Dispersion coefficients of sewers from tracer experiments

Article *in* Water Science & Technology · February 2005

Source: PubMed

5 authors, including:

Jörg Rieckermann
Eawag: Das Wasserforschungs-Institut des ET…
69 PUBLICATIONS **1,065** CITATIONS
SEE PROFILE

Marc B Neumann
BC3-Basque Centre for Climate Change
50 PUBLICATIONS **528** CITATIONS
SEE PROFILE

Christoph Ort
Eawag: Das Wasserforschungs-Institut des ET…
86 PUBLICATIONS **3,021** CITATIONS
SEE PROFILE

Jacco L Huisman
Eawag: Das Wasserforschungs-Institut des E…
15 PUBLICATIONS **395** CITATIONS
SEE PROFILE

DISS. ETH NO. 16975

UNCERTAINTY ANALYSIS FOR PERFORMANCE EVALUATION AND DESIGN OF URBAN WATER INFRASTRUCTURE

A dissertation submitted to
ETH ZURICH

for the degree of
Doctor of Technical Sciences

presented by
MARC BENJAMIN NEUMANN
Dipl. Ing. ETH
born 04.03.1974
citizen of Zurich and Schwändi (GL)

accepted on the recommendation of

Prof. Dr. Willi Gujer, examiner
Prof. Dr. Urs von Gunten, co-examiner
Prof. Dr. Konrad Hungerbühler, co-examiner

2007

Table of Contents

Abstract

Environmental engineers use mathematical models to design, assess and optimise infrastructure for a wide range of transport, storage and reaction processes. The modelling procedures are associated with considerable uncertainties which are often either ignored or dealt with in an implicit way. This can lead to systematic over- or under-design of infrastructure.

The thesis analyses various aspects of uncertainty for exemplary case studies: combined sewer overflow detention, ozone decay during water purification, prediction of disinfection performance in drinking water treatment, estimation of dispersion coefficients in sewers (co-authored), characterisation of sludge properties in wastewater treatment (co-authored).

The goal of the thesis is to explore the potential of explicitly considering uncertainty involved in the modelling process and assess the feasibility of applying probabilistic procedures to systems from environmental technology. The focus lies on eliciting meaningful uncertainty estimates and not on a further development of statistical methods or on an optimised use of computational resources.

Uncertainty introducing sources are identified and investigated. They consist of model structure uncertainty, parameter uncertainty, errors in sampling, errors in chemical analysis and uncertainty due to computational and numerical procedures. Quantitative assessments are performed applying statistical procedures from nonlinear regression, Monte Carlo simulation and sensitivity analysis. Methods are explored to attribute variation in results either to reducible epistemic uncertainty or to irreducible natural variability. Additionally, the transferability of information during up-scaling in engineering design is investigated in the presence of uncertainty. It is demonstrated how diverse approaches of accounting for uncertainty and variability lead to different designs.

For results from an uncertainty analysis to be meaningful, critical factors are identified: the framing of the task, the quality of expert knowledge, the quality of data, the fulfilment of assumptions that statistical tools are based upon and the clarity of the procedure.

A main benefit of performing a thorough uncertainty analysis is that it enables comprehensive model diagnostics and allows one to estimate the accuracy of traditional deterministic approaches. In combination with regional sensitivity analysis (variance decomposition) it is a powerful tool to manage (e.g. reduce) uncertainty and forms the basis for probability based design.

Probabilistic design procedures have various implications for the involved engineers, regulators and system owners. The risk allocation during planning, design and operation phases of infrastructure heavily determines how uncertainty is managed and is a topic that should be further explored.

Zusammenfassung

Umweltingenieure nutzen mathematische Modelle für die Bemessung, Leistungsbeurteilung und Optimierung von Infrastrukturen der Siedlungswasserwirtschaft welche für den Transport und Speicherung von Wasser und Stoffen, sowie für Umwandlungsprozesse eingesetzt werden. Die Modellierung dieser Prozesse ist mit Unsicherheiten verbunden, die häufig entweder gar nicht oder nur implizit mitberücksichtigt werden. Dieser Umstand kann dazu führen dass Anlagen unter- oder überdimensioniert sind.

Die Dissertation untersucht unterschiedliche Aspekte von Unsicherheit anhand von mehreren Fallstudien: Bemessung von Regenüberlaufbecken, Charakterisierung des Ozonzerfalls in der Ozonung von Seewasser, Vorhersage der Desinfektionsleistung einer Trinkwasser-Aufbereitungsanlage, Untersuchung der Dispersion in Kanalisationen (Mitverfasser), Charakterisierung von Schlammeigenschaften in der Abwasserreinigung (Mitverfasser).

Das Ziel der vorliegenden Arbeit ist es die Möglichkeiten der expliziten Berücksichtigung von Unsicherheit abzuschätzen sowie das Potential von wahrscheinlichkeitsbasierten Bemessungsansätzen in der Umwelttechnologie zu untersuchen. Im Mittelpunkt dieser Arbeit steht die Bemühung aussagekräftige Angaben zur Unsicherheit zu machen. Die Weiterentwicklung von statistischen Methoden, oder eine optimierte Nutzung der Rechenleistung stehen nicht im Vordergrund.

Die verschiedenen Quellen von Unsicherheit werden identifiziert und untersucht. Diese umfassen die Unsicherheit über die Angemessenheit der Modellstruktur, Unsicherheit über Parameterwerte, Fehler in den Probenahmen, Fehler in der Messanalytik und Unsicherheiten aufgrund numerischer Effekte sowie der programmtechnischen Umsetzung. Statistische Methoden aus den Bereichen der nichtlinearen Regression, der Monte Carlo Simulation und der Sensitivitätsanalyse werden angewandt. Methoden werden vorgeschlagen um die Streuung in Resultaten die aufgrund reduzierbarer Unsicherheiten auftritt von solcher zu unterscheiden deren Ursprung auf die Variabilität natürlicher Prozesse zurückzuführen ist. Zusätzlich wird die Übertragbarkeit von Information bei der Hochskalierung von Reaktorsystemen in Gegenwart von Unsicherheit untersucht. Es wird gezeigt wie Ansätze die Unsicherheiten unterschiedlich einbeziehen zu unterschiedlichen Dimensionierungen führen.

Um aussagekräftige Unsicherheitsangaben zu erhalten werden kritische Erfolgsfaktoren identifiziert: Eine klare Eingrenzung der Fragestellung, die Qualität des Expertenwissens, die Qualität von Daten, die Einhaltung von Annahmen auf denen die Anwendung der statistischen Methoden fundiert ist, sowie die Klarheit in der Durchführung der Analyse.

Unsicherheitsanalysen ermöglichen eine umfassende Modelldiagnose und erlauben es die Genauigkeit von traditionellen deterministischen Modellen abzuschätzen. Kombiniert mit Methoden der regionalen Sensitivitätsanalyse (Varianzbasierte Masse) bilden sie leistungsfähige Werkzeuge um mit Unsicherheiten umzugehen (z.b. zu reduzieren) und bilden eine Grundlage für risikobasierte Bemessung.

Probabilistische Ansätze haben verschieden Auswirkungen auf die beteiligten Ingenieure, Aufsichtsbehörden und Eigentümer. Die Verteilung von Risiken in Planung und Betrieb bestimmt wesentlich den Umgang mit Unsicherheiten und ist ein Thema das in Zukunft vertieft untersucht werden sollte.

Acknowledgements

From Willi Gujer I learnt about the nature of scientific exploration, the importance of striving for objectivity and the joy obtained from insights when iteratively removing layers of ignorance.

Urs von Gunten provided genuine support with his encouraging attitude and his background in the chemical sciences. I am grateful to Konrad Hungerbühler for offering his expertise as external co-examiner.

I am thankful to Peter Reichert for having been able to learn his rigorous statistical approaches to mathematical modelling.

The urban water engineering group of Eawag was a platform of many fruitful interactions allowing us to support and encourage each other. Thank you again to Jörg Rieckermann, Helge Daebel, Oliver Kracht, Damian Dominguez and Markus Gresch.

I very much enjoyed the company of the PhD students from the other Engineering groups during the Doktorandenseminare and the sabbaticals on Rigi and in Linescio (Reto Manser, Stefano Gianella, Thomas Hug, Christian Abegglen, Silvana Velten, Doris Jermann and Maryna Peter).

My roomies Mariska Ronteltap, Christoph Ort and Dongbin Huang were great fellow hikers to have on the long and winding PhD path.

Workmates who have become friends are constant reminders of the good times at Eawag.

Finally I would like to thank my family in England and Switzerland for their continuous support during the PhD, especially my parents Helen and Fritz and my brother Matthew.

Introduction

Introduction

In the past decades models applied for design and performance analysis in the field of environmental engineering have developed from empirical, heuristic approaches to more sophisticated state space representations characterising storage, transport and reaction processes. Natural scientific system descriptions from hydromechanics, aquatic chemistry and biology have been integrated into the field. The uncertainty involved in the model results however is seldom directly acknowledged. The challenges that arise when explicitly considering uncertainty are the central theme of the thesis. The main contribution lies in exploring and differentiating the involved sources of uncertainty, in their appropriate quantification and the assessment of their relative importance in driving uncertainty of model results. The case studies focus on different aspects of uncertainty analysis and involve combined sewer overflow tanks, reactors for the disinfection and pollutant removal in drinking water treatment, dispersion in sewer pipes, and the analysis of sludge properties in waste water treatment.

UNCERTAINTY ANALYSIS

Uncertainty analysis as a prerequisite for decision making

In scientific applications the assessment of uncertainty is prerequisite for hypothesis testing, i.e. stating that an effect is significant at a certain level of confidence. In engineering practice uncertainty analysis serves as basis for taking actions and making investment decisions. Uncertainty estimates are a basic requirement when applying quantitative risk analysis (Faber and Stewart, 2003) or multi attribute utility theory (Eisenführ, 1999; Clemen and Reilly, 2001). In this context Reckhow (1994) asks “Why are simulation models and scientific assessments, which are known to be uncertain, used for management decisions without consideration of the “goodness” (uncertainty) of the information?” The possible reasons he lists include: Limits on resources, lack of training in probability and statistics, engineer perceived to have failed when the true level of uncertainty is publicly acknowledged and that it is unclear if decisions are improved by considering uncertainty. A recent study on contaminated soil highlights the significance of uncertainty in environmental engineering applications: Remediation costs ranged from $7.5M to $55M due to uncertainty about the octanol - water partition coefficient K_{OW} (Linkov *et al.*, 2005).

Sources of uncertainty in modelling

The types of uncertainty considered in the thesis are summarised by the following four sub-categories:

Uncertainty about model structure

Mathematical models are necessarily simplifications of reality. There is uncertainty involved in selecting the appropriate variables and processes, in specifying a reasonable level of aggregation and the spatial and temporal resolution and in setting the boundary conditions. Since the 1990’s this controversial topic has received increased attention in the statistical community (Chatfield, 1995; Draper, 1995).

Uncertainty about values of parameters

Parameter uncertainty arises when eliciting subjective expert knowledge (e.g. asking an expert to give an estimate for a decay rate constant of a pollutant at a specific location) (Ayyub, 2001)

or originates in statistical inference where random errors of observations are mapped to parameters (Bates and Watts, 1988).

Uncertainty due to errors in data
Measurements are subject to systematic and random errors. Systematic measurement errors are due to non-perfect calibration of the measurement device (Thomann Haller, 2002) whereas random errors arise from statistical fluctuations in the measured data due to the precision limitations of the measurement device. An additional source of bias in data arises due to samples being non-representative (Gy, 1999).

Uncertainty due to numerical, computational and statistical procedures
Errors may arise due to inappropriate numerical and computational implementation of mathematical formulations. Procedures of statistical inference require the selection of objective functions (e.g. least squares) and minimization algorithms (e.g. simplex).

In the scientific literature various suggestions can be found on the classification of uncertainty (Carstensen *et al.*, 1997; Ayyub, 2001).

In the course of the entire thesis the distinction between epistemic uncertainty and variability is made. Uncertainty (often termed "epistemic uncertainty") arises due to incomplete knowledge and is therefore reducible whereas variability (often termed "aleatory uncertainty") characterises the spread of values from a well specified population and is irreducible (Kelly and Campbell, 2000). Barnett and O'Hagan (1997) introduce the concept of a "statistically verifiable ideal standard" which acknowledges variability (frequency of exceedance of the standard) and uncertainty (probability of exceedance of the standard).

Characterisation of uncertain information
The most common approaches for characterising uncertain information are the frequentist and the Bayesian schools which are both founded on probability theory. The frequentist description of uncertainty is based on the notion of replication of hypothetical experiments (Fisher, 1959) and is the most widely used approach in the natural sciences. Confidence regions of parameters are obtained from error information of the estimator (e.g. least squares). In the Bayesian context (O'Hagan, 1994) parameters are characterised as random variables which are successively updated. Other methods offering alternative descriptions of uncertain knowledge have been developed such as imprecise probabilities (Walley, 1991), interval arithmetic (Alefeld and Herzberger, 1983), evidence theory (Dempster, 1967), Fuzzy logic (Zadeh, 1965). As none of these alternative methods have established themselves and the challenge of reliable quantification of the uncertainty bounds is similar for the different approaches this study utilizes the traditional approaches founded on probability theory.

Components of uncertainty analysis in modelling
The goal of *parameter estimation* is to gain information on model parameters from data. Both the frequentist and Bayesian approaches to parameter estimation make use of the maximum likelihood principle (Fisher, 1925). *Identifiability analysis* investigates how well parameters (or sets of parameters) can be estimated due to collinearity (Walter *et al.*, 1987). *Sensitivity analysis* assesses how model results are affected by changes of model parameters (Saltelli, 2000; Frey and Patil, 2002). *Uncertainty propagation* examines how uncertainty about the parameters affects the uncertainty associated with the model results. *Model testing* is concerned with validation of a given model whereas *model structure selection* is required when competing models are available.

Uncertainty analysis in modelling of environmental systems

In the 1960's the structural engineering community increasingly integrated statistics and probability theory into their field and developed an extensive range of methods for failure and reliability analysis such as the first- and second- order reliability methods (FORM, SORM) (Benjamin and Cornell, 1970; Ang and Tang, 1975). In 1979 a meeting was held at the International Institute for Applied Systems Analysis (IIASA) in Laxenburg (Austria) addressing problems associated with uncertainty and forecasting of water quality. The proceedings of the conference (Beck and van Straten, 1983) contain articles by scientists from environmental sciences, groundwater management and hydrology (B. Beck, P. Young, G. Hornberger, R. Spear, L. Somlyódy, R. Gardner, R. O'Neill, K. Reckhow, D. McLaughlin). Taking an approach with it's roots in control theory Beck (1987) reviews much of this work. More recent studies combining uncertainty-, sensitivity-, and identifiability analysis for environmental modelling can be found in (Brun *et al.*, 2001; Dochain and Vanrolleghem, 2001; Omlin *et al.*, 2001; Brun *et al.*, 2002; Reichert, 2002).

Methods for risk analysis, optimization, artificial intelligence, probability theory, statistics, and decision theory are being applied over different scientific fields such as ecology, chemical engineering, climate sciences and socioeconomics making a comprehensive review of uncertainty a difficult task. Indeed the 21st century can be viewed as a paradigm transition from certainty to uncertainty across a wide range of areas such as physics, mathematics, philosophy, art and literature (Peat, 2002).

Why is uncertainty analysis still an issue?

Society increasingly demands transparency concerning the costs, benefits and risks of public infrastructure (Beck, 1986; Slovic, 1987). As projects are under sharp scrutiny and more stakeholders are involved in the decision processes the explicit acknowledgement of the involved uncertainties in planning, design and maintenance of these infrastructure systems becomes a necessary requirement. With increased awareness towards environmental and health related effects more elaborate quality standards are developed closer at the actual endpoints. Examples can be found across all sectors of urban water management: According to new European legislation sewers should be designed to obtain optimal tradeoffs between costs and expected damage. In contrast to the previous procedures involving the normative approach of the "rational method", the new criteria require the hydrodynamic modelling of sewers and an estimate of damages resulting from flooding caused by surcharge of manholes (EN 752). Also water quality modelling is increasingly directed towards immission based criteria as is intended within the European Water Framework Directive (Pollard and Huxham, 1998) or as proposed for future guidelines in Swiss storm water management (Rossi *et al.*, 2005). An example from water supply are the recently introduced disinfection criteria for inactivation of *Cryptosporidium parvum* oocysts in the US (USEPA, 2006). Accross all fields of environmental engineering there is increased use of computer modelling to prove compliance, either for existing systems where the criteria cannot be directly assessed or for design studies. Modern sensors applied in water infrastructure produce a large amount of data from which information can be extracted for performance analysis, optimisation and redesign of systems. The integration of natural scientific system descriptions into the engineering community has lead to modelling being an integrated part of the practitioners work (Gujer, 2006). The growth of computer resources and the availability of commercial software have opened up new possibilities for modelling engineering systems. The involved uncertainty in the model results and its consequences however is seldom directly acknowledged. Engineers and regulators lack quantitative tools to deal with these issues.

What are remaining challenges in uncertainty assessment?
Challenge problems of epistemic uncertainty were identified at a workshop organised by the Sandia National Laboratories (Helton and Oberkampf, 2004). A main topic was to find appropriate descriptions for parameter uncertainty. Experts with different scientific backgrounds solved two example problems using different methods for representation of epistemic uncertainty (Oberkampf *et al.*, 2004). There was disagreement between probabilists (arguing that frequentist and Bayesian frameworks are the only appropriate way of dealing with epistemic uncertainty) and opponents proposing the use of alternative forms of uncertainty representation (Ferson *et al.*, 2004). Specific challenges for real world engineering systems were identified: It was acknowledged that the information typically available for the analysis of an engineering system is a) developed for different facilities or sites, b) produced by multiple experimental programs with different experimental techniques, objectives and levels of resolution, c) extracted from analogues, d) gathered on different temporal or spatial scale than needed. It was recognized that the entire chain of uncertainty analysis for full scale systems requires deeper scientific exploration to obtain a meaningful and useful representation of uncertainty: i) elicitation and conversion of available information into a mathematical structure, ii) aggregation of information from multiple sources into a single representation of uncertainty iii) propagation of uncertainty to the model results, iv) presentation, interpretation and communication of results, v) sensitivity analysis.

RESEARCH QUESTIONS

Research Contribution

In this thesis comprehensive uncertainty analyses are performed for selected systems from the field of urban water infrastructure. The goal is threefold:

I: Identify gains and limitations when applying statistical methodologies to models of full scale urban water systems.

II: Enhance transparency and quality of information by specifying the uncertainty associated with model results.

III: Assess the feasibility and value of explicitly considering uncertainty in the assessment and design of urban water infrastructure.

This thesis limits the analysis of uncertainty to the application of models for performance analysis and design of urban water systems. The systems range from hypothetical systems to laboratory reactors, pilot and full scale systems.

General Research Questions

A list of general research questions was set in the outset of the thesis:

1. How should different sources of uncertainty be classified and separated for the modelling of urban water systems?
   (Chapter 1, 2, 3)

2. How are uncertainty ranges for parameters of full scale systems obtained?
   (Chapter 2, 3)

3. How does model structure uncertainty manifest itself and can it be quantified?
   (Chapter 2, 4, 5)

4. How is the reliability of a model prediction assessed?
   (Chapter 3, 5)

5. How should results from uncertainty analysis be presented and communicated?
   (Chapter 1, 3)

6. Is the explicit consideration of uncertainty (e.g. for probability based design) feasible in urban water engineering?
   (Chapter 1, 3, 5)

7. What are ideal legislative standards for environmental quality or human health when proving compliance with models? *(Chapter 1, 3)*

THESIS OUTLINE AND SUMMARY OF RESULTS

Chapter 1 "Legal Criteria, Variability and Uncertainty in the Design of Combined Sewer Overflow Detention" shows how the current trend of environmental criteria being set at higher levels of complexity increases the required modelling efforts for combined sewer overflow structures (CSO). It is demonstrated how different approaches of accounting for uncertainty and variability lead to different designs. The importance of differentiating between parameter uncertainty and natural variability is demonstrated. The model output is visualised as a 2-dimensional design surface with which the compliance is assessed with respect to temporal frequency (due to yearly rainfall variability) and probability (due to parameter uncertainty) of meeting legal criteria. It is shown that explicitly considering uncertainty in the design stage augments the degrees of freedom of the analysis compared to traditional approaches and that accountability of the engineer increases. The study is performed for a hypothetical urban catchment and the uncertainty ranges of the parameters are based on expert knowledge from real catchments.

The focus of **chapter 2 "Sources of Parameter Uncertainty in Predicting Treatment Performance – The Case of Pre-ozonation in Drinking Water Engineering"** is set on investigating the factors that determine parameter uncertainty when applying models to a full scale system from environmental engineering. The analysis is performed for ozonation of surface water, a technology applied in drinking water treatment for disinfection and degradation of micro pollutants. The pseudo 1st order rate constant of ozone decay k_{O3} is characterised as a time dependent parameter and estimated from data obtained from three experimental setups representing up-scaling stages in engineering design. The uncertainty estimate of the rate constant is found to be dependent on the uncertainty about the model structure, on uncertainty of other model parameters, on systematic errors in sampling and chemical analysis. In the given case study the transferability of information in up-scaling is found to be critical due to the considerable uncertainty associated with the estimation of the rate constant. Optimal strategies for reducing uncertainty when performing on-line estimation of ozone exposure are determined. The results are further applied for the probabilistic assessment of the disinfection performance in chapter 3.

In **chapter 3 "Uncertainty in Prediction of Disinfection Performance"** a micro modelling approach to predict disinfection performance of a full scale reactor is developed based on direct stochastic sampling from density distributions. Comparing the variability of ozone exposures to the variability of lethal ozone doses the inactivation of organisms (*Cryptosporidium parvum* oocysts) is predicted. By merging on-site information about the uncertainty of reactor properties with information from scientific literature about the uncertainty of the disinfection model a meaningful uncertainty estimate of the inactivation level is obtained. The reliability of the prediction is investigated and the relative contribution of the parameters to the output variance is assessed by regional sensitivity analysis. Scenario analysis is conducted to explore robustness of results. The scheme proposed for probabilistic assessment of disinfection reactors is suggested to be a valuable alternative to traditional approaches where safety levels are incorporated within normative procedures.

In **chapter 4 (co-authored) "Dispersion Coefficients of Sewers from Tracer Experiments"** the longitudinal dispersion of sewer pipes is investigated. Dispersion coefficients are estimated for tracer experiments from 37 sewer reaches of different cities applying the Advection Dispersion Equation (ADE). The variability across sewer reaches is found to be small with mean 0.16 m^2s^{-1} and 10% and 90% quantiles of 0.05 m^2s^{-1} and 0.36 m^2s^{-1}. The established database specifying the variability of the dispersion coefficient across many systems is of value for future projects when a stochastic representation of this parameter is satisfactory. Additionally model structure selection is performed with 12 predictive models that require only information on flow and geometry.

Chapter 5 (co-authored) "Towards Estimating Parameter Uncertainty under Model Structure Deficits" deals with a dilemma faced for the sequencing pilot batch reactor of chapter 2: When applying statistical regression to models of environmental engineering system the underlying statistical assumptions are often not fulfilled. Fitting semi-empiric models to data obtained with modern sensor technology at high temporal resolution and precision leads to residuals that are not independent and identically distributed (non-iid). In these cases uncertainty propagation with standard errors leads to underestimation of uncertainty in the model results. Until now this fact has either not been acknowledged or taken into account in an ad-hoc fashion. The study suggests three general strategies for obtaining meaningful error information on parameters in the presence of non-iid residuals: i) artificially enlarging the sum of squares obtained from ordinary least squares regression ("whitening"), ii) reducing the number of original data points ("sub-sampling") and iii) model structure extension.

REFERENCES

Alefeld G. and Herzberger J. (1983) "Introduction to Interval Computations", Academic Press, New York etc.

Ang A.H. and Tang W.H. (1975) "Probability Concepts in Engineering Planning and Design - Volume 1: Basic Principles", John Wiley and Sons, Inc., New York.

Ayyub B.M. (2001) "Elicitation of Expert Opinions for Uncertainty and Risks", CRC, Boca Raton.

Barnett V. and O'Hagan A. (1997) “Setting Environmental Standards - the Statistical Approach to Handling Uncertainty and Variation”, Chapman & Hall, London.

Bates D.M. and Watts D.G. (1988) “Nonlinear Regression Analysis and Its Applications”, Wiley, New York etc.

Beck M.B. (1987). Water-Quality Modeling - a Review of the Analysis of Uncertainty. *Water Resour. Res.*, **23**(8), 1393-1442.

Beck M.B. and van Straten G. (1983) “Uncertainty and Forecasting of Water Quality”, Springer, Berlin.

Beck U. (1986) “Risikogesellschaft auf dem Weg in eine andere Moderne”, Suhrkamp, Frankfurt am Main.

Benjamin J.R. and Cornell C.A. (1970) “Probability, Statistics, and Decision for Civil Engineers”, McGraw-Hill, New York etc.

Brun R., Kuhni M., Siegrist H., Gujer W., and Reichert P. (2002). Practical Identifiability of Asm2d Parameters - Systematic Selection and Tuning of Parameter Subsets. *Water Research*, **36**(16), 4113-4127.

Brun R., Reichert P. and Kunsch H.R. (2001). Practical Identifiability Analysis of Large Environmental Simulation Models. *Water Resour. Res.*, **37**(4), 1015-1030.

Carstensen J., Vanrolleghem P., Rauch W., and Reichert P. (1997). Terminology and Methodology in Modelling for Water Quality Management - a Discussion Starter. *Water Science and Technology*, **36**(5), 157-168.

Chatfield C. (1995). Model Uncertainty, Data Mining and Statistical-Inference. *Journal of the Royal Statistical Society Series a-Statistics in Society*, **158**, 419-466.

Clemen R.T. and Reilly T. (2001) “Making Hard Decisions”, Duxbury.

Dempster A.P. (1967). Upper and Lower Probabilities Induced by a Multivalued Mapping. *Annals of Mathematical Statistics*, **38**(2), 325-&.

Dochain D. and Vanrolleghem P. (2001) “Dynamical Modelling and Estimation in Wastewater Treatment Processes”, IWA, London.

Draper D. (1995). Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society Series B-Methodological*, **57**(1), 45-97.

Eisenführ (1999) “Rationales Entscheiden”. 3. edn., Springer, Berlin.

Faber M.H. and Stewart M.G. (2003). Risk Assessment for Civil Engineering Facilities: Critical Overview and Discussion. *Reliab. Eng. Syst. Saf.*, **80**(2), 173-184.

Ferson S., Joslyn C.A., Helton J.C., Oberkampf W.L., and Sentz K. (2004). Summary from the Epistemic Uncertainty Workshop: Consensus Amid Diversity. *Reliab. Eng. Syst. Saf.*, **85**(1-3), 355-369.

Fisher R.A. (1925). Theory of Statistical Estimation. *Proceedings of the Cambridge Philosophical Society*, **22**, 700-725.

Fisher R.A. (1959) "Statistical Methods and Statistical Inference". 2 edn., Oliver and Boyd, Edinburgh.

Frey H.C. and Patil S.R. (2002). Identification and Review of Sensitivity Analysis Methods. *Risk Anal.*, **22**(3), 553-578.

Gujer W. (2006). Activated Sludge Modelling: Past, Present and Future. *Water Science and Technology*, **53**(3), 111-119.

Gy P. (1999) "Sampling for Analytical Purposes". Repr. edn., Wiley, Chichester.

Helton J.C. and Oberkampf W.L. (2004). Alternative Representations of Epistemic Uncertainty. *Reliab. Eng. Syst. Saf.*, **85**(1-3), 1-10.

Kelly E.J. and Campbell K. (2000). Separating Variability and Uncertainty in Environmental Risk Assessment -- Making Choices. *Hum. Ecol. Risk Assess.*, **6**(1), 1-13.

Linkov I., Ames M.R., Crouch E.A.C., and Satterstrom F.K. (2005). Uncertainty in Octanol-Water Partition Coefficient: Implications for Risk Assessment and Remedial Costs. *Environ. Sci. Technol.*, **39**(18), 6917-6922.

Oberkampf W.L., Helton J.C., Joslyn C.A., Wojtkiewicz S.F., and Ferson S. (2004). Challenge Problems: Uncertainty in System Response Given Uncertain Parameters. *Reliab. Eng. Syst. Saf.*, **85**(1-3), 11-19.

Omlin M., Brun R. and Reichert P. (2001). Biogeochemical Model of Lake Zurich: Sensitivity, Identifiability and Uncertainty Analysis. *Ecol. Model.*, **141**(1-3), 105-123.

Peat F.D. (2002) "From Certainty to Uncertainty", Joseph Henry Press, Washington.

Pollard P. and Huxham M. (1998). The European Water Framework Directive: A New Era in the Management of Aquatic Ecosystem Health? *Aquatic Conservation-Marine and Freshwater Ecosystems*, **8**(6), 773-792.

Reckhow K.H. (1994). Importance of Scientific Uncertainty in Decision-Making. *Environ. Manage.*, **18**(2), 161-166.

Reichert P. (2002) Umweltsystemanalyse - (Environmental Systems Analysis). Lecture Notes. Eawag. Dübendorf.

Rossi L., Krejci V., Rauch W., Kreikenbaum S., Fankhauser R., and Gujer W. (2005). Stochastic Modeling of Total Suspended Solids (TSS) in Urban Areas During Rain Events. *Water Research*, **39**(17), 4188-4196.

Saltelli A. (2000) “Sensitivity Analysis”, John Wiley & Sons, Chichester.

Slovic P. (1987). Perception of Risk. *Science*, **236**(4799), 280-285.

Thomann Haller M. (2002) Datenkontrolle von Abwasserreinigungsanlagen mit Massenbilanzen, Experimenten und statistischen Methoden. Dissertation, ETH Zürich.

USEPA (2006) National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule; Final Rule. Federal Register - National Archives and Records Administration. January 5, 2006.

Walley P. (1991) “Statistical Reasoning with Imprecise Probabilities”, Chapman and Hall, London etc.

Walter E., International Federation of Automatic Control and International Federation of Operational Research Societies (1987) “Identifiability of Parametric Models (Edited, Updated, and Expanded Papers from the 7th Ifac/Ifors Symposium on Identification and System Parameter Estimation; York, July 1985)”, Pergamon Press, Oxford etc.

Zadeh L.A. (1965). Fuzzy Sets. *Information and Control*, **8**(3), 338-353.

Chapter 1

Marc B. Neumann, Christoph Ort, Helge Daebel, Willi Gujer

Legal Criteria, Variability and Uncertainty in the Design of CSO Detention

5th International Conference on Sustainable Techniques and Strategies in Urban Water Management

NOVATECH'2004, Lyon, France

Proceedings, Vol.1, p. 391- 398
ISBN 2-9509337-5-0

Legal Criteria, Variability and Uncertainty in the Design of CSO Detention

Neumann M. B., Ort C., Daebel H., Gujer W.

Swiss Federal Institute for Environmental Science and Technology (EAWAG), and Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

ABSTRACT

The legal criteria for the design of CSO structures can be set at different levels of sophistication. Applying emission based criteria the performances of three design alternatives are compared by using different strategies to account for variability and parameter uncertainty. The importance of separating these two effects is highlighted and the decision processes are explored. Additionally, the consequences of different kind of criteria are discussed with respect to the stakeholders involved.

KEYWORDS

Design Criteria, Uncertainty, Variability, Combined Sewer Overflow, Decision Support, Environmental Regulation

1. INTRODUCTION

The legal requirements for combined sewer overflow (CSO) structures are recently being set closer to the actual “endpoints” in the receiving water using immission based criteria. An engineer designing for compliance of proposed works with such criteria is fundamentally confronted with issues of uncertainty and variability when predicting outcomes with mathematical models.

For impacts of CSO spills Hauger et al. (2002) modelled dissolved oxygen concentrations following a risk-based approach. Kreikenbaum et al. (2002) modelled hydraulic stress and ammonia concentrations in the receiving water specifying frequency and probability of critical events. Concerning the development of immission based criteria Borsuk et al. (2002) discussed a probabilistic approach for water quality prediction in an estuary.

In this paper a case study is presented that deals with the design of CSO detention using emission based criteria. We explore strategies of selecting an appropriate design alternative in the presence of uncertainty and variability. Ongoing questions are raised and discussed: How do different ways of considering uncertainty affect the choice of technology? Which type of criteria is therefore adequate for decision making? What are the implications for stakeholders like engineers, system owners, regulators and the public?

2. CRITERIA FOR CSO PERFORMANCE

Legal criteria for CSO detention are derived from overall goals, such as ecological protection, hygienic safety and aesthetic targets. Three examples of criteria at different levels of sophistication are given below:

- Fixed: $20\ m^3 / ha_{impervious}$
- Emission based: Less than 45 spills/year for a given return period
- Immission based: Concentration-duration-frequency values for NH_3-N

The current Swiss guidelines specify fixed criteria for five different river classes. These guidelines are under consideration, with emission and immission based criteria being discussed for future use.

In the U.K. Urban Pollution Management framework (UPM) the modeller must prove compliance with concentration-duration thresholds for given return periods. These criteria (for dissolved oxygen and un-ionised ammonia) have been derived for three types of ecosystems using ecotoxicological information (Crabtree and Morris, 2002; Blanksby, 2002).

In the presence of fixed criteria, as specified above, only the impervious area needs to be estimated. If emission or immission based criteria are present, then the expected outcomes for design alternatives need to be modelled. These model predictions are strongly influenced by the natural *variability* of the involved physical phenomena. Furthermore, *uncertainty* about model structure and parameters determines the probability of compliance with frequency based criteria.

3. TERMINOLOGY

The need to distinguish between *variability* and *uncertainty* in probabilistic modelling has been recognized e.g. by Kelly and Campbell (2000) or Grum and Aalderink (1999).

Variability is defined as the “true” spread of values (in time or space) of a well-specified population. The spread of these values is not reducible.

Uncertainty on the other hand results from lack of knowledge about the system. Model structure uncertainty includes the uncertainty about the relevance of variables and processes considered. In addition, the choice of model equations and the temporal and spatial resolution fall into this category. Parameter uncertainty is the uncertainty about the values of model parameters. Unlike variability, uncertainty is partly reducible: e.g. through further measurements or deeper investigations into the relevant processes.

The modeller needs to make an informed choice of what he classifies to represent variability or uncertainty depending on the problem being addressed. Not distinguishing between the two may lead to false conclusions and impaired management actions.

Fig. 1 shows qualitatively how variability and parameter uncertainty might affect a predicted outcome in different ways. In both cases, the vertical lines present outputs from deterministic models for different events. These vertical lines show the effect of event-to-event variability on the outcome. In probabilistic modelling these values are not fixed, due to parameter uncertainty. The resulting uncertain outputs are represented as empirical density distributions. In case A the range of predicted outcomes is dominated by variability. Reducing parameter

uncertainty will not significantly reduce the spread of predicted outcomes. In case B parameter uncertainty is the overriding factor. If it were reduced with additional investigations, the spread of predicted outcomes would decrease significantly.

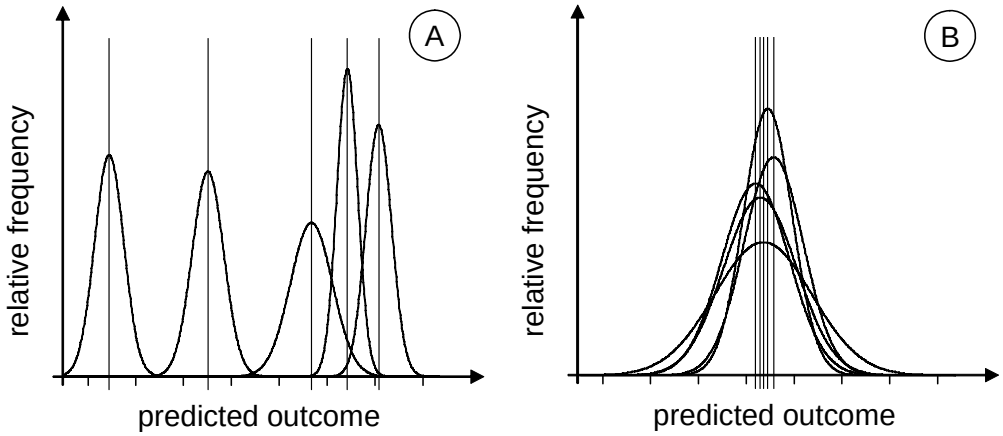

Figure 1. Effects of variability and parameter uncertainty on predicted model outcome. Case A: Variability has a larger influence than parameter uncertainty. Case B: Parameter uncertainty has a larger influence than variability

4. CASE STUDY – SYSTEM

For the purpose of this paper, the case study is intentionally kept simple. The rainfall-runoff process is modelled using a linear reservoir, considering an initial loss and a runoff coefficient. The ongoing flow to the Wastewater Treatment Plant (WWTP) is limited to the double amount of maximum dry weather flow. When the incoming flow drops below 1.5 times the amount of dry weather flow a pump begins to empty the CSO tank with a pumping rate of 0.5 times dry weather flow. Parameter uncertainty is characterised according to expert knowledge. The marginal distributions are specified as normal distributions (see Tab. 1). Covariance for all parameter combinations is zero.

Table 1. Parameter Uncertainty described with Normal Distributions

Parameter	Units	Expected Value	Standard Deviation
Area	ha	20	0.5
Runoff Coefficient	-	0.75 / 0.6	0.05
Storage Coefficient	min	20	2.5
Initial Loss	mm	4	0.5
Average Dry Weather Flow (DWF)	l/s	20	2
Maximum onflow to WWTP (as a factor of DWF)	-	2	0.1

The local 20-year time series of rain data (10 min resolution), which is considered to represent the local rain pattern, expresses temporal variability. The rainfall is not parameterised as in other studies (Grum and Aalderink 1999). Naturally the rainfall time series is also subject to uncertainties. These arise from measurement errors or from the fact that effects of spatial variability are neglected when modelling the catchment with a single linear reservoir and thus

overestimating discharges for events with high rain intensities (Willems and Berlamont 1999). These uncertainties are not quantified in our case study.

5. CASE STUDY – DESIGN ALTERNATIVES AND STRATEGIES

Three design alternatives (see Tab. 2) for CSO detention are compared using the emission based criteria "number of spills per year".

Table 2. Design Alternatives

Alternative	Description	Tank Volume	Runoff Coefficient	Resulting Storage
a	CSO tank (base case)	300 m^3	0.75	2 mm
b	Larger CSO tank	375 m^3	0.75	2.5 mm
c	CSO tank and infiltration sites	300 m^3	0.6	2.5 mm

In the following, we discuss the results of four hypothetical engineers who take variability and parameter uncertainty into account differently when evaluating these alternatives.

Engineer 1: considering neither the effects of variability nor uncertainty
Engineer 2: considering effects of variability
Engineer 3: considering the combined effects of variability and parameter uncertainty
Engineer 4: differentiating between effects of variability and parameter uncertainty

Taking a deterministic approach, Engineers 1 and 2 use the expected parameter values from Tab. 1. Engineers 3 and 4, both following a probabilistic approach, run Monte Carlo simulations by sampling parameter values from the probability distributions specified in Tab. 1.

Engineer 1, who is only in possession of one year of rain data, will be confronted with two difficulties: Depending on the year he has available, he will argue differently as to which alternative is superior (see Fig. 2, left). If only data for the year 1983 is available, he will argue that there is no difference between the three alternatives, whereas if 1988 or 1995 data is available he may conclude that two of the alternatives perform equally well. Secondly, if he uses 1983 as the basis of his design he would heavily underestimate the expected number of spills per year.

Engineer 2 uses the data from 1980 to 1999 (Fig. 2, left). He is able to compare the efficiency of the designs reasonably well as the effects of variability are slightly larger than the effects from parameter uncertainty (Fig. 1, Case A). He can give a frequency based prediction and specify absolute values for the expected differences of the three alternatives.

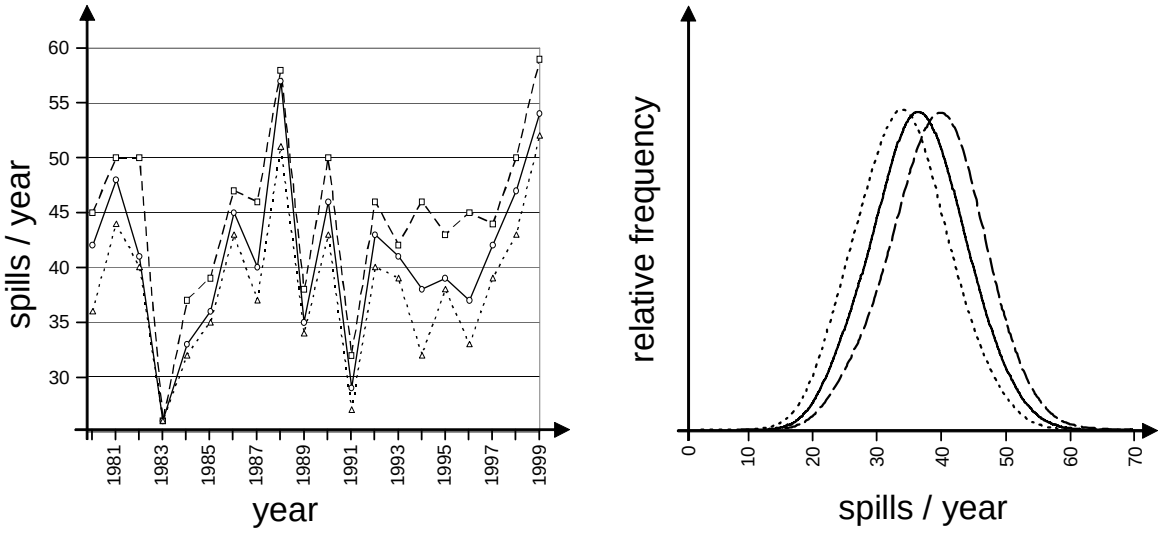

Figure 2. Comparing design alternatives; left: Engineers 1 and 2; right: Engineer 3; Alternatives: a) dashed, b) solid, c) dotted.

Engineer 3, taking an overall view, presents the results by pooling uncertainty and year-to-year rainfall variability (Fig. 2, right). The wide distributions may lead to the false conclusion that, as the overlap of the three output distributions is so large, the three alternatives must perform almost equally well. To make a more qualified interpretation Engineer 3 would need to explore the degree of dependency between the alternatives (Reichert and Borsuk 2003). Because most of the parameters are totally dependent between alternatives and the rain pattern is the same, this would lead to narrow distributions for the differences between alternatives.

Engineer 4 also calculates the combined effects of variability and parameter uncertainty on the outcome, but, unlike Engineer 3, he distinguishes the two aspects in the analysis. It is possible to present the results from the Monte Carlo simulation as a matrix: The columns represent the 20 years and the rows represent the number of Monte Carlo simulations. Each cell value contains the model outcome of "spills per year" resulting from the rainfall of a specific year (e.g. 1992) and a sampled parameter set (e.g. sample no. 121). The columns reflect the effects of year-to-year variability and the rows show the effects of parameter uncertainty. The matrix is sorted consecutively, first within the columns and then within the rows. Plotting the cell values of the sorted matrix (number of spills per year) on the z-axis results in a "design surface" containing all relevant information (Fig. 3). We observe how the effects of year-to-year rainfall variability compare to the effects of parameter uncertainty through exploring the design surface. Engineer 4 is able to give absolute predictions with respect to frequency and confidence of complying with the criteria.

The information contained in the "design surface" is equivalent to the two dimensional cumulative distributions often used for presenting results from embedded (or two dimensional) Monte Carlo simulations e.g. in Frey and Rhodes (1998).

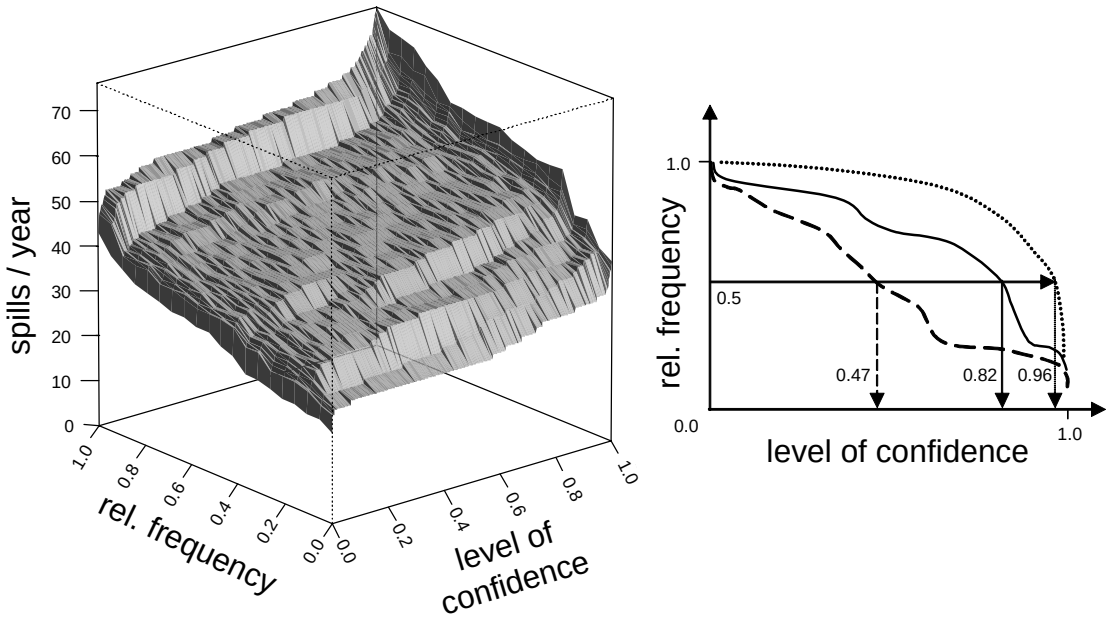

Figure 3. Visualisation of results from Engineer 4; left: design surface of alternative a); right: contours for the criteria "45 spills per year" from the design surfaces of the three alternatives. Alternatives: a) dashed, b) solid, c) dotted.

The horizontal contours for the three alternatives, resulting from the "design surfaces", are explored at the criteria level of "45 spills per year" (Fig. 3, right). Engineer 4 can inform the decision maker that the criteria of "*less than 45 spills in 50% of the years"* will be fulfilled with probabilities of 0.47 for alternative a), 0.82 for alternative b) and 0.96 for alternative c).

The required probability of meeting the criteria must either be set in the legal framework (together with the frequency) or a decision maker must specify it. If the latter is the case, he will choose one of the proposed design alternatives or demand a recalculation to reach his required level of confidence. There is some difficulty in interpreting the derived probabilities. How will the decision maker compare a probability of 0.8 with a probability of 0.6? He should use a formalized approach balancing cost of infrastructure vs. cost of failure according to his risk profile. A decision theoretic framework would structure the analysis, especially when more than one attribute is considered and several technologies are available. The decision maker will then be able to specify the optimal design by maximizing expected utility.

6. DISCUSSION AND CONCLUSIONS

In the case study we explored the effects of variability and uncertainty when evaluating performance of proposed designs with an emission based criteria. Following a deterministic approach, we observed that including information on rainfall variability is a necessity for quantifying performance and choosing among alternatives. However, no statement is possible on the confidence of fulfilling the criteria. Following a probabilistic approach we saw that the effects of variability and parameter uncertainty must be separated to make a qualified interpretation with regard to frequency and probability of fulfilling the criteria. We observed that a probabilistic framework implies a new approach for the involved engineers and decision makers.

The “design surface” derived in such a probabilistic (Fig. 3) approach must be robust, meaning that second-order uncertainty (uncertainty about probabilities) is small. This is still a shortcoming of probabilistic approaches. Possible correlations between parameters must be characterized and marginal distributions chosen appropriately. Model structure must reliably mirror the involved processes at an adequate resolution. Assumptions pertaining to variability like the uniform rain distribution over the catchment, or the representativeness of the 20 year series must hold. One must specify the regions of the design surface for which the assumptions are valid. If these issues are not addressed, then the certainty one pretends to have by presenting probabilistic results may turn out to be very misleading. This is especially the case when using conceptual models like the ones used for assessing CSO performance.

In contrast to our rather simple example of applying emission based criteria, current trends in research and politics demand a decision, based on evaluating the “endpoints” of urban drainage systems, e.g. water quality in receiving waters, fish, benthos or overall river ecology. Consequently, immission based criteria have to be applied. In order to prove their fulfilment, an integrated modelling approach is necessary. These models are often highly complex containing many parameters and several input and state variables. Therefore, the need to consider and distinguish variability and uncertainty appropriately becomes even more important.

However, this implies several consequences for the different stakeholders compared to the use of simple fixed criteria:

Engineers: Consider an engineer being confronted with a fixed criteria of 20 m^3 per $ha_{impervious}$. She needs only to estimate the impervious area. No further modelling is required. Engineers may prefer this type of criteria as this involves the least work and does not involve dealing with uncertainty. On the other hand, as coping with immission based criteria is more demanding, modelling and analysis naturally makes up a larger share in the engineering project than in traditional approaches. Engineers need to have a good knowledge of statistics and systems analysis. They have to be capable of expressing the involved uncertainties as well as spatial and temporal variability. This requires new tools and methods.

System owners might prefer immission based criteria as they assume less chance of overdesign, leading to lower costs in construction. However, they are more involved in the decision process as they need to specify probabilities and take on more responsibility than in the case of fixed criteria.

Regulators will argue that immission based criteria lead to site-specific solutions and thus to optimal design. On the other hand, they have to develop the criteria and establish suitable monitoring strategies. A well laid out monitoring process is necessary to prove compliance with frequency based criteria after a structure has been built. The sampling procedure must hereby take into account the variability of the processes involved (Bertrand-Krajewski et al. 2002).

The **Public** will expect overall goals to be reached at the least expense.

In general more responsibilities are moved from the legal framework to the engineers, the system owners the authorities. The higher degree of freedom in the decision process requires a risk based approach. Robustness in design will only be gained through the experience of

practical engineers applying probabilistic methods to comply with immission based criteria and a monitoring of the performance of structures built.

It remains an open question whether the different stakeholders are capable of coping with these new tasks.

It would be worthwhile to explore how heavily the final design actually depends on the type of criteria used. Do designs determined with fixed criteria differ much from designs derived with immission based criteria? If yes, what is the difference with respect to safety and costs involved in the approaches?

LIST OF REFERENCES

Bertrand-Krajewski, J.-L., S. Barraud, and J.-P. Bardin. 2002. Uncertainties, performance indicators and decision aid applied to stormwater facilities. Urban Water **4**:163-179.

Blanksby, J. 2002. A review of the UK approach to measuring the performance of combined sewer overflows. Urban Water **4**:191-198.

Borsuk, M. E., C. A. Stow, and K. H. Reckhow. 2002. Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development. Environmental Science & Technology **36**:2109-2115.

Crabtree, B., and G. Morris. 2002. Effective environmental regulation to maximise the benefits of integrated wastewater management. Water Science and Technology **45**:211-218.

Frey, H. C., and D. S. Rhodes. 1998. Characterization and Simulation of Uncertain Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty, and Parameter Dependence. Human and Ecological Risk Assessment **4**:423-468.

Grum, M., and R. H. Aalderink. 1999. Uncertainty in return period analysis of combined sewer overflow effects using embedded Monte Carlo simulations. Water Science and Technology **39**:233-240.

Hauger, M. B., W. Rauch, J. J. Linde, and P. S. Mikkelsen. 2002. Model Based Risk Assessment of CSO Induced DO Depletion in Urban Rivers *in* E. W. Strecker and W. C. Huber, editors. 9th International Conference on Urban Drainage. American Society of Civil Engineers, Portland, Oregon.

Kelly, E. J., and K. Campbell. 2000. Separating variability and uncertainty in environmental risk assessment - Making choices. Human and Ecological Risk Assessment **6**:1-13.

Kreikenbaum, S., V. Krejci, W. Rauch, and L. Rossi. 2002. Probabilistic Modeling as a New Planning Approach to Stormwater Management. *in* E. W. Strecker and W. C. Huber, editors. 9th International Conference on Urban Drainage. American Society of Civil Engineers, Portland, Oregon.

Reichert, P., and M. E. Borsuk. 2003. Does High Forecast Uncertainty Preclude Effective Decision Support? Environmental Modelling and Software *in Press.*

Willems, P., and J. Berlamont. 1999. Probabilistic modelling of sewer system overflow emissions. Water Science and Technology **39**:47-54.

Chapter 2

Marc B. Neumann, Urs von Gunten, Willi Gujer

Sources of Parameter Uncertainty in Predicting Treatment Performance – The Case of Pre-ozonation in Drinking Water Engineering

accepted for publication in *Environmental Science and Technology (2007)*

Sources of Parameter Uncertainty in Predicting Treatment Performance – The Case of Pre-ozonation in Drinking Water Engineering

Marc B. Neumann[1,2], Urs von Gunten[1,3], Willi Gujer[1,2]

[1]*Swiss Federal Institute of Aquatic Science and Technology, Eawag, 8600 Dübendorf, Switzerland*

[2]*Institute of Environmental Engineering, ETH Zurich, 8093 Zurich, Switzerland*

[3]*Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland*

ABSTRACT

This paper explores the assessment of parameter uncertainty in a case study from the field of drinking water treatment. For a full scale pre-ozonation reactor we quantify uncertainty related to ozone decay which is a crucial parameter when modelling micro-pollutant oxidation or disinfection. The pseudo first order rate constant of ozone decay is estimated for three experimental setups representing the planning, piloting and operation of a reactor. Hereby we analyze how various uncertainty sources such as sampling procedures, chemical analysis, model parameters, model structure and temporal variability of the underlying processes contribute to uncertainty in the result. Although uncertainty analysis enhances the understanding of the system, it is also shown to be a subjective process depending on the knowledge and assumptions of the modeller and the availability and quality of data.

KEYWORDS

Uncertainty analysis, sensitivity analysis, Monte Carlo simulation, drinking water treatment, ozone, risk assessment

INTRODUCTION

As risk based regulations are introduced across many sectors of environmental management and technology, engineers are required to quantify frequency and probability of meeting health and environmental quality standards (*1*).

A quantitative uncertainty analysis usually consists of the following steps: identification and characterisation of uncertain model parameters, propagation of the uncertain parameters (Monte Carlo simulation) followed by the analysis and further use of the results e.g. for quantitative risk analysis or decision making under uncertainty (*2*,*3*).

The performance of reactors in the field of water technology is typically characterized as a function of inputs, kinetics and hydraulic contacting. For the analysis of reactors from environmental engineering, compartment modelling is the widely applied tool. Batch reactors, completely stirred tank reactors, and plug flow reactors are the building blocks for the modelling of full scale systems (*4*).

The case study of ozonation in water treatment is selected due to the wide interest in optimizing the involved processes. Ozone dosage must be high enough to guarantee prescribed levels of disinfection and micro-pollutant removal and simultaneously be low enough to minimize the production of potentially harmful disinfection by-products (*5, 6*).

In aquatic chemistry, reaction rates of ozone with many chemical compounds have been measured (*7, 8*). Most of this basic research has been conducted in synthetic waters for single compounds. For natural waters contributions have been made to quantifying ozone decay kinetics and its dependencies on temperature, pH, alkalinity, and type and concentration of natural organic matter (*9, 10*). Due to the complexity of the natural water matrix, empirical models are applied to describe decay behaviour: Ozone decay in a batch system is characterized by a rapid initial decay lasting in the range of a minute followed by a second, slower phase which is typically approximated by first-order kinetics: $r_{O3} = -k_{O3} \cdot S_{O3}$ where r_{O3} is the rate of ozone decay, k_{O3} the decay rate constant and S_{O3} the ozone concentration in the batch reactor (*11*).

The rate constant k_{O3} is estimated together with its associated uncertainty for three consecutive design stages of a pre-ozonation reactor treating lake water. The study is conducted within a framework of frequentist statistics applying methods from nonlinear least squares regression, time series analysis and regional sensitivity analysis (*12-14*). The contribution of this study is not to identify more appropriate model structures but to explicitly characterise and take into account uncertainty when using typically applied models and available data. Additionally this study is able to identify a strategy for uncertainty reduction in view of future on-line estimation of ozone exposure.

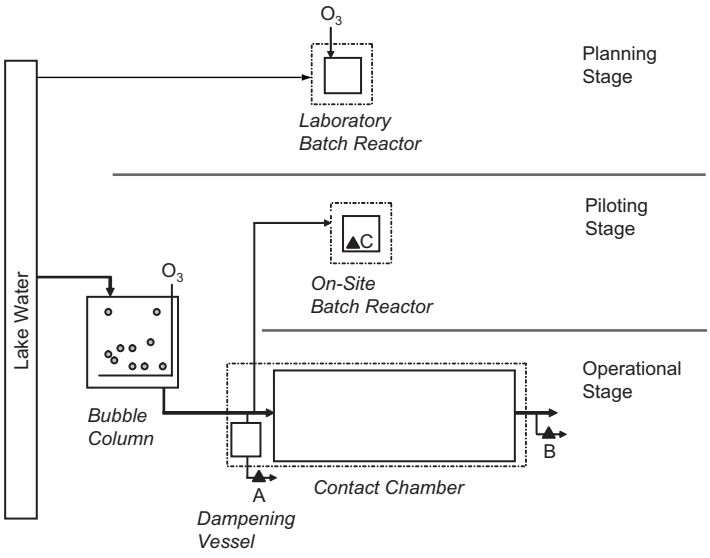

Figure 1. Scheme for comparing rate constants obtained at different stages in engineering design. The system boundaries for estimating the rate constants (dot-dashed boxes) and the location of ozone sensors A, B and C are indicated. The full scale system consists of a counter current bubble column followed by a contact chamber. Signal A is obtained at the outlet of a dampening vessel which is installed to gain a flow independent and smooth signal for optimized control of ozone dosage in the bubble column.

MATERIALS AND METHODS

General Experimental Setup and Procedure

The full-scale pre-ozonation reactor treats Lake Zurich water extracted at a depth of 30 m. The reactor consists of a counter current bubble column where ozone transfer takes place followed by a contact chamber for disinfection and oxidation (*15*). Three levels of scale are investigated representing consecutive stages in engineering design (Figure 1). In the planning stage the rate constant is predicted by coupling online data with an empirical model derived from laboratory experiments. For the piloting stage the rate constant is estimated from ozone decay curves (sensor C) obtained from an on-site automated sequenced batch reactor sampling water from the outlet of the bubble column. In the operational stage the rate constant of ozone decay in the contact chamber is derived from online ozone concentration data (sensors A and B) and a hydraulic model of the full scale reactor. The entire analysis is limited to a time span of 100 hours during which the full scale plant was operated at a constant flow.

Planning Stage

Based upon expert knowledge and previous research conducted on this specific surface water (*10*) a multiplicative model with exponential dependencies for pH, temperature and DOC-concentration is hypothesized. It is made up of four parameters: three factors for the exponential dependencies f_{pH}, f_T, f_{DOC} and a normalizing rate constant $k_{O3,0}$ (eq 1). This model does not intend to be a universally valid description of decay kinetics; rather its intention is to act as an adequate response surface to predict the temporal variation of the rate constant as a function of varying raw water characteristics for the specific surface water under study.

$$k_{O3} = k_{O3,0} \cdot e^{f_{pH} \cdot (pH-8)} \cdot e^{f_T \cdot (T-20°C)} \cdot e^{f_{DOC} \cdot (C_{DOC}-2.4mg/l)} \quad (1)$$

The model is calibrated with data from laboratory experiments which were performed on water obtained at the inlet to the water treatment plant (*16*). The temporal decay of ozone in the batch reactor (initial dosage of 1 mg L^{-1}) was measured by the indigo method (*17*). The reactor configuration together with the methodology for obtaining 1st order rate constants is described in Hoigné and Bader (*11*). The batch experiments were repeated at different levels of pH, temperature and DOC (Supporting Information Table S1).

Piloting Stage

A water sample is extracted from the centre of the inlet to the full-scale contact chamber once every 45 minutes and fed to an on-site batch reactor (V = 1 L). The ozone concentration is measured by a gold microelectrode array (Au-SenSys, CSEM, Neuchâtel) at a temporal resolution of 28 seconds. For every batch cycle the rate constant k_{O3} is estimated together with the initial concentration $C_{O3,ini}$ (eq 2, 3) applying nonlinear least squares (*nls*) regression (*12, 18*):

$$\frac{dC_{O3}}{dt} = -k_{O3} \cdot C_{O3} \quad (2)$$

$$C_{O3}(t=0) = C_{O3,ini} \quad (3)$$

Operational Stage

In the full scale system the raw lake water runs through a counter current bubble column (V = 210 m^3) where the ozone transfer (dosage of 1.1 mg O_3 L^{-1}) takes place with an efficiency of ~ 99%. A continuous side stream (Q_{DV} = 0.3 L min^{-1}) is extracted from the centre of the inlet to the contact chamber. The ozone concentration (sensor A, Figure 1) is measured at the outlet of the dampening vessel (mean retention time = 11 min), which is installed by the water works for control purposes. At the outlet of the contact chamber the ozone concentration is obtained directly from the side stream (sensor B, Figure 1). Both sensors are ozone electrodes (EIT model no. 8422, M.A.S Consulting GmbH, Bern).

We assume the point samples to be representative of the cross-sectional average concentration. The temporal sampling uncertainty is negligible as the concentration data measured at a resolution of 20 seconds are averaged and stored as hourly values. We further assume that the initial fast decay of ozone has taken place in the bubble column and apply a steady state model with first order decay of ozone.

Both the contact chamber and dampening vessel are modelled as a series of completely stirred tank reactors (CSTR's) leading to eq 4:

$$\frac{\left(C_{O3,A}+err_{C_{O3,A}}\right)\cdot\left(1+\frac{k_{O3}\cdot V_{DV}}{Q_{DV}\cdot n_{DV}}\right)^{n_{DV}}}{\left(1+\frac{k_{O3}\cdot V_{CC}}{Q_{CC}\cdot n_{CC}}\right)^{n_{CC}}}-\left(C_{O3,B}+err_{C_{O3,B}}\right)=0 \qquad (4)$$

where C_{O3} denotes the ozone concentrations and err_{CO3} the systematic measurement error for the two sensors A and B. The hydraulic properties for both the contact chamber (CC) and the dampening vessel (DV) are characterised by the volume V, the flow measurement Q and the number of CSTR compartments n. Eq 4 is solved for the pseudo 1st order rate constant k_{O3}.

RESULTS AND DISCUSSION

Planning Stage

The model coefficients of eq 1 are estimated applying nonlinear least squares (*nls*) regression (*12, 18*) to the data supplied by the laboratory (Supporting Information Table S1). The best estimates and error information are listed in Table 1, the measured and fitted rate constants are displayed in Figure 2 together with the settings of pH, temperature and DOC from the laboratory batch experiments.

Table 1. Summary of parameters (equation 1) obtained from nonlinear least squares estimation.

Parameter	Best Estimate	Standard Deviation	Correlation Matrix		
$k_{O3,0}$	28.5	0.7	$k_{O3,0}$		
f_{pH}	0.83	0.05	-0.20	f_{pH}	
f_T	0.10	0.02	0.21	-0.00	f_T
f_{DOC}	0.74	0.05	-0.34	-0.14	-0.25

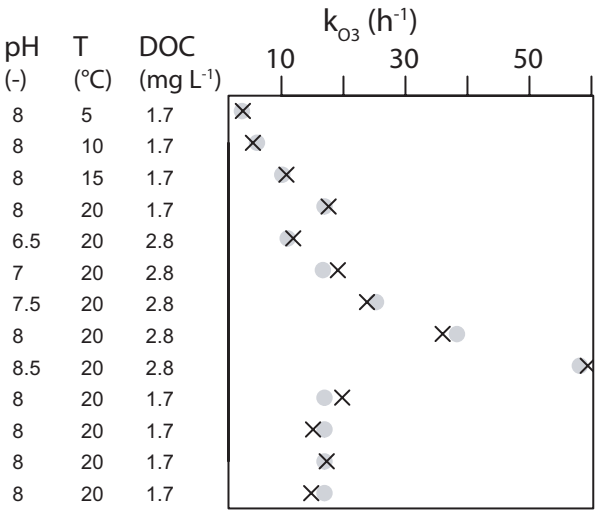

Figure 2. Measured (×) rate constants obtained from the laboratory batch experiments and fitted values (●) from calibrated model (eq 1) together with settings of pH, temperature and DOC of the laboratory batch experiments.

The time dependent rate constant for the 100 hour time window of interest is predicted by coupling the calibrated model (eq 1) with pH-, temperature- and DOC- data. For pH and temperature hourly values are available for the selected time window of interest. As the resolution of available DOC data is much lower the average DOC concentration from 155 monthly measurements is used (Supporting Information Tables S2, S3). In this way a 100 hour time series of rate constants (reference analysis) is predicted (Figure 5).
To assess the uncertainty (due to uncertainty in the model coefficients and uncertainty about the assumption of constant DOC- concentration), the time series are repeatedly computed in a MC simulation (uncertainty analysis). Random samples are drawn for the model coefficients of eq 1 from a multivariate normal distribution which approximates the covariance structure in Table 1. A DOC-concentration is re-sampled from the available 155 values in every MC simulation. Figure 6 indicates the mean and 90% confidence interval obtained for the time dependent rate constant.

Piloting Stage

For the first batch cycle the best estimates (k_{O3}, $O3_{ini}$), standard errors (sd k_{O3}, sd $O3_{ini}$) and correlation (cor(k_{O3}, $O3_{ini}$)) resulting from the nonlinear least squares estimation are supplied in Table 2. In case 1 the estimation is performed with the ozone raw data obtained with the gold microelectrode array. In case 2 the estimation is performed with calibration-corrected data: the ozone raw data are transformed with a linear calibration curve obtained by comparing indigo reference measurements with microelectrode measurements (Supporting Information Table S4). For case 2 the modelled concentrations (lines) are shown together with the data (circles) in Figure 3.

The pattern of the model discrepancies (differences of measured and modelled concentrations) exposes systematic effects (Figure 3). Three successive data points are clustered reflecting systematic electrode behaviour. To a lesser degree there is structural model deficiency across the entire reaction span indicating slightly faster decay during the initial phase. Both effects are observed for all of the batch cycles and lie within similar absolute ranges.

Table 2. Estimation of the rate constant k_{O3} and the initial concentration $O3_{ini}$ for the first batch cycle.

	k_{O3}	sd k_{O3}	$O3_{ini}$	sd $O3_{ini}$	cor(k_{O3}, $O3_{ini}$)	basis for calculations
	h^{-1}	h^{-1}	mg L^{-1}	mg L^{-1}	-	
case 1	1.51	0.023	0.49	0.003	0.79	with raw data
case 2	1.78	0.026	0.51	0.004	0.78	accounting for calibration curve
case 3	1.84	0.066	0.52	0.010	0.77	accounting for calibration curve and addition of normal random process

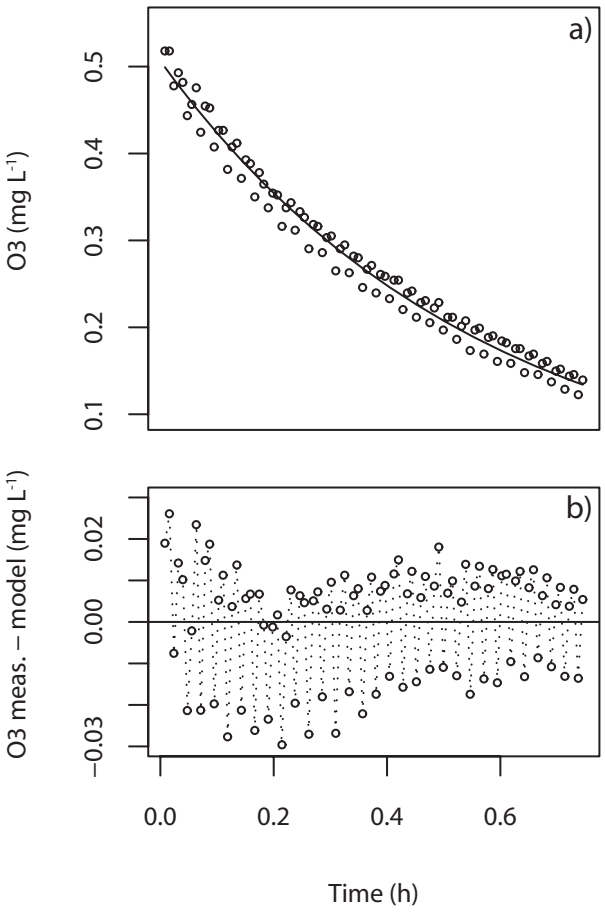

Figure 3. a) Measured (circles) and modeled (line) ozone concentrations for the first pilot batch cycle (Table 2, case 2). b) Model discrepancy (measured minus modelled concentration values) indicates short term autocorrelation due to sensor configuration and systematic model deviation across the reaction span.

As the discrepancies from the single batch run are not independent and identically distributed the standard error sd k_{O3} underestimates uncertainty about the rate constant. To obtain a more meaningful uncertainty measure a normally distributed random process is added to measurement data. The standard deviation of this normal distribution is chosen in such a way that the first 20 coefficients of both the autocorrelation- and partial autocorrelation functions (*acf* and *pacf*) (*18*) of the discrepancies are classified as being independent. This is reached

when the number of significant lag coefficients of the acf and pacf lying outside the 95% confidence bands is not larger than two and they do not deviate from the confidence band by a large margin (*19*). The latter requirement necessitates a visual examination of the acf and pacf plots. In this study it was found that the requirements are reached in all batch cycles when adding a normally distributed random process with standard deviation sd = 0.03 mg L^{-1} to the data of the calibration-corrected data from case 2 (results not shown). For the first batch cycle this procedure leads to an increase of the standard error from sd k_{O3} = 0.026 to 0.066 h^{-1} (Table 2, case 3).

The reference analysis is performed with the calibration-corrected (case 2) data of every batch cycle (Figure 5). For the uncertainty analysis the calibration curve and model structure deficiency are taken into account (case 3). Assuming the parameter covariance (Table 2) to be approximated by a multivariate normal distribution 5% and 95% quantiles are estimated for every batch cycle as k_{O3} +/- 1.65 sd k_{O3} (Figure 6).

Operational Stage

The factors influencing the estimation of k_{O3} in the operational stage are summarised in Table 3. Concentration data of sensors A and B for the time window of interest are available from the water works database (Supporting Information Table S2). Random measurement errors in the concentration data are not present due to the averaging of high-frequency data (20 second resolution). However systematic measurement errors are detected in an analysis of the water works calibration files (err.$C_{O3,A}$ and err.$C_{O3,B}$). The calibration adjustments to the two sensors are correlated and do not exhibit any drifts (no temporal trends) indicating an erroneous recalibration procedure. A bi-variate normal distribution is fitted to 14 calibration adjustments evaluated at the average concentrations of the 100 hr time window (Supporting Information Table S5).

The number of required CSTR's for the contact chamber is derived from tracer experiments; models with number of tanks between 2 and 5 were found to match the tracer data equally well. As no tracer data is available for the dampening vessel, a wide range of hydraulic models between one and 10 CSTR's in series is considered possible. For the systematic errors of the flow measurements uniform uncertainty ranges are assumed based on expert knowledge of the plant operator. Reactor volumes are not considered to be uncertain.

For the reference analysis eq 4 is solved for each of the 100 concentration values with model parameters (Table 3) set at reference values (Figure 5). For the uncertainty analysis this is repeated in an MC simulation (1000 runs) by drawing random samples from the distributions in Table 3 and summarised by mean and the 90% confidence interval (Figure 6).

Table 3. Factors influencing estimation of k_{O3} in the full scale system: Concentrations time series of ozone C_{O3}, systematic errors of measurement devices $err.C_{O3}$, uncertainty about appropriate hydraulic model structure n, systematic errors for flow Q and volume V of the contact chamber (CC) and the dampening vessel (DV).

Factor	Units	Reference Analysis	Uncertainty Analysis	Description	Source
$C_{O3,A}$	mg L^{-1}			O3 concentration sensor A	database water works 100 h time series
$C_{O3,B}$	mg L^{-1}			O3 concentration sensor B	database water works 100 h time series
$err.C_{O3,A}$	mg L^{-1}	0	Normal(0,0.056) $cor(err.C_{O3,B}) = 0.92$	bias sensor A	analysis of calibration protocols
$err.C_{O3,B}$	mg L^{-1}	0	Normal(0,0.054) $cor(err.C_{O3,A}) = 0.92$	bias sensor B	analysis of calibration protocols
n_{CC}	-	2	Uniform(2,..,5)	no. of CSTR's contact chamber	calibration with tracer data
n_{DV}	-	1	Uniform(1,..,10)	no. of CSTR's dampening vessel	assumption analyst
Q_{CC}	$m^3 h^{-1}$	1180	Uniform(1120, 1240)	flow contact chamber	MID - flow device, calibration water works
Q_{DV}	L min^{-1}	0.3	Uniform(0.255, 0.345)	flow dampening vessel	flow device, assumption analyst
V_{CC}	m^3	420	420	volume contact chamber	construction plans
V_{DV}	L	3.3	3.3	volume dampening vessel	construction plans

Systematic measurement errors are characterised by mean, standard deviation and correlation. Uniform distributions are given by minimum and maximum. The number of CSTR's can only assume discrete values.

The uncertainty distribution of the rate constant for the first 1 hour concentration values is shown in Figure 4a. Additionally, the variance contribution of the parameters to the rate constant is assessed by computing 1st order sensitivity indices, a measure of regional sensitivity (*13*). Figure 4b shows the percentage contribution of the parameters to the total variance of k_{O3} for the first time step. Over 80% of the total variance of k_{O3} is explained through systematic measurement errors of ozone sensors. The reason for this factor to be so influential is the small difference between the ozone concentration in the effluent of the dampening vessel and the effluent of the reactor relative to the considerable measurement error. Although there is considerable uncertainty associated with the hydraulic model structure of the dampening vessel ($n_{DV} \sim$ Uniform(1,..,10)), the variance decomposition reveals that this uncertainty is responsible for less than 5% of the variance of the rate constant k_{O3} (Figure 4b). As a consequence the uncertainty about the rate constant cannot be significantly reduced by improving the model structure of the dampening vessel (e. g. by performing tracer studies).

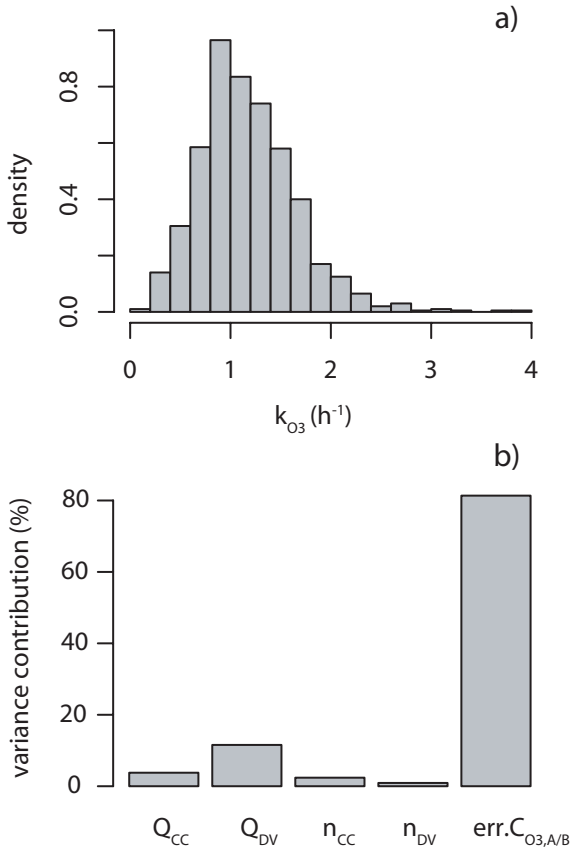

Figure 4. a) Uncertainty of the rate constant k_{O3} for the first 1 hr samples. b) Variance contribution of the parameters to the variance of k_{O3} for the first 1 hr samples. For explanation of Q_{CC}, Q_{DV}, n_{CC}, n_{DV} and err.$C_{O3,A/B}$ see Table 3.

Comparison of rate constants: Reference Analysis

The reference analysis characterises the pseudo first order rate constant as a time dependent parameter for the three engineering stages (Figure 5). Visual examination shows the variability occurring at a time scale of several hours to be similar for the three approaches. For the laboratory batch system temporal variability is induced by the variability in the pH- and temperature data. The structure of the time series in the other two systems reflects the superposition of sampling uncertainty and natural variability of the water composition and temperature. On a relative scale the underlying process (natural water variability) is well reproduced, but the offsets between the three time series are significant indicating that systematic effects outweigh random effects.

Comparison of rate constants: Uncertainty Analysis

In the uncertainty analysis the mean, 5% and 95% quantiles of the rate constant are computed for each engineering stage at each time step (Figures 6a - c). Note the dependency of the results on the availability of information: For the operational stage, the analysis of long-term calibration data allowed us to formulate a sensor error model whereas for the piloting stage no such error model could be derived due to the limited time span of operation. The procedures of arriving at the uncertainty bands differ: In the planning stage the sources of uncertainty are the parameter covariance of the response surface model (major contribution) and uncertainty due to the unknown DOC concentration (minor contribution). In the piloting stage the uncertainty results from the nonlinear parameter estimation. For the operational stage the uncertainty stems from applying expert knowledge on the uncertainty of model parameters

(Table 3). Interestingly the enlarged confidence band of the piloting stage (resulting from the addition of the normally distributed process to account for model discrepancy) remains an order of magnitude smaller than in the other two systems.

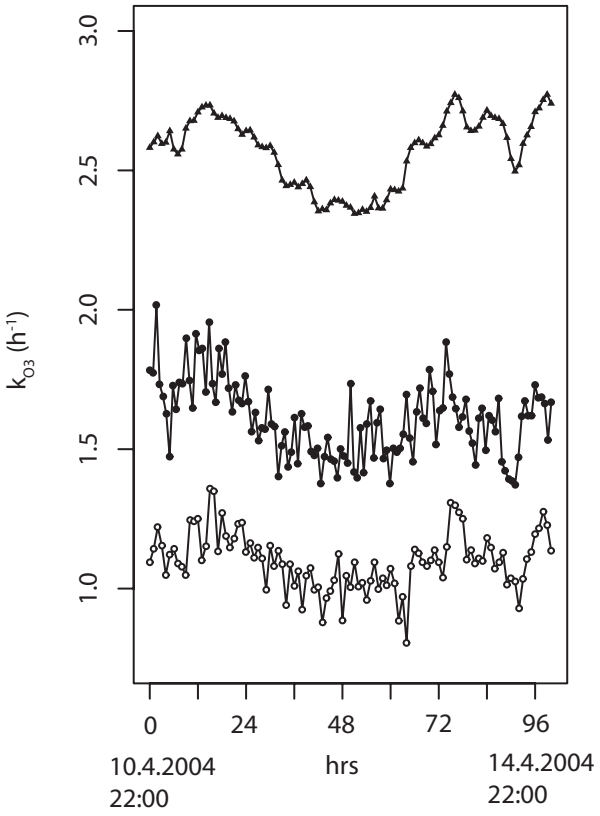

Figure 5. Time series of rate constants from the reference analysis: Planning Stage (▲), Piloting Stage (●) and Operational Stage (○).

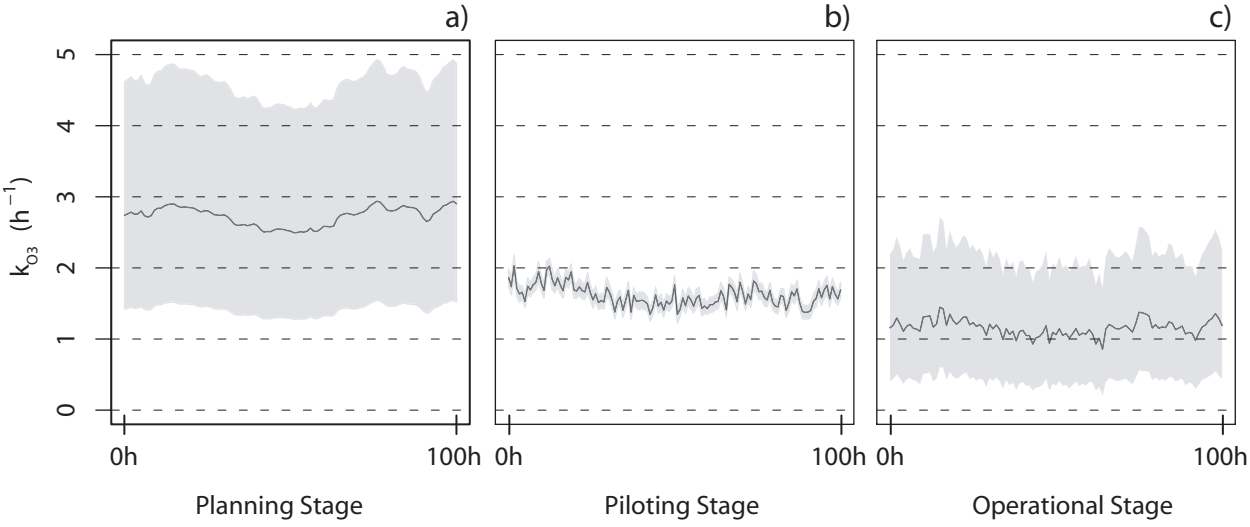

Figure 6. Time series of the ozone rate constant with uncertainty estimates for Planning Stage a), Piloting Stage b) and Operational Stage c). Uncertainty analysis resulting in time series of mean, lower quantile (5%) and upper quantile (95%) of the rate constant for the 100 hour time window of interest.

IMPLICATIONS FOR WATER WORKS

How do the results from the uncertainty analysis affect predicted reactor performance? Assuming rate constants of k_{O3} = 0.5, 2 and 4 h^{-1} one obtains outlet to inlet ozone concentration ratios of 0.85, 0.56 and 0.36, respectively, when modelling the contact chamber with two CSTR compartments. As most micro-pollutant removal and micro-organism inactivation processes follow first order kinetics in ozone this uncertainty about k_{O3} is expected to have a considerable effect on the performance prediction.

If the main interest of the water works is to quantify the kinetics of the full scale system we propose that this is done with an on-site batch system and not with measurements from the full scale system. Uncertainty about the hydraulic model structure is smaller and when 1st order kinetics is appropriate to describe ozone decay, the rate constant obtained from a batch reactor is only affected by a possible offset in the calibration curve but is insensitive to errors on the slope.

Ideally the water works should couple information from "hazard analysis and critical control points" protocol (HACCP) (*20*) - which consists of warning and alarm levels for ozone concentrations at sensor B - with kinetic information extracted from a sequencing batch reactor running in parallel to the full scale system. This information on the rate constant is of significant value in view of online performance prediction of micro-pollutant oxidation and disinfection.

The rate constants derived from the laboratory (planning stage) are critical for several reasons: Choosing later time sections in the laboratory batch data would result in lower rate constants (due to an apparent decrease of the rate constant during the reaction). To adequately capture the temporal variation of rate constants and to determine the dependency on changing water characteristics it is recommended that batch experiments are performed on raw water samples taken at different times. The required temporal resolution of these experiments depends on the intended use and the time scale at which the water characteristics change. To obtain seasonal rate constants monthly experiments should be sufficient, for an analysis comparable to the one conducted in this study hourly experiments are required.

IMPLICATIONS FOR UNCERTAINTY ANALYSIS

Uncertainty about model structure and parameters

In order to reduce model structure uncertainty one might suggest that a more mechanistic characterization of ozone decay would be appropriate, such as the modelling of different fractions of NOM and the involved ·OH radical reactions (*21*). However, for natural systems this is not feasible due to the unknown and complex character of NOM. Such aggregation effects are present in many full scale environmental engineering systems. For applications where the decay process needs to be characterised with high accuracy (e.g. for the laboratory batch experiments) the model structure should be extended: e.g. a time dependent rate constant (e.g. exponentially decaying) or a mixed order decay of ozone and degradable DOC. As neither sensor data (typically not just containing random errors) nor models (due to their semi-empirical nature) are ideal, error information from regression analysis typically underestimates uncertainty due to correlated model discrepancies (measured - modelled concentrations). For the on-site batch reactor of the piloting stage this effect was accounted for by adding a random Gaussian process to measurement data increasing the standard error of the estimated parameter by a factor of 2.5 (Table 2). The authors believe that using

uncertainty estimates (e.g. for uncertainty propagation) when statistical assumptions are not fulfilled is a critical issue that requires further exploration.

Uncertainty due to non-representative sampling

Comparison of sampling procedures is crucial when up-scaling information. In the pilot batch system water packages are sampled once every 45 minutes. Sampling uncertainty results from extracting water packages that have experienced different "histories" in the bubble column. In the full scale system the concentration measurements are averages (1 hour) of high-frequent samples (20 seconds) and thus capture integral information of the water property with respect to time. The appropriate analysis of sampling uncertainty is difficult as concentration profiles across the inlet structure would need to be measured. When possible this effect should be eliminated in a precautionary way ensuring homogenous conditions.

Uncertainty due to errors in chemical analysis

Slope and offset of calibration curves may be subject to drifts resulting in systematic measurement errors. The calibration curves themselves might not be linear. This uncertainty can only be reduced through online calibration of devices and rigorous quality control procedures (*22*). For the full scale system of this case study miscalibration was identified as a significant uncertainty introducing factor.

Uncertainty due to temporal variability of underlying processes

At the interface of natural and technical systems matrix variability is commonly encountered. In our case study, temporal changes of the water matrix lead to a change of ozone decay behaviour. This source of uncertainty is taken into account by explicitly characterising k_{O3} as a time dependent rate constant.

Uncertainty due to numerical and computational procedures

If numerical algorithms are correctly implemented we suppose uncertainties from these sources to be negligible for many applications in environmental engineering. The parameter estimation for the pilot batch system was repeated in two other software environments (*23, 24*) under different minimization algorithms leading to exactly the same results at 4 significant digits. This may be different for systems exhibiting highly nonlinear effects or when identifiability of parameters is low (*12*).

Subjectivity of Uncertainty Analysis

An uncertainty analysis remains a subjective task. Assumptions need to be made: definition of system boundaries, spatial and temporal resolution, etc. When applying estimates of parameters and their uncertainty ranges crucial meta-information should be added that makes the process leading to results transparent and open to peer review: Definition of goal, giving supporting information on the data used, indicating the statistical framework applied and other assumptions made in the analysis.

ACKNOWLEDGEMENTS

The authors would like to thank Lisa Salhi (Eawag) and Hans-Peter Kaiser (WVZ) for fruitful discussions and providing information on the pilot stage and the operational stage respectivley.

LITERATURE CITED

(1) Barnett, V.; O'Hagan, A., Setting environmental standards - The statistical approach to handling uncertainty and variation. Chapman & Hall: London, 1997.

(2) Morgan, M. G.; Henrion, M.; Small, M., Uncertainty - A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press: Cambridge, 1990.

(3) Clemen, R. T.; Reilly, T., Making hard decisions. Duxbury: 2001.

(4) Levenspiel, O., Chemical reaction engineering. Third ed.; Wiley: New York, 1999.

(5) Kim, J. H.; von Gunten, U.; Marinas, B. J., Simultaneous prediction of Cryptosporidium parvum oocyst inactivation and bromate formation during ozonation of synthetic waters. Environmental Science & Technology 2004, 38, (7), 2232-2241.

(6) Tang, G.; Adu-Sarkodie, K.; Kim, D.; Kim, J. H.; Teefy, S.; Shukairy, H. M.; Marinas, B. J., Modeling Cryptosporidium parvum oocyst inactivation and bromate formation in a full-scale ozone contactor. Environmental Science & Technology 2005, 39, (23), 9343-9350.

(7) von Gunten, U., Ozonation of drinking water: Part I. Oxidation kinetics and product formation. Water Research 2003, 37, (7), 1443-1467.

(8) von Gunten, U., Ozonation of drinking water: Part II. Disinfection and by- product formation in presence of bromide, iodide or chlorine. Water Research 2003, 37, (7), 1469-1487.

(9) Elovitz, M. S.; von Gunten, U., Hydroxyl radical ozone ratios during ozonation processes I: The R-ct concept. Ozone-Science & Engineering 1999, 21, (3), 239-260.

(10) Elovitz, M. S.; von Gunten, U.; Kaiser, H. P., Hydroxyl radical/ozone ratios during ozonation processes II: The effect of temperature, pH, alkalinity, and DOM properties. Ozone-Science & Engineering 2000, 22, (2), 123-150.

(11) Hoigne, J.; Bader, H., Characterization of Water-Quality Criteria for Ozonation Processes 2: Lifetime of added ozone. Ozone-Science & Engineering 1994, 16, (2), 121-134.

(12) Bates, D. M.; Watts, D. G., Nonlinear regression analysis and its applications. Wiley: New York, 1988.

(13) Saltelli, A.; Ratto, M.; Tarantola, S.; Campolongo, F., Sensitivity analysis for chemical models. Chemical Reviews 2005, 105, (7), 2811-2827.

(14) Venables, W. N.; Ripley, B. D., Modern applied statistics with S. 4th ed.; Springer: New York, 2002.

(15) von Gunten, U.; Elovitz, M.; Kaiser, H. P., Calibration of full-scale ozonation systems with conservative and reactive tracers. Journal of Water Supply: Research and Technology - Aqua 1999, 48, (6), 250-256.

(16) Jaeggi, P. Dependency of R_{ct} and k_{O3} on raw water variables; Eawag: Dübendorf, 2003.

(17) Bader, H.; Hoigne, J., Determination of Ozone in Water by the Indigo Method. Water Research 1981, 15, (4), 449-456.

(18) R Development Core Team, R: A language and environment for statistical computing. 2006.

(19) Chatfield, C., The Analysis of Time Series - An Introduction. Fifth ed.; Chapman & Hall: Boca Raton, 1999.

(20) Ropkins, K.; Beck, A. J., Evaluation of worldwide approaches to the use of HACCP to control food safety. Trends in Food Science & Technology 2000, 11, (1), 10-21.

(21) Staehelin, J.; Hoigne, J., Decomposition of ozone in water in the presence of organic solutes acting as promoters and inhibitors of radical chain reactions. Environmental Science & Technology 1985, 19, (12), 1206-1213.

(22) Rieger, L.; Thomann, M.; Gujer, W.; Siegrist, H., Quantifying the uncertainty of on-line sensors at WWTPs during field operation. Water Research 2005, 39, (20), 5162-5174.

(23) Reichert, P., Aquasim - a tool for simulation and data-analysis of aquatic systems. Water Science and Technology 1994, 30, (2), 21-30.

(24) Macey, R. I.; Oster, G. F. Berkeley Madonna, Berkeley CA: University of California, 2005.

SUPPORTING INFORMATION CHAPTER 2

Contents:

Table S1. Pseudo first order rate constants k_{O3} derived from laboratory batch experiments in the planning stage.

Experiment	pH	T	DOC	start time	end time	start O_3 conc.	end O_3 conc.	k_{O3} Laboratory	k_{O3} Model
	[-]	[°C]	[mg L^{-1}]	[s]	[s]	[mg L^{-1}]	[mg L^{-1}]	[h^{-1}]	[h^{-1}]
1	8	5	1.7	120	1200	0.53	0.17	3.78	3.63
2	8	10	1.7	120	1200	0.47	0.09	5.4	6.07
3	8	15	1.7	120	360	0.26	0.13	10.8	10.1
4	8	20	1.7	15	540	0.49	0.04	17.6	17
5	6.5	20	2.8	15	600	0.47	0.07	11.9	11
6	7	20	2.8	15	240	0.37	0.12	19.1	16.7
7	7.5	20	2.8	30	600	0.4	0.01	23.8	25.2
8	8	20	2.8	30	300	0.4	0.03	36	38.3
9	8.5	20	2.8	30	240	0.28	0.01	59.4	58.1
10	8	20	1.7	NA	NA	NA	NA	19.8	17
11	8	20	1.7	NA	NA	NA	NA	15.1	17
12	8	20	1.7	NA	NA	NA	NA	17.3	17
13	8	20	1.7	NA	NA	NA	NA	14.8	17
	batch settings			time and concentration ranges of batch data used to estimate k_{O3} Laboratory				rate constants estimated from Laboratory Batch Data	rate constants of fitted model (eq 1)

Table S2. pH and Temperature of raw water and ozone concentrations at sensors A and B during the investigated 100 hour time window.

date and time	pH	Temperature [°C]	sensor A [mg L^{-1}]	sensor B [mg L^{-1}]
10.04.2004 22:00	8.025	4.9644	0.4405	0.3704
10.04.2004 23:00	8.0311	4.9903	0.4441	0.3713
11.04.2004 00:00	8.0367	5.0266	0.4439	0.3665
11.04.2004 01:00	8.0252	5.0164	0.4418	0.3688
11.04.2004 02:00	8.0264	5.0236	0.444	0.3758
11.04.2004 03:00	8.0397	5.069	0.4466	0.3745
11.04.2004 04:00	8.0179	4.9999	0.4482	0.3743
11.04.2004 05:00	8.0103	4.9947	0.4531	0.3813
11.04.2004 06:00	8.0148	5.0246	0.4505	0.3797
11.04.2004 07:00	8.0386	5.1114	0.4375	0.3703
11.04.2004 08:00	8.0473	5.1382	0.4386	0.3614
11.04.2004 09:00	8.0479	5.1382	0.4415	0.3635
11.04.2004 10:00	8.0593	5.1547	0.4456	0.3664
11.04.2004 11:00	8.0672	5.1549	0.4348	0.3652
11.04.2004 12:00	8.0697	5.1574	0.4382	0.3659
11.04.2004 13:00	8.0677	5.1755	0.438	0.3547
11.04.2004 14:00	8.0586	5.1461	0.4437	0.3603
11.04.2004 15:00	8.0528	5.1369	0.4393	0.3678
11.04.2004 16:00	8.0552	5.1405	0.4408	0.3614
11.04.2004 17:00	8.0528	5.1354	0.4407	0.3656
11.04.2004 18:00	8.0517	5.1314	0.4402	0.3678
11.04.2004 19:00	8.0494	5.118	0.4405	0.3659
11.04.2004 20:00	8.0414	5.0815	0.4429	0.3651
11.04.2004 21:00	8.0348	5.0601	0.4531	0.3733
11.04.2004 22:00	8.0388	5.0765	0.4404	0.3683
11.04.2004 23:00	8.0388	5.0856	0.4457	0.3711
12.04.2004 00:00	8.0309	5.0561	0.4379	0.3673
12.04.2004 01:00	8.0226	5.0177	0.4454	0.3722
12.04.2004 02:00	8.0207	5.01	0.4497	0.3779
12.04.2004 03:00	8.0188	5.0111	0.443	0.3779
12.04.2004 04:00	8.0218	5.0158	0.4493	0.3751
12.04.2004 05:00	8.015	4.9792	0.4513	0.3802
12.04.2004 06:00	8.0016	4.9191	0.4574	0.3823
12.04.2004 07:00	7.9841	4.8436	0.4527	0.3816
12.04.2004 08:00	7.9761	4.8293	0.4522	0.3889
12.04.2004 09:00	7.9795	4.8195	0.458	0.3855
12.04.2004 10:00	7.9834	4.8185	0.4515	0.3844
12.04.2004 11:00	7.9772	4.8033	0.4521	0.3825
12.04.2004 12:00	7.9812	4.8182	0.4504	0.3883
12.04.2004 13:00	7.9857	4.8331	0.4474	0.3788
12.04.2004 14:00	7.978	4.8001	0.4536	0.3831
12.04.2004 15:00	7.9605	4.7225	0.4513	0.385
12.04.2004 16:00	7.9504	4.6703	0.4581	0.3908
12.04.2004 17:00	7.9535	4.6764	0.4491	0.3899
12.04.2004 18:00	7.9534	4.6629	0.453	0.3881
12.04.2004 19:00	7.9615	4.6981	0.4504	0.385
12.04.2004 20:00	7.966	4.7116	0.4467	0.3792
12.04.2004 21:00	7.9653	4.7073	0.4528	0.3791
12.04.2004 22:00	7.9638	4.7032	0.4446	0.3855
12.04.2004 23:00	7.958	4.6918	0.4511	0.3825
13.04.2004 00:00	7.9554	4.685	0.4501	0.3834
13.04.2004 01:00	7.9467	4.6633	0.4502	0.3791
13.04.2004 02:00	7.9474	4.6684	0.4526	0.386

13.04.2004 03:00	7.9512	4.6863	0.4477	0.3805
13.04.2004 04:00	7.9477	4.6854	0.4446	0.3814
13.04.2004 05:00	7.9518	4.709	0.4499	0.3825
13.04.2004 06:00	7.9643	4.7761	0.444	0.3739
13.04.2004 07:00	7.9514	4.7072	0.4476	0.3817
13.04.2004 08:00	7.9507	4.7079	0.4577	0.3886
13.04.2004 09:00	7.9602	4.7547	0.4469	0.3803
13.04.2004 10:00	7.9711	4.8223	0.4484	0.3783
13.04.2004 11:00	7.97	4.8215	0.4796	0.4078
13.04.2004 12:00	7.9703	4.799	0.4689	0.4067
13.04.2004 13:00	7.9733	4.818	0.4624	0.3964
13.04.2004 14:00	8.0026	4.9611	0.4414	0.3874
13.04.2004 15:00	8.019	5.0118	0.4451	0.375
13.04.2004 16:00	8.0244	5.0332	0.447	0.3734
13.04.2004 17:00	8.0281	5.0408	0.4492	0.3765
13.04.2004 18:00	8.0258	5.0205	0.4506	0.3794
13.04.2004 19:00	8.0228	5.0001	0.4507	0.3797
13.04.2004 20:00	8.0244	5.02	0.4501	0.3781
13.04.2004 21:00	8.0307	5.0472	0.4521	0.3783
13.04.2004 22:00	8.0329	5.0699	0.4427	0.3728
13.04.2004 23:00	8.0422	5.119	0.4439	0.3763
14.04.2004 00:00	8.0592	5.167	0.4399	0.367
14.04.2004 01:00	8.0703	5.1851	0.4468	0.365
14.04.2004 02:00	8.0795	5.216	0.4378	0.3581
14.04.2004 03:00	8.0745	5.2149	0.4469	0.3662
14.04.2004 04:00	8.0594	5.1678	0.4524	0.3719
14.04.2004 05:00	8.0408	5.1091	0.4522	0.3803
14.04.2004 06:00	8.0365	5.0969	0.4578	0.3825
14.04.2004 07:00	8.0375	5.094	0.4522	0.3805
14.04.2004 08:00	8.0423	5.1076	0.4525	0.3797
14.04.2004 09:00	8.0528	5.138	0.4473	0.3759
14.04.2004 10:00	8.0606	5.1671	0.4393	0.3653
14.04.2004 11:00	8.0549	5.1447	0.4484	0.3736
14.04.2004 12:00	8.0528	5.136	0.4514	0.3808
14.04.2004 13:00	8.0519	5.1296	0.4527	0.3807
14.04.2004 14:00	8.0476	5.0994	0.4536	0.38
14.04.2004 15:00	8.0328	5.0353	0.4496	0.3825
14.04.2004 16:00	8.0096	4.9376	0.4617	0.3915
14.04.2004 17:00	7.996	4.874	0.4597	0.3904
14.04.2004 18:00	8.0035	4.9011	0.4566	0.3934
14.04.2004 19:00	8.0283	4.9917	0.4476	0.3802
14.04.2004 20:00	8.0387	5.0209	0.4493	0.3772
14.04.2004 21:00	8.0466	5.068	0.4448	0.372
14.04.2004 22:00	8.0609	5.1454	0.4408	0.3658
14.04.2004 23:00	8.0647	5.1632	0.4408	0.3642
15.04.2004 00:00	8.0717	5.2128	0.4417	0.3618
15.04.2004 01:00	8.076	5.2443	0.4327	0.3575
15.04.2004 02:00	8.0673	5.2016	0.4442	0.3718

Table S2 continued

Table S3. 155 measurements of raw water DOC concentration

date	DOC [mg L^{-1}]	date	DOC [mg L^{-1}]	date	DOC [mg L^{-1}]
13.11.1990	1.10	13.06.1995	1.10	09.08.1999	1.26
11.12.1990	1.25	11.07.1995	1.05	06.09.1999	1.26
15.01.1991	1.20	15.08.1995	1.08	11.10.1999	1.31
12.02.1991	1.20	12.09.1995	1.21	08.11.1999	1.23
12.03.1991	1.10	10.10.1995	1.23	13.12.1999	1.21
16.04.1991	1.05	14.11.1995	1.24	24.01.2000	1.16
07.05.1991	1.15	12.12.1995	1.23	14.02.2000	1.28
11.06.1991	1.20	15.01.1996	1.21	13.03.2000	1.21
09.07.1991	1.10	12.02.1996	1.20	03.04.2000	1.30
13.08.1991	1.15	11.03.1996	1.18	15.05.2000	1.27
10.09.1991	1.10	11.04.1996	1.26	19.06.2000	1.22
15.10.1991	1.10	13.05.1996	1.30	10.07.2000	1.24
14.01.1992	1.20	10.06.1996	1.35	14.08.2000	1.14
11.02.1992	1.25	08.07.1996	1.30	04.09.2000	1.12
10.03.1992	1.15	12.08.1996	1.27	17.10.2000	1.19
14.04.1992	1.15	12.09.1996	1.20	13.11.2000	1.23
12.05.1992	1.10	14.10.1996	1.22	11.12.2000	1.25
09.06.1992	1.10	11.11.1996	1.27	15.01.2001	1.14
14.07.1992	1.35	09.12.1996	1.25	12.02.2001	1.25
11.08.1992	1.40	09.12.1996	1.25	12.03.2001	1.30
15.09.1992	1.40	13.01.1997	1.28	09.04.2001	1.27
13.10.1992	1.45	13.01.1997	1.28	14.05.2001	1.34
10.11.1992	1.30	10.02.1997	1.10	18.06.2001	1.23
08.12.1992	1.40	10.03.1997	1.36	09.07.2001	1.29
12.01.1993	1.50	07.04.1997	1.23	13.08.2001	1.28
09.02.1993	1.55	07.04.1997	1.23	03.09.2001	1.21
09.03.1993	1.30	12.05.1997	1.28	08.10.2001	1.29
13.04.1993	1.25	12.05.1997	1.28	12.11.2001	1.38
11.05.1993	1.15	09.06.1997	1.35	10.12.2001	1.28
15.06.1993	1.20	14.07.1997	1.35	14.01.2002	1.15
13.07.1993	1.20	11.08.1997	1.22	11.02.2002	1.24
10.08.1993	1.30	08.09.1997	1.21	11.03.2002	1.19
14.09.1993	1.10	13.10.1997	1.33	02.04.2002	1.22
12.10.1993	1.00	10.11.1997	1.22	13.05.2002	1.22
09.11.1993	1.10	08.12.1997	1.18	10.06.2002	1.25
14.12.1993	1.15	12.01.1998	1.21	08.07.2002	1.24
11.01.1994	1.15	09.02.1998	1.23	12.08.2002	1.29
15.02.1994	1.10	09.03.1998	1.18	02.09.2002	1.27
15.03.1994	1.05	06.04.1998	1.17	14.10.2002	1.18
12.04.1994	0.95	11.05.1998	1.16	11.11.2002	1.37
10.05.1994	1.00	15.06.1998	1.18	04.12.2002	1.21
14.06.1994	1.05	13.07.1998	1.24	13.01.2003	1.37
12.07.1994	1.10	10.08.1998	1.17	10.02.2003	1.26
16.08.1994	1.10	12.10.1998	1.23	10.03.2003	1.25
13.09.1994	1.05	14.12.1998	1.56	14.04.2003	1.28
11.10.1994	1.15	11.01.1999	1.28	12.05.2003	1.21
15.11.1994	1.10	08.02.1999	1.25	16.06.2003	1.29
17.01.1995	1.05	08.03.1999	1.28	14.07.2003	1.27
14.02.1995	1.05	12.04.1999	1.23	11.08.2003	1.17
14.03.1995	1.13	10.05.1999	1.20	08.09.2003	1.25
11.04.1995	1.10	14.06.1999	1.23	13.10.2003	1.27
16.05.1995	1.10	12.07.1999	1.28		

Table S4. Data for eliciting calibration curve for ozone sensor C of the piloting stage. The linear calibration curve is obtained from linear regression: **C.indigo = 1.12*C.CSEM - 0.05.**

time [s]	C.CSEM [mg L^{-1}]	C.indigo [mg L^{-1}]
0	0.66	0.68
10	0.65	0.68
60	0.62	0.66
120	0.61	0.65
300	0.575	0.59
600	0.52	0.49
900	0.45	0.44
1200	0.41	0.39
1500	0.36	0.35
1800	0.34	0.33
2130	0.29	0.28
2700	0.25	0.25

Table S5. Total effect of calibration adjustments on average concentrations of sensors (sensor A: 0.45 mg L^{-1}; sensor B: 0.38 mg L^{-1})

date	adjustment sensor A [mg L^{-1}]	adjustment sensor B [mg L^{-1}]
04.09.2003	-0.023	0.000
15.10.2003	-0.033	-0.012
11.12.2003	0.049	0.063
15.01.2004	-0.066	-0.082
12.02.2004	-0.016	-0.014
11.03.2004	-0.065	-0.054
07.04.2004	0.112	0.109
06.05.2004	0.031	-0.020
15.07.2004	0.030	0.044
12.08.2004	0.057	0.032
09.09.2004	0.000	0.000
07.10.2004	-0.052	-0.075
18.11.2004	0.000	0.000
13.01.2005	-0.092	-0.067

Chapter 3

Marc B. Neumann, Urs von Gunten, Willi Gujer

Uncertainty in Prediction of Disinfection Performance

in Press *Water Research (2007)*

Uncertainty in Prediction of Disinfection Performance

Marc B. Neumann[1,2], Urs von Gunten[1,3], Willi Gujer[1,2]

[1]*Swiss Federal Institute of Aquatic Science and Technology, Eawag, 8600 Dübendorf, Switzerland*

[2]*Institute of Environmental Engineering, ETH Zurich, 8093 Zurich, Switzerland*

[3]*Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, 8092 Zurich, Switzerland*

ABSTRACT

Predicting the disinfection performance of a full scale reactor in drinking water treatment is associated with considerable uncertainty. In view of quantitative risk analysis this study assesses the uncertainty involved in predicting inactivation of *Cryptosporidium parvum* oocysts for an ozone reactor treating lake water. A micromodel is suggested which quantifies inactivation by stochastic sampling from density distributions of ozone exposure and lethal ozone dose. The ozone exposure distribution is computed with a tank in series model which is derived from tracer data and measurements of flow, ozone concentration and ozone decay. The distribution of lethal ozone doses is computed with a delayed Chick Watson model which was calibrated by Sivaganesan and Marinas (2005) utilizing a large number of inactivation studies. Parameter uncertainty is propagated with Monte Carlo simulation and the probability of attaining given inactivation levels is assessed. Regional sensitivity analysis based on variance decomposition ranks the influence of parameters in determining the variance of the model result. The lethal dose model turns out to be responsible for over 90% of the output variance. The entire analysis is re-run for three exemplary scenarios to assess the robustness of the results in view of changing inputs, differing operational parameters or revised assumptions about the appropriate model. We argue that the suggested micromodel is a versatile approach for characterization of disinfection reactors. The scheme developed for uncertainty assessment is optimal for model diagnostics and effectively supports the management of uncertainty.

KEYWORDS

Micromodel, stochastic model, variability, uncertainty, variance based sensitivity analysis, disinfection, *Cryptosporidium parvum*, ozone

1. INTRODUCTION

In this paper we investigate the uncertainty involved in predicting disinfection performance of a drinking water treatment plant. The capacity of a full scale pre-ozonation reactor with respect to inactivation of *Cryptosporidium parvum* oocysts is assessed. The reactor which is described in more detail in von Gunten *et al.* (1999) consists of a counter current bubble column for ozone transfer followed by two contact chambers for which the inactivation performance is predicted. Gujer and von Gunten (2003) developed a stochastic model in which they demonstrated that the deterministic approach typically applied leads to erroneous results at low inactivation levels in continuously fed reactors.

This study is a further development of this previous work and includes three significant new aspects:

Firstly, the stochastic micromodel developed in this paper does not require numerical integration (Gujer and von Gunten, 2003) but relies on stochastic sampling from density distributions of ozone exposure (reactor model) and lethal ozone dose (disinfection model).
Secondly, the uncertainty involved in the prediction of the inactivation level is quantified by coupling on-site information with uncertainty estimates from scientific literature. Uncertainty about the parameters of the ozone exposure model is gained from on-site data (Neumann *et al.*, 2007). Uncertainty associated with the disinfection model is obtained from Sivaganesan and Marinas (2005) who estimated parameters of a delayed Chick Watson model from a large number of inactivation studies.

Thirdly, based on the findings and experiences of this study, a general scheme is proposed for uncertainty assessment of a disinfection reactor (Figure 1). The "model building" section involves the selection of appropriate model structures and the characterization of parameter uncertainty ranges. "Variability analysis" performed for the micromodel quantifies the fraction of organism remaining active at the reactor outlet (Y_{Active}). In the "uncertainty analysis" the uncertainty propagation and quantile analysis of the cumulative density distribution $F(Y_{Active})$ are performed. "Sensitivity analysis" by variance decomposition ranks the parameters according to their influence in determining the variance of Y_{Active}. Finally a scenario analysis is suggested to test for the robustness of results towards changes in the model building section.

All calculations are performed within the statistical modelling environment of R (R Development Core Team, 2006).

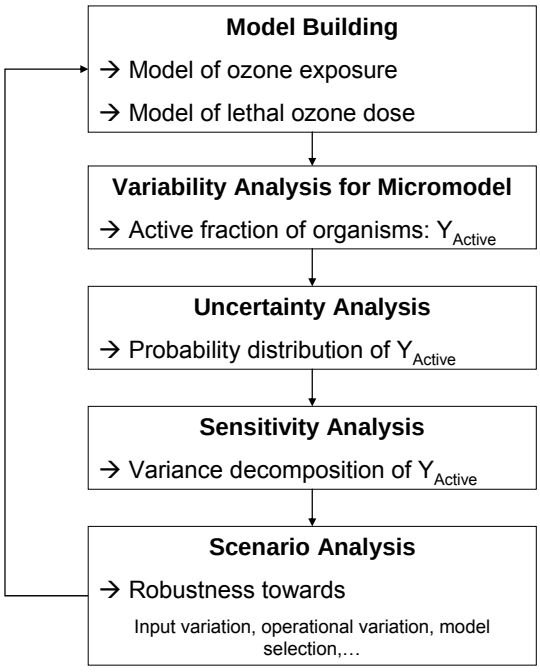

Figure 1. Scheme proposed for assessing uncertainty involved in predicting the fraction of organisms Y_{Active} remaining active at the reactor outlet.

2. MODEL BUILDING

2.1 Model of Ozone Exposure

The ozone exposure distribution characterizes the variability of possible exposures an organism may experience on its way through the contactor. It is obtained by coupling a stochastic transport model with a model for the ozone concentration profile. Applying a completely stirred tank reactor (CSTR) in series approach with equal sizing of compartments and 1st order decay of ozone the ozone concentration profile is quantified as

$$C_i = \frac{C_{inlet}}{\left(1 + k_{O3} \cdot \theta_{mean,\,i}\right)^i} \tag{1}$$

C_i is the steady state ozone concentration in CSTR compartment *i*, C_{inlet} the concentration at the inlet to the contact chambers, k_{O3} the first order decay constant of ozone and $\theta_{mean,\,i}$ the mean residence time of compartment *i*:

$$\theta_{mean,\,i} = \frac{V}{Q \cdot n_R} \tag{2}$$

where V is the total reactor volume, n_R the number of CSTRs and Q the flow through the reactor.

The residence time $\theta_{i,k}$ of organism *k* in compartment *i* is randomly drawn from the exponential residence time distribution:

$$f_i(\tau) = \left(\theta_{mean,i}\right)^{-1} \cdot \exp\left(-\frac{\tau}{\theta_{mean,i}}\right) \tag{3}$$

where $f_i(\tau)$ is the density function of residence time of compartment *i*.

The total ozone exposure $Ct_{Exposure,k}$ of organism *k* is the sum of exposures experienced in each of the n_R compartments:

$$Ct_{Exposure,k} = \sum_{i=1}^{n_R} \theta_{i,k} \cdot C_i \tag{4}$$

Computing total ozone exposure for a large number of organisms allows to characterize the ozone exposure distribution $f(Ct_{Exposure})$ of the reactor.

The analysis is performed for the reactor running at a temperature of 5°C at a given point in time assuming constant flow and ozone properties (Table 1). An extensive analysis on the uncertainty about the modelling of the ozone concentration and decay properties can be found in Neumann *et al.* (2007). The cited study was conducted for a pre-ozonation plant treating Lake Zurich water and analysed different sources of uncertainty including temporal variability of water temperature, pH and the water's DOC concentration. In contrast the present study is performed for a single moment in time so that only systematic measurement errors are included.

Table 1. On-site reactor properties at a water temperature of 5°C. Uncertainties about parameters are characterized by uniform (minimum and maximum) and normal (mean and standard deviation) distributions.

system property	symbol	units	mean	probability distribution	source of information
ozone concentration	C_{inlet}	mg L^{-1}	0.6	Uniform(0.5,0.7)	O_3 sensor
first order decay rate constant	k_{O3}	h^{-1}	2	Normal(2,0.2)	on-site sequencing batch reactor
volume	V	m^3	860	-	construction plans
flow	Q	m^3 h^{-1}	1200	Normal(1200,60)	MID - flow measurement
number of CSTR's	n_R	-	6	Uniform(4,8) Discrete	tracer experiments

Due to systematic calibration errors of the ozone sensor, the true concentration at the inlet of the first contact chamber (C_{inlet} = 0.6 mg L^{-1}) is specified as uniformly distributed between 0.5 and 0.7 mg L^{-1} (probability distributions in Table 1). The rate constant k_{O3} is obtained from an online sequencing batch reactor running on water extracted at the inlet to the contact chambers. The standard deviation of 0.2 h^{-1} accounts for uncertainty due to a possible offset in the calibration curve of the ozone sensor of the batch reactor. For the uncertainty of flow measurement Q a standard deviation of 5% of the measured value is assumed based on expert knowledge of the plant operator. Information on volume V is obtained from construction plans and is assumed to be known exactly. Based on the results from tracer experiments chains of n_R = 4 - 8 CSTR's are found to be equally good representations of the hydraulic reactor properties (results not shown).

2.2 Model of Lethal Ozone Dose

In the course of the USEPA's revision of the Surface Water Treatment Rule (USEPA, 2006) many studies on inactivation of *Cryptosporidium parvum* oocysts have been conducted. Inactivation is quantified as a function of disinfectant exposure (Ct) mainly by "in vitro" excystation or animal infectivity (Rennecker *et al.*, 1999; Driedger *et al.*, 2000; Li and Haas, 2004). Various models are proposed for characterizing inactivation. Among the most frequently selected models for inactivation of *Cryptosporidium parvum* oocysts with ozone is the delayed Chick-Watson model which is described by a lag phase Ct_{lag}, the minimal ozone dose required before inactivation begins and an inactivation rate constant k_D. The delayed Chick Watson model can be expressed as a random variable following a shifted exponential distribution. This distribution, expressing the variability of lethal ozone doses, is characterized in Equations 5 and 6 by it's density function $f(Ct_{Lethal\ Dose})$.

$$f\left(Ct_{Lethal\ Dose}\right) = 0 \qquad Ct < Ct_{lag} \tag{5}$$

$$f\left(Ct_{Lethal\ Dose}\right) = k_D \cdot \exp\left(-\left(Ct - Ct_{lag}\right) \cdot k_D\right) \qquad Ct \geq Ct_{lag} \tag{6}$$

Comprehensive analysis of parameter uncertainty for the delayed Chick Watson model was conducted by Sivaganesan *et al.* (2003) using data from several inactivation studies. Sivaganesan and Marinas (2005) performed a two step Bayesian analysis identifying temperature dependent parameters. They differentiated three types of uncertainty: i)

uncertainty in the regression, ii) predictive uncertainty due to oocyst lot variability and iii) predictive uncertainty due to experimental and measurement errors in inactivation studies. In accordance with USEPA (2006) and Rosen *et al.* (2004) we are not concerned about predicting the outcome of a future inactivation experiment but are interested in actual inactivation at the treatment plant. This is why we consider only i) uncertainty in the regression and ii) predictive uncertainty due to oocyst lot variability to be relevant (Equations 7-11).

$$k_D = \exp(34.9 - 10176 \cdot T_0 + N(0, S_1)) \qquad [mg^{-1}\ L\ min^{-1}] \qquad (7)$$

$$Ct_{lag} = \exp(-37.9 + 11064 \cdot T_0 + N(0, S_2)) \qquad [mg\ L^{-1}\ min] \qquad (8)$$

$$S_1 = \left(0.1 + \left(1.22 - 350 \cdot T_0\right)^2\right)^{0.5} \qquad (9)$$

$$S_2 = \left(0.3 + \left(2.82 - 816 \cdot T_0\right)^2\right)^{0.5} \qquad (10)$$

$$T_0 = (273.2 + T)^{-1} \qquad (11)$$

Parameters k_D and CT_{lag} are uncorrelated and temperature dependent random variables (eq 7, 8). The normally distributed random variables N(0, S_1) and N(0, S_2) have mean 0 and temperature dependent standard deviations S_1 and S_2 (eq 9, 10). T is the temperature in °C and T_0 is the inverse of the temperature in Kelvin (eq 11).

3. VARIABILITY ANALYSIS FOR MICROMODEL

The distribution of ozone exposure f($Ct_{Exposure}$) (histogram in Figure 2) is computed for the mean parameter values for Q, C_{inlet}, k_{O3}, n_R (Table 1) of the stochastic exposure model (eq 1-4). The distribution of lethal ozone doses required for inactivation (line in Figure 2) is derived for expectancies of k_D and Ct_{lag} (eq 7-11) of the delayed Chick Watson model (eq 5-6).

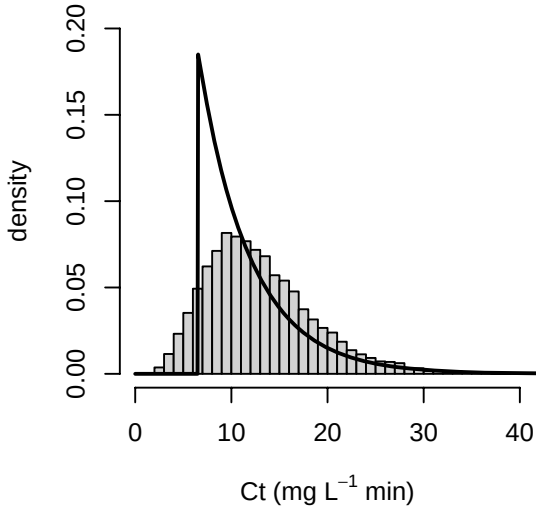

Figure 2. Ozone exposure distribution f($Ct_{Exposure}$) of reactor (histogram) and lethal ozone dose distribution f($Ct_{Lethal\ Dose}$) from delayed Chick-Watson model (line). Both density functions are obtained for mean parameter values of the micromodel at a temperature of 5°C: Q = 1200 m^3h^{-1}, C_{inlet} = 0.6 mg L^{-1}, k_{O3} = 2 h^{-1}, n_R = 6, k_D = 0.19 mg^{-1} L min^{-1}, Ct_{lag} = 6.5 mg L^{-1} min.

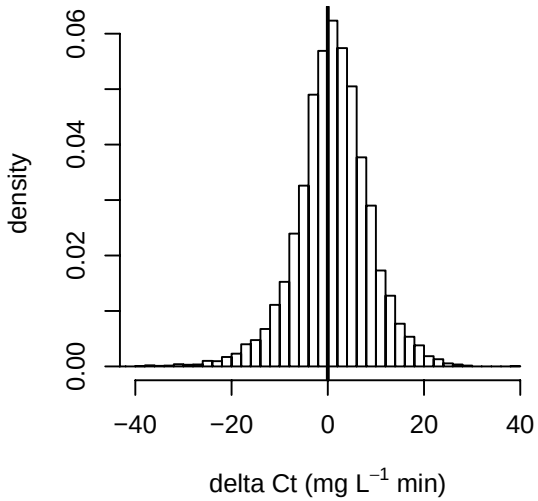

Figure 3. Density distribution of "exposure minus lethal dose" for n_{org} = 10000 organisms. Organisms that have experienced an ozone exposure smaller than the required lethal dose remain active. The fraction of organisms remaining active at the reactor outlet (fraction of organisms with negative delta Ct) is Y_{Active} = 0.42.

The fraction of organisms remaining active Y_{Active} is computed by Monte Carlo simulation assigning exposures and lethal ozone doses to a large number of organisms. Figure 3 shows the distribution differences (exposure minus lethal dose) for n = 10000 organisms. A sample size of n = 10000 is sufficient as the 95% confidence interval for any percentile estimate is within the +/- 1 percentile estimates (Morgan *et al.*, 1990). Thus the fraction of organisms remaining active (organisms with negative delta Ct) is computed as Y_{Active} = 0.42.

4. UNCERTAINTY ANALYSIS

Uncertainty analysis is performed to investigate the uncertainty associated with Y_{Active}. A 2D Monte Carlo (MC) simulation is run, with the stochastic inactivation model described above (n_{org}) as the inner MC loop and the uncertainty analysis as the outer MC loop (n_{UA}). Parameter sets are sampled from the distributions specified for Q, C_{inlet}, k_{O3}, n_R (Table 1) and for k_D and Ct_{lag} (Eq. 7, 8). Every parameter set produces its own density distribution for ozone exposure and lethal dose thus leading to different inactivation levels Y_{Active}. The presence of parameter uncertainty transforms Y_{Active} into a random variable. It is approximated by the empirical cumulative density distribution obtained with n_{UA} = 10000 runs (Figure 4). The expectancy for the active fraction is $E(Y_{Active})$ = 0.48 with standard deviation $V(Y_{Active})^{0.5}$ = 0.22. The analysis for the 5% and 95% quantile yields active fractions of 0.16 and 0.90, respectively.

The next section investigates which parameters are the most influential in determining the variance of Y_{Active}. It identifies which model components would need to be further explored to reduce overall uncertainty and which uncertain factors do not contribute to the output variance.

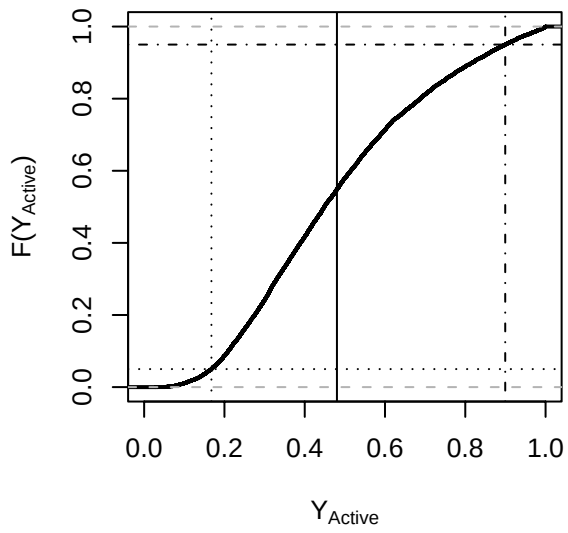

Figure 4. Empirical cumulative density distribution of Y_{Active} with mean (solid line) and 5% (dotted line) and 95% (dash-dotted line) quantiles. Obtained with n_{UA} = 10000 runs of the stochastic micro-model (n_{org} = 10000).

5. SENSITIVITY ANALYSIS

Typically local sensitivity measures are applied to assess the influence of parameters on output variables. They specify the linear dependency of the model output y(x) towards an incremental change of jth parameter x_j. An example of a local sensitivity measure is given in eq 12. It is based on calculating the derivative of the model variable in respect to each parameter. The derivative is taken at a given value of the parameter x_j^0 (typically the mean value) and normalized:

$$S_{j,local} = x_j^0 \cdot \left. \frac{\partial y(x)}{\partial x_j} \right|_{x_j = x_j^0} \tag{12}$$

In this study we propose to apply regional (often also termed global) sensitivity analysis. There are several reasons to prefer a regional over a local approach. As Y_{Active} is derived from a stochastic model the derivatives required for local sensitivity measures are not available and would need to be approximated by regression on results from a number of simulations. Regional sensitivity analysis is model free, the dependencies between variable and parameters can be non-linear and non-monotonous. It takes into account scale and shape of parameters density functions and allows for multidimensional averaging (the effect of the variation of a parameter is evaluated while all others are also varying). The main drawback is often the computational effort as these methods typically require MC simulations. There exist various approaches to regional sensitivity analysis: graphical analysis, correlation analysis, response surface analysis and variance decomposition (Saltelli, 2000). We perform a decomposition of the output variance and calculate 1st order sensitivity indices (McKay, 1997; Saltelli *et al.*, 2005). Eq 13 is the formal representation for the first order sensitivity index S_{Xj} of the jth parameter x_j. V and E denote the variance and expectancy operators with the subscripts making explicit over which parameters they are applied (-j standing for "all except parameter j") and Y and X are the model result and the parameters expressed as random variables:

$$S_{X_j} = \frac{V_{X_j}\left(E_{X_{-j}}\left(Y\middle|X_j\right)\right)}{V(Y)} \qquad (13)$$

The ranking of the S_{Xj}'s is the correct answer to the question: Which parameter, when set to its true value, will reduce output variance the most? In Saltelli (2004) this specific setting is termed "Factor Prioritization". A sampling based scheme for obtaining S_{Xj} is highlighted in Figure 5.

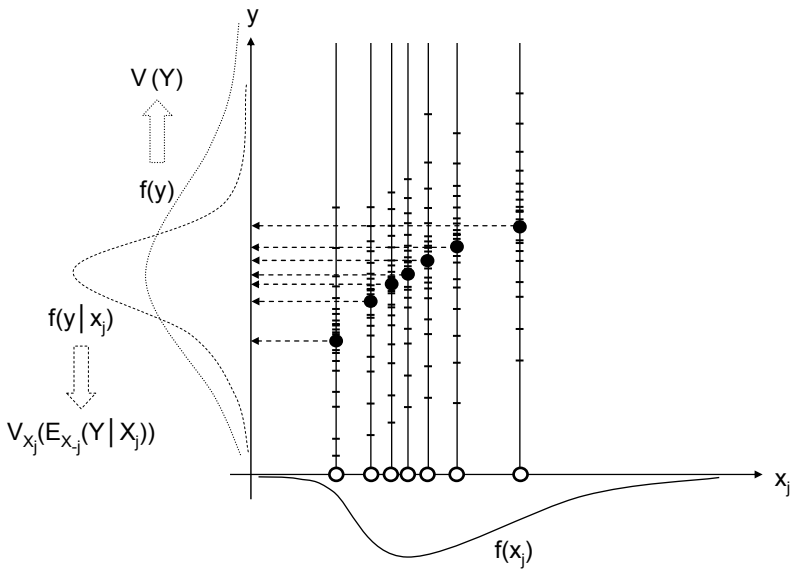

Figure 5. Scheme for regional sensitivity analysis by variance decomposition for jth model parameter:
1. Draw m_{sens} values (○) from marginal density distribution $f(x_j)$ of jth parameter x_j
2. For each (○) run n_{sens} MC simulations (-) by drawing samples for all other parameters
3. For each (○) calculate expectancy (●) of the n_{sens} model results (-)
4. Calculate variance of the m_{sens} expectancies (●): $V_{X_j}\left(E_{X_{-j}}\left(Y\middle|X_j\right)\right)$
5. Calculate variance of model results V(Y) from normal MC simulation
6. Calculate 1st order sensitivity index of parameter j: $S_{X_j} = \frac{V_{X_j}\left(E_{X_{-j}}\left(Y\middle|X_j\right)\right)}{V(Y)}$

The numerator of eq 13 is approximated by running a 2D Monte Carlo simulation which involves $n_{sens} \cdot m_{sens}$ ($n_{sens} = m_{sens} = 200$) simulations of the micromodel ($n_{org} = 10000$).

The resulting 1st order sensitivity indices $100\% \cdot S_{X_j}$ of the six parameters are presented as variance contributions (Figure 6). The model exhibits additive behaviour as $\sum_j S_{X_j} \approx 1$, i.e. all variance is explained by 1st order sensitivity indices.

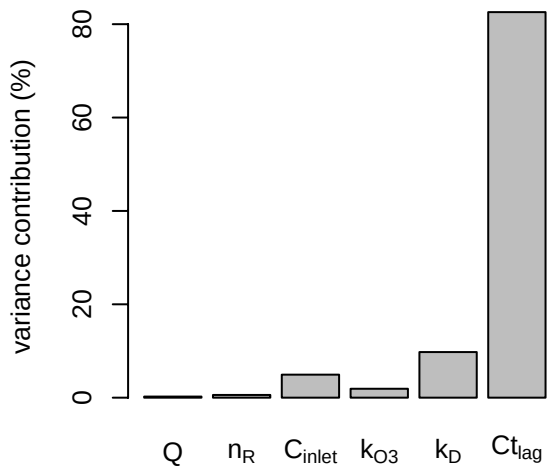

Figure 6. Variance decomposition: Contribution of uncertain parameters to the variance of Y_{Active}. The uncertain parameters of the disinfection model Ct_{lag} and k_D determine most of the output uncertainty: Ct_{lag}: 82.6%, k_D: 9.8%, k_{O3}: 1.9%, C_{inlet}: 4.9%, n_R: 0.6%, Q: 0.2%.

The variance of Y_{Active} is clearly determined by the uncertainty of the parameters from the disinfection model. Ct_{lag} is the most influential parameter contributing to over 80% of output variance. The result from the variance decomposition (Figure 6) indicates that reducing uncertainty of the parameters from the reactor model will not significantly reduce overall uncertainty.

6. SCENARIO ANALYSIS

By changing initial assumptions we analyse how results from uncertainty and sensitivity analysis are affected. It is typically the task of the analyst to creatively develop meaningful scenarios dependent on the context of the study. To show how this may be tackled we define three arbitrary scenarios. Scenario 1 examines how seasonal temperature changes affect the predicted performance. Scenario 2 investigates the impact of applying higher ozone dosages to obtain higher inactivation. In scenario 3 model diagnosis is performed: it is investigated how results are changed when variability between different oocyst strains (lot variability) is not considered. The results are summarized in table 2 with mean and quantiles (5% and 95%) of Y_{Active} together with the 1st order sensitivity indices.

6.1 Scenario 1: Effect of temperature increase from 5°C to 10°C

Effects of temperature change are automatically captured in the disinfection model whereas for the ozone exposure model the rate constant for ozone decay k_{O3} is expected to increase by a factor of $\exp(0.1 \cdot \Delta\text{Temp}) = \exp(0.1 \cdot (10°\text{C} - 5°\text{C})) = 1.65$ (Neumann *et al.*, 2007). The uncertainty distribution of the rate constant is expected to change to k_{O3} ~ Normal(3.3,0.3). Compared to the original setting $E(Y_{Active})$ is reduced by about 50% whereas the result from the variance decomposition remains practically unchanged (Figure 7, Table 2).

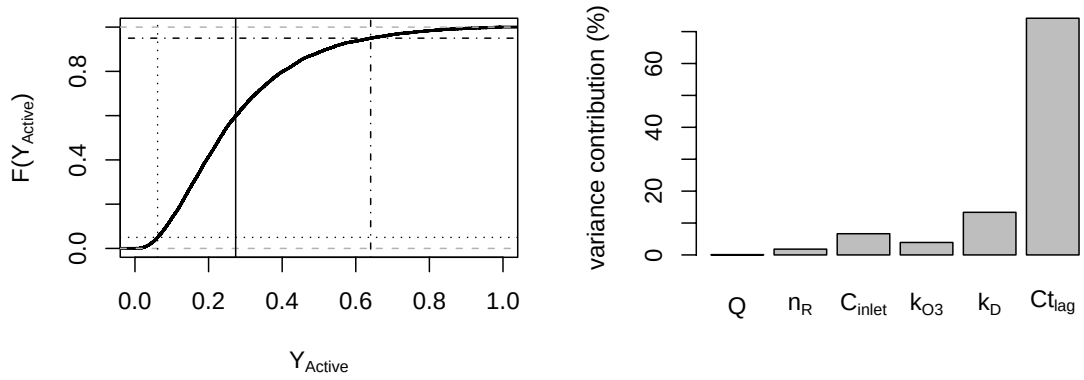

Figure 7. Cumulative density function for Y_{Active} with mean (solid) and 5% (dotted) and 95% (dash-dotted) percentiles and variance decomposition for a temperature increase of 5°C (Scenario 1).

6.2 Scenario 2: Effect of raising ozone dosage: C_{inlet} ~ Uniform(2.1, 2.3)

The absolute measurement error of the inlet ozone concentration is assumed to remain the same as in the original setting. Y_{Active} is now heavily skewed and its mean reduced to $E(Y_{Active}) = 0.036$ (Figure 8, Table 2). Most of the variance is still determined by the parameters of the disinfection model. For the reactor model the main influence is now to be found in the uncertainty about the hydraulic model structure n_R. To increase the probability of reaching a certain inactivation e.g. $F(Y_{Active} < 0.1) > 0.99$ one can either apply still more conservative operational settings (e.g. higher ozone dosage) to shift the empirical density distribution to smaller values or reduce variance at the present configuration by reducing the variance of one of the input parameters.

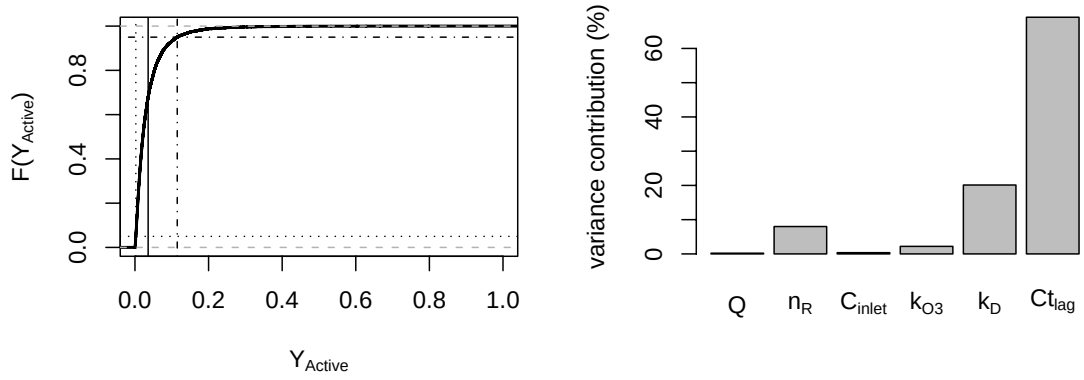

Figure 8. Cumulative density function for Y_{Active} with mean (solid) and 5% (dotted) and 95% (dash-dotted) percentiles and variance decomposition for an ozone dosage C_{inlet} ~ Uniform(2.1, 2.3) (Scenario 2).

6.3 Scenario 3: Effect of not considering oocyst strain variability

Scenario 3 explores the results when ignoring uncertainty between different oocyst strains. Eq 9 and 10 are replaced by eq 14 and 15 (Sivaganesan and Marinas, 2005).

$$S_1 = 1.22 - 350 \cdot T_0 \quad (14)$$

$$S_2 = 2.82 - 816 \cdot T_0 \quad (15)$$

The variance of Y_{Active} is significantly reduced showing that most of the uncertainty in the disinfection model is due to variation between oocyst strains. In this case the variance contributions of the exposure model (parameters Q, n_R, C_{inlet}, k_{O3}) are in a similar range as the contributions from the disinfection model (k_D and CT_{lag}). In this case the improvement of on-site information, i.e. reducing the error of the inflow concentration C_{inlet}, would significantly reduce the variance of Y_{Active} (Figure 9, Table 2).

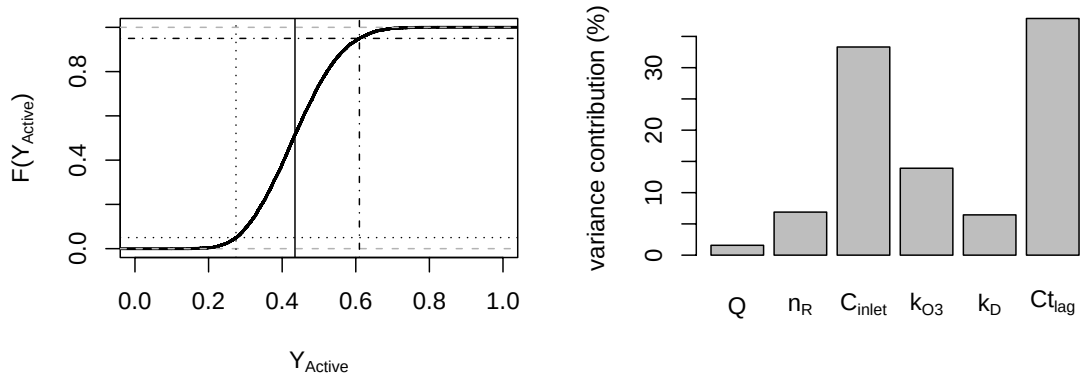

Figure 9. Cumulative density function for Y_{Active} with mean (solid) and 5% (dotted) and 95% (dash-dotted) percentiles and variance decomposition when not considering uncertainty due to lot variability of oocysts (Scenario 3).

Table 2. Summary of results from uncertainty analysis and variance decomposition

scenario	**characteristics of Y_{Active}**			**variance contribution** $100\% \cdot S_{X_j}$ [1)]					
	5% quantile	mean	95% quantile	Q	n_R	C_{inlet}	k_{O3}	k_D	Ct_{lag}
scenario 0: original setting	0.16	0.48	0.90	0.2	0.6	4.9	1.9	9.8	83
scenario 1: 10°C	0.06	0.27	0.64	0.1	1.8	6.6	3.9	13	74
scenario 2: higher ozone dosage	0.002	0.036	0.12	0.2	8.0	0.4	2.2	20	69
scenario 3: no oocyst strain variability	0.27	0.43	0.61	1.6	6.9	33	14	6.4	38

[1)] For all scenarios the model exhibits additive behaviour (the sum of 1st order sensitivity indices S_{Xj} add up to 1), i.e. all the variance of the result is explained by 1st order effects. This is not the case when interaction of model parameters cause higher order sensitivity indices to contribute to the variance. In the case of non-additive behaviour the ranking of 1st order sensitivity indices still remains the correct method to determine which parameter when set to its true value will reduce variance the most (Factor Prioritization) (Saltelli, 2004).

7. IMPLICATIONS FOR ENGINEERING PRACTICE AND REGULATION

The proposed methodology has implications for operators, designers and regulators.
Operators are able to run the scheme (Figure 1) for different operational conditions. They will in this way obtain time varying confidence levels and sensitivity indices, e.g. for seasonally varying temperatures (scenario 1). They can determine the expected decrease in uncertainty by reducing uncertainty of on-site parameters.

The designer is able to run the proposed scheme (Figure 1) as an optimization loop to determine ozone dosages required to obtain prescribed inactivation at specific confidence levels (scenario 2). Using the results from variance decomposition he can determine permissible uncertainty ranges of parameters. As an example, if uncertainty from sources that are outside the design domain dominate output uncertainty (such as uncertainty arising from the disinfection models) it may not be reasonable to invest in highly accurate measurement devices.

Scenario 3 examines the effect of oocyst strain variability on the uncertainty of the predicted inactivation. Such studies are of use to regulators when taking decisions on which databases or disinfection models to use for developing criteria.

Design or performance analysis for disinfection is typically conducted by normative procedures. The US-EPA provides guidelines for determining disinfectant exposure such as the "t_{10}" - or "CSTR" - method (USEPA, 2003). Tables specify log-inactivation of *C. parvum* oocysts as a function of ozone exposure and temperature (USEPA, 2006). Safety levels are implicitly included in these procedures. In contrast the approach suggested in this study explicitly assesses the uncertainty involved in the prediction of the inactivation level by taking the uncertainty of all model components into account simultaneously.

8. UNCERTAINTY OF MODELLING APPROACH

An uncertainty analysis is a subjective task. The assumptions in the model building section depend on the knowledge of the analyst and the availability of information, data, tools and resources. To analyse the effect of the assumptions made in the model building step one would need to compare uncertainty analyses conducted by different experts. Where could one expect deviations for this study? What would alternative modelling approaches be and how would they influence the results?

Concerning the inactivation kinetics the delayed Chick-Watson model is only one of an ensemble of possible models. Competing models such as Hom-, Selleck-, Series Event - models etc. (Gyurek and Finch, 1998) might also be applied to explain the inactivation data for which Sivaganesan and Marinas (2005) suggested the Delayed Chick-Watson model. A value of explicitly considering uncertainty is that by taking uncertainty into account with one modelling approach (here the delayed Chick-Watson) one automatically accounts for most of the possible outcomes if a competing model were used (assuming the same data is applied and statistical assumptions are fulfilled).

Concerning the hydraulic modelling two aspects are relevant: macromixing and micromixing. Macromixing is related to the appropriate characterisation of advection and turbulence in the reactor. Based on the reactor setup of this case study one might also be inclined to introduce some degree of circulation which however cannot be easily identified from the tracer curves. As shown by Craik (2005) a CSTR in series approach assumes an inherent degree of

micromixing lying between the extremes of completely segregated and completely micro-mixed flow and thus might overestimate inactivation if the flow tends towards complete micro-mixing. The sensitivity analysis shows however that for the setting in this study the uncertainty about the hydraulic model structure does not significantly contribute to the output variance (n_R in Figure 6), however at higher ozone dosages (resulting in higher inactivation) the hydraulic model structure becomes more important (Figure 8) and the relevance of micromixing is expected to increase.

9. CONCLUSIONS

- The stochastic micromodel is a versatile characterization allowing to couple arbitrary density distributions of ozone exposure and lethal ozone dose. Exposure distributions resulting from measurements with micro-spheres (Tang *et al.*, 2005) or from models based on computational fluid dynamics can be integrated into this framework. The approach can easily be extended to any segregated or particle related model.

- We merged site specific knowledge containing information on uncertainty about the exposure model with generally available knowledge from literature about uncertainty of the lethal dose model to obtain an informed uncertainty estimate of the predicted inactivation.

- Regional sensitivity analysis by variance decomposition is recognized to be more powerful than the traditionally applied local methods. It does not rely on derivatives and no assumptions on linearity, monotonousity or additivity need to be made.

- We demonstrate that the procedure developed for uncertainty assessment which consists of uncertainty analysis, variance decomposition and scenario analysis is useful for model diagnostics and enables an effective management of uncertainty.

REFERENCES

Craik S.A. (2005). Effect of Micro-Mixing Conditions on Predictions of Cryptosporidium Inactivation in an Ozone Contactor. *Ozone-Science & Engineering,* **27**(6), 487-494.

Driedger A.M., Rennecker J.L. and Marinas B.J. (2000). Sequential Inactivation of Cryptosporidium Parvum Oocysts with Ozone and Free Chlorine. *Water Research,* **34**(14), 3591-3597.

Gujer W. and von Gunten U. (2003). A Stochastic Model of an Ozonation Reactor. *Water Research,* **37**(7), 1667-1677.

Gyurek L.L. and Finch G.R. (1998). Modeling Water Treatment Chemical Disinfection Kinetics. *J. Environ. Eng.-ASCE,* **124**(9), 783-793.

Li L.J. and Haas C.N. (2004). Inactivation of Cryptosporidium Parvum with Ozone in Treated Drinking Water. *Journal of Water Supply Research and Technology-Aqua,* **53**(5), 287-297.

McKay M.D. (1997). Nonparametric Variance-Based Methods of Assessing Uncertainty Importance. *Reliab. Eng. Syst. Saf.,* **57**(3), 267-279.

Morgan M.G., Henrion M. and Small M. (1990) "Uncertainty a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis", Cambridge University Press, Cambridge etc.

Neumann M.B., von Gunten U. and Gujer W. (2007). Sources of Parameter Uncertainty in Predicting Treatment Performance – The Case of Pre-ozonation in Drinking Water Engineering. *Environ. Sci. Technol. (accepted for publication)*

R Development Core Team (2006). R: A Language and Environment for Statistical Computing.

Rennecker J.L., Marinas B.J., Owens J.H., and Rice E.W. (1999). Inactivation of Cryptosporidium Parvum Oocysts with Ozone. *Water Research,* **33**(11), 2481-2488.

Rosen J.S., Najm I., Sobrinho J., and Via S. (2004). Role of Variability in Design, Implementation and Interpretation of Microbial Inactivation Studies. *Water Science and Technology: Water Supply,* **2**(3), 205-211.

Saltelli A. (2000) "Sensitivity Analysis", John Wiley & Sons, Chichester.

Saltelli A. (2004) "Sensitivity Analysis in Practice a Guide to Assessing Scientific Models", John Wiley, Chichester.

Saltelli A., Ratto M., Tarantola S., and Campolongo F. (2005). Sensitivity Analysis for Chemical Models. *Chemical Reviews,* **105**(7), 2811-2827.

Sivaganesan M. and Marinas B.J. (2005). Development of a Ct Equation Taking into Consideration the Effect of Lot Variability on the Inactivation of Cryptosporidium Parvum Oocysts with Ozone. *Water Research,* **39**(11), 2429-2437.

Sivaganesan M., Rice E.W. and Marinas B.J. (2003). A Bayesian Method of Estimating Kinetic Parameters for the Inactivation of Cryptosporidium Parvum Oocysts with Chlorine Dioxide and Ozone. *Water Research,* **37**(18), 4533-4543.

Tang G., Adu-Sarkodie K., Kim D., Kim J.H., Teefy S., Shukairy H.M., and Marinas B.J. (2005). Modeling Cryptosporidium Parvum Oocyst Inactivation and Bromate Formation in a Full-Scale Ozone Contactor. *Environ. Sci. Technol.,* **39**(23), 9343-9350.

USEPA (2003) "LT2ESWTR Toolbox Guidance Manual - Proposal Draft".

USEPA (2006) National Primary Drinking Water Regulations: Long Term 2 Enhanced Surface Water Treatment Rule; Final Rule. Federal Register - National Archives and Records Administration. January 5, 2006.

von Gunten U., Elovitz M. and Kaiser H.P. (1999). Calibration of Full-Scale Ozonation Systems with Conservative and Reactive Tracers. *Journal of Water Supply: Research and Technology - Aqua,* **48**(6), 250-256.

Chapter 4

Co-Authored

Joerg Rieckermann, Marc B. Neumann, Christoph Ort, Jacco L. Huisman , Willi Gujer

Dispersion Coefficients of Sewers from Tracer Experiments

Water Science and Technology 2005, 52(5), 123-133.

Dispersion Coefficients of Sewers from Tracer Experiments

J. Rieckermann[*]*, M. Neumann*[*]*, C. Ort*[*]*, J.L. Huisman*[**] *and W. Gujer*[*]

[*]*Swiss Federal Institute of Environmental Science and Technology (EAWAG) and Swiss Federal Institute of Technology (ETH) 8600 Dübendorf, Switzerland*

[**]*Paques Biosyst BV, Met Grp, T Boerstr 24, POB 52, NL-8560 AB Balk, Netherlands*

ABSTRACT

In this paper, 60 tracer experiments in 37 different sewer reaches have been analyzed for longitudinal dispersion under dry weather flow conditions. It was found that dispersion coefficients of sewers are two to three orders of magnitude smaller than those measured in rivers and do not differ much from system to system. Suitable equations were identified to predict reasonable dispersion coefficients in sewer reaches with uniform geometry and stable flow conditions. For engineering applications that require a high degree of accuracy the performance of tracer measurements is recommended.

KEYWORDS

Transport; dispersion; tracer; mixing; sewer; modelling

INTRODUCTION

Taylor (1954) was the first to relate longitudinal mixing of solutes in flow to a Fickian-Type dispersion. In sewer transport modelling the 1D Advection-Dispersion equation (ADE) is generally applied as

$$\frac{\partial(Ac)}{\partial t} + \frac{\partial(Qc)}{\partial x} = \frac{\partial}{\partial x}\left(AK \frac{\partial c}{\partial x} \right) + AJ \qquad (1)$$

where A= wetted area, c= cross-sectional mean concentration, Q= discharge, K= longitudinal dispersion coefficient, J= sum of reaction rates of transformation processes, t= time, x= direction of mean flow.

Theoretically, the ADE is only valid in the *equilibrium zone*, where an equilibrium has become established between the effects of velocity shear and cross-sectional turbulent diffusion. In order to properly represent mixing in the so called *initial zone*, more complex models have to be applied ((Fischer et al. 1979), (Rutherford 1994), (Reichert and Wanner 1991)).
In spite of the rigorous framework that restricts its practical application, the ADE often proves useful for a wide range of applications such as integrated modelling ((Langeveld *et al.*, 2003), (Meirlaen *et al.*, 2001)), sewer process modelling ((Almeida *et al.*, 1999), (Huisman, 2001)) stochastic forecasting of specific substances (Ort *et al.*, 2004) and the experimental design of tracer experiments (Rieckermann *et al.*, 2004).

Although the choice of a proper dispersion coefficient is often crucial to obtain reliable results, sewer-specific data have only very recently been published ((Garsdal *et al.*, 1995), (Huisman *et al.*, 2000), (Boxall *et al.*, 2003) and (Langeveld *et al.*, 2003). However, no established

equation for the prediction of dispersion coefficients for urban sewer applications is presently available and in most cases predictions are questionable as will be explained further below.

The aim of this study is to suggest optimal methods for the analysis of tracer experiments and to propose a simple formula to rapidly predict dispersion coefficients from bulk data.
We will first describe the data base, then comment on the choice of models for data analysis and finally discuss the results with special regard to their uncertainty and possible limitations.

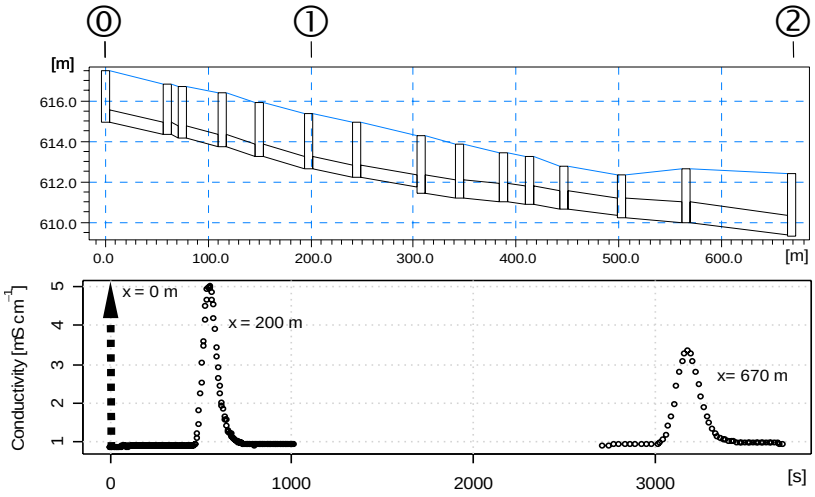

Figure 1. (top) Longitudinal profile of a sewer section. 0= dosing point of tracer (slug injection), 1= first monitoring point (outside the initial zone), 2= second measuring point at the end of the reach; (bottom) tracer curves at dosing and monitoring stations.

MATERIAL AND METHODS

In this section we discuss the data on which the study is based and present appropriate models for the estimation of dispersion coefficients.

Data base

A total of 60 experiments, which were all performed under dry weather conditions, was analysed with regard to dispersion. Mostly, sewers with diameters smaller than 2m were investigated in different European cities under a variety of flow conditions. We distinguish between two classes of data:

The first class contains all experiments which were conducted with the explicit intention of measuring dispersion coefficients. As mixing in the initial zone is not represented by the ADE, the tracer concentrations were measured at two points downstream from the injection point (Figure 1). These data are herein referred to as type "①-②". In our study we have 25 such datasets from 8 reaches with mostly constant diameters and stable flow conditions.

The second class contains tracer experiments which were not originally been designed for dispersion measurements. These datasets, which were mostly collected in the APUSS[1] project to quantify exfiltration with tracers (Rieckermann *et al.*, 2004), contain concentration profiles from the downstream measuring point but not from point ① and will be referenced by "⓪-②". For a better representativeness, well-documented data from literature (Johnson, 1944) was incorporated to give a total of 35 type ⓪-② investigations from 29 different sewer sections.

[1] APUSS (Assessment of the Performance of Urban Sewer Systems) homepage: www.insa-lyon.fr/apuss

All experiments (type ①-② and ⓪-②) were performed with slug injections of the tracer and in the majority of experiments the tracer curves were monitored using inline devices with a time resolution of a few seconds. Most experiments were performed with NaCl using electric conductivity measurements. Different fluorescent dyes were also used as tracers, but from a number of experiments only 8 of these datasets were identified as suitable for this study (Rhodamin WT: #52-55 and Uranine: #22,23, #44,45). When deriving dispersion coefficients from NaCl experiments, corrections for baseline conductivity were necessary.
For 34 of the experiments discharge measurements by the Area-velocity method (Doppler velocity measurements or MID) were available at the downstream measuring point. For the other experiments, the flow was estimated from the tracer data. In these cases the loss of tracer due to exfiltration or erroneous baseline corrections might well cause some degree of uncertainty, but this is assessed to be in the same range as the uncertainty induced by errors in the Area-velocity flow measurements.

All datasets are presented in Table 2 (Appendix) together with the estimated dispersion coefficients. The fact that for ⓪-② data only downstream concentration profiles were recorded complicates the data analysis and raises the question as to whether comparable information on longitudinal mixing can be derived from these.

In the following section we will first discuss different models for the estimation of dispersion coefficients from tracer curves and then present equations for the prediction of dispersion where no such data are available.

Estimating dispersion coefficients from tracer data
Dispersion coefficients are estimated from type ①-② experiments straightforward by applying a routing procedure. Here, the recently proposed method of Singh and Beck (2003) is used, which characterizes the downstream concentration profile as a function of the upstream concentration profile by a convolution integral

$$c_2(x,t) = \int_0^t c_1(\tau)\,\lambda(x,t-\tau)\,d\tau \qquad (2)$$

where c_1= upstream concentration at time τ and $\lambda(x,t-\tau)$= downstream response due to an instantaneous unit concentration upstream. The upstream concentration profile is described as a series of discrete input signals and the convolution represents the transformation during transport. The modelled downstream concentration curve (c_2) is composed from the superposition of the downstream response pulses which are full analytical solutions to the ADE. Functionally, the dispersion coefficient is contained in the downstream response function λ and can be estimated by fitting the modelled curve to the data. In this study, we use the notation K_S for longitudinal dispersion coefficients which were derived with the routing procedure by Singh and Beck (2003).

In contrast, experiments of type ⓪-② can only be used for the derivation of dispersion coefficients under the assumption that the impact of the initial period on the transport in the investigated reach is negligible. If the tracer mixes reasonably rapid over the whole cross section, the following analytical solution to (1), which is also known as the “Taylor solution”, (Rutherford, 1994) can be used as

$$c_2(x,t) = \frac{M}{A\sqrt{4\pi K_T t}} \exp\left(-\frac{(x-ut)^2}{4K_T t} \right) \quad (3)$$

where M= mass of injected tracer, K_T= longitudinal dispersion coefficient, derived with the Taylor solution.

These models are very useful for quantifying dispersion when field measurements are available. However, in the majority of cases tracer data is lacking and we may want to predict dispersion using information on sewer flow and geometry.

Predicting dispersion coefficients from velocity distributions

Transport in open channels is characterized by the full turbulent 3D velocity profile. Therefore, dispersion in itself is not a fundamental physical process, but arises from averaging the non-uniform velocity over channel depth and width. Consequently, analytical solutions for K can be found for assumptions on the mean velocity distributions in the direction of the flow and cross-sectional mixing. Elder (1959) was the first to transfer Taylor's result of turbulent flow (Taylor, 1954) to open channels; subsequently many other examples have followed.

Table 1 contains selected formulae that have since been developed (mostly cited from Seo and Baek (2004) and Seo and Cheong (1998)). Most explanatory comments are also taken from the same sources.

In order to properly apply these formulas to urban sewers, important differences between sewers and rivers need to be considered. As rivers are generally much wider than deep, the hydraulic radius R is often replaced by d (e.g. in the computation of the shear velocity u_*). In contrast we were using the general notation $u_* = \sqrt{gRS_0}$. Furthermore, all formulae except equation (6) were originally derived for rectangular cross-sections. In order to apply these formulae we transformed the maximum water depth from flow measurements in a circular pipe to the average water depth of a corresponding rectangular profile by $d= H= A/W$, where A is the wetted area and W the surface width.

Comparing predicted with estimated dispersion coefficients

In order to evaluate the performance of the equations presented in Table 1 we compared the measured and predicted dispersion coefficients for each dataset. For those experiments, which were performed in uniform reaches this is straightforward. However, to be able to apply the predictive formulae to non-uniform reaches we transformed them into "equivalent uniform reaches" with length-weighted mean diameter and discharge.

One must note that it is conceptually problematic to predict dispersion coefficients from average reach characteristics, because transport processes are non-linear and therefore the measured dispersion coefficient from a highly non-uniform reach is functionally neither related to a specific cross-section in this reach, nor to its length-weighted mean characteristics. Despite of this difficulty, we assume the transport properties (e.g. degree of partial filling, mean flow velocity) in sewers not to vary substantially during dry weather conditions.

Table 1. Selected formulas for the Longitudinal Dispersion Coefficient

Author	Eq.	Model	Comments
Elder (1959)	(4)	$K = 5.93 Hu_*$	Based on the assumption of a logarithmic vertical velocity profile. Neglects transversal velocity shear.
Parker (1961)	(5)	$K = 14.28 R^{3/2} \sqrt{2gS_0}$	Originally derived for pipelines. Parker also analyzed data from tracer experiments in sewers.
Sooky (1969)	(6)	$K = C_{l,c} R u_*$	Based on the assumption of a circular profile and logarithmic velocity profile. Coefficient $C_{l,c}$ is a complex integral that must be solved numerically.
McQuivey (1974)	(7)	$K = 0.058 \frac{Hu}{S_0}$	Based on similarity assumptions between the dispersion of solutes and that of hydraulic waves.
Fischer (1975)	(8)	$K = 0.011 \frac{u^2 W^2}{u_* H}$	Based on approximations of the transversal velocity profile and transversal turbulent mixing for irregular channels. Neglects vertical velocity shear.
Liu (1977)	(9)	$K = 0.18 \left(\frac{u}{u_*}\right)^{1/2} \left(\frac{W}{H}\right)^2 Hu_*$	Modification of equation (8) which is based on field data.
Magazine (1988)[2]	(10)	$K = 75.86 Ru \left(0.4 \frac{u}{u_*}\right)^{-1.632}$	Based on assumptions on roughness parameters.
Iwasa (1991)	(11)	$K = 2.0 \left(\frac{W}{H}\right)^{3/2} Hu_*$	No influence of friction term. Apparently tends to underestimate dispersion in streams.
Seo (1998)	(12)	$K = 5.92 \left(\frac{u}{u_*}\right)^{1.43} \left(\frac{W}{H}\right)^{0.62} Hu_*$	Modification of equation (8) based on dimensional analysis and field data.
Koussis (1998)	(13)	$K = 0.6 \left(\frac{W}{H}\right)^2 Hu_*$	No influence of friction term. Apparently tends to underestimate dispersion in streams.
Huisman (2000)	(14)	$K = 0.003 \frac{u^2 W^2}{Hu_*}$	Modification of equation (8) with assumptions for regular channels to better represent sewer mixing.
Deng (2001)	(15)	$K = \frac{0.15}{8\varepsilon_{t0}} \left(\frac{u}{u_*}\right)^2 \left(\frac{W}{H}\right)^{5/3} Hu_*$ $\varepsilon_{t0} = 0.145 + \left(\frac{1}{3520}\right)\left(\frac{u}{u_*}\right)\left(\frac{W}{H}\right)^{1.38}$	Modification of equation (8) with an extended term for transversal mixing.

A= wetted area and W= surface width, $H=d=A/W$ average water level in a circular pipe, R = hydraulic radius, S_0 = mean slope, u= average velocity, u_*= friction velocity

With regard to measured dispersion coefficients from non-uniform reaches, a dilution factor had to be introduced into the parameter estimation procedure that allowed for the scaling of the concentration profiles (measured under maximum flow at the downstream measuring point) to the average flow conditions of the equivalent uniform reach.

[2] Note: Eq.10 presented in Singh, S. K. and Beck, M. B. (2003) Dispersion coefficient of streams from tracer experiment data. *J. Environ. Eng.-ASCE*, **129**(6), 539-546 is incorrect, because the exponent -1.632 was omitted. This is the reason why the authors predict values of K which are two orders of magnitude bigger than those predicted with the other formulae.

As it is desirable to know where non-uniformities cause predicted dispersion coefficients to be meaningless, we computed the Manning-Strickler coefficient as an indicator to identify problematic datasets

$$k_{st} = uR^{-\frac{2}{3}} S_0^{-\frac{1}{2}} \tag{16}$$

where k_{st}= friction coefficient after Manning Strickler [$m^{1/3}$ s^{-1}]. Where k_{st} differs from reasonable values [$50 < k_{st} < 100$], it indicates either that the reach is highly non-uniform or that the data contain considerable error.

RESULTS AND DISCUSSION

Can the effect of the initial zone be neglected?

Selected data sets which best represent different flow conditions and fill grades were analysed for both the routing procedure and the Taylor solution. As sewer geometry and flow conditions are mostly constant, any significant difference in estimated dispersion coefficients should be related to the initial zone. We found the average relative deviation between K_T and K_S to be only -12% with no apparent bias, which justifies the application of the Taylor solution to the ⓪-② data (Table 2). Consequently, the estimated dispersion coefficients of both methods are considered equivalent and the impact of the initial zone can be neglected (Figure 2).

What are typical values for the dispersion coefficient in sewers?

Dispersion coefficients in urban sewer systems have been found to be very small and they are in the same order of magnitude for all datasets. The average value of the skewed distribution of K is 0.16 m^2s^{-1} [K_{10}= 0.05 m^2s^{-1}; K_{50}= 0.10 m^2s^{-1}, K_{90}= 0.36 m^2s^{-1}] (Figure 3). Impacts of biofilm, rain-events and sediments on the dispersion coefficient could not be evaluated from this database. However, it is observed that some tracer curves show pronounced tailing, which might correlate with the structural state. In rivers and canals, which are generally more irregular, observed dispersion coefficients are one or two orders of magnitude larger (Seo and Cheong, 1998).

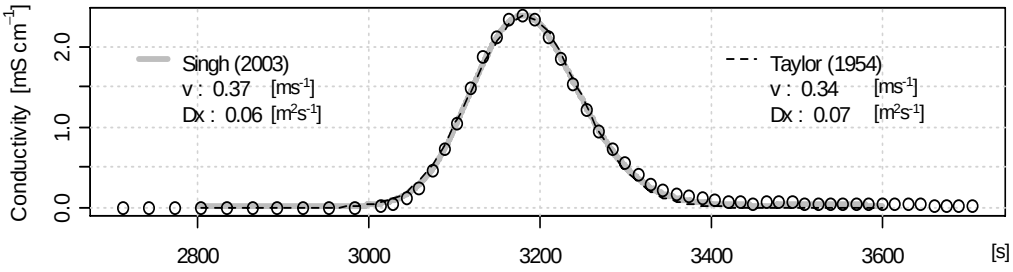

Figure 2. Observed (circles) and modelled (lines) concentrations that were obtained with the routing procedure after Singh and Beck (2003) and the Taylor solution for dataset 03

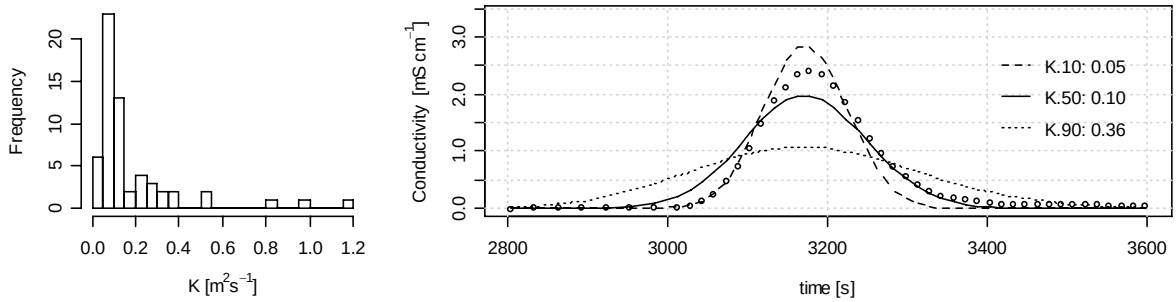

Figure 3. (left) Empirical distribution of *K* (n=60) from experimental data. Average dispersion coefficient: 0.16 $m^2 s^{-1}$ [K_{10}= 0.05; K_{50}= 0.10; K_{90}= 0.36]; (right) The predicted concentration profiles for different quantiles of *K* show the sensitivity of the concentrations to the dispersion coefficient for data set 03.

Can *K* be predicted from geometry and flow data?

Where a sewer reach has uniform characteristics, equations (11), (14) and (13) predict dispersion coefficients (K_p) that agree very well with the measurements (K_m) (Figure 4, black symbols). Where diameters, flow conditions or slopes are considerably different within a reach (grey symbols), predictions may differ by one or two orders of magnitude. However, the predictions are mostly accurate to ±0.3 [m^2s^{-1}] if a reasonable formula is chosen (Figure 4, right graphs). Whether this degree of variability is acceptable or not strongly depends on the individual application and the desired accuracy. The sensitivity of the downstream tracer concentrations of dataset 03 to the dispersion coefficient (Figure 3) indicates that peak concentration change moderately for the 10% and 90% quantiles of *K*. A reasonable alternative to the use of a predictive equation might be to simply choose a dispersion coefficient according to our database, because they include the irregularities (sediments, etc.) of real-life sewers. However, where the response of a reach has to be characterized precisely, the authors strongly advocate the performance of tracer measurements.

Discussion

First, our results were obtained under dry weather conditions and must not be used to describe transport during rain events. Boxall et al. (2003) reported dispersion coefficients under storm conditions that are an order of magnitude higher than in dry wheather. As transport processes under such extreme flow conditions might be completely different from the dry weather situation, it must even be questioned whether the ADE is a useful model for these situations ((Guymer and Obrien 1995), (Guymer et al. 1996)).

Second, it must be taken into account that concentrated solutions of NaCl show a greater density than wastewater, which complicates mixing and makes NaCl a non-conservative tracer in a hydraulic sense. Although we observed that the dispersion coefficients estimated from experiments with NaCl do not differ systematically from those obtained with dyes, the use of NaCl as a tracer clearly increases the mixing length which limits the investigation of short reaches.

Third, dispersion in a heterogeneous sewer line cannot be predicted from mean reach characteristics. The correct solution to this problem would be to subdivide the reach into locally uniform units, for which the suggested formulae should give reasonable results. However, on the one hand this would require a GIS based approach, which is quite the opposite of a simple formula to rapidly estimate dispersion. On the other hand this would still not take into account effects of eventual sediment deposits, bends, and other disturbances.

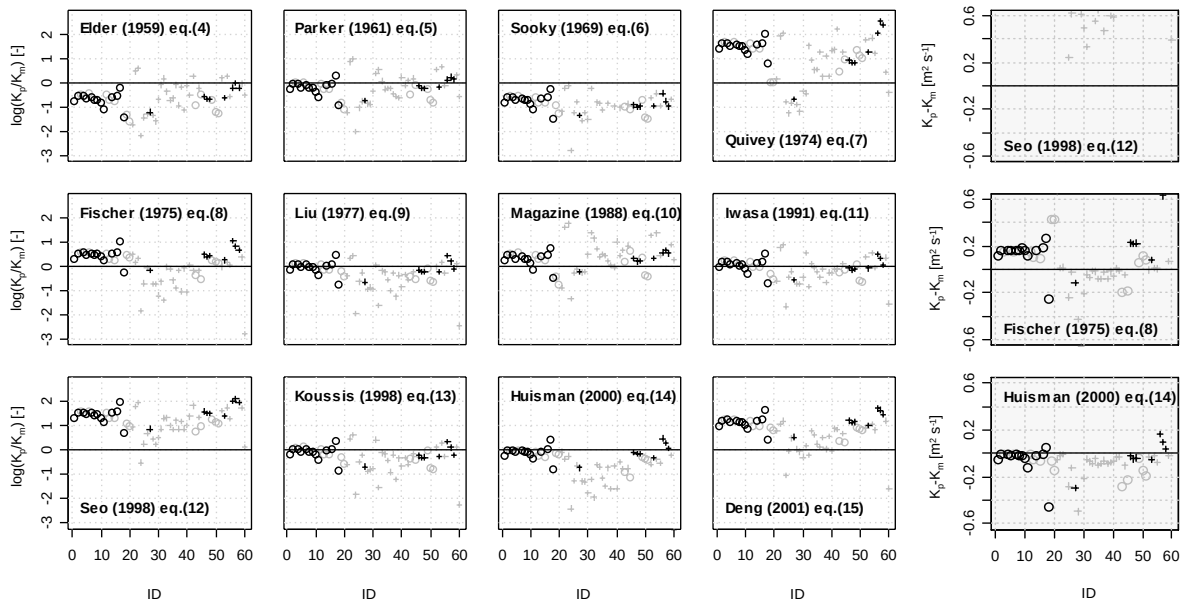

Figure 4. Discrepancy between predicted dispersion coefficients and values derived from field data. (left) relative differences; (right) absolute differences for three selected models. Circles label datasets which were analysed by the routing procedure and crosses datasets to which the Taylor solution was applied. Black symbols mark datasets with k_{st} values between 50 and 100 $m^{1/3}s^{-1}$, whereas grey colour is used for datasets with unreasonable k_{st} values. This points to considerable non-uniformities in a reach or bad data quality.

Such disturbances might be the reason why the equations of Huisman (2000), which was based on the assumption of a well-maintained straight sewer, and Sooky (1969) perform well for uniform reaches but tend to underestimate measured values of *K* for the majority of datasets. One could argue that therefore the estimated *K* from sewer data must be taken as an "operational" dispersion parameter which "contains" more information than only the effects of velocity shear and transversal mixing. Following this reasoning, large dispersion coefficients from tracer experiments might also indicate poor sewer performance.

Why was no predictive formula developed?

First of all, we suggest that the above mentioned formulae already perform satisfactorily in their range of application. In addition, searching the database for explanatory variables for *K*, only very weak relationships were found with the ratio *W/H* and the mean velocity in the reach. Besides collecting more datasets, future work towards the development of a predictive equation could concentrate on the analysis of sewer velocity profiles to obtain dispersion from vertical and lateral mixing processes in a similar way as Fischer did for rivers.

CONCLUSIONS

- 60 datasets from tracer experiments in different urban sewer systems have been analysed with regard to longitudinal dispersion. In the case that upstream and downstream concentration profiles were measured for a sewer reach, we suggest the routing procedure of Singh and Beck (2003) for data analysis. Where only downstream tracer measurements are available, the classical Taylor solution for the ADE can be a useful model.
- Dispersion in sewers was generally found to be very small and to show little variation in an absolute sense, although different sewer systems were investigated. We observe a skewed distribution of the longitudinal dispersion coefficient *K* with an average value of 0.16 m^2s^{-1} [K_{10}= 0.05 m^2s^{-1}; K_{50}= 0.10 m^2s^{-1}, K_{90}= 0.36 m^2s^{-1}].

- In situations of uniform geometry and stable flow, equations (11), (14) or (13) are suitable to predict the longitudinal dispersion from structural and hydraulic data. For non-uniform reaches, predictions may differ from measured values by a factor of 100, because the measured dispersion coefficient is not necessarily related to length-weighted average characteristics.
- Where the transport properties of a specific reach has to be described precisely, the authors advocate the performance of field experiments. As the dispersion coefficient contains integrated information on a reach, it is hypothesized that remarkable large dispersion coefficients could indicate poor sewer performance.
- For the development of a specific equation for the prediction of dispersion coefficients more datasets and detailed information on sewer velocity profiles are needed.

REFERENCES

Almeida, M. C., Butler, D. and Davies, J. W. (1999) Modelling in-sewer changes in wastewater quality under aerobic conditions. *Water Sci. Technol.,* **39**(9), 63-71.

Boxall, J., Shepherd, W., Guymer, I. and Fox, K. (2003) Changes in water quality parameters due to in-sewer processes. *Water Sci. Technol.,* **47**(7-8), 343-350.

Elder, J. W. (1959) The Dispersion of Marked Fluid in Turbulent Shear Flow. *J. Fluid Mech.,* **5**(4), 544-560.

Garsdal, H., Mark, O., Dorge, J. and Jepsen, S. E. (1995) Mousetrap - Modeling of Water-Quality Processes and the Interaction of Sediments and Pollutants in Sewers. *Water Sci. Technol.,* **31**(7), 33-41.

Huisman, J. L. (2001) Transport and transformation processes in combined sewers. Thesis, ETH Zürich, http://e-collection.ethbib.ethz.ch/diss/

Huisman, J. L., Burckhardt, S., Larsen, T. A., Krebs, P. and Gujer, W. (2000) Propagation of waves and dissolved compounds in sewer. *J. Environ. Eng.-ASCE,* **126**(1), 12-20.

Johnson, C. Frank (1944) Determination of Kutter's n for Sewers Partly Filled. *Transactions of the ASCE,* **109**(223-247.

Langeveld, J. G., Clemens, Fhlr and van der Graaf, Jhjm (2003) Interactions within the wastewater system: requirements for sewer processes modelling. *Water Sci. Technol.,* **47**(4), 101-108.

Meirlaen, J., Huyghebaert, B., Sforzi, F., Benedetti, L. and Vanrolleghem, P. (2001) Fast, simultaneous simulation of the integrated urban wastewater system using mechanistic surrogate models. *Water Sci. Technol.,* **43**(7), 301-309.

Ort, C., Schaffner, C., Giger, W. and Gujer, W. (2004) Modeling stochastic load variations in sewer systems. *In* "UDM04", Dresden.

Reichert, P. and Wanner, O. (1991) Enhanced One-Dimensional Modeling of Transport in Rivers. *J. Hydraul. Eng.-ASCE,* **117**(9), 1165-1183.

Rieckermann, Joerg, Borsuk, Mark, Reichert, P. and Gujer, W. (2004) A novel tracer method for quantifying sewer exfiltration. *Water Resources Research (submitted)*.

Rutherford, J.C. (1994). “River Mixing”. Wiley.

Seo, I. W. and Baek, K. O. (2004) Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams. *J. Hydraul. Eng.-ASCE,* **130**(3), 227-236.

Seo, I. W. and Cheong, T. S. (1998) Predicting longitudinal dispersion coefficient in natural streams. *J. Hydraul. Eng.-ASCE,* **124**(1), 25-32.

Singh, S. K. and Beck, M. B. (2003) Dispersion coefficient of streams from tracer experiment data. *J. Environ. Eng.-ASCE,* **129**(6), 539-546.

Taylor, G. (1954) The Dispersion of Matter in Turbulent Flow through a Pipe. *Proceedings of the Royal Society of London Series a- Mathematical and Physical Sciences,* **223**(1155), 446-468.

Acknowledgements
First of all the writers would like to thank everybody who contributed with their data, in particular the colleagues from Technical University of Dresden, IRSA-CNR, INSA Lyon and Middlesex University. This study has been carried out with financial support of the Swiss Federal Office for Education and Science (BBW) and within the framework of the European research project APUSS (Assessing Infiltration and Exfiltration on the Performance of Urban Sewer Systems) which partners are INSA de LYON (FR), EAWAG (CH), Technical University of Dresden (DE), Faculty of Civil Engineering at University of Prague (CZ), DHI Hydroinform a.s. (CZ), Hydroprojekt a.s. (CZ), Middlesex University (UK), LNEC (PT), Emschergenossenschaft (DE) and IRSA-CNR (IT). APUSS is supported by the European Commission under the 5th Framework Programme and contributes to the implementation of the Key Action "Sustainable Management and Quality of Water" within the Energy, Environment and Sustainable Development Contract n° EVK1-CT-2000-00072. The project APUSS is part of the CityNet, the network of European research projects on integrated urban water management.

Appendix

Table 2 Summary of hydraulic and dispersion data measured at 37 different sewers in Europe and the USA

ID	ID_S	Loc	$L_{(1-2)}$ [m]	$L_{(0-2)}$ [m]	Q [$m^3 s^{-1}$]	S0 [‰]	D_S [m]	d [m]	W [m]	R [m]	u [m s-1]	u_* [$m^2 s^{-1}$]	K_S [$m^2 s^{-1}$]	K_T [$m^2 s^{-1}$]	k_{st} [$m^{1/3} s^{-1}$]
1	A	ZH	1671	1788	0.028	0.90	0.90	0.11	0.68	0.10	0.37	0.03	0.12	0.11	60
2	B	ZH	469	574	0.028	0.90	0.90	0.11	0.69	0.10	0.38	0.03	0.07	0.08	60
3	B	ZH	974	1079	0.028	0.90	0.90	0.11	0.69	0.10	0.37	0.03	0.06	0.07	58
4	B	ZH	1564	1669	0.028	0.90	0.90	0.11	0.68	0.10	0.38	0.03	0.07	0.07	61
5	B	ZH	2015	2120	0.028	0.90	0.90	0.11	0.68	0.09	0.39	0.03	0.09	0.09	63
6	B	ZH	505		0.028	0.90	0.90	0.11	0.69	0.10	0.36	0.03	0.06	--	56
7	B	ZH	1095		0.028	0.90	0.90	0.11	0.68	0.10	0.38	0.03	0.08		61
8	B	ZH	1546		0.028	0.90	0.90	0.11	0.68	0.09	0.40	0.03	0.09		64
9	B	ZH	590		0.028	0.90	0.90	0.10	0.67	0.09	0.41	0.03	0.09		66
10	B	ZH	1041		0.028	0.90	0.90	0.10	0.67	0.09	0.42	0.03	0.13		69
11	B	ZH	451		0.028	0.90	0.90	0.10	0.66	0.09	0.44	0.03	0.21		73
12	C	ZH	469	574	0.031	0.90	0.90	0.12	0.70	0.10	0.38	0.03	0.07	0.07	58
13	C	ZH	974	1079	0.031	0.90	0.90	0.12	0.70	0.10	0.38	0.03	0.12	0.07	57
14	C	ZH	1564	1669	0.031	0.90	0.90	0.11	0.70	0.10	0.39	0.03	0.08	0.07	61
15	C	ZH	505		0.031	0.90	0.90	0.12	0.71	0.10	0.37	0.03	0.12		57
16	C	ZH	1095		0.031	0.90	0.90	0.11	0.69	0.10	0.40	0.03	0.07		62
17	C	ZH	590		0.031	0.90	0.90	0.11	0.68	0.10	0.43	0.03	0.03		68
18	C	ZH	1041		0.031	0.90	0.90	0.11	0.68	0.10	0.42	0.03	0.53		66
19	D	DI	850	1300	0.012	8.08	0.71	0.04	0.37	0.03	0.91	0.05	0.26	0.23	97
20	E	SA	680	1410	0.013	5.35	0.78	0.04	0.38	0.03	0.88	0.04	0.36	0.23	120
21	F	BA		340	0.009	3.33	0.80	0.03	0.37	0.03	0.79	0.03		0.32	142
22	G	DU		1350	0.040	1.84	0.70	0.21	0.69	0.15	0.27	0.05		0.02	22
23	G	DU		1350	0.040	1.84	0.70	0.21	0.69	0.15	0.28	0.05		0.02	23
24	H	DR		764	0.002	2.98	0.40	0.05	0.21	0.02	0.18	0.02		0.99	44
25	I	DR		1341	0.002	38.09	0.30	0.02	0.21	0.02	0.47	0.08		0.29	36
26	J	DR		1297	0.002	28.12	0.20	0.02	0.21	0.02	0.62	0.07		0.16	52
27	K	DR		1288	0.010	39.00	0.25	0.03	0.20	0.03	1.43	0.11		0.36	75
28	L	DR		655	0.008	32.69	0.25	0.04	0.21	0.03	0.92	0.10		0.52	49
29	M	DR		564	0.006	31.99	0.25	0.03	0.23	0.04	0.93	0.12		0.80	42
30	N	DR		1446	0.007	18.75	0.22	0.07	0.22	0.05	0.43	0.10		0.22	23
31	O	DR		1091	0.004	48.20	0.35	0.05	0.28	0.04	0.31	0.14		0.05	12
32	P	DR		730	0.016	4.00	0.10	0.23	0.10	0.15	0.68	0.08		0.07	38
33	Q	BE		1035	0.039	4.14	0.60	0.13	0.55	0.10	0.53	0.06		0.12	38
34	R	BE		1058	0.004	4.20	0.35	0.04	0.27	0.04	0.31	0.04		0.06	42
35	S	BE		640	0.008	3.90	0.40	0.08	0.33	0.06	0.30	0.05		0.11	31
36	T	BE		1955	0.032	3.20	0.40	0.28	0.35	0.06	0.32	0.04		0.09	37
37	U	BE		706	0.003	4.90	0.40	0.04	0.27	0.03	0.30	0.04		0.08	42
38	V	BE		1435	0.024	4.70	0.40	0.27	0.26	0.03	0.34	0.04		0.09	50
39	W	BE		789	0.003	4.00	0.30	0.03	0.27	0.03	0.35	0.04		0.10	56
40	X	BE		1220	0.012	1.80	0.30	0.18	0.30	0.12	0.22	0.05		0.06	22
41	Y	BE		1450	0.010	1.56	0.35	0.08	0.33	0.06	0.36	0.03		0.05	57
42	Z	BE		1460	0.041	1.24	0.45	0.32	0.39	0.07	0.33	0.03		0.03	57
43	AA	ZH	105	2130	0.029	2.88	0.90	0.11	0.69	0.10	0.38	0.05	0.32	0.35	34
44	AB	OP		1840	0.069	2.36	0.75	0.22	0.73	0.16	0.44	0.06		0.13	31
45	AB	OP	1576	1840	0.061	2.37	0.77	0.22	0.75	0.16	0.37	0.06	0.25	0.13	26
46	AC	EF		763	0.035	5.41	0.45	0.09	0.40	0.07	1.02	0.06		0.12	81
47	AC	EF		763	0.034	5.41	0.45	0.08	0.39	0.07	1.03	0.06		0.15	84
48	AC	EF		763	0.032	5.41	0.45	0.08	0.39	0.07	1.01	0.06		0.14	84
49	AD	ZH	469	574	0.016	0.90	0.90	0.09	0.64	0.08	0.28	0.03	0.09	0.09	49
50	AE	RM	225	724	0.022	0.90	1.20/ 2.10	0.10	0.50	0.08	0.57	0.03	0.24	0.19	103
51	AE	RM	225	724	0.022	0.90	1.20/ 2.10	0.09	0.50	0.08	0.58	0.03	0.29	0.23	105
52	AF	LO		150	0.040	1.73	0.60	0.21	0.60	0.14	0.32	0.05		0.04	28
53	AF	LO		660	0.031	1.73	0.60	0.11	0.53	0.09	0.53	0.04		0.11	62
54	AF	LO		150	0.034	1.73	0.60	0.19	0.59	0.13	0.31	0.05		0.03	29
55	AF	LO		450	0.027	1.73	0.60	0.12	0.54	0.10	0.42	0.04		0.11	48
56	AG	LV		473	0.206	0.87		0.24	1.47	0.15	0.58	0.04		0.09	69
57	AH	LV		272	0.515	0.49		0.48	1.75	0.25	0.61	0.03		0.11	68
58	AI	LV		323	0.648	0.50		0.51	1.60	0.33	0.80	0.04		0.20	75
59	AJ	LV		80	0.024	32.40		0.07	0.27	0.05	1.19	0.12		0.06	52
60	AK	LY		906	0.039	33.00	0.58	0.37	0.16	0.09	0.68	0.17		1.18	19

ZH= Zurich (CH), DI= Dillhaus (CH), SA= Saaland (CH), BA=Bauma (CH), DU= Dübendorf (CH), DR= Dresden (D), BE= Berlin (D), OP=Opfikon (CH), EF= Effretikon (CH), RO= Rome (I), LO= London (GB), LV= Louisville (USA), LY= Lyon (F)
ID= ID of dataset, ID_S= ID of sewer, Loc= Location of sewer, L= Length, D_S= Diameter of sewer, S_0= slope, Q= Discharge, R= Hydraulic radius, W= width, d= average water level (in a circular pipe), u= average velocity, u_*= shear velocity, K_S = Dispersion coefficient (Singh and Beck, 2003) K_T= Dispersion coefficient (Taylor, 1954) k_{st}= Manning-Strickler friction coefficient.

Chapter 5

Co-Authored

Helge Daebel, Joerg Rieckermann, Marc B. Neumann, Willi Gujer

Towards estimating parameter uncertainty under model structure deficits

submitted to *Water Research*

Towards estimating parameter uncertainty under model structure deficits

Helge Daebel, Jörg Rieckermann, Marc B. Neumann, Willi Gujer

Swiss Federal Institute of Aquatic Science and Technology, Eawag, 8600 Dübendorf, Switzerland
and Institute of Environmental Engineering, ETH Zurich, 8093 Zurich, Switzerland

ABSTRACT

Reliable information on parameter uncertainty is vital when conducting uncertainty analysis for environmental models. Yet, there appears to be a general dilemma: modern technology permits the recording and storing of accurate data with a time resolution that is larger than the time constants of established models. This mismatch leads to autocorrelated residuals, i.e. that assumptions of traditional inference methods are violated. As a consequence, parameter uncertainty is underestimated. Model structures should be extended until the assumptions are met. However, in engineering, this is often only feasible to a certain degree. In this paper, we introduce five methods that extend ordinary least-squares regression and potentially lead to a conservative estimate of parameter uncertainty in the face of autocorrelated residuals. We investigate the practical applicability of these methods in a representative case study - a respirometric experiment to estimate a Monod-kinetic parameter. We conclude that most of the methods give reliable results and that they are mostly well applicable in practice. Yet, the approaches are pragmatic, partly lacking a formal statistical basis. But our contribution should stimulate a necessary scientific debate.

KEYWORDS

autocorrelation, parameter uncertainty, model structure uncertainty, Monod-kinetics, residual analysis, respirometry

1. INTRODUCTION

Environmental decision making is often based on computer aided simulation studies. Since poor decisions may lead to major investments or be the cause of serious ecological damages, uncertainty analysis has become an important constituent of environmental modeling (Beck, 1987; Goodman, 2002; Reckhow, 1994; Reichert and Vanrolleghem, 2001; Reichert and Borsuk, 2005; Vanrolleghem and Keesman, 1996). A prerequisite for obtaining interpretable results from such an analysis is the reliable estimation of uncertainties associated to each modeling step. Model parameters are crucial in this respect and their uncertainty shall be the focus of this study.

The preferable way of estimating parameter values and their uncertainty is by applying statistical inference between model and measured data. If standard methods, e.g. least-squares regression, are applied, their results are only meaningful if underlying assumptions hold: (1) a correct model, (2) random observation errors and thus (3) independent and identically distributed (iid) residuals (Bates and Watts, 1988).

In environmental modeling, however, these assumptions are often violated but consequences neglected. Modern sensor and storage technology allow for very accurate measurements with nearly unlimited temporal resolution. This mostly conflicts with the time constants of our mostly empirical models, leading to systematic patterns (autocorrelation) in the residuals (see e.g. Rieckermann et al., 2005). Hence, there appears to be a general dilemma in environmental modeling and uncertainty analysis: The more accurate and frequent measurements become, the more often standard regression fails. The critical aspect of this situation is that parameter uncertainties are underestimated. Because of autocorrelation in the residuals, the information content of the data is overestimated by the regression algorithm. Fox, 1997) nicely illustrates this fact with an example: the variance of a sample mean m is usually estimated by $\sigma_m^2 = \sigma^2/n$, where σ^2 is the variance of n independent random samples. If instead observations are sampled from a first-order autoregressive process (ρ = first-lag autocorrelation parameter) the correct variance of m is

$$\sigma_m^2 = \frac{\sigma^2}{n} \cdot \frac{1+\rho}{1-\rho}$$

In case of strong autocorrelation ($\rho \to 1$), "real" uncertainty about m is substantially larger than estimated by σ^2/n. In other words, the "effective" number of "independent" observations is $n(1-\rho)/(1+\rho)$ rather than n. In engineering reality, autoregressive patterns in residuals are mostly of much higher order. The resulting severe underestimation of parameter errors may result in overconfidence regarding the accuracy of the model result - a situation that ought to be avoided in engineering applications.

While statisticians argue that model structures should be improved until iid residuals are reached, we see significant limitations for doing so in engineering practice: (i) Limited resources (time/money and knowledge) often impede elaborate model structure adaptation, e.g. by adding a sensor error model. (ii) More complex models easily have too many parameters to be identifiable. (iii) Finally, environmental engineers are often bound to use certain model structures that are either well established, implemented in software tools or even accepted as state-of-the-art from a juridical point of view. Consequently, there is a need for approaches that ensure conservative estimation of parameter uncertainty under imperfect model structures causing systematic patterns in residuals.

In this paper, we introduce five methods that potentially lead to a conservative estimate of parameter uncertainty in the face of autocorrelated residuals. All approaches extend ordinary least-squares (OLS) which is the most popular regression method (Hastie, 2003). We investigate the practical applicability of the methods in a representative engineering case - a respirometric batch experiment to estimate the oxygen half-saturation constant of ammonium oxidizing bacteria using Monod kinetics. Our goal is not to present methods that assign "true" levels of uncertainty to a parameter. Instead, they are rather pragmatic approaches that should lead to a conservative result in the sense that uncertainties are not underestimated according to risk-averse attitude in engineering.

The article is structured as follows: First, a short review is given on existing methods reported from different research areas and terminology is defined with respect to least- squares regression. Second, we present the five methods, which can roughly be differentiated in three categories: (i) artificially enlarging the sum of squares obtained from the OLS regression, (ii) reducing the number of original data points or (iii) reflecting systematic observation "errors"

directly onto the parameter of interest. Third, we describe the experimental setup and model of the respiration experiment. In the fourth section, we present the results of the five estimation procedures and conduct a critical evaluation of the methods regarding their usefulness for engineering applications. Finally, the main findings from this work are concluded.

2. METHODS

Literature reviews as well as expert interviews were conducted in the field of statistics, econometrics and geodesy. The consulted statistician acknowledged the above stated problem (Stahel, 2000 p. 349) and argued that model structure should be extended until residuals are iid. An expert in econometrics was interviewed because the general idea of bootstrap (Efron and Tibshirani, 1993) or the recently developed subsampling methods (Politis et al., 1999) appeared to be very promising. The expert was not convinced of the direct use of the methods to our purposes since they only allow for nearly iid residuals and weak instationarity. Still, their basic concepts were adapted for the development of the formalized subsampling approach presented in section 2.4. A geodesist pointed us towards the *collocation* method which is equivalent to the generalized least-squares approach presented in section 2.6. Another approach that was reviewed was robust regression promoted by Rousseeuw, 1984; Rousseeuw and Leroy, 2003) suggesting least median squares for robust parameter estimation. But the term is used in the common statistical sense of robustness towards outliers and thus does not solve the current problem. To conclude, except for the generalized least-squares approach, our quest in the areas of expertise of uncertainty estimation did not convey appropriate methods. As a result, we developed or enhanced the least-squares extending methods presented in the following sections.

2.1 Ordinary least-squares regression

In this subsection we briefly summarize the statistical framework of ordinary least-squares regression, defining the terminology used in the paper (compare Bates and Watts, 1988; Draper and Smith, 1998). Equation 1 is the abstraction of an arbitrary time-related model.

$$\mathbf{Y}_M = f(\boldsymbol{\Theta}, \mathbf{t}) \qquad \text{(Eq.1)}$$

with $\mathbf{Y}_M$ output vector from the deterministic model f
$\boldsymbol{\Theta}$ parameter vector of the deterministic model f
$\mathbf{t}$ time vector, the independent variable of model f

Measured values are assumed to be represented by this deterministic model plus an additional error which is purely random, i.e. independent and identically (LS: normally) distributed.

$$\mathbf{Y}_{meas} = f(\boldsymbol{\Theta}, \mathbf{t}) + \mathbf{E} \qquad \text{(Eq.2)}$$

with $\mathbf{Y}_{meas}$ observation vector
$\mathbf{E}$ observation error vector

Residuals are the difference between observed and calculated values of the dependent variable.

$$\mathbf{R} = \mathbf{Y}_{meas} - \mathbf{Y}_M \qquad \text{(Eq.3)}$$

where **R** is the residual vector representing the best estimate of the observation error vector **E** if unknown (Stahel, 2000). In case of least-squares regression where the model is assumed to be true, the elements of **R** ought to have the same characteristics as those of **E** - independent and identically (normally) distributed.

Parameter estimates are derived through minimizing the least-squares objective function.

$$\boldsymbol{\Theta}_{est,\,LS}\,(\mathbf{Y}_{meas}): \qquad SS\,(\mathbf{Y}_{meas}\,,\,\boldsymbol{\Theta}_{LS}) = \min \qquad \text{(Eq.4)}$$

where SS stands for the sum of squared residuals ($\Sigma\, R_i^2$), $\boldsymbol{\Theta}_{\mathbf{LS}}$ represents the parameter vector to be altered until the optimum is reached and $\boldsymbol{\Theta}_{\mathbf{est,\,LS}}\,(\mathbf{Y}_{\mathbf{meas}})$ is the set of best parameter estimates based on measurements, the given model structure and the least-squares objective function. The covariance matrix of the estimated parameters needs to be approximated numerically (compare Checchi and Marsili-Libelli, 2005; Dochain and Vanrolleghem, 2001; Liu and Zachara, 2001). According to the so called FIM (Fisher Information Matrix)-Method it can be calculated as

$$\mathbf{COV}\,(\boldsymbol{\Theta}_{est}) \approx (\,\mathbf{J}^{\,T}\,\mathbf{V}^{\,-1}\,\mathbf{J}\,)^{\,-1} \qquad \text{(Eq.5)}$$

where **J** represents the Jacobian matrix (derivatives of calculated output states towards estimated parameters) and **V** the covariance matrix of observation errors. In case of LS and with measurement errors unknown but assumed to be iid (no off-diagonal elements of **V** and V_{ii} equal for all i data points), the covariance matrix can be approximated by

$$\mathbf{COV}\,(\boldsymbol{\Theta}_{est,\,LS}) \approx SS \cdot (n - m)^{-1} \cdot (\,\mathbf{J}^{\,T}\,\mathbf{J}\,)^{\,-1} \qquad \text{(Eq.6)}$$

where n is the number of measured values (data points) and m the number of parameters to be estimated. The standard error se of parameter j is

$$se(\boldsymbol{\Theta}_{est,\,LS,\,j}) = \mathbf{COV}\,(\boldsymbol{\Theta}_{est,\,LS})_{\,j,j}^{\;\;1/2} \qquad \text{(Eq.7)}$$

Especially equations 6 and 7 will be referred to frequently during the following introduction of methods that extend LS regression in case of autocorrelation: whitening, sum of squares correction, formalized subsampling, mirroring and generalized least-squares.

2.2 Whitening

***Idea*:** Additional noise is used as a penalty function for systematic patterns in residuals (SS is increased in equation 6). The variance of the noise is increased until resulting residuals $\mathbf{R}_{W,\,LS}$ meet statistical iid assumptions.

***Realization*:** *Pre-whitening (PrW):* Noise is added to the data before parameter estimation is performed. A sample of size n (= number of measured values) is drawn from a normal distribution with $\mu_{noise} = 0$ and $\sigma_{noise,start} \approx sd(\mathbf{R}_{LS}) = [(n-1)^{-1}\,\Sigma(R_{LS,i} - R_{LS,mean})^2\,]^{-1/2}$. The noise is added to the observed values $y_{meas,\,1\,..\,n}$ and least-squares regression is performed leading to $\mathbf{R}_{PrW,\,LS}$. The value of σ_{noise} is raised until iid assumptions in $\mathbf{R}_{PrW,\,LS}$ are met. This state is defined by the following: If the addition of a random sample from the noise distribution $N(0, \sigma_{noise,opt})$ to the original data points is repeated often enough (e.g. u = 10'000 times), then the fraction of all u $\mathbf{R}_{PrW,LS}$ for which independence and normality tests are fulfilled reaches the same fraction (e.g. 90% or 9000) as if the same tests were applied to u random samples of size n from any normal distribution.

Post-whitening (PoW): The noise is added to the residuals, i.e. after parameter estimation. The procedure makes use of the fact that the best estimate from least-squares regression is already an unbiased estimator (Stahel, 2000) and therefore does not have to be re-estimated. Thus, only a new standard error is calculated by a posteriori replacing SS_{LS} in equation 6 and 7 with SS_{PoW} by the following equation.

$$se_{PoW,LS} = se_{LS}\sqrt{\frac{SS_{PoW}}{SS_{LS}}} \qquad \text{(Eq.8)}$$

The level of noise (σ_{noise}) to be added is found with the same concept as outlined for the pre-whitening approach. Note, however, that no further parameter estimations are required as the analysis adds the noise to $\mathbf{R}_{LS}$ instead of to the data. This makes it much less computationally demanding.

2.3 Sum of squares correction for autocorrelation

***Idea*:** In the presence of systematic patterns in the residuals $\mathbf{R}_{LS}$ (eq. 3) the standard error from LS (eq. 7) is augmented by "correcting" SS in equation 6. A new sum of squares SS_{SSC} is calculated by treating errors at regions of autocorrelation as dependent and not as random. The rationale is that data points contain less information if they lead to autocorrelated residuals and are therefore treated rather as clusters. The method is founded on the "Estimated probable error" approach from Berar and Lelann, 1991.

***Realization*:** Equations 9 to 12 have to be executed to receive the corrected standard error se_{SSC} (Θ_{est}) for estimated parameters. The corrected sum of squares SS_{SSC} treats residuals with apparent local correlation differently (square of sum) than the ones with none (sum of squares = LS). If there is a sign shift between residuals they are regarded as independent and therefore in a common LS sense incorporated into SS_{SSC}. Otherwise, they are treated as linearly dependent to a certain degree z representing a "probability of local correlation". This is done by applying the χ^2- distribution with two degrees of freedom representing the sum of squares of two independent, normally distributed values. Factor z is set equal to the probability that this ideal distribution takes a value smaller or equal to the actual sum of two consecutive residuals (compare eq. 12). Thus, if this actual sum is large, the estimated probability z is close to one and residuals are treated as almost entirely linear dependent and if very small, almost completely independent because z is close to zero.

$$se_{SSC} = se_{LS}\sqrt{\frac{SS_{SSC}(\mathbf{R}_{norm})}{SS_{LS}(\mathbf{R}_{norm})}} \qquad \text{(Eq.9)}$$

with

se_{SSC}	corrected standard error
se_{LS}	standard error from ordinary LS regression
SS_{LS} ($\mathbf{R}_{norm}$)	sum of squares of normalized residuals from LS (eq. 10 and 11)
SS_{SSC} ($\mathbf{R}_{norm}$)	corrected sum of squares of normalized residuals (eq. 10 and 12)

$$\mathbf{R}_{norm} = \frac{\mathbf{R}_{LS}}{sd(\mathbf{R}_{LS})} \qquad \text{(Eq.10)}$$

with $\quad sd(\mathbf{R}_{LS}) = \sqrt{\frac{1}{n-1}\sum\left(r_{LS,i} - \bar{r}_{LS}\right)^2}$

$\mathbf{R}_{LS}$ LS residuals according to eq. 3

$$SS_{LS}(\mathbf{R}_{norm}) = \sum_{i=1}^{n} r_i^2 \qquad \text{(Eq.11)}$$

$$SS_{SSC}(\mathbf{R}_{norm}) = \sum_{i=1}^{n}\left(1 - z_i\right)\cdot r_i^2 + \sum_{\mathbf{Q}}\left(\sum_{i=q}^{q+h_q} z_i r_i\right)^2 \qquad \text{(Eq.12)}$$

with $\quad z_i = \begin{cases} 0 & \text{if} \quad r_i \cdot r_{i-1} \le 0 \quad \text{(case 1)} \\ F_{\chi^2(v=2)}\left(r_i^2 + r_{i-1}^2\right) = 1 - e^{-\left(r_i^2 + r_{i-1}^2\right)/2} & \text{if} \quad r_i \cdot r_{i-1} > 0 \quad \text{(case 2)} \end{cases}$

z_i correction factor at i^{th} residual of $\mathbf{R}_{norm}$ (Def.: $z_1 = 0$)
r_i i^{th} element of normalized residual vector $\mathbf{R}_{norm}$
n number of data points
Q vector of indices i where z switches from case 1 at z_{i-1} to case 2 at z_i
q a single element of **Q**
h_q number of consecutive cases 2 in z_i with i > index q
$F_{\chi^2(v=2)}$ χ^2-cumulative distribution function with 2 degrees of freedom

2.4 Formalized subsampling

***Idea*:** Autocorrelated residuals indicate that the data resolution exceeds model resolution. In other words, data show sub-scale effects that are not reproducible with the given model structure. Therefore, the resolution of data points is reduced by drawing subsamples until LS assumptions are fulfilled. In this approach, the penalty for systematic patterns is the reduction of independent observations (compare equation 6).

In the field of environmental modeling, Dochain and Vanrolleghem, 2001) already proposed subsampling to overcome the problem of autocorrelation in residuals. However, a formal method was not presented which left the following questions unanswered: (1) How is the subsample formally chosen? (2) How to use the information contained in all data points? In the following, we introduce a formal method that responds to these questions.

***Realization*:** Figure 1 outlines the procedure of determining the optimal subsample size $n_{sub,pot}$. In order to avoid non-identifiability, we draw a subsample from the data at equidistant time steps of the original time vector **t**.[1] This equidistance is only feasible, if original data points are regarded as a circular sequence (t_1, t_2, t_3 ... t_n, t_1, t_2 ...) according to Politis et al., 1999. The optimal subsample size $n_{sub,opt}$ is found by a stepwise reduction of n_{sub} until the resulting

[1] Note, that this is a restriction purposely taken. Correctly speaking it depends on the model structure and measured dynamics whether "equidistance" should relate to **t** and / or / and not to **Y**.

residuals from least-squares regression with n_{sub} data points ($\mathbf{R}_{sub,LS}$) fulfill the tests for independence and normality. Ideally, these tests should be met at all possible "placings" (subsample sets $num_{sub.sets,\ ideal} = n / n_{sub}$) of the subsample of size n_{sub} or at a reasonable rate, e.g. 95% of $num_{sub.sets,\ ideal}$. If model complexity, amount of original data and/or computer resources do not permit this ideal number of sets, random placing has to be performed with $num_{sub.sets} < num_{sub.sets,\ ideal}$ that ensures a representative outcome statistic. The algorithm of Figure 1 not only leads to optimal subsample size $n_{sub,pot}$. It is already a basic formalized subsampling approach.

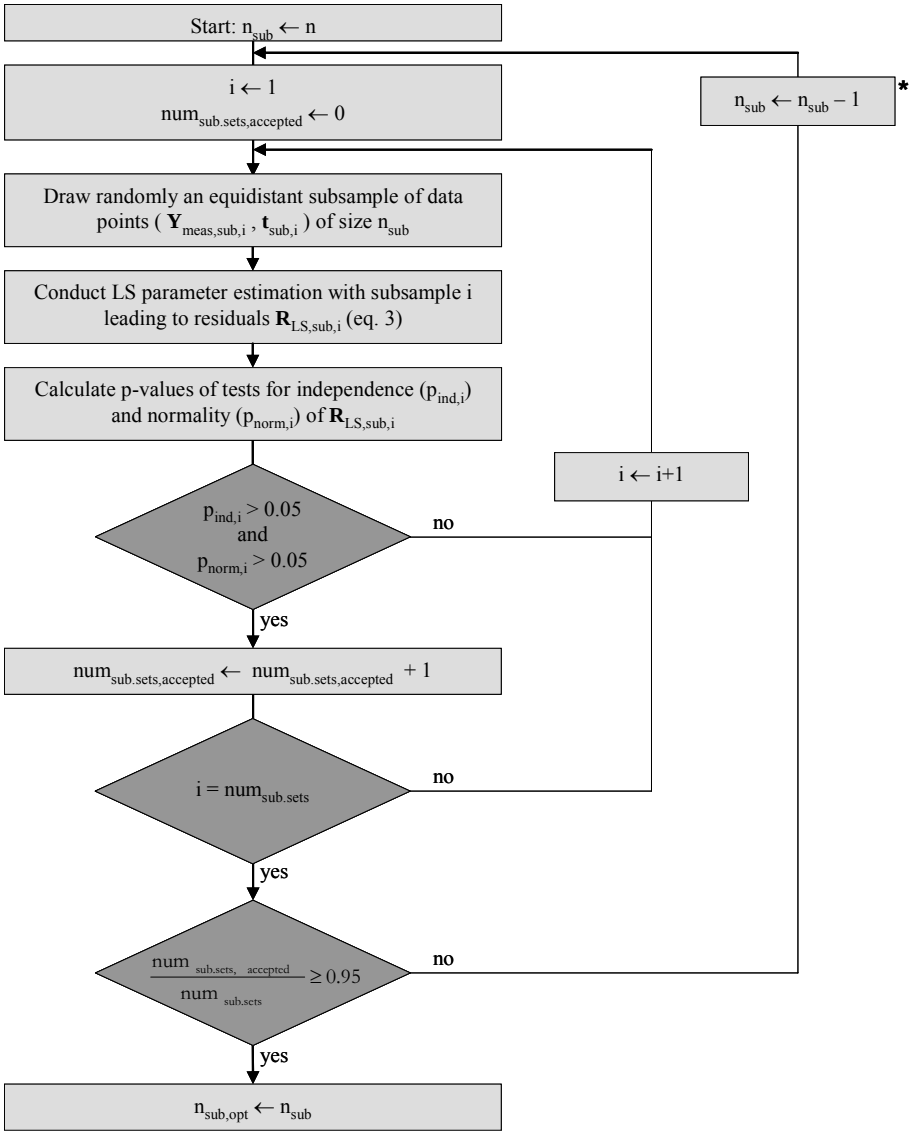

Figure 1. Formalized Subsampling approach: Determination of optimal subsample size $n_{sub,opt}$. (*) For simplicity, a stepwise reduction of n_{sub} is indicated which may be replaced by a more sophisticated search algorithms.

In addition to the formalized approach of Figure 1, we suggest to use further information contained in the data in an enhanced subsampling approach (see Figure 2): Repetitively, $n_{sub,opt}$ virtual data points are created that are then used for LS parameter estimation. These virtual observations are formed by calculated model outputs at subsampled time steps $\mathbf{t}_{sub}$ and an additional error term. This error is characterized by a bootstrap of the neighboring residuals. Neighbors are defined as the data points within the block $b = \pm n / 2n_{sub,opt}$ surrounding each subsampled time step of $\mathbf{t}_{sub}$ in the original data set. Consequently, the final number of

repetitions of parameter estimation with $n_{sub,opt}$ data points is f × g (f subsampling sets and g bootstrap runs). The number has to be in a range that the resulting statistic is representative. Since this is highly case specific, e.g. dependent on model nonlinearities, amount and range of data, no general statement can be made and f and g might have to be estimated iteratively.

The results of the formalized subsampling approach are statistics that have to be interpreted (see also Figure 5). The expected value of the distribution of best estimates be regarded as the overall best estimate of the parameter. The width of this distribution characterizes the bias for a situation where the parameters would be estimated from a single subsample only. The distribution of standard errors contains information about the expected standard error for the best estimate and again about 2nd order uncertainty (= distribution width) when applying only one subsample.

The method proposed is consistent. If LS residuals already fulfill iid assumptions with $n_{sub,opt} = n$, distributions of best estimate and standard error show zero variance, i.e. best estimates and uncertainties of the parameters correspond to an ordinary LS regression.

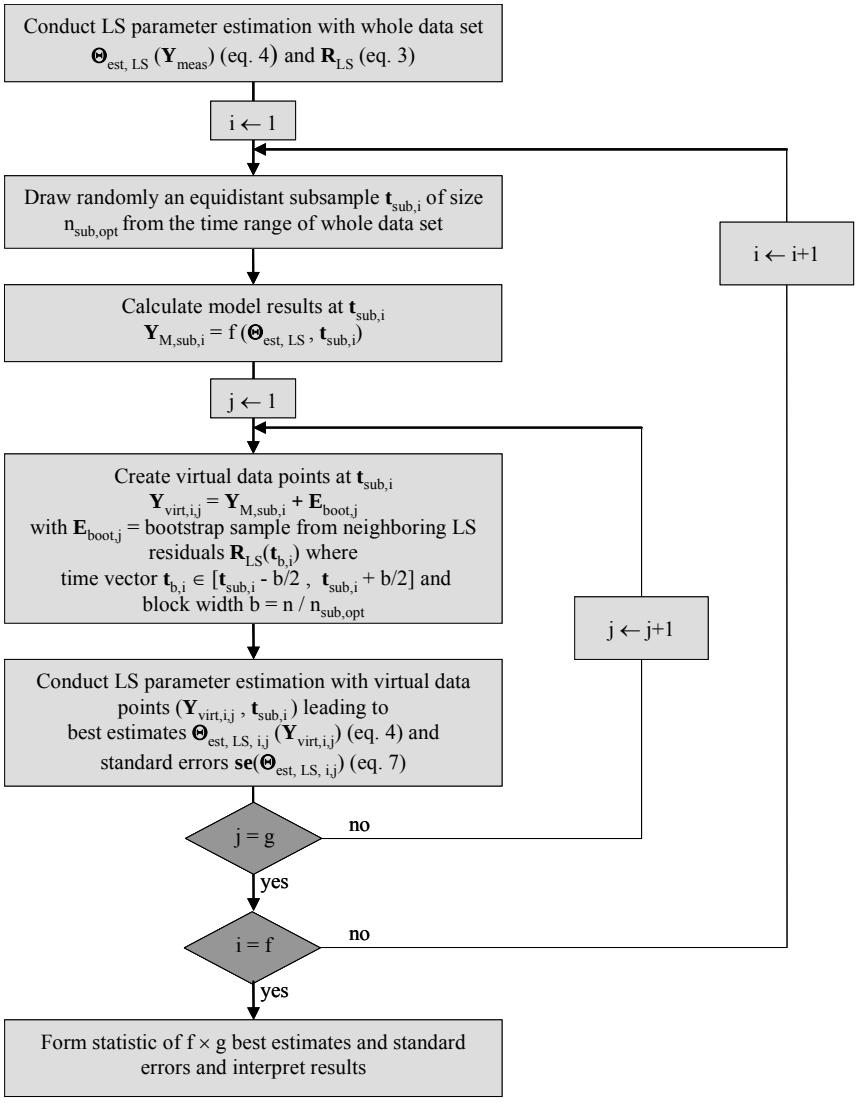

Figure 2. Formalized Subsampling approach: enhanced subsampling by using information of LS residuals. Estimation of $n_{sub,opt}$, f and g according to Figure1 and section 2.4 respectively.

2.5 Mirroring systematic patterns in residuals onto parameter of interest

***Idea*:** At some point of data analysis we have no reasonable basis anymore to ascribe the systematic patterns in residuals entirely either to a device characteristic (when the model should be extended accordingly) or to system inherent processes that should be represented by the lumped parameter of interest. Limited resources, knowledge or other constraints impede sophisticated model structure extension (compare section 1). Since our goal is to avoid underestimation of uncertainty, we have to choose the worst case at "this point": ascribe (mirror) systematic patterns in the residuals to the parameter of interest by replacing it with a high-order polynomial function of the variables Y or t. Whether the independent or dependent variable is chosen does not matter for the above reasons. However, one of them might facilitate a "quasi mechanistic interpretation" which is favorable.

***Realization*:** First a LS regression is performed. Then the parameter of interest is replaced by a polynomial function of t (or Y respectively).

$$\theta_i = \sum_{n=0}^{m} a_n t^n \qquad \text{(Eq.13)}$$

A new fit is executed for all $a_{0..m}$ only, i.e. all other parameters are kept fixed at their formerly LS estimated values, meaning that all systematic patterns a solely assigned to the parameter of interest. The fit is repeated with increasing orders m of the polynomial until iid. residuals are reached. In order to compute marginal distributions and best estimates for θ_i, standard errors of $a_{0..m}$ have to be taken into account in a multivariate manner, i.e. by considering the estimated correlation matrix of $a_{0..m}$ (compare section 4.5).

2.6 Generalized least-squares

All previously introduced methods are empirical approaches developed by the authors. In contrast, the generalized least-squares approach is a well established statistical method (Fox, 1997).

***Idea*:** The method allows for autocorrelation in observation errors (off-diagonal elements in matrix **V,** eq. 5). From this error matrix **V** a transformation matrix **Γ** is calculated and then used to transform data in such manner that ordinary LS can be applied.

***Realization*:** Libraries for many programming languages (JAVA, C++, S , R, ... Venables and Ripley, 2002; R Development Core Team, 2005) offer generalized least-squares packages for both linear and non-linear models (*gls* and *gnls* respectively). For filling the measurement error matrix **V**, at least a correlation structure must be provided, e.g. an autoregressive/moving-average model (ARMA) or a symmetric matrix (for details see Jenkins et al., 1994; Venables and Ripley, 2002). The according parameters of the correlation structure can either be estimated from *gnls* directly or beforehand, e.g. by fitting an appropriate function to the autocorrelation function (ACF) of $\mathbf{R}_{LS}$ - so called empirical generalized least-squares.

3. CASE STUDY: ESTIMATING THE OXYGEN HALF-SATURATION CONSTANT FOR NH_4 - OXIDISING BACTERIA

The oxygen half-saturation constant of ammonium oxidizing bacteria ($K_{O,AOB}$) is a decisive parameter for aeration control in wastewater treatment. It is typically determined by respirometric batch experiments. Respirometry is a state state-of-the-art procedure to estimate parameters for activated sludge modeling (Henze et al., 2000; Sin et al., 2005). The case study should therefore be a representative example to demonstrate the application of the introduced methods. In the following, a brief summary is given of the experimental setup and the model (for further details see Daebel et al., 2006).

3.1 Experiment

Sludge is withdrawn from a conventional denitrifying activated sludge plant, filled into a batch reactor and aerated until all readily degradable substances are consumed ($MLSS_{final}$ = 5048 $g_{COD}m^{-3}$). Temperature and pH are controlled at 20 ± 0.1°C and pH = 7.2 ± 0.1. Unlimited substrate conditions for autotrophic bacteria are ensured. After full activity of the bacteria has been reached, aeration is stopped. The decreasing oxygen concentration is measured by a WTW oxygen probe (Oxi330) with a temporal resolution of 1.5 sec. See the upper row of Figure 3 for a visualization of the measured data together with the calculated values of the fitted models A and B.

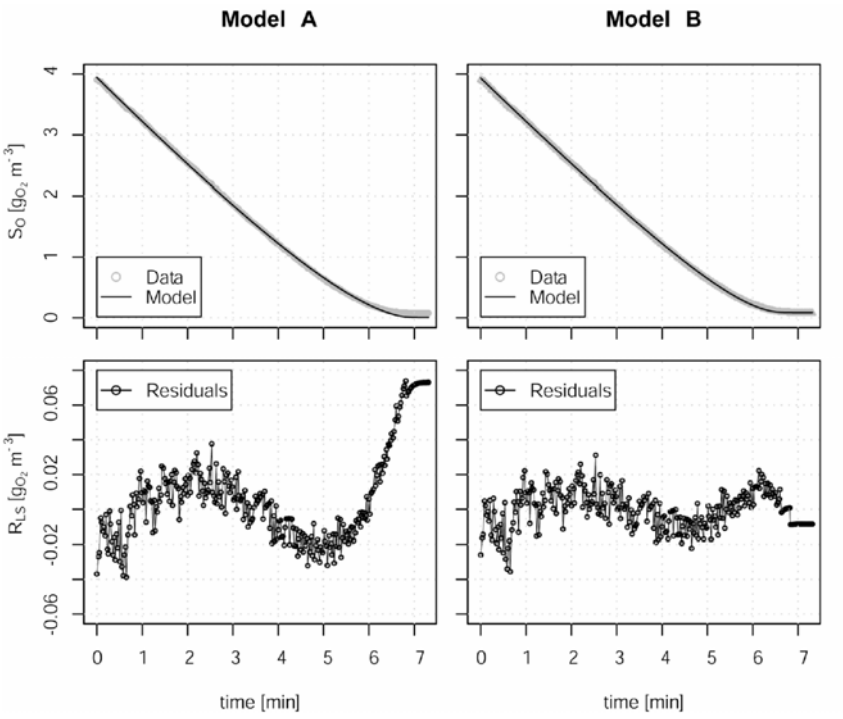

Figure 3. Experimental data from batch experiment (temporal resolution = 1.5 sec), fitted models and resulting residuals

3.2 Model

Model A

Experimental conditions are such that oxygen should be the only limiting factor for the activity of nitrifying bacteria. Accordingly, a Monod-type model structure is applied (compare eq. 14): decreasing oxygen concentration as a result of the activity of ammonium oxidizing bacteria (AOB), nitrite oxidizing bacteria (NOB) and heterotrophic bacteria (HET) each characterized by a maximum potential oxygen uptake rate (r_{max}) and a half-saturation constant (K_O). Short experimental time and initial aeration allow for the assumption of constant maximum oxygen uptake rates r_{max}. Surface aeration is considered as a function of an aeration coefficient (K_La) and oxygen saturation deficit ($S_{O,sat}$ - S_O).

$$\frac{dS_O}{dt} = -r_{max,AOB}\frac{S_O}{K_{O,AOB}+S_O} - r_{max,NOB}\frac{S_O}{K_{O,NOB}+S_O} - r_{max,HET}\frac{S_O}{K_{O,HET}+S_O} + K_La\left(S_{O,sat}-S_O\right) \qquad \text{(Eq.14)}$$

All eight parameters were estimated by the authors. For identifiability reasons three additional experiments had to be performed after the experiment described above, employing ATU to inhibit AOB. For the sake of simplicity, we will skip their analysis and assume here all parameters to be known[2] except for $r_{max,AOB}$ and $K_{O,AOB}$. Our discussion could easily be applied by dropping this assumption. However, since we want to focus on illustrating the proposed methods it makes sense to limit our discussion.

Model B

The modeled oxygen concentration $S_{O,A}(t)$ does not fit the data perfectly (compare Figure 3). At the end of the experiment, where the measured concentration should tend to zero, an asymptotic behavior of the data towards a concentration $S_{O,end} > 0$ $g_{O2}m^{-3}$ is observable (offset). This clearly probe related effect should be added to the model structure. Model B represents a corresponding model extension (Eq. 15). An analysis of the device revealed that the offset at 0 $g_{O2}m^{-3}$ and a drift leading to underestimated concentrations towards oxygen saturation. Therefore a linear correction function $c_{probe}(t)$ is introduced.

$$S_{O,B}(t) = S_{O,A}(t) + c_{probe}(t) \qquad \text{(Eq.15)}$$

$$c_{probe}(t) = S_{O,probe,0} + \frac{\left(S_{O,probe,sat}-S_{O,theo,sat}\right)-S_{O,probe,0}}{S_{O,theo,sat}}\cdot S_{O,A}(t)$$

$S_{O,A}$	calculated oxygen concentration with model A (with $S_{O,sat} := S_{O,probe,sat}$)
$S_{O,B}$	calculated oxygen concentration of the combined model B (sludge processes + probe effects)
$S_{O,probe,0}$	probe offset at $S_{O,A} = 0$ $g_{O2}m^{-3}$ (estimated directly from experimental data shown in Figure 3)
$S_{O,probe,sat}$	independently estimated saturated oxygen concentration = 8.38 $g_{O2}m^{-3}$ (measured in water at $T_W = 20°C$ and $p_{atm} \approx 1013$ mbar)
$S_{O,theo,sat}$	theoretical saturated oxygen concentration = 9,07 $g_{O2}m^{-3}$ (in water at $T_W = 20°C$ and $p_{atm} \approx 1013$ mbar)

[2] $r_{max,NOB} = 231 g_{O2}m^{-3}d^{-1}$, $K_{O,NOB} = 0.40 g_{O2}m^{-3}$, $r_{max,HET} = 460 g_{O2}m^{-3}d^{-1}$, $K_{O,HET} = 0.02 g_{O2}m^{-3}$, $K_La = 12.7 d^{-1}$, $S_{O,sat} = 9.07 g_{O2}m^{-3}$

Further model extension
Model structure alteration by the correction function is a good example of a first appropriate step in dealing with autocorrelated residuals. A more sophisticated sensor model was developed that included an additional response-time model. However, our investigations showed that no clear distinction was possible whether a probe or sludge-related effect would be modeled by this alteration. The danger then is that system inherent processes are assigned to the probe that should rather be reflected by the lumped parameter $K_{O,AOB}$. Such a state is nearly always reached in environmental data analysis - we simply do not know better! Our task is then to reflect this state of imperfect model structure in our reported parameter uncertainty.

3.3 Numerical implementation
Depending on the type of analysis models A and B were implemented in *AQUASIM* (Reichert, 1994a) or in the ***R*** programming environment using *lsoda* from the *odesolve*-package for solving differential equations, *nls* for nonlinear least-squares parameter estimation and *runsTest* and *normalTest* from the *fBasics* package (R Development Core Team, 2005) for testing independence and normality. Further details are given in the following section directly linked to the results of each method.

4. APPLICATION AND DISCUSSION
The results of applying the introduced methods to our case study are summarized in Table 1 and Figure 7 at section 4.7. They are followed by a meta-level comparison of the approaches (see Table 2). Beforehand, we specifically comment on the important aspects of the application of each method to the case study.

4.1 Ordinary least-squares regression
The standard error is reduced from 0.05 $g_{O2}m^{-3}$ for model A to 0.01 $g_{O2}m^{-3}$ for model B. These relatively small standard errors compared to the difference between best estimates ($K_{O,AOB,A}$ = 0.71$g_{O2}m^{-3}$ and $K_{O,AOB,B}$ = 0.40$g_{O2}m^{-3}$) lead to non-overlapping confidence regions (compare Figure 7). The residuals in Figure 3 reveal that their range can be reduced to an acceptable degree: for Model B they are in the order of absolute errors stated by the probe manufacturer (0.01$g_{O2}m^{-3}$). But, it is also visible that the systematic pattern, though attenuated, is persistent. This means that reasonable model adaptation alone does not solve the problem of autocorrelated residuals. The remaining questions are: How influential are these patterns on the uncertainty of $K_{O,AOB}$? And can imperfect model structures be reflected by the parameter uncertainty? The LS extending methods should offer adequate answers.

4.2 Whitening
Corrected standard errors were estimated from the post-whitening method using 10'000 repetitions. Tests for independence and normality were performed by applying a runs test of monotony (Chatfield, 2004; Dochain and Vanrolleghem, 2001) and a Jarque-Bera test (Jarque and Bera, 1980) respectively at significance levels of 5%. In simulations with normally distributed noise only, the highest possible fraction of repetitions, at which both tests were fulfilled, was found to be 90%. For Model A this level was reached at $\sigma_{noise,A} = 5.5 \cdot sd(\mathbf{R}_{LS,A})$ and for Model B at $\sigma_{noise,B} = 5.0 \cdot sd(\mathbf{R}_{LS,B})$.

Post-whitening has several advantages over pre-whitening and other approaches. The main advantage is the low processing time as only one parameter estimation is performed. Additional programming is only minor and identifiability problems do not occur. The

enlarged sum of squares can directly be applied to the uncertainty estimate of other parameters (compare eq. 6).

4.3 Sum of squares correction for autocorrelation

Sum of squares correction for autocorrelation was executed at several data resolutions, starting with the original data set until only every 30th data point was used (see x-axis of Figure 4). Different possible data sets for each resolution were considered and the corresponding results aggregated to their median values. Figure 4 shows the course of these median values as a function of data resolution for both conventional standard errors se_{LS} from LS (dotted line) and sum of square corrected se_{SSC} (solid line). The results for Model A, the strong monotonic increase of se_{LS} with fewer data points confirms that the resolution of data exceeds the limits of the model and that the estimated parameter uncertainty will greatly be underestimated with ever increasing data resolution (effect of large n in calculating se). In contrast, the nearly horizontal se_{SSC} graph indicates independency of the a posteriori corrected standard error from data resolution, supporting a reliable result for our goal. Interestingly, the resolution at which both lines intersect corresponds well to the optimal subsample sizes $n_{sub,opt}$ of the formalized subsampling approach presented in section 4.4 (Model A: 290/25 = 11.6 ≈ 12 = $n_{sub,opt,A}$; Model B: 290/18 = 16.1 ≈ 15 = $n_{sub,opt,B}$). We choose to assign the corrected standard errors where this overlap between se_{LS} and se_{SSC} lines begins, i.e. $se_{SSC,A}$ = 0.27 $g_{O2}m^{-3}$ and $se_{SSC,B}$ = 0.06 $g_{O2}m^{-3}$.

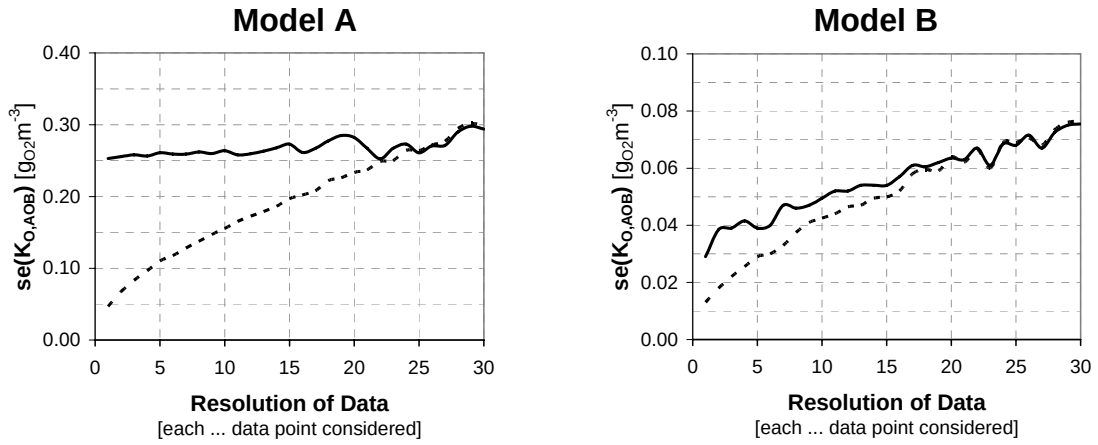

Figure 4. Estimated standard errors of $K_{O,AOB}$ as a function of data resolution. Dotted lines : estimated standard errors from conventional LS $se_{LS}(K_{O,AOB})$. Solid lines: sum of squares corrected standard error $se_{SSC}(K_{O,AOB})$.

However, Figure 4 (primarily for Model B) reveals a dependence of the corrected standard error se_{SSC} from data resolution. This indicates that the model structure is very good. To support this statement, we calculated a new sum of squares corrected se_{SSC} but this time analyzing normalized white noise $\mathbf{WN}_{norm}$ at different resolutions instead of normalized autocorrelated residuals $\mathbf{R}_{norm,LS}$. For both model A and model B, the calculated curves (not shown) correspond almost exactly to the se_{LS} curves - a very convincing result in terms of reliability of the method.

Yet, one drawback remains: the non-horizontal $se_{SSC,A/B}$ ($K_{O,AOB}$) lines introduce a degree of subjectivity as uncertainty levels have to be assigned from a certain point in the graph. Using the point where se_{LS} and se_{SSC} curves start to overlap appears reasonable and leads to results

that match the subsampling results. But whether this is reproducible for other cases remains an open question.

Since SS is also enlarged a posteriori, the method has computational advantages similar to those of the post-whitening approach. But programming is more demanding due to the mathematically intense character of the method (see eq. 12).

4.4 Formalized subsampling

Following the algorithm depicted in Figure 1, the optimal subsample sizes $n_{sub,opt}$ were found to be 12 for Model A and 15 for Model B of originally 290 data points. Independence of residuals was tested by applying a runs test for monotony and normality by the Jarque-Bera test, both at a significance level of 5%. With $n_{sub,opt}$ the procedure outlined in Figure 2 was run 8000 times (f = 200, g = 40). Figure 5 contains a basic statistical summary of the formalized approach (e.g. for Model A the best estimate for $K_{O,AOB}$ is $0.73 g_{O2} m^{-3}$ with an expected standard error of $0.27 g_{O2} m^{-3}$). The widths of the two distributions are an indicator for bias if only one arbitrary subsample were used. Note, that the distributions for best estimate and standard error may be highly correlated which needs to be considered for estimating confidence regions.

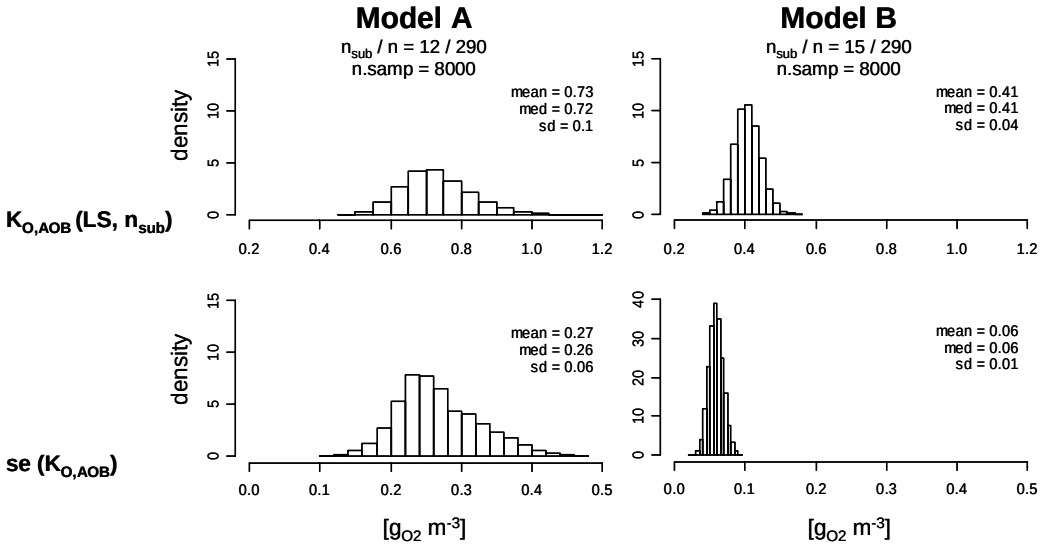

Figure 5. Results for $K_{O,AOB}$ from formalized subsampling approach based on 8000 repeating paramter estimations according to the procedure outlined in section 2.4 ($n_{sub,opt,A}$ = 12, $n_{sub,opt,B}$ = 15, f = 200, g = 40). **Upper row:** histograms of best estimates. **Lower row:** histograms of estimated standard errors.

The method is meaningful for two reasons. First, the amount of data is reduced to a degree where data and model resolution coincide. Second, the potential subjectivity (bias) in extracting such a subsample is avoided by random repetitions and a bootstrap approach which ensure the usage of the whole information contained in the data.

But, at the same time the application of the method requires considerable efforts: The model structure and required routines from Figure 2 must be implemented in a statistical software environment. High values of significant lags in autocorrelated residuals (leading to low $n_{sub,opt}$) and a high number of original data points demand numerous loops of calculations to obtain representative histograms as in Figure 5. Moreover, the reduction of data points to the subsample size $n_{sub,opt}$ might cause identifiability problems.

4.5 Mirroring systematic patterns in residuals onto parameter of interest

A polynomial function of 4th order was employed to assign systematic patterns of residuals $\mathbf{R}_{LS}$ to $K_{O,AOB}$ as a function of oxygen concentration S_O. Polynomial parameters a_n (n = 0, 1, .. , 4) were estimated with LS while all other parameters ($r_{max,AOB}$ and $S_{O,ini}$) were fixed at their formerly LS estimated values. Results are presented in Figure 6. Uncertainties of the estimated parameters a_n were taken into account by Monte Carlo simulations based on *Hammersley* sampling with n_{samp} = 100 (Hammersley, 1960; Handscomb and Hammersley, 1964) from the formerly estimated multivariate distribution of a_n's. The best estimate of $K_{O,AOB}$ stated in Table 1 was derived by calculating the median of the marginal distribution of all calculated Monte Carlo outputs (histograms in Figure 6). Employing the median for the best estimate seems most reasonable in the face of "outliers" at the limits of the range of S_O.

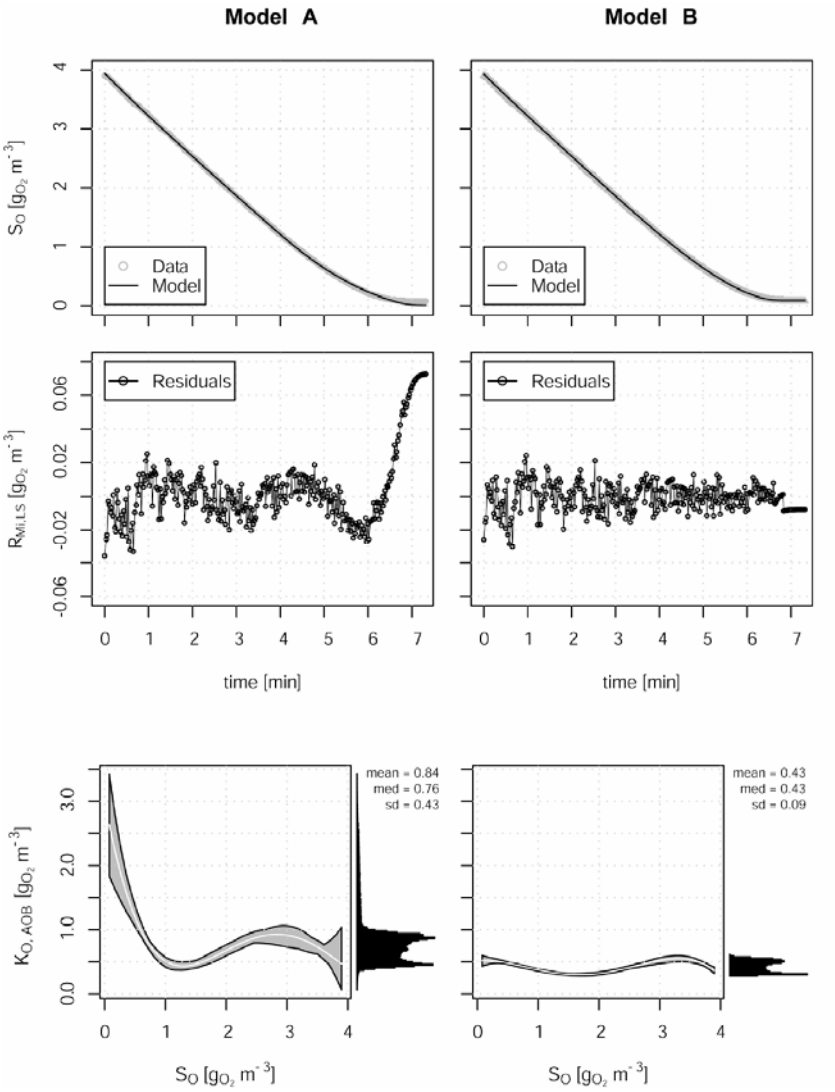

Figure 6. Results from the mirroring approach. Upper part: model results from least-square regression with $K_{O,AOB} = \Sigma\ a_n S_O^n$, n = 0, 1, 2, 3, 4 and resulting residuals. Lower diagrams: *white lines:* $K_{O,AOB} = \Sigma\ a_n S_O^n$ with best estimates of $a_{0..4}$; *grey areas:* envelope of Monte Carlo simulations of $K_{O,AOB} = \Sigma\ a_n S_O^n$ with 100 Hammersley samples from the estimated multivariate distribution of $a_{0..4}$; *black histograms:* marginal distribution of $K_{O,AOB}$ values.

The very pragmatic approach is closely related to recently proposed methods that introduce time-variable parameters to cope with model structure limitations (see e.g. Brun, 2002). Reflecting systematic patterns onto parameters, these approaches allow for either (i) detecting

mechanistic backgrounds or (ii) for the assumption of a truly stochastic behavior of a parameter which can then be modeled accordingly. For a conservative estimate, our simple approach seems to be sufficient, moreover as we do not have a sound basis to assume (i) or (ii) for our activated sludge parameter.

However, there are four major disadvantages of the approach: (a) The method has to be repeated in case of multiple parameters. (b) In cases of strong systematic patterns the method seems to be problematic: the resulting residuals $\mathbf{R}_{Mi,LS}$ for Model A in Figure 6 still show strong autocorrelation. Such results are therefore unreliable. (c) Data distribution might have a strong influence. (d) Finally, the interpretation of results in terms of a "standard error" is rather vague due to a missing underlying statistical model. An "interval" as uncertainty estimate seems to be more appropriate. We recommend this method therefore only for first checks on what influence autocorrelated residuals actually have on the parameter of interest.

4.6 Generalized least-squares

The application of *gnls* (*nlme*-package, R Development Core Team, 2005) failed for both models A and B. An identifiability analysis revealed that the reason for the unsuccessful application is related to the high correlation between $r_{max,AOB}$ and $K_{O,AOB}$ ($\approx$0.95). Certainly this high correlation is a major cause for uncertainty in $K_{O,AOB}$ at any method, but the non-linear least-squares function (nls) was still able to handle this situation in the previously presented extensions. High parameter correlation is often encountered in environmental models. The failure of the method in our case study is therefore very critical.

Testing *gnls* in a simple example worked well. We applied the function to virtual data created from a non-linear model and a third-order autoregressive error model, i.e. we knew the "truth". There, the method was able to produce uncertainty estimates of the non-linear model parameters that were reliable in contrast to the ones from ordinary least-squares.

Indeed, the generalized least-squares method works well in principal and inherits the most sophisticated statistical background compared to the other methods. It should therefore be investigated further.

4.7 Comparison of the approaches

Table 1 summarizes the results from applying the introduced methods. Additionally, Figure 7 visualizes them by assuming that uncertainty about $K_{O,AOB}$ can be approximated by a normal distribution. Results from the Mirroring approach were not plotted due to less interpretable results.

Table 1. Results of the methods applied for estimating the parameter uncertainty of $K_{O,AOB}$ given autocorrelated residuals

	Model A					
	ordinary least squares	whitening	sum of squares correction	formalized sub-sampling	mirroring	general. least squares
$K_{O,AOB}$ [1)]	0.71	0.71 [3)]	0.71 [3)]	0.73	0.76 [4),5)]	-
se [2)]	0.05	0.26	0.26	0.27	0.43 [5)]	-
	Model B					
	ordinary least squares	whitening	sum of squares correction	formalized sub-sampling	mirroring	general. least squares
$K_{O,AOB}$ [1)]	0.40	0.40 [3)]	0.40 [3)]	0.41	0.43 [4)]	-
se [2)]	0.01	0.06	0.06	0.06	0.09	-

[1)] best estimate in [$g_{O2}m^{-3}$]
[2)] estimated standard error in [$g_{O2}m^{-3}$]
[3)] method does not provide a best estimate; the result from ordinary least-squares is used
[4)] using median values
[5)] unreliable results, see discussion in section 4.5

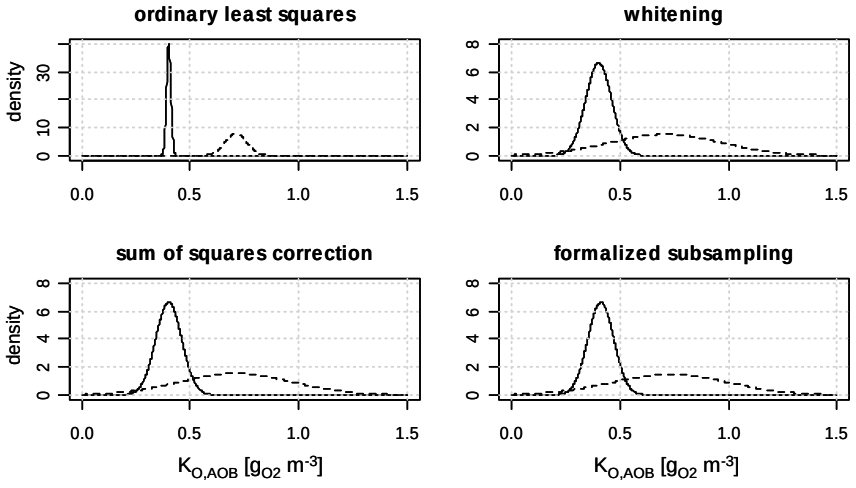

Figure 7. Visualization of results: Uncertainty of $K_{O,AOB}$ plotted as normal distributions (dashed line = estimation from Model A, solid line = estimation from Model B).

In contrast to uncertainty estimates from ordinary LS, the applied methods (except for the gnls approach) provide consistent results for three reasons: (1) $K_{O,AOB}$ values estimated from the more sophisticated model B are well represented by the uncertainty range estimated under model A. Best estimates B differ approximately one standard deviation sd_A from best estimate A. (2) The methods are also able to reflect the introduction of additional information which in our case was the offset characteristic of the oxygen probe. Introducing this knowledge led to a significant reduction of parameter uncertainty ($se(K_{O,AOB})_B \approx se(K_{O,AOB})_A / 5$). (3) All methods led to similar results.

Apart from checking compliance in our case study, a comparison of the methods on a meta level is an important evaluation step. To this aim, Table 2 summarizes the answers to the following specific questions: (1) Does the method require an additional programming environment or is common software, e.g. individual modeling software and spread sheets, enough? (2) All methods require programming, but to which degree? (3) How computationally demanding are they? (4) Can the results directly be used for all estimated parameters or does the procedure have to be repeated for every parameter? (5) How robust are the approaches towards highly correlated parameters? (6) What degree of reliability with respect to the results would the authors ascribe to each method? (7) How solid is the formal statistical basis of the approaches? Sections 4.2 to 4.6 contain the details to the scores listed in Table 2.

Table 2. Comparison of different approaches to conservatively estimate parameter uncertainty in case of model structure deficits. Scores range from "+++" (strong compliance with criteria statement) to "- - -" (no compliance).

	Methods	whitening		SS	formalized	mirroring	generalized
		pre	post	correction	subsampling		least squares
	Criteria						
1	Common software environment is applicable	+/-	+++	+/-	- -	+/-	- - -
2	Little additional programming	+/-	++	- -	- - -	+/-	+/-
3	Fast computation	- -	++	++	- - -	++	+++
4	No repetition for multiple parameters	++	+++	+++	++	- - -	+++
5	Robustness towards low identifiability of parameters	+/-	+++	+	+	++	- -
6	Expected reliability of method	++	++	+/-	++	- -	+++
7	Degree of solid statistical basis	+	+	- -	+/-	- -	+++

5. CONCLUSIONS

In this paper, we focused on the estimation of parameter uncertainty in the face of autocorrelated residuals. It is a situation that is increasingly encountered as modern measuring devices record data with ever higher resolution and accuracy while most of our models, mostly empirically grown, cannot reproduce the newly visible effects. As a consequence parameter uncertainties are underestimated by standard regression methods.

We showed that qualified model structure extensions (here: by a sensor error model), are a first appropriate step to tackle systematic patterns in residuals. In order to consider remaining model structure deficiencies, five least-squares extending methods were introduced to ensure a conservative estimate under the given circumstances: (i) pre- and post-whitening, (ii) sum of squares correction for autocorrelation (iii) formalized subsampling, (iv) mirroring and (v) generalized least-squares. Methods (i) to (iv) were developed for time-dependent models but could easily be extended to any other independent variable. Applying the approaches to a representative engineering case study, a respirometric experiment to estimate a Monod-type saturation parameter, reveals that they seem to produce reliable results as their outcomes nearly correspond. Post-whitening for its limited demand of resources and generalized least-squares for its sophisticated statistical background are the most recommendable methods.

According to the risk-averse attitude in engineering, the approaches serve the need for a conservative estimate of parameter uncertainty given model structure deficits. This uncertainty background must be acknowledged and quantified, if the conclusions drawn from uncertainty analysis are applied in practice (sensitivity rankings, decision making under uncertainty as optimal experimental design or risk-based design of engineering structures).

However, it has to be stated that four of the five presented methods are pragmatic approaches of scientific engineers seeking reliable uncertainty estimates under the constraint of imperfect model structures. Experts in statistics could not provide us with an alternative solution. Thus, none of these methods is built upon a sound formal statistical basis which makes a clear interpretation and further usage, e.g. for error propagation, difficult. This article is, therefore, a call for a scientific debate. It is a fact that statistical tools for inference are commonly applied nowadays (key word “automated calibration”), interpretations drawn, but underlying assumptions mostly violated and the consequences left untouched.

Finally, scientists have to be careful in stating terms like “robust”, “reliable” or “conservative” uncertainty estimate without further reference. Otherwise, they suggest a degree of objectivity which does not exist. Beside the often neglected violation of statistical assumptions discussed in this paper, many subjective decisions are taken during parameter estimation consciously or not: (i) objective functions are defined (see Dochain and Vanrolleghem, 2001; Yapo et al., 1998), (ii) search algorithms are chosen for finding the optimum of the objective function (Dochain and Vanrolleghem, 2001; Reichert, 1994b) - especially in case of weakly identifiable parameters the type of algorithm may influence results strongly - and (iii) numerical approximations for the solution of differential equations and covariance estimations are selected (Checchi and Marsili-Libelli, 2005; Marsili-Libelli et al., 2003). Last but not least, the transfer of via experiment estimated parameters to full scale model applications is often the primary source of uncertainty - easily outranging many others. Thus, there is a need for holistic studies, including all uncertainty introducing factors.

ACKNOWLEDGEMENTS

We thank Michael Wolf, Werner Stahel, Alessandro Carosio, Dante Salvani, Jack W. Baker, Jochen Köhler and especially Peter Reichert for stimulating discussions.

REFERENCES

Bates, D. M. and Watts, D. G. (1988). "Nonlinear regression analysis and its applications". Wiley, New York etc. 0-471-81643-4

Beck, M. B. (1987) Water-Quality Modeling - a Review of the Analysis of Uncertainty. *Water Resour. Res.* **23**, 1393-1442.

Berar, J. F. and Lelann, P. (1991) Esds and Estimated Probable-Error Obtained in Rietveld Refinements with Local Correlations. *J. Appl. Crystallogr.* **24**, 1-5.

Brun, R. (2002). "Learning from data parameter identification in the context of large environmental simulation models", PhD thesis. Zürich.

Chatfield, C. (2004). "The analysis of time series an introduction". Chapman & Hall, Boca Raton. 1-58488-317-0

Checchi, N. and Marsili-Libelli, S. (2005) Reliability of parameter estimation in respirometric models. *Water Res.* **39**, 3686-3696.

Daebel, H., Manser, R. and Gujer, W. (2006) Exploring temporal variations of oxygen saturation constants of nitrifying bacteria by using residual analysis. *Submitted to Water Research*.

Dochain, Denis and Vanrolleghem, Peter. (2001). "Dynamical modelling and estimation in wastewater treatment processes". Iwa, London. 1-900222-50-7

Draper, N. R. and Smith, H. (1998). "Applied regression analysis". Wiley-Interscience Publication, New York. 0-471-17082-8 (acid-free paper)

Efron, B. and Tibshirani, R. (1993). "An introduction to the bootstrap". Chapman & Hall, New York [etc.]. 0-412-04231-2

Fox, J. (1997). "Applied regression analysis, linear models, and related methods". Sage Publications, Thousand Oaks etc. 0-8039-4540-X alk. paper

Goodman, D. (2002) Extrapolation in risk assessment: Improving the quantification of uncertainty, and improving information to reduce the uncertainty. *Hum. Ecol. Risk Assess.* **8**, 177-192.

Hammersley, J. M. (1960) Monte-Carlo Methods for Solving Multivariable Problems. *Ann.NY Acad.Sci.* **86**, 844-874.

Handscomb, D. C. and Hammersley, J. M. (1964). "Monte Carlo methods". Methuen etc., London.

Hastie, T. (2003). “Elements of statistical learning data mining, inference, and prediction”. Springer, New York. 0-387-95284-5

Henze, M., Gujer, W., Mino, T. and Van Loosdrecht, M. (2000). “Activated sludge models ASM1, ASM2, ASM2d and ASM3”. IWA, London. 1-900222-24-8

Jarque, C. M. and Bera, A. K. (1980) Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. *Econ. Lett.* **6**, 255-259.

Jenkins, G. M., Reinsel, G. C. and Box, G. E. P. (1994). “Time series analysis forecasting and control”. Prentice Hall, Englewood Cliffs. 0-13-060774-6

Liu, C. X. and Zachara, J. M. (2001) Uncertainties of monod kinetic parameters nonlinearly estimated from batch experiments. *Environmental Science & Technology* **35**, 133-141.

Marsili-Libelli, S., Guerrizio, S. and Checchi, N. (2003) Confidence regions of estimated parameters for ecological systems. *Ecol. Model.* **165**, 127-146.

Politis, D. N., Romano, J. P. and Wolf, M. (1999). “Subsampling”. Springer, New York. 0-387-98854-8

R Development Core Team. (2005). “R: A language and environment for statistical computing”. R Foundation for Statistical Computing, Vienna, Austria. 3-900051-07-0

Reckhow, K. H. (1994) Water-Quality Simulation Modeling and Uncertainty Analysis for Risk Assessment and Decision-Making. *Ecol. Model.* **72**, 1-20.

Reichert, P. (1994a) Aquasim - a Tool for Simulation and Data-Analysis of Aquatic Systems. *Water Sci. Technol.* **30**, 21-30.

Reichert, P. and Vanrolleghem, P. (2001) Identifiability and uncertainty analysis of the River Water Quality Model No. 1 (RWQM1). *Water Sci. Technol.* **43**, 329-338.

Reichert, P. (1994b). “Concepts underlying a computer program for the identification and simulation of aquatic systems”. Swiss Federal Institute for Environmental Science and Technology (EAWAG), Dübendorf. 3-906484-08-4

Reichert, P. and Borsuk, M. E. (2005) Does high forecast uncertainty preclude effective decision support? *Environmental Modelling & Software* **20**, 991-1001.

Rieckermann, J., Borsuk, M., Reichert, P. and Gujer, W. (2005) A novel tracer method for estimating sewer exfiltration. *Water Resour. Res.* **41**.

Rousseeuw, P. J. (1984) Least Median of Squares Regression. *Journal of the American Statistical Association* **79**, 871-880.

Rousseeuw, P. J. and Leroy, A. M. (2003). “Robust regression and outlier detection”. Wiley, Hoboken, N.J. 0-471-48855-0

Sin, G., Van Hulle, S. W.H., De Pauw, D. J.W., van Griensven, A. and Vanrolleghem, P. A. (2005) A critical comparison of systematic calibration protocols for activated sludge models: A SWOT analysis. *Water Res.* **39**, 2459-2474.

Stahel, W. A. (2000). "Statistische Datenanalyse eine Einführung für Naturwissenschaftler". Vieweg, Braunschweig. 3-528-26653-8

Vanrolleghem, P. A. and Keesman, K. J. (1996) Identification of biodegradation models under model and data uncertainty. *Water Sci. Technol.* **33**, 91-105.

Venables, W. N. and Ripley, B. D. (2002). "Modern applied statistics with S". Springer, New York. 0-387-95457-0 (alk. paper)

Yapo, P. O., Gupta, H. V. and Sorooshian, S. (1998) Multi-objective global optimization for hydrologic models. *J. Hydrol.* **204**, 83-97.

General Conclusions and Outlook

General Conclusions and Outlook

GENERAL CONCLUSIONS

Framing of uncertainty analysis

One of the most critical tasks is recognised to be the framing of the uncertainty analysis. The challenge lies in defining a clearly set goal for the modelling study which determines how to set the boundary conditions, which processes to consider and which data to acquire. Morgan *et al.* (1990) suggest the clarity test of Howard and Matheson (1984) to remove uncertainty due to vagueness: "Imagine a clairvoyant who could know all the facts about the universe, past, present and future. Given the description of the event or quantity could she give the exact numerical value of the quantity?"

Classifying and separating sources of uncertainty

A second critical task is the separation of different types of uncertainty. In chapter 1 a design surface is proposed which distinguishes between the effects induced from rainfall variability and effects due to parameter uncertainty of the catchment model. In chapter 2, temporal variability is accounted for by characterising the pseudo first order rate constant of ozone decay as a time varying parameter and the associated uncertainty is described with confidence bands. In chapter 3 the ozone exposure and lethal dose are expressed as random variables characterising variability across organisms whereas uncertainty is expressed through probability distributions for the parameters characterising these random variables. It is not feasible to differentiate between more than two levels of uncertainty/variability as the representation of results becomes burdensome. Scenario analysis is recommended to test alternative model structures, higher order uncertainty (uncertainty about parameter ranges) or the effect of differing operational settings (chapter 3).

Quantifying uncertainty

When performing uncertainty analysis for full scale systems the notion of obtaining a "reliable" uncertainty estimate is problematic. Without being aware of the fact, engineers typically follow a Bayesian approach relying on information from previous analysis or from expert knowledge (chapter 3). Probabilistic representations of parameter uncertainty express degrees of belief. It is of high importance to recognise that the results from uncertainty analysis e.g. empirical density distributions are not objective statements of the variable under consideration but dependent on the framing, the quality of the available information and the experience of the analyst.

When applying pure frequentist approaches we find that the assumptions, upon which the statistical procedures rely, are not fulfilled (chapter 2, chapter 5). In a pure frequentist approach parameter uncertainty is the mapping of random errors through a "correct" model onto the parameters. However as neither sensor data (typically not just random error) nor models (due to their semi-empiric nature) are ideal, studies conducted within a pure frequentist setting typically underestimate uncertainty due to correlated residuals. This is especially critical if the uncertainty information (standard errors and correlation structure) is used for uncertainty propagation. For risk analysis within these contexts statistical procedures must be further developed to acknowledge this fact. Possible approaches such as whitening, sub-sampling or model structure extension are developed in chapter 5.

Uncertainty propagation and quantile analysis
As the models of the case studies are not computationally expensive and the probabilities of non-compliance are not very low (typically $p > 0.001$) uncertainty propagation with Monte Carlo simulation was found to be adequate. Some of the studied systems are specifically modelled in a way that reduces simulation time: e.g. steady state descriptions are chosen when the time scales of temporal input variability are larger than 2-3 times the residence time.

Regional Sensitivity Analysis
Within the context of environmental and stochastic models regional sensitivity analysis is a useful approach. Variance decomposition which is applied in chapter 2 and 3 is a powerful tool as it is "model-free" and no assumptions on linearity, monotonicity or additivity need to be made.

Information transfer of pseudo-parameters across temporal and spatial scales
The transfer of information across life stages of an engineering project or across reactor systems is viewed as critical (chapter 2). Due to the semi-empirical nature of most models in environmental technology the parameters are of aggregate nature (such as the ozone decay constant in chapter 2, the dispersion coefficient in chapter 4 and the half saturation constant in chapter 5) . Estimated parameters convey site-specific and event dependent information. However when a large number of systems has been analyzed distributions specifying variability of best estimates can be utilized for the analysis of comparable systems (as is recommended in chapter 4).

Value of uncertainty analysis
It is demonstrated that uncertainty analysis is especially useful for model diagnostics. It forces the analyst to define the modelling problem clearly and allows obtaining an estimate of accuracy of results from deterministic models. Uncertainty and sensitivity analysis forces one to apply models with more care and generally increases robustness and transparency of model results. Combining uncertainty analysis with regional sensitivity analysis enables informed decisions for managing (e.g. reducing) uncertainty. There is however also a risk of overrating the goodness of results obtained from uncertainty analysis. For instance probabilistic results from environmental engineering systems given with a high level of precision (such as $p = 0.092$) must be critically viewed.

There remain some challenging implications for practice: Traditional design approaches rely on simplified descriptions taking uncertainty into account implicitly whereas the approaches used to comply with more sophisticated criteria require information on frequency and probability of meeting target values. The regulator faces important questions: At which level of sophistication does compliance need to be proven and how is the quality of results from modelling studies judged? Only through experience on a wider range of systems and in different contexts can methods for best practice for probabilistic analysis and design be developed.

The use of easily accessible statistical tools should be accompanied by the necessary statistical rigour as not to create delusion and to cause the opposite effect of the intended increase of transparency. This can be enhanced by stating required meta-information such as source and quality of data or the statistical methodology applied.

OUTLOOK

Uncertainty and sensitivity analysis in environmental engineering

Concerning regional sensitivity analysis a deeper exploration of numerically efficient algorithms is recommended such as FAST (Cukier *et al.*, 1977; Saltelli *et al.*, 1999). Recently developed entropy based methods (Liu *et al.*, 2006) might prove to be a valuable alternative.

Frameworks should be developed that allow for consistent integration of uncertain information stemming from different sources (such as aggregation across different experts (Ayyub, 2001) or coupling site specific information with information from literature). This may be reached by following a rigorous Bayesian approach.

Another important source of uncertainty not specifically analyzed in this study is the variability in results obtained by different analysts (Refsgaard *et al.*, 2006). Experiments where analysts are given the same problem to solve would allow the exploration of the tacit processes involved in model building.

Optimal use of increasing computational resources should be tested by comparison of results from probabilistic or stochastic descriptions with results from mechanistic based approaches such as Computational Fluid Dynamics (CFD) (Greene *et al.*, 2004). The modelling of disinfection is viewed as an ideal application for conducting such model comparisons.

Optimal decisions for design of public infrastructure

How sensitive are investment decisions on the choice of considering or not considering uncertainty or on the sophistication of criteria applied (Reckhow, 1994)? This can only be answered by comparing the different approaches for a number of real-world case studies (Linkov *et al.*, 2005). The analysis of uncertainty should hereby be conducted within a decision theoretic framework such as multi-attribute-utility theory.

The valuation of environmental and health effects is itself subject to considerable uncertainty. Research for developing legislative criteria for systems from environmental engineering should be conducted in a benefit-cost-risk framework. Within this context a crucial aspect is the specification of engineering codes. To obtain optimal incentive structures tender procedures and the corresponding risk allocation need to be reconsidered (Flyvbjerg *et al.*, 2003). Environmental engineering scientists should take an active role in the formulation of optimal criteria together with ecologists and economists.

A long term goal would be the development of an uncertainty toolbox for practising engineers with appropriate methods for the different stages of engineering design. Especially the exploration of tools for dealing with uncertainty across larger time scales (Dominguez and Gujer, 2006) such as real options (Pindyck and Dixit, 1994) or scenario analysis (Schoemaker, 1995) are expected to add value to the work of civil and environmental engineers.

REFERENCES

Ayyub B.M. (2001) “Elicitation of Expert Opinions for Uncertainty and Risks”, CRC, Boca Raton.

Cukier R.I., Levine H.B. and Shuler K.E. (1977). Nonlinear Sensitivity Analysis of Multiparameter Model Systems. *Journal of Physical Chemistry,* **81**(25), 2365-2366.

Dominguez D. and Gujer W. (2006). Evolution of a Wastewater Treatment Plant Challenges Traditional Design Concepts. *Water Research,* **40**(7), 1389-1396.

Flyvbjerg B., Bruzelius N. and Rothengatter W. (2003) "Megaprojects and Risk an Anatomy of Ambition", Cambridge University Press, Cambridge.

Grayman W.M. and Males R.M. (2002). Risk-Based Modeling of Early Warning Systems for Pollution Accidents. *Water Science and Technology,* **46**(3), 41-49.

Greene D.J., Farouk B. and Haas C.N. (2004). Cfd Design Approach for Chlorine Disinfection Processes. *J. Am. Water Work Assoc.,* **96**(8), 138-150.

Howard R.A. and Matheson A.J. (1984) Readings in the Principles and Practice of Decision Analysis. *In* "Strategic Decision Systems", Menlo Park, Calif.

Jakeman A.J., Letcher R.A. and Norton J.P. (2006). Ten Iterative Steps in Development and Evaluation of Environmental Models. *Environ. Modell. Softw.,* **21**(5), 602-614.

Linkov I., Ames M.R., Crouch E.A.C., and Satterstrom F.K. (2005). Uncertainty in Octanol-Water Partition Coefficient: Implications for Risk Assessment and Remedial Costs. *Environ. Sci. Technol.,* **39**(18), 6917-6922.

Liu H.B., Chen W. and Sudjianto A. (2006). Relative Entropy Based Method for Probabilistic Sensitivity Analysis in Engineering Design. *Journal of Mechanical Design,* **129**(2), 326-336.

Morgan M.G., Henrion M. and Small M. (1990) "Uncertainty a Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis", Cambridge University Press, Cambridge etc.

Pindyck R.S. and Dixit A.K. (1994) "Investment under Uncertainty", Princeton University Press, Princeton, New Jersey.

Reckhow K.H. (1994). Importance of Scientific Uncertainty in Decision-Making. *Environ. Manage.,* **18**(2), 161-166.

Refsgaard J.C., van der Sluijs J.P., Brown J., and van der Keur P. (2006). A Framework for Dealing with Uncertainty Due to Model Structure Error. *Advances in Water Resources,* **29**(11), 1586-1597.

Saltelli A., Tarantola S. and Chan K.P.S. (1999). A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics,* **41**(1), 39-56.

Schoemaker P.J.H. (1995). Scenario Planning - a Tool for Strategic Thinking. *Sloan Management Review,* **36**(2), 25-40.

Curriculum Vitae

2007 -	Postdoctoral Researcher at Eawag, Dübendorf
2002 - 2006	Research Assistant at Eawag, Dübendorf
	Dissertation at the Swiss Federal Institute of Science and Technology (ETH), Zurich
	Uncertainty Analysis for Performance Evaluation and Design of Urban Water Infrastructure.
1999 - 2001	Full-Time Teaching Assistant ETH Zurich (Chair of Urban Water Management)
1994 - 1999	Studies at the Department of Civil Engineering, Environmental Engineering and Geomatics, ETH Zurich, dipl. Ing. ETH
1989 - 1994	Grammar School, Typus C, Im Lee, Winterthur
March 4, 1974	Born in Bern, Switzerland