How can we currently include uncertainty and variability in model-based projects?

Marc B. Neumann[1], Peter A. Vanrolleghem[1], Lorenzo Benedetti[2], Stefan Weijers[3], Sudhir Murthy[4], Bruce R. Johnson[5], Evangelina Belia[6]

[1]model*EAU*, Département de génie civil et de génie des eaux. Université Laval, 1065, Avenue de la Médecine. Québec G1V 0A6, QC, Canada (E-mail: *marc.neumann@gci.ulaval.ca*; *peter.vanrolleghem@gci.ulaval.ca*)
[2]WATERWAYS srl, Via del Ferrone 88, 50023 Impruneta (FI), Italy (E-mail: *lorenzobenedetti@waterways.it*)
[3]Waterschap De Dommel, 5280 Boxtel, The Netherlands (E-mail:*sweijers@dommel.nl*)
[4]DCWater, DWT, 5000 Overlook Ave., SW Washington, DC 20032, U.S.A.
(E-mail: *Sudhir.Murthy@dcwater.com*)
[5]CH2M HILL, 9193 South Jamaica Street, Englewood, CO 80112, U.S.A. (E-mail: *Bruce.Johnson2@ch2m.com*)
[6] Primodal Inc., 145 Aberdeen, Quebec, G1R 2C9, QC, Canada (E-mail: *belia@primodal.com*)

Abstract
We propose methods that practitioners can currently include in modelling-based projects to account for uncertainty and variability.

Keywords
Design, modelling, operation, risk, uncertainty, variability

INTRODUCTION

Models are increasingly used in wastewater engineering for a multitude of purposes ranging from process optimisation to design. However, although models are widely used in engineering practice, it often remains unclear how to adequately account for the associated uncertainty and variability. *Variability* is hereby defined as the "real spread" of values (in time or space) of a well-specified population (such as historic daily COD load variability in the influent of a specific treatment plant). The spread of these values is not reducible. *Uncertainty*, on the other hand, results from a lack of knowledge. *Parameter uncertainty* is the uncertainty about the values of model parameters (e.g. half-saturation constants). *Model structure uncertainty* pertains to the adequacy of the model equations and the model resolution in view of the modelling objective. Unlike variability, uncertainty is partly reducible: e.g. through further measurements or deeper investigations into the relevant processes. Although the topic on how to adequately account and deal with variability and uncertainty is an active research topic in academia this contribution highlights which tools and methods are already available for adoption in engineering practice.

APPROACHES TO INCLUDE UNCERTAINTY AND VARIABILITY

Accounting for variability

Variability in WWTPs occurs in time and space. Whereas accounting for temporal variability is the central aspect of dynamic modelling, spatial variability has typically only been coarsely resolved using compartmental models such as tanks-in-series.

Tools to account for *temporal variability* include probability distributions, dynamic modelling, multivariate time series analysis, and influent generators. *Probability distributions* can be used to describe the frequency of dynamic variables such as flows or loads. This is useful when using a steady-state solution of the model: E.g. when describing average monthly behaviour the influent concentrations and flows can be sampled from the (joint) probability distributions to capture meaningful scenarios (Bixio *et al*. (2002), McCormick *et al*. (2007)). In addition, cumulative distributions are often used to characterise plant performance. They summarise the information contained in a time-series and allow extracting the frequency of exceedance of effluent concentration limits. *Dynamic solutions* capture how dynamic influents affect the state variables of the system and predict a dynamic effluent profile from which the desired statistics can be extracted. An advantage of using historic time series and dynamic models is that temporal dependence (auto-correlation) is appropriately and explicitly accounted for. Multivariate time series models can be used to generate synthetic time series with the same characteristics as historic time series retaining the cross-correlation between variables. If synthetic time series are required that represent future load scenarios influent generators need to be used (Gernaey *et al*., 2010).

Concerning the description of *spatial variability* the rapidly growing CFD (computational fluid dynamics) applications allow investigating spatial phenomena at high resolution (e.g. Gresch *et al.*, 2011). Such analyses are critical for multiphase systems (settling) or systems that need to guarantee a certain contact time (disinfection). To decrease the computational burden methods have been developed that allow translating a CFD model to a compartment model (Gresch *et al*., 2009).

Accounting for uncertainty

Parameter uncertainty can be addressed by assigning probability distributions to parameters to reflect the knowledge of the engineer. In applications where no data is available, a priori uncertainty estimates are obtained from expert knowledge. The effects of parameter uncertainty on model outputs can be quantified by the use of Monte Carlo simulation techniques (Benedetti *et al*., 2008; Sin *et al*., 2009). In the case of an existing system, some parameters may be measured directly (with experiments) while others can be estimated by means of calibration.

For the practitioner, *model structure uncertainty* can be addressed in various ways. In contrast to a scientist who is developing a model from first principles to test a research hypothesis, the engineer will typically select a model from a model library. The engineer is restricted to the models implemented in commercial simulators. To address model structure uncertainty, the analyst may want to repeat the modelling exercise with a different model or integrate his/her own model structure extensions or reductions.

Uncertainty due to *implementation errors* or numerical errors can be captured through the use of multiple simulators, but this may not be possible in practice due to time restrictions. Numerical accuracy can be checked by changing the solver properties (such as time step size or solver type and accuracy).

Scenario uncertainty can be accounted for by applying scenario analysis techniques to account for multiple possible futures in terms of loads, requirements, etc. (Dominguez *et al*. 2009). If quantifiable, scenario uncertainty can also be accounted for in model applications by Monte Carlo simulation.

Typically, a *sensitivity analysis* is required to prioritize the sources of uncertainty. For instance, the uncertainty of one of the parameters may be very large; however, due to non-linearity the effect on the model variable might be very small. The same holds for the opposite case: a parameter with little uncertainty might cause large uncertainty in one of the design variables. Some of the recently proposed methods in scientific literature are available and easy to implement (e.g. Benedetti *et al.*, 2011; Sin *et al.*, 2011).

WHAT ARE WE LACKING?

Although there are many approaches available to explicitly account for uncertainty and variability, open questions remain which are addressed in current research, e.g.:

- How to move from guidelines with the safety factor approach to probabilistic model-based design?
- Determination of prior uncertainty ranges (e.g. in design).
- Parameter (uncertainty) estimation in systems with poor identifiability.
- How to adequately deal with model structure uncertainty?

CONCLUSION

For practitioners: there are many tools available to account for uncertainty and variability – their use should be encouraged. For scientists: there are still problems to solve, especially regarding the applicability of methods in practice (e.g. determination of prior uncertainty) and the exploration of promising methods (e.g. Bayesian estimation, Artificial Intelligence).

REFERENCES

Benedetti, L., Bixio, D., Claeys, F. and Vanrolleghem, P.A. (2008) Tools to support a model-based methodology for emission/immission and benefit/cost/risk analysis of wastewater systems that considers uncertainty. *Environmental Modelling & Software* **23**(8), 1082-1091.

Benedetti, L., Claeys, F., Nopens, I. and Vanrolleghem P.A. (2011) Assessing the convergence of LHS Monte Carlo simulations of wastewater treatment models. *Water Science and Technology* **63**(10), 2219-2224.

Bixio, D., Parmentier, G., Rousseau, D., Verdonck, F., Meirlaen, J., Vanrolleghem, P.A. and Thoeye, C. (2002) A quantitative risk analysis tool for design/simulation of wastewater treatment plants. *Water Science and Technology* **46**(4-5), 301-307.

Dominguez, D., Worch, H., Markard, J., Truffer, B. and Gujer, W. (2009) Closing the capability gap: Strategic planning for the infrastructure sector. *California Management Review* **51**(2), 30-50.

Gernaey K.V., Benedetti L., Corominas Ll., Langergraber G., Jeppsson U., Flores-Alsina X., Rosen C. and Vanrolleghem P.A. (2010) Wastewater treatment plant influent disturbance models. In: Proceedings 2nd IWA/WEF Wastewater Treatment Modelling Seminar (WWTmod2010). Mont-Sainte-Anne, Québec, Canada, March 28-30 2010. 283-287.

Gresch, M., Armbruster, M., Braun, D. and Gujer, W. (2011) Effects of aeration patterns on the flow field in wastewater aeration tanks. *Water Research* **45**(2), 810-818.

Gresch, M., Brugger, R., Meyer, A. and Gujer, W. (2009) Compartmental models for continuous flow reactors derived from CFD simulations. *Environmental Science & Technology* **43**(7), 2381-2387.

Mc Cormick, J. F., B. Johnson and A. Turner (2007) Analyzing risk in wastewater process design: using Monte Carlo simulation to move beyond conventional design methods. In: Proceedings 80th Annual Water Environment Federation Technical Exhibition and Conference, San Diego, California, USA, October 13-17, 2007.

Sin, G., Gernaey, K.V., Neumann, M.B., van Loosdrecht, M.C.M. and Gujer, W. (2009) Uncertainty analysis in WWTP model applications: A critical discussion using an example from design. *Water Research* **43**(11), 2894-2906.

Sin, G., Gernaey, K.V., Neumann, M.B., van Loosdrecht, M.C.M. and Gujer, W. (2011) Global sensitivity analysis in wastewater treatment plant model applications: Prioritizing sources of uncertainty. *Wat. Res.* **45**(2), 639-651.