



Jul 1st, 12:00 AM

Influence of Uncertainty Analysis Methods and Subjective Choices on Prediction Uncertainty for a Respirometric Case

Katrijn Cierkens

Stijn Van Hoey

Bernard De Baets

Piet Seuntjens

Ingmar Nopens

Follow this and additional works at: <https://scholarsarchive.byu.edu/iemssconference>

Cierkens, Katrijn; Van Hoey, Stijn; De Baets, Bernard; Seuntjens, Piet; and Nopens, Ingmar, "Influence of Uncertainty Analysis Methods and Subjective Choices on Prediction Uncertainty for a Respirometric Case" (2012). *International Congress on Environmental Modelling and Software*. 147.
<https://scholarsarchive.byu.edu/iemssconference/2012/Stream-B/147>

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Influence of Uncertainty Analysis Methods and Subjective Choices on Prediction Uncertainty for a Respirometric Case

**Katrijn Cierkens^{1,2}, Stijn Van Hoey^{1,3}, Bernard De Baets², Piet Seuntjens^{3,4,5},
Ingmar Nopens¹**

¹*BIOMATH, Model-based biosystem analysis and optimisation, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, B-9000, Ghent, Belgium (Katrijn.Cierkens@UGent.be, Ingmar.Nopens@UGent.be, Stvhoey.VanHoey@UGent.be)*

²*KERMIT, Research Unit Knowledge-based Systems, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure links 653, B-9000, Ghent, Belgium (Bernard.DeBaets@UGent.be)*

³*Flemish Institute for Technological Research (VITO), Boeretang 200, B-2400 Mol, Belgium*

⁴*Department of Soil Management, Ghent University, Coupure links 653, 9000 Ghent, Belgium (Piet.Seuntjens@UGent.be)*

⁵*University of Antwerp, Faculty of Bioscience Engineering, Groenenborgerlaan 171, 2020 Antwerp, Belgium*

Abstract: Uncertainty analysis in integrated wastewater treatment modelling is still in its infancy, although techniques from different fields are increasingly used in research for design and control of wastewater treatment plants (WWTPs). However, results should be interpreted with care. This paper shows explicitly the influence of different methodologies and subjective choices on prediction uncertainty for a simple environmental modelling case: a respirometric experiment for acetate degradation without storage. This case uses experimental data and can be used to estimate kinetic parameters of activated sludge in full-scale wastewater treatment modelling. These are subsequently used in more complex activated sludge models (ASMs) for dynamic modelling studies, in order to reduce the number of parameters to calibrate in these overparameterised models. Three uncertainty analysis methodologies are compared: (1) classical parameter estimation (Fisher Information Matrix or FIM-based), which has been performed extensively for respirometric models. The derived parameter confidence intervals are then propagated to the output through Monte Carlo (MC) simulations with and without correlation based sampling; (2) MC simulations from expert-based probability density functions; (3) the generalised likelihood uncertainty estimation (GLUE) method, a well-known uncertainty analysis method applied in hydrological modelling, recently also applied on full-scale WWTPs. The output uncertainty boundaries on the model output are very specific to the method used and to subjective choices like probability density functions in the expert-based method and threshold values in the GLUE method. Besides, classical parameter estimation and GLUE are useful to observe and easily handle correlated parameters, which is very important when using these methods on full-scale models. Appropriate validation experiments are needed to judge the applicability of the different applied methods. Whereas this seems not possible at full-scale level, it should be feasible to perform validation respirometric experiments at lab-scale.

Keywords: GLUE, parameter estimation, bioprocess model

1. Introduction

1.1. Uncertainty in wastewater treatment modelling

According to Belia et al. (2009), uncertainty analysis methods can be divided into three groups: 1) methods used to characterize and prioritize uncertainty, 2) methods used to increase the quality of information, 3) methods used to quantify the sources of uncertainty, propagate these uncertainties to the model output and evaluate the resulting output uncertainty. In this paper, the term uncertainty analysis only refers to the last group.

Most studies on the propagation of uncertainty in WWTP are conducted for design or control of WWTPs (Benedetti et al., 2010; Benedetti et al., 2012; Bixio et al., 2002, Mannina, 2011). Indeed, an uncertainty analysis offers a quantitative basis to justify safety factors (used in design) and to give rise to better informed decision making in cost saving engineering projects (Sin et al., 2009). In WWTP control strategy evaluations, uncertainty can be used to assess the robustness of a certain control strategy against deviating biological, operation and design parameters, to assess the importance of uncertainty in multi-criteria analysis and to assess the probability of exceeding legal effluent standards (Benedetti et al., 2010)

Although uncertainty analysis revealed to be a powerful tool in previous mentioned applications, the need for an easily applicable and scope-specific comprehensive protocol that incorporates uncertainty identification and analysis in the modelling procedure was raised recently. Moreover, further research to model structure uncertainty and the influence of correlations between uncertain input parameters on uncertainty is recommended (Sin et al., 2009).

However, since the complexity of full-scale models, it is difficult to interpret or track the origin of model prediction uncertainty. Simple respirometric experiments are typically used in 'theoretical' parameter estimation studies or to estimate biological parameters of the activated sludge in order to reduce the number of parameters to calibrate in the overparameterised activated sludge models (ASM) used in full-scale modelling studies (Vanrolleghem et al. 1999, Gernaey et al. 2002). This widely known experiment and model can also be used to investigate output uncertainty because the model predicts the oxygen uptake rate of the bacteria, which can be indirectly measured from the dissolved oxygen profile in the reactor.

1.2. Uncertainty analysis methodologies

Three prediction uncertainty methodologies are compared: classical parameter estimation, Monte Carlo (MC) simulations from a multi-dimensional parameter space set by expert knowledge and the Generalised Likelihood Uncertainty Estimation method (GLUE).

Classical parameter estimation determines parameter confidence intervals (CIs) based on linear approximations of the parameter estimation covariance matrix, which can be approximated by inverting the Fisher Information Matrix (FIM). Classical parameter estimation is extensively used in 'theoretical' respirometric studies. Different methodologies for propagating parameter uncertainty to the model output (OUR_{ex}) are compared: (1) MC simulations from a parameter space defined by uniform distributions with a range equal to the 95% CIs without correlation-based sampling; (2) same as (1) but with correlation-based sampling; (Iman and Conover, 1982) and (3) linear error propagation (Omlin and Reichert, 2000).

The second method consists of performing MC simulations from expert-based probability density functions without taking into account parameter correlations (Benedetti et al., 2010) The method is increasingly being used in research for practical applications of uncertainty in design and control of WWTP. Most studies are performed by defining a probability density function (PDF) for each uncertain

parameter. Distribution type, mode and variability from the mean value are defined. A commonly used method to set boundaries is the division of the selected parameters into three groups according to their 'expected' uncertainty, each corresponding with a predefined variability around the default value, typically 5%, 25% and 50% and typically a uniform, triangular or truncated normal distribution is implied (Reichert and Vanrolleghem, 2001). However, the determined uncertainty when sampling from an expert-based parameter space is directly linked to the choice of these PDFs (clearly shown by Benedetti (2008))

GLUE is an uncertainty analysis method mainly applied in hydrological modelling (Beven, 2006) and recently also used in full-scale WWTP modelling (Mannina et al., 2011). In the GLUE method, a large number of random simulations from a prior multi-dimensional parameter space are run, after which the 'behavioural' simulations are selected based on a predefined likelihood measure (cf. Nash-Sutcliffe criterion (NS) and sum of squared errors (SSE)) and a user-defined threshold. 'Behavioural' parameter combinations are conditioned to the parameters and weighted proportional to their corresponding likelihood value. This results in a posterior cumulative distribution function, which allows deriving 95% predictive uncertainty on the model output. The GLUE method was proposed as a method to handle potential equifinality (non-identifiability or non-uniqueness) of parameter combinations during parameterization. It is related to Bayesian estimation techniques, but it doesn't depend on an explicit statistical error (residual) model. This error model is a key point in the Bayesian method (Beven, 2006), but sometimes difficult to specify as a result of the composite effect of multiple error sources (Beven et al., 2008). On the other hand, when applying the GLUE method, the prediction limits and the posterior distribution of the parameters will be sensitive to the choice of the likelihood function and the choice of the threshold value (Li et al., 2010; Montanari, 2005).

2. Materials and methods

2.1. Respirometric experiment

The flowing gas-static liquid respirometer consists of a reactor with a volume of 2L filled with sludge, taken from the aerobic tanks of the municipal WWTP of Ossemeersen (Gent, Belgium) which was aerated overnight to ensure endogenous state. Temperature is controlled at 20 °C (± 0.05) and pH at 7.5 (± 0.1). Dissolved oxygen and pH are recorded every second with an LDO sensor (Mettler Toledo, Inpro 6870i) and a pH-sensor (Mettler Toledo HA 405-DXK-S8/225). An acetate pulse of 60 mg COD/L was added according to Gernaey et al. (2002). Exogenous oxygen uptake rate (OUR) profiles are calculated similar to Petersen (2000).

2.2. Respirometric model

A simple respirometric model (Eqs. (1)-(3)) for aerobic degradation of acetate without storage was used (Gernaey, 2002). It predicts the model output variable exogenous oxygen uptake rate: OUR_{ex} ($\text{mg L}^{-1} \text{d}^{-1}$).

$$\frac{dS}{dt} = -\left(1 - e^{-\frac{t}{\tau}}\right) \mu_{max} \frac{1}{Y} \frac{S}{S + K_S} X \quad (1)$$

$$\frac{dX}{dt} = \left(1 - e^{-\frac{t}{\tau}}\right) \mu_{max} \frac{S}{S + K_S} X - b X \quad (2)$$

$$OUR_{ex} = \left(1 - e^{-\frac{t}{\tau}}\right) \mu_{max} \frac{1-Y}{Y} \frac{S}{S + K_S} X \quad (3)$$

with S : substrate concentration (mg COD L^{-1}), X : biomass concentration (mg COD L^{-1}), τ : retardation on biomass activity (d^{-1}), Y : yield of the biomass (-), μ_{max} : maximum growth rate (d^{-1}), K_S : half-saturation Monod constant (mg COD L^{-1}).

3. Results

3.1. Respirometric experiment and parameter estimation

From the performed experiments, a profile that matched expectations (no storage, biomass adaptation or noticeable experimental errors) was used as test case. The sum of squared errors of the OUR_{ex} profile was minimised with the Shuffled-Complex Evolution (SCE) algorithm (Duan et al., 1992) and resulted in an excellent model prediction (Figure 1). Estimated parameter values are shown in Table 1.

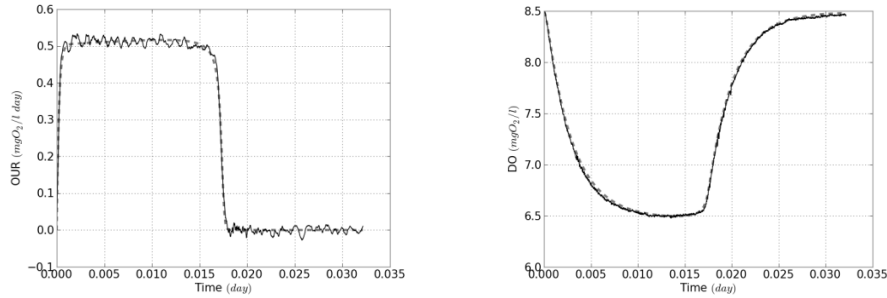


Figure 1. OUR_{ex} (left) and DO (dissolved oxygen; right) profiles of the selected respirometric experiment. Measurements are indicated with black lines, model predictions with grey dotted lines.

3.2. Subjective choices and assumptions

The classical parameter estimation uncertainty method is based on statistical properties of the estimation error. It assumes white noise, i.e., the deviations between model prediction and measurements, is only caused by measurement errors. Parameter CIs can be propagated to the output by linear error propagation (assuming linearity of the model near the optimal point in parameter space) or by MC simulations (no assumption on linearity). It should be emphasised that correlation-based sampling strategies are needed to avoid propagation of impossible parameter combinations (Figure 2).

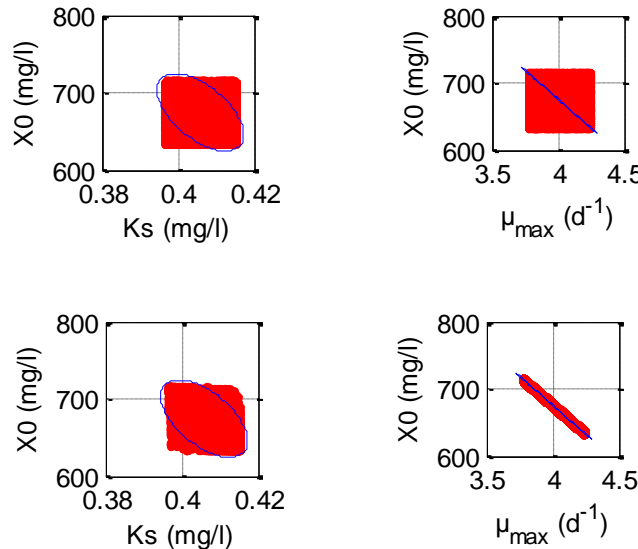


Figure 2. Uncorrelated sampling (top) and correlated sampling (bottom). Blue: statistically calculated 95% CI. Red: sampling points for 10,000 MC simulations from uniform distributions over the parameter CIs.

PDFs for the expert-based method were defined similarly to Benedetti et al. (2010) and are given in Table 1. It is trivial to see that both the selected variation and distribution may have a huge impact on the results. Convergence of the cumulative distributions was checked to ensure a representative sampling.

Table 1. Characteristics of the expert-based PDFs. Nominal parameter values are equal to the calibrated parameter values

Name	Nominal value	Variation	Distribution
μ_{\max}	4 (d^{-1})	20%	Triangular
Y	0.79 (gX/gS oxidised)	5%	Triangular
K_s	0.41 (mg/l)	50%	Triangular
τ	2.26e-4 (d)	50%	Triangular
X_0	675.63 (mg/l)	50%	Triangular

For the GLUE method, prior parameter distributions are set to the PDFs in Table 1 and following fit criteria were selected: SSE and NS. Thresholds are set by visually comparing simulations with different criterion values. Convergence of the cumulative distributions was checked.

3.3. Dealing with correlations

Classical parameter estimation and GLUE are useful to observe and easily handle correlated parameters, which is very important when using these methods on full-scale models. In classical parameter estimation, confidence intervals that tend more to lines than to circles indicate correlated parameters (Figure 2), whereas in the GLUE method, scatter plots of posterior parameter combinations can show interdependency between two parameters (Figure 3).

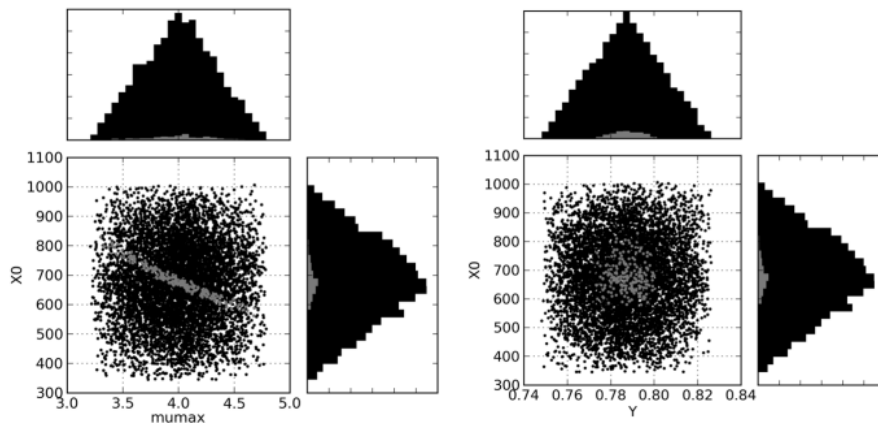


Figure 3. Prior (black) and posterior (grey) parameter distributions of the GLUE method. Correlated parameters show a trend in 'behavioural' parameter combinations (left). Uncorrelated parameters result in scattered 'behavioural' parameter combinations (right)

The relation between initial biomass concentration (X_0) and maximum heterotrophic growth rate (μ_{\max}) is visualized in Figure 4, illustrating the absence of a real optimum in the objective surface due to this relationship. It also hampers a reliable uncertainty calculation and should be removed in advance. One way to do this is by determining the active biomass concentration before the experiment. However, it is difficult to divide biomass in heterotrophic/autotrophic and active/non-active biomass. Applicable methods are currently under investigation.

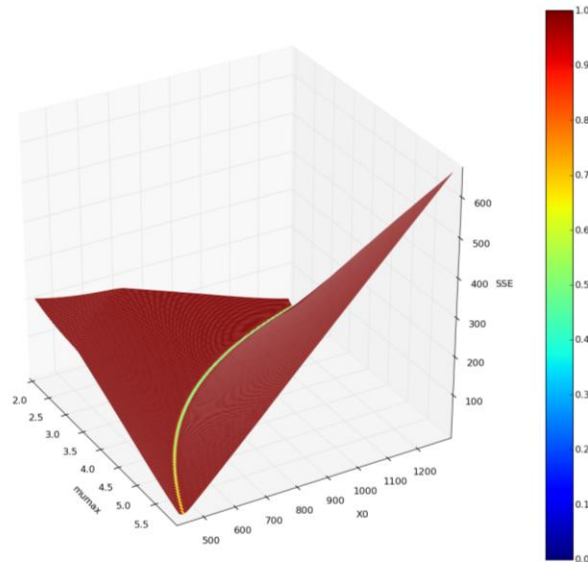


Figure 4. Objective function (SSE) for different combined values of the growth rate (μ_{max}) versus the initial biomass concentration (X_0), and fixing the remaining parameter values.

3.4. Uncertainty boundaries

Uncertainty boundaries on the model output (here 95%) are very specific to the method used and to subjective choices like probability density functions in the expert-based method and selected criteria and threshold values in the GLUE method (Figure 4). The FIM-based method with either correlation-based sampling or linear error propagation (the latter is not shown) results in almost coinciding prediction uncertainty boundaries. Similarity between these methods is logical because they are both based on the parameter estimation correlation matrix.

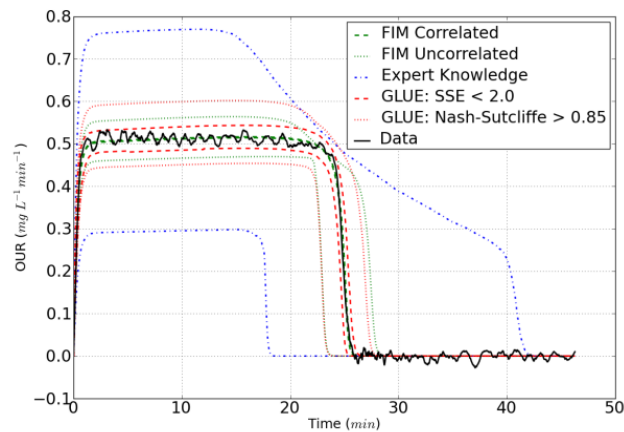


Figure 4. 95% prediction uncertainty boundaries of the exogenous OUR determined with different uncertainty methods.

4. Conclusions and recommendations

This article did not focus on the uncertainty boundaries as such, but on the difference in assumptions, subjectivity and interpretation of three uncertainty analysis methods. There is a clear need for an overall understanding of what uncertainty boundaries determined by different scientists actually mean and a unified approach would be even more preferred. The lack of coherent terminology and systematic approaches is also mentioned by Montanari (2007) and Pappenberger et al. (2006).

In this simple example, the correlation between X_0 and μ_{\max} cannot be disregarded. Since the 'behavioural' parameters of the GLUE method are selected based on a fitting criterion, these correlations are inherently included in the method. This is a major advantage of GLUE over the expert-based method. In classical parameter estimation, correlated sampling is a good alternative. However, more complex models may result in singular matrices, which hamper CI determination.

Different interpretations of parameter CIs are given in Omlin and Reichert (1999). PDFs and resulting CI boundaries should represent frequency distributions of measurements from an 'infinite' number of reproducible experiments. However, PDFs in the applied Bayesian methods are also used to mimic the present knowledge about a certain parameter.

Another emerging need is validation. Indeed, it can be seen that uncertainty boundaries can be obtained in any range you want by the currently applied methods. More experiments and well-thought validation experiments should be performed. Indeed, it is impossible to estimate 'frequentist' PDFs by performing only one experiment.

REFERENCES

- Belia, E., Amerlinck, Y., Benedetti, L., Johnson, B., Sin, G., Vanrolleghem, P.A., Gernaey, K.V., Gillot, S., Neumann, M.B., Rieger, L., Shaw, A. and K. Villez, Wastewater treatment modelling: dealing with uncertainties, *Water Science and Technology*, 60(8), 1929-1941, 2009.
- Benedetti, L., Bixio, D., Claeys, F. and P.A. Vanrolleghem, 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, 2008.
- Benedetti, L., De Baets, B., Nopens I. and P.A. Vanrolleghem, Multi-criteria analysis of wastewater treatment plant design and control scenarios under uncertainty, *Environmental modelling & software*, 25(5), 616-621, 2010.
- Benedetti, L., Batstone, D., De Baets, B., Nopens, I. and P.A. Vanrolleghem, Uncertainty analysis of WWTP control strategies made feasible, *Water Quality Research Journal of Canada*, accepted Feb 2012.
- Beven, K., A manifesto for the equifinality thesis, *Journal of hydrology*, 320(1-2), 18-36, 2006.
- Beven, K., Smith, P. and J. Freer, So Just Why Would a Modeller Choose to Be Incoherent? *Journal of Hydrology* 354,15-32, 2008.
- Bixio, D., Parmentier, G., Rousseau, D., Verdonck, F., Meirlaen, J., Vanrolleghem, P.A. and C. Thoeys, A quantitative risk analysis tool for design/simulation of wastewater treatment plants, *Water Science and Technology*, 46(4-5), 301-307, 2002.
- Duan, Q., Sorooshian, S. and V. Gupta, Effective and efficient global optimization for conceptual rainfall-runoff models, *Water resources research*, 28(4), 1015-1031, 1992.
- Gernaey, K., Petersen, B., Nopens, I., Comeau Y. and P.A. Vanrolleghem, Modeling aerobic carbon source degradation processes using titrimetric data and combined respirometric-titrimetric data: experimental data and model structure, *Biotechnology and bioengineering*, 79(7), 741-753, 2002.
- Iman, R.L. and W.J. Conover, A distribution-free approach to inducing rank correlation among input variables, *Communications in statistics part b-simulation and computation*, 11, 311-334, 1982.

- Li, L., Xia, J., Xu, C.-Y and V.P. Singh, Evaluation of the Subjective Factors of the GLUE Method and Comparison with the Formal Bayesian Method in Uncertainty Assessment of Hydrological Models, *Journal of Hydrology*, 390, 210-221, 2010
- Mannina, G., Cosenza A., Vanrolleghem P.A.. and G. Viviani, A practical protocol for calibration of nutrient removal wastewater treatment models, *Journal of hydroinformatics* 13(4), 575–595, 2011.
- Montanari, A, Large Sample Behaviors of the Generalized Likelihood Uncertainty Estimation (GLUE) in Assessing the Uncertainty of Rainfall-runoff Simulations. *Water Resources Research*, 41, W08406, 2005.
- Montanari, A, What do we mean by ‘ uncertainty ’? The need for a consistent wording about uncertainty assessment in hydrology, *Hydrological Processes* 21, 841-845, 2007.
- Omlin, M. and P. Reichert, A comparison of techniques for the estimation of model prediction uncertainty, *Ecological modelling*, 115, 45-59, 1999.
- Pappenberger, F., and K. J. Beven, Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resources Research* 42, W05302, (2006).
- Petersen, B., Calibration, identifiability and optimal experimental design of activated sludge models, *Phd thesis, Universiteit Gent*, 320p, *gent*, 2000.
- Reichert, P. and P.A. Vanrolleghem, Identifiability and uncertainty analysis of the river water quality model No. 1 (RWQM1), *Water science and technology*, 43, 329-338, 2001.
- Sin, G., Gernaey, K.V., Neumann, M.B., van Loosdrecht, M.C.M. and W. Gujer, Uncertainty analysis in WWTP model applications: A critical discussion using an example from design. *Water Research*, 43(11), 2894-2906, 2009.
- Vanrolleghem, P.A., Spanjers, H., Petersen, B., Ginestet, P. and I. Takacs, Estimating (combinations of) Activated Sludge Model No. 1 parameters and components by respirometry, *Water Science and Technology*, 39(1), 195-214, 1999.