Appendix A

Terms and definitions – application and discussion

A.1 INTRODUCTION

Discussions on the topic of uncertainty can be confusing as similar terms are often used interchangeably for different concepts. This Appendix defines and discusses the concepts and terms related to uncertainty within the context of model-based design and operation of wastewater treatment systems. This Appendix is organised in four sections.

The first section includes terms commonly used in modelling.

The second section covers general terms relating to basic statistical concepts and metrics that form the basis for all uncertainty evaluations. The section lists a set of general terms that describe relationships between measured and simulated quantities. Knowledge of these terms is necessary as they form the basis of all the key concepts in the field of uncertainty. For each term, a general definition has been included and how the term is applied to measurements, model structure and parameter values and results of model simulation and prediction.

The third section covers the most essential concepts and terms regarding uncertainty.

The final (fourth) section presents comparative discussions on terms and concepts confounded with uncertainty.

Editor(s): Evangelia Belia, Lorenzo Benedetti, Bruce Johnson, Sudhir Murthy, Marc Neumann, Peter Vanrolleghem and Stefan Weijers
doi: 10.2166/9781780401034_0129

A.2 MODELLING

Table A1 Terms and definitions relating to model architecture.

Area	Term	Definition	Example
Model architecture	Model	Abstract mathematical description of a system.	The ASM1 model.
	Variable (of system)	Changing characteristic of a system.	A biomass or substrate concentration
	State (of model)	Model states describe system variables.	Bulk liquid ammonia concentration, symbolised as SNH.
	Dynamic model	Model describing the evolution in time of variables of interest ($dx/dt = f(x, t)$).	The ASM1 model describes temporal changes in variables such as COD, oxygen and ammonia (among others).
	Steady-state model	Model describing the expected values of variables of interest under fixed conditions for process inputs and operational variables ($dx/dt = 0$).	Total COD = C1 + C2 + C3, pH = log(H +)
	Algebraic state	State which is computed from other states by means of an algebraic (non-dynamic) equation	Acid−base equilibrium equations in some applications
	Parameter	Value that specifies the behaviour of a system or system model. A parameter is considered to have a single true value only for a given system and/or model, although it may be unknown a priori, and can be determined separately for each application of a given model structure.	Maximum growth rate, affinity constants, kLa
	Stoichiometric parameter	Numerical relationship between compounds participating in a given (bio)chemical reaction or the conversion rates of a compound relative to those of another compound.	Yield
	Kinetic parameter	A parameter included in the equation describing the rate by which a given (bio)chemical reaction occurs as a function of one or more model states	Maximum growth rate, affinity constant
	Constant	Value that specifies the behaviour of systems or models in a universal and invariant way. A constant remains unchanged throughout all applications of a model	Gravitational constant, molar masses

Table A2 Terms and definitions relating to acts of modelling.

Area	Term	Definition	Example
Acts of modelling	Simulation	Generation of a model response	Effluent ammonia concentrations are simulated
	Forecasting	Generation of a model response for future conditions based on currently available information	Expected future nitrogen removal rates are evaluated by means of prediction of nitrogen concentrations in the effluent
	Model set-up or structure identification	Definition or adjustment of the model structure including: reactor design, reactor configurations, reactions and biokinetic model structure	Fixed-film or suspended growth reactor, anoxic and aerobic zones, ASM-type or model using other growth/inhibition terms.
	Calibration or parameter optimisation	Adjustment of parameters for a given model structure so to better reflect observations made in a specific set of conditions (operation)	Observations made during a measurement campaign are used to fit the model to the data
	Validation	Comparison of the simulation results of a calibrated model and an independent set of observations. In theory, the data used for validation contains no information contained in the data used for calibration.	Industrial practice: validation is executed for a wide range of purposes. Good statistical practice: validation is exclusively aimed at the selection of a better/best model structure.
	Objective (function)	Degree of performance of a given model. An optimal model is a model with the best performance. The same objective is used throughout calibration, validation and testing steps.	Often a least-squares objective is used (e.g., mean-squared residual, MSR), in which the objective (MSR) needs to be minimised for model improvement.
	Identifiability	Ability to assign a unique, optimal value to the model parameters under reasonable expectations for data availability and data quality. A model lacks structural identifiability when one cannot assign unique values to the model parameters even in the hypothetical case with an infinite amount of data that is representative (e.g., including dynamics) and is perfectly accurate (no bias, no variability). Parametric identifiability concerns the idea that the parameters of a structurally identifiable model are not necessarily identifiable in practically realisable situations with finite resources and practical limits on available dynamics and data quality.	
	Domain of validity or generalisation	A set or range of situations, either foreseen or not, under which a given model is (still) applicable. A good quality of the model is the ability to stretch its use or extrapolate it further compared to other models. The domain of validity can also be seen as the set of conditions within which a model will give results reliable enough to serve as a basis for a decision, despite its uncertainty. This ability can be defined as generalisation properties or domain of validity. The term domain of validity suggests that a distinct quantifiable boundary for generalisation exists.	
	Optimality	An optimal model is a model that is the best among those available in reaching a certain objective. Such objectives may range from describing a certain phenomenon qualitatively to predicting concentrations of interest.	

Table A3 Terms and definitions relating to model evaluation.

Area	Term	Definition	Example
Model evaluation	Least-squares (LS)	Least-squares objectives/optimality refer to the practice of using a penalty function which is a sum of squared prediction errors/residuals used to calibrate, validate and select models. The use of a LS objective is usually motivated based on the assumption that measurement errors are independently and identically distributed according to the normal distribution.	Sum of squared residuals (SSR), mean -squared residual (MSR) and root mean-squared residual (RMSR).
	Sum of squared residuals (SSR)	Sum of squared residuals.	$SSR = \sum_{i=1}^{m} (r_i)^2 \quad \text{for } i = 1 \ldots m \text{ residuals}$
	Mean-squared residual (MSR)	Average of squared residuals.	$MSR = \frac{1}{m} \times SSR \quad \text{for } i = 1 \ldots m \text{ residuals}$
	Root mean-squared residual (RMSR)	Average of squared residuals.	$RMSR = \left(\frac{1}{m} \times SSR\right)^{1/2} \quad \text{for } i = 1 \ldots m \text{ residuals}$
	Sum of absolute residuals (SAR)	Sum of absolute residuals.	$SAR = \sum_{i=1}^{m} \lvert r_i \rvert \quad \text{for } i = 1 \ldots m \text{ residuals}$
	Mean absolute residual (MAR)	Average of absolute residuals.	$MAR = \frac{1}{m} \times SAR \quad \text{for } i = 1 \ldots m \text{ residuals}$
	Independently and identically distributed (i.i.d.)	Notion that a set of given outcomes (e.g., prediction residuals) are distributed independently and characterised by the same distribution. Independence practically means that the value of one outcome is not informative about another outcome. I.i.d. conditions are typically assumed for measurement errors.	
	Normal distribution	The normal or Gaussian distribution is a widely assumed and applied distribution for residuals and errors and can be characterised by two parameters, namely mean and standard deviation. The probability density function follows a symmetric bell-shape.	

A.3 STATISTICS

Table A4 Terms and definitions relating to basic statistical concepts and metrics.

Term	General Definition	Measurement	Model Structure/Parameters	Model Simulation/Prediction
Outcome	Result of measurement, experimentation, simulation or modelling.	For example, a nitrate concentration.	–	For example, an OUR estimate, a biomass concentration prediction, a kinetic parameter estimate.
Measurement	Assessment of the value of a variable of interest by means of an analytic experiment or on-line signal generation.	For example, a dissolved oxygen measurement.	–	–
Error	Deviation between an outcome and its true value.	Numerical difference between a measurement and the true corresponding value in the sampled system.	Difference between the true system and the model representation. This can be in structure and parameters (separately or simultaneously).	Difference between a predicted value and the reference value in the modelled system or a reference value.
Residual	Deviation between an outcome and its reference value.	–	–	Difference between a predicted and a measured concentration.
Credibility	Probability or degree of belief that a given outcome corresponds to its true, usually unknown, value.	Probability or degree of belief that a given measurement reflects the true underlying variable well.	Probability or degree of belief that a model is representative for the true system.	Probability or degree of belief that a simulated result corresponds well to the true corresponding value.
Credible interval/region	Interval within which an outcome or the region within which a set of simultaneous outcomes are believed to lie, with a given probability.	Range around a measurement within which the true value is expected to lie with a given level of confidence.	Range of model structures and/or parameters within which the true system is expected to be with a given level of confidence.	Range around a simulated result within which the true corresponding value is expected to lie with a given level of confidence.
Confidence interval/region	Interval in which an outcome or the region in which a set of simultaneous outcomes are to be found with a given frequency, when repeated many times	Range within which a repeated measurement is expected to lie with a given level of probability	Range of model structures and/or parameters that will be reached with a given frequency upon repetition of the applied data collection, model identification, and/or calibration procedures	Range of model structures and/or parameters that will be reached with a given frequency upon repetition of the applied data collection, model identification, calibration, and/or simulation procedures

(*Continued*)

Table A4 Terms and definitions relating to basic statistical concepts and metrics (*Continued*).

Term	General Definition	Measurement	Model Structure/Parameters	Model Simulation/Prediction
Bias or systematic error	Bias is the consistent deviation of the measured value from an accepted reference value (ISO 15839:2003). In statistical texts, bias and systematic error are considered to be one and the same.	Bias is introduced into measured variables by means of consistent error (s). It is recommended to use measurement bias to explicitly refer to bias in a measurement device or outcome.	Bias in model structure selection or model parameter identification is the systematic deviation between the real system and the model representation. This concurs when the considered model(s) are not representative of the system or when the data used for model identification, calibration, and validation is biased.	Bias in model simulation is the systematic over- or under-prediction of a variable of interest as the model is unable to sufficiently predict the observations made.
Trueness	Antithesis of bias, that is, the degree of how close an outcome is to an accepted reference value.	How close the measurement is to its reference value.	–	–
Precision	Precision is the closeness of agreement between independent measured values obtained under stipulated conditions (ISO 15839:2003). Precision is a qualitative concept and not a number.	–	Degree to which repeated modelling exercises will deliver a similar model.	Closeness of independently reproduced outcomes under the same, specified conditions to each other.
Variability	Spread of 'true' values of a quantity. In measurements, variability is the opposite of precision (ISO 5725-1:1994). Variability is an expression of random error and is a property of the population, not of our state of knowledge (Kelly and Campbell, 2000).	Degree to which repeated measurements show different or dissimilar results; also, the degree of being far from each other.	–	–
Accuracy	Comprises trueness and precision and is therefore a single expression for systematic and random error.	Closeness of agreement between a measured value and the accepted reference value (ISO 5725-1:1994, ISO 15839:2003).	–	–

A.4 UNCERTAINTY

Table A5 Terms and definitions relating to essential concepts regarding uncertainty.

Term		Definition
Uncertainty		Degree of inability to determine or predict the exact behaviour of a system or process both now and in the future. Uncertainty relates to (1) the inability to determine truly and precisely what has happened in the past because several possibilities lead to similar observations and (2) the inability to predict truly and precisely what will happen in the future. Uncertainty results from lack of knowledge and is *partly* reducible through the acquisition of additional knowledge, for example, more data or further understanding of a process.
Nature of uncertainty	Aleatory (irreducible) uncertainty	Aleatoric uncertainty is representative of unknowns that differ each time the same experiment is run. It is due to the inherent variability of a system and cannot be reduced with any further research (e.g., rainfall, toxic spills). It is classified as irreducible and called variability.
	Epistemic (reducible) uncertainty	Epistemic uncertainty is due to things that could be known in principle but are not known in practice. This may be because a quantity has not been measured sufficiently accurately, or because the model neglects certain effects. It can be reduced with further research or measurements (e.g., experimental determination of kinetic parameters) in which case it is classified as reducible and called epistemic uncertainty.
Level of uncertainty	Quantifiable uncertainty	Can be quantified and described with statistical methods and can be attributed to uncertainties such as a random measurement error of a sensor.
	Scenario uncertainty	Can be described with qualitative estimations of possible outcomes that may develop in the future. Realistic assumptions about relationships and/or driving forces within the model can be established. It is not possible, however, to derive the probabilities of the scenarios taking place.
	Recognised ignorance	State where fundamental uncertainty exists, and the scientific basis is insufficient to develop functional relationships, statistics or scenarios.
	Total ignorance	State where the actors are not aware of uncertainty. It is unknown what is unknown.
Location (source) of uncertainty	Context uncertainty	Context refers to the economic, political, social and technical conditions and circumstances that influence the model boundaries and frame the issues that the model is to address. Context uncertainty also relates to the suitability of a model for its intended purpose.
	Input uncertainty	Includes system data uncertainty and external driving force uncertainty. Data uncertainty includes uncertainty in, for example, the influent flow and concentrations to a model. External driving force uncertainty relates to uncertainty associated with changes in conditions that are outside the model boundaries but rather are inputs describing the reference system and external forces driving changes in the current system.
	Model uncertainty	Both model structure uncertainty and model numerical uncertainty arising from computer implementation of the model.
	Parameter uncertainty	Parameter uncertainty is associated with the lack of knowledge regarding the true value of the model parameters as well as uncertainty associated with parameter optimisation technique used during model calibration (e.g., lack of convergence, parameter selection for optimisation).
	Model output uncertainty	The total uncertainty assessed by uncertainty propagation taking all model uncertainties into account.

A.5 DISCUSSION OF TERMS OFTEN CONFOUNDED WITH UNCERTAINTY

A.5.1 Precision and variability

Precision is the quality of a repeated process or procedure to deliver similar results. Variability is the degree of absence of that quality. The larger the variability, the lesser the precision. The term variability is recommended to describe the concept qualitatively while standard deviation and variance are measures used to quantify variability (Taylor & Kuyatt, 1994).

A measurement is variable when subjected to random disturbances or fluctuations. An example that is easily demonstrated is that of a noisy sensor. Even in lab conditions one expects different values for repeated measurements. A less obvious example is sampling error which can induce variability. Indeed, one does not always sample the exact same volume of water or at the exact same spot. Heterogeneity of the medium may cause variation as well.

Precision of simulation results is the degree to which several simulation results are similar to each other. For model quantities (e.g., influent) that are variable, a simulation result can be generated for each possible value of that quantity. A distribution for the simulated variable can be generated based on a single simulation only, usually based on the mean parameter value and transformed into a confidence interval for interpretation. As such, the confidence interval quantifies the precision/variability quality of the simulation.

A.5.1.1 Quantification of precision and variability

Precision should not be defined as the inverse of standard deviation (Taylor & Kuyatt, 1994). Precision is therefore a qualitative concept and not a number. It is most typical to quantify variability (imprecision) rather than precision (ISO 15839:2003). To this end, it is common to estimate the standard deviation. Root mean square residual (RMSR) is the most popular approach to estimate the standard deviation. For this, one subtracts reference values and/or (estimated) bias from the measurements and then computes the averaged square of these residuals. This is then a measure of spread of the measurements.

For simulation results, variability is obtained in a similar way as for measurements. One subtracts reference values and (estimated) bias and then measures the spread, for example by computing RMSR.

A.5.2 Accuracy and uncertainty

An outcome that is close to its reference/true value is more accurate; one that is further away is inaccurate or uncertain. Accuracy comprises trueness and precision and is therefore a single expression for systematic and random error. For this reason, accuracy should only be used as a qualitative concept (Taylor & Kuyatt, 1994) and one should avoid quantifying it. Instead, accuracy/uncertainty should be described with separate measures of bias and variability.

Improving both trueness and precision simultaneously to any desired degree is generally impossible, thereby resulting in a necessary compromise. Since it is not defined how this compromise should be made a priori and since measures for trueness and precision are typically in different scales, it is most often left to the end-user to make this trade-off.

The accuracy of a measurement is the closeness of the given value to the true value. Measurement uncertainty is thus the degree of inability of the measurement to describe the true corresponding value.

The accuracy of a model is the closeness of the model to the described true system. Naturally, model uncertainty is then the inability of the model to describe the targeted system well. Model accuracy/uncertainty can be quantified in several ways depending on the information available. In the purest sense, it is quantified based on the mismatch between the model (structure and parameters) and the true system from which data was derived. More practical measures are based on the ability of the model to predict the true behaviour of the system. A model validation or test step (see below) can serve this purpose.

A.5.2.1 Quantification of accuracy and uncertainty

Accuracy and uncertainty are difficult to quantify for several reasons. Firstly, because the true values are not available for real systems. This can be accommodated in practice by using reference values (see Section A.5.6). Secondly, because accuracy/uncertainty encompasses both systematic as well as random deviations from the truth. Quantifying both by means of one single measure is difficult and has little value in view of model improvement as systematic and random deviation requires different actions for model improvement. As such, accuracy should be decomposed into trueness and precision when attempting quantification. Similarly, uncertainty should be decomposed into bias and variability for purposes of quantification.

A.5.3 Error and residual

The term error is recommended to describe the difference between the obtained value and the corresponding true value and residual for the difference between the obtained value and the reference value if this reference is different than the true value. Except for well-designed laboratory experiments, only residuals are available in practice. Figure A.1 illustrates this notion where the obtained value is given by model simulation (ySIM) and the reference value is a measurement (yREF) for a true value (yTRUE).

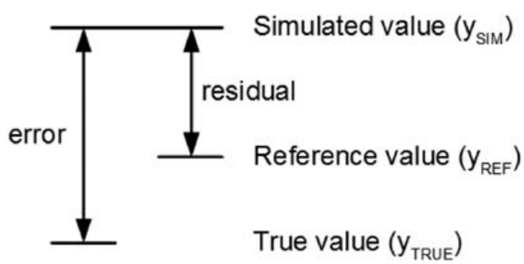

Figure A.1 Suggested separation of true value and reference value and, correspondingly, error and residual.

A.5.4 Trueness and bias

An outcome that is generally or systematically close to its reference/true value is truer; one that is generally or systematically further away is biased. In this definition, the terms generally or systematically are important as they define the difference with accuracy/uncertainty. Conceptually, trueness/bias describes the consistent, general, or long-term characteristics of an outcome.

Upon repetition (of a measurement or experiment), an averaged outcome will (typically) converge to a value called the expected value. The difference between this expected value and the reference value is an estimate of the bias. As such, bias expresses systematic error. The larger the bias, the smaller the trueness. Bias is introduced into measured variables by means of consistent error(s). In Taylor and Kuyatt (1994), bias is reserved exclusively to measurements and systematic error is considered a generally applicable term. [ISO 5725-1:1994; ISO 15839:2003] do not make such a distinction. This may be due to sampling at a location that systematically results in values too high or too low for the overall system. Other possibilities include sensor calibration errors, systematic error in the analytic protocol, and systematic error in handling of samples and/or results.

Bias in model structure selection or model parameter identification results when the considered model(s) are not representative for the system. In other words, insufficient flexibility of the model(s) to be fit to a given data set results in a systematic deviation between the real system and the model representation. For example, kinetic reaction rate coefficients may be far from reality, a one-step nitrification model may be consistently off if the second step nitrification is relatively slow. Another reason for model bias is that the data used for model identification, calibration, or validation was biased.

Bias in model simulation is recognised as the systematic over- or under-prediction of a variable of interest as the model is unable to sufficiently predict the observations made. This may be due to errors introduced throughout measurement campaigns or in modelling or, alternatively, introduced in the measurement of the variable of interest. It may also be due to the extrapolation of a model to a situation which it was not calibrated for or it could be that the system behaviour has changed since the measurement campaign used for modelling.

A.5.4.1 Quantification of trueness and bias

It is most common to quantify bias, rather than trueness. Bias can be quantified relatively easily for measurements. For this, one repeats the measurement (same sampling and measurement procedure) several times and compares to a reference value. The distance between reference and the average of the measurements is an (least-squares optimal) estimate of the bias. The reference value may consist of a trusted reference measurement (e.g., standard protocol) or the set concentration in lab-made standards or samples.

For model structure/parameters, it is difficult to estimate bias, as it requires a true or reference model, which generally is not available. As a result, bias in the model is usually ignored in practice, implicitly assuming that model structure and used data are unbiased. One way of assessing model bias, is to compare simulation results to real measurements to evaluate whether consistent deviations are present. The validation or testing step (see below) may serve that purpose.

For simulation results, one compares simulation results with reference values and computes the consistent deviation between the two. A practical measure is the mean deviation between simulation result and reference value.

A.5.5 Note on true values

Many of the definitions presented in Appendix A include the term 'true value'. True values are generally not known and by virtue of this, definitions based on the knowledge of 'true' values are of little practical use. It is therefore common to replace 'true values' with 'accepted reference values' (ISO 5725-1:1994; ISO 15839:2003; Taylor and Kuyatt, 1994). It is however crucial to realise that one then deliberately ignores the mismatch between accepted reference values and true values as a source of uncertainty. Whether this is of importance will depend on the quality of the accepted reference values. According to [ISO 5725-1:1994; ISO 15839:2003], an accepted reference value is:

(a) An assigned or certified value based on experimental work of some national or international organisation;
(b) A consensus or certified value based on collaborative experimental work;
(c) A theoretical or established value based on scientific principles;
(d) When (a), (b) and (c) are not available, the expectation of the (measurable) quantity, that is, the mean of a number of measurements.

The use of accepted reference values is common for sensor calibration. For example, one obtains laboratory standard measurements and uses these to check a sensor which is based on a different measurement principle. This is unlikely to be useful for model parameters. For predictions, one may be able to obtain standardised measurements as a reference. Importantly, when one uses reference values, one inherently assumes that these show no bias or variability of their own, that is, one considers the reference values perfect.

A.5.6 Note on repetitions

Variability of measurements, parameter estimates and model predictions can be described by means of repetitions. It is assumed that these repeated values are produced in the exact same way each time (identically distributed) and that they are independent of each other (independent sampling). In practice, this may not be the case for the following reasons:

- There is simply no repetition made;
- The measured/parameter/prediction value depends on other values, for example, through redundant relationships or dynamic relationships. For example, it is common that two parameters are correlated. Also, consecutive measurements of a variable in a dynamic system are likely auto-correlated.

Without repetitions, one can do little to obtain separate descriptors of bias and variability. In the rare case where the true value is known, one can only obtain an overall measure of uncertainty based on a single measurement. In such a case, one has little clue on how one can reduce this uncertainty since the reduction of bias and variability require different actions.

One can try to explicitly account for dependencies between distinct variables through redundancy or dynamic relationships when one has measurement or values for these variables. For example, data reconciliation techniques may be used. To obtain corrected values which satisfy the assumed relationships, for example, a mass balance over a process unit, one uses the corrected values as reference values. Following that, one can compute residuals between measurements and corrected values, which then serve to characterise the residuals in terms of bias and variance. Here, one will now typically assume that the residuals are independent and identically distributed. Available techniques belong to the fields of statistical process control and data reconciliation. Note that one typically assumes that the relationships are given without error, that is, they are a perfect representation of reality.

A.5.7 Bias, variability and uncertainty: a graphical example

Consider that one aims at a target in a shooting game and one has multiple chances to try. At each trial, one may or may not be close to the target. After a series of trials, one can characterise the distribution of the result in the different trials. Suppose four people are participating in the game. For each individual, one obtains results as in Figure A.2. It is typical to describe these results in terms of bias and variance. The combination of both represents the uncertainty.

In the result on the top right of the figure, one has the results for the best player. This player has a low bias, that is, the average of all trials is close to the target. This player has also low variability since all trials are close to each other. The second player (top left) also has low bias (average of trials is on target) but shows much more variability.

At the bottom-right, the third player's results are shown. In this case, the average of the results is far from the off target. One says that the player shows bias. However, the trials are very close to each other, meaning that the player shows low variability. Finally, the average of results for the fourth player is off target while the trials are rather far from each other. This player thus shows high bias and high variability.

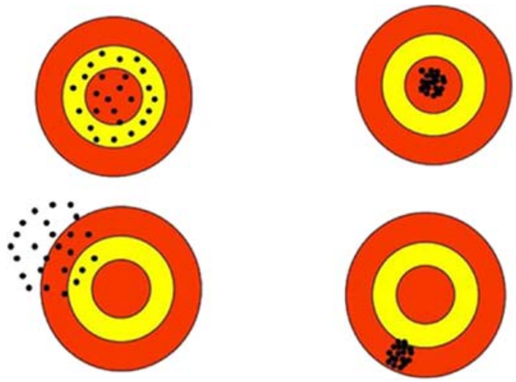

Figure A.2 Top left: low bias, high variability. Top right: low bias, low variability. Bottom-left: high bias, high variability. Bottom-right: high bias, low variability. (Source: http://www.minedesignwiki.org/index.php/Sampling_QAQC).

Uncertainty is now a combination of both bias and variability. As such, one can say that the first player shows low uncertainty. One also says that this person is accurate. The fourth player shows high uncertainty, or in other words, low accuracy. The second and third players show uncertainty levels between the first and fourth players' uncertainty. Importantly, it is not always clear how one should weigh accuracy against precision. It is difficult to gauge whether player 2 is better or worse than player 3.

A.5.8 Link between measurement, modelling and prediction

One can also characterise measurements, model (parameters) and prediction in terms of uncertainty, bias and variability. In these cases, a bias is the general tendency of a measurement, model parameter estimates or prediction to have a different value than the true value. Variability is the descriptor for how far repeated values for measurements, model parameter estimates or predictions are from each other. In Figure A.3, each square represents the average of three measurements demonstrating measurement error, the vertical dotted line represents the influent variability of the daily average, the horizontal dotted line shows the annual average and the continuous line shows the interannual variability of the measurement.

A.5.9 Qualitative model performance criteria

A.5.9.1 Identifiability

In certain situations, it may not be possible to obtain (good) estimates of parameters. This is usually the result of the combination of (1) a model with a rather large number of parameters which one aims to identify, (2) lack of representative, dynamic data or (3) lack of data quality. According to Dochain *et al.* (1995), identifiability can be defined as follows:

Assume that a certain number of the state variables are available for measurement; on the basis of model structure (structural identifiability) or on the type and quality of available data (practical identifiability), can we expect to give via parameter estimation a unique value to the model parameters?

A model lacks structural identifiability when one cannot assign unique values to the model parameters even in the hypothetical case with an infinite amount of perfect data that is representative (for example including dynamics) and is perfectly accurate (no bias, no variability).). In this case, additional data collection cannot aid in the modelling process. In contrast, parametric identifiability concerns the idea

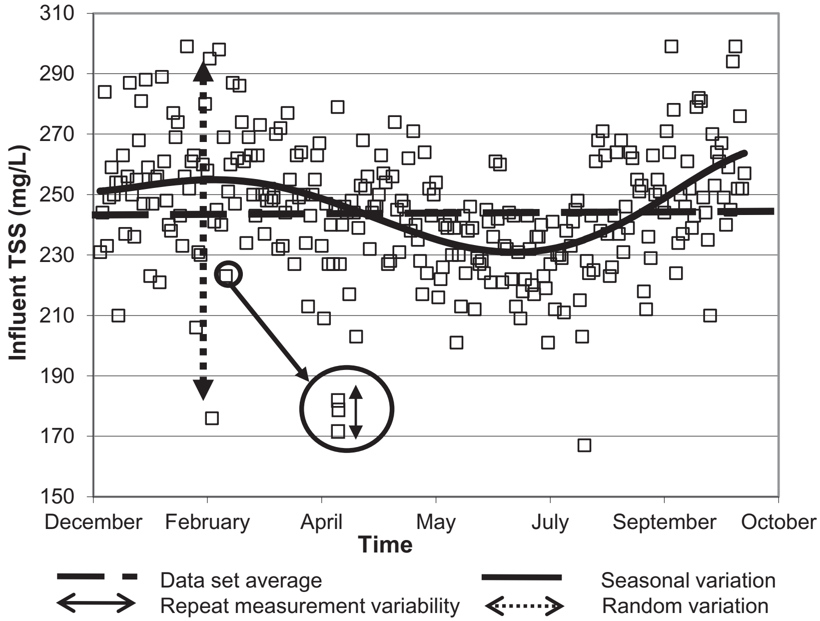

Figure A.3 Measurement error, daily average variability and inter-annual variability of influent TSS.

that the parameters of a structurally identifiable model are not necessarily identifiable in practically realisable situations with finite resources and practical limits on data quality.

A.5.9.2 Generalisation and domain of validity

As one may expect that during modelling, one has no access to data representative of all behaviours of the system, it is important to consider the capacity to extrapolate the model beyond the conditions covered in the calibration and validation set also, one may expect that the true system itself may change over time. Consequently, one will likely arrive in a situation where the model is used still while not exactly representative of the true system. Such extrapolation may be valid or not. A good quality of the model is thus the ability to stretch its use or extrapolate it further compared to other models. The domain of validity can also be seen as the set of conditions within which a model will give results reliable enough to serve as a basis for a decision, despite its uncertainties. This ability can be defined as generalisation properties or domain of validity. The latter suggests that a certain and quantifiable boundary for generalisation exists.

A.5.9.3 Optimality

An optimal model is a model that is the best among those available in reaching a certain objective. Such objectives may range from describing a certain phenomenon qualitatively to predicting concentrations of interest.

A.5.10 Reliability and redundancy

According to IEEE (1990), reliability is the ability of a system or component to perform its required functions under stated conditions for a specified period of time. For example, reliability of a model is the degree to which one is certain that a given model will support its intended task over a time horizon, that is, the extent to which one can rely on the model in the future. As such, reliability includes the notion that the performance of a model will degrade over time due to the inability to incorporate unforeseen changes in the represented systems or inherent incompleteness.

The basic method to assess reliability is to define the risk, define the likelihood, and define the consequence (preferably as a cost). This method has been used to assess criticality, which is used for capital improvements planning and prioritisation. The criticality of a project is the product of the likelihood/frequency of failure and the consequence of failure.

Redundancy in an engineered system can be defined as a design practice to include backup components in a system or incorporate interchangeable components so that the system can be repaired quickly. In either case, the intention is that the system can operate at an acceptable performance level without interruption when a piece of equipment fails or must be taken out of service. This is important in safety-critical systems (e.g., plane, nuclear reactor) to avoid damage in the case of failures in a single part of the system. Data redundancy is the degree to which multiple measurements contain the same information. This property is what one uses to remove gross outliers from data sets by means of mass balances. Most statistical techniques for fault detection are based on redundancy (e.g., Frank, 1990).

Wastewater treatment redundancy should not be confused with process reliability. Reliability refers to the inherent dependability of a piece of equipment, a unit process, or the overall treatment process in meeting the design objective (Tanaka *et al.*, 1998). In terms of effluent quality, Niku *et al.* (1979) and Bott and Parker (2011) referred to process reliability as the ability to meet the specified effluent requirements free from failure, or as the probability of success, where failure is the probability that the effluent concentration is greater than the discharge permit limit. McBride and Ellis (2001) defined reliability as the percentage of time a wastewater treatment plant remained in compliance with discharge standards. Reliability analysis has been used to predict the performance of a technology over time and to determine the strategies that improve performance and reduce risks of failure (Etnier *et al.*, 2005). Redundancy on the other hand, can be viewed as a subject of reliability. Redundancy is practiced in the design of wastewater treatment plants to improve reliability through the provision of standby equipment or processes that reduce the risk of failure to meet water quality regulations or guidelines (Palmer *et al.*, 2003).

Although the definition of redundancy does not include regulatory compliance or reliability standards, some data show a direct relationship between treatment process reliability, redundancy design and regulatory compliance. Bott and Parker (2011) concluded in a comprehensive study of nutrient removal plants, that one of the main causes affecting the performance of treatment plants was the reliability and redundancy of important unit processes or pieces of equipment in the wastewater treatment plants.

A.5.11 Robustness and resiliency

In general, one desires that any engineered system can handle disturbances for quite some time before (large) deviations in operation are seen. Robust systems do not easily break down or disintegrate, that is, they can withstand extreme conditions without visible changes in structure or functionality. This capacity is usually described as robustness which can be defined as the property that permits a system to maintain its functions against internal and external perturbations (Kitano, 2004). Robustness is a characteristic which becomes apparent when imposing extreme or potentially harmful conditions (stress) onto a good-working system. For example, concrete is a robust material as it does not change shape under considerable stress. Rubber is not a robust material as it changes form and shape under slight stresses already. Naturally, a larger robustness is generally a good thing. However, robustness generally comes at a cost. For example, in order to make a buffer tank more robust to extreme flows, one needs to design a larger tank.

Resiliency is a term used broadly and differently in different contexts. In the field of process control, it signifies a property very similar to robustness, for example, resilient control systems are those that tolerate fluctuations via their structure, design parameters, control structure and control parameters (Mitchell & Mannan, 2006). More recent definitions provided in the context of cyber-security, include a non-random component to the cause of disturbances and a goal of awareness: A resilient control system is one that maintains state awareness and an accepted level of operational normalcy in response to disturbances, including threats of an unexpected and malicious nature (Rieger *et al.*, 2009). Thus, to achieve resiliency, it is not sufficient to tolerate or cope with a disturbance, one must also be aware of it. This makes a clear distinction with robustness, which is a passive approach to handle (random) disturbances.

A general definition of resiliency could be the degree to and/or rate at which a system can recover from disturbances or upsets, caused by random causes or wilful actions. In contrast to robustness, which is characterised under stress conditions, resilience is, in addition, characterised by means of a recovery process or period. For example, can one get the initial performance again? Does the system return quickly to normal behaviour? Concrete can be considered a non-resilient material. Indeed, once broken it is not easy to fix and one will generally replace it with new concrete. Rubber is a resilient material. Indeed, as one releases stress, rubber generally goes back to its original shape and form. Resilience can also be regarded as a capacity of self-healing. As with materials, it is expected that robustness and resilience are to be bargained against each other. In addition, resilience also comes at a certain cost. For a buffer tank to recover faster from an extreme flow event, one may need pipes with a larger diameter.

REFERENCES

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.

Dochain D., Vanrolleghem P. A. and Van Daele M. (1995). Structural identifiability of biokinetic models of activated sludge respiration. *Water Research*, **29**(11), 2571–2578.

Etnier C., Willetts J., Mitchell C., Fane S. and Johnstone D. (2005). Decentralised Wastewater System Reliability Analysis Handbook. Project No. WU-HT-03–57. Prepared for the National Decentralised Water Resources Capacity Development Project, Washington University, St. Louis, MO, by Stone Environmental, Inc., Montpelier, VT.

Frank P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. *Automatica*, **26**(3), 459–474.

IEEE. Institute of Electrical and Electronics Engineers. (1990). IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries. New York, NY ISBN 1559370793

ISO 5725-1. (1994). Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions. International Organisation for Standardisation, ISO 5725-1:1994.

ISO 15839. (2003). Water quality – On-line sensors/analysing equipment for water – Specifications and performance tests. International Organisation for Standardisation, ISO 15839:2003.

Kelly E. J. and Campbell K. (2000). Separating variability and uncertainty in environmental risk assessment – making choices. *Human and Ecological Risk Assessment: An International Journal*, **6**(1), 1–13.

Kitano H. (2004). Biological robustness. *Nature Reviews Genetics*, **5**(11), 826–837.

McBride G. B. and Ellis J. C. (2001). Confidence of compliance: a Bayesian approach for percentile standards, *Water Research*, **35**(5), 1117–1124.

Mitchell S. M. and Mannan M. S. (2006). Designing resilient engineered systems. *Chemical Engineering Progress*, **102**, 39–45.

Niku S., Schroeder E. and Samaniego F. (1979). Performance of activated sludge processes and reliability-based design. *Journal (Water Pollution Control Federation)*, **51**(12), 2841–2857.

Palmer T. M., Ross M. C. and Jank B. (2003). Efficient Redundancy Design Practices. International Water Association (IWA), London, UK.

Rieger C., Gertman D. and McQueen M. (2009). Resilient control systems: Next generation design research. 2nd Conference on Human System Interactions, pp. 632–636.

Tanaka H., Asano T., Schroeder E. D. and Tchobanoglous G. (1998). Estimating the safety of wastewater reclamation and reuse using enteric virus monitoring data. *Water Environment Research*, **70**(1), 39–51.

Taylor B. N. and Kuyatt C. E. (1994). Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results. National Institute of Standards and Technology Technical Note 1297, MD, USA. https://www.nist.gov/pml/nist-technical-note-1297 (accessed 2021).