Parameter, Data Input and Structural Uncertainty Propagating through Coupled Models in a River Water Quality Modelling System

K.-E. Lindenschmidt[a]

[a] *GeoForschungsZentrum Potsdam, Section 5.4 – Engineering Hydrology, Telegrafenberg, 14473 Potsdam, Germany, email: kel@gfz-potsdam.de*

Abstract: This paper investigates three sources of uncertainty in a river water quality modelling system: parameters, input data and model structure. Emphasis is placed on structural uncertainty since this aspect has received less attention in the literature compared to parameter and input data uncertainty. Model structure is understood to be the equations and algorithms used to describe processes of substance transport and transformation in a model. Focus is given to two processes: i) sorption of heavy metals to suspended solids as a function of the fraction of organic carbon constituting the solids and ii) phytoplankton growth limitation by light extinction through the water column as a function of the concentration of suspended solids and chlorophyll-a in the water. Both processes are described by empirical relationships using regression curves derived from field data. The variance in the y-axis interceptions of the curves (also called regressor uncertainty) is taken to be the uncertainty bounds for the Monte Carlo simulations used for the uncertainty analyses. The three models of the WASP5 (Water quality Analysis Simulation Package), DYNHYD (hydrodynamics), EUTRO (dissolved oxygen, nutrient and phytoplankton dynamics) and TOXI (transport and transformation of sediments and micro-pollutants) were used for the study and were coupled in the HLA (High Level Architecture) platform to allow interactions between the models during simulations. The results show that structural uncertainty can be more significant than parameter and input data uncertainty to the overall uncertainty ranges of the model output, especially if the processes are very sensitive to the state variables. In addition, the behaviour of uncertainty propagation through a system of models is investigated.

Keywords: Model structure uncertainty; Saale River; WASP5, Water quality modelling.

1. INTRODUCTION

Model uncertainty is a necessary consideration when modelling due to errors in sampled data and model structure. Sources of uncertainty include i) the values and applicable range of values for the parameter settings used to calibrate the model results, ii) the data used as initial and boundary conditions in the model and iii) the structure of the model which includes the equations and algorithms used for the simulations. This paper investigates aspects of all three and is drawn from excerpts from Lindenschmidt et al. (2005a, 2005b), Lindenschmidt (2006).

One aim of this research is to simulate the hydrodynamics and water quality of the regulated river Saale (Germany) and to conduct an uncertainty analysis of the hydrodynamic parameters and input data on the water quality output state variables. Hence, the influence of the morphological status of the river (described by the hydrodynamic parameters) and the hydrological condition of the basin catchment (described by the discharges of the tributary and most-upstream boundary) on the water quality (e.g. dissolved oxygen content and chlorophyll-a concentrations) can be determined. A literature search indicates that uncertainty analyses have been carried out on a number of water quality modelling studies, such as the transport of organic (Giri et al., 2001) and inorganic (Carroll and Warwick, 2001) pollutants, simple Streeter-Phelps oxygen dynamics (Maihot and Villeneuve, 2003), phytoplankton - nutrient dynamics using QUAL2E (McIntyre et al., 2003) and the Biebrza river model (van der Perk and Bierkens, 1997) and solute transport (Whelan et al., 1999). However, all these modelling exercises were run assuming steady state flow and concentrated on the uncertainty in the water quality parameters, not on hydrodynamic parameters. Sincock et al. (2003) conducted an uncertainty analysis with hydrodynamic parameters in a river water quality model under unsteady flow conditions but they used a

simplified routing approach based on the kinematic wave with a lag-cascade model extension. The parameters that they incorporated in their analysis were the storage coefficient and a corresponding flow exponent. The model used in this study, DYNHYD, a module in the WASP5 simulation package (Ambrose et al., 1993), uses the full dynamic wave equation and Manning's equation to simulate river flow for which the roughness and weir overflow coefficients become important for parameter uncertainty. These parameters better describe the morphological condition of the river.

A second aim of this research is to quantify another important source of uncertainty in river water quality modelling, model structure. In this study, the structure of the model is understood to be the algorithms and equations used to describe and calculate the processes (Radwan et al., 2004). Structural uncertainty is very difficult to quantify and only a few attempts are found in the literature. For example, Engeland et al. (2005) calculated both total and parameter uncertainty in a hydrological model and found that "the uncertainties in the simulated stream flow due to parameter uncertainty are less important than uncertainties originating from other sources". Reference is made to model structure as one of the sources of uncertainty but it was not explicitly calculated. Brugnach (2005) conducted a sensitivity analysis of the processes in complex ecological models. The variation in parameter sets determined the uncertainty of each process. Todd et al. (2001) investigated the effect different demographic equations have on calculating the population dynamics of the eastern barred bandicoot, a small marsupial endemic to south-eastern Australia. However, each equation has a different set of parameters which needs to be calibrated; hence a strict division of parameter and structural uncertainty cannot be quantified.

2. METHODS

The modelling package used for the uncertainty analysis is WASP5 (Water quality Analysis Simulation Program - Version 5), developed by the US Environmental Protection Agency (Ambrose et al., 1993), was used for the simulation of the Saale's water quality and consists of three models for the water quality simulation of a water body: DYNHYD (hydrodynamics), EUTRO (dissolved oxygen, phytoplankton and nutrient dynamics) and TOXI (sediment and micro-pollutant transport). The models of the WASP5 package were embedded in the HLA (High Level Architecture) coupling platform in order to improve interactive transfer of information between the models (see Lindenschmidt et al. (2005b) for details).

2.1 DYNHYD

"The hydrodynamic model solves one-dimensional equations describing the propagation of a long wave through a shallow water system while conserving both momentum (energy) and volume (mass). The equation of motion, based on the conservation of momentum, predicts water velocities and flows. The equation of continuity, based on the conservation of volume, predicts water heights (heads) and volumes" (Ambrose et al., 1993). These two equations constitute the St. Venant equations and are integrated numerically on a discretized network of the river course using a finite difference "link-node" approach. Links solve the momentum equation outputting the required velocities and discharges and nodes solve the continuity equation processing the water levels and volumes of each discretization unit. The Manning equation which relates flow velocity U with water levels H is used to close the solution:

$$U = \frac{R^{2/3}}{n}\sqrt{\frac{\partial H}{\partial x}}$$

where: n - Manning's roughness coefficient which characterises the roughness of the river bottom and is used for calibration; R -hydraulic radius (cross-sectional area / wetted perimeter); x - distance along longitudinal direction; $\partial H / \partial x$ - energy gradient.

Discharge Q over a weir has been included in the model by Warwick et al. (1999) and is based on the Poleni equation, which simplified reduces to:

$$Q = \alpha \cdot b \cdot h^{1.5}$$

where: α - weir discharge coefficient which ranges between 1.6 and 2.0 for the weirs along the Saale and is used for calibration; b -breadth of weir crest; h - height between weir crest and water level upstream of the weir.

2.2 EUTRO

This model simulates the oxygen balance and phytoplankton and nutrient dynamics of a water body using eight state variables and their interactive processes in various combinations and complexities. Processes include reaeration, oxidation of carbonaceous matter in the water column and bottom sediments, mineralization and

settling of organically-bound nutrients (phosphorus and nitrogen), nitrification and denitrification, and phytoplankton growth and losses (respiration, death, grazing, settling). A Petersen matrix of the most complex model configuration is included in Lindenschmidt (2006) showing the dependencies between state variables and process stoichiometric yields.

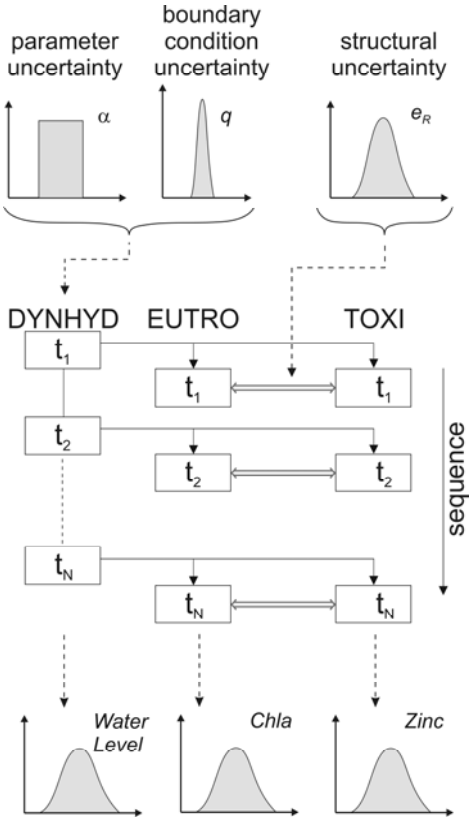

Figure 1. Model system in the HLA environment for Monte-Carlo simulations (modified from Lindenschmidt et al., 2005b).

2.3 TOXI

This model was implemented to simulate the transport of suspended solids, salts and heavy metals. A combination of three dissolved and three particulate substances can be modelling with varying degrees of complexity. Salts were modelled as conservative substances hence, no reaction terms are required for their transport. The transport of suspended solids requires additional sink and source terms to describe the movement of particles to and from the bottom sediments. Settling, deposition and resuspension rates are described by velocities and surface areas. Diffusion of dissolved substances from the bottom sediments into the water column is driven by the gradient between substance concentrations found in the sediments and in the overlying water and its rate is controlled by a diffusion coefficient. Sorption processes must also be included in the reaction term when the transport of heavy metals is simulated. The process is described using a partition coefficient K_D which represents the fraction of dissolved and particulate fractions of the heavy metals in relation to the concentration of suspended solids.

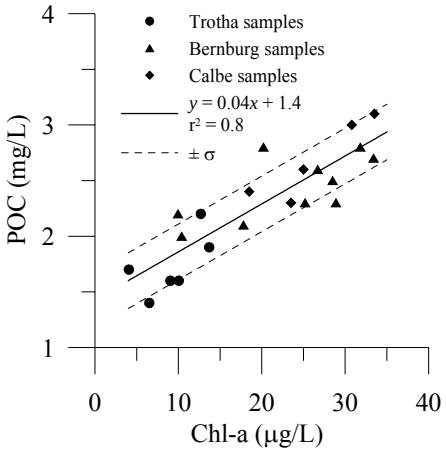

Figure 2. Correlation between chlorophyll-a *Chla* and particulate organic carbon *POC* (unpublished data)

2.4 MOCA (Monte Carlo Analysis)

Figure 1 shows how the MOCA is implemented for interactive communication between the WASP5 models using the HLA environment. Shown is the variation in the hydrodynamic parameters, boundary conditions and a coefficient for the coupling equation between EUTRO and TOXI (structure, described next).

Coupling EUTRO and TOXI together in the HLA environment allows ease of interactive communication between the two models. Chlorophyll-a concentrations *Chla* correlate well with particulate organic carbon *POC* content in the water Figure 2 and brings forth the structure for the EUTRO → TOXI coupling using the equation:

$$POC = 0.04 \cdot Chla + 1.4$$

By dividing *POC* with the concentration of suspended sediment *SS* simulated in TOXI, the weight fraction of organic carbon in suspended matter *foc* is obtained (Ambrose et al., 1993):

$$foc = \frac{POC}{SS}$$

In this study a dependence of K_D on the fraction of organic carbon f_{OC} of the particulate matter was used:

$$K_D = f_{OC} \cdot K_{OC}$$

where K_{OC} is a constant and represents the organic carbon partition coefficient and is calibrated for each heavy metal separately. The partitioning of heavy metals in the water can now occur between the dissolved and particulate phases.

Information is also transferred from TOXI to EUTRO. In the original WASP5 version, the extinction coefficient K_E of light passing through the water column is a constant parameter implemented for each discretized unit of the modelled river. With the communication between TOXI and EUTRO, K_E can now vary depending on *Chla* (µg/l) and phytoplankton biomass *Phyto* (mg/l) computed in EUTRO and *SS* (mg/l) calculated in TOXI (equation modified from Schöl et al. (2002)):

$$K_E = 0.052 \cdot (SS - Phyto) + 0.013 \cdot Chla + 1.06$$

3. STUDY SITE AND MODEL SETUP

The modelling exercise was conducted for the Saale River. It has a total length of 413 km but only the lower reach of 90 kilometres was considered in this study (see Figure 3). The investigated course was discretized using 186 segments, each approximately 500 m in length. Cross-sectional profiles every 100 m along the river were available from which initial hydraulic radii and segment water volumes (calculated from mean water levels) of each segment were determined.

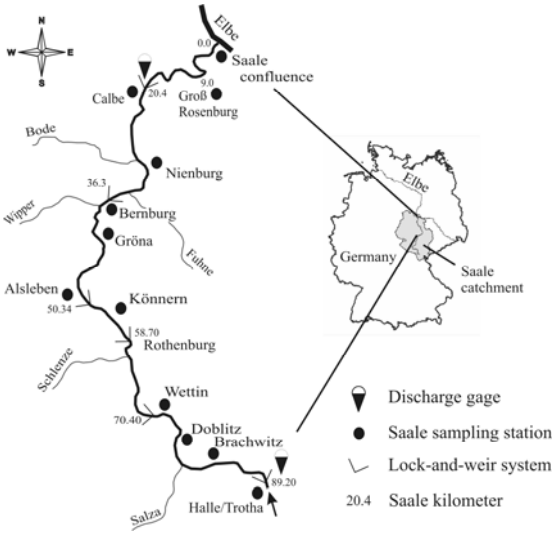

Figure 3. Lower reach of the Saale River.

The river has been heavily modified by channelling and the construction of locks and weirs. The mean discharge of this river course varies from 99 m^3/s at Trotha to 115 m^3/s at Calbe. Five tributaries, Salza, Schlenze, Wipper, Fuhne and Bode, drain into the lower Saale of which the Bode is the most significant in regards to water volume. Many wastes from industrial plants and WWTPs (wastewater treatment plants) are discharged into the Saale. The WWTPs have all been extended with facilities for secondary and tertiary treatment, reflected in the decreasing trend in nitrogen loading.

Two data sets sampled along the lower Saale and its tributaries were available for the calibration (sampled 5.–18. June 2001) and validation (sampled 8.–10. September 2003) of the modelling system.

4. RESULTS AND DISCUSSION

Calibration and validation of the models have been carried out in previous studies: hydrodynamics (von Saleski et al., 2004), eutrophication (Lindenschmidt, 2006) and transport of sediments and inorganic pollutants (Lindenschmidt et al., 2006).

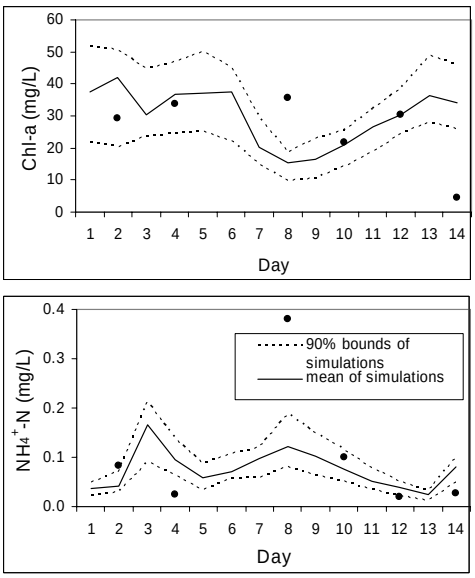

Figure 4. Measured and range of simulated chlorophyll-a and ammonium nitrogen at Groß Rosenburg for the 5. – 18. June 2001 sampling campaign (source: Lindenschmidt et al., 2005a).

4.1 Effect of Hydrodynamic Parameter Uncertainty on Water Quality Variables

The DYNHYD → EUTRO/TOXI coupling was used for this first part of the investigation. Figure 4 shows the results of the MOCA analysis for three eutrophication variables, *Chl-a* and ammonium NH_4^+-*N* based on the uncertainty of the hydrodynamic parameters α and *n*. The 5% and 95% percentiles stem from a Monte Carlo analysis (MOCA) of 2000 simulations extracting α (weir

discharge coefficient) and *n* (roughness coefficient) randomly from a normal probability distribution for each 14-day simulation. The variation in the output distributions increased with distance along the flow direction of the river course.

The simulations of these two water quality state variables were the most difficult to fit to sampled data. An improvement in the fit was attained from the pool of MOCA runs in which the hydrodynamic parameters were considered. The first peak of NH_4^+-*N* at Day 3 is due to a surge load at Day 1 into the most upstream portion of the reach at Halle. Sampling was carried out at most every two days and, unfortunately, an observation of the surge at Groß Rosenburg was missed. The surge is, however, verified at other upstream sampling points along the river. The second peak on Day 8 is due to the mineralization of particulate organic nitrogen *ON* produced by the strong algal growth prior to this day. The growth of phytoplankton is also reduced so that fewer nutrients are being taken up by the algae.

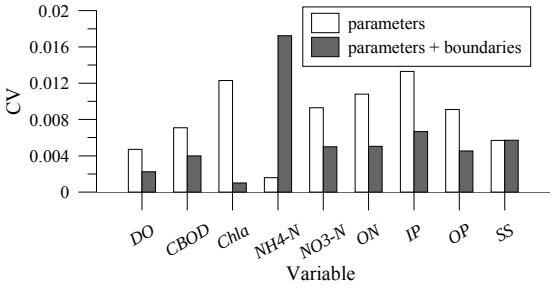

Figure 5. CVs of variables using parameter-only (α and *n*) and parameter + boundary MOCAs uncertainty at Calbe (modified from Lindenschmidt, 2005a, 2005b).

4.2 Effect of Boundary Condition Uncertainty on Water Quality Variables

Surprisingly, all CV (coefficient of variation) at the lower gage at Calbe halved for all the variables with the exception of NH_4^+-*N*, whose CV increased by a factor of ten (see Figure 5). This is primarily due to the input of the Bode River, the largest and most immediate upstream tributary from the Calbe weir, which has a diluting effect of the river water since its concentrations of most of the sampled substances are lower than in the Saale River. The exception is NH_4^+-*N*, which has concentrations up to five times of those found in the Saale. Hence, dilution of substances will decrease the deviation of the simulated output distributions; concentrating will increase the deviation. For further details see Lindenschmidt et al. (2005a, 2005b). The variation in *Chla* decreased substantially, which is not the case for the upper gage.

4.3 Model Structure Uncertainty

Here we describe the MOCA using parameters, boundaries and model structure. The first two are also included to give more variations in the MOCA and to compare the difference between structural uncertainty and uncertainty propagating from parameters and boundary conditions through the modelling system.

A MOCA was carried out for zinc using three configurations of the source of uncertainty: i) parameters only, ii) parameters and boundaries and iii) parameters, boundaries and model structure. TOXI results are summarized in Figure 6 for the Calbe station which shows that the most significant uncertainties for the dissolved and particulate zinc fractions (*ZnD* and *ZnP*) are found in the model system structure, followed by the parametric uncertainty and to a minimal extent, the uncertainty stemming from the boundary conditions. The uncertainty for total zinc *ZnT* stems solely from the parameter sources and is not influenced by boundary or structural uncertainty.

Structural uncertainty using the equation for the TOXI → EUTRO coupling had less effect, due to the equation's low sensitivity on EUTRO variables. In this case, the source of uncertainty is in the following order of decreasing significance: parameter, boundary and structure. Comparison of the uncertainty ranges indicates that parameters and boundary conditions are a more important source of uncertainty for EUTRO variables whereas model system structure is a significant source of uncertainty for TOXI variables.

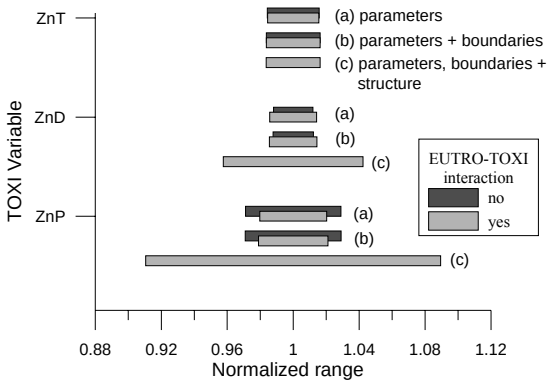

Figure 6. Ranges from MOCAs normalized to the mean for total zinc *ZnT* and its dissolved *ZnD* and particulate *ZnP* fractions (source: Lindenschmidt, submitted).

5. CONCLUSIONS

As is the case for parameters and boundary condition data, structural uncertainty becomes relevant when the corresponding equations react sensitive to the output results. In our case, uncertainties propagating through a chain of models in a modelling system can be accentuated. This was particularly the case for the process of heavy metal sorption to suspended sediments, not so for light extinction limiting phytoplankton growth. More modelling examples and research are required before generalisations can be made.

6. REFERENCES

Ambrose, R.B., T.A. Wool, and J.L. Martin, The Water Quality Simulation Program, WASP5: model theory, user's manual, and programmer's guide, U.S. Environmental Protection Agency, Athens, GA, 1993.

Brugnach, M., Process level sensitivity analysis for complex ecological models, *Ecological Modelling 187*, 99-120, 2005.

Carroll, R.W.H., and J.J. Warwick, Uncertainty analysis of the Carson River mercury transport model, *Ecological Modelling 137*, 211-224, 2001.

Engeland, K., C.-Y. Xu, and Gottschalk, L., Assessing uncertainties in a conceptual water balance model using Bayesian methodology, *Hydrological Sciences Journal 50*(1), 45-63, 2005.

Giri, B.S., I.A. Karimi, and M.B. Ray, Modeling and Monte Carlo simulation of TCDD transport in a river, *Water Research 35*(5), 1263-1279, 2001.

Lindenschmidt, K.-E., The effect of complexity on parameter sensitivity and model uncertainty in river water quality modelling, *Ecological Modelling 190*(1-2), 72-86, 2006.

Lindenschmidt, K.-E., K. Poser, and M. Rode, Impact of morphological parameters on water quality variables using a computer modelling simulation of a regulated, lowland river, *Water Science and Technology 52*(6), 187-193, 2005a.

Lindenschmidt, K.-E., J. Rauberg, and F. Hesser, Extending uncertainty analysis of a hydrodynamic – water quality modeling system using High Level Architecture (HLA), *Water Quality Research Journal of Canada 40*(1), 59-70, 2005b.

Lindenschmidt, K.-E., R. Wodrich, and C. Hesse, The effects of scaling and model complexity in simulating the transport of inorganic micro-pollutants in a lowland river reach, *Water Quality Research Journal of Canada 41*(1), 24-36, 2006.

Mailhot, A., and J.-P. Villeneuve, Mean-value second-order uncertainty analysis method: application to water quality modeling, *Advances in Water Resources 26*, 491-499, 2003.

McIntyre, N.R., T. Wagener, H.S. Wheater, and S.C. Chapra, Risk-based modelling of surface water quality: a case study of the Charles River, Massachusetts, *Journal of Hydrology 274*, 225-247, 2003.

Radwan, M., P. Willems, and J. Berlamont, Sensitivity and uncertainty analysis for river quality modelling, *Journal of Hydroinformatics 6*, 83-99, 2004.

Schöl, A, V. Kirchesch, T. Bergfeld, F. Schöll, J. Borcherding, and D. Müller, Modelling the chlorophyll a content of the river Rhine – interrelation between riverine algal production and population biomass of grazers, rotifers and zebra musel, Dreissena polymorpha. *Internationaler Revue der gesamten Hydrobiologie 87*(2–3), 295-317, 2002.

Sincock, A.M., H.S. Wheater, and P.G. Whitehead, Calibration and sensitivity analysis of a river quality model under unsteady flow conditions, *Journal of Hydrology 277*, 214-229, 2003.

Todd, C.R., P. Inchausti, S. Jenkins, M.A. Burgman, and M.P. Ng, Structural uncertainty in stochastic population models: delayed development in the eastern barred bandicoot, Perameles gunnii, *Ecological Modelling 136*, 237-254, 2001.

van der Perk, M., and M.F.P. Bierkens, The identifiability of parameters in a water quality model of the Biebrza River, Poland, *Journal of Hydrology 200*, 307-322, 1997.

von Saleski, M., J.J. Warwick, R. Hohmann, and K.-E. Lindenschmidt, Parameterunsicherheit eines hydrodynamischen Flussmodells mit Wehren. *Gas- und Wasserfach – Wasser/Abwasser 145*(5), 310-317, 2004.

Warwick, J.J., D. Cockrum, and A. McKay, Modeling the impact of subsurface nutrient flux on water quality in the Lower Truckee River, Nevada, *Journal of the American Water Resources Association 35*(4), 837-851, 1999.

Whelan, M.J., C. Gandolfi, and G.B. Bischetti, A simple stochastic model of point source solute transport in rivers based on gauging station data with implications for sampling requirements, *Water Research 33*(14), 3171-3181, 1999.