biological parameter specification. Here we apply the Sequential Importance Resampling filter (Rubin, 1988) for the parameter optimization problem. Several nonparametrical statistics criteria are presented and used for estimating "goodness" of the model to data fit.

1 Model

The version of a nitrogen based four-compartment (NPZD) model has been developed by *I. Kriest* and *A. Oschlies* (WP 7.2.1). In the new model, phytolpankton is implicitly presented by a spectrum of different sizes. Thus, some of the described biogeochemical process are size-dependent. (We will refer to the model as SD NPZD)

2 Data

The model is constrained by monthly mean data of

the Bermuda Atlantic Time-series Study (BATS 32^0N, 65^0W), Ocean Weather stationPAPA (50^0N, 145^0W), the North Atlantic Bloom Experimen (NABE, 47^0N, 20^0W), the Arabian Sea C station (AS-C, 10^0N, 65^0E), Equatorial Pacific Ocean (EqPac, 0^0S, 140^0W), the Ross Sea (63.2^0S, 170^0W)

particularly, by measurements of dissolved inorganic nitrogen and chlorophyll concentrations.

3 Parameter optimization method

(WP 7.3.1.2, in colaboration with P.J. van Leeuwen, Universiteit Utrecht)

In Sequential Importance Resampling (see Figure 1.) the probability density function (pdf) of the ecosystem model is represented by an ensemble ψ of K members, generated for model state variables X and parameters P from a prior distribution Each ensemble member evolves in accordance to the model equations subject to model noise.

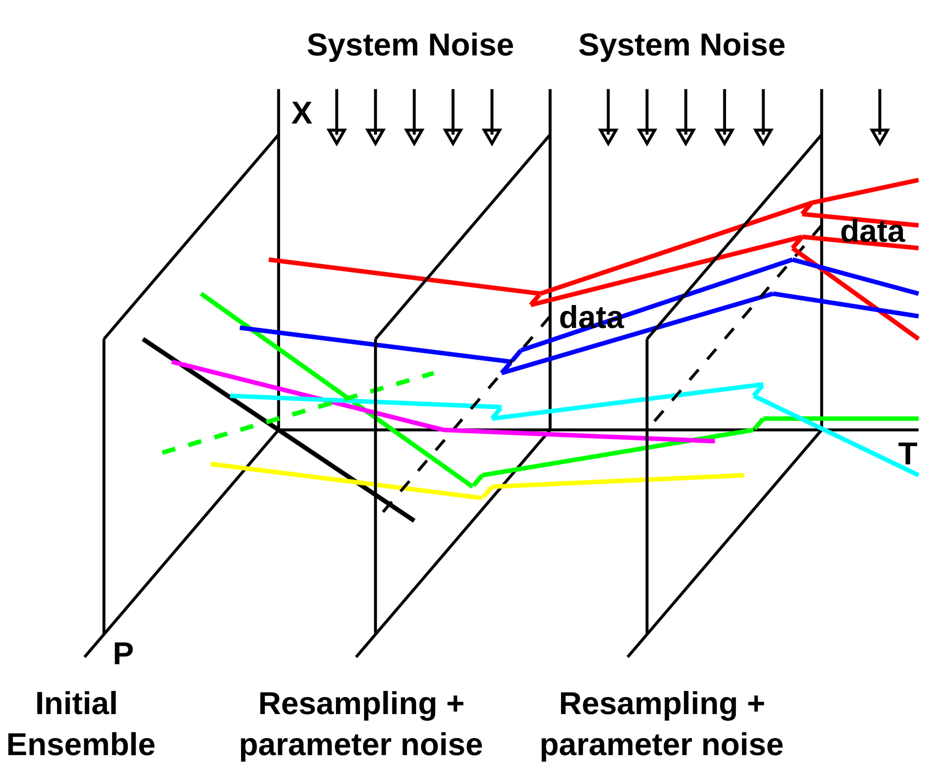

Figure 1: SIRF scheme

As soon as observations are present each ensemble member is weighted by the 'distance' to these observations. The weight w^k for ensemble member ψ_k is calculated from :

$$w^k = p(d|\psi_k) / \sum p(d|\psi_k) \qquad (1)$$

which follows directly from Bayes theorem (Kivman, 2003, Van Leeuwen, 2003). w^k is considered as probability in the k member to be resampled. The relative weights (or conditional probability) might be calculated under the assumption of Gaussian

$$p(d|\psi_k) = C \exp(-0.5\,\sigma_{X_{obs}}^{-2}(X^k - X^k_{obs})^2),$$

or Lorenz data errors

$$p(d|\psi_k) = \frac{1}{1 + (X^k - X^k_{obs})^2 \sigma_{X_{obs}}^{-2}} \qquad (2)$$

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

4 Model validation

Nonparametrical (distribution free) rank statistics (based on "rank order") is used as criteria of "goodness" of model-to-data fit:

r_{sp} – Spearman Rank Correlation Coefficient, a measure of the strength of the associations between model components and data

r_{sp} = 1-6 $\sum \frac{d^2}{N(N^2-1)}$
d are differences in statistical ranks of respective variables

W – Wilcoxon test shows whether the model solution and data are of the same distributions (have same median value)

W = (T - $0.25N(N$-1) ± 0.5)$\sqrt{\frac{24}{N(N-1)(2N+1)}}$
T is a sum of negative (or positive) ranks

MW – Mann-Whitney U criterion tests whether all modes of data distribution and distribution of model results are similar and sampled equally well

$U = N^2 + \frac{N(N+1)}{2}$ - R_1,
MW = (U - $0.5N^2$) $\sqrt{\frac{12}{N^2(2N+1)}}$
R1 is the sum of model (or data) ranks
N is the number of data

Table 1. Agreement between model (simple and SD versions) and observed chlorophyll "a"

Stat. criteria	PAPA		NABE		BATS		AS-C		EqPac.		Ross Sea	
	simple	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*
r_{sp}	0.72	0.41	0.88	0.38	0.05	0.45	-0.19	0.74	0.79	0.78	-0.07	-0.07
MW	-3.98	-0.68	-6.65	-5.91	-18.58	-1.05	-1.95	-1.29	-0.16	-2.16	-0.84	-0.84
W	2.92	0.10	4.77	4.75	11.97	3.17	1.77	2.4	1.53	4.14	0.91	1.19
notes					*improved*		*improved*					

Table 2. Agreement between model (simple and SD versions) and observed DIN concentrations

Stat. criteria	PAPA		NABE		BATS		AS-C		EqPac.		Ross Sea	
	simple	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*	*simple*	*SD*
r_{sp}	0.52	0.67	0.51	0.027	0.59	0.77	-0.60	0.095	0.56	0.57	-0.90	-0.90
MW	-9.15	-4.47	-7.30	-7.30	-9.85	-10.7	-4.55	-2.95	-2.41	-1.40	-3.50	-3.57
W	7.77	5.25	5.22	5.22	10.30	14.85	3.28	2.04	6.27	2.45	2.60	2.60
notes	*improved*						*improved*		*improved*			

5 Results *of*

simultaneous tuning the SD NPZD model for all the noted locations

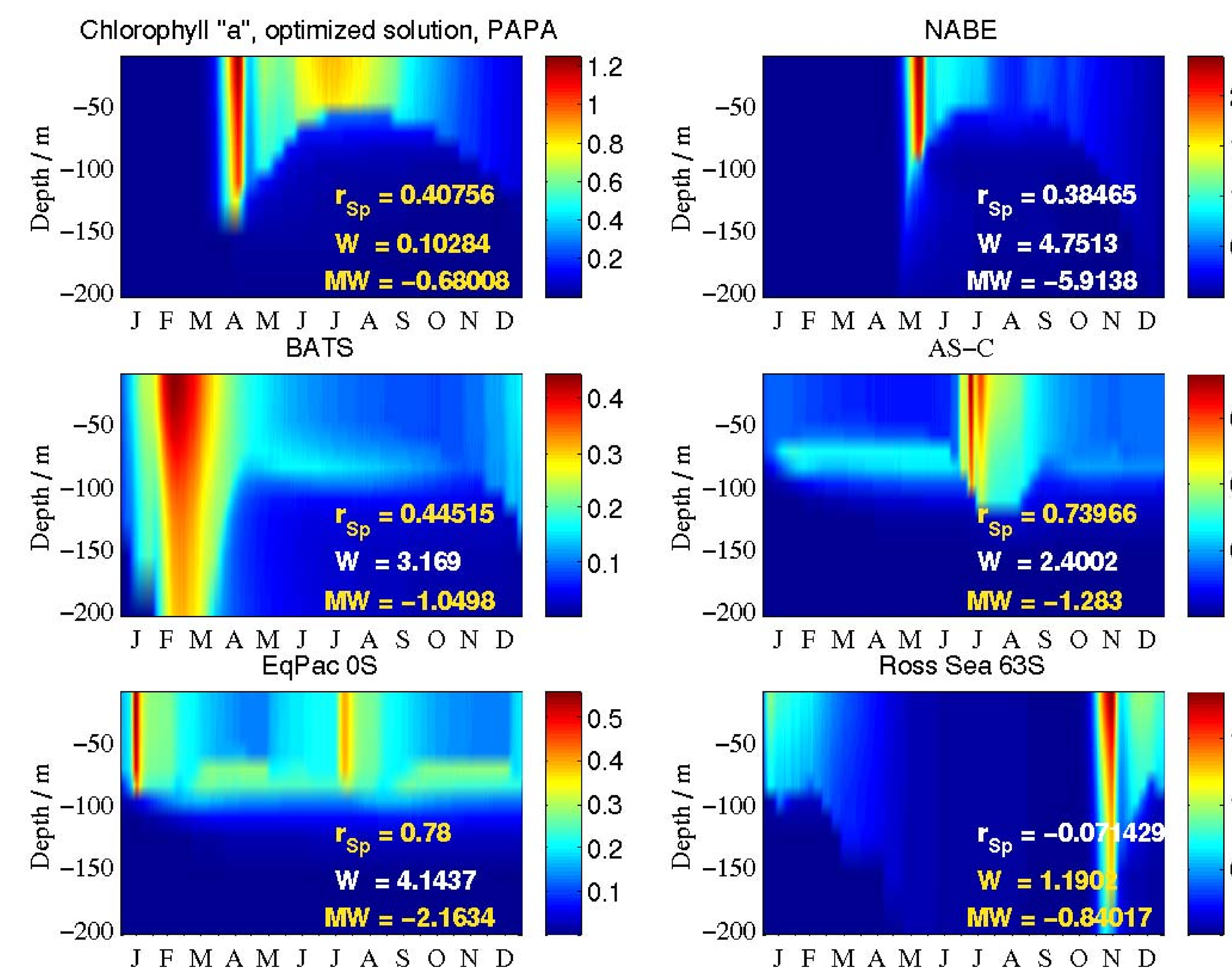

Figure 2: Chlorophyll "a" simulated by the size-dependent NPZD eco model at 6 sites. The statistics criteria values of yellow color indicate sufficient agreement between model and observed chlorophyll, with respect to a certain criterion.

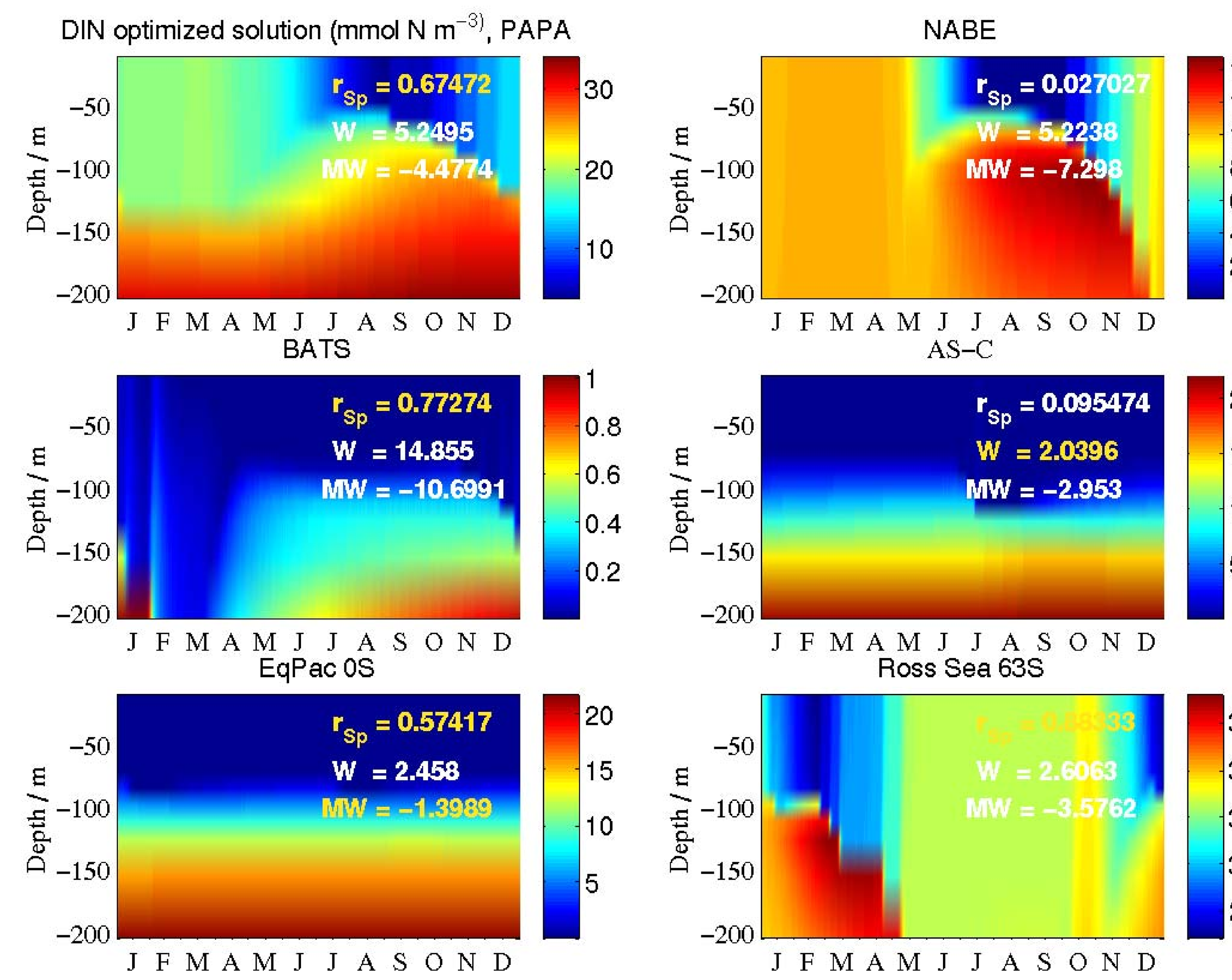

Figure 3: The time evolution of dissolved inorganic nitrogen simulated by the size-dependent NPZD eco model at 6 sites. The statistics criteria values of yellow color indicate sufficient agreement between model and observed concentrations, with respect to a certain criterion.

See Table 1 and Table 2 to compare the quality of SD NPZD ecosystem model performance against a similar experiment carried out with a simple (not account for phytoplankton size structure) NPZD model version. (Light brawn color indicates that some improvement is achieved with the new SD version, blue color means worse results)

6 Conclusions

The results make us concluding that with the present version of the ecosystem model it is not possible to reproduce the dynamics under different environmental conditions given one biological parameter set.

References

[1] Brasseur, P., Bahurel, P., Bertino, L., Birol, F., Brankart, J.-M., Ferry, N., Losa, S., Remy, E., Schroeter, J., Skachko, S., Testut, C.-E., Tranchant, B., van Leeuwen, P.-J., Verron, J., 2005. Data assimilation in operational ocean forecasting systems: the MERCATOR and MERSEA developments (submitted).

[2] Kivman, G.A., 2003. Sequential parameter estimation for stochastic systems. *Nonlinear Process. Geophys*. 10, 253-256.

[3] van Leeuwen, P.J., 2003. A truly variance minimizing filter for large-scale applications,*Mon. Weather Rev*. 131, 2071-2084.

[4] 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.

http://www.awi-bremerhaven.de/People/show?sloza