A TWO-STEP APPROACH TO PROPAGATE RATING CURVE UNCERTAINTY IN ELBE DECISION SUPPORT SYSTEM

YUEPING XU[1], HARRIETTE HOLZHAUER[2] and MARTIJN.J. BOOIJ[3]

[1]Phd candidate, Water Engineering and Management, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE, Enschede, the Netherlands.
(Tel: +0031-53-4891094, Fax: +0031-53-4895377, e-mail: y.p.xu@ctw.utwente.nl)
[2]Researcher, Water Engineering and Management, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE, Enschede, the Netherlands.
(Tel: +0031-53-4892585; Fax: +0031-53-4895377, e-mail: h.holzhauer@ctw.utwente.nl)
[3]Assist professor, Water Engineering and Management, Faculty of Engineering Technology, University of Twente, P.O. Box 217, 7500 AE, Enschede, the Netherlands.
(Tel: +0031-53-4892564; Fax: +0031-53-4895377, e-mail: m.j.booij@ctw.utwente.nl)

Abstract

The relationships between water level and discharge along the river channel are vital for decision support systems in river basin management. Normally the reliability of the so-called rating curves along the river channel depends greatly on the accuracy and duration of the measured discharge and water level data. In the Elbe Decision Support System (DSS), the rating curves are combined with the HEC-6 model to investigate the effects of river engineering measures on the Elbe river system. Under such situations, the uncertainty originated from the HEC-6 model is of great importance for the reliability of the rating curves. This paper presents a two-step approach to analyze the uncertainty in the rating curves and propagate it into the vegetation model used in the Elbe DSS. The first step is to identify the uncertainty sources. An analytical method is adopted to propagate the uncertainty sources into the final rating curves. The second step is to propagate the uncertainties in the rating curves into the model outputs of the vegetation model using Monte Carlo Analysis. By this two-step uncertainty analysis approach, the uncertainty in the rating curves is successfully propagated into the vegetation model in the Elbe DSS. The final Monte Carlo Simulation results show large uncertainties in the model outputs.

Keywords: Elbe Decision Support System; Rating curves; Uncertainty propagation; Vegetation model; Monte Carlo Analysis; HEC-6

1. INTRODUCTION

Integrated river basin management involves issues like water quality, water supply, hydropower, flood risk, and ecology. In recent years, integrated river basin management is becoming more complicated because of conflicting water management issues, large amount of information, and changing environmental conditions. Decision support systems are one of the possible solutions to aid decision makers in dealing with such complicated management issues (Loucks and Da Costa 1991; Jamieson and Fedra 1996; Todini 1999; Salewicz and Nakayama 2004). To make a sound decision, decision makers need to be aware of the existence of uncertainty in the evaluation of river engineering measures in the DSSs. It is therefore proposed that the modellers provide model outputs with uncertainty information to decision makers. Uncertainty assessment is regarded as one of the most important issues in a decision support system. However, in most DSSs in river basin management, the uncertainty has not been taken into account.

A decision support system is currently under way for the Elbe River which is located in Central Europe. This decision support system focuses on the German part of the Elbe — the river section between the Czech Border (river km 0) up to weir Geesthacht (river km 568) (see Fig. 1). The DSS integrates important issues like water quality, flood risk, navigation, and ecology (De Kok et al. 2000). To cope with these different issues, several models are adopted, such as 1D hydraulic models (the rating curves and the HEC-6 model), a hydrological model HBV (Bergström 1995), a flood risk model, a shipping model, and a vegetation model. Among these models, the rating curves are one of the substantial components of the Elbe DSS. They serve as inputs into other models, such as the shipping and vegetation models.

The rating curves are often used to produce hydrological data. The reliability of the rating curves is essential to model relationships between the water level and discharge along the river channel. In this paper, the uncertainty in the rating curves is identified and a two-step uncertainty analysis approach is used to propagate this uncertainty into the Elbe DSS. The vegetation model is selected as an example.

2. MODELS

2.1 RATING CURVES AND HEC-6

In the Elbe DSS, two 1D hydraulic models are used to provide inputs to the vegetation model. One is based on the fitted rating curves and the other one is the HEC-6 model (US Army Corps of Engineers 1993).

The rating curves describe the relationships between the discharge (Q) and water level (H) along the river normally based on measurements at gage stations. It is a simple but useful method. The disadvantage of this method is that the rating curves based on measurements cannot model the effects of river engineering measures on the water levels. One such measure is channel dredging, which changes the geometry of the river channel and may change the rating curves along the river after implementation. The rating curves in the Elbe DSS are represented by (Shaw 1994)

$$H = a * Q^b \tag{1}$$

Where Q is the discharge (m^3/s), H is the water level (m), and a and b are location dependent parameters which can be found by least-square fitting.

The main function of the HEC-6 model is to compute the water levels for steady river flow in the main channel. It is a one-dimensional open channel flow model. The advantage of HEC-6 is its capability to take into account the effects of river engineering measures. One disadvantage is that this model cannot be used directly to compute the flood duration in the Elbe DSS while the flood duration is one dominant component for the vegetation model.

In order to investigate the effects of different river engineering measures, the two models are combined in the Elbe DSS. First the HEC-6 model is used to produce the discharge and water level data for different measures for each location along the Elbe River. Then these discharge and water level data are analyzed to derive the rating curves along the river by regression analysis. By doing this, new rating curves can be obtained to account for the possible measures that affect the geometry of channel and floodplains. The rating curves are finally used as inputs into the vegetation model in the DSS.

2.2 VEGETATION MODEL

The vegetation model is used to produce maps for the dominant groups of vegetation (so called biotypes) in the floodplain area along the Elbe (Fuchs et al. 2002). The vegetation model uses the flood duration (i.e. the total number of flooding days per year), the distance to the main channel, and land use to determine the presence or absence of biotypes. This model has been developed by the German Federal Institute (one of the main decision makers for the Elbe DSS). Eleven different biotypes are identified in the vegetation model, which are shown in Tab. 1.

One important concept in this vegetation model is the flood duration. The lognormal distribution is used to model the daily discharge statistics (Shaw 1994). The number of flooding days based on the critical discharge in the floodplain area is calculated for each cell (x, y) in the area using the approximation of error function. The approximation is given by

$$N_{flood}(x,y) = \frac{365}{2}\left(1 - erf\left[\frac{\log(Qcrit(x,y)) - \mu(x)}{\sigma(x)\sqrt{2}}\right]\right) \tag{2}$$

Where the discharge parameters μ and σ are location dependent and given by the daily discharge range (Matthies et al. 2003). The critical discharge Q_{crit} is defined as the discharge at which a certain piece of land (x, y) starts to flood. In order to determine this critical discharge, the elevation of the land z(x, y) and the rating curves are needed. The error function erf can be found in Abramowitz and Stegun (1972).

The decision variables here are the frequencies of 11 biotypes in Tab. 1. They are calculated as the individual number of cells of each biotype divided by the total number of cells of all biotypes in the floodplains. The decision variables are to characterize the biotype diversity in the floodplains along the Elbe. The equation is

$$Per_i = \frac{N_i}{N_{total}} \tag{3}$$

Where N_i is the number of cells of ith biotype in the floodplains and N_{total} is the total number of cells of all biotypes.

3. UNCERTAINTY SOURCES

Two linear regression analyses are involved in estimating the uncertainty in the rating curves. As stated before, the rating curves along the Elbe main channel are fitted from the discharge and water level data calculated from the HEC-6 model by linear regression analysis. This is the first linear regression analysis. In addition, the water levels upstream and downstream are highly dependent on each other. This is presented by a high dependence of the parameters a / b upstream and downstream (Eq. (1)). The dependence is modelled by another linear regression analysis in order to estimate the final uncertainty in the rating curves (uncertainty in the regression parameters). This is the second regression analysis. Due to data availability, the uncertainty analysis is only applied to the river section 345-425 km.

The reliability of the rating curves depends greatly on the accuracy of the water levels calculated by the HEC-6 model. One source of uncertainty in the rating curves is therefore the error in the computed water levels from HEC-6. In the Elbe DSS, HEC-6 calculates the water levels for 10 different discharges for every 100 meters along the concerned river section. These 10 values correspond to the discharges of different return periods, which represent a whole range of the river flows in the Elbe. A rough estimation of the error is obtained from the calibration of HEC-6 along the Elbe river sections 252-272 km, 291-299 km, and 332-343 km (Nestman and Buchele 2002). According to them, the differences between the measured and calculated water levels are from 4 cm to 10 cm. A maximum value of 10 cm is adopted in this paper as the error in the calculated water levels in the concerned river section. This error will be propagated into the parameters a and b in Eq. (1) for each river location. Assume there is no correlation between the parameters a and b.

4. A TWO-STEP UNCERTAINTY PROPAGATION APPROACH

In order to investigate the effect of the uncertainty in the rating curves on the vegetation model, a two-step approach is proposed. First an analytical method is used to investigate the uncertainty in the rating curves. Then this uncertainty is propagated into the model outputs of the vegetation model by using Monte Carlo Analysis.

4.1 STEP 1: ANALYTICAL UNCERTAINTY PROPAGATION

The analytical approach is to investigate the uncertainty in the parameters a and b in Eq. (1) and then propagate this uncertainty into the parameters in the second regression analysis. Assume the error in the water levels is normally distributed and the standard deviation is estimated to be $\sigma_H = 5$ cm. To apply an analytical method, the equation of the rating curves (Eq. (1)) is transformed to

$$\log H = \log a + b \log Q \tag{4}$$

Assume x=log Q and y=log H. Then have $\sigma_y = \sqrt{\sigma_H^2 (\frac{\partial y}{\partial H})^2} = \frac{\sigma_H}{H}$. Furthermore, assume A=log a and B= b. According to Sabatelli et al. (2002), the estimation of A and B are shown in Eq (5)

$$A = \frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_H^2} \sum \frac{y_i}{\sigma_H^2} - \sum \frac{x_i}{\sigma_H^2} \sum \frac{x_i y_i}{\sigma_H^2}\right)$$

$$B = \frac{1}{\Delta}\left(\sum \frac{1}{\sigma_H^2} \sum \frac{x_i y_i}{\sigma_H^2} - \sum \frac{x_i}{\sigma_H^2} \sum \frac{y_i}{\sigma_H^2}\right) \tag{5}$$

$$\Delta = \sum \frac{1}{\sigma_H^2} \sum \frac{x_i^2}{\sigma_H^2} - \left(\sum \frac{x_i}{\sigma_H^2}\right)^2$$

A first- order error propagation equation is then used to propagate the error originated from HEC-6 into the uncertainty in A and B. This equation can be found in Bevington and Robinson (1992). The standard deviations of A and B are computed

$$\sigma_A = \sqrt{\frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_H^2}\right)} \qquad \sigma_B = \sqrt{\frac{1}{\Delta}\left(\sum \frac{1}{\sigma_H^2}\right)} \tag{6}$$

So the uncertainty in the original parameters a and b is

$$\sigma_a = \sigma_A e^A \qquad \sigma_b = \sqrt{\frac{1}{\Delta}\left(\sum \frac{1}{\sigma_H^2}\right)} \tag{7}$$

The second linear regression analysis is used to model the dependence of upstream and downstream a and b. The equations to model the dependency are:

$y = f_1 + e_1 x$ for parameter a

$y = f_2 + e_2 x$ for parameter b (8)

The uncertainty in a and b that is calculated by Eq. (7) is propagated into the regression parameters e_1, f_1, e_2, and f_2 in Eq. (8). The method to calculate the uncertainty in regression parameters can be found in Sabatelli et al. (2002).

4.2 STEP 2: MONTE CARLO ANALYSIS

The second step is to propagate the uncertainty in the regression parameters e_1, f_1, e_2, and f_2 into the vegetation model by Monte Carlo Analysis. The sampling scheme used in Monte Carlo Analysis is Latin Hypercube Sampling (LHS) (Saltelli et al. 2000). This sampling scheme first segments the assumed probability distributions into a number of intervals, each having equal probability. Then, from each interval, a value is selected at random according to the probability distributions within the interval. Latin Hypercube Sampling is generally more precise for producing random samples than conventional Monte Carlo sampling, because the full range of the probability distribution is sampled more evenly.

5. RESULTS

5.1 UNCERTAINTY IN REGRESSION PARAMETERS

The uncertainty in the rating curves is propagated into e_1, f_1, e_2, and f_2 using the analytical method described in Section 4. Tab. 2 shows the computed mean values and corresponding uncertainties (expressed by standard deviations) of the regression parameters e_1, f_1, e_2, and f_2.

The numbers in these four regression parameters have quite different orders of magnitude. The order of magnitude of the uncertainty in f_1 is relatively high compared to those of other parameters. So does the mean values.

5.2 UNCERTAINTY IN THE FREQUENCIES OF 11 BIOTYPES

The uncertainty in the regression parameters shown in Tab. 2 is then propagated into the vegetation model. Assume these four parameters are normally distributed. 100 Latin Hypercube simulations are used in this uncertainty analysis. As stated before, the decision variables here are the frequencies of 11 different biotypes in the floodplains. Fig. 2 and 3 show examples of the scatter plots expressing the relationships between the four regression parameters and the frequencies of Biotype 3 and 4. The effects of the regression parameters are different for these two biotypes. For Biotype 3, the scatter plots show a high screwed relationship between the four parameters and the vegetation model outputs. Most model outputs center around the mean value 0.04. The frequency of Biotype 3 decreases monotonously with the parameter f_1. For Biotype 4, the model outputs distribute more or less evenly and a monotonously increase with f_1 can be observed. These graphs show that the effect of regression parameter f_1 on the model outputs and its uncertainty is more important than those of the other parameters. As expressed in Eq. (8), f_1 is the regression parameter related to the parameter a in Eq. (1). This indicates that the parameter a may be more important than the parameter b in this case.

Fig. 4 shows the uncertainties in the frequencies of 11 biotypes in the floodplains along the Elbe. ‘Bio 0’ in this figure indicates the situation with no data available. The error bars show the mean values, 10, and 90 percentiles of the model outputs. 10 percentile of the model outputs indicates the value that is greater than 10 percent of the values in the frequencies of biotypes along the concerned river section. The values of 10 and 90 percentiles indicate the amount of uncertainty in the frequencies of 11 biotypes. From this figure, high uncertainties can be observed in the model outputs of this vegetation model. For example, for Biotype 4, the mean value is around 0.06. The 10 and 90 percentiles are 0.02 and 0.13 respectively, which shows high variability of the value of the frequency of Biotype 4.

Some facts can be figured out from the uncertainty results shown in Fig. 4 as well. For example, the error bar of Biotype 2 shows that Biotype 2 is likely to disappear in the future under the current uncertainty analysis. If more diverse biotypes in the floodplains are hoped, measures need to be identified by relevant decision makers to increase the frequency of Biotype 2. As shown in Tab. 1, Biotype 2 is soft wood. Therefore renaturation in the floodplains may be a good solution to increase the biotype diversity. The error bars shown in Fig. 4 can actually provide very useful information of the uncertainty in the biotype diversity in the floodplains along the Elbe in the future to decision makers for making a better decision.

6. CONCLUSIONS

The rating curves are a vital component of the Elbe DSS in producing the hydrological data and modelling the effects of river engineering measures (combined with other hydraulic models). The example of the vegetation model in the Elbe DSS in this paper demonstrated the propagation of uncertainty originated from the rating curves into the model outputs by a two-step uncertainty analysis approach. Although the uncertainty propagated into the frequencies of 11 biotypes is high, it provides useful insights on the uncertainty information for further decision making in river basin management. As it is known, models are never perfect. Knowing about the uncertainty can help decision makers understand the gaps in the current knowledge. Uncertainty analysis can thus provide the possibility to collect additional information to reduce the uncertainty and achieve better decision making.

The two-step approach proposed in this paper is a quantitative method to propagate the uncertainty into model outputs. This approach increases the accuracy of uncertainty analysis by applying an analytical error propagation equation. The example of the vegetation model in the Elbe DSS shows a successful application of this approach. The extension of its applicability can be a full analysis of uncertainty in the data and models in the decision support system and thus to assist decision makers with the complicated management problem.

ACKNOWLEDGEMENTS

The authors wish to acknowledge people involved in the development of the Elbe DSS. Special thanks are sent to the German Federal Institute for Hydrology in Germany for data supply. The authors also thank Jean-Luc de Kok of the University of Twente for extensive and useful discussions.

REFERENCES

Abramowitz, M. and Stegun, I.A. (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover Publications, New York.

Bergström, S. (1995). “The HBV model,” Computer Models of Watershed Hydrology, V.P.Singh (Eds.), Water Resources Publish, Highlands Rangch, pp. 443-476.

Bevington, P.R. and Robinson, D.K. (1992). Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill.

De Kok, J.L., Wind, H.G., Delden, H. and Verbeek, M. (2000). Towards a Generic Tool for River Basin Management. Feasibility Assessment for a Prototype DSS for the Elbe. Feasibility Study – report 2/3, Final report, University of Twente, Enschede, the Netherlands.

Fuchs, E., Giebel, H., Hettrich, A., Huesing, V., Rosenzweig, S. and Theis, H.J. (2002). Applications of Ecologicla Models in Water and Navigation Management, the Integrated Model INFORM, NfG Mittelung NR 25, Koblenz (in German).

Jamieson, D.G. and Fedra, K. (1996). “The ‘WaterWare’ decision support system for river basin planning. 1. Conceptual design,” Journal of Hydrology, Vol. 177, pp.163-175.

Loucks, D.P. and Da Costa, J.R. (1991). Decision Support Systems, Water Resources Planning, Springer, Berlin.

Matthies, M., Berlekamp, J., Lautenbach, S., Graf, N., Reimer, S., Hahn, B., Engelen, Van der Meulen, G. M., De Kok, J.-L., Van der Wal, K.U., Holzhauer, H., Huang, Y., Nijeboer, M., Boer, S. (2003). Pilot Phase for Building a DSS for River Basin Management with the Elbe as an Example, Institut fur Umweltsystemforschung (USF) der Universitat Osnabruck, Department of Water Engineering and Management of the University of Twente, Research Institute for Knowledge Systems (RIKS), Infram International BV, Zwischenreport Phase 1, im Auftrag der Bundesanstalt fur Gewasserkunde, Projektgruppe Elbe Okologie, BMBF Forschungsvorhaben FKZ 339542A, Koblenz-Berlin, 102 S (in German).

Nestman, F. and Buchele, B. (2002). Morphodynamics of the Elbe, Final Report from BMBF – Project with Separate Contributions and Appendix– CD, Kapitel III-2, K., Institut Fur Wasserwirtschaft und Kulturtechnik der Universitat Karlsruhe, Karlsruhe (in German).

Sabatelli, V., Marano, D., Braccio, G. and Sharma, V.K. (2002). "Efficiency test of solar collectors: uncertainty in the estimation of regression parameters and sensitivity analysis," Energy Conversion and Management, Vol. 43, pp.2287-2295.

Salewicz, K. A. and Nakayama, M. (2004). "Development of a Web-based Decision Support System (DSS) for Managing Large International Rivers," Global Environmental Change, Vol.14, No. 1, pp. 25-38.

Saltelli, A, Chan, K. and Scott, E.M. (2000). Sensitivity Analysis, John Wiley & Sons Ltd., England.

Shaw, E.M. (1994). Hydrology in Practice, T.J. International Ltd, Padstow, Cornwall, Great Britain.

Todini, E. (1999). "An operational decision support system for flood risk mapping, forecasting and management," Urban Water, Vol. 1, pp.131-143.

U.S. Army Corps of Engineers. (1993). CPD-6, HEC-6, Scour and Deposition in Rivers and Reservoirs, User's Manual.

Table 1. Biotypes in the floodplains along the Elbe River

Biotype number	Biotype description
0	no data
1	Seasonally flooded grassland
2	Softwood floodplain forest
3	Hardwood floodplain forest
4	Reed
5	Herb fringes and herb meadows
6	Grassland of wet to moist sites
7	Intensively used, species-poor, moist grassland
8	Other reeds
9	Herby flood banks and -plains near the water
10	Dry and warm ruderal sites with dense vegetation
11	Moist ruderal sites

Table 2. Uncertainty in e_1, f_1, e_2, and f_2

Regression parameters	Units	Means	Standard deviations
e_1	-	-1.51E-1	1.16E-3
f_1	-	79.81	0.46
e_2	-	2.18E-4	7.26E-6
f_2	-	-2.41E-2	2.78E-3

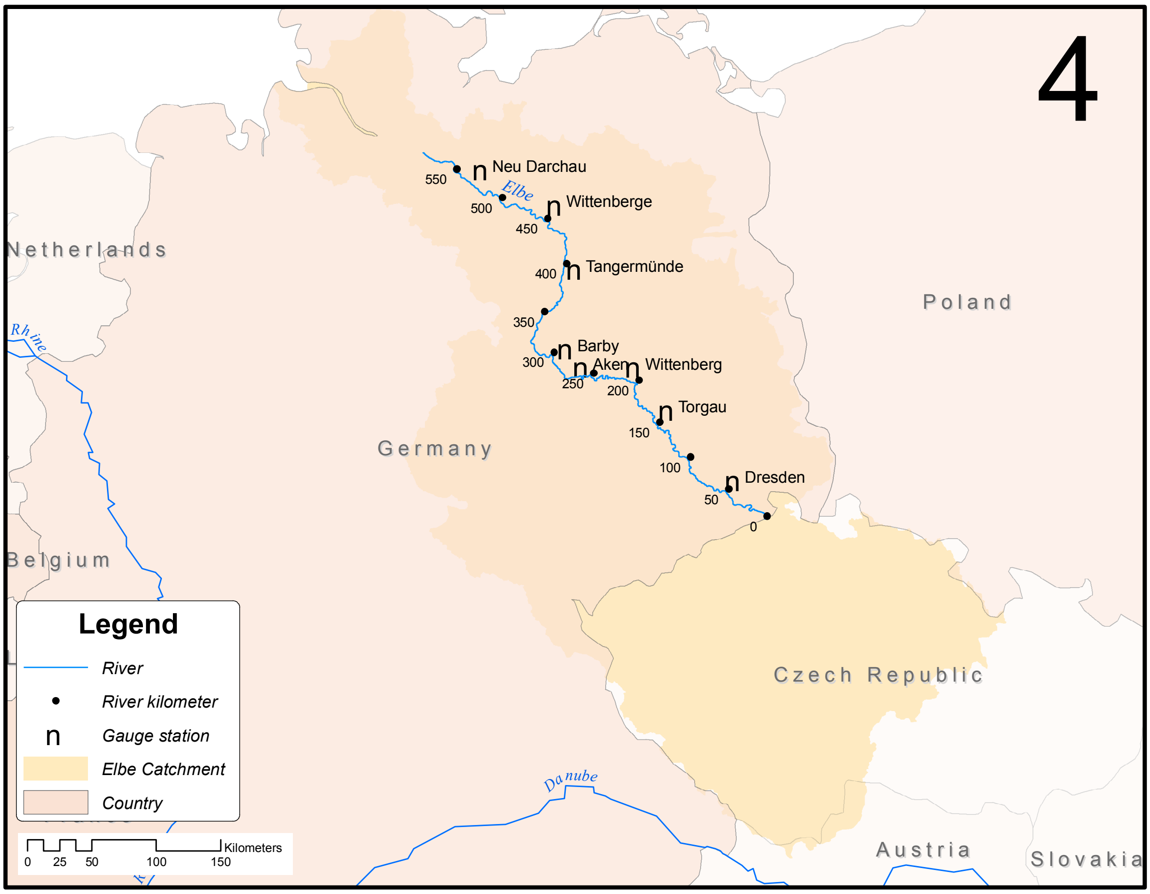

Fig. 1 Elbe River in Germany

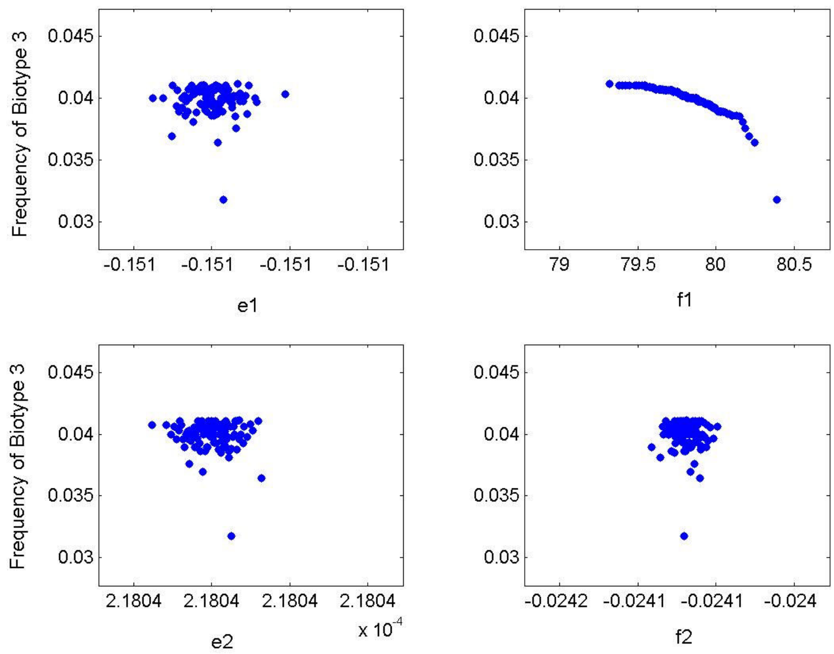

Fig. 2 Regression parameters vs. model outputs for Biotype 3

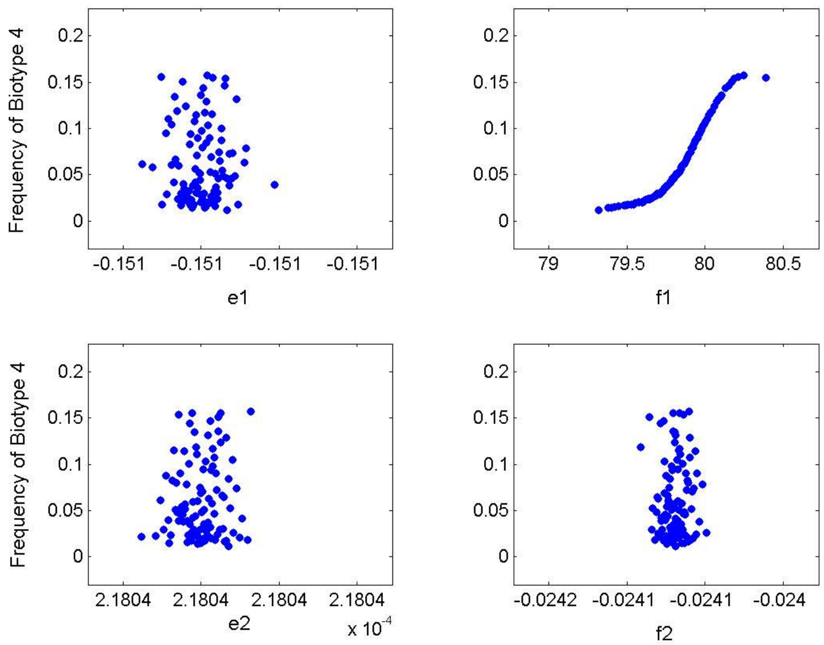

Fig. 3 Regression parameters vs. model outputs for Biotype 4

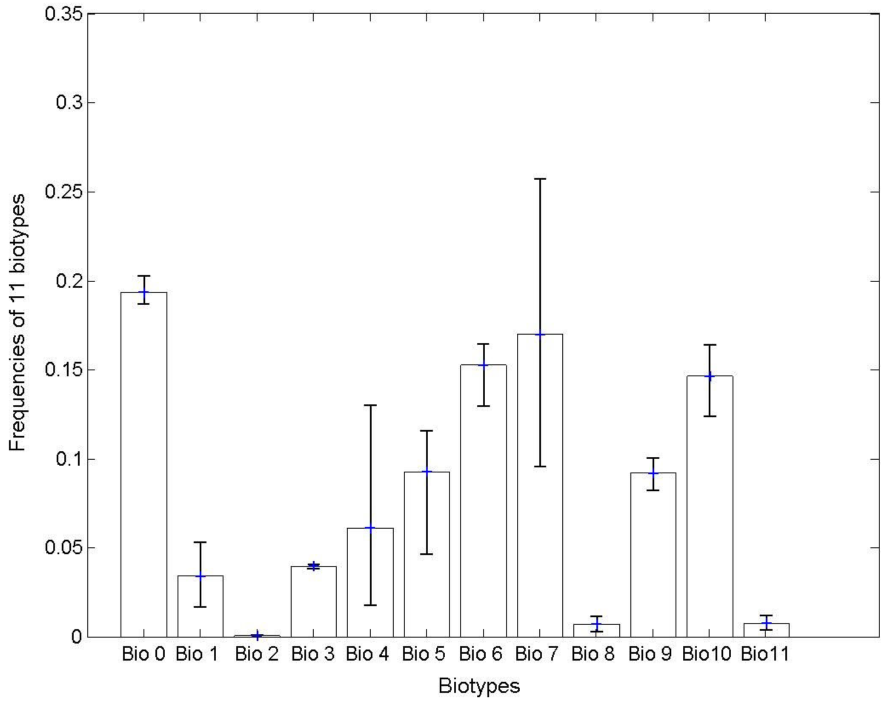

Fig. 4 Error bars for the frequencies of 11 dominant biotypes in the Elbe floodplains ('Bio 0' is the situation without data, 'Bio 1' ~ 'Bio 11' are the 11 dominant biotypes in the Elbe floodplains)