Parameter Dependences Arising from Calibration of a Riverine Diatom Model - Representation in Terms of Posterior Conditional Distributions

Ulrich Callies (ulrich.callies@hereon.de)
Helmholtz-Zentrum Hereon https://orcid.org/0000-0003-4837-1006

Christopher G. Albert
Max-Planck-Institute for Plasma Physics

Udo von Toussaint
Max-Planck-Institute for Plasma Physics

Research Article

Keywords: Model calibration, Overparametrization, Posteror parameter dependence, Markov Chain Monte Carlo, Bayesian network

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-445734/v1

Abstract

We address the analysis and proper representation of posterior dependence among parameters obtained from model calibration. A simple water quality model for the Elbe River (Germany) is referred to as an example. The joint posterior distribution of six model parameters is estimated by Markov Chain Monte Carlo sampling based on a quadratic likelihood function. The estimated distribution shows to which extent model parameters are controlled by observations, highlighting issues that cannot be settled unless more information becomes available. In our example, some vagueness occurs due to problems in distinguishing between the effects of either growth limitation by lack of silica or a temperature dependent algal loss rate. Knowing such indefiniteness of the model structure is crucial when the model is to be used in support of management options. Bayesian network technology can be employed to convey this information in a transparent way.

Full Text

This preprint is available for download as a PDF.

Figures

Figure 1

The Elbe River with station Geesthacht where the chlorophyll a and silica observations under study were taken. Some aspects of model forcing were obtained from stations Neu Darchau (river discharge), Schnackenburg (temperatures in 1997) and Schmilka (silica). The map also indicates the four most important tributaries. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

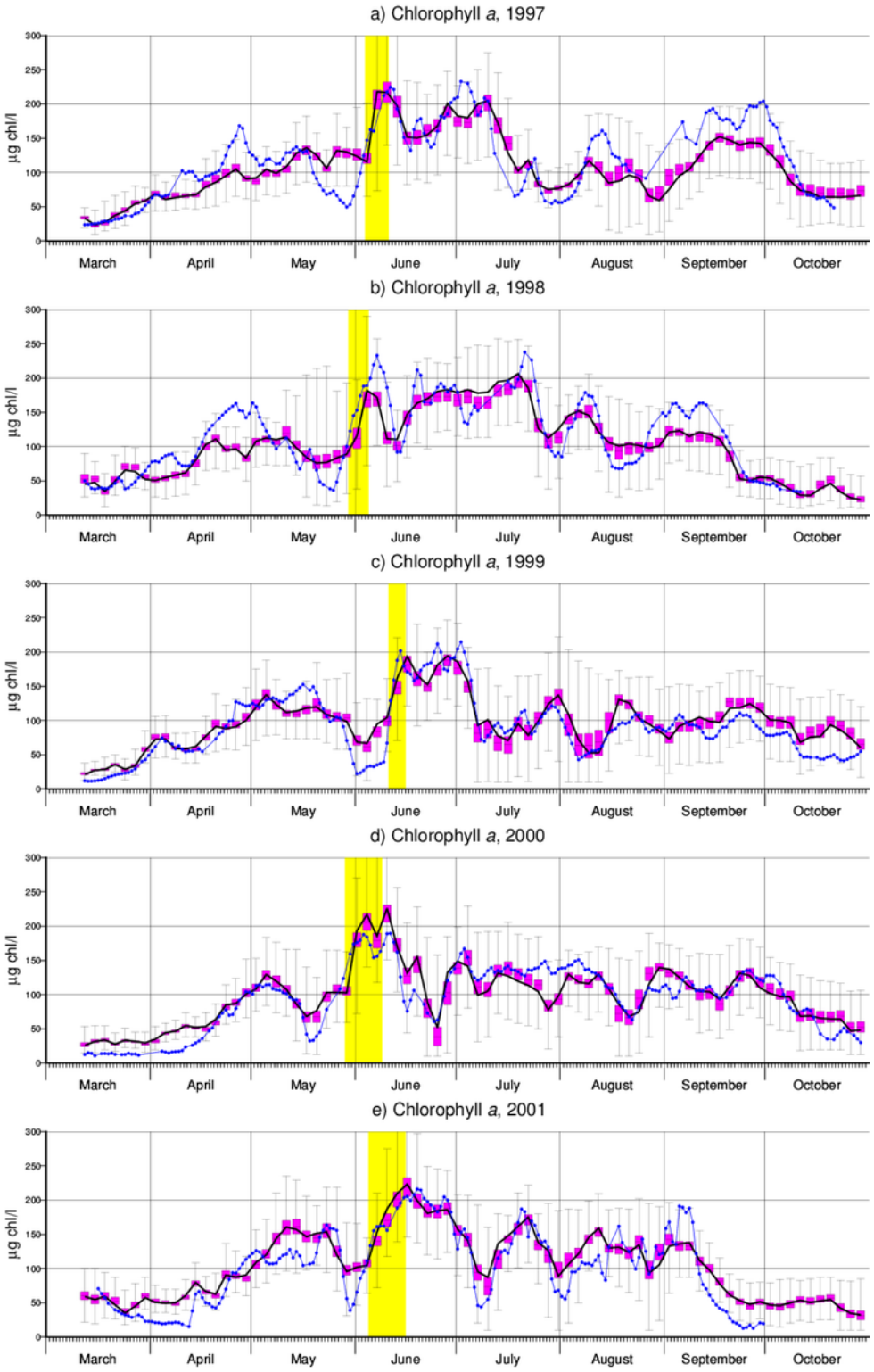


Figure 2

Chlorophyll a observations (blue) and corresponding simulations optimized to reproduce chlorophyll a observations in the five year period 1997-2001. Black lines represent the simulation for which the cost function (Eq. (10)) with σchl = 5 µg chl/l assumes a minimum value. Box-Whiskers-Plots represent the spread among simulations based on the 106 feasible parameter sets obtained from MCMC. Yellow bars indicate periods during which in the model assimilation of silica was abandoned (see Section 2.2).

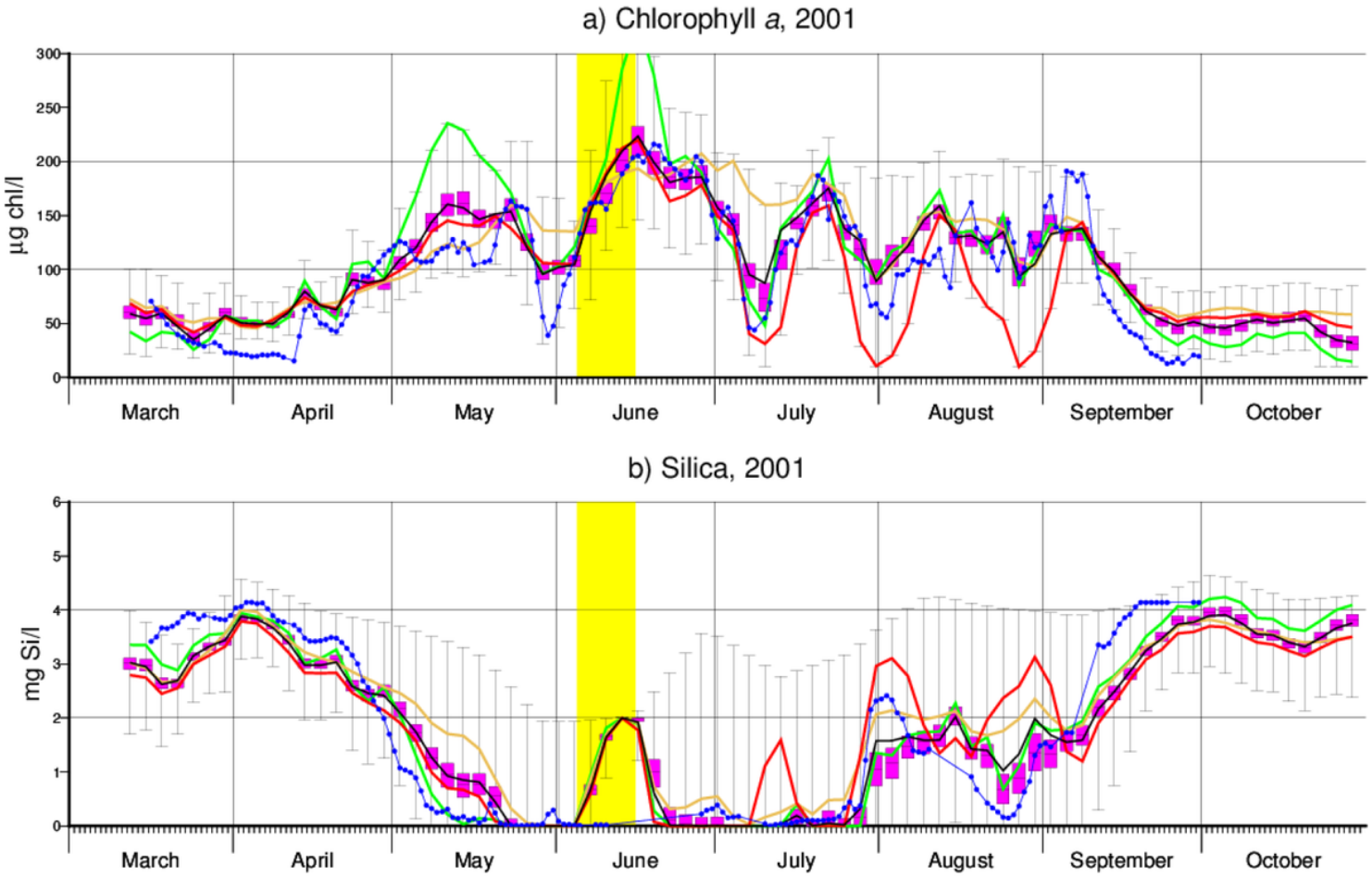
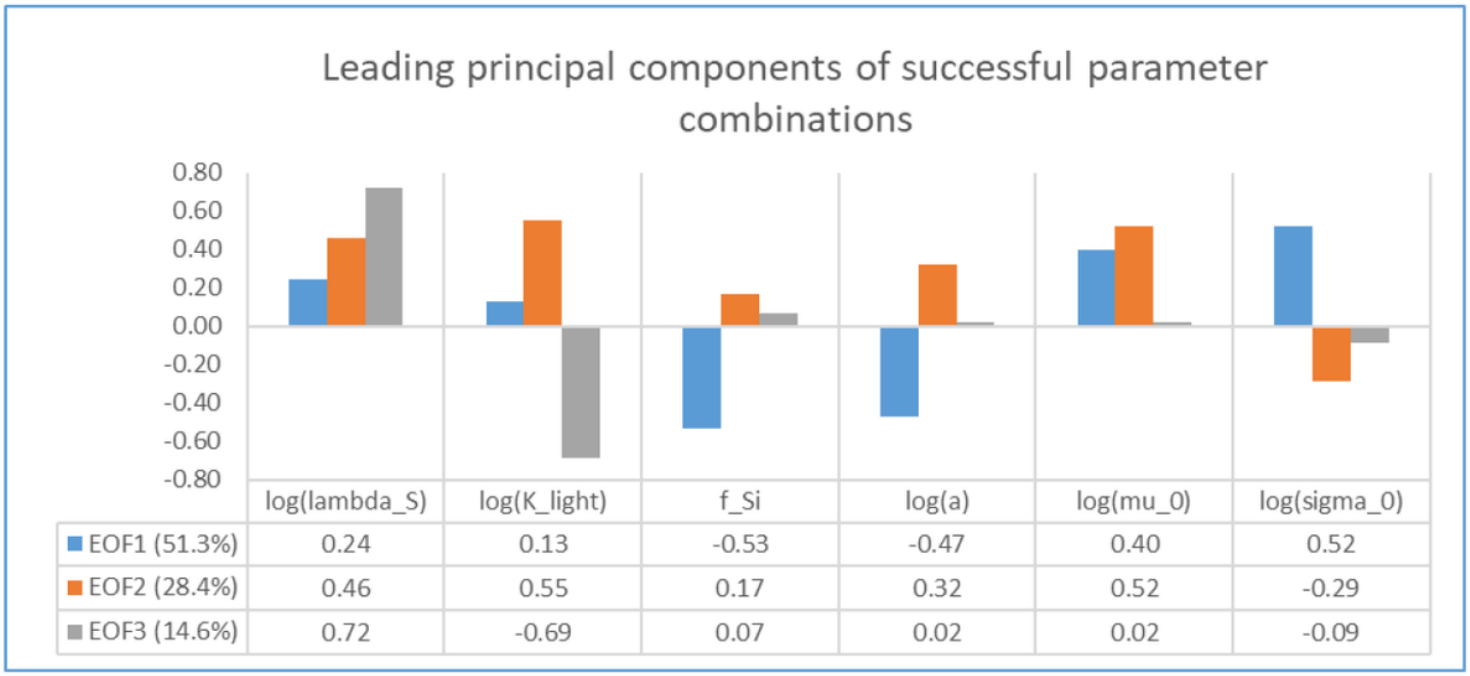


Figure 3

Upper panel: Data (blue) and best fitting simulation (black) of chlorophyll a including uncertainties (Box-Whiskers-Plots), copied from Fig. 2e. Additionally, three simulations are shown that produce the maximum simulation at 11 May (green), 10 July (brown) or the minimum value on 31 July (red). The underlying parameter sets are listed in Table 2. Lower panel: Corresponding data, simulations and simulation uncertainties for SiO2. Yellow bars indicate periods during which in the model assimilation of silica was abandoned (see Section 2.2).



	log(lambda_S)	log(K_light)	f_Si	log(a)	log(mu_0)	log(sigma_0)
EOF1 (51.3%)	0.24	0.13	-0.53	-0.47	0.40	0.52
EOF2 (28.4%)	0.46	0.55	0.17	0.32	0.52	-0.29
EOF3 (14.6%)	0.72	-0.69	0.07	0.02	0.02	-0.09

Figure 4

Principal component analysis applied to feasible parameter combinations obtained for the period 1997-2001. The graph shows three leading empirical orthogonal functions (EOFs) with corresponding PCs jointly accounting for approximately 94% of total parameter variability.

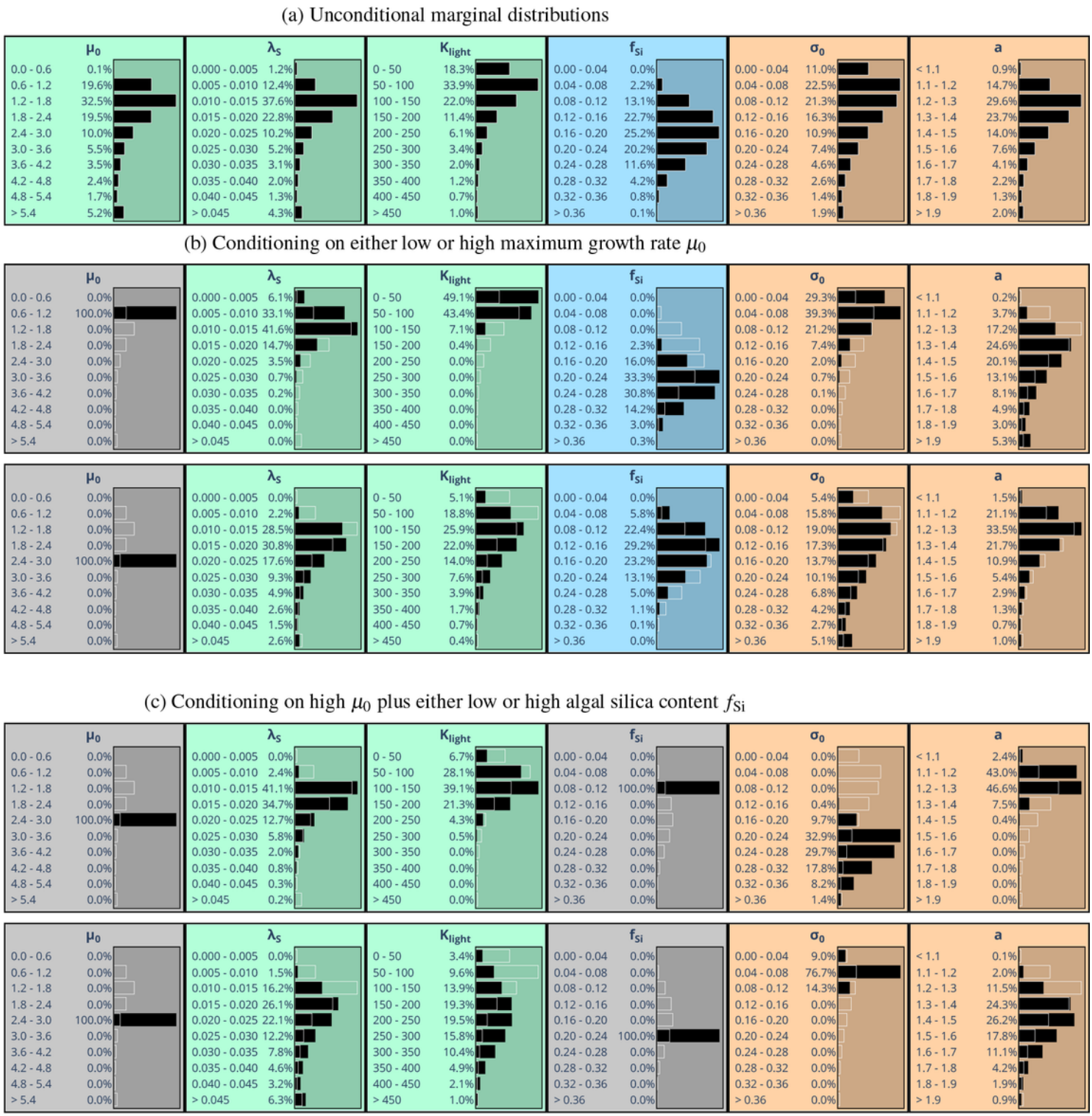


Figure 5

Each line combines 6 histograms that represent posterior marginal distributions of calibrated parameters (black bars). Background colours are used for grouping parameters into those related to algal growth (green), silica content (blue) and algal loss (brown). Grey coloured histograms indicate that specific evidence regarding the respective parameter has been entered. To ease comparison, white contours in conditional distributions reproduce the unconditional distributions.

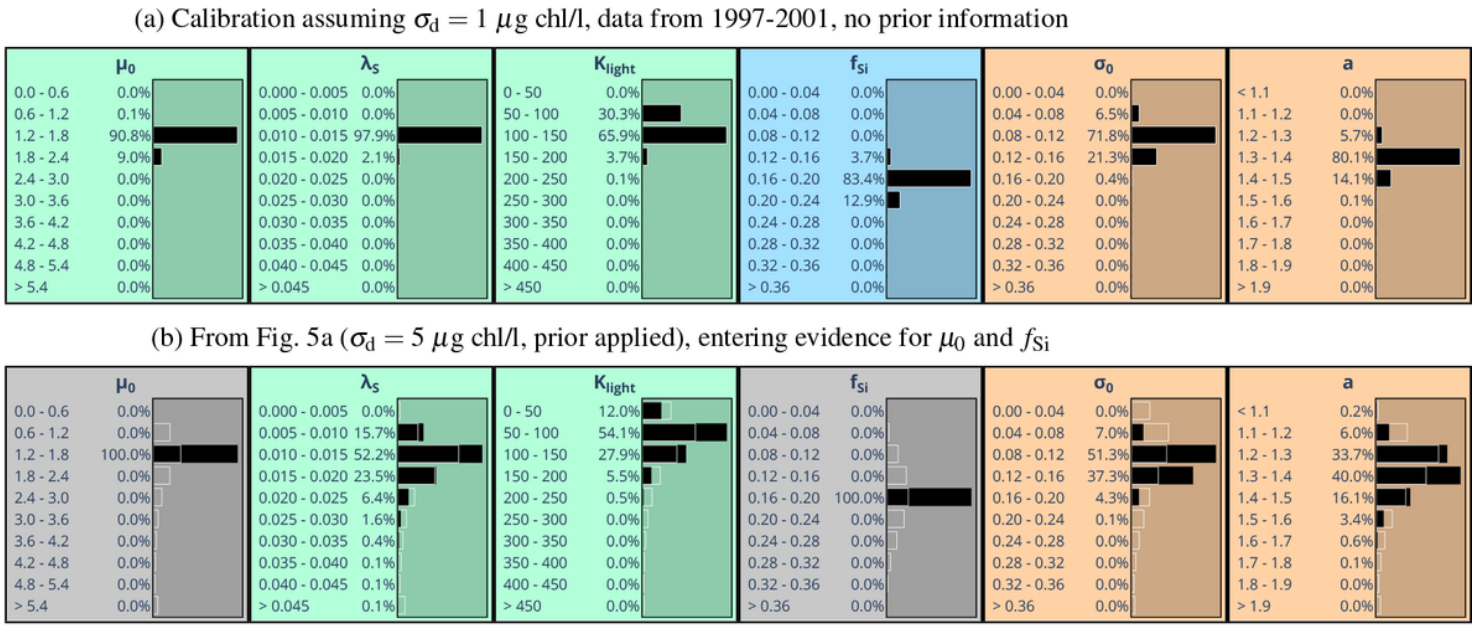


Figure 6

Narrow marginal distributions obtained assuming a small observational error schl = 1 µg chl/l, without provision of prior information on feasible parameter values. (b) Distributions obtained from Fig. 5a (σchl =5 µg chl/l, prior applied), evidence for µ0 and fSi being entered. White contours indicate unconditional distributions.

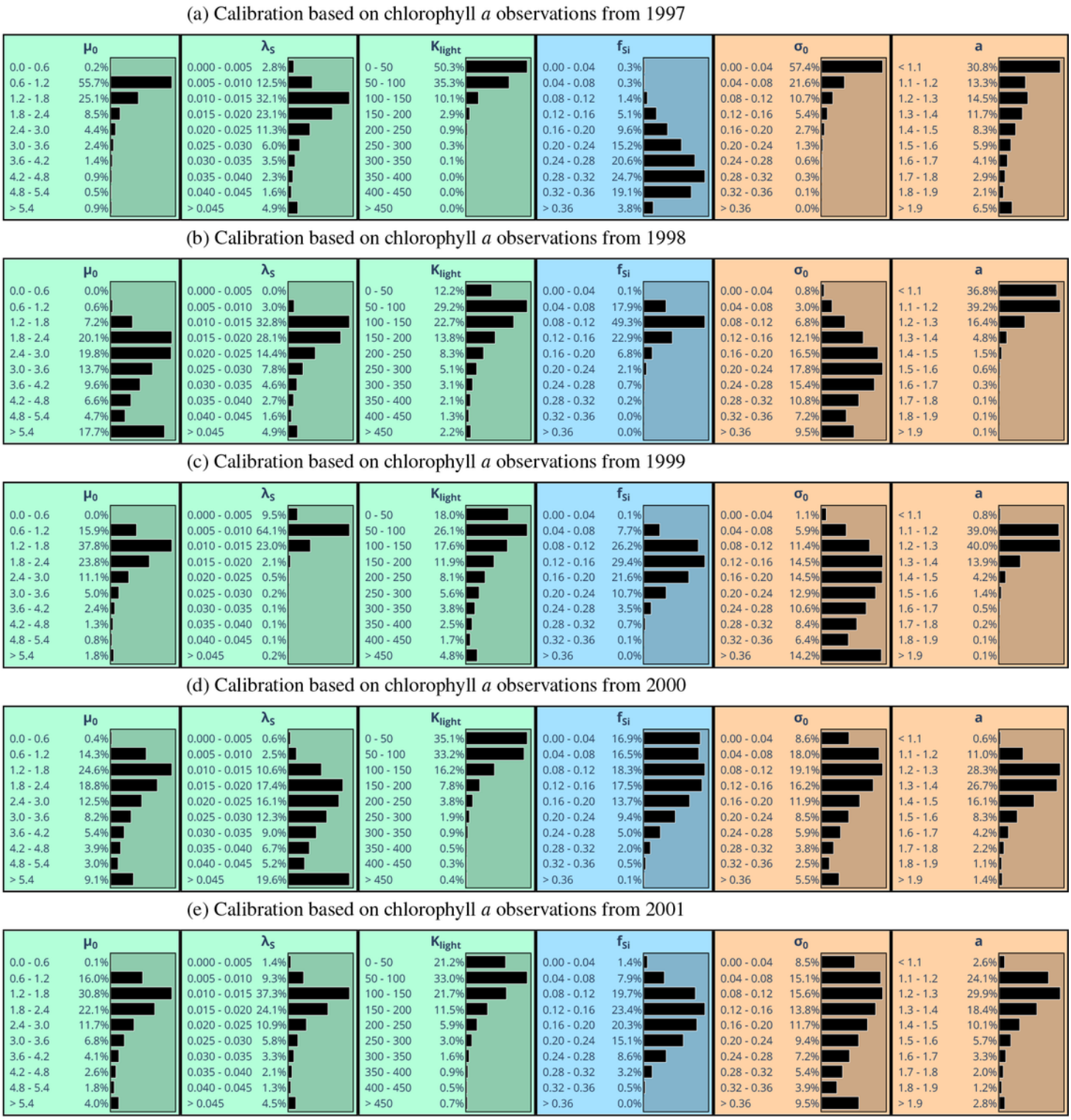


Figure 7

Marginal distributions of BNs calibrated using chlorophyll a data from different years. The overall setup agrees with that underlying Fig. 5a, apart from the different time periods model calibration refers to.

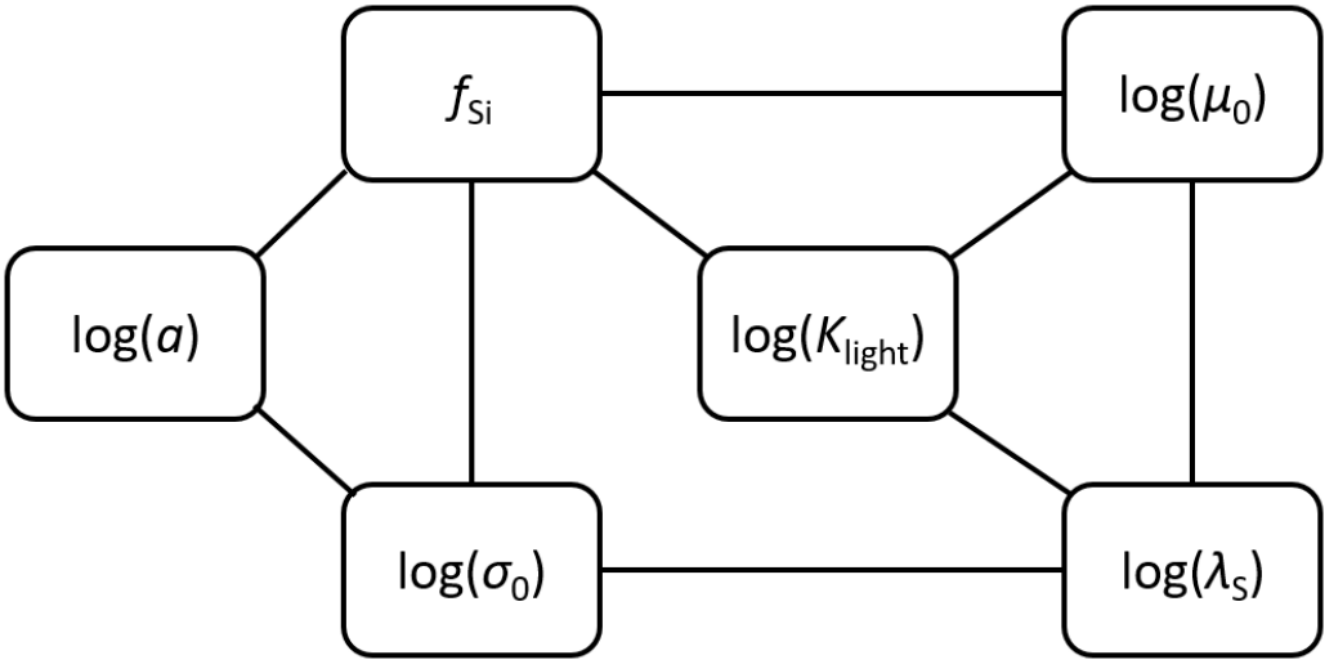


Figure 8

GGM fitted to parameter combinations that proved successful for the years 1997-2001. In the GGM, 6 out of 15 undirected edges representing partial correlation were discarded.

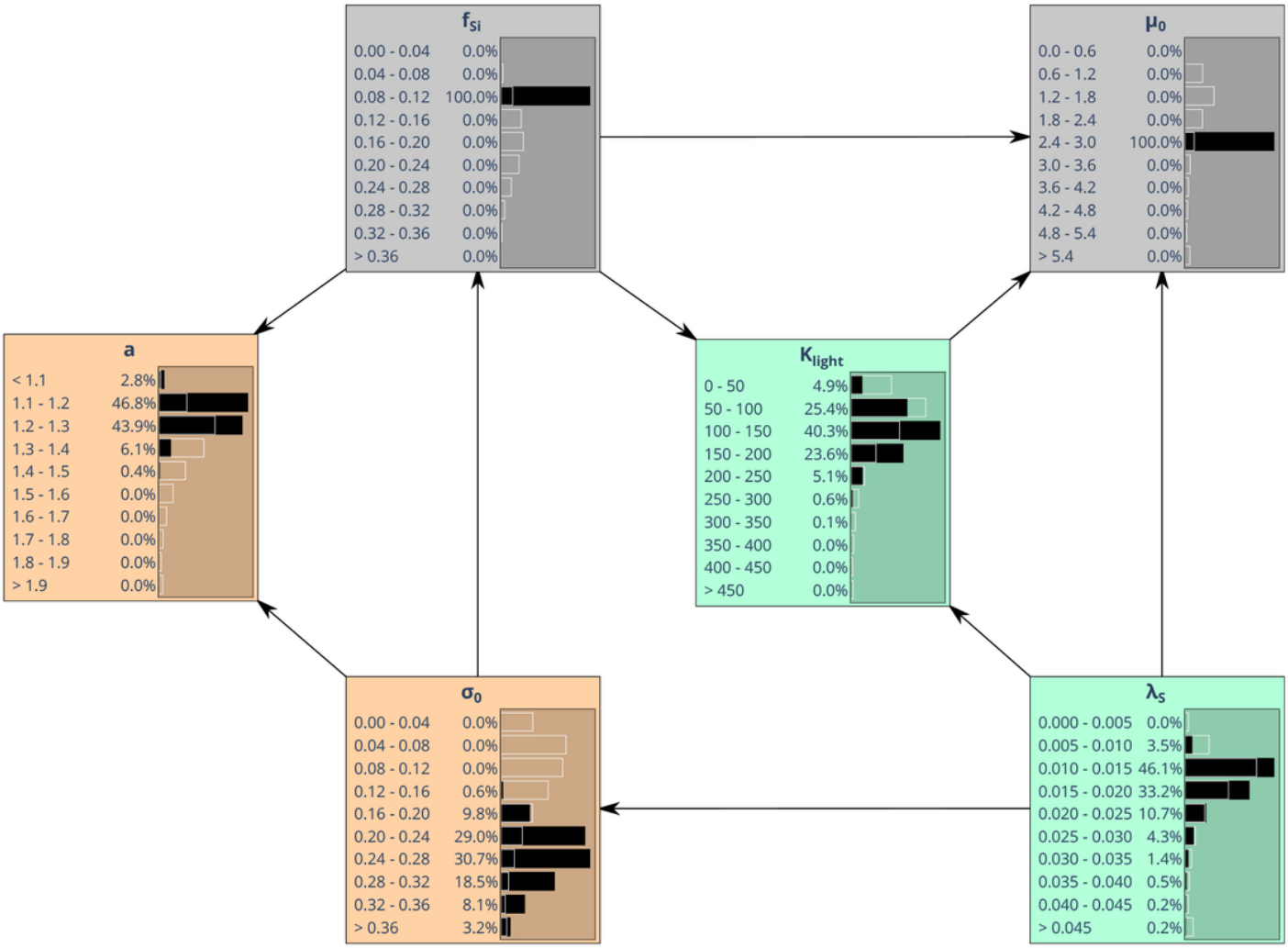


Figure 9

A BN with directed edges only where undirected edges exist in the GGM (Fig. 8). The BN is shown in a state after evidence for both μ0 and fSi was entered. Calculations were performed using Netica. Conditional marginal distributions obtained from the truncated BN well reproduce those shown in Fig. 5c. White contours indicate unconditional distributions.

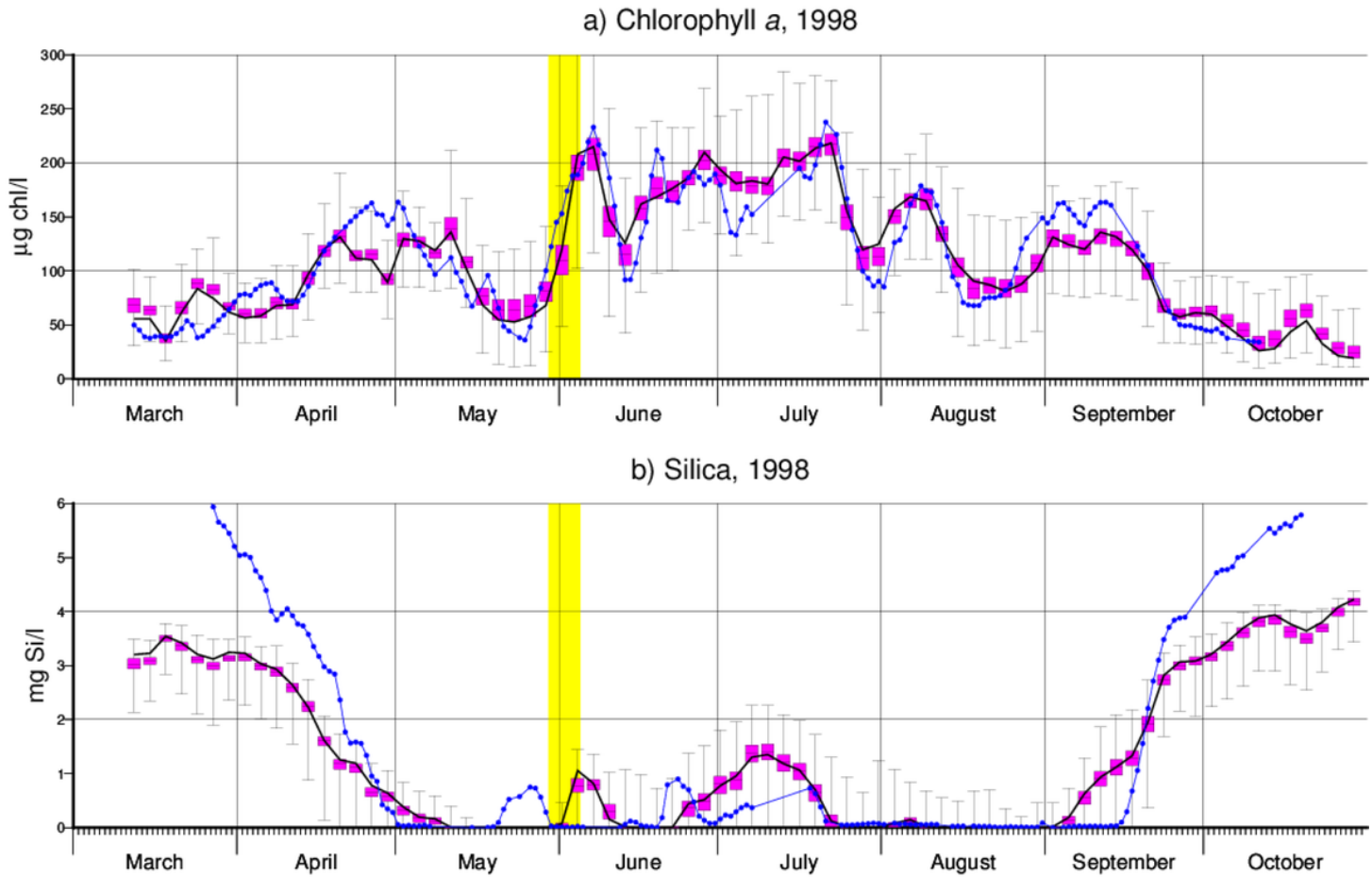


Figure 10

Chlorophyll a observations (blue) and corresponding simulations optimized to reproduce chlorophyll a observations in the year 1998. The silica observations shown were not used for model calibration. Box-Whiskers-Plots represent the spread among simulations based on the 106 parameter sets obtained from MCMC. Black lines represent the simulation for which the quadratic cost function for chlorophyll a data assumes a minimum value. Yellow bars indicate periods during which in the model assimilation of silica was abandoned (see Section 2.2).