On the Dominance of Parameters in Structural Models of Ill-Defined Systems

Karel Keesman

Department of Chemical Engineering
University of Twente
P.O. Box 217
7500 AE Enschede, The Netherlands

Transmitted by John Casti

ABSTRACT

For ill-defined systems (like environmental and economical systems) with sparse and uncertain data, Homberger and Spear have proposed, as a variation on Monte Carlo simulation analysis, the so-called regionalized sensitivity analysis (RSA) to determine dominant model parameters. In this paper a minor modification to the RSA is proposed to mitigate the effect of arbitrarily chosen initial parameter intervals. This modification concerns the application of a parameter space identification method prior to the RSA to offer more reliable (with respect to the observed systems' behavior) initial parameter intervals. An alternative procedure to the RSA is also proposed. Herein, the results from a parameter-space identification method are analyzed directly in terms of eigenvectors (principal axes) and eigenvalues characterizing the associated identified parametric subspace. These methods are applied to a simple water-quality model with sufficient as well as sparse data. Comparison with the results from period-average analysis and maximum-likelihood estimation reveals that consistent information is obtained.

1. INTRODUCTION

In setting up a model of an ill-defined system, information from observed system inputs and outputs will be combined with some prior hypothetical knowledge about the system under research. Within the modeling context, an ill-defined system is particularly characterized by a high degree of uncertainty in the model structure and sparse information in the observed system inputs and outputs [11]. In practice, it appears that an ultimate structural
model is found after passing through a number of loops, thereby adjusting the presumed model structure. It is therefore important to have information about the dominant processes and their related parameters, which appear to control the system behavior. That is, the model output that corresponds with the observed systems' behavior is sensitive to these dominant parameters. In addition, for complex models with a large number of parameters, parameter selection is necessary when a formal parameter estimation procedure is applied. It appears, in practice, that only a small number (\(< 6-10\)) of parameters can be handled satisfactorily.

Different techniques have been proposed to obtain this information about parameter dominance. The most conventional method is sensitivity analysis (see [9]), where sensitivity functions are specified analytically or represented by their numerical approximations. Sensitivity analysis approaches the problem from a deterministic point of view by perturbating the system through a small variation in a particular parameter and then evaluating its effect on some objective function (usually the state variables). It is recognized that in practice this method has some limitations due to violation of some implicit assumption. Such assumptions concern the unbiasedness of the nominal parameters, which are responsible for the expected model behavior; the independent contribution of each parameter; and the first-order effect of small perturbations [2].

Another method to determine the dominant parameters uses information from a parameter estimation process, where the parameters are still handled as deterministic variables. Herein the criterion function space in the neighborhood of the optimal parameter estimate is evaluated by means of Taylor series expansion (for instance, up to second-order terms). If the parameter estimate is an optimum of the criterion function, then the first-order term will diminish. So the region of interest in the criterion function space is approximated by an ellipsoid (see [1]). The eigenvectors (principal axes) and the associated eigenvalues of the Hessian, appearing in the second-order term, represent the orientation and shape of the ellipsoid. The largest axis, corresponding to the smallest eigenvalue of the Hessian, indicates the worst-determined direction in the parameter space.

Alternatively, one can approach the problem from a stochastic point of view by considering each parameter as a random variable. Within this context the covariance matrix of the estimation error can be used to evaluate the dominance of parameters or linear combinations of parameters. Because of the close relationship between the covariance matrix and the Hessian in view of the Taylor series expansion of the criterion function in the neighborhood of the optimum for a wide class of estimates, it is evident that similar (principal-component) analyses can be performed to detect dominant parameters.
In the stochastic setting Monte Carlo simulation analysis is also applied. Unlike the sensitivity analysis, where the parameters are usually varied independently in a deterministic way, in a Monte Carlo approach the parameters are selected randomly from predefined frequency distributions. This information can be provided by the parameter estimation process, but an exact formulation of the frequency distribution can only be obtained under very severe assumptions. It turns out then that most of the time the distributions are chosen by engineering judgement. It is worth noting that in the Monte Carlo approach higher-order effects are not ignored. If, however, a linear relationship between objective function (state variables) and parameters is valid, then the estimated derivatives of the objective function with respect to the parameters are analogous to the sensitivity coefficients of the aforementioned sensitivity analysis. If on the other hand higher-order effects cannot be ignored, correlations between the objective function and the parameters reveal sensitivities in a more reasonable way (see [2]).

Hornberger and Spear [4, 5] stressed the fact that traditional parameter sensitivity analysis pertains to a particular point in the parameter space, i.e. a nominal parameter vector from a priori information of the system or an optimal (a posteriori) parameter estimate. It is evident that the choice of such a specific parameter vector in models of ill-defined systems is questionable, because of lack of knowledge about the model parameters. On the other hand, sparse data will preclude optimal parameter estimation.

Under these conditions Hornberger and Spear [4, 5, 8] proposed, as a variation on the Monte Carlo simulation analysis, a so-called regionalized sensitivity analysis (RSA) to identify dominant parameters in an early stage of the modeling process. Herein a particular point is replaced by predefined individual parameter distributions, from which randomly parameter vectors are selected. Also, a “nominal” trajectory of the state variable is replaced by constraint conditions in the observation space. A criterion function is defined then in terms of a classification algorithm, which selects well- or ill-behaved simulation runs. This classification ultimately results in two classes of behavior-giving and non-behavior-giving parameter vectors, respectively. Statistical procedures are then employed to indicate whether a single parameter or a linear combination of parameters is dominant. RSA is an attractive procedure because it is robust, but the results do depend on the a priori defined distributions. Alteration of a distribution can significantly influence the dominance ranking of the parameters [8].

This paper offers a minor modification of RSA to mitigate the effect of the predefined distributions. An alternative procedure is developed wherein a posteriori set-theoretic parametric uncertainty is evaluated to identify dominant parameters. Both procedures will be demonstrated by application to a simple (for illustration only) water-quality model with sufficient data—so
that the results can also be compared with information resulting from a (more conventional) maximum-likelihood estimation procedure—and with sparse data.

2. REGIONALIZED SENSITIVITY ANALYSIS ACCORDING TO HORNBERGER AND SPEAR

The regionalized sensitivity analysis procedure consists of the following steps:

1. Select an appropriate model structure.
2. Specify parameter distributions from data, literature, and experiments. (If only parameter ranges are available, choose a uniform distribution.)
3. Define the systems’ behavior in terms of constraint conditions.
4. Perform a number of Monte Carlo simulations on the basis of the predefined parameter distributions.
5. Classify the simulation run and the associated parameter vector as behavior- or non-behavior-giving.
6. Analyze the resulting sample cumulative distribution functions of the behavior- and non-behavior-giving classes, which will result in a sensitivity ranking of the parameters.

The last step will be considered in more detail. The sensitivity ranking is based on utilizing the Kolmogorov-Smirnov statistic $d_{m,n}$, which represents the maximum vertical distance between the two sample distribution functions. The indices $m$ and $n$ denote the classes of behavior- and non-behavior-giving parameter vectors, respectively. The key idea is that large values of $d_{m,n}$ indicate dominant parameter vectors with respect to the behavioral classification. To decide whether the separation is significant this value of $d_{m,n}$ is compared with a 90 or 99% confidence bound value. The behavioral classification will not only result in a separation along the original axes, but can also result in model induced covariance. However, this classification is disturbed by mean shifts and induced covariance resulting from the Monte Carlo method, because only a finite number of parameter vectors are sampled. To avoid wrong interpretations of the results, the sampled vectors are normalized. The value of $d_{m,n}$ is then determined for the normalized set of samples, which will not differ too much from the original set, when the number of simulations $(m + n)$ is chosen to be large enough.

After having removed the Monte Carlo induced covariances and mean shifts by application of the normalization procedure, there can still be
correlations left between the parameters. As was noted before, these parameter correlations originate then from model induced covariances. Hornberger and Spear [5] suggested performing a linear transformation (rotation) of the normalized coordinate system in such a way that the aforementioned univariate analysis for indicating separation can still be utilized. In this transformed coordinate system the sensitivity analysis is continued by determination of \( d_{m,n}^* \), which is then associated with well- or ill-determined linear parameter combinations instead of individual parameters.

However, as was recognized in [8], the results pertain only to the predefined parameter space defined by the limits of the uniform distributions usually chosen. It is evident that this parameter space represents the parametric uncertainty in the sensitivity analysis in a more realistic way than a particular parameter vector. But one should still be alert to the effects of wrong assumptions with respect to the predefined space. To illustrate a rather extreme effect of a predefined parameter range on the results, let the parameter \( p_i \) be defined by \( p_i \in U[-\infty, \infty] \). Unlike the non-behavior-giving range, the behavior-giving range of \( p_i \) will then be extremely small with respect to the predefined range. It can be seen from the associated cumulative distribution functions that \( d_{m,n} \) approaches 0.5 for a large number of simulations, which means that any parameter can be selected as a dominant parameter. On the other hand it is also clear that, besides the range, the position of the parameter bounds is of great importance.

To avoid this problem, Whitehead and Hornberger [10] placed the regionalized sensitivity procedure within an iterative setting, adjusting the parameter space for each new iteration in a rather heuristic way. In this way the Monte Carlo simulations are used as a crude estimation procedure (see also Hornberger and Cosby [3]), so that the final information results from a posteriori defined parameter intervals.

A more systematic approach for selecting proper parameter intervals is provided by Keesman and Van Straten [7]. The a posteriori defined parameter intervals are then determined in a less heuristic way by an algorithm using random scanning and principal-component analysis. It is worth noting then that the parameter intervals are not only adjusted by translation, as was done by Whitehead and Hornberger [10], but also by rotation. The effect of rotation is particularly important when there are parameter correlations.

So, by incorporating a formal parameter-space identification algorithm prior to RSA, a better starting point, from the viewpoint of bias due to less subjective guesses, for the sensitivity analysis is obtained. It must be noted that some extra uncertainty must then be injected by means of enlarged parameter ranges, to explore the sensitivity analysis more fully. It must be emphasized that RSA is performed completely within a statistical setting, unlike the following approach.
3. A POSTERIORI SET-THEORETIC PARAMETRIC UNCERTAINTY ANALYSIS

As an alternative to the previous approaches, information about dominant parameters is obtained now from a set-theoretic point of view, which means that we assume that the a priori parametric as well as the observation uncertainty can be characterized by an upper and a lower bound. On the basis of these weaker assumptions with respect to the uncertainties, it is not a single "optimal" parameter vector that is found, but a set of equally acceptable parameter vectors expressing the a posteriori parametric uncertainty. Within this context Keesman and Van Straten [7] proposed a parameter-space adjustment algorithm to find the so-called behavior-giving parameter space, which is consistent with the behavior space (spanned by a number of bounds in the observation space), the assumed model structure, and the prior knowledge of the parameter ranges. Herein, four steps can be distinguished:

(1) Standardization of the original parameter space to stress the effect of rotation in step 3.
(2) Identification of behavior-giving parameter vectors.
(3) Rotation of the coordinate-axis system of standardized variables according to the principal axes of the behavior-giving subspace.
(4) Extension of the behavior-giving subspace.

Standardization of the original space is obtained via the a priori transformation of the separate parameters $p_j$,

$$\theta_j = \sqrt{3} - 2\sqrt{3} \frac{\bar{p}_j - p_j}{\bar{p}_j - \bar{p}_j}, \quad j = 1, \ldots, s$$

where $p_j$ and $\bar{p}_j$ represent the predefined lower and upper bounds on $p_j$, and $s$ denotes the number of parameters. It can be derived easily that this transformation is identical to the statistical standardization assuming independent uniform distributions. In this standardized space a number of behavior-giving parameter vectors are identified and analyzed to obtain the rotation matrix for a subsequent iteration. The basic idea is that information about dominant parameter is available from this rotation matrix containing the eigenvectors of the covariance matrix of behavior-giving parameter vectors in the standardized space. Note that the behavior-giving parameter space is then approximated (via the eigenvectors of the covariance matrix) by an ellipsoid to utilize information from the set of behavior-giving parameter vectors. As in the stochastic case, eigenvalues associated to the eigenvectors
will indicate the worst- and best-determined directions in the parameter space.

It is worth noting that the identified rotation matrix in the standardized space will depend on the predefined parameter bounds in the original space according to the transformation of (1). To filter out this effect it is desirable to determine a rotation matrix with respect to the bounds of the behavior-giving parameter space. The covariance matrix $\Sigma_m$ of behavior-giving parameter vectors in the standardized space is then transformed to

$$
\Sigma_m^+(i, j) = \Sigma_m(i, j) \frac{\delta p_i \delta p_j}{\delta p_i^+ \delta p_j^+}, \quad i = 1, \ldots, s \text{ and } j = 1, \ldots, s,
$$

where $\delta p$ represents the parameter interval, and the indices $m$ and $+$ denote respectively the behavior-giving class and the dependence of the corresponding variable on the behavior-giving class. Well-determined directions in the standardized parameter space are indicated by those eigenvectors of $\Sigma_m^+$ which are associated to the smallest eigenvalues.

From the viewpoint of consistency of the results, the eigenvectors of $\Sigma_m^+$ are compared with the eigenvalues of the correlation matrix, which is the covariance matrix of standardized variables, resulting from a stochastic parameter estimation process. This comparison should reveal that the sets of eigenvectors are similar, assuming uniform distributions and correct estimation results.

4. EXAMPLE

The regionalized sensitivity analysis using either a priori or a posteriori parametric information and the newly proposed method, analysis of the a posteriori set-theoretic parametric uncertainty, are now applied to the determination of dominant parameters in a simple water-quality model. This model describes the diurnal dissolved oxygen (DO) patterns in a well-mixed lake. In addition, the analysis of parametric uncertainty from a maximum-likelihood estimation is applied for comparison.

The DO model,

$$
\dot{c}(t) = K_r [C_s(t) - c(t)] + aI(t) - R,
$$

(3)
where

\[ c(\cdot) = \text{DO concentration (g/m}^3\text{)}, \]

\[ C_s(\cdot) = \text{saturation concentration (g/m}^3\text{)}, \]

\[ I(\cdot) = \text{radiation (W/m}^2\text{)}, \]

\[ K_r = \text{reaeration coefficient (day}^{-1}\text{)}, \]

\[ a = \text{photosynthetic rate coefficient (g/m}^3\text{ day W)}, \]

\[ R = \text{sink term (g/m}^3\text{ day)}, \]

describes the rate of change of DO as a function of reaeration exchange with the atmosphere, photosynthetic production from algae and water plants, and consumption due to respiration, biodegradation, and sediment processes. With respect to the availability of the data two situations are distinguished.

### 4.1. Analysis with Sufficient Data

For the situation with sufficient data, hourly observations are available for a period of eight days (Figure 1). It must be noted that the observed DO concentration at instant \( k = 169 \) is most likely unreliable, so that this observation is regarded as an outlier (see Keesman and Van Straten [6]). The system's behavior is defined in terms of observed DO concentrations plus or minus an assumed observational error bound of 1.5 g/m\(^3\).

![Figure 1. Observed DO concentrations (Lake De Poel and 't Zwet, 21–30 April 1983).](image-url)
For the application of the original RSA, initial parameter intervals (Table 1) are specified from literature and engineering judgement (see [6]). On the basis of these predefined parameter intervals, the defined system's behavior, and the DO model, only three behavior-giving parameter vectors are found from 200 simulations. This number of behavior-giving parameter vectors is insufficient to perform further statistical analysis. At this point additional simulations could be performed in order to obtain a sufficient number. But from the viewpoint of computational efficiency it is more attractive at this point to obtain additional behavior-giving parameter vectors from a more efficient estimation procedure.

From preceding investigations (see [6, 7]), a posteriori information for initial parameter intervals is available. In this application of the modified RSA the individual parameter intervals are enlarged by 20% (Table 2) to take account of incomplete coverage of the parameter space due to the random selection procedure.

Two hundred Monte Carlo simulations then result in 43 behavior-giving and 157 non-behavior-giving parameter vectors under the specified classification. The RSA results are presented in Table 3. The parameter vector \( [\theta_1 \theta_2 \theta_3]^T \) denotes the standardized parameter vector \( [K_r, a, R]^T \) according to Equation (1). For a better understanding of the results presented in Table 3 it is worth noting that the elements of the eigenvectors are the weights of the original parameters in a linear combination of these parameters, on rotating the original axis system. Due to significant correlations between the parameters \( a \) and \( R \) (see [6, 7]), a univariate analysis is incompetent to

### Table 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_r )</td>
<td>0.0 - 2.0</td>
</tr>
<tr>
<td>( a )</td>
<td>0.0 - 0.1</td>
</tr>
<tr>
<td>( R )</td>
<td>0.0 - 5.0</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( K_r )</td>
<td>0.98 - 1.71</td>
</tr>
<tr>
<td>( a )</td>
<td>0.036 - 0.072</td>
</tr>
<tr>
<td>( R )</td>
<td>1.2 - 3.57</td>
</tr>
</tbody>
</table>
identify dominant parameters [8]. Therefore a principal-component transformation of the "covariance matrix" of either the behavior-giving or the non-behavior-giving class (see [5]) was performed to determine the statistics $d_{m,n}$. This maximum vertical distance between the two sample distribution functions in the transformed coordinate system is then associated to a principal axis. The resulting statistics indicate that only the linear combination $-0.015\theta_1 - 0.737\theta_2 - 0.676\theta_3$ (representing a principal axis) in the standardized space is dominant. The associated $d_{m,n}$ equals 0.433, which is larger than 0.24, the critical $d_{m,n}$ at 99% confidence level (see Table 3), indicating a significant difference in the distribution functions.

The relevant set-theoretic parameter-estimation results, according to Keesman and Van Straten [7], are presented in Table 4. The eigenvector associated to the smallest eigenvalue of the covariance matrix indicates the

### Table 3

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Eigenvectors in standardized space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 (K_r)$</td>
<td>$-0.853$ $-0.497$ $-0.015$</td>
</tr>
<tr>
<td>$\theta_2 (a)$</td>
<td>$0.328$ $-0.608$ $-0.737$</td>
</tr>
<tr>
<td>$\theta_3 (R)$</td>
<td>$0.407$ $-0.619$ $0.676$</td>
</tr>
<tr>
<td>$d_{m,n}^*$</td>
<td>$0.151$ $0.152$ $0.433$</td>
</tr>
</tbody>
</table>

Number of simulations: 200
% behavior-giving: 21.5%

Critical values of Kolmogorov-Smirnov test statistic:

$d_{m,n}(0.90) = 0.19$, $d_{m,n}(0.95) = 0.20$, $d_{m,n}(0.99) = 0.24$

"Columns represent the eigenvectors (principal axes) of the "covariance matrices".

### Table 4

<table>
<thead>
<tr>
<th>Updated rotation matrix in standardized space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.168$ $0.986$ $0.021$</td>
</tr>
<tr>
<td>$0.863$ $-0.127$ $0.747$</td>
</tr>
<tr>
<td>$0.739$ $-0.112$ $-0.665$</td>
</tr>
</tbody>
</table>

Eigenvalues: $1.69$ $1.05$ $0.03$

"Columns represent eigenvectors of $\Sigma_m^+$. 

---

TABLE 3
MODIFIED RSA RESULTS

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Eigenvectors in standardized space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1 (K_r)$</td>
<td>$-0.853$ $-0.497$ $-0.015$</td>
</tr>
<tr>
<td>$\theta_2 (a)$</td>
<td>$0.328$ $-0.608$ $-0.737$</td>
</tr>
<tr>
<td>$\theta_3 (R)$</td>
<td>$0.407$ $-0.619$ $0.676$</td>
</tr>
<tr>
<td>$d_{m,n}^*$</td>
<td>$0.151$ $0.152$ $0.433$</td>
</tr>
</tbody>
</table>

Number of simulations: 200
% behavior-giving: 21.5%

Critical values of Kolmogorov-Smirnov test statistic:

$d_{m,n}(0.90) = 0.19$, $d_{m,n}(0.95) = 0.20$, $d_{m,n}(0.99) = 0.24$

"Columns represent the eigenvectors (principal axes) of the "covariance matrices".

---

TABLE 4
SET-THEORETIC PARAMETER-ESTIMATION RESULTS

<table>
<thead>
<tr>
<th>Updated rotation matrix in standardized space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0.168$ $0.986$ $0.021$</td>
</tr>
<tr>
<td>$0.863$ $-0.127$ $0.747$</td>
</tr>
<tr>
<td>$0.739$ $-0.112$ $-0.665$</td>
</tr>
</tbody>
</table>

Eigenvalues: $1.69$ $1.05$ $0.03$

"Columns represent eigenvectors of $\Sigma_m^+$. 

---
best-determined direction \([1]\). So, from Table 4 it can be seen that the best-determined direction in the standardized parameter space is represented by the vector \([0.021 \ 0.747 \ -0.665]^T\). Note that the worst-determined direction \(([0.168 \ 0.653 \ 0.739]^T)\) is perpendicular to the best-determined direction. It is also worth noting that the orthogonality of the axes in the standardized space will be lost after mapping these axes to the original parameter space according to Equation (1).

It appears from Tables 3 and 4 that the results of the modified RSA and \emph{a posteriori} set-theoretic parametric uncertainty analyses yield consistent information about the dominance of linear parameter combinations in the standardized parameter space. These results are also confirmed by results from \emph{a priori} analysis.

From previous data analysis (see \([6]\)) the following linear relationship between the separate model parameters was found: \(R \approx -0.07Kr + 51.5a - 0.17\). Incorporation of this relationship into Monte Carlo simulation analysis will increase the percentage of behavior-giving parameter vectors. In the standardized space (with respect to the behavior-giving parameter space) this relationship is transformed to

\[
\theta_3 \approx -0.021\theta_1 + 0.773\theta_2 + 1.16,
\]

using Equation (1) and information about the upper and lower limits from Table 2. So \(\theta_3\)'s sampled far away from this plane will contribute to a high percentage of non-behavior-giving parameter vectors. In other words, the direction perpendicular to the plane represented by Equation (4) in the standardized space will be well determined, i.e., behavior-giving parameters can be found in a relatively small range. It can also be seen from (4) that these considerations are particularly valid in the \(\theta_2\)-\(\theta_3\) plane. The well-determined direction in this plane can be represented by the normalized vector \([-0.79 \ 0.61]^T\), which agrees with results presented in Tables 3 and 4. Note that the relationship of (4) is not exact because of errors in the data (see \([6]\)).

Finally, the results of the proposed methods (Tables 3 and 4) are compared with those resulting from a maximum-likelihood estimation (Table 5). Although structural model error is present in the defined DO model \([6]\), it is expected that the results will not differ much. The dominant directions in the standardized space (represented by the eigenvector associated to the smallest eigenvalue) from the set-theoretic (Table 4) and from the stochastic uncertainty (Table 5) modeling approach are in agreement.

### 4.2. Analysis with Sparse Data

For the situation with sparse data it is supposed that 16 (characteristic) observations are available (Figure 2), reflecting the dissolved oxygen patterns.
It is expected that the dominant direction from these sparse data will not differ too much from the one resulting from the complete data set. Again, the system’s behavior is expressed in terms of the observations and an assumed observation error bound of 1.5 g/m². Due to failure of the sensor (a quite common practical problem), a number of these observations are only represented by their upper or lower bound, which is indicated by the vertical lines in Figure 2.

Under these (hypothetical) circumstances the RSA is applied using the a priori information from Table 1. Of 200 sampled parameter vectors, 15 are classified as behavior-giving. The RSA results are presented in Table 6.

Note that the dominant parameter combination of the preceding analysis with sufficient data (Table 3) is indicated as dominant at a lower than 99% confidence level (second eigenvector in Table 6). In addition, another linear combination is indicated as dominant at a 99% confidence level. The differences in these results can be to a large extent attributed to the choice of the

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Correlation matrix</th>
<th>Eigenvectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_c$</td>
<td>1</td>
<td>0.968 0.139</td>
</tr>
<tr>
<td>$a$</td>
<td>0.149 1</td>
<td>-0.048 -0.711</td>
</tr>
<tr>
<td>$R$</td>
<td>0.033 0.587 1</td>
<td>-0.245 0.690</td>
</tr>
<tr>
<td>Eigenvalues:</td>
<td></td>
<td>1.164 0.984 0.401</td>
</tr>
</tbody>
</table>

![Fig. 2. Reduced information about the observed DO concentrations (Lake De Poel and 't Zwet, 21–30 April 1983).](image-url)
TABLE 6
RSA RESULTS

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Eigenvectors in standardized space</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1(K_r)$</td>
<td>-0.658 -0.175 0.711</td>
</tr>
<tr>
<td>$\theta_2(a)$</td>
<td>0.437 0.722 0.559</td>
</tr>
<tr>
<td>$\theta_3(R)$</td>
<td>0.613 -0.670 0.426</td>
</tr>
<tr>
<td>$d^\ast_{m,n}$</td>
<td>0.151 0.389 0.477</td>
</tr>
</tbody>
</table>

Number of simulations: 200

% behavior-giving: 7.5%

Critical values of Kolmogorov-Smirnov test statistic:

$d_{m,n}(0.90) = 0.30$, $d_{m,n}(0.95) = 0.34$, $d_{m,n}(0.99) = 0.40$

Columns represent the eigenvectors or principal axes of the "covariance matrices".

TABLE 7
SET-THEORETIC PARAMETER-ESTIMATION RESULTS

<table>
<thead>
<tr>
<th>Updated rotation matrix in standardized space</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.231 0.962 -0.141</td>
<td></td>
</tr>
<tr>
<td>0.669 0.052 -0.741</td>
<td></td>
</tr>
<tr>
<td>0.706 0.265 0.657</td>
<td></td>
</tr>
</tbody>
</table>

Eigenvalues: 1.38 0.70 0.10

Columns represent eigenvectors of $\Sigma^+_m$.

initial parameter intervals, for modified RSA results do not show these differences.

Set-theoretic parameter-estimation results using the same initial parametric information from Table 1 can be found in Table 7. The information in Table 7 agrees to a large extent with the information in Table 4 on the basis of the complete data set. From the computational point of view it is worth noting that the behavior-giving percentage increased from 7.5% in the RSA to 51% on application of the parameter-space adjustment algorithm.

A stochastic parameter-estimation evaluation is not possible on the basis of the set-theoretic (unknown but bounded) information that is present in the 16 observations.

5. CONCLUSIONS

Conventional methods for determination of dominant parameters have some severe limitations. In particular, the fact that these sensitivity analyses
pertain to a single point in the parameter space, preferably an optimal parameter estimate, can be an especially severe drawback when dealing with (complex) ill-defined systems. The region-ized sensitivity analysis proposed by Hornberger and Spear avoids this problem. However, it appears that their analysis still possesses some subjective elements. This is the starting point for two alternative methods, modified RSA and analysis of the set-theoretic parametric uncertainty, proposed in this paper. These methods are essentially based on a posteriori information, which can be extracted from very limited information about the system's behavior. Under circumstances with sufficient data these methods will yield information about the dominance of parameters that is consistent with maximum-likelihood estimation results, assuming that reliable estimation results are available. This will not always, however, be the case. So the proposed methods will cover a wider area of application.

I am indebted to Gerrit van Straten for his comments on preliminary versions of the paper.

This research is supported by the Netherlands Technology Foundation (STW).

REFERENCES

