Chapter 4

Available methods for uncertainty analysis in model-based projects – critical review

4.1 INTRODUCTION

This chapter reviews and summarises the uncertainty analysis methods described in the published literature from the wastewater treatment field. The objective of the review is to capture the breadth of the state of uncertainty analysis within wastewater treatment and does not attempt to focus on a detailed evaluation of individual methods. Most publications reviewed date between 1958 and 2009, with key papers added between 2009 and 2013. Some of the more popular methods referenced in this chapter are illustrated in more detail in Appendix B. Appendix C includes the full list of papers reviewed by the Task Group (including about two topics not covered in this chapter: on-line control and regulatory issues).

Appendix C also includes more recent publications (2011–2019) covering a wide range of uncertainty topics (not reviewed in this chapter).

The presentation in this chapter is of a highly technical nature. Much of the discussion assumes a familiarity with the mathematics that underlie probability theory. The chapter is not meant to be a statistics compendium; it is rather structured as a review article to point interested and theoretically inclined individuals to the body of literature containing examples and discussion of how uncertainty analysis has been applied to wastewater treatment problems. Many practitioners that use treatment plant models for design, or to analyse operational issues, may not find that the material within provides them with guidance that can be applied to their day-to-day work. The connection between this material, and its potential practical applications is the subject of Chapter 5.

Table 4.1 lists the main methods available for uncertainty assessment in model inputs, model parameters, model structure and model-based decision-making. The methods covered in this chapter address **variability and quantifiable uncertainty**. Additional details on the methods referenced in Table 4.1 can be found in Appendix B.

© IWA Publishing 2021. Uncertainty in Wastewater Treatment Design and Operation: Addressing Current Practices and Future Directions
Editor(s): Evangelia Belia, Lorenzo Benedetti, Bruce Johnson, Sudhir Murthy, Marc Neumann, Peter Vanrolleghem and Stefan Weijers
doi: 10.2166/9781780401034_0047

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Table 4.1 Uncertainty assessment methods (see Appendix A for definitions).

Category	Method
Model inputs	Summary statistics Statistical tests Outlier detection Data reconciliation Principal components analysis (PCA)
Model parameters	Estimation of inference (confidence) region Bayesian statistics
Model structure	Compartmental modelling
Model-based decision making	Monte Carlo simulation

4.2 METHODS AND LITERATURE REVIEW RESULTS SUMMARY

Literature searches were conducted on ISI/Web of Knowledge (Science), Compendex, Scopus, and Pollution Abstracts and Toxicology Abstracts. The list of wastewater treatment uncertainty references collected is included in Appendix C.

Table 4.2 summarises the results of the search and provides a framework to synthesise the considerable breadth of the topic and size of the literature search results into discrete topics. The categorisation in Table 4.2 is based on subjective judgement and there are several references assigned to one category that address other categories.

It must be noted that not all of the above categories could be covered comprehensively in a succinct manner within this chapter. Specifically, categories 4 and 5 are not covered in this chapter. However, the

Table 4.2 Literature search results.

No.	Category	Description	Number of Papers
1	Input and parameters	References that provide information on the uncertainty in model parameters (single values for steady-state models) or in input time series (inputs for dynamic models)	51
2	Model structure	References that address uncertainty generated from the structure of a wastewater model or references that address mathematical concepts related to uncertainty within the context of a wastewater treatment model	58
3	Propagation of uncertainty	References that address uncertainty evaluation of one or more different treatment trains or plant-wide alternatives in model-based decision-making	119
4	On-line control signals and strategies	References that consider the uncertainty of an on-line measurement or references dealing with the use of on-line signals within a real-time control loop	92
5	Fate of pollutants in the environment	References that address the uncertainty associated with the presence of pollutants in the environment and resulting regulatory (WRRF effluent standards) issues	85

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

reader may review the references under these categories in Appendix C as a starting point for further information on these subjects.

This chapter focuses on the first three categories listed in Table 4.2. Within each category, abstracts were further screened, and a number of papers was selected for detailed review.

4.3 ASSESSMENT OF INPUT AND PARAMETER UNCERTAINTY

4.3.1 Input uncertainty (measurement errors)

This area of research focuses on quantifying the uncertainty in model inputs and on the use of techniques to minimise the model input uncertainty before performing further analysis. Input uncertainty is due to measurement errors and here the methods implemented in the literature to quantify them are discussed. Measurements contain uncertainty due to random, systematic and gross errors (for a definition of measurement errors, see Chapter 3, Section 3.6.2.2 and Appendix A).

Methods for quantifying the uncertainty stemming from the assumption that the measured data are an unbiased estimate of the underlying population have not been included in this review.

4.3.1.1 Overview of statistical techniques used in measurement error detection

The topics addressed in the reviewed papers include quantifying the uncertainty in data and data collection methods and identifying systematic errors in data. The uncertainty in measured data is typically assessed using statistical techniques. Standard Methods for the Examination of Water and Wastewater (1998) provides background on measurement uncertainty and basic statistical techniques used to quantify this uncertainty as applied to the examination of wastewater. Skoog *et al.* (1995) describe the use of these uncertainty evaluation techniques in the field of analytical chemistry.

The precision of a measurement can be determined by replicate measurements and by the calculation of the standard deviation and variance of the replicates. A common technique for assessing the accuracy of measurements and detecting systematic errors in an analytical method or an instrument is to analyse a sample whose composition is accurately known (i.e., a calibration check standard). Statistical hypothesis testing is then used to assess whether the difference between the measured value and the known calibration check standard value could be caused by random error or systematic error. Outlier detection tests are available for detection of large biases but must be used cautiously (Skoog *et al.*, 1995). More advanced techniques, such as data reconciliation, may be more suitable for outlier detection where applicable (see discussion below).

4.3.1.2 Error propagation

In cases where results are computed from multiple sources of experimental data or a calibration curve is used to provide the measured value, it is necessary to determine how the error in the measured values is transmitted to the results. For general non-linear functions, a few uncertainty assessment approaches are available such as the law of propagation of uncertainty, Monte Carlo simulation, and empirical sensitivity studies based on designed experiments (Standard Methods for the Examination of Water and Wastewater, 1998). A common approach is to use the law of propagation of uncertainty where the function of interest (e.g., the calibration equation) is linearised using a first-order Taylor-series expansion about the key variables and then the variance formula for a linear sum of variables is used to calculate the total variance (Box *et al.*, 1978).

In the wastewater treatment literature, these statistical techniques have been used to quantify the uncertainty in various types of wastewater measurements. Friedler and Butler (1996) used basic

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

statistical techniques to quantify the inherent uncertainty in the quantity and quality of wastewater discharged from domestic appliances. The analysis was based on data from two surveys conducted in the United Kingdom. The appliance volumes, pollutant loads, and frequency of use were not found to be normally distributed variables. The authors suggest that Monte Carlo analysis could be used to help quantify the combined effects of the uncertainties at the household level on the overall uncertainty in the wastewater flow rate and concentrations within a sewer system.

4.3.1.3 Examples of measurement error detection

Joannis *et al.* (2008) studied the uncertainty in wastewater turbidity measurements. They found that the major sources of uncertainty were in the standard solutions used for calibration and the nonlinearity of the calibration curve. Bertrand-Krajewski *et al.* (2007) compared the uncertainties in COD measurements between standard laboratory techniques, small tube tests (STT) employing a photometer, and field UV−visible spectrometry. They found that standard laboratory methods and small tube tests had similar levels of uncertainty but had different mean values, indicating that specific calibration functions are needed to correct systematic errors if high accuracy is required or the methods are compared. Bertrand-Krajewski *et al.* (2007) found that the use of low-frequency sampling is the major source of uncertainty with standard laboratory methods and methods for COD determination. UV−visible spectrometry was found to have a similar level of uncertainty as standard laboratory methods but only under strictly controlled conditions.

Rieger *et al.* (2005) evaluated the uncertainty of on-line sensors at WRRFs using comparisons between independent measurements of the same sample (i.e., sensor and a reference laboratory method). The comparison is based on a linear regression fitted between the sensor and reference measurements. The authors assess whether the linear regression is applicable by considering a relationship between the variables (using an F-test), checking the linearity between the sensor and reference measurements (using an F-test), checking for outliers, and checking the homogeneity of the variances. If the linear regression is applicable, statistical tests on the regression predictions and the regression slope and intercept are used to assess whether the regression equation is significantly different from the perfect correlation (i.e., slope = 1 and intercept = 0), indicating the presence of systematic errors. If no systematic errors are detected, the total uncertainty is represented by a confidence interval for the regression predictions, assuming a perfect correlation between the sensor and reference measurements. If systematic errors are detected, the bias is quantified by the linear regression fit between the sensor and reference measurements. The random errors are quantified by the confidence interval for the regression predictions.

4.3.1.4 Multivariate statistical methods

Robinson *et al.* (2005) used multivariate statistical methods to identify outliers in water quality data. They advocated the use of multivariate statistical methods due to the correlation between plant variables. Methods discussed include: Mahalanobis distance, jack-knife distance, and Hadi's method. These methods are more applicable than univariate methods but do not account for the serial correlation of the variables over time. Multivariate statistical control methods are available that can address serial correlation, as discussed below.

There are also more advanced statistical methods available for detecting and removing systematic errors and gross errors. These methods range from statistical process control and fault detection methods to data reconciliation.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

4.3.1.5 Statistical process control and fault detection methods

Statistical process control techniques involve monitoring process variables over time using statistical control charts. The variables of interest are charted over time and compared to control limits to determine if the process is within control (i.e., its correlation structure is unchanged). The methodology is used to distinguish between common cause variability and special causes. Typically, these methods are used to study process variability but can also be used in the current context to detect sensor or measurement process faults leading to large systematic or gross errors.

Because WRRF measurements can exhibit autocorrelation, seasonality, and non-constant variance (Berthouex, 1989), it can be difficult to apply traditional control charts such as Shewhart or cumulative sum (CUSUM) charts to the measured process variables themselves. One option discussed by Thomann *et al.* (2002) and Thomann (2008) is to create a control chart that tracks the difference between the sensor values and reference values at a WRRF. Unfortunately, this approach may not always be practical if reference measurements are not taken at a suitable interval. Another approach is to fit a time-series model such as an auto-regressive integrated moving average (ARIMA) model to normal operating data and then use the model as a charting tool (Berthouex, 1989). The model would be used to continually predict process data given the previous data and the difference between these predictions and the actual measurements (i.e., residuals), are plotted on a conventional control chart. When the measurements are collected normally, the residuals will be independent, random, and have constant variance.

A simpler alternative is to construct exponentially weighted moving average (EWMA) charts. The one-step-ahead prediction errors (i.e., residuals between predictions and actual measurements), calculated using the EWMA statistic, can be plotted on a traditional control chart. As discussed by Montgomery and Mastrangelo (1991), the EWMA approach can be a reasonable approximation of the ARIMA model approach in many cases. For suitably selected value of the EWMA filter constant, the EWMA statistic is an excellent one-step-ahead predictor for processes where the mean does not shift too rapidly, and the observations are positively auto-correlated.

The limitation of univariate control charts is that they do not consider the correlation between the variables within the process. Some researchers have looked at the use of multivariate statistical techniques such as principal components analysis (PCA) and partial least squares (PLS) to monitor process data. PCA involves projecting multivariate data into a lower dimensional or latent variable space. The variables in the lower dimensional space are uncorrelated and explain the majority of the variance in the data. PLS is a latent variable regression method used when multivariate input and/or output data are available. The PLS regression model captures the correlation between the inputs and outputs in a lower dimensional variable space. PCA and PLS models used for monitoring are built using data from normal operation so that they model the normal measurement variability and correlation.

Using a PCA model for example, one can create control charts for outputs from the model including t-scores, the Hotelling T^2 statistic, and the squared prediction error (SPE) to detect unusual measurements in a multivariate context (Kourti & MacGregor, 1995). Although PCA and PLS consider static covariance relationships, they can be adapted to the analysis of dynamic data by including time-lagged data into the data matrices so that the correlation over time is included into the models. In a wastewater treatment context, some examples of latent variable monitoring are discussed by Lennox and Rosen (2002) and Tomita *et al.* (2002).

4.3.1.6 Data reconciliation

Data reconciliation is a technique used to adjust process measurements so that they are consistent with known conservation laws and other process constraints. The procedure requires a set of redundant

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

measurements to verify that the conservation laws have been obeyed. Optimal data reconciliation is a constrained least-squares problem requiring the minimisation of a weighted sum of the measurement adjustments subject to the process constraints. The weighting matrix is typically the inverse of the variance−covariance matrix of the errors in the measurements. The weighting may be selected based on previous experience, calculated using the sample variance for the data, or using robust estimators (Chen *et al.*, 1997). The measurements in the data reconciliation procedure are weighted inversely to their variance so that measurements with large variance are adjusted more than those with a smaller variance. Therefore, the success of the method is dependent on the use of reasonable variance estimates.

In the context of biochemical reactions, data reconciliation has been studied by Van der Heidjen *et al.* (1994a, b, c). Mass balancing has been discussed in a WRRF context by Nowak *et al.* (1999) and Barker and Dold (1995), while more formal data reconciliation analyses have been reported by Meijer *et al.* (2002), Puig *et al.* (2008), and Thomann (2008). Recently, Rieger *et al.* (2010) discussed data reconciliation for WRRF simulation studies. Their focus was on planning measurement campaigns so that high-quality data can be collected. They discuss the use of basic reliability checks and manual checking of mass balances to verify the quality of the data and to identify systematic errors.

In a WRRF context, it is most common to reconcile flow and total phosphorus measurements across the plant, and suspended solids measurements around clarifiers and thickening and dewatering processes. COD and total nitrogen measurements could be also potentially reconciled using mass balances, but this typically requires measurements not typically collected such as the oxygen utilisation rate (OUR), oxygen transfer parameters, and nitrogen gas flows.

Data reconciliation can be performed using either a steady-state or dynamic analysis. Steady-state data reconciliation is commonly performed using averaged measurements over a period of approximate steady-state or zero accumulation. Examples in a WRRF context are provided by Meijer *et al.* (2002) and Puig *et al.* (2008). Puig *et al.* (2008) suggest that the data be averaged over a period of at least two to three sludge retention times. In the case of steady-state reconciliation, the process constraints (typically mass balances) are assumed to be known and the measurements are considered to be stochastic.

In dynamic data reconciliation, the process constraints are typically dynamic process models so that uncertainty is considered to exist in the model structure and parameters, and the measured data. Dynamic data reconciliation can be conveniently performed in a simulation environment by minimising a weighted least-squares function of the measurement adjustments, subject to the WRRF model, over successive time horizons or windows. This method is known as the horizon method (Romagnoli & Sanchez, 2000). In the horizon method, the initial values of the model states for each time horizon are the optimisation variables.

Dynamic data reconciliation can also be performed using a filtering approach based on the extended Kalman filter (Romagnoli & Sanchez, 2000). In this context, the filter acts as a state estimator which takes the states predicted by the model and adds the filtered difference between the measured and predicted model outputs. The filtering approach has an advantage in that its calculations are recursive and do not require iteration as in the horizon method. The horizon method is thought to be better suited to slower processes (Cameron *et al.*, 1992), such as biological growth, while the filtering method is thought to be better suited to faster processes.

Following data reconciliation, gross error detection techniques can be used to identify and eliminate systematic errors caused by sensors and other faults. Gross error tests involve statistical hypothesis testing on the least-square's objective function (which is a Chi-square variable), and on the ratios of the measurement adjustments and the mass balances errors to their standard deviations. Meijer *et al.* (2002) and Thomann (2008) illustrate the use of simple gross error detection techniques in a WRRF context. More sophisticated techniques are discussed by Crowe (1996) and involve the use of tests of maximum

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

power in detecting gross errors for the case of a single gross error and the use of PCA in cases where multiple gross errors exist.

The use of formal data reconciliation in the wastewater treatment field has been limited due to a lack of data, the complexity of the solution procedure, and the lack of availability of software dedicated to the solution of the data reconciliation problem. For simulation studies, most model calibration protocols typically recommend basic reliability checks and manual evaluation of mass balances to verify the quality of the data and to identify systematic errors (Rieger *et al.*, 2013). This is expected to change in the future as online instrumentation becomes common in WRRFs, and modelling and other software vendors add the necessary tools to their products. Data may be taken at different intervals, contain missing measurements, have redundancy, and could sometimes be erroneous. Adaptation of existing mathematical tools into the databases and SCADA systems of full-scale plants will be required in order to promote the progressive incorporation of advanced monitoring systems, decision support systems, and plant-wide controllers.

4.3.2 Parameter uncertainty

Whole plant models are complex and have hundreds of parameters, all with some uncertainty. While most of those parameters can be assumed to be fixed, others require to be considered uncertain, given their importance towards the results of the model use.

Uncertainty in model parameters arises from many sources such as the model structure, their measurement error (in case they are directly or indirectly measured), the choice of experimental conditions used for model calibration, the calibration data, and the objective function or criterion used for parameter estimation.

4.3.2.1 Inference vs. confidence regions

Uncertainties in model parameters are typically assessed during the process of parameter estimation. See Bard (1974), Draper and Smith (1989), and Bates and Watts (1988) for the theory of nonlinear parameter estimation. Parameter estimation problems are often solved using maximum likelihood estimation. Depending on the assumptions on the error structure, the number of response variables, and the available information on the variance and covariance of the errors, the objective function minimised during the procedure ranges from ordinary least squares, to weighted least squares, to the Box−Draper criterion (Box & Draper, 1965).

Parameter uncertainty is typically assessed following parameter estimation, through the calculation of approximate joint confidence regions for the parameters and approximate confidence limits on individual parameters. The inference regions and limits or bands are often estimated by extending linear regression theory. The model residuals are linearised using a Taylor-series expansion and analogous formulas as those used for linear regression inference regions and bands are developed (Draper & Smith, 1989).

The inference region formulas require the calculation of the variance−covariance matrix for the parameters. The variance−covariance matrix is often approximated as the inverse of the Hessian matrix of the objective function (i.e., matrix of second derivatives of the objective function) multiplied by a scale factor at the solution to the parameter estimation problem (Bard, 1974). In the linearisation approach, the Hessian matrix is calculated using first-order model sensitivity coefficients only (Gauss−Newton approximation). First-order model sensitivity coefficients, which express the local sensitivity of the process model to infinitesimally small changes in the model parameters, are defined as the partial derivatives of the model with respect to the model parameters. The sensitivity coefficients can be determined using finite-difference approximations, by solving the model sensitivity equation (Leis &

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Kramer, 1988), using variational methods, or by automatic differentiation (De Pauw & Vanrolleghem, 2003). Alternatively, it is also possible to approximate the variance−covariance matrix using the full Hessian matrix which requires the calculation of second-order sensitivity coefficients which can be calculated as shown by Guay and Maclean (1995).

4.3.2.2 Application to wastewater treatment models

In the context of biokinetic models of the activated sludge process, the calculation of approximate inference regions for the model parameters has been discussed by numerous researchers including Vanrolleghem *et al.* (1995), Vanrolleghem and Keesman (1996), Brouwer *et al.* (1998), Petersen (2000), Petersen *et al.* (2000), Dochain and Vanrolleghem (2001), Marsili-Libelli and Tabani (2002), Sin (2004), Checchi and Marsili-Libelli (2005), and De Pauw (2005). Parameter estimation techniques are most often applied in a wastewater treatment context when fitting biokinetic models to respirometric batch experiments. Formal parameter estimation (e.g., using maximum likelihood estimation) is not recommended for calibrating entire WRRF models to historical plant data due to the lack of data, the complexity of the models, and the correlation between the model parameters (Petersen, 2000; Vanrolleghem *et al.*, 2003). Historical plant data are rarely suitable for estimating complex model parameters and for assessing their uncertainty due to missing data, inconsistencies in the data, limitations in the ranges of the variables due to process control, confounding effects between variables, and variations in unmeasured variables (Box *et al.*, 1978). Typically, model calibration focuses on influent characterisation, accurate modelling of plant hydraulics and aeration, and manual adjustment of some model parameters to achieve a reasonable fit between the measured data and model outputs.

In the context of parameter estimation, it is possible to design experiments that minimise the uncertainty in the estimated parameters. One common approach, introduced by Box and Lucas (1959), is to minimise the volume of the parameter confidence region. This involves minimising the determinant of the inverse of the variance−covariance matrix. A sequential strategy is often used, as the best set of experimental conditions depends on the parameter values (Box *et al.*, 1978). Vanrolleghem *et al.* (1995) discuss the use of optimal experimental design in the context of activated sludge models (ASMs) and list alternative optimal design criteria.

The main drawbacks of these approximate parameter uncertainty assessment methods are that they assume that only the response variables in the parameter estimation procedure contain uncertainties, they use an approximation to the variance−covariance matrix, and they are specific to the local solution to the parameter estimation problem.

4.3.2.3 More sophisticated methods

It is possible to reformulate the parameter estimation problem using an error-in-variables (EIV) approach so that both the independent and dependent variables (i.e., all model inputs) are considered to contain uncertainties (Bard, 1974; Romagnoli & Sanchez, 2000). This becomes a simultaneous data reconciliation and parameter estimation problem.

Better estimates of the parameter uncertainties can be obtained using the Monte Carlo method where the parameter estimation problem is solved repeatedly for different simulated samples of the measured model inputs leading to a distribution of parameter estimates (Bard, 1974). Another option is to use the technique known as profiling (Bates & Watts, 1988), after parameter estimation, to obtain exact marginal likelihood intervals for the model parameters. Cox (2004) used Bayesian statistics to develop uncertainty distributions for the parameters in the ASM1 model. That procedure involves combining expert opinion (prior distribution) and measured or calibrated parameter values into a single posterior distribution known as a

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

universal distribution. The method is promising but the specific application given by Cox may not be useful given that many of the calibrated values are taken from subjective calibration studies involving historical plant data.

4.4 ASSESSMENT OF MODEL STRUCTURE UNCERTAINTY

A portion of the references identified by the literature search addressed the fundamental wastewater process model uncertainty issues of (1) model structure and (2) mathematical methods. These areas of the literature search may be of most interest to researchers investigating fundamental information and approaches to uncertainty assessment in wastewater treatment engineering. Some of the most compelling works published in these areas are discussed in the following sections.

4.4.1 Macroscopic vs. microscopic mixing scales

Several researchers have investigated structural issues at the core of activated sludge models (ASMs) that arise from the conceptual basis of some state variables and assumptions used in modelling of the completely stirred tank reactor (CSTR) configuration. While Danckwert's (1958) and Zwietering's (1959) seminal publications on residence time distribution and reactor modelling identified the influence of the nature of the reactants, reaction rates, and local mixing scale on CSTR reactor analysis, and while chemical engineering text books (e.g., Levenspiel, 1999; Rawlings & Ekerdt, 2002) have further formalised the concepts to recognised limiting cases of 'complete segregation' and 'maximum mixedness', it was perhaps not until Gujer (2002) that the importance of these concepts to ASMs was noted.

Gujer (2002) observed that ASM state variables representing cell internal storage products are conceptually linked to a local environment (an individual cell), and that reaction rates in ASMs are not necessary first order. There are, therefore, resulting consequences of applying ASMs on a macroscopic and microscopic mixing scales that impact the applicability of kinetic parameters determined for and by different reactor configurations. Gujer (2002) considered a model with a simple single substrate with cell storage product and applied it in both the typical macroscopic fashion and in a microscopic fashion which tracked individual bacteria and used a probabilistic rule to control the residence time of a bacterium within a CSTR zone. While Gujer (2002) notes that this simple model is not directly applicable to any relevant system, the results he presents identify the nature of this basic model structure issue and lead to important conclusions regarding the differences between the determination and applicability of kinetic parameters for sequencing batch reactors and flow through systems.

Gujer (2002) suggests that (1) there may be an inherent amount of quantifiable uncertainty in ASM results associated with reactants' local environment and residence time and (2) that there may be uncertainty or inaccuracy induced into modelling efforts by application of kinetic parameters estimated from differing flow schemes.

Schuler (2005, 2006) extended the concepts of Gujer (2002) to a model that included the competition between phosphorus accumulating organisms and heterotrophic organisms, without nitrifiers and the potential interference of nitrate, and applied the model to reactor configurations relevant to activated sludge treatment systems. Schuler (2005, 2006) illustrated the differences in results that occur between a lumped parameter (macroscopic) model structure and a (microscopic) model structure that includes the distributed state of reactant residence times and concluded that the lumped parameter approach consistently predicted better effluent phosphorus performance.

Curlin *et al.* (2004) also reported on this issue and applied formal concepts from the field of chemical engineering. They used activated sludge model (ASM) No. 1 for a laboratory-scale membrane bioreactor. They established the macroscopic mixing characteristics of their system through tracer studies

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

and arrived at a CSTR combination that fits the experimental tracer results. They then solved the ASM not only using the CSTR assumption typically employed (the ideal mixing case, an assumption of constant reactant concentration over the reactor volume) but also microscopic mixing scale limiting cases of complete segregation and maximum mixing.

Curlin *et al.* (2004) conclude that (1) in ASMs that do not include cell internal storage product state variables, the non-first-order reaction rates may result in un-quantified model structure uncertainty if only the ideal mixing CSTR assumption is considered, (2) the model structural uncertainty generated by imperfect knowledge of microscopic mixing might be quantified through consideration of the limiting cases and (3) the magnitude of the uncertainty may be significant. A validation of the simplified transport models is suggested, compared to more sophisticated approaches.

It is worthwhile to note that other researchers, including Lee *et al.* (1999a, b) and Makinia and Wells (2000a, b) have considered the impact of mixing conditions and residence time distribution in ASMs, using an advection and dispersion equation approach. This line of development, as well as combined unit process modelling and computation fluid dynamics modelling, may also be useful for the identification or reduction of non-ideal mixing contributions to model structure uncertainty.

4.4.2 Unquantified model structure uncertainty

Other researchers have reported on a variety of specific structural issues related to the modelling of wastewater treatment processes that may contribute to what is, at this time, un-quantified model structure uncertainty. Abusam and Keesman (2002) carried out a factorial sensitivity analysis on the use of the double-exponential function in secondary clarifier models and concluded that the model had a structural problem related to the prediction of solids in the underflow stream. Haider *et al.* (2003) reported on experimental results that illustrated that the characterisation of modelled influent non-biodegradable substrate was not independent from the system sludge age and concluded that two such influent biodegradable state variables may be required for models of short sludge age systems. Lavallee *et al.* (2005) similarly noted that observed kinetic parameters depend on substrate, process configuration and sludge, and introduce an ASM within the ASM framework that mimics enzyme induction and may lead towards models applicable over more widely varying conditions. Sin and Vanrolleghem (2006) observed that, even with constant influent conditions, the ASM2d model structure had to be adapted in response to changes in system behaviour observed for three different operational scenarios to match experimental findings. Neumann and Gujer (2008) provided an analysis of model structure uncertainty by generating synthetic data with one model structure (using Tessier rate equation) and fitting it with another putative model (using the Monod rate equation). They illustrated the application of a range of methods for analysing model fit and for propagation of parameter uncertainty to modelling results. This made possible a comparison of model predictions with parameter uncertainty addressed to be compared to the 'true' result and illustrated that the propagation of parameter uncertainty was not adequate to address an error in model structure. Neumann and Gujer (2008) concluded that uncertainty estimates obtained from regression of time-continuous environmental systems should be used with caution.

This is a small sampling of reports in the literature that may be taken to illustrate the degree of structural uncertainty present in wastewater treatment models. In general, these reports suggest recognition that the current wastewater models may be calibrated only to narrow and specific ranges of influent, operating and process configurations conditions, and that any extension of the use of a model outside its specific range of calibration may induce what is, at this time, an un-quantified degree of uncertainty.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

4.4.3 Mathematical methods for quantification of model structure uncertainty

While some researchers have focused on and provided useful information on issues of uncertainty within wastewater treatment model structure, other researchers have considered issues related to the mathematics and numerical methods of the quantification of uncertainty.

4.4.3.1 Monod growth model

Tenno and Uronen (1995) applied an ASM1-like model structure within a stochastic model to arrive at a method of carbon removal process control using on-line instrumentation. Kops and Vanrolleghem (1996) investigated the incorporation of uncertainty analysis into wastewater modelling predictions through consideration of the Monod growth model. They compared three methods for approximating prediction uncertainty: Monte Carlo simulation, Monte Carlo simulation with stochastic parameters, and stochastic differential equations. They identified the Monte Carlo simulation as having the disadvantage that for a dynamic simulation all the parameters must stay constant during one model run. The other two alternatives permit time-varying parameters within one model run. They identified the disadvantages of the stochastic differential equation alternative as requiring parameters to have Gaussian distributions and that stable solutions may be limited to a certain range of parameter values. The Monte Carlo simulation with stochastic parameters did not have the disadvantages of the stochastic differential equation alternative. For the situation examined, Kops and Vanrolleghem (1996) found the stochastic differential equation alternative to give a higher predicted variable variance than either of the Monte Carlo-based methods. In their case, the stochastic differential equation method generated a variance of almost 40% while the Monte Carlo with stochastic parameters simulation resulted in a variance of 1%. They concluded that the comparison of these two alternatives becomes a question of which result is more realistic. This appears to remain a crucial, valid and unresolved issue.

Omlin and Reichert (1999) provide a comparison of parameter estimation methods and their related model prediction uncertainty for a simple Monod equation model. They concluded that classical frequentist (i.e., least squares) technique is superior in the case of identifiable model parameters but in the case of poor parameter identifiably, a Bayesian approach is recommendable. Rauh *et al.* (2004, 2007) and Krasnochtanova *et al.* (2009) discuss incorporation of parameter uncertainty at the numerical simulation time step interval level. They note that the efficiency of Monte Carlo methods decreases significantly for higher dimensional systems and present numerical solution algorithms that generate upper and lower bounds on variable uncertainty at each step interval.

4.4.3.2 Non-linear dynamical and chaotic behaviour

The final area of the literature review addressed in this subsection is published work which has considered the potential role of non-linear dynamical and chaotic behaviour in the variability and randomness of wastewater treatment process observations and process models. Graham *et al.* (2007) report the experimental demonstration of chaotic instability in biological nitrification. They operated three highly controlled aerobic chemostats. They indicate that their experimental results and analysis suggest broad chaotic behaviour and conclude that nitrification is prone to chaotic behaviour because of a fragile ammonia oxidising bacteria and nitrite oxidising bacteria mutualism. Zhang and Henson (2001) demonstrated the possibility of multiple steady-state solutions for several continuous biochemical reactor models and advocate the use of bifurcation analysis to aid in obtaining more efficient and complete characterisation of model behaviour. Saikaly and Oerther (2004) and Stroot

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

et al. (2005) investigated the potential dynamical nature of competition between several species for several resources within activated sludge treatment systems. They report on modelling results as well as fluorescence in-situ hybridisation determination of biomass composition within the pilot reactors. They found the potential for population oscillations within the bacterial community under some operating conditions and conclude that such a dynamical nature could contribute to system stability while confounding the tracking of activated sludge system composition when done by limited grab sampling. Ibrahim *et al.* (2008) present results of static and dynamic bifurcation investigations of an activated sludge system model using a rate equation that includes an inhibition term. Their analyses show a complex variety of dynamic results when inhibition is significant, including periodic attractors, point attractors, and chaotic attractors for realistic ranges of parameter values.

Thus, several publications suggest that the non-linearity and dynamical nature of wastewater treatment process models and, perhaps, the systems themselves result in inherent randomness, periodicity or chaotic behaviour. The tools of this branch of mathematics may, therefore, be useful in characterising uncertainty for some conditions and model applications.

4.5 PROPAGATION OF UNCERTAINTY FOR MODEL-BASED DECISIONS

4.5.1 Review of uncertainty propagation methods

Models are used in process engineering to configure and size facilities to reliably meet effluent quality requirements. The research reviewed for this effort included a variety of biokinetic models used for process engineering including ASM1, ASM2, and ASM3 models as well as 2-D clarifier modelling. Amongst all consulted literature sources, there was a common understanding that the models are both relatively complex and the inputs used for modelling include significant uncertainty. With this basis, the researcher's general goals were then to understand the sensitivity of model results to uncertainty in model inputs, and the evaluation of methods for generating robust designs.

The literature reviewed focused on three main subject areas: model calibration, sensitivity analysis, and design optimisation.

The review was structured following a number of criteria that involve the purpose, method, accuracy, difficulty/simplicity, time to do the analysis, data requirement, applicability and stakeholder interest (e.g., researchers, control decision support, design, operation). The review is summarised in Table 4.3 and in the following sections.

In addressing the input uncertainty for modelling, the majority of the researchers relied on a Monte Carlo approach both for determining uncertainty in model outputs and for calibrating models. Variations on the Monte Carlo approach included:

- Using Spearman's rank correlation to determine model sensitivity to input data (Griborio *et al.*, 2007);
- Applying the Hooke−Jeeves direct search technique for design optimisation (Tansel, 1999);
- Applying various statistical analyses techniques with Monte Carlo outputs for model calibration.

In addition to Monte Carlo techniques, researchers used genetic algorithms and the 'ϵ constraint method' to determine optimal designs. Although these methods appear to have their strengths, the special expertise needed to apply them may not make them useful for most process engineers.

The information provided did not allow for a detailed assessment of the time required for each method. Furthermore, because the methods were applied to several different models with varying levels of complexity, a direct comparison is not possible.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Table 4.3 Review of uncertainty propagation methods and analysis works applied for model-based evaluation.

MODEL CALIBRATION

Purpose	Method	Accuracy	Difficulty	Time	Data	Engineering Task	Reference
Model calibration of ASM3	Monte Carlo (MC)	$\pm$ 20% for denitrification with sludge digestion	Medium	Not quantified		Estimation of N removal and sludge production	Koch *et al.* (2001)
Model Calibration of ASM2d	MC + results analysis with MAE, RMSE and Janus coefficient		Medium	2 weeks for 500 MC runs on Pentium IV 3 GHz	Plant data (assumed uniformly distributed), uncorrelated kinetic parameters	Calibration of dynamic models	Sin *et al.* (2008)
Estimate enclosures of state variables in a simple ASM model	MC		More difficult	Not quantified, but would be significant for complex model			Kletting *et al.* (2007)

SENSITIVITY ANALYSIS

Purpose	Method	Accuracy	Difficulty	Time	Data	Engineering Task	Reference
Sensitivity analysis for control strategies	MC + input variables classification		More difficult	Not quantified, but MC $\times$ 10 Not completely automated			Von Sperling (1993)
Uncertainty and sensitivity with ASM1	MC + Spearman's rank correlation		Medium		Only kinetic variables, not influent characteristics		Huo *et al.* (2004)
Estimate secondary clarifier performance with 2Dc	MC		Relatively simple statistical approach	Uncertain	Floc and settling parameters	CFD modelling of SC and some info on SVI and settling relationships	Griborio *et al.* (2007) Kletting *et al.* (2007)

(*Continued*)

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Table 4.3 Review of uncertainty propagation methods and analysis works applied for model-based evaluation (*Continued*).

DESIGN OPTIMISATION							
Purpose	**Method**	**Accuracy**	**Difficulty**	**Time**	**Data**	**Engineering Task**	**Reference**
Dependability of design using non-ASM model	MC + Hooke-Jeeves optimisation	Lower (less sophisticated model)	Simple	Not quantified			Tansel (1999)
Optimised design using a precursor to ASM	Genetic algorithm also briefly discusses coupling with MC	Better probability than non-linear programming approach	High	Not quantified but faster than non-linear programming	Varied process sizing – did not look at variability in kinetic parameters or influent	May be applicable to optimising or calibrating	Doby *et al.* (2002)
Risk-based design (ASM1)	MC	Good (comparing model results to ammonia data)			Plant data and kinetic variables from Cox (2004)		Huo *et al.* (2006)
Low-cost design using ASM3 for BNR	ε constraint method to generate Pareto optimality		More difficult	Not quantified	ASM3/ASM2d standard		Afonso and da Conceição Cunha (2007)
Predict failures of wastewater pipe systems (utility)	Generalised likelihood uncertainty estimation (GLUE)	Sensitive to errors in model and data (subjective)	High (GLUE algorithm more complex than MC)	Not quantified but expected to be long (GLUE requires thousands of model evaluations)	Past failure rate of pipes Hydrological/ climate/geological data of the area flows	Operators, control decision support	Franks (1999)
Compute extreme event statistics in the water quality field (e. g., after pollutant load discharge from CSO to lake) (utility)	MC first order reliability model (FORM) + importance sampling (IS) or LHS Random directional sampling (RDS)	FORM/LHS most accurate	High (FORM algorithms quite complex)	Thousands of simulations (e.g., 15 × 2500 for some methods)	Rainfall COD Water quality model Extreme events Input uncertainty characterisation	Researchers	Portielje *et al.* (2000)

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Table 4.3 Review of uncertainty propagation methods and analysis works applied for model-based evaluation (*Continued*).

DESIGN OPTIMISATION

Purpose	Method	Accuracy	Difficulty	Time	Data	Engineering Task	Reference
Rank stormwater control strategies under uncertainty	MC ranking methods (mean + sd) uniform distributions	Sensitive to type of ranking method used Depends on the scenario	Medium (MC engineering standard)	500 MC simulations (NO CPU time reported)	Flows, BOD, DO, temperature Hydraulic/biochemical data Control strategies (seven total)	Decision makers	Duchesne *et al.* (2001)
Quantify uncertainty for WRRF design / retrofit	MC	Relative1 measure	Medium	Long (days), depends on computational power and case	Plant data (influent load/composition, size, layout, …)	General applicability (researchers, designers, operators, control)	Rousseau *et al.* (2001)
Uncertainty in estimating the cost of WRRF constructions	Linear regression Fuzzy linear regression Fuzzy goal regression	Sensitive to database used for building regression models	Medium (regression is standard practice)	3 simulations	Database on construction cost of domestic/industrial plants in Taiwan	Decision makes	Chen and Chang (2002)
Risk-based WRRF design (replace safety factors)	MC (method of Rousseau *et al.*, 2001)	Relative1 measure for decision-making	Medium (see data requirement)	Long (days), depends on computational power and the plant in question	Influent load/composition Rainfall Plant model Temperature	Researchers and designers (highly relevant) Dedicated software to generate MC samples + run them	Bixio *et al.* (2002)
Integrated process design and control via global optimisation	Non-linear programming (NLP) Mixed integer optimal control problem (MIOPC) Global optimisation methods	Depends on initial layout and starting point for optimisation	Complicated (NLP programming tedious)	1000 CPU seconds (1 s is worth of the computer's processing time) (depends on different solvers)	Plant layout with initial design values Ranges for different design and control parameters Plant model Influent characteristics	Design and control engineers	Moles *et al.* (2003)

(*Continued*)

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Table 4.3 Review of uncertainty propagation methods and analysis works applied for model-based evaluation (*Continued*).

DESIGN OPTIMISATION

Purpose	Method	Accuracy	Difficulty	Time	Data	Engineering Task	Reference
Screen WRRF technologies (emerging + state of the art) with a decision-making framework	Stochastic dynamic programming Latin hypercube sampling Orthogonal arrays	Sensitive to performance of emerging technology (i.e., data) Influent characteristics	Decision-making framework clear yet numerical solution complicated	12 167 model evaluations (CPU time not quantified but expected to be long)	Influent data Emerging WRRF technologies and performance data Uncertainty in data	Technology screening purposes hence for design engineers and decision makers	Tsai *et al.* (2004)
Evaluate WRRF system design/upgrade options	MC (method of Rousseau *et al.*, 2001)	Relative measure for comparison1	Medium	Long (weeks), depends on the computation power and the scenario (9 × 100 MC runs)	Models for WRRF configurations Yearly influent profile/load Influent fractions Climate cost index	Researchers and designers (highly relevant) Dedicated software to generate MC samples + run them	Benedetti *et al.* (2006)
Challenge the traditional design approaches in view of future uncertainty of WRRFs: scenario planning for accounting	Historical plant data analysis		Simple	Pending data collection issues	Historical data on influent load, performance, modifications, changes, … Socio-economic development data	Designers	Dominguez and Gujer (2006)
Control alternatives evaluation for WRRF operation	Monte Carlo + multi-criteria decision-making framework	Relative	Medium	Long (weeks)	Models for WRRF configurations Influent profile/load Influent fractions Cost index	Designers and operators, control engineer; for decision making	Flores-Alsina *et al.* (2009); Flores-Alsina *et al.* (2008)

Note: Monte Carlo (MC).

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

4.5.2 Discussion

4.5.2.1 Model calibration

With respect to how uncertainty affects model calibration, the publications that were reviewed focused on determining wastewater composition and kinetic variables based on available plant data. They present methods relevant to determining the parameters that result in the best fit of the models along with whether the model results are statistically significant.

4.5.2.2 Sensitivity analysis

Several authors focused on the sensitivity of models to the uncertainty of inputs and how that affects the design and performance of treatment facilities. In the simplest approach to sensitivity analysis, the change in model result for a selected output was measured by individually varying input parameters by 10%. Most of these sensitivity analyses were completed using the Monte Carlo approach to varying input parameters. This approach is generally considered more rigorous because it may show the interaction between multiple parameters. As would be expected based on the ASM-based models, each output had a unique set of input parameters that it was most sensitive to. Even when model inputs fell within the expected range, ASMs showed that the uncertainty in the output was significant enough that the ability to meet discharge limits could be uncertain. Similarly, the 2-D clarifier model showed significant uncertainty in secondary clarifier performance due to uncertainty in model inputs.

4.5.2.3 Design optimisation

The studies that focused on robust designs, all generally defined their goal as a design that had the lowest cost, while reliably capable of meeting discharge requirements. The results of these studies generally try to illustrate how increasing or decreasing the amount of money spent changes the risk of being capable of meeting permit requirements.

From the reported studies, Monte Carlo emerges as the most commonly used method of uncertainty analysis when evaluating different WRRF plant design and controller alternatives. While there is no explicit mentioning of how the procedure is applied in these studies, one can infer the following requirements for the uncertainty analysis: (i) a mathematical model describing the system, (ii) uncertainty range and distribution of the parameters in the system (that could be influent data or biochemical parameters). Mostly a uniform distribution is assumed with the upper and lower ranges adopted from literature. There is no standard on the upper and lower range of ASM parameters, (iii) uncertainty analysis typically represented by a cumulative distribution function (CDF) or by a mean accompanied by a standard deviation.

Besides the Monte Carlo method, the following methods are alternatively used (i) generalised likelihood uncertainty estimation (GLUE) which is a Bayesian approach, (ii) fuzzy linear regression method, (iii) stochastic dynamic programming. All these alternative methods add complexity since the user is expected to have some skills and expertise in statistical and numerical programming. It is the authors' opinion that Monte Carlo simulation is intuitively simple hence can be understood by a larger number of practitioners.

4.5.2.4 Computational demand

About the computational demand of uncertainty analysis methods, it depends on the method used. The number of Monte Carlo simulations ranges from 100 to 1000 model evaluations. On the other hand, the GLUE method requires a number of simulations on the order of 10 000. It should be remarked that

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Monte Carlo simulations are used just for the purpose of propagating input uncertainty (assumed from expert knowledge) to output uncertainty, however, GLUE method aims to first identify the posterior distribution of parameters (a step which requires many model evaluations in the order of 10 000's) and then propagate this to output uncertainty (this step will be comparable to running a Monte Carlo simulations). Similarly, the stochastic dynamic programming required also on the order of 10 000 simulations. The computationally simplest method appears to be fuzzy linear regression as it involves formulation of linear programming problem with fuzzy inequality constraints for which many effective LP solvers are available.

4.5.2.5 Method accuracy

About the accuracy of the methods, this is difficult to comment on but it is clear that the outcome of an uncertainty analysis depends on how the scenario for the uncertainty analysis is defined, on the framing of the analysis (Sin *et al.*, 2009). This sets the objective (what is the question to be answered) and the boundaries for the analysis, that is which system parameters are included as uncertain, what are the upper and lower ranges selected for each uncertain parameter. In other words, the framing of the analysis reminds the analyst to ask the right question and to set-up the right framework to do so. If the scope of the analysis is set too narrow, the outcome will also be narrow, hence missing out the important implications on the design decisions (the outcome being the right question asked, but the answer is biased). On the other hand, if one sets the scope of the uncertainty analysis too large, then the outcome is likely to be too complex to make sense (as there are too many sources contributing to the decision variable), hence uninformative. While there are still needs for better ways to frame uncertainty analysis, there are already available useful examples on how to setup an uncertainty analysis (see Benedetti *et al.*, 2012; Sin *et al.*, 2009, 2011).

In terms of the purpose of uncertainty analysis, most studies reported in this category aimed at providing decision support for comparing different design alternatives, operation (control) alternatives or technology selection alternatives.

4.6 SUMMARY

4.6.1 Input and parameter uncertainty assessment

- Random errors are characterised using statistical measures of precision such as standard deviation and variance. A common technique for detecting systematic errors in an analytical method or an instrument is to analyse a sample whose composition is accurately known. Another option is to use two independent methods to analyse the same sample as shown by Rieger *et al.* (2005).
- Alternative techniques for detecting and removing systematic errors include multivariate outlier detection methods, statistical process control methods, and data reconciliation. Statistical process control methods and data reconciliation are well suited to on-line applications as they can be easily automated, can simultaneously consider a number of variables, and do not require comparison to a reference method or sample, which may not always be practical. In addition, they can account for auto- and cross-correlation.
- The uncertainty in model parameters is typically assessed as part of parameter estimation. Parameter inference or confidence regions can be developed after parameter estimation based on an approximate variance−covariance matrix for the parameters. This often provides a sufficient approximation of the uncertainty, which can be better approximated using other, more sophisticated techniques.
- Other potentially more powerful techniques include profiling (Bates & Watts, 1988), Monte Carlo analysis, and the use of Bayesian statistics to create a parameter distribution based on prior and

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

current knowledge (Cox, 2004). Combining Monte Carlo simulation with a parameter estimation algorithm is recommended when a more detailed evaluation of parameter uncertainty is required as it is a powerful and reasonably easy to understand method. It has the disadvantage of requiring a considerable number of simulations.

4.6.2 Model structure uncertainty assessment

- It is largely recognised that wastewater treatment models have structural uncertainty, but methods for quantifying this are generally not available or addressed. This may be an area where considerable additional work is required.
- The level of sophistication and the nature of the state variables included in wastewater treatment models require that those that use them understand and address the limitations of the mathematical approaches used in the models. The completely mixed stirred tank reactor uniform concentration assumption is widely employed in wastewater treatment models but its limitations, which may be more pronounced with the nature of state variables representing storage products and with non-first-order rate expressions, are not generally addressed. In this case, however, the work within the chemical engineering field provides tools to consider the structural uncertainty of this model assumption and the limiting cases of maximum mixing and complete segregation may need to be addressed more often by wastewater treatment modellers.
- The wastewater treatment modelling profession should not become complacent with the Monte Carlo approach to quantify uncertainty. The tools and knowledge of fundamental and applied mathematics should be considered. There is some indication in the literature that different methods for model uncertainty quantification generate different uncertainty results and, therefore, additional understanding and work are required to determine what meaningful uncertainty results are and how they are truly achieved. Research should focus on finding the best methodologies for specific cases and types of analyses.
- There is some indication in the literature that inherent random variability, and hence uncertainty, in wastewater treatment processes may arise from the nature of the systems. The complex, non-linear nature of the systems and the numerous potential competitive and cooperative populations in systems modelled by the wastewater professional may result in dynamical, so-called chaotic, behaviour. The tools of this discipline of mathematics may play a useful role in describing some wastewater treatment systems.

4.6.3 Propagation of uncertainty in model-based decision making

- Although detailed modelling is useful for process engineering, the uncertainty in the inputs and the complexity of the models still result in significant uncertainty in the model outputs.
- Monte Carlo type techniques can be used for design, sensitivity analyses, and calibration. How the Monte Carlo techniques are applied and how the results are interpreted has been approached differently within all of the publications reviewed with no consensus on the best approach.
- Some more mathematically advanced techniques have been applied to process engineering. Although these approaches may improve results, it is unlikely that they can easily be adopted by practicing engineers due to their complexity.
- All approaches for 'robust designs' showed that the uncertainty in some inputs, both wastewater characterisation as well as kinetic parameters, can significantly affect the model predicted rate of

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

failure. Therefore, the critical inputs need to be identified, and the variability or uncertainty in their values should be defined and used in the modelling exercise.

This section has reviewed methods that have been used to assess and propagate uncertainty in wastewater treatment analyses. Together, these reviews converge to several overarching conclusions:

(1) There is uncertainty in wastewater treatment process model structure, parameters and inputs. Currently, the profession does not have a comprehensive understanding of the extent, impact or relative importance of these contributions.
(2) The Monte Carlo method is the engineering standard method for uncertainty analysis. It is understood widely and works well. However, the community may benefit from the development of further advanced methods and tools.
(3) Framing of uncertainty analysis by asking the right question and setting up the right framework (boundaries of the analysis) are key to arriving at a meaningful and useful result.
(4) Some successful industrial applications were found, but more case studies are needed to realise the benefits of uncertainty analysis methods.

REFERENCES

Abusam A. and Keesman K. (2002). Sensitivity analysis of the secondary settling tank double-exponential function model. *European Water Management Journal*. Official Publication of the European Water Association (EWA), **2002** (7), 1–9.

Afonso P. M. and da Conceição Cunha M. (2007). Robust optimal design of activated sludge bioreactors. *Journal of Environmental Engineering*, **133**(1), 44–52.

Bard Y. (1974). Nonlinear Parameter Estimation. Academic Press, New York and London. ISBN-13: 978-0120782505.

Barker P. and Dold P. (1995). COD and nitrogen mass balances in activated sludge systems. *Water Research*, **29**(2), 633–643.

Bates D. and Watts D. (1988). Nonlinear Regression Analysis and Its Applications. Wiley New York, New York, NY, USA.

Benedetti L., Bixio D. and Vanrolleghem P. A. (2006). Benchmarking of WWTP design by assessing costs, effluent quality and process variability. *Water Science and Technology*, **54**(10), 95–102.

Benedetti L., Batstone D. J., De Baets B., Nopens I. and Vanrolleghem P. A. (2012). Uncertainty analysis of WWTP control strategies made feasible. *Water Quality Research Journal of Canada*, **47**(1), 14–29.

Berthouex P. (1989) Constructing control charts for wastewater treatment plant operation. *Research Journal of the Water Pollution Control Federation*, **61**(9), 1534–1551.

Bertrand-Krajewski J.-L., Winkler S., Saracevic E., Torres A. and Schaar H. (2007). Comparison of and uncertainties in raw sewage COD measurements by laboratory techniques and field UV-visible spectrometry. *Water Science and Technology*, **56**(11), 17–25.

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.

Box G. and Draper N. (1965). The Bayesian estimation of common parameters from several responses. *Biometrika*, **52** (3–4), 355.

Box G. and Lucas H. (1959). Design of experiments in non-linear situations. *Biometrika*, **46**(1–2), 77.

Box G., Hunter J. and Hunter W. (1978). Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building. John Wiley and Sons Inc., Hoboken, NJ, USA. ISBN-13: 978-0471093152.

Brouwer H., Klapwijk A. and Keesman K. J. (1998). Identification of activated sludge and wastewater characteristics using respirometric batch-experiments. *Water Research*, **32**(4), 1240–1254.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Cameron D. B., Paterson W. R. and Morton W. (1992). Dynamic Data Reconciliation in Process Simulation. Technical Association of the Pulp and Paper Industry (TAPPI) Engineering Conference, Boston, MA, USA.

Checchi N. and Marsili-Libelli S. (2005). Reliability of parameter estimation in respirometric models. *Water Research*, **39**(15), 3686–3696.

Chen H. W. and Chang N. B. (2002). A comparative analysis of methods to represent uncertainty in estimating the cost of constructing wastewater treatment plants. *Journal of Environmental Management*, **65**(4), 383–409.

Chen J., Bandoni A. and Romagnoli J. A. (1997). Robust estimation of measurement error variance/covariance from process sampling data. *Computers & Chemical Engineering*, **21**(6), 593–600.

Cox C. D. (2004). Statistical distributions of uncertainty and variability in activated sludge model parameters. *Water Environment Research*, **76**(7), 2672–2685.

Crowe C. (1996). Data reconciliation – progress and challenges. *Journal of Process Control*, **6**(2–3), 89–98.

Curlin M., Matic M., Matosic M., Mijatovic I. and Kurtanjek Z. (2004). Effects of hydraulic residence time and mixing on wastewater treatment in a membrane bioreactor. *Chemical and Biochemical Engineering Quarterly*, **18**(2), 97–104.

Danckwerts P. (1958). The effect of incomplete mixing on homogeneous reactions. *Chemical Engineering Science*, **8**(1–2), 93–102.

De Pauw D. (2005). Optimal Experimental Design for Calibration of Bioprocess Models: A Validated Software Toolbox. BIOMATH, University of Ghent. Belgium, http://hdl.handle.net/1854/LU-472062 (accessed 2021).

De Pauw D. and Vanrolleghem P. A. (2003). Practical aspects of sensitivity analysis for dynamic models. Proceedings of the 4th IMACS Symposium on Mathematical Modelling (MATHMOD), Vienna, Austria.

Doby T. A., Loughlin D. H., de los Reyes III F. L. and Ducoste J. J. (2002). Optimisation of activated sludge designs using genetic algorithms. *Water Science and Technology*, **45**(6), 187–198.

Dochain D. and Vanrolleghem P. A. (2001). Dynamical Modelling and Estimation in Wastewater Treatment Processes. IWA Publishing, UK, pp. 360. ISBN13: 9781900222501.

Dominguez D. and Gujer W. (2006). Evolution of a wastewater treatment plant challenges traditional design concepts. *Water Research*, **40**(7), 1389–1396.

Draper N. and Smith H. (1989). Applied Regression Analysis, 3rd edn. Wiley, New York, NY, USA.

Duchesne S., Beck M. B. and Reda A. (2001). Ranking stormwater control strategies under uncertainty: the river Cam case study. *Water Science and Technology*, **43**(7), 311–320.

Flores-Alsina X., Rodriguez-Roda I., Sin G and Gernaey K. V. (2008). Multi-criteria evaluation of wastewater treatment plant control strategies under uncertainty. *Water Research*, **42**(17), 4485–4497.

Flores-Alsina X., Rodriguez-Roda I., Sin G and Gernaey K. V. (2009). Uncertainty and sensitivity analysis of control strategies using the benchmark simulation model No1 (BSM1). *Water Science Technology*, **59**(3), 491–499.

Franks S. W. (1999). Disaggregation of environmental factors affecting sewer pipe failures. *Journal of Infrastructure Systems*, **5**(4), 150–158.

Friedler E. and Butler D. (1996). Quantifying the inherent uncertainty in the quantity and quality of domestic wastewater. *Water Science and Technology*, **33**(2), 65–78.

Graham D., Knapp C., Van Vleck E. S., Bloor K., Lane T. B. and Graham C. E. (2007). Experimental demonstration of chaotic instability in biological nitrification. *The ISME Journal*, **1**(5), 385–393.

Griborio A., McCorquodale J., Pitt P., Vinci P. and Wang T. (2007). Uncertainty analysis in wastewater engineering modelling: understanding the limitations. *Proceedings of the Water Environment Federation*, **2007**(17), 1584–1596.

Guay M. and McLean D. (1995). Optimization and sensitivity analysis for multiresponse parameter estimation in systems of ordinary differential equations. *Computers and Chemical Engineering*, **19**(12), 1271–1286.

Gujer W. (2002). Microscopic versus macroscopic biomass models in activated sludge systems. *Water Science and Technology*, **45**(6), 1–11.

Haider S., Svardal K., Vanrolleghem P. A. and Kroiss H. (2003). The effect of low sludge age on wastewater fractionation (SS, SI). *Water Science and Technology*, **47**(11), 203–209.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Huo J., Robinson R. B. and Cox C. D. (2004). Uncertainty and Sensitivity Analysis of Activated Sludge Model No. 1 by Monte Carlo Simulation for Single CSTR with Universal Distribution Parameters. EWRI'04 World Water & Environmental Resources Congress, June 27–July 1, 2004, Salt Lake City, Utah, USA.

Huo J., Jiang Y., Seaver W., Robinson R. B. and Cox C. D. (2006). Statistically based design of wastewater treatment plants (WWTPs) using Monte Carlo simulation of Activated Sludge Model No.1 (ASM1). World Environmental and Water Resources Congress, Omaha, Nebraska, USA, ASCE.

Ibrahim G., Habib H. and Saleh O. (2008). Periodic and chaotic solutions for a model of a bioreactor with cell recycle. *Biochemical Engineering Journal*, **38**(2), 124–137.

Joannis C., Ruban G., Gromaire M. C., Cheboo G. and Bertrand-Krajewski J. L. (2008). Reproducibility and uncertainty of wastewater turbidity measurements. *Water Science and Technology*, **57**(10), 1667–1673.

Kletting M., Rauh A., Aschemann H. and Hofer E. P. (2007). Consistency tests in guaranteed simulation of nonlinear uncertain systems with application to an activated sludge process. *Journal of Computational and Applied Mathematics*, **199**(2), 213–219.

Koch G., Kuhni M. and Siergrist H. (2001). Calibration and validation of an ASM3-based steady-state model for activated sludge systems-part I: prediction of nitrogen removal and sludge production. *Water Research*, **35**(9), 2235–2245.

Kops S. and Vanrolleghem P. A. (1996). An evaluation of methodologies for uncertainty analysis in biological waste water treatment. 9th European Junior Scientist Workshop, 11–14 April, Klive, UK.

Kourti T. and MacGregor J. (1995). Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemometrics and Intelligent Laboratory Systems*, **28**(1), 3–21.

Krasnochtanova I. A. R., Kletting M., Aschemann H., Hofer E. P. and Schoop K. M. (2009). Interval methods as a simulation tool for the dynamics of biological wastewater treatment processes with parameter uncertainties. *Applied Mathematical Modelling*, **34**(3), 744–762.

Lavallee B., Lessard P. and Vanrolleghem P. A. (2005). Modelling the metabolic adaptations of the biomass under rapid growth and starvation conditions in the activated sludge process. *Journal of Environmental Engineering and Science*, **4**(6), 533–548.

Lee T. T., Wang F. Y. and Newell R. B. (1999a). Distributed parameter approach to the dynamics of complex biological processes. *AIChE Journal*, **45**(10), 2245–2261.

Lee T. T., Wang F. Y. and Newell R. B. (1999b). On the modelling and simulation of a BNR activated sludge process based on distributed parameter approach. *Water Science and Technology*, **39**(6), 79–88.

Leis J. and Kramer M. (1988). The simultaneous solution and sensitivity analysis of systems described by ordinary differential equations. *ACM Transactions on Mathematical Software (TOMS)*, **14**(1), 45–60.

Lennox J. and Rosen C. (2002). Adaptive multiscale principal components analysis for online monitoring of wastewater treatment. *Water Science & Technology*, **45**(4), 227–235.

Levenspiel O. (1999). Industrial and Chemistry Engineering Research. ACS Publications, DC, USA.

Makinia J. and Wells S. (2000a). A general model of the activated sludge reactor with dispersive flow – I. Model development and parameter estimation. *Water Research*, **34**(16), 3987–3996.

Makinia J. and Wells S. A. (2000b). A general model of the activated sludge reactor with dispersive flow – II. Model verification and application. *Water Research*, **34**(16), 3997–4006.

Marsili-Libelli S. and Tabani F. (2002). Accuracy analysis of a respirometer for activated sludge dynamic modelling. *Water Research*, **36**(5), 1181–1192.

Meijer S., Van Der Spoel H., Susanti S. and Heijne J. J. (2002). Error diagnostics and data reconciliation for activated sludge modelling using mass balances. *Water Science and Technology*, **45**(6), 145–156.

Moles C. G., Gutierrez G., Alonso A. A. and Banga J. R. (2003). Integrated process design and control via global optimisation: a wastewater treatment plant case study. *Chemical Engineering Research and Design*, **81**(5), 507–517.

Montgomery D. and Mastrangelo C. (1991). Some statistical process control methods for autocorrelated data. *Journal of Quality Technology*, **23**(3), 179–193.

Neumann M. B. and Gujer W. (2008). Underestimation of uncertainty in statistical regression of environmental models: influence of model structure uncertainty. *Environmental Science and Technology*, **42**(11), 4037–4043.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

Nowak O., Franz A., Svardal K., Müller V. and Kühn V. (1999). Parameter estimation for activated sludge models with the help of mass balances. *Water Science and Technology*, **39**(4), 113–120.

Omlin M. and Reichert P. (1999). A comparison of techniques for the estimation of model prediction uncertainty. *Ecological Modelling*, **115**(1), 45–59.

Petersen B. (2000). Calibration, identifiability and optimal experimental design of activated sludge models. PhD thesis, Department of Applied Mathematics, Biometrics and Process Control (BIOMATH) University of Ghent, Belgium.

Petersen B., Gernaey K. and Vanrolleghem P. A. (2000). Improved theoretical identifiability of model parameters by combined respirometric–titrimetric measurements. A generalisation of results. Proceedings IMACS 3rd Symposium on Mathematical Modelling, Vienna, Austria.

Portielje R., Hvitved-Jacobsen T. and Schaarup-Jensen K. (2000). Risk analysis using stochastic reliability methods applied to two cases of deterministic water quality models. *Water Research*, **34**(1), 153–170.

Puig S., van Loosdrecht M., Colprim J. and Meijer S. C. F. (2008). Data evaluation of full-scale wastewater treatment plants by mass balance. *Water Research*, **42**(18), 4645–4655.

Rauh A., Kletting M., Aschemann H. and Hofer E. P. (2004). Application of interval arithmetic simulation techniques to wastewater treatment processes. Proceedings of the IASTED International Conference on Modelling, Identification, and Control (MIC 2004), Grindelwald, Switzerland, pp. 287–293. https://www.actapress.com/Content_of_Proceeding.aspx?proceedingID=247 (accessed 2021).

Rauh A., Kletting M., Aschemann H. and Hofer E. P. (2007). Reduction of overestimation in interval arithmetic simulation of biological wastewater treatment processes. *Journal of Computational and Applied Mathematics*, **199**(2), 207–212.

Rawlings J. and Ekerdt J. (2002). Chemical Reactor Analysis and Design Fundamentals. Publisher: Nob Hill Publishing, Madison, Wisconsin, USA.

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

Rieger L., Takacs I., Villez K., Siegrist H., Lessard P., Vanrolleghem P. A. and Comeau Y. (2010). Data reconciliation for wastewater treatment plant simulation studies planning for high-quality data and typical sources of errors. *Water Environment Research*, **82**(5), 426–433.

Rieger L., Gillot S., Langergraber G., Ohtsuki T., Shaw A., Takacs I. and Winkler S. (2013). Guidelines for Using Activated Sludge Models, Scientific and Technical Report No. 22, International Water Association (IWA), London, UK.

Robinson R. B., Cox C. D. and Odom K. (2005). Identifying outliers in correlated water quality data. *Journal of Environmental Engineering*, **131**(4), 651–657.

Romagnoli J. and Sanchez M. (2000). Data Processing and Reconciliation for Chemical Process Operations. Academic Press, San Diego, CA, USA.

Rousseau D., Verdonck F., Moerman O., Carrrette R., Thoeye C., Meirlaen J. and Vanrolleghem P. A. (2001). Development of a risk assessment based technique for design/retrofitting of WWTPs. *Water Science and Technology*, **43**(7), 287–294.

Saikaly P. and Oerther D. (2004). Bacterial competition in activated sludge: theoretical analysis of varying solids retention times on diversity. *Microbial Ecology*, **48**(2), 274–284.

Schuler A. (2005). Diversity matters: dynamic simulation of distributed bacterial states in suspended growth biological wastewater treatment systems. *Biotechnology and Bioengineering*, **91**(1), 62–74.

Schuler A. (2006). Process hydraulics, distributed bacterial states, and biological phosphorus removal from wastewater. *Biotechnology and Bioengineering*, **94**(5), 909–920.

Sin G. (2004). Systematic calibration of activated sludge models. PhD thesis, Department of Applied Mathematics, Biometrics and Process Control (BIOMATH), University of Ghent, Belgium.

Sin G. and Vanrolleghem P. A. (2006). Evolution of an ASM2s-like model structure due to operational changes of an SBR process. *Water Science and Technology*, **53**(12), 237–245.

Sin G., De Pauw D. J. W., Weijers S. and Vanrolleghem P. A. (2008). An efficient approach to automate the manual trial and error calibration of activated sludge models. *Biotechnology and Bioengineering*, **100**(3), 516–528.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest

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: prioritising sources of uncertainty. *Water Research*, **45**(2), 639–651.

Skoog D. A., West D. M. and Holler F. J. (1995). Fundamentals of Analytical Chemistry, 7th edn. Saunders College Publishing, USA. ISBN-13: 978-0-03-005938-4.

Standard Methods for the Examination of Water and Wastewater. (1998). 20th edn. American Public Health Association/American Water Works Association/Water Environment Federation, Washington DC, USA.

Stroot P., Saikaly P. and Oerther D. B. (2005). Dynamic growth rates of microbial populations in activated sludge systems. *Journal of Environmental Engineering*, **131**, 1698.

Tansel B. (1999). Propagation of parameter uncertainties to system dependability. *Civil Engineering and Environmental Systems*, **16**(1), 19–35.

Tenno R. and Uronen P. (1995). State and parameter estimation for wastewater treatment processes using a stochastic model. *Control Engineering Practice*, **3**(6), 793–804.

Thomann M. (2008). Quality evaluation methods for wastewater treatment plant data. *Water Science and Technology*, **57**(10), 1601–1609.

Thomann M., Rieger L., Frommhold S. and Siegrist H. (2002). An efficient monitoring concept with control charts for on-line sensors. *Water Science and Technology*, **46**(4–5), 107–116.

Tomita R., Park S. and Sotomayor O. (2002). Analysis of activated sludge process using multivariate statistical tools – a PCA approach. *Chemical Engineering Journal*, **90**(3), 283–290.

Tsai J. C. C., Chen V. C. P., Beck M. B. and Chen J. (2004). Stochastic dynamic programming formulation for a wastewater treatment decision-making framework. *Annals of Operations Research*, **132**(1–4), 207–221.

Van der Heijden R., Heijnen J., Hellinga C., Romein B. and Luyben K. C. (1994a). Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. *Biotechnology and Bioengineering*, **43**(1), 3–10.

Van der Heijden R., Heijnen J., Hellinga C., Romein B. and Luyben K. C. (1994b). Linear constraint relations in biochemical reaction systems: II. Diagnosis and estimation of gross errors. *Biotechnology and Bioengineering*, **43**(1), 11–20.

Van der Heijden R., Heijnen J., Hellinga C., Romein B. and Luyben K. C. (1994c). Linear constrain relations in biochemical reaction systems III. Sequential application of data reconciliation for sensitive detection of systematic errors. *Biotechnology and Bioengineering*, **44**(7), 781–791.

Vanrolleghem P. A. and Keesman K. J. (1996). Identification of biodegradation models under model and data uncertainty. Proceedings of the 1995 IAWQ Interdisciplinary Symposium on Uncertainty, Risk and Transient Pollution Events, 26–28 July 1995, Pergamon Press Inc., Exeter, UK.

Vanrolleghem P. A., Van Daele M. and Dochain D. (1995). Practical identifiability of a biokinetic model of activated sludge respiration. *Water Research*, **29**(11), 2561–2570.

Vanrolleghem P. A., Insel G., Petersen B., Sin G., Nopens I., Dovermann H., Weijers S. and Gernaey K. (2003). A comprehensive model calibration procedure for activated sludge models. Proceedings of the Water Environment Federation, Alexandria, VA, USA. Water Environment Federation 76th Annual Technical Exhibition and Conference (WEFTEC) 2003, Los Angeles, CA, USA.

Von Sperling M. (1993). Parameter estimation and sensitivity analysis of an activated sludge model using Monte Carlo simulation and the analyst's involvement. Proceedings of the 6th JAWO Workshop on Instrumentation, Control and Automation of Water and Wastewater Treatment and Transport Systems, 17–25 June 1993, Banff, Canada, Published by Pergamon Press Inc.

Zhang Y. and Henson M. (2001). Bifurcation analysis of continuous biochemical reactor models. *Biotechnology Progress*, **17**(4), 647–660.

Zwietering T. (1959). The degree of mixing in continuous flow systems. *Chemical Engineering Science*, **11**(1), 1–15.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/1117956/9781780401034_0047.pdf
by guest