A statistical framework for risk assessment of environmental effects of pollutants from diffuse sources: The single site case

SAMBA/26/02
Gudmund Høst
Thorjørn Larssen
David Hirst
Ragnar Bang Huseby
Jack Cosby
Tore Høgåsen

NR Norsk Regnesentral
ANVENDT DATAFORSKNING

Tittel/Title:

A statistical framework for risk assessment of environmental effects of pollutants from diffuse sources: The single site case

Forfatter/Author:
Gudmund Høst, Thorjørn Larssen[1], David Hirst, Ragnar Bang Huseby, Jack Cosby[2] and Tore Høgåsen[1]

[1]Norsk institutt for vannforskning, Postboks 173 Kjelsås, 0411 Oslo
2) University of Virginia, Dept. of Environ. Sciences. Charlottesville, Virginia 22903 USA

Sammendrag/Abstract:

In this report we describe the development of a common statistical framework for risk analysis of diffuse pollutants with space-time characteristics, restricted to the single site case. This framework is/will be tested with available biogeochemical models and field data, and it will form the basis for scaling up to regional risk assessment. Our model provides a link from pollutant exposure to the probability of adverse effects, thereby forming a quantitative basis for decision-makers. The methods will be general, enabling use within a broad range of applications. We give results for some pilot runs of the framework applied to surface water acidification (and effects on fish mortality) for the Birkenes site.

1. Introduction

Environmental risk assessment is the science that identifies stressors that may alter ecosystems and quantifies the probability of adverse effects on those ecosystems and impacts to humans. In environmental risk assessment, the uncertainties concerning potential environmental effects are explicitly recognised and quantified (Hunsaker et al. 1990).

Traditional concepts and methodology in risk assessment are developed mainly for application to single sites or small geographical areas. Furthermore, the current methodology is usually applied in a static mode and does not account for temporal dynamics of pollutant loads or its effects. Many important environmental problems are caused by diffuse pollutants, which vary in space and time. Examples include effects of long-range transportation of air pollutants, radioactive substances, and persistent organic compounds. Our knowledge of diffuse pollutants and their environmental effects is uncertain due to limited sampling and inherent space-time dependencies of pollutant fields. Such uncertainties need to be quantified and propagated through relevant models to generate space-time specific probability statements for the environmental effects.

The development of a statistical framework for risk assessment of such diffuse pollutants is the primary focus of the project "Statistical techniques for risk assessment of environmental effects of pollutants from diffuse sources", funded by the Research Council of Norway, grant 148307/720. Within the CLRTAP/UN-ECE, the current environmental standards, such as critical loads, do not fully reflect probabilities of harmful effects. Amongst policy makers, there is an increasing focus on uncertainties and the limitations of the current concepts. Within the current project, we propose to develop an alternative way of defining environmental standards, based on spatial-temporal specific probabilities of harmful effects or recovery. A risk assessment framework may also give a better separation of science and decision-making.

In the proposed single site framework, we regard the diffuse pollutants as random processes in time. These are input to a biogeochemical model for water and soil chemistry. The results from the biogeochemical model are then used as input to an exposure model to quantify the harmful effects. By using a probabilistic framework, we obtain a probability distribution for the harmful effects given all the available information and its related uncertainties. We illustrate the methodology with some test runs for the MAGIC biogeochemical model (Cosby e al., 1985; Cosby, et al., 2001).

2. Statistical framework

Denote by f the biogeochemical model, which maps inputs α to outputs β. Data related to the inputs α are denoted by x, and data related to the outputs β are denoted by y. The outputs $\beta = f(\alpha)$ of the biogeochemical model are input to the exposure model $g()$. We denote the environmental effect by γ, and data on these effects are denoted by z. Thus, the environmental effect induced by pollutant input α is $\gamma = g(\beta) = g(f(\alpha))$. In some situations, the data may depend on (nuisance) parameters that are independent of the models. We decompose the parameter vectors into structural and nuisance by subscripting, i.e. $\alpha = \alpha_s + \alpha_n$, etc. The structure is depicted graphically in Figure 1.

We assume that there is prior knowledge available on the parameters in the form of prior probability distributions. Such knowledge could in practise be taken from related studies at the current site or in similar biogeochemical environments. The information in the data may be quantified through likelihood functions. Our purpose is to infer the posterior distribution of the inputs α, given the models f, g and the data x, y, z. By Bayes' theorem, this posterior probability density is

$$p(\alpha \mid x, y, z, f, g) \propto p(\alpha) L(x, y, z; \alpha, f, g)$$

From this posterior distribution, we may generate all information of interest regarding the effects model. For example, a random value α' drawn from $p(\alpha|x,y,z,f,g)$ gives a random exposure $\gamma' = g(f(\alpha'))$. By drawing a large number of such α's we may fit a probability density curve to the resulting γ's. This, in turn,

characterises all the statistical properties of the exposure (mean, standard deviation, median and percentiles). There are several possible alterations to the structure described above, which may have to be involved in practise. This includes nuisance parameters, as described above, and additional parameters specific to either of the models.

The models f and g are run sequentially, and, given α, the data x, y and z are conditionally independent. Thus, we may decompose the posterior distribution

$$p(\alpha \mid x, y, z, f, g) \propto p(\alpha)L(x; \alpha)L(y; f(\alpha))L(z; g(f(\alpha)))$$

The posterior distribution of α may be complicated or impossible to write down analytically. Fortunately, it may be possible to obtain samples from the posterior by Monte Carlo simulation. One option is to use the sampling importance sampling algorithm (SIR) (Rubin, 1988). This involves sampling from a known proposal distribution, then resampling according to some SIR weights. Another alternative is to use a Markov Chain Monte Carlo (MCMC) method. This algorithm generates samples from a random walk, which eventually converges to the true posterior. In the current example, we have used the Metropolis-Hastings MCMC algorithm (Metropolis et al., 1953).

3. Example

The critical load concept is established as a helpful tool for policy makers in the international cooperation against acid deposition. However, there is an increasing awareness of, and interest in, the uncertainties involved in the applications, such as spatial uncertainty, as well as the importance of the lack of a time component (dynamics). A modification of the critical load concept, placing it in a risk analysis context, has been suggested (Barkman et al., 1995). This will lead to a more clear reference to probability and political acceptability, as compared to the present methodology where critical loads are simple cut-off values (Skeffington, 1999). In order to address the time component a dynamic acidification model must be applied. In the present application, we use the predictive acidification model MAGIC. Through the application of MAGIC we may link the future development in deposition of acidifying compounds to the future water chemistry by taking specific chemical and physical characteristics into account. By accounting for uncertainties, water chemical parameters can be related to biological response such as probability of fish survival. To address the spatial uncertainty, we will need to run the acidification model repeatedly for a large number of spatial locations. This calls for automatic calibration of the model, which will be described in the following.

We have applied the framework described above to calibrate the MAGIC model to data from Birkenes. MAGIC is a deterministic model which is usually set up with 30-50 input parameters and a similar number of output parameters. Also available is measurement data corresponding to the model output parameters. Currently, no automatic method is available for model calibration that can include information on time trends in the observations. We do this by applying MCMC techniques. This involves specifying prior and likelihood, and writing a computer code that runs MAGIC repeatedly with parameters suggested by the MCMC scheme. Results may be analysed after the algorithm has converged.

In the present application, we show results only for the biogeochemical model f. Furthermore, we omit the likelihood for the input data, and specify these directly through prior distributions. Finally, nuisance parameters are fixed and do not enter the analysis in this example. The dependency structure for the example is shown graphically in Figure 2. The simplified posterior reads

$$p(\alpha \mid y, f) \propto p(\alpha)L(y; f(\alpha))$$

A detailed description of all input parameters α and their prior distributions is given in Table 1. Field measurements, expert judgements or experience from related studies are used to suggest minimum and maximum values in independent uniform distributions for each input parameter. Parameters that are fixed do not enter the statistical analysis and such parameters will not be referred to in the following. Currently, the software can only use input from a single year, in this case 2000. To mitigate the effects of measurement

error we have smoothed each wet deposition concentration value by fitting a linear trend to the last 5 years of observations, then using the fitted value for 2000.

A description of the output parameters β is given in Table 2. The output parameters are generated by the MAGIC model f. Thus, the prior for the output parameter vector is induced by $\beta = f(\alpha)$ when $\alpha \sim p(\alpha)$. The prior distribution for the output may be a complicated distribution with dependencies between the individual outputs, due to the effect of the model f.

The data used in the calibration is described in Table 3. For the data, we use a Gaussian likelihood. The likelihood function involves a variance-covariance matrix for the observations. This matrix is a nuisance parameter, in which we have no particular interest. In principle, these parameters should be assigned priors and be integrated out of the posterior. In the present calibration exercise we have fixed these parameters to allow for easier interpretation of the results. Soil properties are only measured once, and the measurement error is specified by expert opinion. Surface water chemistry is measured weekly and volume weighted annual mean values calculated yearly through 1974-2000. The variance of each variable is fixed at 10% of the empirical variance for the observation period. Finally, we set all correlations to zero. The likelihood is

$$L(y; \beta) \propto \exp\left\{ -\frac{1}{2} \sum_{j=1}^{5} \sigma_j^{-2} (y_{1j} - \beta_{1j})^2 - \frac{1}{2} \sum_{i=2}^{28} \sum_{j=6}^{14} \sigma_j^{-2} (y_{ij} - \beta_{ij})^2 \right\}$$

Here, the first sum is over the five soil observations, while the second sum is over the 14 surface water measurement for all 27 years.

The posterior of the input parameters $p(\alpha|y)$ was obtained by MCMC sampling of α, using f to match the observations y. A general description of the MCMC algorithm can be found in numerous references, see for example Besag et al., 1999.

Inference on the output parameters β can be derived from its posterior distribution, which is the distribution of $\beta = f(\alpha)$ when $\alpha \sim p(\alpha|y)$.

4. Results

The algorithm was run with 2000 iterations, giving a 31-dimensional posterior for the input parameters α. This posterior induces a 14-dimensional posterior for the outputs β. The full results are by far too extensive for presentation in the present report. Thus, we only show the results for some selected parameters and combination of parameters. These are

Input parameters:
- Sum of base cation weathering in soil (WeCa+WeMg+WeK+WeNa)
- Initial base saturation in soil (ExCai+ExMgi+ExKi+ExNai)

Output parameters:
- Estimated base saturation in soil (ExCa+ExMg+ExK+ExNa)
- Estimated charge balance alkalinity in surface water (ANC)
- Estimated sulphate concentration in surface water (SO_4)
- Estimated calcium concentration in surface water (Ca)

Figure 3 shows a trace of base cation weathering in soil. This is a sum of 4 input parameters. Initially, the algorithm suggests values from the full range of the prior, but the parameter seems to stabilize after roughly 1000 iterations of the MCMC algorithm. The initially large fluctuations are typical for the burn-in of MCMC algorithms, and the 1000 iterations burn-in phase is characteristic for all parameters in the present application. Thus, we may take the last 1000 iterations as samples from the joint posterior distribution. Figure 4 shows the prior distribution for base cation weathering and the fitted posterior density. The rather flat prior with support on values between 0 and 400 arises as the convolution of 4 uniform distributions. The estimated posterior distribution is much more peaked with probability mass concentrated around 48.

Figure 5 shows the MCMC trace for initial base saturation. After som initial fluctuations, the values seem stuck around 100% for many iterations before they stabilize after some 1000 iterations. From Figure 6, we see that the prior gives much weight to the larger values, which may be part of the explanation of the trace plot in the previous figure.

Figure 7 shows the MCMC-trace for estimated base saturation in soil. Again, we see that the values stabilize after about 1000 iterations. In Figure 8, the estimated values for the years 1970-2000 are shown, with an estimated 90% confidence. The filled square at year 2000 indicates the measured value. The mismatch between estimated and measured values is in part related to choice of priors, but also to measurement error, spatial variability and effects from other data used in the calibration. In a later phase of this project this mismatch will be subject to further investigations.

For the surface water parameters, the output is a multiple time series. Figure 9 shows the trace of ANC for 2000, and similar traces may be produced for other years in the calibration period. In Figure 10, the posterior mean of ANC (blue line) is shown together with a 90% confidence band (broken red lines) on these parameters. Also shown is the measured values for each year (filled squares). It is seen that the calibrated values are systematically lower than the observed values. Furthermore, the confidence band seems too narrow and includes very few observed values.

Figures 11-14 give trace plots and time series for SO_4 and Ca. Again, we see stabilization of parameter values after 1000 iterations (Figures 11 and 13). Figure 12 suggests systematic error in SO_4, while Figure 14 suggests that Ca gives a fairly good fit to the observations. The confidence bands for the parameters seem too narrow (Figures 12 and 14). The mismatch for SO_4 may reflect a mismatch between the prior for the inputs and the observations used in model calibration. Based on such an observation the prior for the inputs should be revisited and evaluated. In this particular case, the mismatch is most likely due to the simplification we did when not taking the time trend in the wet deposition concentrations into account, while we did so for the observations. Since 2000 was a year with extremely high precipitation amounts and hence high sulphur deposition flux, the resulting outputs will be overestimated as compared to the long term trend.

Figure 15-17 shows pairwise dependence between selected parameters. Figure 15 suggests that there is little dependence between soil weathering parameters. This is reasonable from a geochemical point of view as release of these cations from weathering depends on the composition of minerals, and the individual base cations may occur in different minerals. Figure 16 suggests that the parameters that sum up to the initial base saturation are much more correlated. In particular, exchangeable calcium and magnesium are positively correlated. This is acceptable in terms of chemical processes. Figure 17 shows posterior estimated output base saturation from MAGIC. In contrast to initial exchangeable Ca and Mg, Figure 17 suggests negligible correlation between the estimated exchangeable Ca and Mg. The exchangeable parameter values will be driven in the same direction due to the positive correlation of the initial values, but these parameters may follow different paths through time. The result is that the calibrated exchange constants in the model will be different.

5. Discussion and concluding remarks

A statistical framework for single site risk analysis of pollutants has been suggested. The framework connects a biogeochemical model with a biological exposure model in terms of probability distributions. We have used the framework for pilot investigations on an acidification problem. A preliminary run of a Markov Chain Monte Carlo method gives promising results on automatic calibration of the acidification model MAGIC. Further studies should include the input of multiple years of observations. Furthermore, the cause of bias in estimated output parameters should be investigated and the model refined in order to reduce the mismatch between model outputs and observations. Important points are to include more input parameters and to allow for additional statistical nuisance parameters for the measurement errors in the likelihood functions. These modifications will make the framework more robust for applications to other diffuse pollutants, which we intend to study in a later stage of the project.

A natural next step is to extend the current application to incorporate a model for the effects on fish survival. The relevant data for this extension has been made accessible to the project. At a later stage, the single site framework will be extended to the regional scale, as described in phase 2 of the project description.

The developed framework provides a basis for quantitative risk assessment, which may later serve as a tool for policy makers. The results from the acidification case illustrate how the uncertainty in model predictions can be quantified as probability distributions or confidence bands for parameters of interest. From the estimated posterior distributions, we may easily run forecasts under different emission scenarios, thereby obtaining uncertainty statements for future environmental effects. Such studies are planned in phase 3 of the project, where we will focus on relevance to policy-making.

6. References

Barkman, A., Warfvinge, P., and Sverdrup, H. 1995. Regionalization of critical loads under uncertainty.Water Air Soil Pollut 85, pp. 2515-2520.

Besag, J., Green, P., Higdon, D. and Mengersen, K. 1995. Bayesian Computation and Stochastic Systems. Statistical Science Vol. 10, No. 1, pp 3-41.

Cosby, B. J., Ferrier, R. C., Jenkins, A., and Wright, R. F. 2001. Modelling the effects of acid deposition: refinements, adjustments and inclusion of nitrogen dynamics in the MAGIC model. Hydrol.Earth System Sci. 5: 499-518.

Cosby, B. J., Hornberger, G. M., Galloway, J. N., and Wright, R. F. 1985. Modelling the effects of acid deposition: assessment of a lumped parameter model of soil water and streamwater chemistry. Water Resourc.Res. 21: 51-63.

Hunsaker, C. T., Graham, R. L, Suter, G. W., O'Neill, R. V., Barnthouse, L. W., and Gardner, R. H. 1990. Assessing ecological risk on a regional scale. Environmental Management 14, pp. 325-332.

Metropolis, N., Rosenbluth, A. N., Rosenbluth, M. N., Teller, A. H., and Teller, E. 1953. Equation of State Calculations by Fast Computing Machines. J. Chem. Phys. 21 1087-92.

Rubin, D. B. 1988. Using the SIR algorithm to simulate posterior distributions. In M.H. Bernardo, K.M. an DeGroot, D.V. Lindley, and A.F.M. Smith, editors, Bayesian Statistics 3. Oxford University Press, Oxford, UK.

Skeffington, R. A. 1999. The use of critical loads in environmental policy making: A critical appraisal. Environ.Sci.Technol 33, A245-A252.

7. Figures

Figure 1: Dependence graph for the general modelling framework

Figure 2: Dependence graph for the acidification example

Figure 3: Trace of base cation weathering in soil for 2000 iterations of the algorithm.

Figure 4: Prior (blue) and estimated posterior density (red) for base cation weathering.

Figure 5: Trace of initial base saturation for 2000 iterations of the algorithm.

Figure 6: Prior (blue) and estimated posterior density (red) for initial base saturation.

Figure 7: Trace of estimated base saturation in soil

Figure 8: Time series of estimated base saturation in soil. Posterior mean (blue line), 5% and 95% quantiles (red broken lines).

Figure 9: Trace of estimated ANC in 2000.

Figure 10: Time series of ANC for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.

Figure 11: Trace of SO_4 in 2000.

Figure 12: Time series of SO_4 for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.

Figure 13: Trace of Ca for 2000.

Figure 14: Time series of Calcium for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.

Figure 15: Scatter plots of samples soil weathering parameters from the posterior distribution.

Figure 16: Scatter plots of parameters of initial soil base cation saturation from the posterior distribution.

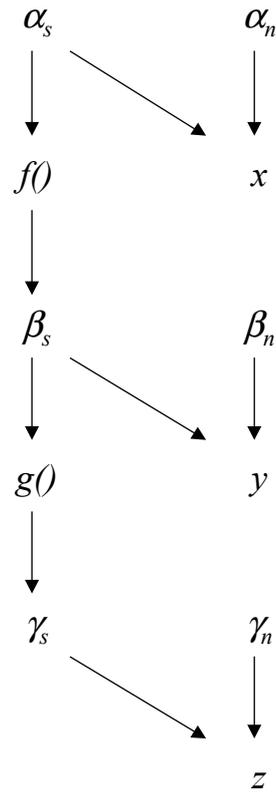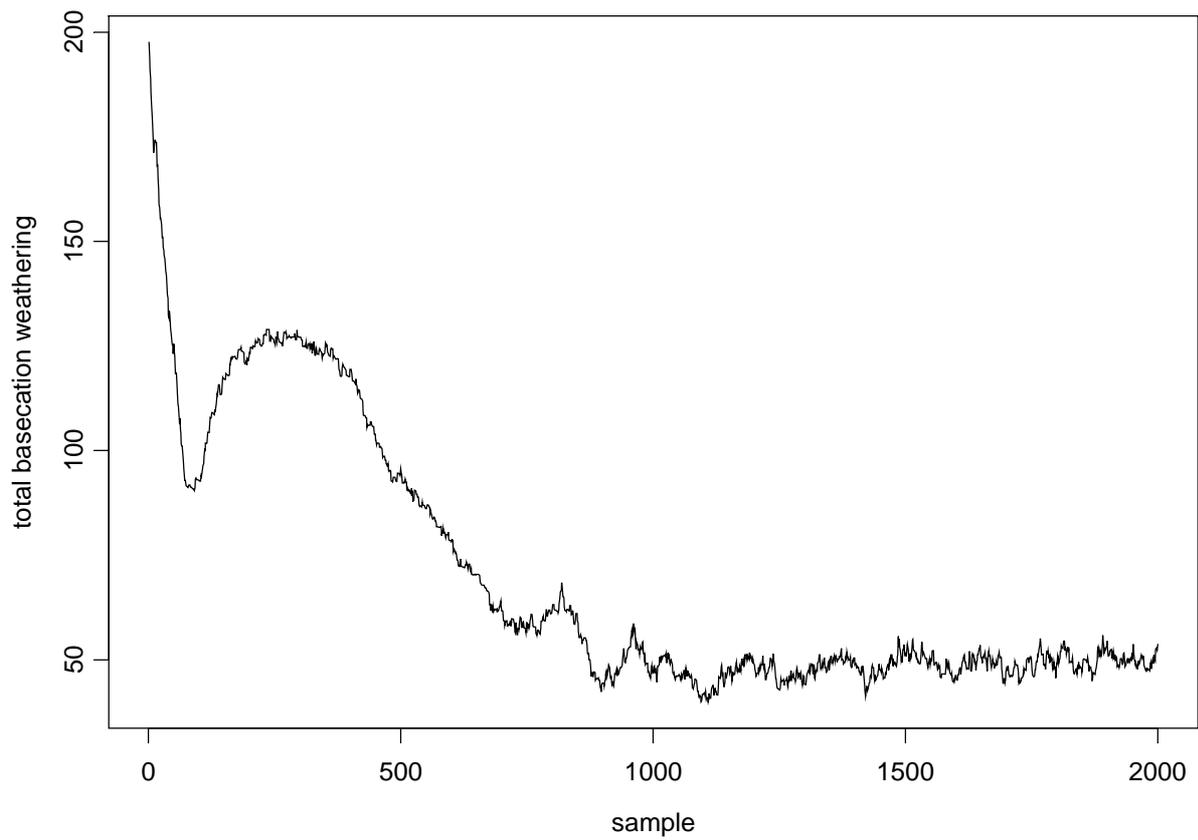Figure 17: Scatter plots of parameters of output soil base cation saturation from the posterior distribution.

$$\alpha_s \qquad \alpha_n$$

$$f() \qquad x$$

$$\beta_s \qquad \beta_n$$

$$g() \qquad y$$

$$\gamma_s \qquad \gamma_n$$

$$z$$

Figure 1: Dependence graph for the general modelling framework

$$\alpha$$

$$\downarrow$$

$$f()$$

$$\downarrow$$

$$\beta$$

$$\downarrow$$

$$y$$

Figure 2: Dependence graph for the acidification example

Total basecation weathering, trace

Figure 3: Trace of base cation weathering in soil for 2000 iterations of the algorithm.

Total basecation weathering, prior density, posterior density
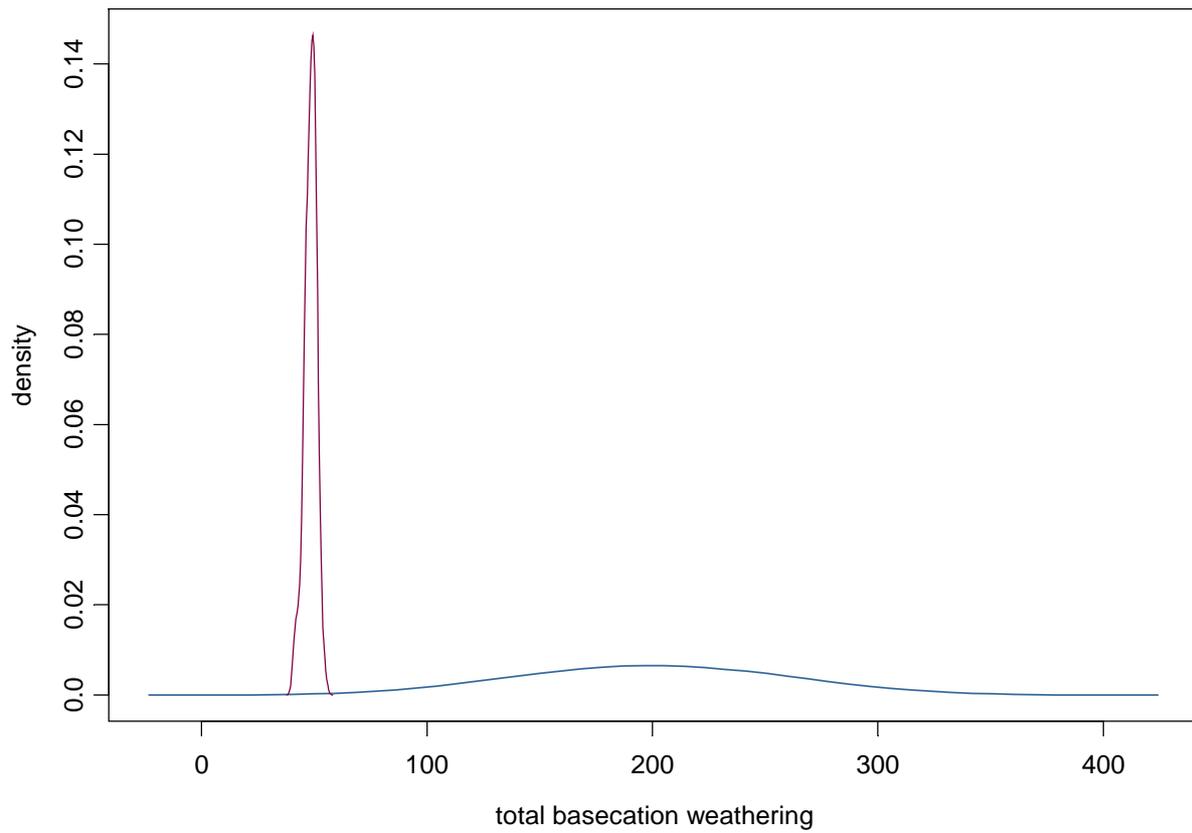
Figure 4: Prior (blue) and estimated posterior density (red) for base cation weathering.

Initial base saturation, trace
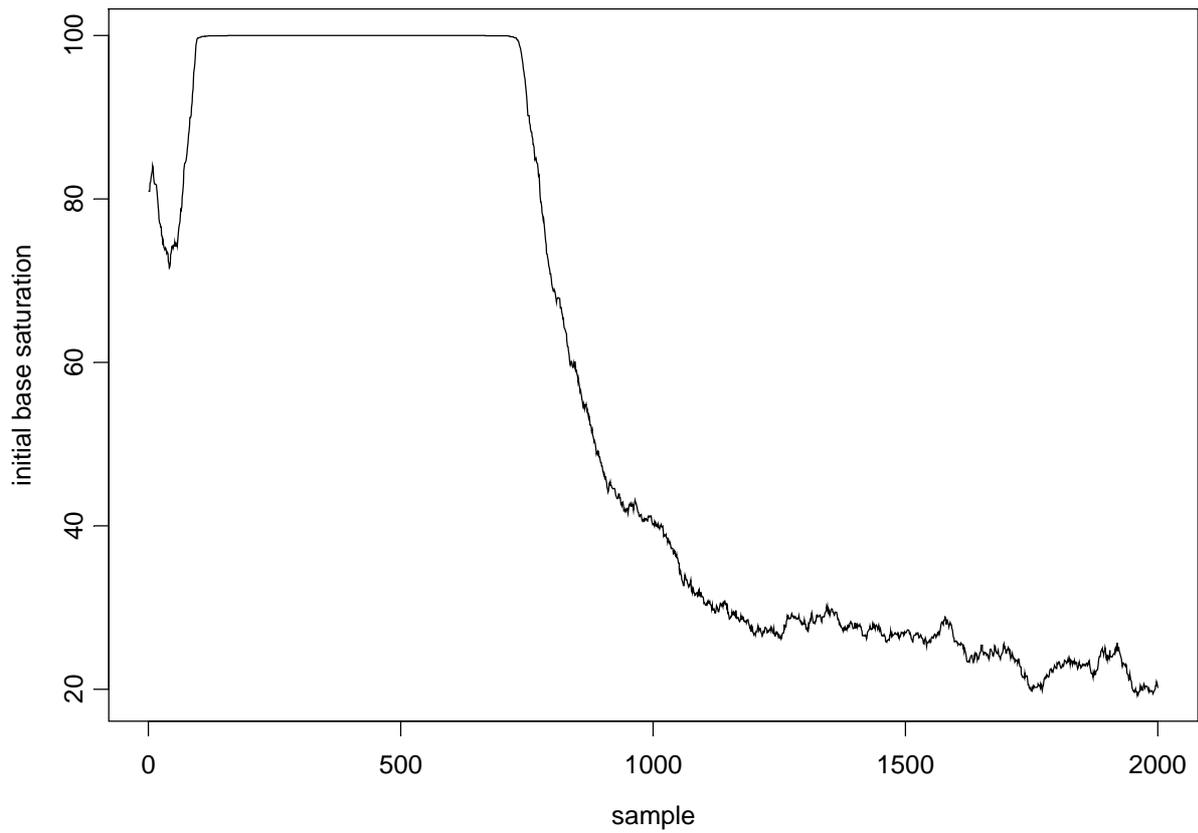
Figure 5: Trace of initial base saturation for 2000 iterations of the algorithm.

Initial base saturation, prior density, posterior density
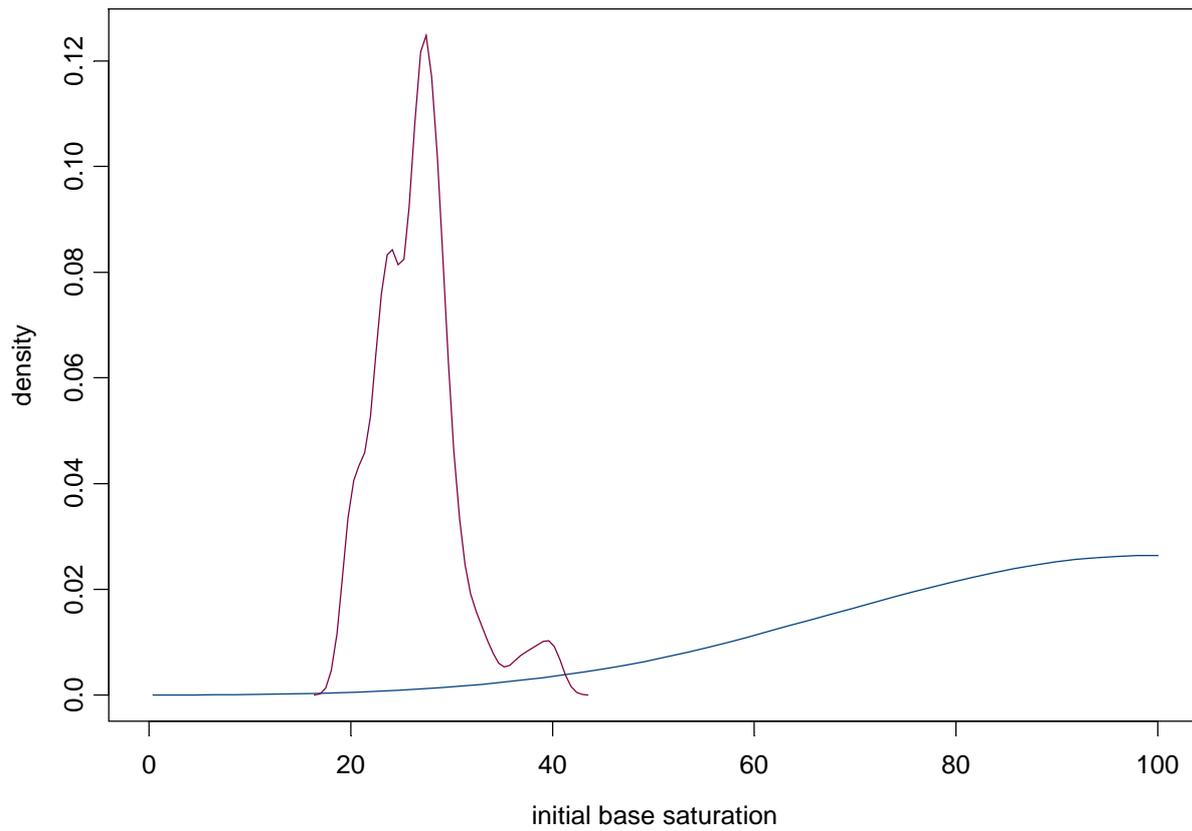
Figure 6: Prior (blue) and estimated posterior density (red) for initial base saturation.
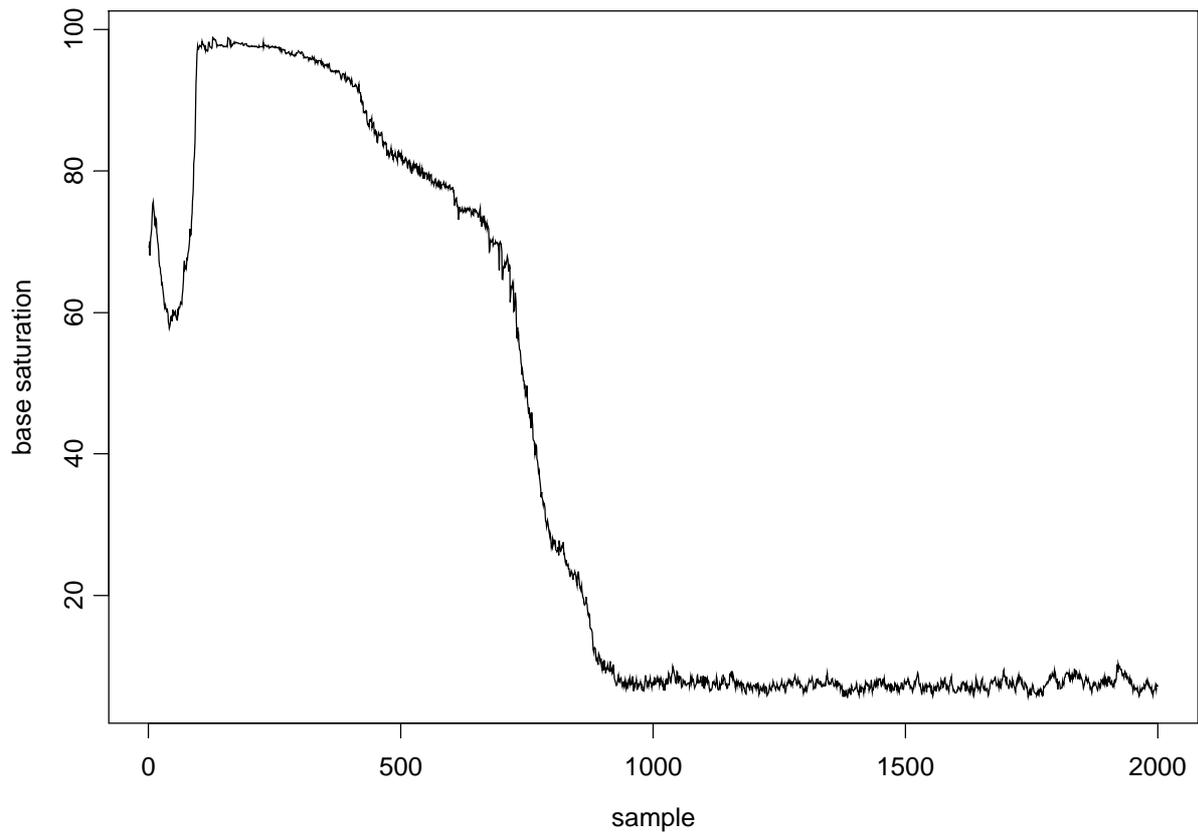
Base saturation, Year 2000, trace

Figure 7: Trace of estimated base saturation in soil
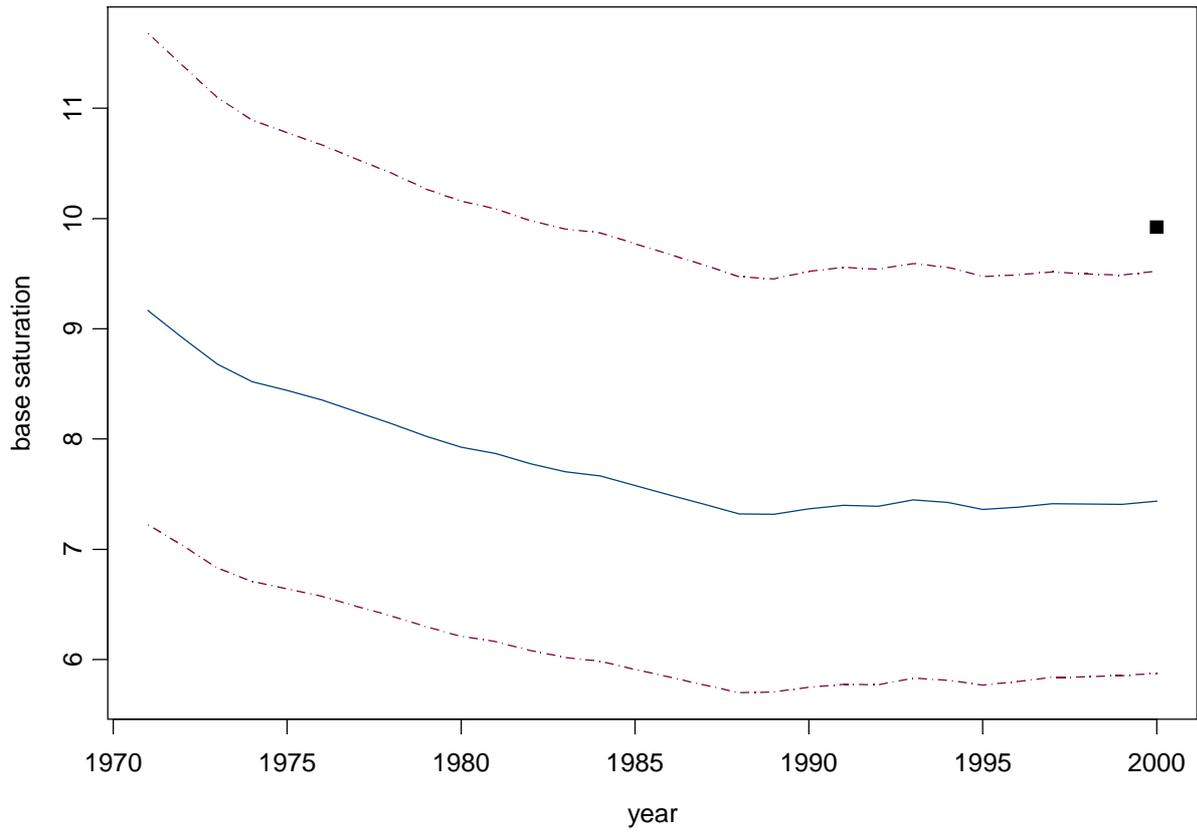
Base saturation, posterior mean

Figure 8: Time series of estimated base saturation in soil. Posterior mean (blue line), 5% and 95% quantiles (red broken lines).
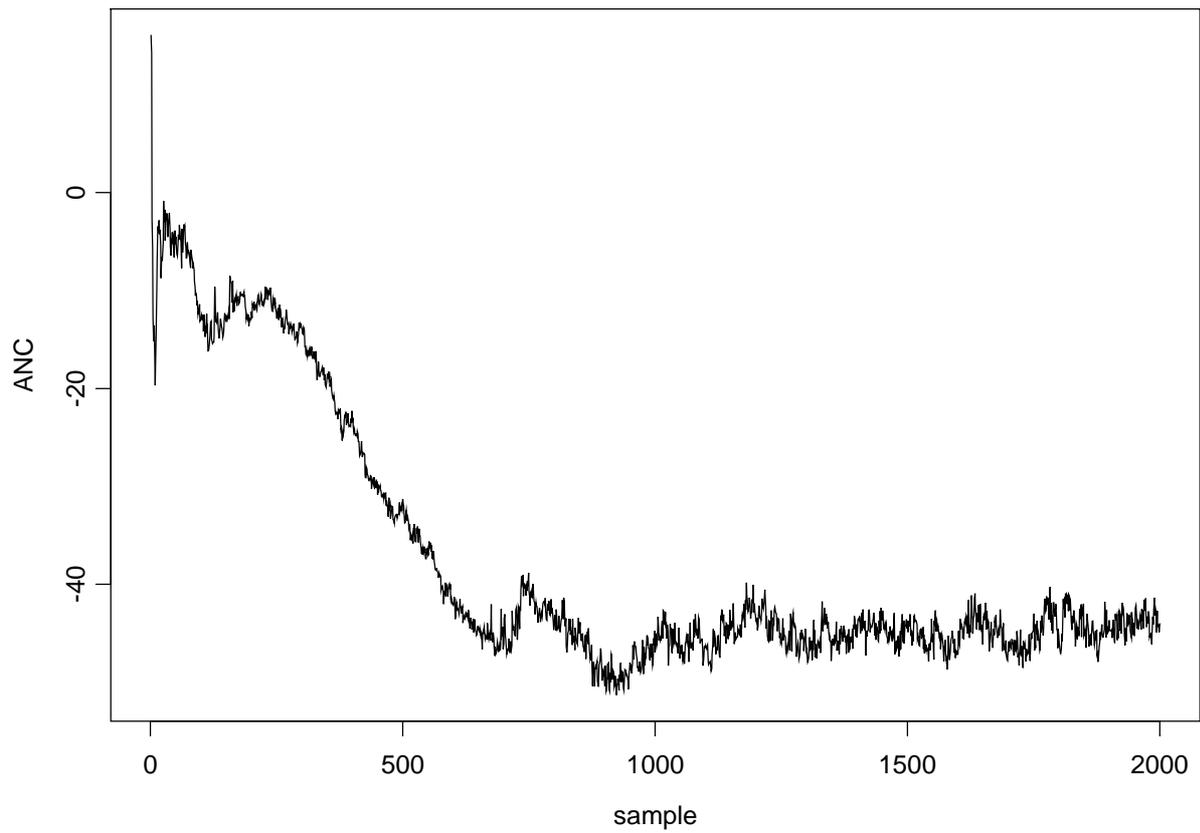
ANC, Year 2000, trace

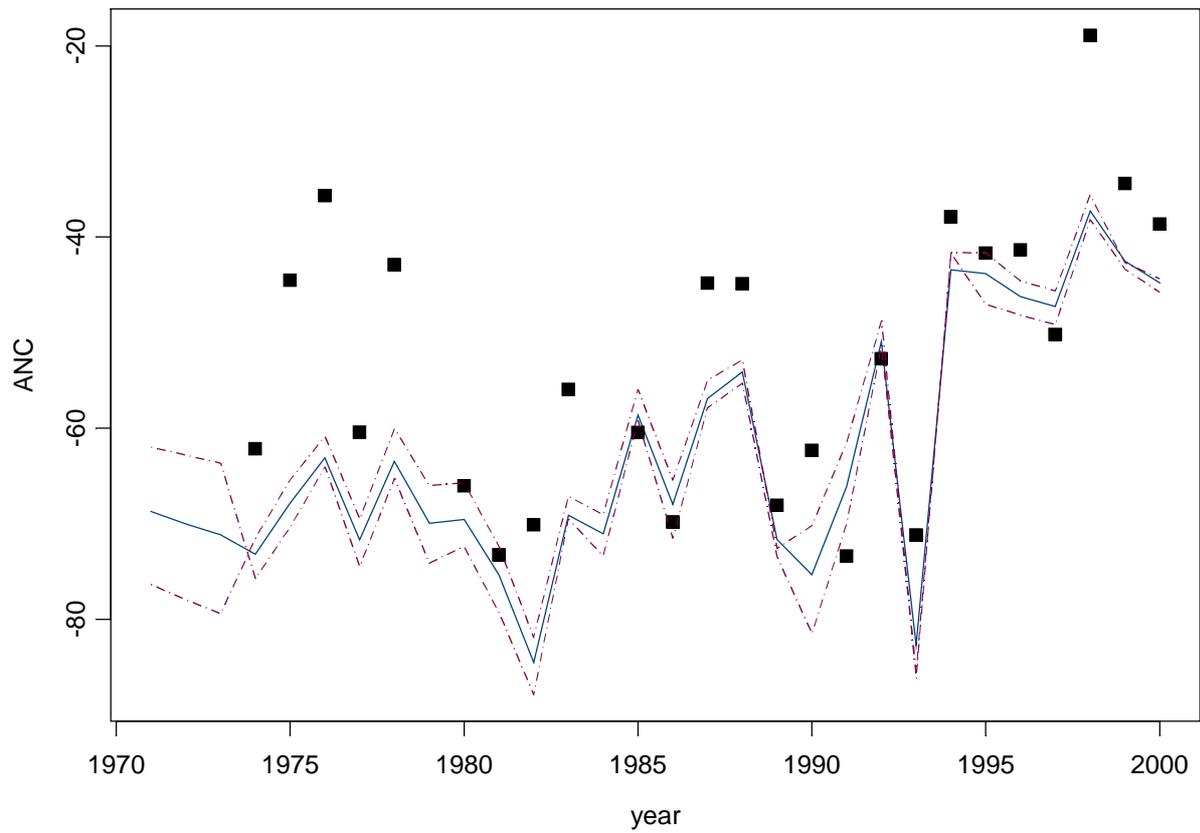**Figure 9: Trace of estimated ANC in 2000.**

Figure 10: Time series of ANC for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.

SO4, Year 2000, trace

Figure 11: Trace of SO$_4$ in 2000.

SO4, posterior mean

Figure 12: Time series of SO$_4$ for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.
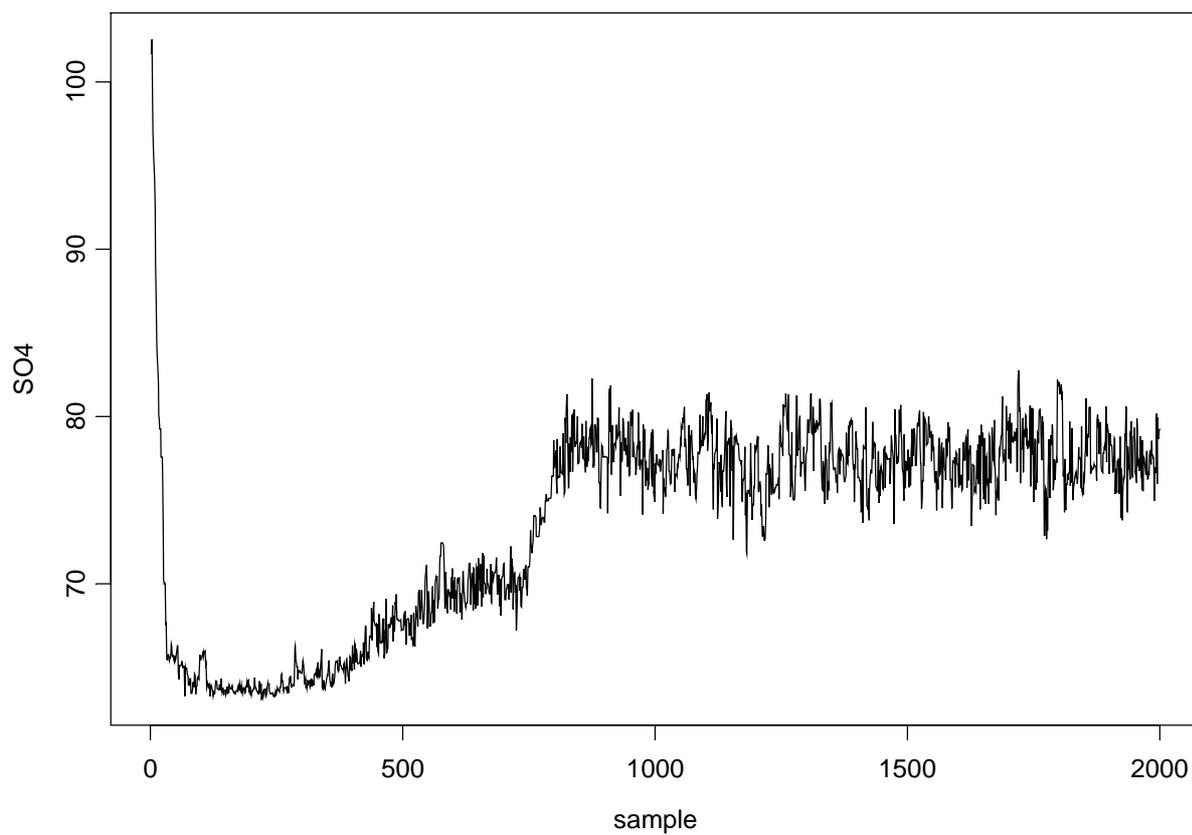
Ca, Year 2000, trace

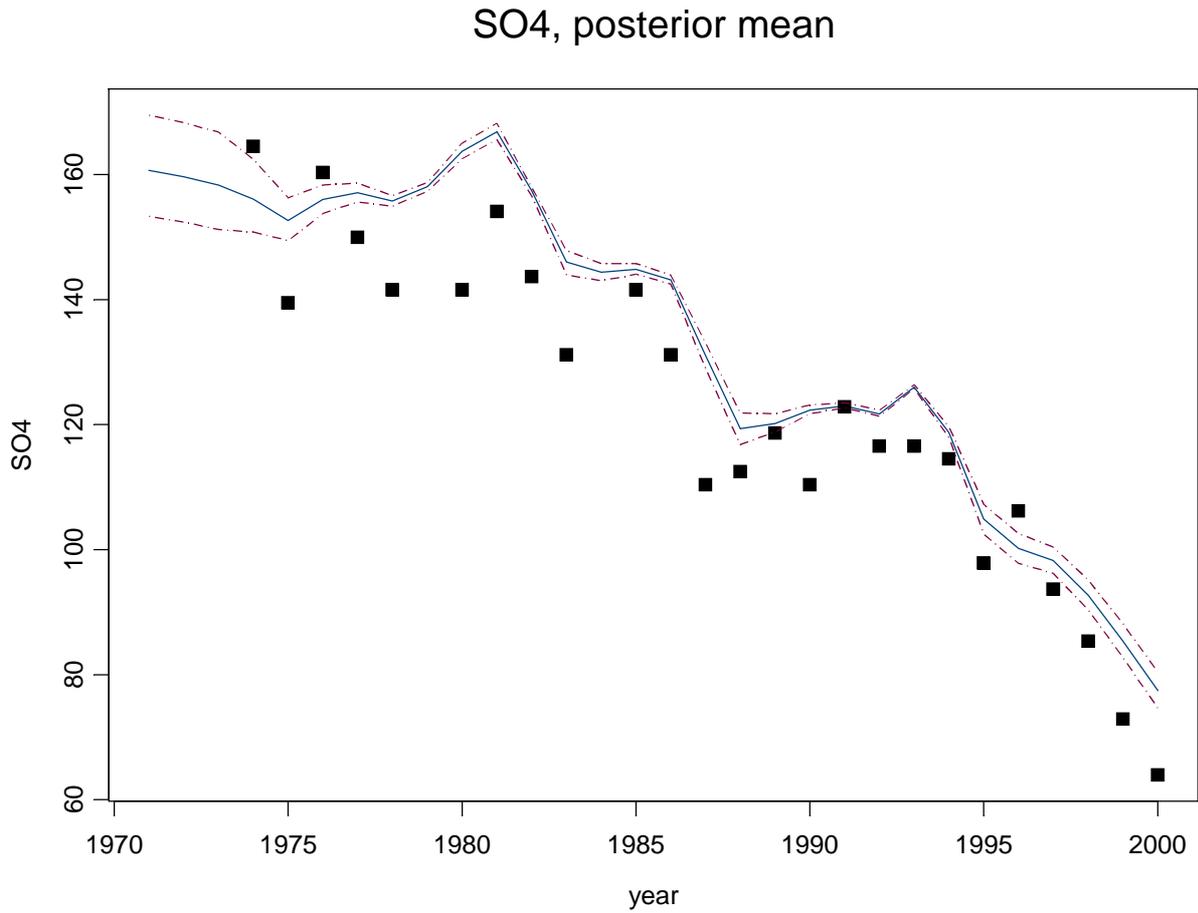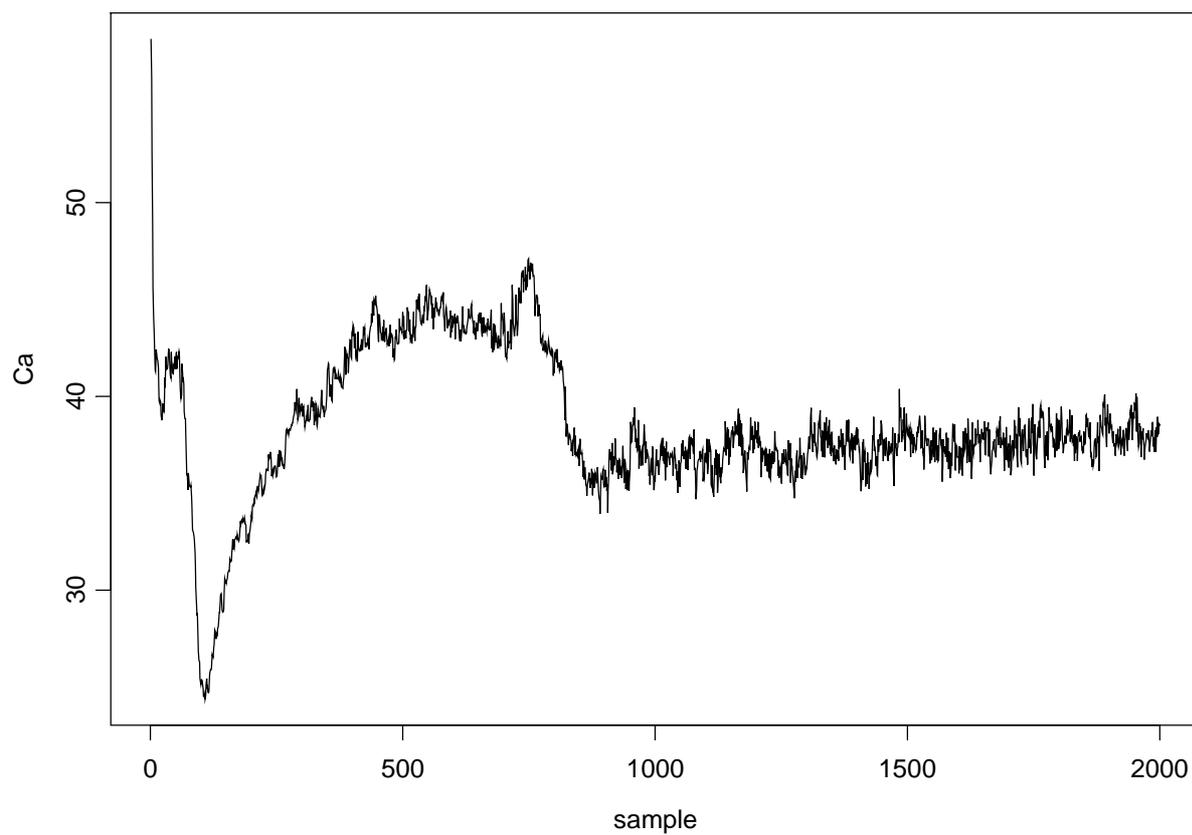**Figure 13: Trace of Ca for 2000.**

Ca, posterior mean

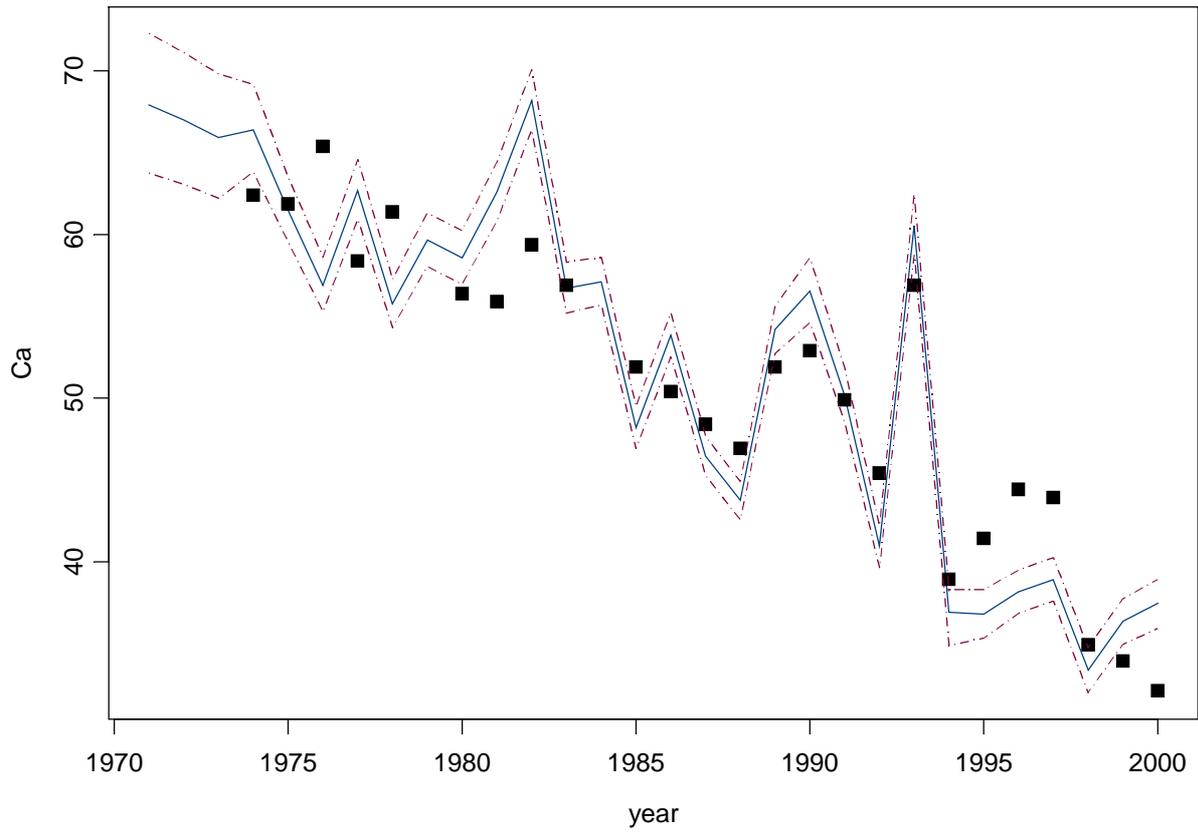**Figure 14: Time series of Calcium for the observation period. The blue line is the posterior mean, the broken lines are the 0.05- and 0.95- quantiles, and the filled squares are the measured values.**
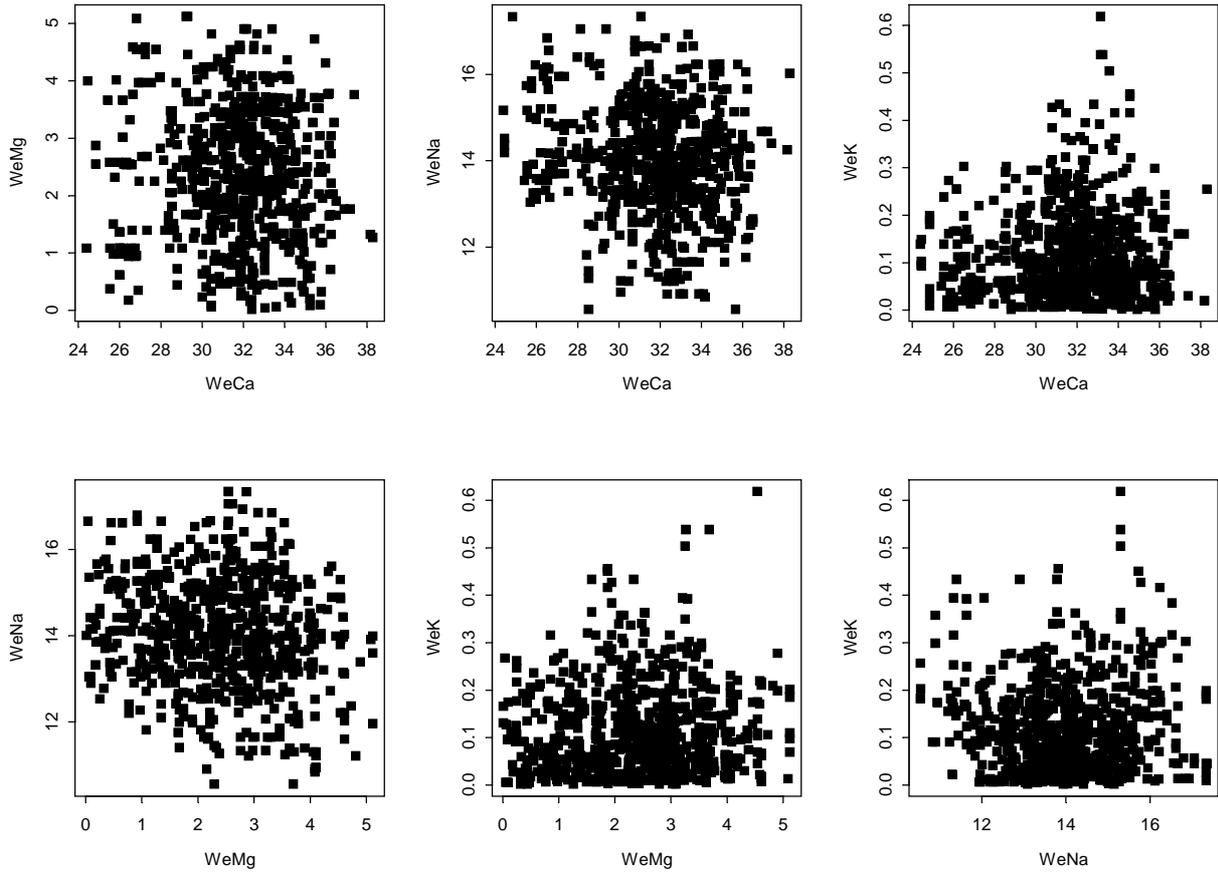
Figure 15: Scatter plots of samples soil weathering parameters from the posterior distribution.

'
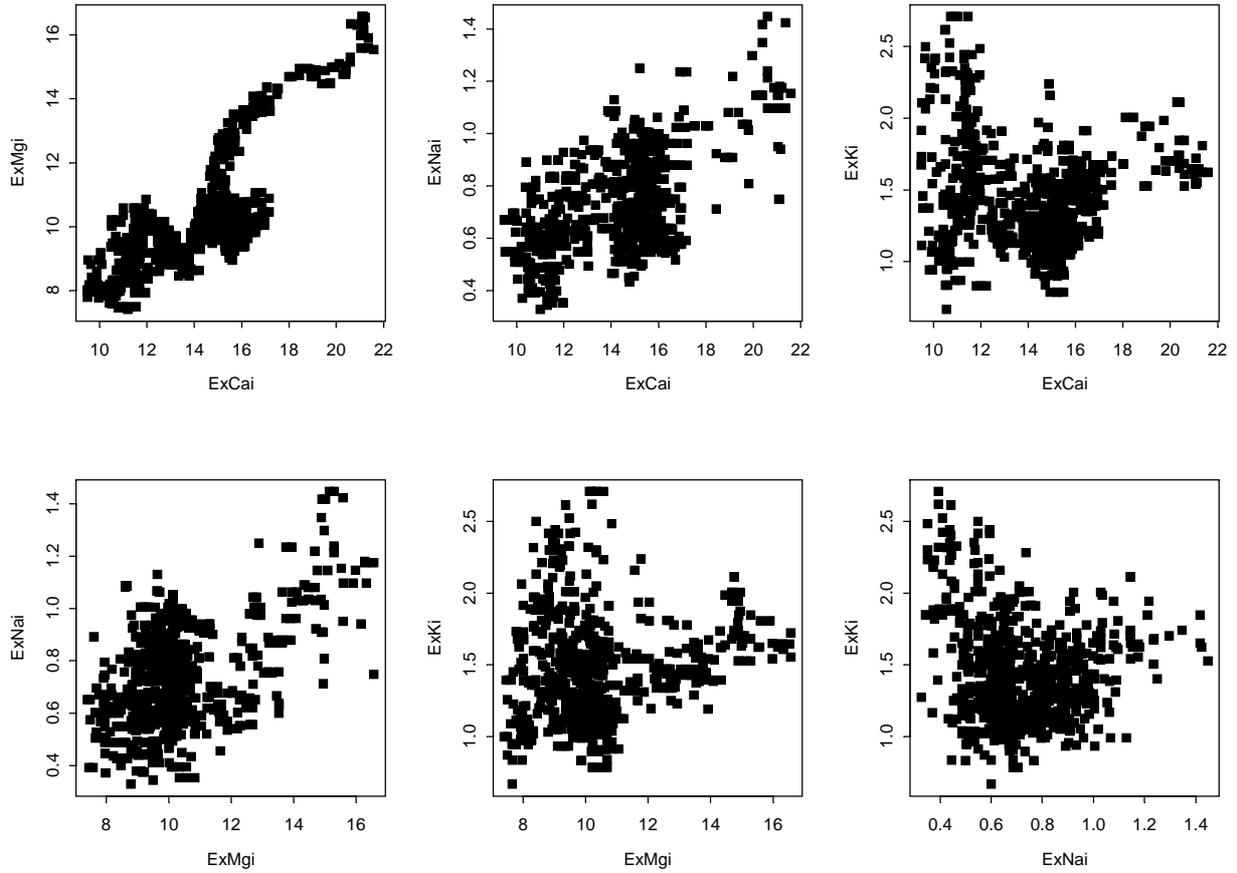
Figure 16: Scatter plots of parameters of initial soil base cation saturation from the posterior distribution.

Figure 17: Scatter plots of parameters of output soil base cation saturation from the posterior distribution.

8. Tables

Table 1: Prior distributions for the input parameters to the MAGIC model. (***** Constrained to ExCai+ExMgi+ExNai+ExKi≤100)

Table 2: MAGIC model output parameters used in the calibration.

Table 3: Observations used in the calibration ([*] The surface water data are time series of observations through the period 1974-2000). The standard deviations are fixed in the calibration.

Input parameters (α)	Parameter name	Unit	Distribution class	Min	Max
Fixed Soil Parameters	Depth	m	Uniform	0,30	0,50
	Porosity	%	Uniform	40,00	60,00
	BulkDens	kg/m3	Uniform	696,00	850,00
	CEC	meq/kg	Uniform	95,00	117,00
	HlfSat	ueq/L	Uniform	1,00	500,00
	Emx	meq/kg	Uniform	0,00	50,00
	KAl	Log10	Uniform	6,00	11,00
	Temp	degC	Fixed	5,00	5,00
	PCO2	%atm	Uniform	0,50	2,00
	DOC	umol/L	Uniform	0,00	250,00
	Nitrif	%	Fixed	100,00	100,00
Soil Uptake	UptNH4	%	Uniform	0,00	100,00
	UptNO3	%	Uniform	0,00	100,00
Soil Weathering	WeCa	meq/m2/yr	Uniform	0,00	100,00
	WeMg	meq/m2/yr	Uniform	0,00	100,00
	WeNa	meq/m2/yr	Uniform	0,00	100,00
	WeK	meq/m2/yr	Uniform	0,00	100,00
Initial soil base cation saturation*	ExCai	%	Uniform	0,1	50
	ExMgi	%	Uniform	0,1	50
	ExNai	%	Uniform	0,1	50
	ExKi	%	Uniform	0,1	50
	Ppt	m/yr	Fixed	2,27	2,27
Wet Concentrations	Ca	ueq/L	Fixed	5,0	5,0
	Mg	ueq/L	Fixed	13,8	13,8
	Na	ueq/L	Fixed	60,66	60,66
	K	ueq/L	Fixed	2,0	2,0
	NH4	ueq/L	Fixed	23,84	23,84
	SO4	ueq/L	Fixed	25,4	25,4
	CL	ueq/L	Fixed	72,68	72,68
	NO3	ueq/L	Fixed	30,32	30,32
Dry Deposition Factors	Ca	DDF	Uniform	1,00	2,00
	Mg	DDF	Uniform	1,00	2,00
	Na	DDF	Uniform	1,00	2,00
	K	DDF	Uniform	1,00	2,00
	NH4	DDF	Uniform	1,00	2,00
	SO4	DDF	Uniform	1,00	2,00
	CL	DDF	Uniform	1,00	2,00
	NO3	DDF	Uniform	1,00	2,00
Lake/Stream Characteristics	RelArea	%	Fixed	0,01	0,01
	RetTime	Year	Fixed	0,00	0,00
	KAl	Log10	Uniform	6,00	11,00
	Temp	degC	Uniform	3,00	10,00
	pCO2	%atm	Uniform	0,05	0,20
	DOC	umol/L	Uniform	0,00	100,00
	Nitrif	%	Fixed	100,00	100,00

Table 1: Prior distributions for the input parameters to the MAGIC model. (* Constrained to ExCai+ExMgi+ExNai+ExKi≤100)

Output parameters (β)	Parameter name	Unit
Soil Output parameters	ExCa	%
	ExMg	%
	ExNa	%
	ExK	%
	SoilpH	
Surface Water Output Parameters	Ca	ueq/L
	Mg	ueq/L
	Na	ueq/L
	K	ueq/L
	NH4	ueq/L
	SO4	ueq/L
	Cl	ueq/L
	NO3	ueq/L
	H+	
	Al	

Table 2: MAGIC model output parameters used in the calibration.

Observations (y)	Name	Unit	Likelihood	Value	Standard deviation
Soil Data	ExCa	%	Gaussian	3,9	0,25
	ExMg	%	Gaussian	2,1	0,25
	ExNa	%	Gaussian	1,1	0,25
	ExK	%	Gaussian	2,8	0,25
	SoilpH		Gaussian	4,2	0,075
Surface Water Data	Ca	ueq/L	Gaussian	*	2,96
	Mg	ueq/L	Gaussian	*	1,84
	Na	ueq/L	Gaussian	*	5,82
	K	ueq/L	Gaussian	*	0,74
	SO4	ueq/L	Gaussian	*	8,27
	Cl	ueq/L	Gaussian	*	9,36
	NO3	ueq/L	Gaussian	*	1,3
	H+		Gaussian	*	1,58
	Al		Gaussian	*	2,58

Table 3: Observations used in the calibration ([*] **The surface water data are time series of observations through the period 1974-2000). The standard deviations are fixed in the calibration.**