Chapter 3

Incorporating uncertainty analysis into model-based decision making – opportunities and challenges

3.1 INTRODUCTION

There is increasing consensus among water reclamation professionals that the predictive power of models is a critical component of plant design and operation. However, models and simulators used by designers and operators are yet to systematically incorporate methods for the evaluation of the uncertainty associated with design and operations. Such approaches have the opportunity to assess the upset resilience of individual processed and the system as a whole.

This chapter presents a general discussion on the implications of uncertainty when using process models for design and optimisation of treatment plants. The discussion highlights both the potential opportunities for explicit accounting of uncertainty as well as open questions that need to be dealt with before such approaches can be integrated into daily design practice.

3.2 INCORPORATION OF SAFETY IN CURRENT MODEL-ASSISTED DESIGN

Common practice for the design of a treatment plant is to use design guidelines. In some cases, the designer/engineer will then run a steady state or dynamic simulation with a process-based biokinetic model and if the predicted effluent concentration is (well) below the effluent requirements, then the design will be judged as appropriate.

In some cases, when a model is directly used to obtain a design, uncertainty and variability are accounted for in several ways as discussed in Chapter 2:

- Use of a higher influent load than the design load;
- Design to more stringent effluent requirements (e.g., to 70% of the effluent limit concentration);
- Choice of conservative values for process parameters (e.g., reduced maximum growth rates, increased sludge volume index);

Editor(s): Evangelia Belia, Lorenzo Benedetti, Bruce Johnson, Sudhir Murthy, Marc Neumann, Peter Vanrolleghem and Stefan Weijers
doi: 10.2166/9781780401034_0033

- Increase of obtained design variables (e.g., multiply a resulting design volume with a safety factor of 1.5);
- A combination of the above.

An alternative would be to use a model-assisted approach where the designer/engineer tries to directly quantify the uncertainty associated with the different model factors (parameters and inputs). Uncertainty propagation can then be used to quantify how the plant performance predicted by the model is affected. Design variables (such as tank volume) can then be iteratively modified until the probability of compliance reaches a specific value, for example, 95% or 99%. This probability ties into the resilience of the process and the overall treatment system.

3.3 OPPORTUNITIES OF EXPLICITLY CONSIDERING UNCERTAINTY AND VARIABILITY

One can view the use of safety factors in a design guideline as an implicit way of dealing with uncertainty. By adding a margin of safety, various sources of uncertainty are accounted for simultaneously. The advantage of this approach lies in its simplicity. Historically developed safety factor approaches have withstood the test of time and are often widely accepted as industry standards. However, guidelines are not available for all configurations or new technologies and thus may limit innovation. This is especially true when highly integrated systems are being conceived (water resource recovery facilities (WRRFs) with many feedback loops due to sophisticated control or interconnected systems, e.g., sewer-WRRF-river). Recent developments in simulation-assisted approaches means that the behaviour of such complex systems can be described.

The challenge when using a process-based model is how to appropriately account for uncertainty, that is, how to appropriately translate the safety factor approach of a guideline. The designer/engineer may explicitly acknowledge the uncertainties present in the modelling exercise in different ways. One possibility is to document how each uncertain input value was determined, to provide the rationale for the decisions made for each of the uncertain inputs. Another possibility would be to express the uncertainty by using quantitative probability distributions for model inputs and model parameters.

The main practical advantage of explicit approaches is increased transparency regarding uncertainties (Flyvbjerg *et al.*, 2003). It is expected that such methods can be refined by conducting post-project audits during which the assumptions made in the design phase are compared with the performance of the built plant and the causes of any discrepancies are identified.

3.4 SCOPE AND LIMITATIONS OF MODELS

Models by definition are not exact replicas of real-world systems. The necessary simplifications inevitably lead to the introduction of uncertainty. When using a model, the model scope needs to be considered. For example:

- Which phenomena were included, and which ones were not included?
- What is the range of values for model inputs and parameters for which the model is expected to give adequate results?

3.4.1 Evolution of wastewater treatment modelling

The scientific development of activated sludge models (ASMs) reflects the improvements in the understanding of the fundamental microbiological transformation processes occurring in biological wastewater treatment and the current ASMs are widely applied in engineering practice as state-of-the-art models and used to predict plant performance.

3.4.2 Desirability criteria for models

The following features are often used to judge the desirability of a model (e.g., Reichert, 2009):

- **Causality**: The model represents the relevant cause−effect relationship for the system response at the required level of resolution.
- **Universality**: The model structure and as far as possible the model parameter values should be transferable to a similar system.
- **Predictive capability**: The model should remain valid for the extrapolation of external influence factors to ranges required for predictive use.
- **Identifiability**: Unknown parameter values should be identifiable with the available data. This means that a model can be fitted (i.e., parameters can be identified) when applying a fitting (optimisation) algorithm. When this is not the case, prior knowledge of non-identifiable parameters must be available.
- **Simplicity**: The model should be as simple as possible.

With respect to the five desirability criteria listed above, an attempt to classify the importance of these features to ASMs is: causality (high), universality (medium−high), predictive capability (medium−high), identifiability (see Box 3.1) (low), simplicity (medium). The strength of the ASM suite lies in characterising the microbial processes. However, when trying to emulate full-scale systems, other factors such as hydrodynamics, mass-transfer (such as oxygen transfer during aeration), varying sludge settling characteristics, precipitation chemistry, sensors and actuators, equipment reliability may become equally important or even more important. The following example shows the relationship between WRRF model limitations and sources of uncertainty.

3.4.3 Example of wastewater treatment plant model limitations

Consider the following scenario: An engineer is given all the details of the configuration of a plant and is then asked to model the (expected) effluent time series of the plant using one year of measured influent data (Figure 3.1). For this example, it is assumed that the measurements do not contain errors and are representative of the true values.

Assuming that the model is representative of the behaviour of the full-scale plant, it is reasonable to expect the modelled and the measured effluent time series to be similar. However, if the plant effluent data include an effluent limit exceedance due to a toxic spill from an industrial source, resulting in inhibition of microorganisms, this will not typically be reflected in the modelled time series, as such processes are not captured in the ASM-based models.

Figure 3.1 Influent, effluent and system state as functions of time. The effluent can be modelled using the measured influent and a model representation of the system state.

The comparison of the predicted (modelled) and the measured effluent time series demonstrates what the model does and does not account for. It identifies the processes that are not included in the model domain.

This simple example demonstrates the challenge faced when using standard ASMs for the prediction of WRRF behaviour. Because models do not **perfectly** emulate real-world plant behaviour, there exist processes and events, such as equipment failures, that impact plant performance but are not captured in the models.

It is therefore essential for the designer/engineer to determine which processes are crucial for each model-based design project. If additional processes or elements (e.g., equipment models) need to be included, the models need to be expanded (e.g., Rosen *et al.*, 2008). However, even in cases where model expansion is required but not possible, models can be used very effectively to compare alternative configurations with the same assumptions.

3.5 WHAT DON'T WE KNOW ABOUT DEALING WITH UNCERTAINTY?

Although there are many ways to include explicit descriptions of uncertainty in model-assisted designs, there remain several challenges. A selection of challenges that the authors believe to be relevant are listed in the following sections.

3.5.1 How conservative are we with the safety factor approach?

To evaluate how conservative the safety factor approach is, we need to answer the question: what do safety factors account for? This question is difficult to answer. On the one hand, safety factors have been developed over time, and the rationale and the data that were used to determine their values are not always known. On the other, it is very difficult to assess the actual quality of a design: ideally one needs the designed plant running at design conditions to determine if it is over- or under-designed. However, by the time that the design conditions are reached the plant has typically undergone substantial changes compared to the start-up configuration.

Deviations from the original planning assumptions could quite easily be detected in post-project audits: for example, number of person equivalents connected to a WRRF. Deviations from design assumptions, for example, critical SVI or growth rates could be detected by long-term monitoring of such parameters. Therefore, long-term post-project audits might be a valuable tool to improve design approaches. Long-term monitoring of influent and effluent quality could provide information for developing more robust design procedures (e.g., Bott & Parker, 2011).

Also, inter-guideline comparisons and the comparison of guidelines with ASM-based process models (Corominas *et al.*, 2010) can help quantify margins of safety inherent in design approaches and enable the exploration of aggregated safety factors.

3.5.2 How to move from guidelines with the safety factor approach to probabilistic model-assisted design?

Often biokinetic models are used to verify designs obtained from guidelines or standards. It is not common practice to design a facility solely with the use of a biokinetic model especially for greenfield plants. There are several reasons for this. In some cases, the engineer is legally protected if he/she uses a 'recognised' design guideline or standard. Model-assisted design may transfer liability to the engineer. Even if there are no liability issues with respect to model-based design, an engineer will need to decide where to add safety and how much to add. In many cases, biokinetic models are much easier to employ for plant upgrades where prior knowledge exists on influent and effluent characteristics, process operations, or verifiable model parameters. In these cases, a probabilistic model-assisted design can more easily be used as a primary approach for design, than in the case of a green field project. Under such conditions innovative technologies that are not already included in guidelines or standards can be incorporated.

3.5.3 Determination of prior uncertainty ranges

Two major issues to address within a probabilistic framework include which elements to select as uncertain and, secondly, how to quantify the uncertainty surrounding them. It is common practice for prior uncertainty ranges to be obtained from experts who have experience in determining typical values. Many models and simulation software also provide helpful prior uncertainty ranges.

3.5.4 Parameter (uncertainty) estimation in systems with poor identifiability

Available data, together with models can be used to estimate parameters and their uncertainty by using parameter estimation techniques. However, poor identifiability (see Box 3.1) of ASM-based models remains an issue (see also Box 3.2).

BOX 3.1 PARAMETER IDENTIFIABILITY

Parameters are identifiable when model fitting (optimisation) algorithms are able to find best estimates for the model parameters. For ASM-type models this is usually not the case due to the high number of parameters which one aims to identify and due to the lack of sufficient data.

BOX 3.2 FREQUENTIST VS. BAYESIAN PARAMETER ESTIMATION

In **Frequentist** parameter estimation, the models are confronted with observed data to obtain parameter estimates. Hereby, random observation errors in the data are mapped to uncertainty about the parameters. As wastewater treatment plant models are typically over-parameterised, optimisation (fitting) algorithms cannot find a unique solution due to compensation (technically termed non-identifiability, see Box 3.1). For example, increasing the value of one parameter can be compensated by decreasing the value of another one. Therefore, most parameters are set to default values (taken from literature) and only a few are estimated with statistical techniques. This leads to biased estimates (i.e., dependent on where the other parameters were fixed).

An alternative is to use a **Bayesian** framework where information from literature and expert knowledge can be combined to define prior value ranges for all the parameters. Then, the probabilistic model is confronted with the data and the parameter ranges are updated, typically narrowed. This framework does not require identifiability of the parameters. However, this framework requires the elicitation of ‘inter-subjective’ (i.e., experts must agree) prior parameter ranges. Also, in this framework, the parameter updating procedure is computationally expensive which may still be a limiting factor for dynamic WRRF models.

Often, the data does not provide enough information to enable a statistical estimation of uncertain parameters. As a result, engineering practice often fixes most of the model parameters at default values and just a few are estimated/calibrated. As a consequence, it is important that the estimated parameter values as well as their uncertainty estimates be assessed in an informed manner. One approach for obtaining uncertainty ranges is by extracting probability distributions for an influent characteristic or model parameter using performance data at brownfield locations (Alikhani *et al.*, 2017; Sharifi *et al.*, 2014).

3.5.5 How to adequately deal with biokinetic model structure uncertainty?

Uncertainty about the biokinetic model structure remains a difficult issue. Conducting statistical inference in the presence of model structure error, leads to biased estimates of the parameter values and unreliable uncertainty assessment (Neumann & Gujer, 2008; Villez *et al.*, 2020). This is especially critical for practising engineers as they cannot be expected to modify predefined bio-kinetic model structures (e.g., ASM). Nevertheless, several model structures are now available in most commercial simulators. In engineering projects, model structure selection often depends on the key process being simulated or the effluent parameter(s) associated with a permit limit.

3.5.6 Full-fledged probabilistic model-based design

When considering the replacement of a design guideline with a probabilistic WRRF simulator it is important to identify which real-world phenomena are included. If the model is expected to be a true emulator of the future WRRF, then current models would need to be significantly enhanced with models that account for operational behaviour of sensors, actuators and other equipment. They would need to be able to re-produce toxic spill events, inhibition events, bulking and foaming events, and operator failures, among other things. If these aspects are not covered by the simulator, alternative ways need to be found to take them into account. Such aspects should be clarified by modellers and process engineers by adding a disclaimer on which kind of processes are included in the model and which ones are not.

3.6 HOW CAN WE CURRENTLY ACCOUNT FOR VARIABILITY AND UNCERTAINTY?

3.6.1 Accounting for variability

Whereas accounting for temporal variability is the central aspect of dynamic modelling, spatial variability has until recently only been coarsely resolved using compartmental models such as tanks-in-series.

3.6.1.1 Temporal variability

Accounting for temporal variability include the use of probability distributions, dynamic modelling (multivariate) time-series analysis, and influent generators.

Probability distributions can be used to characterise the variability of dynamic variables such as flows, concentration or loads. This approach is also useful when using a steady-state solution of the model: for example, when describing average monthly behaviour, the influent concentrations and flows can be sampled from the probability distributions to capture meaningful scenarios (Bixio *et al.*, 2002; Mc Cormick *et al.*, 2007).

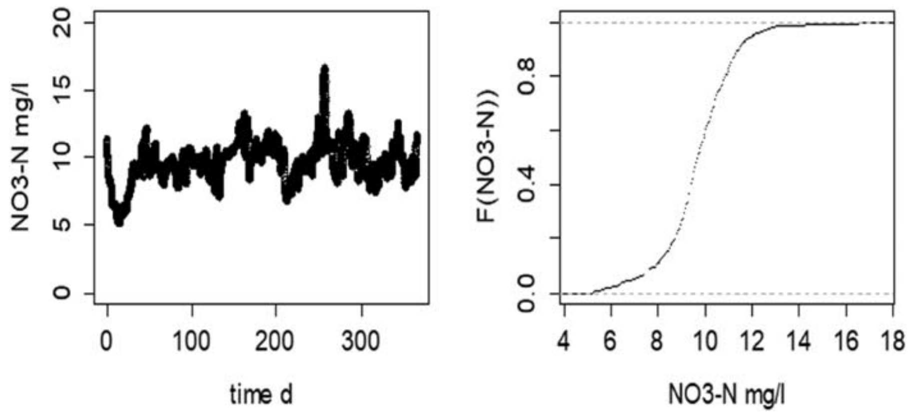

Figure 3.2 365 days of NO_3-N effluent concentration data (sampling interval: 1.2 hours). Left: Time series. Right: The corresponding empirical cumulative density distribution.

In addition, empirical cumulative distributions are often used to characterise plant performance. They condense the information contained in a time series and can, for example, extract the frequency of exceedance of effluent concentration limits. An advantage of directly using time-series analysis over distributions, is that temporal dependence (auto-correlation) is appropriately and explicitly accounted for. A typical time-series analysis identifies:

- trends;
- periodic phenomena;
- autocorrelation.

An example is given in Figure 3.2 for one year of nitrate effluent data with the original time series in the left panel and the corresponding cumulative distribution in the right panel. The y-axis in the cumulative distribution quantifies the percentage of time that the concentration is below the value on the x-axis.

To account for correlation between variables (cross-correlation) the same procedure can be followed with multivariate techniques. Dynamic simulators capture how dynamic influents affect the state variables of the system and predict a dynamic effluent profile from which desired statistics can be extracted.

If synthetic time series are required that represent future load scenarios, then influent generators can be used (Gernaey *et al.*, 2011; Martin & Vanrolleghem, 2014). Influent generators are typically either based on (stochastic) catchment models or are derived from black-box models that are calibrated with historic time series.

3.6.1.2 Spatial variability

Concerning the description of spatial variability, the rapidly growing computational fluid dynamics (CFD) field enables the investigation of spatial phenomena at high resolution (e.g., Alvarado *et al.*, 2013; Gresch *et al.*, 2011; Rehman *et al.*, 2017). Such analyses are critical for multiphase systems (e.g., settling), systems with spatial heterogeneity (e.g., anaerobic digestion) or systems that need to guarantee a certain contact time (e.g., disinfection). To decrease the computational burden, methods have been developed that enable the translation of a CFD model to a compartment model (Gresch *et al.*, 2009).

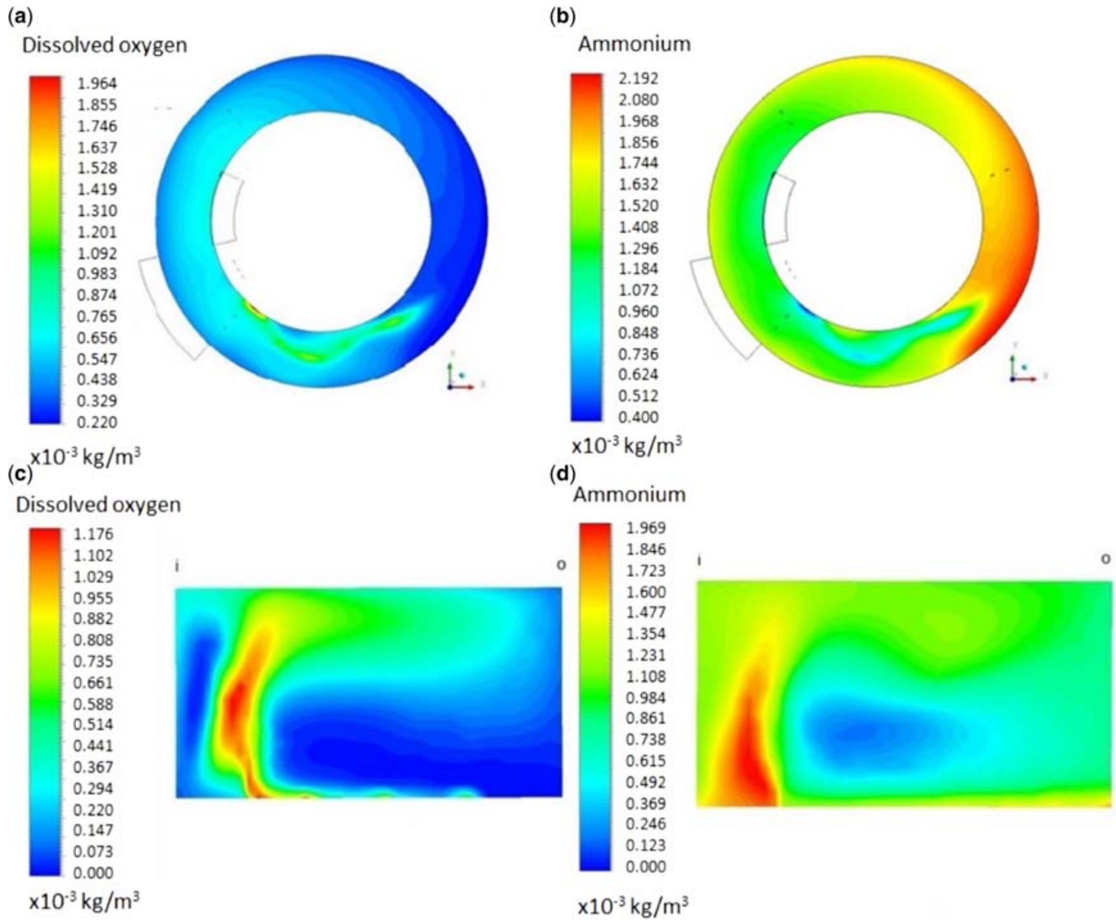

Figure 3.3 2-D concentration profiles in a reactor (Rehman *et al.*, 2017).

An example of a typical CFD model output illustrating the distribution of mean residence time in a reactor is shown in Figure 3.3.

3.6.2 Accounting for uncertainty

3.6.2.1 Uncertainty related to design scenarios

When planning a treatment plant, significant uncertainty is associated with defining the appropriate design loads. This uncertainty can be accounted for by applying foresighting tools such as scenario analysis techniques. These techniques enable multiple possibilities of future loads or other requirements to be accounted for (e.g., Dominguez *et al.*, 2009). They often involve expert interviews and participatory methods. A systematic use of such techniques within the water resources field is not yet widespread, but is increasingly an imperative, especially in cities where urbanisation is rapid or, in service areas where changes to wastewater characteristics are anticipated due to reductions in infiltration and inflow within sewers, due to changes to septage management practices or due to industrial development within the service area.

3.6.2.2 Uncertainty related to data

Measurements contain uncertainty due to both random errors, systematic and gross errors.

Random errors are a consequence of the many uncontrollable and unpredictable errors that exist in the measurement process. They are the effect of many small errors added together. Random errors are indeterminate and can be potentially minimised but never completely removed. They arise in any measurement process and can only be reduced by improving the precision of the measurement.

Systematic errors are non-random errors caused by miss-calibration or malfunction of instruments or the improper location or method for manual or automated sampling. Calibration errors can be reduced through prevention (regular calibration of instrumentation) and partly through data analysis, for example, mass balancing or fault detection (e.g., Lee *et al.*, 2004). Sampling errors can be reduced through better knowledge of the underlying process and increasing of sampling frequency (e.g., Ort *et al.*, 2010). Systematic errors are determinate and can be detected and removed thereby reducing the uncertainty in the measured model inputs. They may be occasional errors or persistent errors.

Gross errors include human oversight and other mistakes while reading, recording, and reading measurements. The most common errors, human errors in the measurement, fall under this category. They can be reduced by the adoption of quality control procedures.

3.6.2.3 Uncertainty related to process modelling

Uncertainty in process modelling arises due to parameter uncertainty (which values to use?), model structure (which model to select?) and errors in implementation.

Parameter uncertainty can be addressed by assigning probability distributions to parameters. In applications where no data are available, a priori uncertainty estimates are obtained from expert knowledge and/or literature. The effects of parameter uncertainty on model outputs can be quantified by the use of Monte Carlo (MC) simulation techniques (Benedetti *et al.*, 2008; Sin *et al.*, 2009).

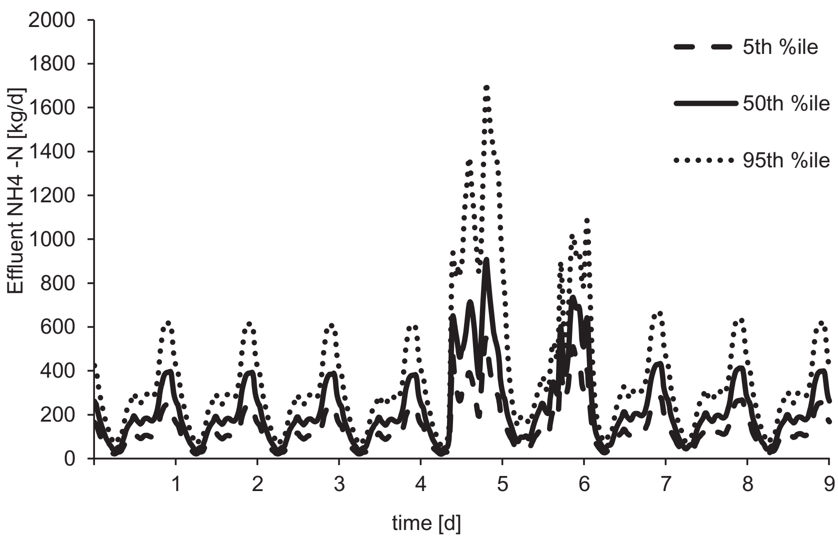

Figure 3.4 Output from an MC simulation with time series for 5th, 50th and 95th percentile values.

Figure 3.4 shows a possible output from such an MC simulation. The plant effluent dynamics are depicted by a range of time series representing percentile values.

In applications where data are available, some parameter values and their uncertainty range can be estimated with the approaches described in Section 3.5.4.

For the practitioner, model structure uncertainty can be addressed in various ways. For example, he/she may want to repeat the modelling exercise with a different model structure or integrate his/her own model structure extensions or reductions (Rieger *et al.*, 2013).

Uncertainty due to modelling errors can be checked by running redundancy checks and elemental balances (e.g., Hauduc *et al.*, 2010). Uncertainty due to software errors can be checked by running a verified model on multiple simulators (reference). Uncertainty due to numerical errors can be captured through the use of multiple simulators, numerical accuracy can be checked by changing the solver properties (such as time step size or solver type and accuracy).

3.6.3 Sensitivity analysis

Typically, a sensitivity analysis is required to prioritise the sources of variability and uncertainty. 'Local' methods analyse how variation in one of the parameters affects the model output while all other parameters are held at the nominal values. 'Global' methods analyse how variation in one parameter affects the model output while all the other parameters are also varying within their uncertainty ranges. Such global methods are quite easy to implement, although some require many simulations (e.g., Benedetti *et al.*, 2011; Neumann, 2012; Sin *et al.*, 2011). An example of such a global analysis is depicted in Figure 3.5.

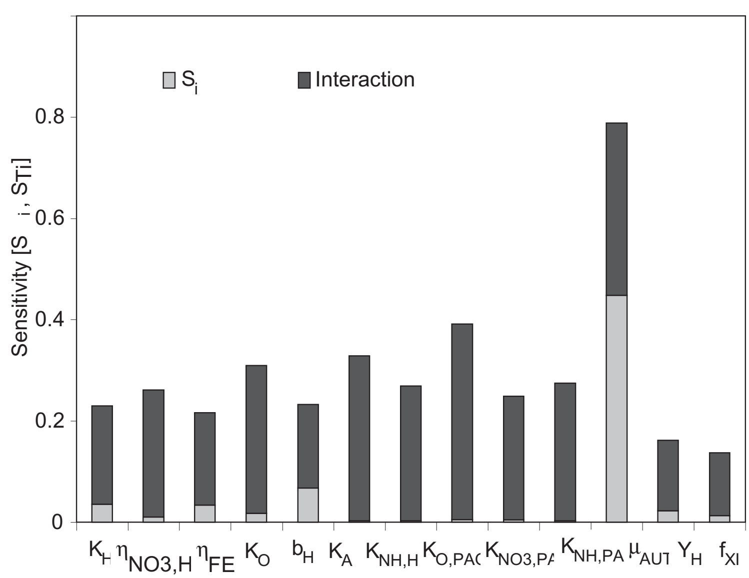

Figure 3.5 Example of a global sensitivity analysis for a membrane bioreactor. Grey section: direct impact of the parameter on the model output uncertainty. Black section: indirect impact of a parameter due to its interaction with all other parameters (Cosenza *et al.*, 2011).

The grey section of the bars quantifies the direct influence of the parameter in determining model output uncertainty (as a fraction of model output variance) and the black section quantifies the interaction effect, which is the indirect influence of a parameter due to its interaction with all the other parameters (Cosenza *et al.*, 2011).

3.7 OPPORTUNITIES OF COMBINING MODELS WITH UNCERTAINTY – EXAMPLE

Figure 3.6 illustrates a typical output of a design example where a mathematical model in combination with uncertainty analysis has been implemented. The x-axis represents concentration and y-axis represents costs (diagonal lines) or probability density.

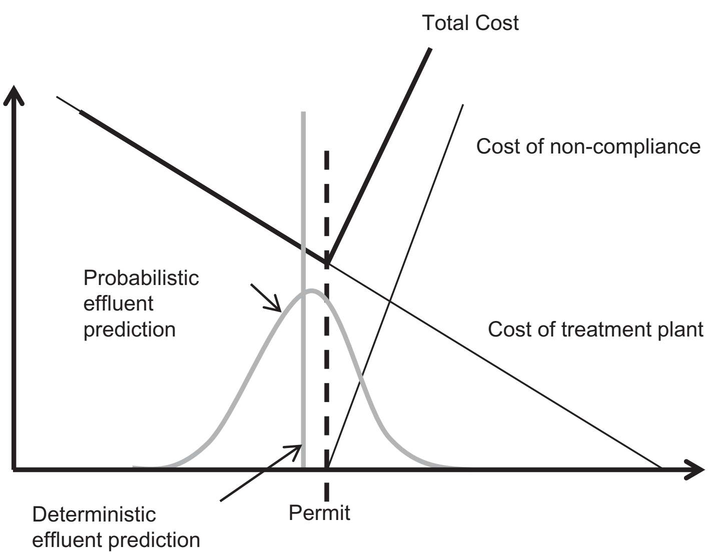

Figure 3.6 Probabilistic design: An optimal design is found by combining probabilistic model predictions with cost functions. X-axis represents concentration and y-axis represents costs (diagonal lines) or probability density in the case of the probabilistic concentration prediction (grey curve). The dashed black vertical line represents the permit limit and the grey vertical line represents the effluent prediction of the deterministic design. Through iteration of the design variables (e.g., tank volume) a design can be found that locates the grey curve in such a way that the expected total cost is minimised.

In this hypothetical example, strict effluent concentration limits (e.g., for maximum month) need to be met. A designer/engineer using a deterministic model will predict a single value for the effluent concentration (vertical grey line in Figure 3.6).

If, however, the engineer expresses her uncertainty about the model parameters with probability distributions and runs an MC simulation, the simulated effluent concentration will become a probability distribution (grey curve in Figure 3.6). Given the permit, the engineer can now either design to a chosen failure probability (e.g., probability of failure = 0.05) or, if cost functions for the treatment plant (capital, operating costs and as non-compliance costs) are available, determine a design that minimises the expected total cost. In Figure 3.6, proposing a smaller design (e.g., smaller tank volumes) would move the entire probability distribution to the right and then the costs due to non-compliance would increase rapidly. Proposing a larger design would move the distribution to the left and then costs would increase due to higher construction- and capital costs. Presumably, the optimal design is the one that leads to a probability distribution that minimises total expected costs, that is, the design with $\int_{-\infty}^{\infty} \{f(predicted\,concentration) \cdot (total\,cost)\} = min$ (see also Reckhow, 1994). This can be seen as an illustrative example of a rational design approach that explicitly deals with uncertainty.

3.8 SUMMARY

Current design approaches still rely heavily on guidelines and on the use of safety factors to account for uncertainty. At the same time, simulators using mechanistic models that capture the details of hydraulic and biochemical dynamics have become common tools in the wastewater engineer's toolbox. If such models are used in design, then uncertainty is typically accounted for in an implicit way, such as designing to stricter standards than those specified. However, these models do also offer the opportunity for explicit considerations of variability and uncertainty. On the one hand, spatial and temporal variability can be examined at higher resolution through CFD and dynamic models, respectively. On the other, (statistical) techniques can be applied that make possible the consideration of measurement uncertainty, parameter uncertainty and model structure uncertainty. This opens up the possibility of moving towards full probabilistic and risk-based designs. At the same time this requires that the limits of predictability are better appreciated by clarifying which real-world phenomena are captured by the models and which are not.

REFERENCES

Alikhani J., Takacs I., Al-Omari A., Murthy S. and Massoudieh A. (2017). Evaluation of the information content of long-term wastewater characteristics data in relation to activated sludge model parameters. *Water Science and Technology*, **75**(6), 1370–1389.

Alvarado A., Vesvikar M., Cisneros J. F., Maere T., Goethals P. and Nopens I. (2013). CFD study to determine the optimal configuration of aerators in a full-scale waste stabilisation pond. *Water Research*, **47**(13), 4528–4537.

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

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

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

Bott C. B. and Parker D. S. (2011). Nutrient Management Volume II: Removal Technology Performance & Reliability, WERF Research Project under Nutrient Challenge, WERF Report NUTR1R06k, Water Environment Research Foundation, Alexandria, VA, USA.

Corominas L., Flores-Alsina X., Muschalla D., Neumann M. B. and Vanrolleghem P. A. (2010). Verification of WWTP design guidelines with activated sludge process models. Proceedings of the 83rd Annual Water Environment Federation Technical Exhibition and Conference (WEFTEC 2010), 4–6 October, New Orleans, Louisiana, USA, pp. 137–146.

Cosenza A., Mannina G., Vanrolleghem P. A., Viviani G. and Neumann M. (2011). Global sensitivity analysis in ASM applications: comparison of the SRC and Extended-FAST method for a UCT-MBR model. Proceedings of the 8th IWA Symposium on Systems Analysis and Integrated Assessment, Watermatex, 2011, San Sebastian, Spain, pp. 543–552.

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

Flyvbjerg B., Bruzelius N. and Rothengatter W. (2003). Megaprojects and Risk: An Anatomy of Ambition. Cambridge University Press., UK, ISBN-13: 978-0521009461.

Gernaey K., Flores-Alsina X., Rosen C., Benedetti L. and Jeppsson U. (2011). Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. *Environmental Modelling & Software*, **26**(11), 1255–1267.

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

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

Hauduc H., Rieger L., Takács I., Héduit A., Vanrolleghem P. A. and Gillot S. (2010). A systematic approach for model verification: application on seven published activated sludge models. *Water Science and Technology*, **61**(4), 825–839.

Lee J. M., Yoo C. K., Choi S. W., Vanrolleghem P. A. and Lee I. B. (2004). Nonlinear process monitoring using kernel principal component analysis. *Chemical Engineering Science*, **59**(1), 223–234.

Martin C. and Vanrolleghem P. A. (2014). Analysing, completing, and generating influent data for WWTP modelling: a critical review. *Environmental Modelling & Software*, **60**, 188–201.

Mc Cormick J. F., Johnson B. and Turner A. (2007). Analyzing risk in wastewater process design: Using Monte Carlo Simulation to move beyond conventional design methods. In Proceedings WEFTEC2007 pp. 1507–1521, San Diego, California, USA.

Neumann M. B. (2012). Comparison of sensitivity analysis methods for pollutant degradation modelling: a case study from drinking water treatment. *Science of the Total Environment*, **433**, 530–537.

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

Ort C., Lawrence M. G., Reungoat J. and Mueller J. F. (2010). Sampling for PPCPs in wastewater systems: comparison of different sampling modes and optimisation strategies. *Environmental Science and Technology*, **44**(16), 6289–6296.

Rehman U., Audenaert W., Amerlinck Y., Maere T., Arnaldos M. and Nopens I. (2017). How well-mixed is well mixed? Hydrodynamic-biokinetic model integration in an aerated tank of a full-scale water resource recovery facility. *Water Science & Technology*, **76**(8), 1950–1965.

Reckhow K. H. (1994). Importance of scientific uncertainty in decision-making. *Environmental Management*, **18**(2), 161–166.

Reichert P. (2009). Environmental Systems Analysis (Lecture Notes). Eawag and ETH, Switzerland.

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.

Rosen C., Rieger L., Jeppsson U. and Vanrolleghem P. A. (2008). Adding realism to simulated sensors and actuators. *Water Science and Technology*, **57**(3), 337–344.

Sharifi S., Murthy S., Takács I. and Massoudieh A. (2014). Probabilistic parameter estimation of activated sludge processes using Markov Chain Monte Carlo. *Water Research*, **50**, 254–266.

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

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

Villez K., Del Giudice D., Neumann M. B. and Rieckermann J. (2020). Accounting for erroneous model structures in biokinetic process models. *Reliability Engineering & System Safety*, **203**, 26.