

NATIONAL INSTITUTE OF PUBLIC HEALTH AND THE ENVIRONMENT
BILTHOVEN - THE NETHERLANDS

Report nr. 733001004

**Rotated-Random-Scanning: a simple method
for set-valued model calibration**

P.H.M. Janssen, P.S.C. Heuberger

December 1995

This study has been carried out on behalf and for the account of the Directorate General Environment (DGM/LE) in the context of the 3rd phase of the Dutch Priority Programme of Acidification (DPPA), theme II: 'Modelling for Policy Analysis', within the framework of project no. 733001 'Acidification Research'.

National Institute of Public Health and the Environment, RIVM, Antonie van Leeuwenhoeklaan 9,
P.O. Box 1, 3720 BA BILTHOVEN, The Netherlands, Tel: ..(31)30-2749111; Telefax:...(31)30-2742971

Mailing list

1. Stuurgroep APV-III (Additioneel Programma Verzuring)
2. Directeur Generaal Milieubeheer
3. Plaatsvervangend Directeur Generaal Milieubeheer
4. Drs. H. Marseille (DGM/LE)
5. Drs. W.S. Smeulders (DGM/LE)
6. Dr. K. Bristow (SWASRU; Australia)
7. Prof. B.J. Cosby (University of Virginia)
8. Dr. B. Gardner (Oak Ridge National Laboratory)
9. Dr. B. Gottesbüren (BASF-AG)
10. Prof. P-E. Jansson (University of Uppsala)
11. Dr. M. Johansson (National Board of Waters and the Environment, Helsinki)
12. Ir. J. van Minnen (IIASA)
13. Dr. H. Wheeler (Imperial College)
14. Ir. T.N. Olsthoorn (Gem. Waterleidingen, Amsterdam)
15. Dr. ir. H. Erbrink (KEMA; Arnhem)
16. Dr. G.W. Heil (Resource Analysis; Delft)
17. Dr. J.J.M. Berdowski (TNO; Delft)
18. Drs. C.I. Bak-Eijsberg (Rijkswaterstaat, DBW/RIZA; Lelystad)
19. Ir. H.J. Vinkers (KIWA NV; Nieuwegein)
20. Ir. A. Seebregts (ECN; Petten)
21. Dr. J. Th. Lubbe (iMGZ, Rotterdam)
22. Dr. ir. G.M.J. Mohren (De Dorschkamp; Wageningen)
23. Drs. J.R. van Veen (De Dorschkamp; Wageningen)
24. Ir. A.A.M. Jansen (GLW, Wageningen)
25. Dr. M.J.W. Jansen (GLW, Wageningen)
26. Dr. H. van der Voet (GLW, Wageningen)
27. Ir. G.J. Reinds (Winand Staring Centrum; Wageningen)
28. Ir. J.E. Groenenberg (Winand Staring Centrum; Wageningen)
29. Ir. J. Kros (Winand Staring Centrum; Wageningen)
30. Drs. C. van der Salm (Winand Staring Centrum; Wageningen)
31. Dr. ir. W. de Vries (Winand Staring Centrum; Wageningen)
32. Dr. P. Herman (DIHO, Yerseke)
33. Drs. H. Scholten (DIHO, Yerseke)
34. Dr. Ir. W. Bouten (Univ. van Amsterdam)
35. Prof. dr. ir. C. van der Akker (TU Delft)
36. Drs. M.S. de Wit (TU Delft)
37. Drs. R. Boerema (Prof. H.C. van Hall Instituut; Groningen)
38. Prof. Dr. F. Berendse (LU Wageningen)
39. Dr. ir. H.W.G. Booltink (LU Wageningen)
40. Dr. D. van Dam (LU Wageningen)
41. Prof. dr. L. Hordijk (LU Wageningen)
42. Dr. ir. K. Keesman (LU Wageningen)
43. Prof. dr. ir. G. van Straten (LU Wageningen)
44. Dr. S.E.A.T.M. van der Zee (LU Wageningen)
45. Depôt Nederlandse Publicaties en Nederlandse Bibliografie
46. Directie RIVM
47. Ir. F. Langeweg
48. Dr. J.M. Alcamo
49. Drs. T. Aldenberg
50. Drs. A.H. Bakema
51. Ir. R. v.d. Berg
52. Ing. G.P. Beugelink
53. Drs. A.H.W. Beusen
54. Dr. L.C. Braat

55. Drs. H.C. Eerens
56. Drs. J.C.H. van Eijkeren
57. Dr. M.G.J. den Elzen
58. Dr. ing. J.W. Erisman
59. Ir. B. Fraters
60. Drs. A. van der Giessen
61. Dr. ir. J.J.M. van Grinsven
62. Ir. B.J. de Haan
63. Drs. C. van Heerden
64. Dr. J.-P. Hettelingh
65. Dr. S.H. Heisterkamp
66. Ir. G.J. Heij
67. Dr. ir. J.A. Hoekstra
68. Ir. R.T. Hoogenveen
69. Drs. J.H. Janse
70. Drs. M.A. Janssen
71. Dr. O. Klepper
72. Drs. P.R.G. Kramer
73. Dr. M.A.J. Kuijpers-Linde
74. Ir. G.M.H. Laheij
75. Prof. Dr. Ir. A. Leijnse
76. Ir. A.M.A. van der Linden
77. Drs. R. Meijers
78. Ir. J.H.C. Mülschlegel
79. Dr. M. Posch
80. Prof. Dr. Ir. J. Rotmans
81. Drs. R. Sanders
82. Ir. F.J. Sauter
83. Dr. W. Slob
84. Dr. Ir. F. Swartjes
85. Drs. A. Tiktak
86. Drs. Th. P. Traas
87. Dr. ir. G.J.M. Uffink
88. Dr. E.J.M. Veling
89. Dr. ir. L.G. Wesselink
90. Drs. W.J. Willems
91. Drs. F.G. Wortelboer
- 92-93. Authors
94. Hoofd Bureau Voorlichting en Public Relations
- 95-96. Bibliotheek RIVM
97. Bureau Rapportenregistratie
- 98-120. Reserve exemplaren t.b.v. Bureau Rapportenbeheer

Table of contents

Mailing list	ii
Table of contents	iv
List of tables	v
List of figures	v
Notation	vi
Preface	vii
Summary	viii
Samenvatting	viii
1 Introduction	1
2 Problem statement	3
3 Features of the Rotated-Random-Scan method	4
4 A practical calibration study	14
5 Discussion	27
References	29
Appendix A: Outline of the Rotated-Random-Scan procedure	32
Appendix B: Theoretical considerations on the influence of working space, scaling and decomposition	37

List of tables

1	Recommended choices for the calibration options	8
2	Parameter ranges, nominal values and associated relevant outputs	16
3	Info on criterion values for various iterations in the first calibration experiment	20
4	Info on criterion values for various iterations in the second calibration experiment	20
5	Statistics for selected accepted parameter values	21
6	Statistics for the misfit-contributions C_i of the various model outputs	21

List of figures

1	An outline of the set-theoretic calibration problem	3
2	Transformation in the parameter space based on the subset of currently available acceptable parameters	5
3	Flow scheme of the iterative rotated-random-scan procedure	7
4	Annual flux-weighted concentration measurements over the period 1973-1989 at the spruce site in Solling, Germany	15
5	Comparison of measurements and results from the nominal model	17
6	Comparison of measurements and results from the nominal model and the optimal model	22
7	Scatterplots showing the selected acceptable parameters after calibration	23
8	Histograms for the selected acceptable parameters after calibration	24
9	Scatterplots showing the trade-offs between the misfits for the various modeloutputs	25
10	Uncertainty band around modeloutputs after calibration	26
11	The set of currently available acceptable samples in the \hat{x}_1 - and the \hat{x}_2 -space, as well as the set of new (uniform, uncorrelated) candidate samples in these spaces	40
12	The sets of new candidate samples $\hat{x}_{1,can}$ and $W^T \cdot \hat{x}_{2,can}$ in the \hat{x}_1 -space, before back transformation to the original space	41

Notation

Symbols

μ : mean

σ : standard deviation

A^T : transposed of matrix A

A^{-1} : inverse of matrix A

$A_{i,j}$: (i, j) -entry (element) of matrix A

$A_{i,*}$: i -th row of matrix A

$A_{*,j}$: j -th column of matrix A

\mathbb{E} : Expectation; mean

ICENT: Centralization option

ISCAL: Scaling option

N_{acc} : Number of currently available acceptable samples

N_{ini} : Number of candidate samples in the initialisation step

N_{lim} : Minimal number of acceptable samples to be obtained

N_{can} : Number of new candidate samples per iteration

Acronyms

CWM Centre for Mathematical Methods (Centrum voor **W**iskundige **M**ethoden) (**RIVM**)

DAS Dutch **A**cidification **S**ystem

DDA Dominant **D**irection **A**nalysis

DOMDIR Computer program for **DOM**inant **DIR**ection analysis

DPPA3 Dutch **P**riority **P**rogramme on **A**cidification; part 3

GENSEN Computer program for **GEN**eralized **SEN**sitivity analysis

GSA Generalized **S**ensitivity **A**nalysis

LHS Latin **H**ypercube **S**ampling

RIVM National Institute of Public Health and Environmental Protection (**R**ijks **I**nstituut voor **V**olksgezondheid en **M**ilieu)

RORASC Computer program for **RO**tated **R**andom **SC**anning

SMART Simulation **M**odel for **A**cidification's **R**egional **T**rends

UNCSAM **UNC**ertainty and **S**ensitivity **A**nalysis by **MON**te Carlo sampling

Acknowledgement

We thank Hans Kros (Staring Centrum; Wageningen) for providing the model SMART and the data, and for giving advice and comments on the reported calibration study,

Preface

In the context of the model-oriented research activities for the Dutch Priority Programme on Acidification (DPPA3; theme II), a project was initialized on the *‘Development and application of methods for calibration of DAS models, both on a regional scale and on the site level’* (P.H.M. Janssen and P.S.C. Heuberger, CWM; RIVM).

One of the objectives of the project was to develop general tools for calibration of models in situations where the available prior information on the unknown parameters is limited, and where measurements of the modelled quantities are incomplete, scarce or inaccurate, which is often the case in environmental applications. Moreover these tools should be applicable in more favorable situations, where e.g. more measurements are available.

Against this background, a simple method for set-valued calibration was proposed as a general tool. This method, the rotated-random-scanning method, uses an efficient Monte Carlo based scan of the parameter space, rotating it in an iterative fashion on basis of the characteristics (covariance structure) of the set of parameters obtained thusfar. In this way it gradually and efficiently zooms in on the part of the parameter space which gives good agreement between model outcomes and measurements (good fit).

Since computational complexity increases very rapidly with the dimensionality of the parameter space to be scanned, the method will benefit greatly from a prior reduction of the number of parameters to be calibrated. For this purpose related tools have been proposed which establish the sensitivity of the parameters. Parameters which appear to be insensitive can typically be excluded from calibration.

In this report the proposed methods for calibration and sensitivity analysis are presented, and results of their use on a soil-acidification model (SMART) are reported. The developed tools have been designed in a model independent way. They have been implemented in a portable code for general use (standard FORTRAN77 embedded in an ANSI-C environment to enable dynamic memory allocation), and are presented in four companion background documents.

Summary

A method is proposed for calibrating models in ill-defined and information-poor situations where data are sparse and uncertain. This method, the so-called rotated-random-scan method, performs an iterative Monte Carlo search of the parameter-space. By using the characteristics of the parameter set obtained thusfar, it applies rotations in combination with random scans in order to gradually zoom in on the subset of acceptable model parameter values which render model results that match satisfactorily with the measurement data. The efficiency of the model largely depends on the applied uniform random scan, which may ask for the future incorporation of a more sophisticated set-theoretic approach in case of non-connected or irregularly shaped parameter sets.

The proposed technique has been implemented in FORTRAN77, and related tools have been developed for sensitivity analysis to significantly reduce the number of parameters to be calibrated. A real-life calibration study of a soil-acidification model (SMART) illustrates the use and the possible limitations of the rotated-random-scan method. The performance of the method in obtaining parameter uncertainty reduction and compliance of model outputs with measurement data may also provide useful indicative information on the merits of model formulations.

Samenvatting

Een methode wordt voorgesteld voor het calibreren van modellen in slecht-gedefinieerde en informatie-arme situaties waarbij de meetgegevens gebrekkig en onzeker zijn. Deze zogenaamde 'rotated-random-scan' methode voert een iteratieve, Monte Carlo gebaseerde, zoekactie uit in de parameter ruimte. Gebruik makend van de karakteristieke eigenschappen van de tot dan toe bepaalde parameterwaarden, worden in de parameter ruimte rotaties uitgevoerd in combinatie met random zoek-acties. Op deze wijze wordt geleidelijk een deelverzameling van acceptabele model parameterwaarden opgespoord, waarvoor de bijbehorende modeluitkomsten in voldoende mate overeenkomen met de meetgegevens. De efficiëncy van de methode hangt voor een groot deel af van de toegepaste uniforme random scan, hetgeen de toepassing van een meer geavanceerde verzameling-theoretische aanpak vereist voor niet-samenhangende of grillig gevormde parameter verzamelingen.

De voorgestelde techniek is in FORTRAN77 geïmplementeerd. Bovendien is programmatuur ontwikkeld voor het uitvoeren van gevoeligheidsanalyses waarmee het aantal te calibreren parameters aanzienlijk gereduceerd kan worden. Het gebruik van de 'rotated-random-scan' methode wordt uiteindelijk geïllustreerd aan de hand van een calibratie van een bodemverzuuringsmodel (SMART). De resultaten van deze calibratie studie verduidelijken de mogelijke beperkingen van de gehanteerde methode, en verschaffen ook nuttige informatie over de waarde van de gehanteerde model formuleringen.

1 Introduction

In modelling ecosystems one is often confronted with the problems of sparse, uncertain or unreliable data and the lack of knowledge on the relevant processes. These problems have direct consequences for building and calibrating models. In particular, it will generally be impossible to model the considered system exactly. The best we can hope for is to model the important features of the system *approximately*, trying moreover to make statements on the *validity* of the obtained model and on the associated *uncertainty* in the model results (e.g. predictions). Pursuing this should involve a judicious trade-off between model-complexity and data-availability, taking in mind the intended objective of the modelling exercise (see Janssen and Heuberger [1995]).

Due to the above mentioned sparsity and unreliability of the data and due to the inevitable presence of *non-random* model (structure) errors, a purely *statistical* approach which is typically based on assumptions on the *random* nature of uncertainty¹, is not always adequate and may give misleading results when calibrating the model. Attempts to discriminate explicitly between (non-random) model errors and (random) measurement errors are a possible alternative approach to tackle these problems. However this will typically require a considerable amount of data or prior information since a detailed description of the model error (bounds) is needed.

Another approach, which is less confined by the amount of available data, has been studied intensively in literature over the last decade. This approach treats the uncertainty/error in a *set theoretic* way (see e.g. the review paper of Milanese, Vicino [1991b]; see also Milanese, Vicino [1991a]; Belforte et al. [1990a,b], Walter [1990]), describing it by means of an additive noise which is known only to have given bounds (*Unknown-But-Bounded* (UBB) uncertainty models). In practice this description is often more realistic and less demanding than a statistical or a combined deterministic/statistical description. Having selected an appropriate model structure, the calibration of the parameters in this context often amounts to determining the *set*² of parameter vectors (the so-called *acceptable parameter set*) leading to model behaviour which is consistent with the data within their associated noise-bounds. The obtained set of acceptable parameter vectors can subsequently be used to assess the ‘uncertainty’ of the model results or predictions in a set-theoretic sense.

Apart from being an adequate approach for calibration in *complex* or *ill-defined, highly uncertain* situations, this set-theoretic approach can also serve as a useful *initial step* to delimit the prior parameter uncertainty range in *conventional* calibration studies where one typically tries to find a parameter vector which renders an optimal fit between model and data.

This set-valued calibration requires the determination of the acceptable parameter set, which can be rather *complex*; see Walter and Pronzato [1994]. Instead of looking for an exact description one therefore often looks for an *approximate* description, using simply shaped sets (e.g. boxes, or ellipsoids) containing or contained in the set of interest. Alternative approaches use e.g. *Monte Carlo sampling ideas* (cf. Keesman and van Straten [1988],

¹E.g. choosing stochastic uncertainty models for the errors, as in maximum likelihood or Bayesian techniques.

²The common calibration techniques often try to find a *single* ‘optimal’ parameter vector, minimizing a certain fit-criterion. In situations where there is considerable variability/uncertainty in the data, the system or the models, this ‘*optimal fit*’ approach is not fully adequate. It will be more appropriate to look for models which have an *acceptable fit* (not necessarily minimal) to the data. This can e.g. be achieved by looking for the *set* of parameter vectors or the set of models for which the associated model-to-data misfit remains under an acceptable level (*maximally tolerated misfit*). This approach automatically results in calibration in a set-theoretic context.

Keesman [1989a,1990]), or global optimisation methods based on (controlled) random search (cf. Walter and Piet-Lahanier [1988], Klepper and Rouse [1991]) to obtain an (approximate) description of this set. A promising alternative approach which applies an efficient way of uniformly covering subsets in the parameter space has recently been proposed by Klepper and Hendrix [1994a,b] in the context of level set determination for accuracy assessment.

The above mentioned Monte Carlo based technique has recently been used in calibration and prediction studies for various ecosystems (Keesman and van Straten [1987,1988,1989,1990], Keesman [1989a,1990]) and is considered in this report in a slightly generalized version. The key idea behind the method is to iteratively apply rotations of the parameter space in combination with a random scan (Monte Carlo sampling) of the rotated parameter space, in order to discover linear interactions (correlations) between parameters and to gradually zoom in on the set of acceptable parameter vectors in an efficient way.

The main features of this so-called *rotated-random-scan method* are discussed in section 3, after having stated the calibration problem more formally in section 2. In section 4 the method is illustrated by results from a real life study, concerning the calibration of the acidification model SMART (cf. Posch et al. [1993]) on a data set of the Solling Spruce Site. This data set has been extensively used to compare the performance of various international acidification models during the international 'Workshop on Comparison for Forest-Soil-Atmosphere Models', May 10-14, 1993, Leusden, the Netherlands (van Grinsven, Driscoll, Tiktak [1995]). Section 5 sums up final remarks concerning the performance of the method. The appendices contain detailed technical information on the presented method and its properties.

The *software* which has been developed for the rotated-random-scan method is described in Janssen [1995a] and is available on request (PC-version). It relies heavily on the recently developed software package UNCSAM which is commercially available (Janssen, Heuberger, Sanders [1992,1994]). Tests with the rotated-random-scan method have been performed and are reported in Janssen and Sanders [1995]. Moreover, related software has been developed to perform *sensitivity analyses*³ in a set theoretic calibration context (see Janssen [1995b,1995c]). This software is also available on request.

³For a successful calibration it will often be necessary to reduce the number of parameters to be calibrated beforehand. A preliminary sensitivity analysis can be helpful in accomplishing this goal.

2 Problem statement

Similar to Klepper [1988], a model is primarily considered as a mapping, for specific input functions (forcing functions), from a parameter space \mathbb{R}^p into the space of associated model outputs⁴. Moreover, it is assumed that prior information on the parameters is available which restricts the parameters to a certain subset of \mathbb{R}^p : the set of *allowable* parameter values (i.e. $\Theta \subset \mathbb{R}^p$). Set-theoretic calibration is now aimed at restricting this parameter set further, by confronting the associated model outputs with the available data/observations on the system, and trying to find parameters for which the model outputs and the data match decently. In this confrontation one should bear in mind that the observations on the system will be uncertain or imprecise; moreover one should take the intended purpose of the model into account. E.g. if the measurements have an accuracy of 10 %, the model results are considered acceptable if their difference with the data is on the average at least this amount; likewise, if the model is e.g. only intended for indicating general trends for certain long-term scenarios, it often suffices to determine models which agree with the data within a certain order of magnitude.

In the set-theoretic context this confrontation between observations and model outputs is typically accomplished by first specifying an *appropriate data range*. This specification can be achieved in various ways (cf. Fedra et al. [1981], Fedra [1983], Whitehead and Hornberger [1984], Klepper and Rouse [1991], Keesman [1990]), and should reflect the scatter in the field data and the desired accuracy for the intended model application at hand⁵. This data range is subsequently compared (and hopefully overlaps) with (part of) the model output range associated to the set of allowable parameter values Θ . This is depicted in figure 1, which indicates that *the final intention* is to identify the subset of Θ which renders model outputs falling in the data range, i.e. *determining the subset Θ_{acc} of acceptable parameter values* (in the sequel denoted briefly as acceptable parameters). See Klepper [1988].

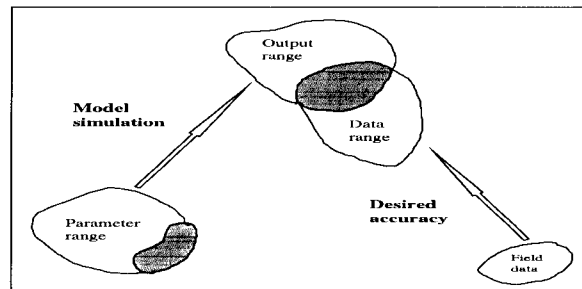


Figure 1: An outline of the set-theoretic calibration problem. The shaded area in the parameter range corresponds to the acceptable parameter values.

⁴ p denotes the number of considered parameters.

⁵For ill-defined and information poor systems one often can not do more than giving a rough classification of what is considered as acceptable behaviour, consistent with the available observations.

3 Features of the Rotated-Random-Scan method

In this section the main properties are described of the so-called *rotated-random-scan* method, a simple method for set-theoretic calibration on basis of rotations in the parameter space in combination with random scanning (see Keesman [1989a,1990], Keesman and van Straten [1988,1989]).

[I] Key Idea:

The aim of the method is to efficiently collect a desired (prespecified) number (N_{lim}) of representative parameters in the acceptable parameter set Θ_{acc} , by updating, in an *iterative/adaptive* fashion, the subset of currently available acceptable parameters on basis of a suitable *rotation or transformation* of the parameter space followed by a uniform *random scan* in this rotated or transformed parameter space.

The rotation or transformation of the parameter space is based on the subset of currently available acceptable parameters, and is intended to *zoom in* on this subset (see figure 2). In this way one hopes to improve the efficiency of the subsequent uniform random scanning step, i.e. one hopes that this ‘rotated’ random sampling yields the desired number of acceptable parameters more quickly. Incorporation of this *learning* aspect clearly improves the random-scan of the original parameter space, which appears to be rather inefficient (Fedra et al. [1981]).

[II] Application Area:

The method is useful for a *crude form of set-valued calibration*, typically to be used in settings where the *information content is poor*, or in the *initial stages of conventional calibration* studies in order to delimit the wide uncertainty range of the unknown parameters (e.g. in the context of regional calibration using binfilling and weighted frequency matching methods; cf. Heuberger et al. [1992]).

Based on a crude dichotomy of the model outcomes in results which are judged as *acceptable* (i.e. behaviour), or as *inacceptable* (i.e. non-behaviour), the procedure aims to find the parameters which lead to acceptable model results (i.e. the *acceptable parameter set*). See figure 1.

In this way more accurate information on the parameters is obtained (i.e. reduction of the parameter uncertainty, e.g. tighter lower- and upper bounds), which consequently leads to more accurate model predictions.

[III] Requirements:

Although in essence no prior limitations are put on the class of models to be calibrated (models can be e.g. dynamic, steady-state, static etc.), application of this procedure requires that:

1. Each parameter vector under consideration can be judged as being *acceptable* or *inacceptable*. Typically, this amounts to evaluating the model for the specific parameter vector, and to deciding whether the associated model results are ‘(in)acceptable’ when compared with the data and/or the a priori knowledge. This comparison can e.g. be based on
 - evaluating whether a certain misfit measure remains under a predefined acceptable tolerance level (Whitehead and Hornberger [1984], Klepper and Rouse [1991], Klepper and Hendrix [1994a,b]);

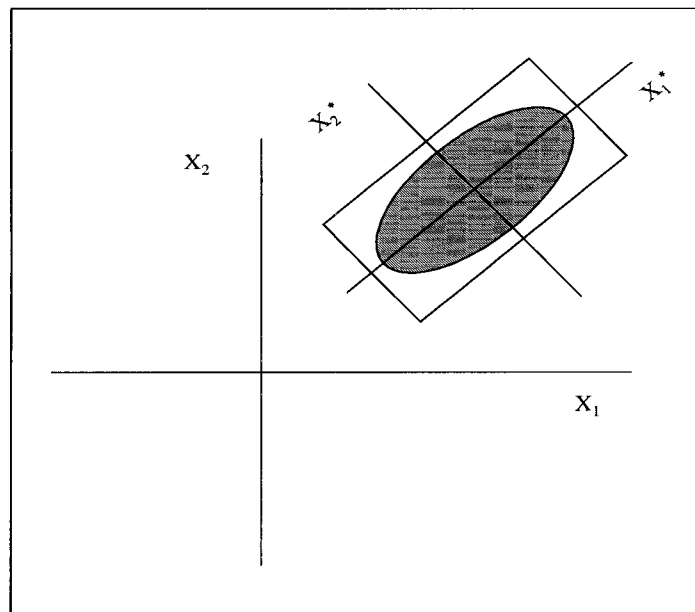


Figure 2: Transformation in the parameter space based on the subset of currently available acceptable parameters (indicated by shaded area).

- evaluating whether the model outcomes are compatible with a predefined error bound on the data or with the prior information (Keesman [1989a; 1990]; Fedra et al. [1981]).

Certainly other alternatives are possible. The specific form according to which this comparison between data and model takes place in practice, should ideally reflect the aspects which are deemed relevant for the study at hand, and the (amount of scatter in the) available data.

2. Prior information on the parameters to be calibrated should be *simple*: lower- and upper bounds on each parameter are required. More detailed information (e.g. correlations between parameters; functional inequalities, e.g. $\theta_1 + \theta_2 \leq 2$) can not be taken *explicitly* into account by the procedure, unless it can be transformed into the specification of lower- and upper bounds by appropriate re-parametrization. Hence, if more intricate constraints are present, the newly generated candidate samples should always be checked afterwards on compliance with the more intricate constraints.

Notice that all acceptable parameters are considered as *equally likely*. In this respect the set-theoretic approach differs from the statistical approach where a probability is associated to the acceptable parameters to denote the disparity in their occurrence.

[IV] Procedure:

The procedure is a slightly generalized version of the method presented in Keesman [1989a,1990], Keesman and van Straten [1988,1989] (more choices are offered for scaling, decomposition, sampling technique; see item [V] later on). It employs an *iterative* search of the parameter space, which is preceded by an *initialisation step* in which an initial subset of sampled parameters is determined, containing N_{ini} parameter combinations which are subsequently judged on their (in)acceptability. Each of the *subsequent iterations* typically consists of two parts (see figure 3):

- **Part A: Generating new candidate samples:**⁶ First all *currently* available *acceptable* parameters are collected. Next it is checked whether the number of currently available acceptable samples already exceeds the (prespecified) number N_{lim} of acceptable parameters which have to be obtained. If this is not the case, then N_{can} ⁷ new candidate samples are generated. This is achieved by:
 1. *Transforming or rotating* the original parameter space to focus on the subset of currently available acceptable parameters. This transformation is based on a decomposition of the covariance matrix of the current acceptable parameters, possibly after scaling (optional) and is updated in each iteration.
 2. Performing a *random scan* on basis of uniform random sampling in this transformed space. Next the randomly sampled ‘transformed’ parameters are *back-transformed* to the original space, and serve as new candidate samples. Due to these actions, the obtained samples can fall outside the range (lower and upper bounds) which was originally specified; therefore they have to be checked explicitly on their compliance with these bounds.
- **Part B: Simulation and acceptance:**⁸ Model simulations are performed for the newly generated candidate samples obtained in part A, and it is decided whether the results are acceptable or unacceptable. Subsequently one returns to the beginning of the next iteration (i.e. starting again with part A).

A detailed outline of the rotated-random-scan procedure is presented in appendix A. The specific progress of the procedure depends on a number of choices and options which have to be specified by the user in the initialisation stage and in part A of a typical iteration (see next item).

[V] Users choices:

In addition to the fundamental choices:

- Which parameters have to be calibrated⁹?
- What are their initial lower- and upper bounds?
- When is a parameter vector (in)acceptable?

the user has to make the following choices when applying the procedure sketched in appendix A:

⁶I.e. step 1-5 of the procedure outlined in appendix A.

⁷See the sequel for information concerning the choice of N_{lim} and N_{can} .

⁸I.e. step 6 of the procedure outlined in appendix A.

⁹See the remark at the end of this section.

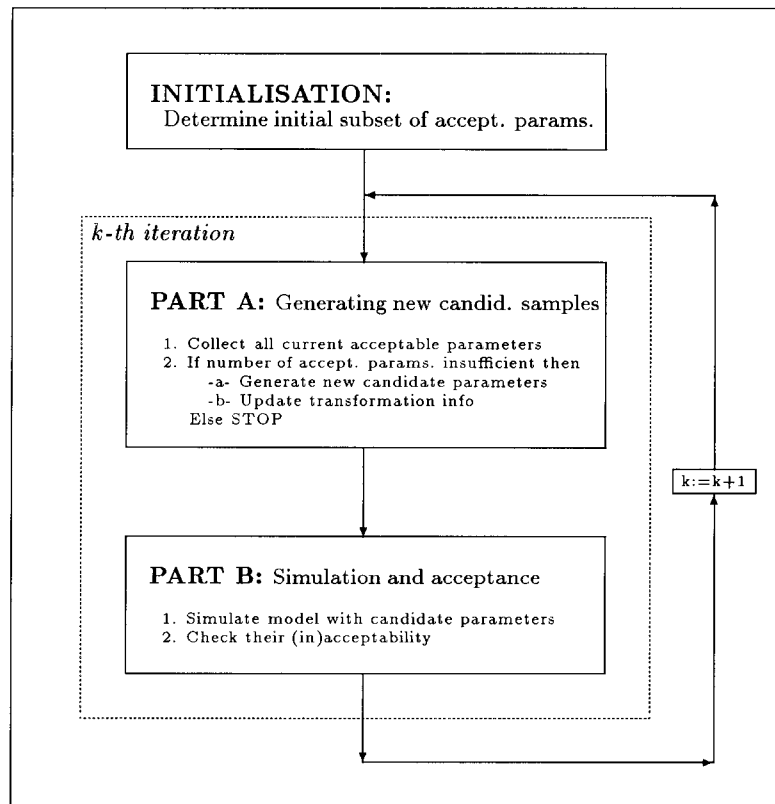


Figure 3: Flow scheme of the iterative rotated-random-scan procedure

1. Choice of the stop-criterion N_{lim} , i.e. the amount of acceptable parameters which one minimally likes to obtain.
2. Choice of the initial number N_{ini} of candidate samples, to be determined in the initialisation stage.
3. Choice of the number of new candidate samples N_{can} to be determined in PART A of each new iteration.
4. Choice of the sampling technique for performing the uniform random scan in the transformed parameter space ([1]: random sampling, or [2]:Latin Hypercube sampling).
5. Choice of whether or not the spurious correlations due to sampling should be corrected.
6. Choice of the *scaling* option: Various options are available to scale the acceptable parameters before computing their covariance matrix on basis of which the desired transformation or rotation in the parameter space is determined. Scaling can diminish the effects of numerical errors (rounding/truncation) due to possible strongly varying sizes of the parameter values¹⁰. Moreover it can serve to focus attention on certain aspects of the parameter range. The user can choose amongst four scaling options: [1] → no scaling; [2] → scaling on basis of the standard deviations of the currently available acceptable parameters; [3] → scaling on basis

¹⁰In calibration studies it is recommended to normalize parameters such that their values have comparable sizes.

of the range of the currently available acceptable parameters, or [4] \rightarrow scaling on basis of the user-specified range of the parameter set Θ .

7. Choice of the *working space* option: The covariance matrix of the (scaled) acceptable parameters which forms the basis for the transformation can be expressed in terms of, [1] \rightarrow the original coordinates, or [2] \rightarrow the transformed coordinates.
8. Choice of the *decomposition* option: In order to determine the transformation matrix for a typical iteration, the covariance matrix of the currently available (scaled) acceptable parameters has to be decomposed to obtain information on the principal directions in the parameter space. Two methods are available for decomposition of this matrix: [1] \rightarrow the eigen-system decomposition, or [2] \rightarrow the Cholesky decomposition; see appendix A for details.

Janssen and Sanders [1995] study the influence of the various choices in more detail, and present guidelines for making these choices: Concerning the choice of N_{can} a value is recommended which is (at least) 10 times the number of parameters ($\geq 10 * p$) (Keesman [1990]). It is, however, difficult to give practical guidelines for the choice of N_{ini} and N_{lim} . Adequate choices of these quantities depend on the specific form and volume of the complete set of acceptable parameters Θ_{acc} in the initial parameter space. This set is however a priori unknown. Consequently, the best one can hope for is to choose the initial candidate sample (N_{ini}) sufficiently large such that its ‘range’ covers, albeit very incomplete, the major part of subset Θ_{acc} . Likewise N_{lim} should ideally be chosen such that the finally obtained subset of acceptable parameters characterizes the complete set Θ_{acc} adequately. Since this, by nature, can not be guaranteed a priori, it is recommended to try various values for N_{lim} and N_{ini} , and to study whether this will influence the results considerably. See also item VII.

For the remaining calibration options 4-8, default choices are recommended (table 1). The tests in Janssen and Sanders [1995] illustrate however that non-default choices lead to comparable results.

Option	Default
Sampling Technique	Latin Hupercube Sampling
Correlation Correction	yes
Scaling	Standard deviation of acceptable parameters
Working Space	Original coordinates
Decomposition	Eigensystem-decomposition

Table 1: Recommended choices for the calibration options

[VI] Software:

Software tools have been developed to perform the rotated-random-scan method presented in appendix A (see Janssen [1995a] for a complete description of this software):

- The *initialisation* step of the procedure (i.e. STEP 0: obtaining an initial uniform random sample in the parameter space and determining which sampled parameter combinations lead to acceptable simulation results), can e.g. be performed with

the tools offered by the software package UNCSAM (Janssen, Heuberger, Sanders [1992]). The application program MCSAMP in UNCSAM can be employed to generate the initial uniform random sample. The program MODGEN is subsequently applied for passing these sampled parameters to the model in order to perform simulations which should be judged on their acceptability (see also part B).

- PART A of a typical iteration (i.e. generating new candidate samples on basis of a transformation and a random scan in the parameter space) can be easily performed with the program RORASC¹¹. RORASC has been developed in a model independent¹² way, thus offering the user the opportunity to treat a large variety of models, irrespective of their specific implementation. On input, RORASC requires data files which contain the currently available acceptable parameters. On output, it generates a data file with new candidate samples. For a complete description see Janssen [1995a].
- PART B of a typical iteration, i.e. simulation of the model for the sampled candidate parameters which have been obtained in PART A, and determination whether these results are acceptable, requires more explicit effort from the user: the various new candidate samples in the file which resulted from application of RORASC in step A should be passed in a suitable format to the model for subsequent simulation. This interfacing task can be easily performed with the input-interfacing program MODGEN of UNCSAM (see Janssen, Heuberger, Sanders [1992]). The subsequent task of simulating the model and judging the acceptability of the results has to be performed completely by the user. Careful processing and administration of the various simulation runs will be required. Much time can be gained if the simulation is stopped as soon as the results appear to be unacceptable. Especially for simulation of computer-intensive dynamic models this can be very beneficial.

After the initialisation step, the various iterations for (gradually) obtaining a sufficient number of acceptable parameter vectors are performed by repeated application of RORASC (PART A) and the subsequent ‘simulation-and-acceptance’ step (PART B). See figure 3.

Most input- and output files generated during these iterations have a structure which corresponds with the file-structure employed in UNCSAM. Therefore it is recommended to use RORASC in combination with UNCSAM to have maximal benefit of the various facilities offered by UNCSAM (e.g. file-handling, sampling, plotting etc.). Moreover UNCSAM plays a prominent role in accomplishing PART B of an iteration.

[VII] Disadvantages:

- Inherent to its nature, the set theoretic approach to calibration employs a rather ‘rough’ way of matching the model to the reality: it considers model results bluntly as ‘acceptable’ or ‘non-acceptable’, without a further differentiation e.g. in terms of a ‘degree of acceptability’. In ill-defined and information poor settings such a ‘rough’ approach is usually very adequate, but in other applications an additional fine-tuning of the model to the reality may be required.

¹¹RORASC = ROTated RAndom SCan. RORASC is a standard FORTRAN77 program which has been embedded in an ANSI-C environment to enable dynamic memory allocation.

¹²This could be achieved since PART A is completely separated from PART B, where all simulations and decisions on acceptability of the model results take place.

- The *initial step* of the procedure (STEP 0) typically relies on a straightforward uniform random scan of the original parameter space, and can therefore be rather *inefficient* when compared to the subsequent iterations performed on basis of an efficient rotated random scan.

The efficiency of this initial step depends on the (a priori unknown!) ratio $\rho = \text{Vol}(\Theta_{acc})/\text{Vol}(\Theta)$ between the volumes of the subset of *acceptable* parameters Θ_{acc} and the *prior* parameter space Θ in which the sampling is performed. If one generates initially N_{ini} sampled points uniformly in the parameter space Θ , then on the average $N_{ini} \cdot \rho$ of these samples will be acceptable¹³. Consequently N_{ini} has to be very large to obtain a sufficiently large set of acceptable samples in case that ρ is very small. This problem aggravates for higher dimensional parameter spaces (p large) principally due to two reasons, i.e. (a) ρ can be very small in these situations, and, (b) more acceptable parameters will be needed to obtain a fair, albeit largely incomplete, cover of the acceptable subset, especially if this subset has strongly irregular (non-ellipsoid, non-box) features. This illustrates the importance and the necessity to *restrict beforehand* the number of parameters to be calibrated, e.g. to maximally 6 à 10 parameters. See the remark at the end of this section for more information on this issue.

- The rotated-random-scan procedure uses rotations/transformations in the parameter space on basis of (scaled) covariance matrices; consequently it mainly focusses on *linear relations* between the parameters or equivalently on *ellipsoid/box* shaped parameter sets. Therefore one can expect that identification of whimsical, ‘non-linear’ (i.e. non-ellipsoidal), disconnected parameter sets can be rather inadequate/inefficient. The tests in Janssen and Sanders [1995] illustrate this.

[VIII] Remark:

For an effective use of the rotated random scan method it is desirable to restrict the number of parameters to be calibrated beforehand. Moreover, due to the scarcity and inaccuracy of the data, often only a small number of parameters ($\leq 6 - 10$) can be handled satisfactorily (see Keesman [1989b]).

This reduction of the number of parameters can e.g. be achieved by applying a preliminary *sensitivity analysis*¹⁴, using the software developed at the RIVM/CWM (Janssen, Heuberger, Sanders [1992], Janssen [1995b,1995c]). Various approaches can be applied, as discussed in the sequel:

1. **Conventional Sensitivity Analysis (CSA)** on basis of *Monte Carlo* techniques (sampling and simulation) in combination with regression analysis. Use can be made of the software package UNCSAM (see Janssen, Heuberger, Sanders [1992]) to perform this type of analysis.

The CSA is aimed at establishing how much the model outputs y (or functions of the model outputs; e.g. misfit criteria) will be affected by variations of the model parameters

¹³Suppose that initially $\theta_i \in [0, 1]$ for $i = 1, \dots, p$, and that the values $\theta_i \in [0, \frac{1}{2}]$ are acceptable. Then $\rho = (\frac{1}{2})^p$. Suppose that we need $2 \cdot p$ acceptable parameters for a representative covering in the initialisation phase. Then N_{ini} should be approximately equal to $2^{(p+1)} \cdot p$. For $p = 9$ this are already more than 9000 samples.

¹⁴In situations where calibration is primarily intended to reduce the *uncertainty in model predictions*, it is recommended to perform also a preliminary *uncertainty analysis* on basis of the currently available information on the uncertainty in the parameters (distributions), e.g. by using UNCSAM. This analysis helps to identify the parameters which contribute substantially to the uncertainty. These parameters should thus be the primary candidates for calibration.

θ around certain nominal values. The variations can be small (*local* sensitivity analysis) or large (*global* sensitivity analysis). Sensitivity is typically expressed in terms of (linear) regression coefficients which approximately characterize the relation between parameters and model outputs, and which can be seen as approximations of the first derivative $\partial y / \partial \theta_i$ of the output y with respect to the parameters θ_i .

This approach can be used to identify those parameters which have (nearly) no effect on the model outputs. These parameters can be excluded from calibration (e.g. by fixing them to a nominal value) without affecting the mismatch between model and data much. Although this approach helps restricting the number of parameters to be calibrated, it does not completely preclude problems in identifying the remaining parameters: it can occur that some sensitive parameter(combination)s affect the model outputs in completely similar fashion, and therefore cancel out their influence and thus can not be determined simultaneously on basis of measured outputs. See Janssen and Heuberger [1992] for more information on this so called *identifiability* issue.

Janssen and Heuberger [1992] also mention an important additional limitation of the conventional sensitivity analysis: the results of such an analysis pertain to a specific nominal point in the parameter space around which the (local or global) variations are considered, and can be *heavily dependent* on this point. A large dependence indicates that strong non-linear effects are present in the relationship between parameters and model outputs; therefore the results of a conventional sensitivity analysis in a specific nominal point need not be representative for the complete parameter space. In this situation the (local) conventional sensitivity analysis has to be repeated for various nominal points in the parameter space. It can however be difficult to determine how many extra analyses should be performed and where the associated nominal points should be chosen, especially in calibration situations where the initial parameter space is large and where information on meaningful and representative (nominal) parameter values is unreliable due to the ill-definedness and information-poor character of the system under study.

2. **Generalized Sensitivity Analysis (GSA)**, which is particularly appropriate for ill-defined and information-poor situations. This technique was originally proposed by Hornberger, Spear and Young (Young et al. [1980], Hornberger and Spear [1980,1981,1983], Hornberger and Cosby [1985], Spear and Hornberger [1980]) as a heuristic tool to identify dominant aspects, parameters or processes during early stages of modelling, thus indicating useful directions for further research.

The method is based on a rough subdivision (dichotomy) of the model output space in a subset of model outputs which are considered as *acceptable (behaviour)*, and a complementary subset of model outputs which are considered as *inacceptable (non-behaviour)*. The prior knowledge on the parameters is expressed in terms of probability distributions (e.g. uniform distributions to characterize the range and parameter bounds); samples are randomly drawn from these distributions and the model is simulated for the sampled parameter vectors. The above mentioned dichotomy of the model output space in ‘behaviours’ and ‘non-behaviours’ now induces a corresponding dichotomy of the sampled parameters in ‘behaviour-giving’ and ‘non-behaviour giving’ parameter vectors. The discrepancy between these subsets is evaluated and parameters or parameter combinations¹⁵ are identified which yield a large discrepancy. These parameter(s) combinations are then considered as sensitive/important for the applied dichotomy, and should thus be included in the calibration.

The above mentioned discrepancy between the ‘behaviour giving’ or ‘non-behaviour giving’ parameter sets can be evaluated in terms of e.g. the associated ‘difference’ in the means, the variances or the cumulative (marginal) distribution functions of the various

¹⁵If the ‘behaviour-giving’ parameters show important mutual correlations, one typically applies a *transformation* of the parameter space (on basis of an eigen-system decomposition of the covariance matrix) to make the transformed parameters uncorrelated. These transformed parameters are combinations of the original parameters. The discrepancy between the ‘behaviour giving’ and the ‘non-behaviour giving’ subset will be expressed in terms of the transformed parameters, i.e. in terms of the established parameter combinations.

parameters or the considered parameter combinations. Especially the ‘difference’ between the (marginal) distribution functions is considered as an important discrepancy measure and is expressed in terms of the well-known Kolmogorov-Smirnov statistic (see Press et al. [1986]).

The above procedure depends on the prior distributions which are specified for the parameters. These distributions affect the ability to obtain samples in the ‘behaviour’ or ‘non-behaviour’ subset, and therefore the results can be significantly influenced by alteration of the distributions (cf. Spear and Hornberger [1980]). To mitigate this effect, Keesman [1989b] proposes a minor modification of the GSA, by applying it to the set of sampled parameter values which has been obtained after performing various iterations of the rotated random scan procedure. This *a posteriori* parameter set renders a more efficient and representative characterization of the ‘behaviour-giving’ subset, and is therefore expected to generate better and more robust results.

Despite the heuristic nature of the GSA (Hornberger and Cosby [1985]) it appears to be a practically useful approach to identify the critical parameters or processes and hence restrict the number of parameters requiring calibration or indicate useful directions for further model development (see also the recent application study of Lence and Takyi [1992]). Its close relationship with the set-theoretic calibration (notice the dichotomy between ‘behaviours’ and ‘non-behaviours’) makes it especially suited in this context.

The generalised sensitivity analysis method has been implemented in a computer program and is available on request (see Janssen [1995b] for more information).

3. **Dominant Direction Analysis (DDA)**, aimed at determining the *dominant directions* (parameter combinations) in the parameter space, on basis of an eigen-system decomposition of the (scaled) covariance matrix of the *acceptable* (i.e. ‘behaviour-giving’) parameters. The eigenvectors associated to the smallest eigenvalues (i.e. smallest uncertainty) indicate the best-determined directions. Notice that the method in fact focusses on *linear* relationships between parameters (dominant directions).

This method, which also fits well into the set-theoretic calibration framework, was proposed by Keesman [1989b]. For applications it requires an (initial) set of *acceptable* parameters, which often can only be obtained after performing some preliminary calibration on the complete, i.e. non-reduced, set of parameters. The ‘dominant direction’ method has also been implemented in a computer program and is available on request (see Janssen [1995c] for more information).

Unlike the CSA, the last two methods (GSA and DDA) in fact require an *initial set of acceptable* parameters. Typically a preliminary calibration of all parameters will be needed to determine this initial set. In order to prevent an excessive number of model runs¹⁶, it is recommended to perform this calibration in a tentative way applying e.g. less strict conditions and using only limited information on the system (see Keesman [1989b]). Later on, when applying a more complete and elaborate calibration, the results of this GSA or DDA can be used to decide on how to restrict the number of parameters which actually should be calibrated more accurately.

Another point of difference between the CSA and the GSA/DDA, is that the latter approaches are oriented towards *global* sensitivity statements (i.e. for the whole parameter space), while the former typically suffers from the drawback that it pertains to a particular (nominal) point and thus may be of limited use for obtaining statements on the parameter space as a whole. This difference in orientation is mainly due to the fact that the GSA and, to a lesser extent, the DDA consider the model in a far more crude way than the CSA: classifying the model outputs as ‘acceptable’ or ‘non-acceptable’,

¹⁶It is obvious that this is imperative, since the very reason for performing the sensitivity analysis is the reduction of the computational load.

instead of looking at their actual values or variations, enables that the ‘sensitivity’ can be assessed for the whole parameter space. Certainly a price has been paid for this ability: the statements of the GSA and the DDA, albeit global, are less informative and differentiated than the statements of a CSA which refer, albeit local, to actual changes in the output *values* due to changes in the parameters.

Summarizing the above discussion on approaches to sensitivity analyses, it can be concluded that the GSA and the DDA are especially useful in the context of set-theoretic calibration for ill-defined and information poor systems. The results of the GSA and DDA refer to the impact of individual parameters and their linear combinations on *acceptable* model behaviour and pertain to the *complete* parameter space (global orientation). The GSA studies the parametric *sensitivity* w.r.t. the (in)-acceptability dichotomy, while the DDA in fact analyses the parametric *uncertainty* of the acceptable parameters.

The CSA, on the other hand, is more appropriate for conventional calibration studies e.g. aimed at *minimizing* a misfit criterion. It renders more differentiated and detailed results than the GSA and DDA, but suffers from the drawback of (possible) local dependence on the nominal point in the parameter space where the CSA is performed. This makes the CSA sometimes less suitable for adequate assessment of the sensitivity in models for ill-defined, information poor systems.

4 A practical calibration study

The rotated-random-scan method is illustrated in this section by a preliminary calibration study which has been performed with the model SMART (see Posch et al. [1993]).

The model

SMART is a one-layer dynamic soil-acidification model which has initially been developed to obtain insight in the impact of different emission scenarios on forest soils in Europe (De Vries et al. [1989]; Posch et al. [1993]). It computes important indicators for the status and effect of soil-acidification, such as the concentrations of Al, monovalent base cations (BC1, i.e. Na+K), divalent base cations (BC2, i.e. Ca+Mg), NH_4 , SO_4 , NO_3 and H in the soil solution, using a temporal resolution of one year. Annual values of hydrological data (drainage fluxes, root uptake fluxes and water contents) serve as input for SMART, as well as annual deposition data of SO_2 , NO_x , NH_3 and base cations (BC1, BC2).

SMART incorporates most of the geochemical processes which contribute to acidification (weathering, cation exchange, sulphate adsorption, dynamic nitrogen immobilization), but only a very limited number of biological processes is taken into account. Nutrient cycling processes (e.g. litterfall, root decay and mineralization) are not included since the model is based on the assumption that the amount of organic matter is in equilibrium. Cation exchange, sulphate adsorption, dissolution of carbonates and Al-hydroxides are treated as equilibrium reactions, while weathering of base cations and (de)nitrification are described as first-order reactions (see De Vries et al. [1989], Posch et al. [1993]).

The data

The model SMART is applied to an intensively monitored spruce site at Solling, Germany (Van der Salm et al. [1995]). Input parameters for SMART have been derived from measured data concerning bulk precipitation, throughfall and soil solution chemistry at this site (Bredemeier et al. [1995], Tiktak et al. [1995]). For calibration, the simulated concentrations and leaching fluxes are compared with the measured values at this strongly acidified site during the period 1973-1989. In the present study, the comparison concerns the flux-weighted values in the *subsoil* (90 cm depth), averaged over a time period of one year. To obtain these values, the weekly measurements have been recasted into annual flux-weighted concentrations¹⁷. These annual data are displayed in figure 4. Notice the strong rise of SO_4 and Al in the 1970s. This is caused by the high sulphate¹⁸ load in the early 1970s onto a soil which had already been saturated with respect to sulphate sorption in the preceding decades. In the 1980s, the concentrations of SO_4 , Al and BC2 decrease due to reductions in the atmospheric deposition of SO_4 and base cations (Ca and Mg) as a result of reduced industrial emissions after 1976. The decrease of SO_4 and Al is more gradual due to the desorption of adsorbed SO_4 , and the buffering effects by Al immobilisation.

The observed H concentrations showed more or less similar tendencies as Al and SO_4 , but the decrease in the 1980s is far more pronounced. Finally notice that the BC1 concentration

¹⁷The flux-weighted annual average concentrations were obtained by dividing the 'measured' leaching fluxes by the (simulated) annual water fluxes. The 'measured' leaching fluxes have been calculated by multiplying measured concentrations with (monthly) simulated water fluxes, obtained by the model SWATRE, see Belmans et al. [1983].

¹⁸The behaviour of Al is very similar to the SO_4 behaviour because large parts of the (acid) SO_4 input are buffered by Al mobilisation.

remains nearly constant, and that NO_3 and NH_4 show a rather fickle behaviour, which is mainly caused by hydrological variations. In the long run both concentrations stay more or less at the same level.

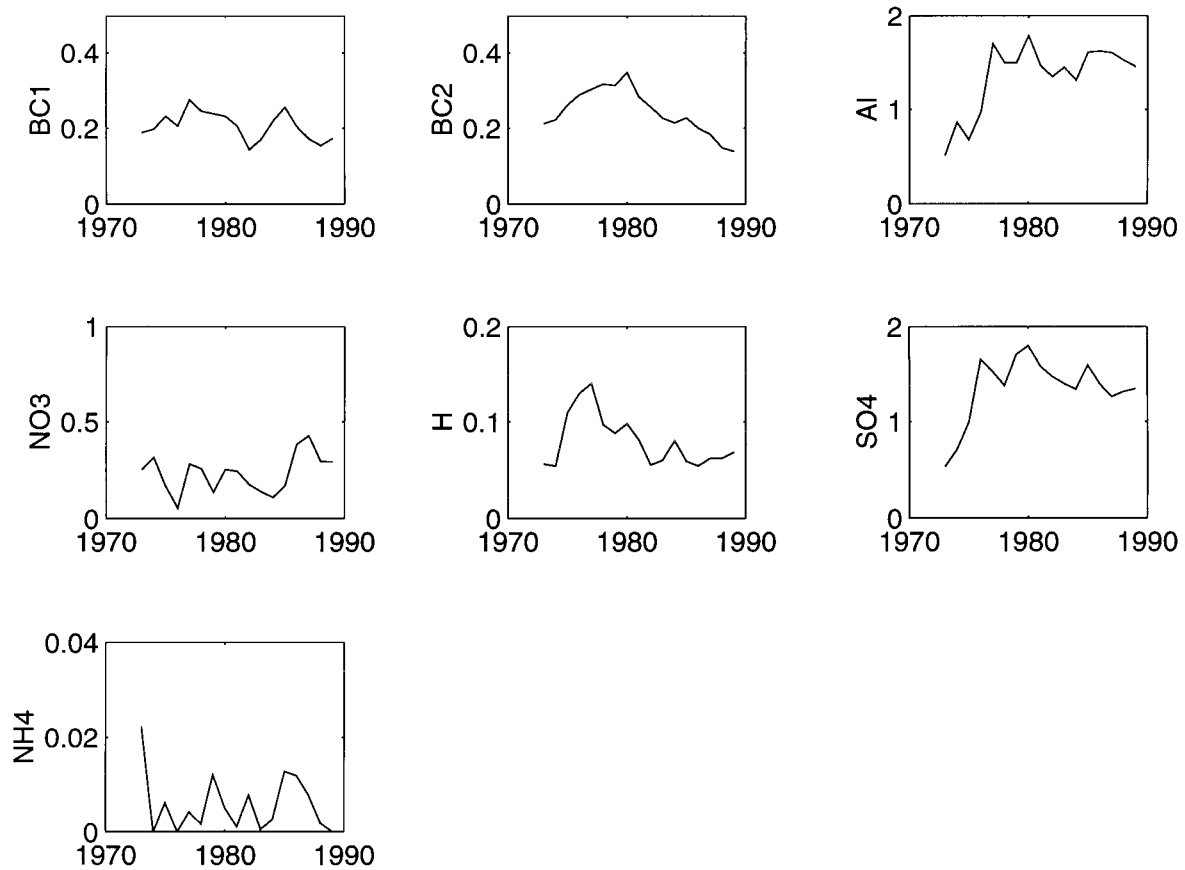


Figure 4: Annual flux-weighted concentration measurements over the period 1973-1989 at the spruce site in Solling, Germany

The parameters

The present calibration study aims at matching the model outputs BC1, BC2, Al, NO_3 , H, SO_4 and NH_4 with the above-mentioned ‘measurements’ from figure 4, and is primarily intended to illustrate the Rotated-Random-Scan method. The choice of the model parameters to be estimated is largely based on expert judgment and previous experience with the related soil-acidification model RESAM (see Kros et al. [1993]), and concerns parameters which are considered highly uncertain and of influence on the chosen model outputs. This leads to the parameters indicated in table 2; nominal values are specified in this table, as well as their ranges (minima, maxima), which characterize the initial uncertainty.

The parameter $k_{Al_{ox}}$ is the gibbsite equilibrium constant, which characterizes the dissolution of $\text{Al}(\text{OH})_3$ in non-calcareous soils. $k_{Al_{ex}}$ is the selectivity constant for the Al-BC exchange. f_{den} and f_{nit} are the (de)nitrification factors. The factors F_{BC1} , F_{BC2} , F_{SO2} are multiplication factors to upgrade/downgrade the annual input-deposition fluxes of BC1, BC2 and SO_2 . The considered parameters thus refer to process parameters as well as deposition input-parameters. The corresponding model simulation for the nominal parameters is displayed in

figure 5. Notice that the dynamic features of observed concentrations of BC1, BC2, Al and SO₄ are mimicked fairly by the model results. The modelled H concentration does not display the observed H increase in the 1970s. This is most likely caused by a large dispersion of the H front, which is inherent to a one layer system, as used by SMART.

Parameters	Unit	Nom.	Min.	Max.	Important effect on:
kAlox	$\log((mol\ l^{-1})^{-2})^3$	9.3	7	10	Al,H
kAlex	$\log(l\ mol^{-1})^3$	0.54	-1	5	Al, H, BC2
F_SO2	-	1	.5	1.5	SO ₄
F_BC1	-	1	.5	1.5	BC1
F_BC2	-	1	.5	1.5	BC2
fden	-	0.10	0.00	0.50	NO ₃
fnit	-	0.98	0.80	1.00	NH ₄ , NO ₃

Table 2: Parameter ranges, nominal values and associated relevant outputs

Remark 1: The model simulations for the nominal parameters (cf. figure 5) indicate, amongst all, a considerable underestimation of the SO₄ and Al concentration. This can be due to several reasons, e.g. a bad specification of associated model parameters (especially the weathering-rates *FBC1we*, *FBC2we* for the monovalent and divalent base cations, and the half-saturation constant *kSO4ad* for SO₄-adsorption), inadequate description¹⁹ of the involved mechanisms (i.e. process-equations), inadequacy of the input-data which are needed for the model (e.g. deposition flux data of BC1, BC2, SO₂ etc. for the period 1973-1989), and, last but not least, inadequate measurement values (i.e. the flux-weighted annual average concentration measurements).

Though a more complete calibration study would require that the process-parameters *FBC1we*, *FBC2we*, *kSO4ad* should also be estimated, it was decided to attribute the mismatch in the SO₄ and Al concentrations solely to flaws in the deposition data, and not to inadequate values of these process-parameters. Three extra parameters *F_BC1*, *F_BC2*, *F_SO2* have therefore been introduced to account for inadequate deposition fluxes. These parameters are multiplication factors for the input-deposition trajectories, to upgrade or downgrade these deposition-fluxes over the entire time period 1973-1989.

Exclusion of the above-mentioned process parameters *FBC1we*, *FBC2we*, *kSO4ad* from calibration has the advantage that extra identifiability problems are avoided: it is expected that the individual effects of the involved process-parameters *FBC1we*, *FBC2we*, *kSO4ad*, and the multiplication-factors *F_BC1*, *F_BC2*, *F_SO2* counterbalance each other, which leads to problems when trying to identify them simultaneously.

Needless to say there is a certain arbitrariness²⁰ in this approach, since one assigns supposedly *certain* (i.e. fixed) values to parameters (e.g. *FBC1we*, *FBC2we*) which are inherently uncertain. Ideally a more complete calibration study would be needed which considers these parameters also as candidates for calibration, and which relies on a thorough sensitivity analysis to select the parameters which will be estimated.

¹⁹It is e.g. expected that the ‘gibbsite-formulation’ employed in SMART falls short in accurately describing the H, Al and SO₄ behaviour.

²⁰The use of (uniform) multiplication-factors to account for deposition-uncertainty is another arbitrary issue.

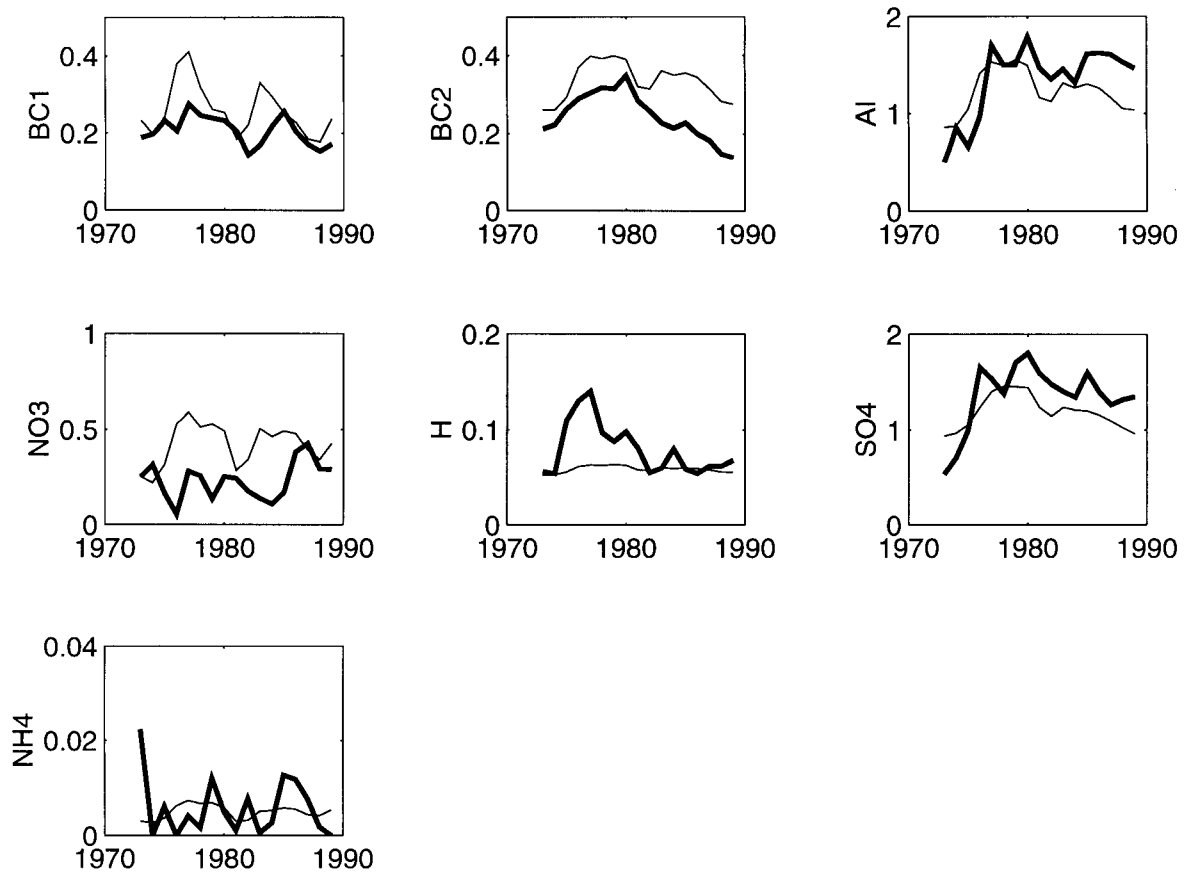


Figure 5: Comparison of measurements (fat, solid) and results from the nominal model (thin, solid)

The criterion

For application of the ‘rotated-random-scan’ method one needs to explicitly specify when the model results are judged as (in)acceptable, when comparing them with observations and accounting for the desired level of accuracy for the intended model application. For this purpose a misfit criterion is employed which expresses quantitatively how much the model results deviate from the measurements. Various choices for such a criterion are possible. It was decided to use a ‘*normalized average absolute deviation*’ criterion, due to its straightforwardness and fair robustness to outliers²¹. It has the form²²:

$$C(\cdot) = \frac{1}{7} \cdot \sum_{i=1}^7 \left(\frac{\sum_{j=1}^{17} |Y_i(j) - M_i(j)|}{\sum_{j=1}^{17} M_i(j)} \right) \quad (1)$$

where $Y_i(j)$ denotes the value of the i -th model output in the year $1972 + j$, and $M_i(j)$ denotes the corresponding measured value. Here $i = 1, \dots, 7$ refers to the 7 model outputs

²¹Expressing the deviations as a sum of absolute values, instead of as a sum-of-squares, renders better robustness properties to outliers.

²²For NH_4 the ‘measurements’ sometimes fall below the detection-limit. In these situations, the evaluation of the absolute deviation is adapted, by replacing $M_i(j)$ in (1) by the value of the detection-limit (i.e. 0.005), and by replacing $Y_i(j)$ by its censored value, i.e. $\max(Y_i(j), 0.005)$.

which are considered (i.e. BC1, BC2, Al, NO₃, H, SO₄, NH₄), while $j = 1, \dots, 17$ indicate the consecutive years in the period 1973-1989.

Notice that this criterion is composed of the normalized average absolute deviations:

$$C_i := \sum_{j=1}^{17} \left[\frac{|Y_i(j) - M_i(j)|}{\sum_{j=1}^{17} M_i(j)} \right] = \frac{\frac{1}{17} \sum_j |Y_i(j) - M_i(j)|}{\frac{1}{17} \sum_j M_i(j)} \quad (2)$$

of the individual model outputs ($i = 1, \dots, 7$). These deviations C_i are weighted equally in the overall criterion, and thus $C(\cdot)$ is the average of the individual misfits (i.e. $C(\cdot) = \frac{1}{7} \sum_i C_i(\cdot)$), allowing for mutual trade-offs²³ between the misfits C_i .

Remark 2: C_i denotes the average (over time) absolute deviation between model output $Y_i(j)$ and measurement $M_i(j)$, expressed in terms of the averaged measurements $\frac{1}{17} \sum_j M_i(j)$. The normalization by $\frac{1}{17} \sum_j M_i(j)$ serves to treat the outputs in a universal way, making the deviations dimensionless. A value of $C_i = .3$ means that the absolute deviation for the i -th output is a factor .3 of the averaged measurements (i.e. an average absolute deviation of 30 %).

■

Since the criterion (1) expresses the overall deviation on the *average*, visual inspection of the individual model-outputs w.r.t. the measurements adds useful additional information on the discrepancy between model results and measurements. Consider e.g. figure 5 and compare this with the numeric information on the individual nominal misfits in table 6: The overall criterion value for the nominal parameters is 0.37 (table 3), while the contributions of the various outputs vary considerably. The mismatch for the nominal model is especially large for NO₃ (about .9, or 90 %). NH₄ and BC2 show a nominal misfit of about 40 % (.4), H and BC1 of 30 % (.3) and Al and SO₄ of 20 %.

The acceptance strategy

The final aim of the present calibration study is to determine a set of at least 70 parameter samples (i.e. $N_{lim} = 70 = 10 \cdot p$, where $p (=7)$ denotes the number of calibrated parameters) which render an ‘acceptable’ misfit, allowing e.g. for a deviation of 25 % from the minimal misfit. To achieve this in each iteration of the rotated-random-scan procedure candidate parameter samples $\hat{\theta}$ are considered as acceptable²⁴ if the associated misfit $C(\hat{\theta})$ is within a factor F_{crit} around the minimal value in that iteration (i.e. $C(\hat{\theta}) \leq F_{crit} \cdot C_{min}$).

Notice that this acceptance strategy is in fact iteration dependent, since it refers to the actual minimal criterion value in each iteration. Consequently an extra final acceptance test will be required to check which samples fulfill the condition $C(\hat{\theta}) \leq F_{crit} \cdot C_{min}^{tot}$ where C_{min}^{tot} is the minimal criterion value for all iterations.

The calibration results

To initialize the calibration (iteration 0) $N_{ini} = 50$ candidate samples are determined by uniformly sampling the parameters between the bounds indicated in table 2. Subsequently 4 iterations of the random-random-scan method were performed. The results are reported in table 3 and indicate that 82 ‘acceptable’ parameters have been generated applying the

²³An alternative approach would be to define $C(\cdot) = \max C_i(\cdot)$. This approach in fact focuses on the worst-case situation.

²⁴The exceedance of the parameter-ranges indicated in table 2 will be no reason for rejection.

above acceptance strategy with the listed values F_{crit} . These values have deliberately been chosen too high²⁵ in this experiment to ensure that still a considerable number of ‘accepted’ candidates could be determined. Stricter (i.e. lower) values of F_{crit} would have resulted in an inefficient zoom-in search.

As a consequence of this liberal strategy only a small part ($N_{acc} = 13$) of the resulting 82 parameter samples show misfits which deviate less than 25 % from the minimal one. These samples are subsequently selected as a starting point (prior information) for a second calibration experiment to extend the set of acceptable parameter samples. The results are summarized in table 4 and show that 84 runs result after 4 iterations, using $F_{crit} = 1.25$. In addition those runs are selected with misfit values that deviate less than 25 % from the *overall*²⁶ minimum. Discarding finally the runs for which the parameter F_BC1 is beyond its prespecified limits (i.e. $\leq .5$), and for which the parameter F_BC2 is negative, 64 parameter samples remain which are considered as acceptable.

The minimum, mean, median and maximum of these selected 64 parameter-values are listed in table 5 and give an indication for the resulting uncertainty after calibration (confidence region). This information is displayed graphically in the scatterplots of figure 7 and in the histograms of figure 8.

To obtain information on the misfit-contributions of the various outputs the normalized average absolute deviations $C_i(\cdot)$ (see equation 2) of the individual model outputs have been evaluated for the selected 64 parameters, and their statistics are summarized in table 6, showing that the misfit is largest for NO_3 , and NH_4 (approximately .5, i.e. 50 %), followed by H (approximately .3, i.e. 30 %), BC2, BC1, Al (approximately .2, i.e. 20 %) and SO_4 (approximately .17, i.e. 17 %). Notice the considerable misfit improvement when comparing the optimal setting with the nominal setting, especially for NO_3 , BC2 and BC1. The fit for Al however deteriorates somewhat, and the fit for NH_4 does not change. See also figure 6.

Discussion

From these results it is obvious that the uncertainty is reduced substantially for the parameter f_{nit} , and to a far lesser extent for the parameters $kAlox$, $kAlex$, F_SO2 , F_BC1 ; see the ranges in table 5 and in figures²⁷ 7, 8.

The scatterplots for the criterion value in figure 7 suggest ‘unimodality’ for most parameters, and seem to indicate a clear and unique optimum for the criterion. No significant correlations between individual parameters can be inferred from figure 7, except some slight connection between $kAlex$ and F_BC2 , and between F_BC2 and f_{den} . Observe moreover that the two latter parameters exceed the pre-specified ranges of table 2, and that the nominal parameter value for f_{den} differs much from the mean, median or optimal values (table 5).

The contribution of the various individual misfits C_i to the overall criterion $C(\cdot)$ for the selected parameters is shown graphically in the scatterplots of figure 9. This renders information on the potential trade-off’s between the misfits C_i . Notice e.g. the negative correlations between various misfits, e.g. between NH_4 and H, SO_4 , NO_3 etc., indicating that the improvement of one individual misfit, tends to lead to a worse fit for the other outputs.

²⁵Properly speaking F_{crit} should be 1.25 to allow for a deviation of 25 % around the minimal misfit value.

²⁶I.e. over all iterations.

²⁷The range of the x-axis in these figures denote for all parameters, except F_BC2 and f_{den} , the initial parametric uncertainty range).

Post-calibration uncertainty analysis

Using the 64 acceptable parameter samples for simulating the model, the ‘uncertainty’ of the model outputs can be assessed which remains after calibration. Figure 10 illustrates that the uncertainty bands can be fairly wide; nevertheless it is still not guaranteed that the measurements fall completely within these bands. The uncertainty band for BC2 is very wide due to the fact that the parameter F_{BC2} takes values over a broad range (0.4-1.2).

Summary

The previous results show that the uncertainty in the parameter f_{nit} and to a lesser extent in k_{Alox} , k_{Alex} , F_{SO2} and F_{BC1} is reduced by calibration. The other parameters f_{den} , F_{BC2} are badly identifiable, and moreover take values over a wide range, which need not conform with the initial parameter bounds.

The resulting uncertainty in the model outputs is large (for BC2 very high), but the measurements still do not fall completely within these bands. This is an indication that the model with the employed input data is not fully able to describe the data.

Iteration	Min $C(\cdot)$	F_{crit}	N_{sam}	N_{acc}	% Accept.
Nominal	0.37	-	1	-	-
0	0.38	1.75	50	13	26 %
1	0.35	1.5	50	13	26 %
2	0.29	1.5	50	11	22 %
3	0.32	1.5	50	20	40 %
4	0.34	1.5	50	25	50 %
Total	0.29	-	250	82	33 %

Table 3: Info on criterion values for various iterations in the first calibration experiment

Iteration	Min $C(\cdot)$	F_{crit}	N_{sam}	N_{acc}	% Accept.
Nominal	0.37	-	1	-	-
0	0.29	1.25	82	13	16 %
1	0.27	1.25	50	14	28 %
2	0.26	1.25	50	18	36 %
3	0.26	1.25	50	16	32 %
4	0.26	1.25	50	23	46 %
Total	0.26	-	282	84	30 %

Table 4: Info on criterion values for various iterations in the second calibration experiment

	kAlox	kAlex	F_SO2	F_BC1	F_BC2	fden	fnit
Min	8.30	-1.01	0.91	0.56	0.02	0.21	0.95
Median	8.98	0.58	1.10	0.83	0.42	0.47	0.98
Mean	8.95	0.67	1.10	0.85	0.50	0.46	0.98
Max	9.64	2.50	1.33	1.14	1.26	0.74	1.01
Nomin.	9.30	0.54	1.00	1.00	1.00	0.10	0.98
Optim.	9.00	0.47	1.25	0.78	0.54	0.47	0.98
Nmb. exceed.	0	1	0	0	38	27	2

Table 5: Statistics for selected accepted parameter values ($C(\cdot) \leq 1.25 \cdot \min C(\cdot)$) after calibration. The nominal and optimal parameter value and the number of exceedances of the parameter bound are also indicated.

	BC1	BC2	Al	NH ₄	NO ₃	H	SO ₄
Min	0.18	0.09	0.16	0.42	0.43	0.23	0.15
Median	0.20	0.21	0.21	0.45	0.49	0.26	0.16
Mean	0.23	0.22	0.21	0.48	0.51	0.31	0.17
Max	0.43	0.45	0.32	0.75	0.71	0.65	0.25
Nomin.	0.27	0.37	0.19	0.42	0.88	0.28	0.19
Optim.	0.18	0.12	0.23	0.42	0.47	0.23	0.18

Table 6: Statistics for the misfit-contributions C_i of the various model outputs; the statistics concern the selected acceptable parameter values after calibration.

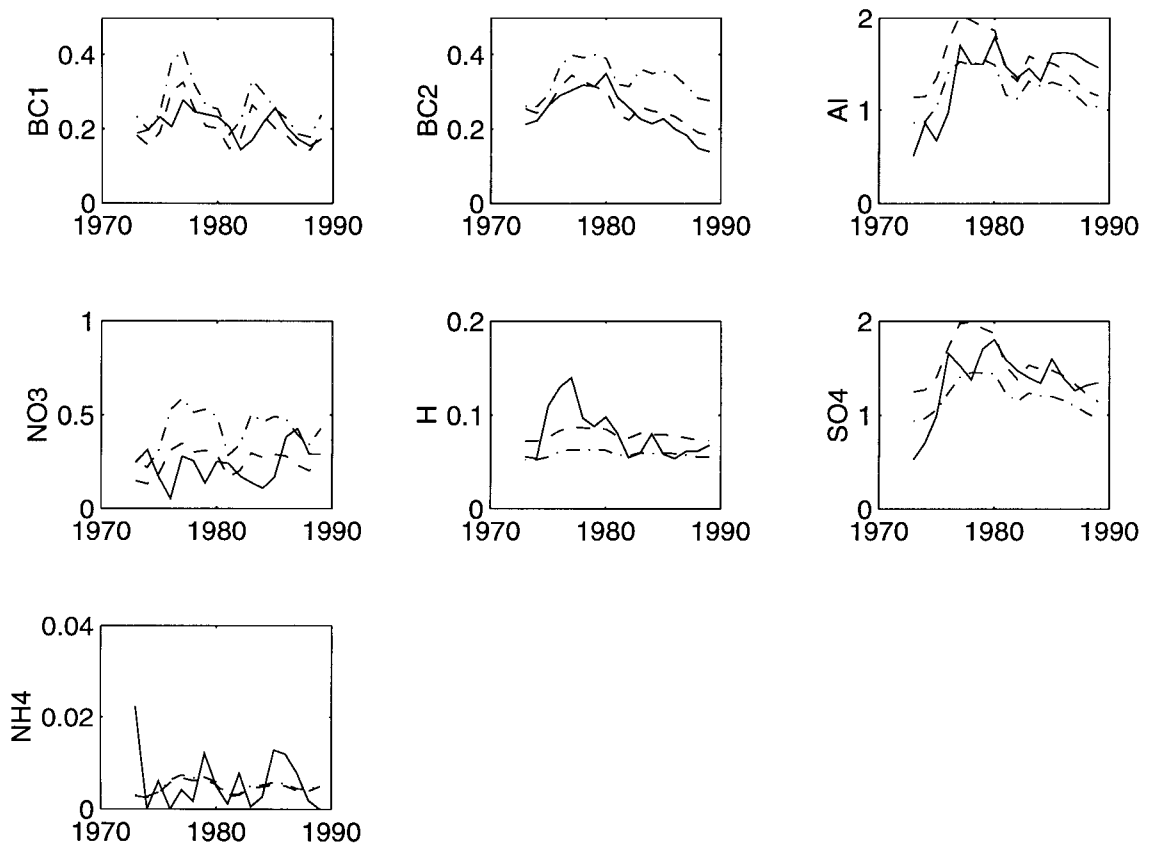


Figure 6: Comparison of measurements (solid) and results from the nominal model (dash-dot) and the optimal model (dash).

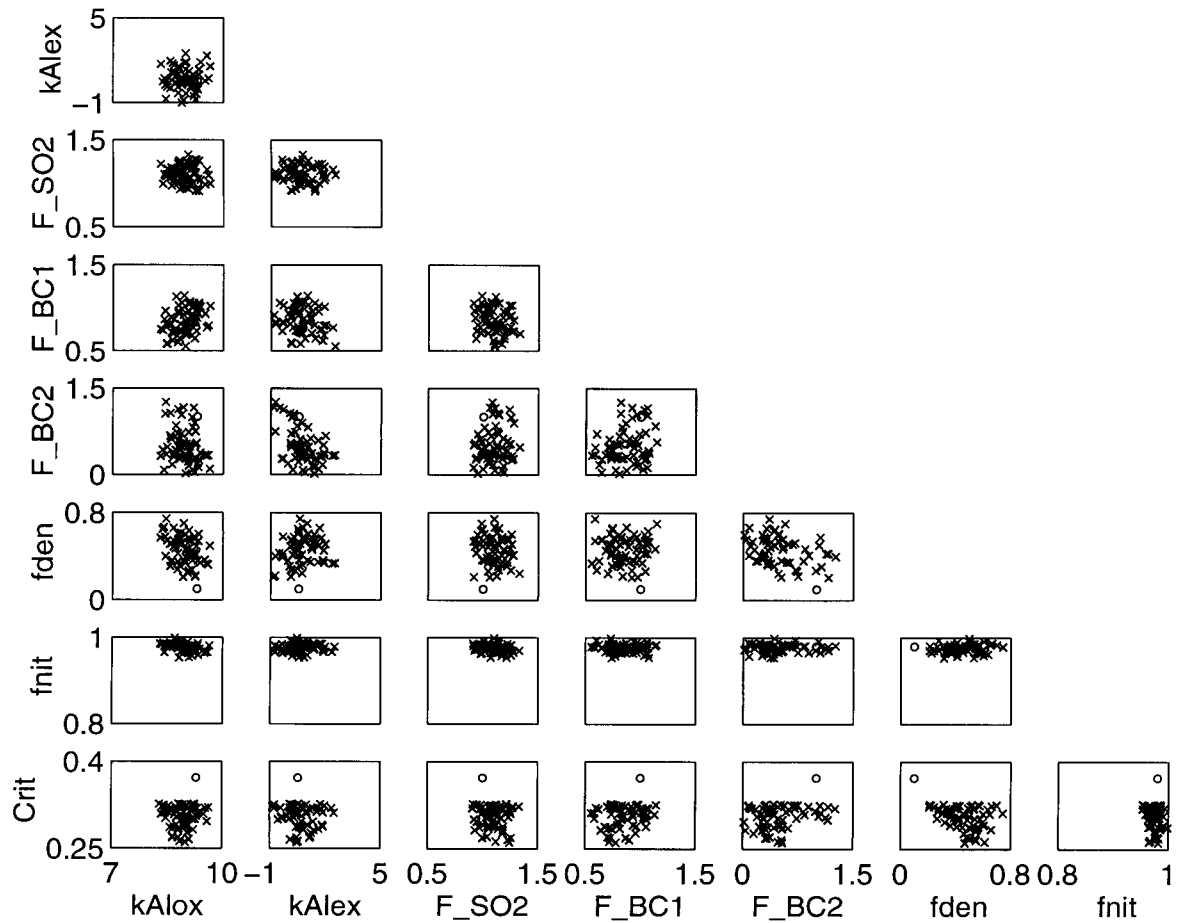


Figure 7: Scatterplots showing the selected ($C(\cdot) \leq 1.25 \cdot \min C(\cdot)$) acceptable parameters after calibration. The nominal values are indicated by a dot (o).

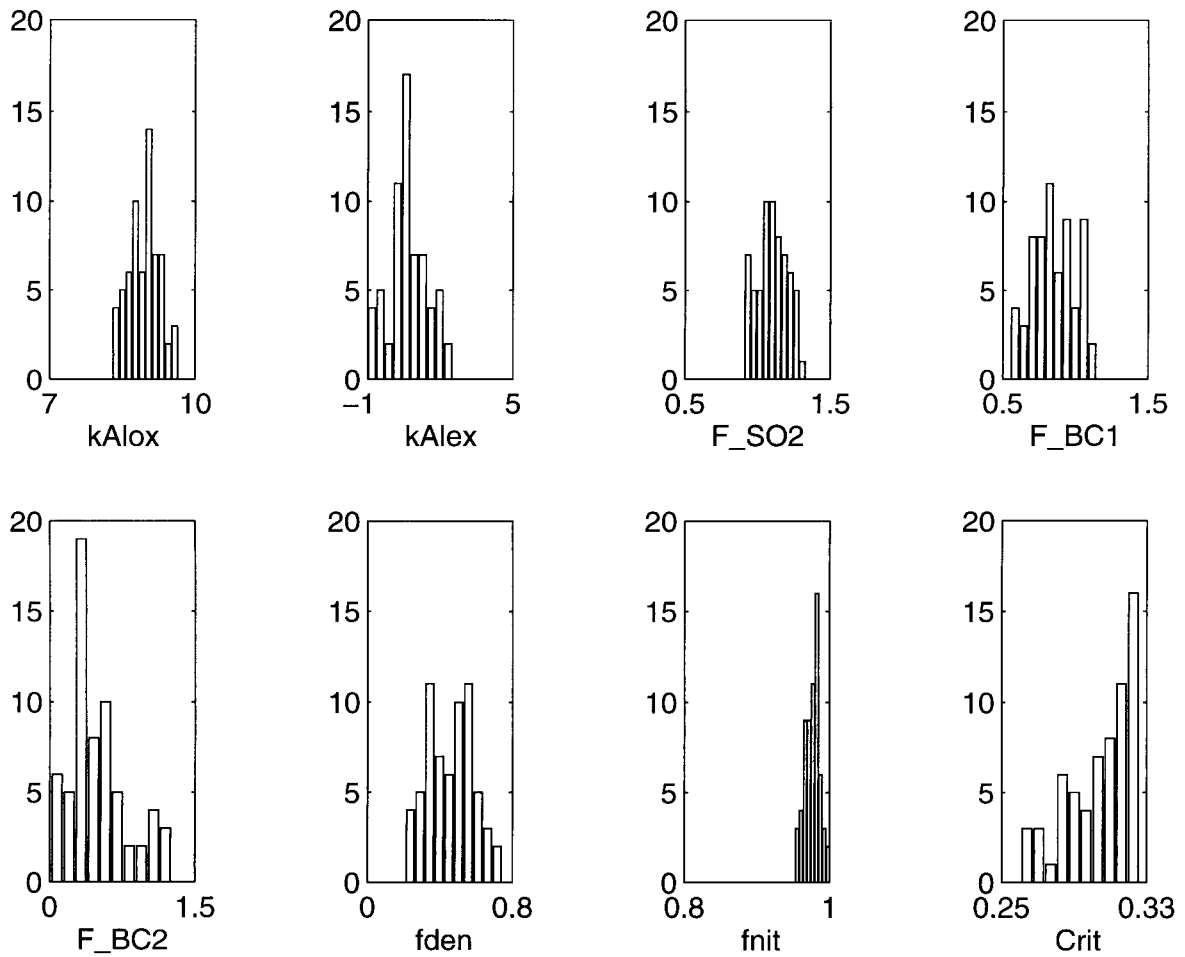


Figure 8: Histograms for the selected ($C(\cdot) \leq 1.25 \cdot \min C(\cdot)$) acceptable parameters after calibration. Also the associated criterion values are indicated (Crit).

The x -axis for all parameters except F_BC2 , $fden$ reflects the initial parametric uncertainty range.

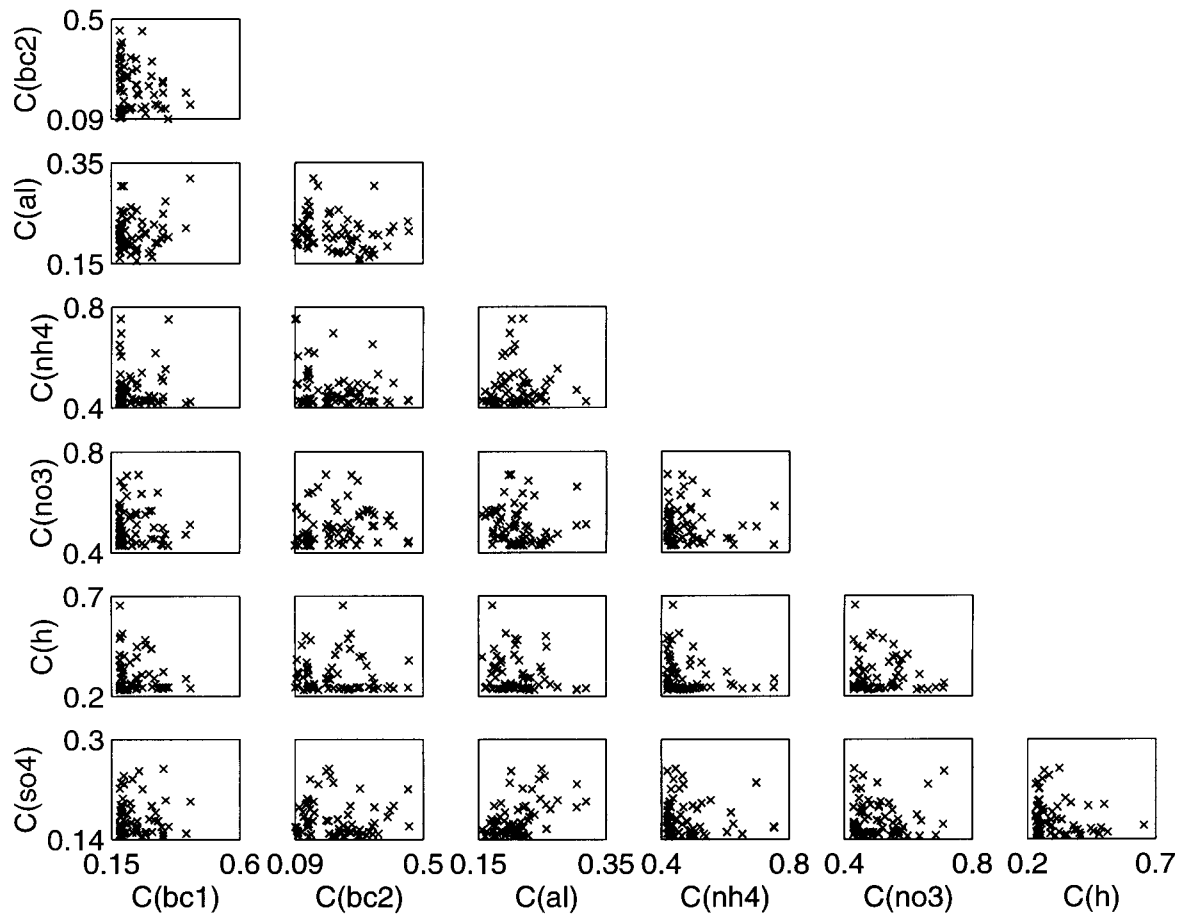


Figure 9: Scatterplots showing the trade-offs between the misfits for the various model outputs. The misfits are evaluated for the selected acceptable parameters after calibration.

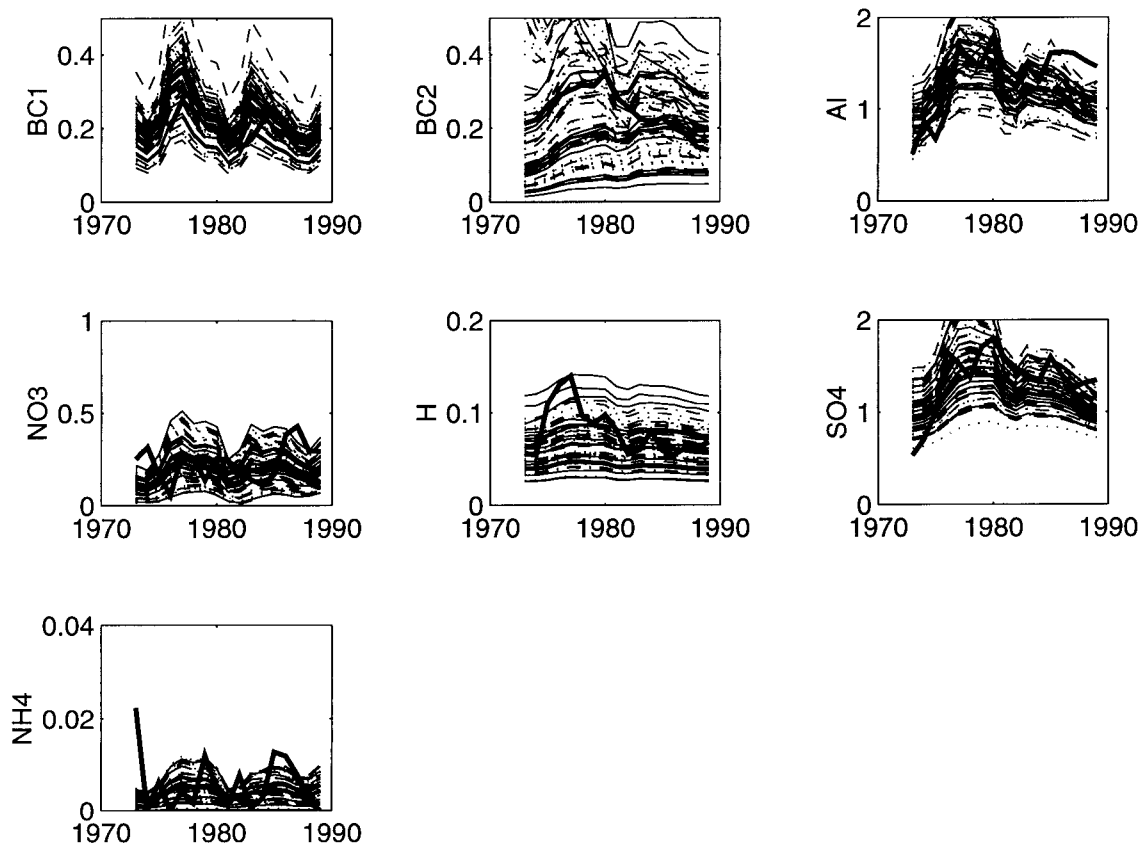


Figure 10: Uncertainty band around model outputs after calibration, simulating the selected acceptable parameters. The fat line refers to the measurement data.

5 Discussion

In information-poor and ill-defined settings, which are typically characterized by scarcity and inaccuracy of data or incomplete and uncertain process knowledge, a statistical approach to calibration is often misplaced. A viable alternative is to use a set-theoretic approach which aims at finding a set of parameter vectors leading to acceptable model results. This requires in the first place an adequate definition of when to judge the model results as acceptable; in the second place it requires appropriate means to select which parameters should be calibrated, and last but not least it demands adequate methods to determine the requested parameter set.

The first task of defining when model results are considered (in)acceptable will certainly depend on the problem at hand: the modelling objectives and the quantity and quality of the available information and data play a major role in this.

Concerning the second task of deciding which parameters should be calibrated it is important to realize that a reduction of the number of calibrated parameters can be very beneficial for the success of calibration. Moreover, due to the scarcity and the inaccuracy of the data, often only a small number (e.g. $\leq 6 - 10$) of parameters can be handled satisfactorily. Such a reduction can be established by performing appropriate sensitivity analyses to detect the most important parameters which should be calibrated. In addition to the conventional type of sensitivity analysis in the software package UNCSAM (Janssen, Heuberger, Sanders [1992]), recently two alternative approaches for sensitivity analysis in a set-theoretic context have been implemented (Janssen [1995b,1995c]).

For the third task of determining the set of acceptable parameters, the use of the rotated-random-scan method was proposed in this report. The rotated-random-scan method is an iterative Monte Carlo search procedure inspired by Keesman and van Straten [1988,1989], Keesman [1989,1990]. It updates the currently available subset of acceptable parameter vectors by iteratively applying rotations or transformations of the parameter space, followed by a uniform random scan in the transformed space. In this way one hopes to gradually and efficiently zoom in on the acceptable parameter set, and to determine an appropriate representative subsample.

The rotated-random-scan procedure appears to be a useful and simple tool for set-valued calibration, and is moreover applicable in a preliminary stage of conventional calibration studies to delimit the prior uncertainty range of parameters to be calibrated. The method, which is available as software (Janssen [1995a]), has been tested on various artificial examples (Janssen and Sanders [1995]), and has been applied for the calibration of the soil-acidification model SMART. The tests have clarified the properties of the proposed procedure, and have led to useful recommendations in making the various user's choices. The method appears to work well in characterizing or covering acceptable parameter sets which are connected and regularly shaped (e.g. box and ellipsoidal shaped sets). The uniform random scan which is used in the initial step determines for a large part the efficiency of the method. For non-connected or irregular shaped sets the performance decreases, since the procedure focusses basically on linear relationships and therefore possibly fails to identify important non-linear dependences. This situation can be improved by reparametrization, or by using more sophisticated set-covering methods. The method proposed recently in Klepper and Hendrix [1994a,b] can be a useful alternative in this context.

The reported case study on SMART showed that the calibration results were somewhat disap-

pointing: only one parameter could be determined rather accurately, while the uncertainty reduction in most other parameters is rather marginal. The resulting uncertainty in the model outputs appears to be fairly large, but the measurements still do not fall completely within these bands, indicating that the model with the employed input data and the associated parametrization is not fully able to describe the data. A more detailed calibration study will be needed to analyse whether this situation can be improved.

References

- Belforte G., Bona B., Cerone V. [1990a]. Parameter estimation algorithms for a set-membership description of uncertainty. *Automatica*, Vol. 26, pp. 887-898.
- Belforte G., Bona B., Cerone V. [1990b]. Identification, structure selection and validation of uncertain models with set-membership error description. *Mathematics and Computers in Simulation*, Vol. 32, pp. 561-569.
- Belmans C., Wesseling J.G., Feddes R.A. [1983]. Simulation model of the water balance of a cropped soil providing different types of boundary conditions: SWATRE. *Journal of Hydrology*, Vol. 63, pp. 27-286.
- Bredemeier M., Tiktak A., Van Heerden K. [1995]. The Solling Norway Spruce site. *Ecological Modelling*, Vol. 83, pp. 7-15.
- De Vries, W., Posch M., Kämäri, J. [1989]. Simulation of the long-term soil response to acid deposition in various buffer ranges. *Water, Air and Soil Pollution*, Vol. 48, pp. 349-390.
- Fedra K. [1983]. *Environmental Modeling Under Uncertainty: Monte Carlo Simulation*. IIASA, Research Report RR-83-28, Laxenburg, Austria.
- Fedra K., Straten van G., Beck M.B. [1981]. Uncertainty and arbitrariness in ecosystems modelling: A lake modelling example. *Ecological Modelling*, Vol. 13, pp. 87-110.
- Golub G.H., van Loan C.F. [1989]. *Matrix Computations*. Second Edition. The John Hopkins University Press, Baltimore, Maryland.
- Heuberger P.S.C., Sanders R., Janssen P.H.M. [1992]. *Calibration for regionalisation studies. Two methods for the estimation of distribution functions on basis of Monte Carlo sampling* (in Dutch). RIVM-report nr. 723001008, Bilthoven, the Netherlands.
- Hornberger G.M., Cosby B.J. [1985] Selection of parameter values in environmental models using sparse data: a case study. *Appl. Math. Comput. Model. Environ.*, Vol. 17, pp. 335-355.
- Hornberger G.M., Spear R.C. [1980]. Eutrophication in Peel Inlet - I. The problem-defining behaviour and a mathematical model for the phosphorus scenario. *Water Research*, Vol. 14. pp. 29-42.
- Hornberger G.M., Spear R.C. [1981]. An approach to the preliminary analysis of environmental systems. *J. Environ. Management*, Vol. 12, pp. 7-18.
- Hornberger G.M., Spear R.C. [1983]. An approach to the analysis of behaviour and sensitivity on environmental systems. In *Uncertainty and Forecasting of Water Quality*; edit. M.B. Beck en G. van Straten, pp. 101-116. Springer Verlag, New York.
- Iman R.L., Conover W.J. [1980]. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. *Communications in Statistics*, Vol. A9, pp. 1749-1842. Rejoinder to comments, *ibid.* pp. 1863-1974.
- Iman R.L., Conover W.J. [1982]. A distribution free approach to inducing rank correlations among input variables. *Communications in Statistics*, Vol. B 11, pp. 311-334.
- Janssen P.H.M. [1995a]. *RORASC: software for the rotated-random-scan calibration procedure*. RIVM report nr. 733001005, RIVM, Bilthoven, the Netherlands.
- Janssen P.H.M. [1995b]. *GENSEN: a program for generalized sensitivity analysis*. RIVM report nr. 733001006, RIVM, Bilthoven, the Netherlands.

- Janssen P.H.M. [1995c]. *DOMDIR: a program to determine the dominant directions in the parameter space*. RIVM report nr. 733001007, RIVM, Bilthoven, the Netherlands.
- Janssen P.H.M., Heuberger, P.S.C. [1992]. *The role of sensitivity analysis and identifiability analysis in model calibration*. (in Dutch). RIVM report nr. 723001007, RIVM, Bilthoven, The Netherlands.
- Janssen P.H.M., Heuberger, P.S.C. [1995]. Calibration of process-oriented models. *Ecological Modelling*, Vol. 83, pp. 55-66.
- Janssen P.H.M., Heuberger, P.S.C., Sanders R. [1992]. *UNCSAM 1.1: a Software Package for Sensitivity and Uncertainty Analysis; Manual*. RIVM report nr. 959101004, RIVM, Bilthoven, the Netherlands.
- Janssen P.H.M., Heuberger, P.S.C., Sanders R. [1994]. UNCSAM: a tool for automating sensitivity and uncertainty analysis. *Environmental Software*, Vol. 9, pp. 1-11.
- Janssen P.H.M., Sanders R. [1995]. *Tests of methods and software for set-valued model calibration and sensitivity analyses*. RIVM report nr. 733001008, RIVM, Bilthoven, the Netherlands.
- Keesman K.J. [1989a]. *A set-membership approach to the identification and prediction of ill-defined systems: Application to a water quality system*. Ph.D. Thesis, Univ. of Twente, Enschede, The Netherlands.
- Keesman K.J. [1989b]. On the dominance of parameters in structural models of ill-defined systems. *Applied Math. Comput.*, Vol. 30, pp. 133-147.
- Keesman K.J. [1990]. Membership-set estimation using random scanning and principal component analysis. *Mathematics and Computers in Simulation*. Vol. 32, pp. 535-543.
- Keesman K., Straten van G. [1987]. Modified set-theoretic identification of ill-defined water quality system from poor data. In *Proceedings of IAWPRC Symposium on Systems Analysis in Water Quality Management*, pp. 297-308, Pergamon, New York.
- Keesman K., Straten van G. [1988]. Embedding of random scanning and principal component analysis in set-theoretic approach to parameter estimation. *Proc. 12th IMACS World Congress*, Paris.
- Keesman K., Straten van G. [1989]. Identification and prediction propagation of uncertainty in models with bounded noise. *Intern. Journal of Control*, Vol. 49, pp. 2259-2269.
- Keesman K., Straten van G. [1990]. Set membership approach to identification and prediction of lake eutrophication. *Water Resources Research*, Vol. 26, pp. 2643-2652.
- Klepper O. [1988]. *A model of carbon flows in relation to macrobenthic food supply in the Oosterschelde estuary (S.W. Netherlands)*. Ph.D dissertation, Agricultural University Wageningen, The Netherlands.
- Klepper O., Hendrix E.M.T. [1994a]. A comparison of algorithms for global characterization of confidence regions for nonlinear models. *J. Environ. Toxicol. Chem.*, Vol. 13, pp. 1887-1899.
- Klepper O., Hendrix E.M.T. [1994b]. A method for robust calibration of ecological models under different types of uncertainty. *Ecological Modelling*. Vol. 74, pp. 161-182.
- Klepper O., Rouse D.I. [1991]. A procedure to reduce parameter uncertainty for complex models by comparison with real system output illustrated on a potato growth model. *Agricultural Systems*, Vo. 36, pp. 375-395.

- Kros J., De Vries W., Janssen P., Bak C. [1993]. The uncertainty in forecasting regional trends of forest soil acidification. *WATER, AIR AND SOIL POLLUTION*, Vol. 66, pp. 29-58.
- Lence B. J., Takyi A.K. [1992]. Data requirements for seasonal discharge programs: an application of a regionalised sensitivity analysis. *Water Resources Research*, Vo. 28, pp. 1781-1789.
- McKay M.D., Beckman R.J., Conover W.J. [1979]. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, Vol. 21, pp. 239-245.
- Milanese M., Vicino A. [1991a]. Estimation theory for nonlinear models and set membership uncertainty. *Automatica*, Vol. 27, pp. 403-408.
- Milanese M., Vicino A. [1991b]. Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*, Vol. 27, pp. 997-1009.
- Morrison D.F. [1984]. *Multivariate Statistical Methods*. 2-nd edition. Mc-Graw Hill, Singapore.
- Posch M., Reinds G.J., De Vries W., [1993]. *SMART - a simulation model for acidification's regional trends: model description and user manual*. Mimeograph Series of the National Board of Waters and the Environment, nr. 477, Helsinki, Finland.
- Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T. [1986]. *Numerical Recipes: the Art of Scientific Computing*. Cambridge University Press.
- Spear R.C., Hornberger G.M. [1980]. Eutrophication in Peel Inlet-II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, Vol. 14, pp. 43-49.
- Tiktak A., Bredemeier M., Van Heerden K. [1995]. The Solling dataset: Site characteristics, monitoring data and deposition scenarios. *Ecological Modelling*, Vol. 83, pp. 17-34.
- Van der Salm C., Kros J., Groenenberg J.E., De Vries W., Reinds G.J. [1995]. Application of soil acidification models with different degrees of process description on an intensively monitored spruce site. To appear in: *Solute Modelling in Catchment Systems*. Ed. S.T. Trudgill.
- Van Grinsven H.J.M., Driscoll C.T., Tiktak A. [1995]. Workshop on Comparison of Forest-Soil-Atmosphere Models: Preface. *Ecological Modelling*, Vol. 83, pp. 1-6.
- Walter E., Piet-Lahanier H. [1988]. Estimation of the parameter uncertainty resulting from bounded-error data. *Mathematical Biosciences*, Vol. 92, pp. 55-74.
- Walter E. [1990]. (editor) Special Issue on Parameter Identifications with Error Bound. *Mathematics and Computers in Simulation*, Vol. 32, no. 5 and 6, pp. 447-607.
- Walter E., Pronzato L. [1994]. Characterizing sets defined by inequalities. In Vol. 2 of the preprints of the *SYSID'94: 10th IFAC Symposium on System Identification*, eds. M. Blanke, T. Söderström. 4-6 July 1994, Copenhagen. pp. 15-26.
- Whitehead P.G., Hornberger G.M. [1984]. Modelling algal behaviour in the river Thames. *Water Res.*, Vol. 18, pp. 945-953.

Appendix A: Outline of the Rotated-Random-Scan procedure

In this appendix the various steps are presented which constitute the rotated-random-scan procedure. Application of this procedure requires a preliminary specification of the numbers $N_{\text{ini}}, N_{\text{can}}, N_{\text{lim}}$, indicating respectively the number of samples to be generated in the initialisation step, the number of candidate samples to be generated in each iteration, and the total number of acceptable samples which one minimally likes to obtain. The various steps of the procedure are now as follows:

Step 0. Initialization:

Determine N_{ini} samples by sampling the parameters *uniformly* between their lower- and upper bounds. Use can be made of the simple random sampling or of the efficient Latin Hypercube Sampling technique (Mc Kay et al. [1979]; Iman and Conover [1980]). Moreover the correction technique of Iman and Conover [1982] can be used to remove spurious correlations in the sampled values. The software package UNCSAM can be applied for this initial sampling (cf. Janssen, Heuberger, Sanders [1992]).

Subsequently, by simulating the model with the sampled parameters, and by comparing the results with the measurements, it is determined which of the N_{ini} sampled parameter combinations are accepted. These parameter vectors are stored in the set Θ_{acc} . N_{acc} denotes the number of currently obtained acceptable parameter vectors.

Step 1. Preparation:

Consider the (updated) set Θ_{acc} of acceptable parameter vectors obtained thus far, and check first whether one has already obtained sufficiently many acceptable parameter vectors, i.e. [”Is $N_{\text{acc}} \geq N_{\text{lim}}$?”], where N_{lim} denotes the number of acceptable parameter vectors which one minimally likes to obtain. If the answer is affirmative, the search can be completed. If not, the (sample) *covariance* matrix Σ^+ of the accepted parameters in the *original* parameter space (θ -space, where $\theta \in \mathbb{R}^p$) should be determined:

$$\Sigma^+ := \sum_{i=1}^{N_{\text{acc}}} \frac{(\theta^+(i) - \bar{\theta}^+)(\theta^+(i) - \bar{\theta}^+)^T}{N_{\text{acc}} - 1} \quad (\text{A.1})$$

$\theta^+(i)$ denotes the i -th acceptable parameter vector; $\bar{\theta}^+$ denotes the sample mean of the acceptable parameter vectors. It is assumed that $N_{\text{acc}} \geq p + 1$ to ensure that Σ^+ is invertible.

Step 2. Scaling:

Scaling is applied primarily to prevent unnecessary numerical errors (rounding/truncation) due to the possible strongly varying sizes of the parameter values. Moreover it can serve to focus attention on certain aspects of the parameter range.

First an appropriate *scaling*-matrix Λ_s for scaling the *centralized*²⁸ parameters $(\theta^+(i) - \bar{\theta}^+)$, is determined. Λ_s is a diagonal matrix with positive diagonal entries which depend on the user-specified *scaling option* (see below). Subsequently the covariance matrix of the *scaled acceptable* centralized parameters $(\theta_{s,c}^+ = \Lambda_s \cdot (\theta^+ - \bar{\theta}^+))$ is determined:

$$\Sigma_s^+ := \Lambda_s \cdot \Sigma^+ \cdot \Lambda_s \quad (\text{A.2})$$

Various options are available for scaling the parameters:

²⁸Centralization is applied also due to numerical reasons, to avoid unnecessary accuracy loss due to truncation and round-off.

- (1) No scaling is used ($\Lambda_s = I$).
- (2) The acceptable parameters are scaled by their *standard deviations* ($\Lambda_s = \text{Diag}([\sigma_i^+]^{-1})$, where $[\sigma_i^+]$ denotes the standard deviation of the i -th component of the set of *acceptable* parameters, i.e. $\sigma_i^+ := \sqrt{[\Sigma^+]_{ii}}$.
This means that the covariance matrix Σ_s^+ of the scaled parameters is equal to the *correlation matrix* of the acceptable parameters. In this situation, the procedure in fact amounts to applying *principal component analysis* (see Morrison [1984]).
- (3) Scaling by the ranges of the *acceptable* parameters in the original parameter space ($\Lambda_s = \text{Diag}([\theta_{i,max}^+ - \theta_{i,min}^+]^{-1})$, where $\theta_{i,min}^+$ and $\theta_{i,max}^+$ denote the minimal and maximal value of the i -th component of the *acceptable* sampled parameter vectors obtained so far.
- (4) Scaling by the (user-specified) ranges of *all* parameters in the original parameter space ($\Lambda_s = \text{Diag}([\theta_{i,max} - \theta_{i,min}]^{-1})$, where $\theta_{i,min}$ and $\theta_{i,max}$ denote the user-specified minimal and maximal value of the i -th component of all sampled parameter vectors.

The scaling affects the subsequent transformation or rotation of the parameter space. In Janssen and Sanders [1995] it is studied how the applied scaling options influence the final results. See also the theoretical results in appendix B.

Step 3. Determination of the transformation:

The scaled (centralized) parameter vectors typically show mutual correlations (i.e. the covariance matrix Σ^+ is not a diagonal matrix). By applying a suitable transformation in the scaled (centralized) parameter space the resulting transformed parameters become uncorrelated. This increases the efficiency of a further exploration of the (transformed scaled) parameter space by the uniform random scanning in the subsequent step.

The desired transformation which renders uncorrelated parameter vectors can be determined in *two* ways (it is assumed in the sequel that the matrix Σ_s^+ is invertible, i.e. there do not exist any (non-trivial) linear relationships between the components of the acceptable parameter vectors):

- (a) Using an *eigen-system* decomposition²⁹ of the matrix Σ_s^+ :

$$\Sigma_s^+ := U \cdot \Lambda \cdot U^T \quad (\text{A.3})$$

where U is the orthogonal matrix³⁰ of eigenvectors, and Λ is a diagonal matrix whose diagonal entries consist of the (non-negative) eigenvalues of Σ_s^+ . The corresponding transformed (scaled) centralized parameter vector is defined by:

$$\tilde{\theta} := U^T \cdot \theta_{s,c} = U^T \Lambda_s \cdot (\theta - \bar{\theta}^+) \quad (\text{A.4})$$

Hence the associated transformation matrix T_c , which relates θ to $\tilde{\theta}$ (i.e. $\tilde{\theta} = T_c \cdot (\theta - \bar{\theta}^+)$), is equal to

$$T_c := U^T \Lambda_s \quad (\text{A.5})$$

The transformed parameter space ($\tilde{\theta}$ -space) is thus obtained by rotating the coordinate axes in the scaled centralized parameter space ($\theta_{s,c}$ -space) according to the rotation matrix U^T . Notice that by applying the above compounded transformation matrix T_c , the corresponding covariance matrix of the transformed acceptable

²⁹If A denotes a $n \times m$ matrix, then A^T denotes its transpose ($m \times n$ matrix).

³⁰A $n \times n$ real matrix W is called *orthogonal* if $W \cdot W^T = I_n$.

(centralized) parameters $\tilde{\theta}^+$ has become equal to the diagonal matrix Λ (Note: the sample mean $\bar{\theta}^+$ is 0, due to centralization):

$$\begin{aligned}\tilde{\Sigma}^+ &:= \sum_{i=1}^{N_{\text{acc}}} \frac{(\tilde{\theta}^+(i) - \bar{\theta}^+)(\tilde{\theta}^+(i) - \bar{\theta}^+)^T}{N_{\text{acc}} - 1} \\ &= U^T \cdot \Sigma_s^+ \cdot U = U^T \cdot U \cdot \Lambda \cdot U^T \cdot U = \Lambda\end{aligned}\quad (\text{A.6})$$

which means that the transformed acceptable parameters $\tilde{\theta}^+(\cdot)$ are *uncorrelated*.

(b) Using the *Cholesky* decomposition³¹ of the $p \times p$ -matrix Σ_s^+ , see Golub and van Loan [1989]:

$$\Sigma_s^+ := L \cdot L^T \quad (\text{A.7})$$

where L is a lower triangular matrix with positive diagonal elements. The transformed parameter³² is defined by:

$$\tilde{\theta} := L^{-1} \cdot \theta_{s,c} = L^{-1} \Lambda_s \cdot (\theta - \bar{\theta}^+) \quad (\text{A.8})$$

Hence the (currently) associated transformation matrix T_c , which relates θ to $\tilde{\theta}$ ($\tilde{\theta} = T_c \cdot (\theta - \bar{\theta}^+)$), is equal to

$$T_c := L^{-1} \Lambda_s \quad (\text{A.9})$$

It is clear that the associated covariance matrix of the transformed acceptable parameters (Note: the sample mean $\bar{\theta}^+$ is 0, due to centralization):

$$\begin{aligned}\tilde{\Sigma}^+ &:= \sum_{i=1}^{N_{\text{acc}}} \frac{(\tilde{\theta}^+(i) - \bar{\theta}^+)(\tilde{\theta}^+(i) - \bar{\theta}^+)^T}{N_{\text{acc}} - 1} \\ &= L^{-1} \cdot \Sigma_s^+ \cdot L^{-T} = L^{-1} \cdot L \cdot L^T \cdot L^{-T} = I\end{aligned}\quad (\text{A.10})$$

is equal to the identity matrix. This means that the transformed acceptable parameters $\tilde{\theta}^+$ are *uncorrelated and have unit variance*.

Although the use of the Cholesky-decomposition is preferable to the use of the eigen-system decomposition with regard to computational effort, it will reveal less clear information on the correlation structure of the acceptable parameters (e.g. the eigen-system decomposition clearly indicates the dominant directions in the parameter space). Notice moreover that application of the eigensystem decomposition results in a *rotation* in the scaled (centralized) parameter space, but application of the Cholesky decomposition does not (L^{-1} is not a rotation matrix). In Janssen and Sanders [1995] the influence of the chosen decomposition option is studied in more detail. See also appendix B.

Step 4. Uniform sampling in the transformed parameter space ($\tilde{\theta}$ -space):

Subsequently new samples (say N_{can}) are generated in the transformed parameter space by sampling each transformed centralized parameter component $\tilde{\theta}_i$ *uniformly* between the bounds $[\tilde{\theta}_{i,\min}^*, \tilde{\theta}_{i,\max}^*]$, which are determined by expanding the minimum and maximum of the transformed accepted parameter components $[\tilde{\theta}_{i,\min}^+, \tilde{\theta}_{i,\max}^+]$ according to:

$$\tilde{\theta}_{i,\min}^* := \tilde{\theta}_{i,\min}^+ - \beta \cdot (\tilde{\theta}_{i,\max}^+ - \tilde{\theta}_{i,\min}^+) \quad (\text{A.11})$$

$$\tilde{\theta}_{i,\max}^* := \tilde{\theta}_{i,\max}^+ + \beta \cdot (\tilde{\theta}_{i,\max}^+ - \tilde{\theta}_{i,\min}^+) \quad (\text{A.12})$$

³¹A similar transformation matrix can be obtained *directly* (i.e. without the need for computing and decomposing the covariance matrix Σ_s^+) on basis of a *QR*-decomposition on a signal matrix of the scaled acceptable parameters. We will not elaborate on this.

³²If Σ_s^+ is invertible, the lower triangular matrix L will also be invertible. Its inverse can be easily determined.

where the expansion factor β is defined by:

$$\beta = \frac{1}{(N_{\text{acc}} - 1)} \quad (\text{A.13})$$

This expansion is used because the sampled parameter points obtained thusfar do not exactly reflect the bounds of the parameter space, since they have been obtained by a finitely sized random sample. The specific choice of the expansion factor β is based on the fact that

$$\hat{z}_{\text{max}} = z_{\text{max}} + \frac{(z_{\text{max}} - z_{\text{min}})}{N - 1} \quad (\text{A.14})$$

can be used as an unbiased estimate³³ of the upper-bound of a *uniform* sample z_1, \dots, z_N of N points.

The above sketched uniform sampling can e.g. be performed by simple *random sampling* or by the more efficient *Latin Hypercube Sampling* technique (Mc Kay et al. [1979]; Iman and Conover [1980]). Moreover the correction technique of Iman and Conover [1982] can be applied to remove spurious correlations which are introduced due to the finite number of samples taken. It is expected that Latin Hypercube sampling generates a more even covering of the transformed parameter space than simple random sampling, and that the correction technique of Iman and Conover is beneficial in obtaining (near) uncorrelatedness. See also Janssen and Sanders [1995].

Step 5. Back-transformation to the original parameter space:

The N_{can} newly sampled transformed parameters are subsequently back-transformed to the original parameter space by using the transformation:

$$\theta := \bar{\theta}^+ + T_c^{-1} \cdot \tilde{\theta} \quad (\text{A.15})$$

where T_c is the transformation-matrix associated to the eigen-system decomposition or the Cholesky decomposition (see formulae (A.3), (A.7)). The thus obtained set of parameters Θ_{can} will serve as a new set of candidate parameters for subsequent model simulations.

Step 6. Simulation with candidate parameters in Θ_{can} :

Each parameter vector in the candidate parameter set Θ_{can} , obtained in the previous step is used to *simulate* the model. On basis of the simulation results it is subsequently determined which parameter vectors are acceptable and which are not. Finally one returns to step 1 of the sketched procedure.

Remark 3: In the above sketched procedure the desired transformation/rotation is always based on the (scaled) covariance matrix of the currently available acceptable parameters in the *original* parameter space (i.e. using the original coordinates).

An alternative approach would be to express the covariance matrix in terms of the coordinates in the current *transformed* parameter space instead, and to base the transformation/rotation on the decomposition of this matrix. This has the following consequences for the steps of the procedure:

- **Step 0:** The current transformation matrix T_c is initially equal to the identity matrix, and the initial bounds in the ‘transformed’ parameter space are set equal to the specified upper and lower bounds on the original parameters.

³³Reason for the unbiasedness is the fact that the order statistics z_{min} and z_{max} of N samples of a uniform distribution on the interval $[a, b]$ have expectation $a + \frac{b-a}{N+1}$ and $b - \frac{b-a}{N+1}$ respectively.

- **Step 1:** Assuming that the current transformation matrix T_c , which has been obtained during the previous iterations, is available, the covariance matrix of the acceptable parameters in the *transformed* coordinates ($\theta_{tr} := T_c \cdot \theta$) is given by:

$$\Sigma_{tr}^+ := T_c \cdot \Sigma^+ \cdot T_c^T \quad (\text{A.16})$$

- **Step 2:** Scaling will be based on the standard deviations or the ranges/bounds of the (acceptable) parameters in the *transformed* coordinates. Denoting the diagonal scaling matrix as $\Lambda_{s,tr}$, the scaled transformed covariance matrix will be equal to:

$$\Sigma_{s,tr}^+ := \Lambda_{s,tr} \cdot \Sigma_{tr}^+ \cdot \Lambda_{s,tr} \quad (\text{A.17})$$

Notice that this matrix is the sample covariance matrix of the acceptable parameters in the scaled transformed (centralized) parameter space ($\theta_{s,c,tr}$ -space; $\theta_{s,c,tr} = \Lambda_{s,tr} \cdot (\theta_{tr} - \bar{\theta}_{tr}^+)$).

- **Step 3:** Determination of the transformation is based on the eigenvector-eigenvalue decomposition or the Cholesky decomposition of the matrix $\Sigma_{s,tr}^+$. The resulting transformed parameter will be

- (a) when using the eigenvector-eigenvalue decomposition ($\Sigma_{s,tr}^+ = U \Lambda U^T$):

$$\tilde{\theta}_{tr} := U^T \cdot \theta_{s,c,tr} = U^T \cdot \Lambda_{s,tr} \cdot (\theta_{tr} - \bar{\theta}_{tr}^+) = U^T \cdot \Lambda_{s,tr} \cdot T_c \cdot (\theta - \bar{\theta}^+) \quad (\text{A.18})$$

This leads to an update of the current transformation matrix T_c , according to:

$$T_c^+ := U^T \cdot \Lambda_{s,tr} \cdot T_c \quad (\text{A.19})$$

- (b) when using the Cholesky decomposition ($\Sigma_{s,tr}^+ = L \cdot L^T$):

$$\tilde{\theta}_{tr} := L^{-1} \cdot \theta_{s,c,tr} = L^{-1} \cdot \Lambda_{s,tr} \cdot (\theta_{tr} - \bar{\theta}_{tr}^+) = L^{-1} \cdot \Lambda_{s,tr} \cdot T_c \cdot (\theta - \bar{\theta}^+) \quad (\text{A.20})$$

This leads to an update of the current transformation matrix T_c , according to:

$$T_c^+ := L^{-1} \cdot \Lambda_{s,tr} \cdot T_c \quad (\text{A.21})$$

- **Step 4:** Subsequently the uniform sampling is performed in the transformed parameter space ($\tilde{\theta}_{tr}$ -space), according to similar bounds as used in (A.11) and (A.12).
- **Step 5:** Back-transformation to the original parameter is performed by:

$$\theta := \bar{\theta}^+ + [T_c^+]^{-1} \tilde{\theta}_{tr} \quad (\text{A.22})$$

where $[T_c^+]^{-1}$ is the inverse of the updated transformation matrix in (A.19) or (A.21).

- **Step 6** remains the same.

In Janssen and Sanders [1995] it is studied by means of test-examples how the use of this alternative approach will affect the results. Their findings are supported by the theoretical results in the subsequent appendix B. ■

The rotated-random-scan procedure presented above, is more general than the procedure proposed in Keesman [1989,1990], Keesman and van Straten [1988,1989]. It offers more choices³⁴ with respect to scaling, transformation determination, sampling technique etc. Moreover the choice of the expansion factor β in eqn. (A.13) differs slightly from Keesman's choice.

³⁴Keesman considers scaling based on standard deviations, and applies the eigensystem decomposition. Uniform sampling of new candidates is performed by means of ordinary Monte Carlo sampling.

Appendix B: Theoretical considerations on the influence of working space, scaling and decomposition

In applying the rotated-random-scan procedure the user has to choose between various options, concerning:

1. **Working space:** Should one work in the original coordinates, or in the transformed coordinates?
2. **Scaling:** Should one apply no scaling to the (transformed) parameters, or should one apply scaling on basis of the *standard-deviations*, the ranges of the *acceptable* parameters or the ranges of *all* parameters?
3. **Decomposition:** Should one apply a transformation on basis of an *eigen-system* decomposition or a *Cholesky* decomposition of the scaled (transformed) covariance matrix?

In order to provide the user with guidelines for making these choices, various *numerical* experiments were performed in Janssen and Sanders [1995] to study the impact of the above options on the final results. In this appendix this issue is investigated *theoretically*, and it is tentatively indicated in what way the outcomes can differ due to these choices. It is however difficult to relate these differences directly to the various individual options; this will be very much case-dependent.

In order to ‘prove’ the above statement, two distinct ways (i.e. $j = 1, 2$) are considered of obtaining new candidate samples in a typical iteration of the rotated random scan procedure, depending on the user-specified options. Let Σ^+ denote the covariance matrix of the acceptable parameters in the current iteration, as expressed in the original coordinate system. Assume that the *transformed* parameter space $x_{tr}^{(j)} := T_c^{(j)} \cdot (x - \bar{x}^+)$ is subsequently considered as working-space, where \bar{x}^+ denotes the sample mean of the acceptable parameters in the original coordinates. The matrix $T_c^{(j)}$ will be equal to the identity matrix $I_{p \times p}$ if the *working space option* is equal to 1, i.e. working in the original coordinates; if this option is 2, i.e. working in the transformed coordinates, $T_c^{(j)}$ indicates the actual transformation matrix associated to these transformed coordinates.

Assume moreover that the scaling matrix $\Lambda_s^{(j)}$ is applied. The corresponding scaled transformed covariance matrix Ω_j (see the notation in appendix A) is then given by:

$$\Omega_j := \Lambda_s^{(j)} \cdot T_c^{(j)} \cdot \Sigma^+ \cdot [T_c^{(j)}]^T \cdot \Lambda_s^{(j)} \quad (j = 1, 2) \quad (\text{B.1})$$

Suppose finally that this matrix is decomposed according to

$$\Omega_j = V_j \cdot \Lambda_j \cdot V_j^T \quad (j = 1, 2) \quad (\text{B.2})$$

This decomposition refers to the eigensystem decomposition as well as to the Cholesky decomposition, depending on the chosen *decomposition option*. If the decomposition option is (1) (i.e. based on eigensystem decomposition) V_j denotes the orthogonal matrix of eigenvectors of Ω_j , and Λ_j is the diagonal matrix of eigenvalues. If the decomposition option is (2) (i.e. based on Cholesky decomposition) V_j is the lower triangular Cholesky factor, and Λ_j is the identity matrix $I_{p \times p}$.

During a typical iteration of the rotated-random-scan procedure, new coordinates are defined

on basis of the above-mentioned decomposition of the scaled transformed covariance matrix Ω_j (cf. the procedure described in appendix A):

$$\tilde{x}_j := \tilde{T}_j \cdot (x - \bar{x}^+) \quad (j = 1, 2) \quad (\text{B.3})$$

where the transformation matrix \tilde{T}_j is given by

$$\tilde{T}_j := V_j^{-1} \cdot \Lambda_s^{(j)} \cdot T_c^{(j)} \quad (j = 1, 2) \quad (\text{B.4})$$

From equation (B.1) it is easily established that the original covariance matrix Σ^+ of the acceptable parameters, will be equal to a diagonal matrix, when expressed in these new coordinates: i.e.

$$\tilde{T}_j \cdot \Sigma^+ \cdot [\tilde{T}_j]^T = \Lambda_j \quad (j = 1, 2) \quad (\text{B.5})$$

The transformed acceptable samples \tilde{x}_j^+ are thus uncorrelated, and therefore an uncorrelated uniform scan is performed in this transformed space to generate new candidate samples. This scan is determined by uniform sampling between the associated upper- and lower bounds of the transformed parameters; these upper- and lower bounds are computed according to the procedure in step (4) of the rotated-random-scan method (see formulae (A.11),(A.12)).

In order to establish relationships between the results obtained for $j = 1$ and $j = 2$, we normalize/standardize the candidate parameters by defining the vector \hat{x}_j by scaling \tilde{x}_j according to:

$$\hat{x}_j := [\Lambda_j]^{-\frac{1}{2}} \cdot \tilde{x}_j \quad (j = 1, 2) \quad (\text{B.6})$$

Notice that uncorrelated uniform sampling in the \tilde{x}_j -space is equivalent to uncorrelated uniform sampling in the \hat{x}_j -space. The relation between the original space and the \hat{x}_j -space is (use equation (B.3)):

$$\hat{x}_j = \hat{T}_j \cdot (x - \bar{x}^+) \quad (j = 1, 2) \quad (\text{B.7})$$

where \hat{T}_j is the associated transformation matrix:

$$\begin{aligned} \hat{T}_j &:= [\Lambda_j]^{-\frac{1}{2}} \cdot \tilde{T}_j \\ &= [\Lambda_j]^{-\frac{1}{2}} \cdot V_j^{-1} \cdot \Lambda_s^{(j)} \cdot T_c^{(j)} \quad (j = 1, 2) \end{aligned} \quad (\text{B.8})$$

As a result of this additional scaling the transformed acceptable parameters have a covariance matrix in the \hat{x}_j -space which is equal to the *identity* matrix, since

$$\hat{T}_j \cdot \Sigma^+ \cdot [\hat{T}_j]^T = I_{p \times p} \quad (j = 1, 2) \quad (\text{B.9})$$

From this equality it follows that:

$$[\hat{T}_1]^{-1} \cdot ([\hat{T}_1]^{-1})^T = [\hat{T}_2]^{-1} \cdot ([\hat{T}_2]^{-1})^T \quad (\text{B.10})$$

and thus

$$W := \hat{T}_2 \cdot [\hat{T}_1]^{-1} \quad (\text{B.11})$$

fulfils

$$W \cdot W^T = I_{p \times p} \quad (\text{B.12})$$

i.e. W is an orthogonal matrix. From this a simple relationship can be obtained between \hat{T}_1 and \hat{T}_2 :

$$\hat{T}_2 = W \cdot \hat{T}_1 \quad (\text{B.13})$$

for the orthogonal matrix W in (B.11). Consequently the \hat{x}_1 - and \hat{x}_2 -space are related as:

$$\hat{x}_2 = W \cdot \hat{x}_1 \quad (\text{B.14})$$

The previous relationships can now be used to indicate how the results of the rotated-random-scan procedure are affected by the chosen options. Suppose therefore that a specific set of currently available acceptable samples Θ_{cur} , with (sample) covariance matrix Σ^+ is taken as starting point. Depicting this set in the \hat{x}_1 -space and the \hat{x}_2 -space respectively, renders the sets $\Theta_{cur}^{(1)}$ and $\Theta_{cur}^{(2)}$, which are rotated versions of each other, due to relationship (B.14). Notice that the parameters in these sets have covariance matrix I . For ease of demonstration, we suppose that the acceptable set $\Theta_{cur}^{(1)}$ in the \hat{x}_1 -space is the shaded square $[-1, 1] \times [-1, 1]$ depicted in figure 11 (a), while $\Theta_{cur}^{(2)}$ is the same square, but rotated over 45 degrees (shaded diamond in figure 11 (b)).

According to step 4 of the rotated-scan procedure (see appendix A), first the (expanded) minima and maxima are determined of the transformed acceptable parameters in the \hat{x}_1 - and \hat{x}_2 -space. Subsequently, uniform sampling is performed between these (expanded) ranges; this results e.g. in the candidate samples $\hat{x}_{1,can}$ and $\hat{x}_{2,can}$ which lie in the enclosing squares indicated in figure 11 (a) and (b). In the example we have employed an expansion factor of 1.1. Notice that the candidate set in the space \hat{x}_2 is larger than in the space \hat{x}_1 .

These candidate samples are subsequently back-transformed into the original coordinates according to

$$x^{(1)} - \bar{x}^+ = \hat{T}_1^{-1} \cdot \hat{x}_{1,can} \quad (\text{B.15})$$

$$x^{(2)} - \bar{x}^+ = \hat{T}_2^{-1} \cdot \hat{x}_{2,can} \quad (\text{B.16})$$

Notice that (B.16) can be rewritten as (use B.13):

$$x^{(2)} - \bar{x}^+ = \hat{T}_1^{-1} [W^T \cdot \hat{x}_{2,can}] \quad (\text{B.17})$$

thus illustrating the relationship with $x^{(1)}$: $[x^{(2)} - \bar{x}^+]$ is obtained by first rotating $\hat{x}_{2,can}$ according to W^T (i.e. the inverse of the rotation characterized by W), and then back-transforming it according to \hat{T}_1^{-1} , while $[x^{(1)} - \bar{x}^+]$ is obtained by directly back-transforming $\hat{x}_{1,can}$ according to \hat{T}_1^{-1} . The difference between the thus obtained new candidates according to the first set of options (i.e. using $T_c^{(1)}, \Lambda_s^{(1)}, V_1, \Lambda_1$), and the second set of options (i.e. using $T_c^{(2)}, \Lambda_s^{(2)}, V_2, \Lambda_2$) is illustrated in figure 12, where the candidates $\hat{x}_{1,can}$ (dotted square) and the back-rotated $W^T \cdot \hat{x}_{2,can}$ (hatched diamond) are depicted in the \hat{x}_1 -space, i.e. before back-transformation to the original space by the matrix \hat{T}_1^{-1} . Notice that the different options render candidate-sets of different sizes and orientation in the \hat{x}_1 -space. It is however difficult to relate these differences directly to the specific choices of working-space option, scaling option and decomposition option, since the resulting matrices W and \hat{T}_1 are determined by these options in an indirect way. The experimental results in Janssen and Sanders [1995] indicate however that the influence of the options is not critical.

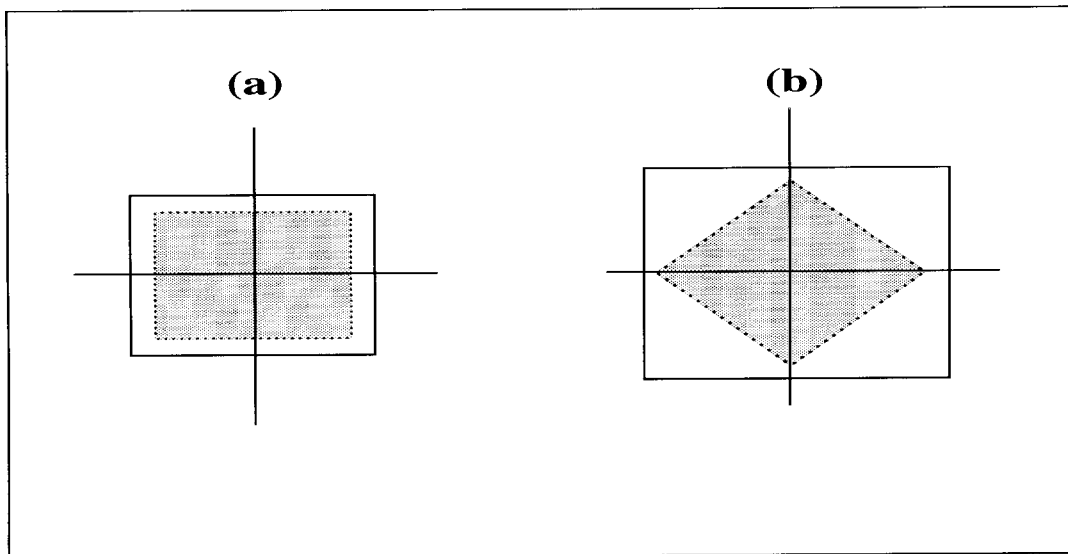


Figure 11: The set of currently available acceptable samples in the \hat{x}_1 - (a) and the \hat{x}_2 -space (b) (shaded figures), as well as the set of new (uniform, uncorrelated) candidate samples in these spaces (enclosing squares).

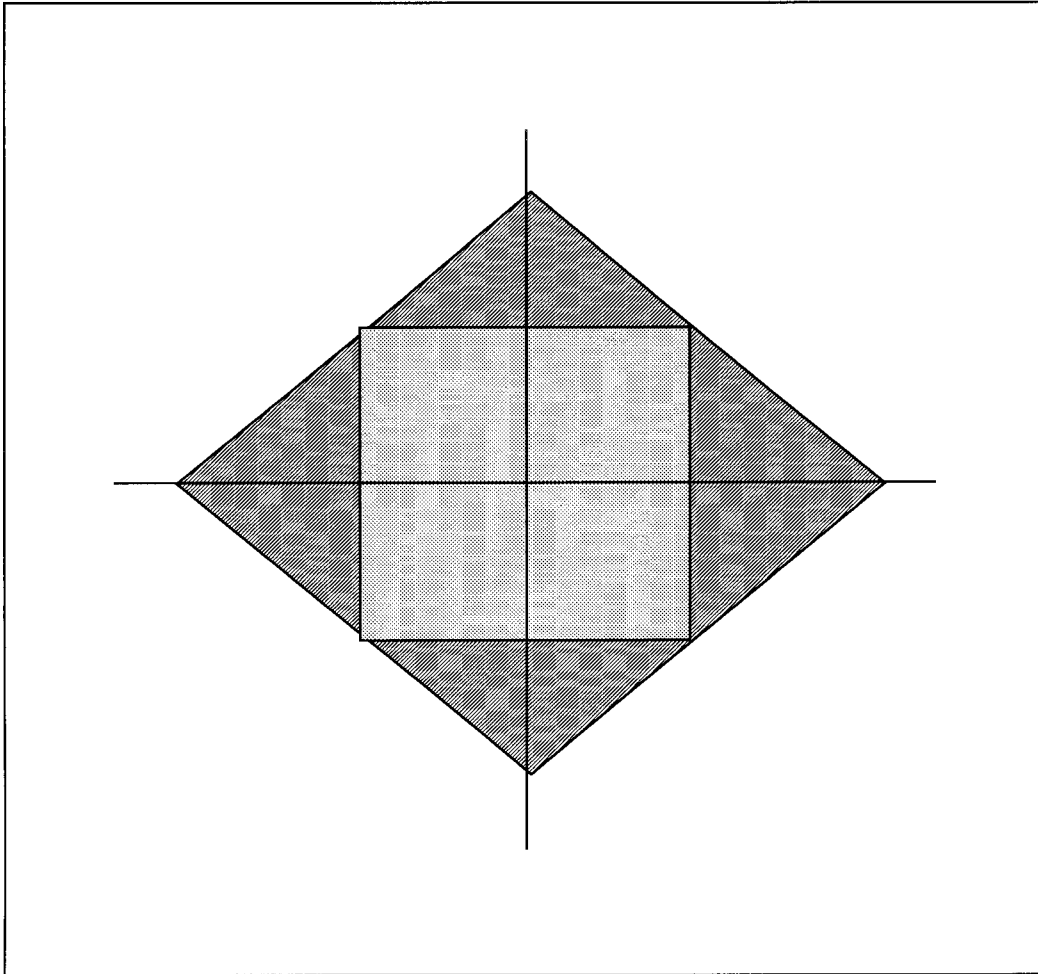


Figure 12: The sets of new candidate samples $\hat{x}_{1,can}$ (shaded square) and $W^T \cdot \hat{x}_{2,can}$ (hatched diamond) in the \hat{x}_1 -space, before back transformation to the original space.