On Model Uncertainty and Simplicity, within the Context of Catchment and Lake Water Quality Modelling

T.J. Cox, Streamlined Environmental Ltd.

"All models are wrong...some are useful"

"Since all models are wrong, the scientist cannot obtain a 'correct" one by excessive elaboration. On the contrary, he/she should seek an economical description of the natural phenomena. Just as the ability to devise simple but evocative models is the signature of the great scientist, so overelaboration and overparameterisation is often the mark of mediocrity."

"Since all models are wrong, the scientist must be alert to what is importantly wrong. It is inappropriate to worry about mice when there are tigers abroad."

- Dr. George Box, Statistician, c. 1976

"*Understanding uncertainty, and being transparent about incomplete knowledge, is essential if policies or regulations based on models are to be credibly defended*."

"*For models used in environmental regulatory decision-making, high standards of transparency are important for a range of reasons. Most fundamentally, those affected by regulations have a right to understand the basis on which the regulations were made.*"

- Simon Upton, Parliamentary Commissioner for the Environment, 2018[1]

The words above span a period of over forty years, which roughly corresponds to the lifetime of computer models in the field of environmental science. Models are used to both make predictions about the future of natural systems and to gain a better understanding of current conditions. Their use in the field of water resources is widespread, for simulating both the quantity and quality of freshwater systems. Models will undoubtedly continue to serve a vital role in tackling our water challenges in the future. That said, models must be used smartly and appropriately in all applications. Good modellers understand the limitations of their tools. They also understand the value added, or lack thereof, by model complexity and work to ensure "fit for purpose" in their developed models. The discussion below touches on the subjects of model uncertainty, transparency, simplicity, and usability. In many instances, the comments represent my own reflections and opinions, based on experience. This is a discussion paper and certainly isn't intended to present as any sort of final or authoritative word on the topics. In other words, consider this "food for thought".

Uncertainty and sensitivity…

As George Box opines, since all models are wrong, it is critically important to understand where the models are likely to be wrong and the potential magnitude of this error. The fact that models are wrong creates uncertainty in their predictions. New Zealand Ministry for the Environment guidelines[2] refer to this type of error as "model and parameter uncertainty", contrasting it with "natural uncertainty" (known variability in nature) and "deep uncertainty" (fundamental, unavoidable, and unrelated to model formulation). More recently, Simon Upton emphasizes the importance of understanding uncertainty in the New Zealand regulatory context specifically. There is a range of methods available to modelers to achieve, and communicate, an understanding of uncertainty. Two primary categories of such methods are sensitivity analyses and uncertainty analyses. Both are briefly defined below.

Sensitivity analyses reveal how model output changes in response to changes in individual input parameters, within a range of input uncertainty. Typically, in such an exercise, input parameters are varied in isolation. The primary result of a sensitivity analysis is the relative *apportionment* of uncertainty to individual input parameters[3]. In other words, the analysis identifies the input parameters that key model outputs are most sensitive to. Such results might, for example, guide data collection efforts, or model refinement, designed to reduce uncertainty. Sensitivity analyses can also be used to evaluate overall prediction uncertainty when a small number of input parameters are known, a priori, to account for the bulk of the total model uncertainty. See Box 1 for a brief discussion of some simple examples.

Box 1 – Sensitivity analyses in lake water quality modelling

In a recent local study of a proposed new constructed lake, we applied an uncalibrated model (there was no lake!) to predict the quality, and ultimate trophic status, of the future lake. Clearly, uncertainty was high! However, a simple sensitivity analysis revealed model predictions to be surprisingly robust within ranges of reasonable lake kinetic parameter uncertainty. Less surprisingly, a major source of uncertainty was shown to be the assumed lake inflow nutrient concentrations from adjacent stormwater wetlands. Therefore, close monitoring of lake inflows became an integral part of a developed adaptive management plan for the lake.

As a second example, a Waikato urban lake was modeled to predict impacts from long-term changes in catchment land use. The magnitude of groundwater infiltration contaminant loads was known, a priori, to be a major source of uncertainty in the analysis. Such loads are very hard to measure. Simple sensitivity analyses, focused on this single input parameter, helped to quantify and communicate this important element of uncertainty in the model predictions.

The term "uncertainty analysis" refers to another category of methodology that is related to, but fundamentally different than, sensitivity analysis[3]. Uncertainty analyses aim to quantify the magnitude of overall uncertainty in model predictions. By itself, an uncertainty analysis doesn't apportion uncertainty, but rather computes a range of model output values resulting from the

combination of multiple uncertain input parameters. Probabilistic, or stochastic, simulations of models are commonly employed to comprehensively capture model uncertainties. In this type of model, input parameters are defined as random variables with a range, or continuous statistical distribution, of values, rather than as a single (deterministic) value. A random sampling regime (e.g. Monte Carlo) is used to sample over the full range of input values, and calculated outputs are stored for each sampling iteration. Final output is generally presented in the form of probability distributions, providing insight to stakeholders on the range of potential outcomes and even the likelihood of each (Figure 1).

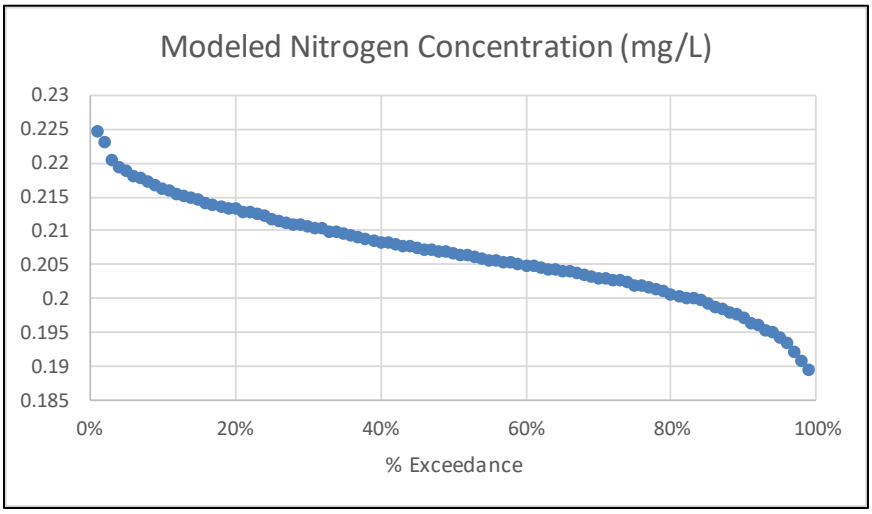


Figure 1 Stochastic model example output (from the CASM catchment water quality modeling software).

On a related note, the output of uncertainty analyses can be particularly useful for decision-makers and regulators who seek to assess "risk". Risk can be viewed as the combination of likelihood and consequences (or effects)[2]. For example, the risk associated with an outcome with low quantified likelihood of occurrence, but dramatic consequences, would likely be considered high. In this way, model uncertainty analysis output can inform risk assessment.

Model calibration exercises can reduce predictive uncertainty and increase confidence. If a model can accurately reproduce past conditions, we have more confidence in its ability to accurately forecast future conditions. Note that calibration is focused on reducing "model and parameter uncertainty", as defined above, but does not generally address either "natural uncertainty" or "deep uncertainty"[2]. Further, it is important to realise that even the best of calibrations does not eliminate the potential for significant predictive error. In other words, a set of parameters may produce an excellent hindcast reproduction of past conditions but not provide for an accurate prediction of future conditions. Anyone who has performed a model "auditing" exercise (which, by the way, is not done enough in the modeling world) knows that this can often be the case. Specific parameter inaccuracies can have little impact on the simulation results for one set of conditions but become significant when simulating a different set of conditions. Similarly, there may exist multiple combinations of input parameters that achieve roughly the same set of hindcast outputs but produce decidedly different output for predictive simulations. Implied here is the existence of non-unique calibration parameter sets.

Also implied here is the fact that predictive simulations may be sensitive to such parameter error, despite calibration output being insensitive.

It can be argued that more complex models, with greater numbers of input parameters, are at greater risk of the type of error described above than simpler models. Depending on the supporting data, models can be "over-parameterised" with the potential for multiple, widely varying calibration parameter sets. Uncertainty analyses should therefore focus on the uncertainty associated with the output that is ultimately most important to final decision making. It should also reflect the full range of plausible values associated with input parameters. Identifying multiple, equally plausible model calibration parameter sets, and incorporating into uncertainty analyses, can be a powerful way of addressing the calibration uncertainty problem.

Model transparency....

Model transparency is important for appropriately interpreting, and using, results and for identifying, and understanding, predictive limitations and potential prediction inaccuracies. In my opinion, transparency, in this context, refers primarily to an understanding of underlying model algorithms, fundamental equations, and conceptual construct (i.e. the relationships between input and output parameters). It also refers to an understanding of what a model does, and doesn't, do well. This basic understanding is imperative for any model user.

Transparency is achieved primarily through model documentation and the user interface experience. Uncertainty and sensitivity analyses, as described above, can also go a long way toward model transparency. For these reasons, it might be argued that transparency is easier to achieve with simple models compared to complex models. Put another way, transparency and complexity typically have an inverse relationship.

Note that, to a lesser degree, open source code can also contribute to transparency. However, this requires a careful review of likely thousands of lines of computer code to achieve any gains in transparency. The typical model user is unlikely to have either the expertise, or inclination, to perform such a review. Code review, of course, has its place in the initial model development peer review process, but I don't believe that is what Simon Upton[1] is referring to.

Model simplicity...

On the subject of simplicity, George Box urges us to seek an "economical description" of natural phenomena and avoid "excessive elaboration". More sharply, he states that overelaboration is often the mark of "mediocrity". Most of us are also familiar with Einstein's imperative that "everything should be made as simple as possible, but no simpler". A good modeler should identify an appropriate level of model complexity based on the fundamental questions to be answered, the availability of supporting data, and even the time and resources available for a given study. Model complexity should also be guided by our understanding of the modeled system and relevant processes. If we know very little about the underlying processes, it makes no sense to construct a complex mechanistic model.

In many, but not all, cases "simpler is better". If added complexity does not clearly provide improvement in predictive power and accuracy, then it should be avoided. Simpler models, while being more economical to develop, also generally provide for greater transparency and better

understanding of uncertainty. In fact, comprehensive uncertainty analyses are often not possible with complex models due to the sheer number of parameters. Simple models can often be developed within a probabilistic framework, allowing for easy and comprehensive uncertainty analysis. Indeed, resources can, importantly, be spent on understanding what we don't know in a model rather than attempting to do more than is practical and/or, potentially, giving the illusion of knowing more than we do. In line with Dr. Box's recommendation, simple models allow us to focus on parameters and relationships that are critical to subsequent predictive simulations and to be "alert to what is importantly wrong" about them. Simple tools are also often more usable, to a broader range of end users. Usability is an important consideration if the models are to be used for decision support by non-expert users, such as in a regulatory context or for government planning (see my parallel white paper on water allocation modeling for more on this topic![4]).

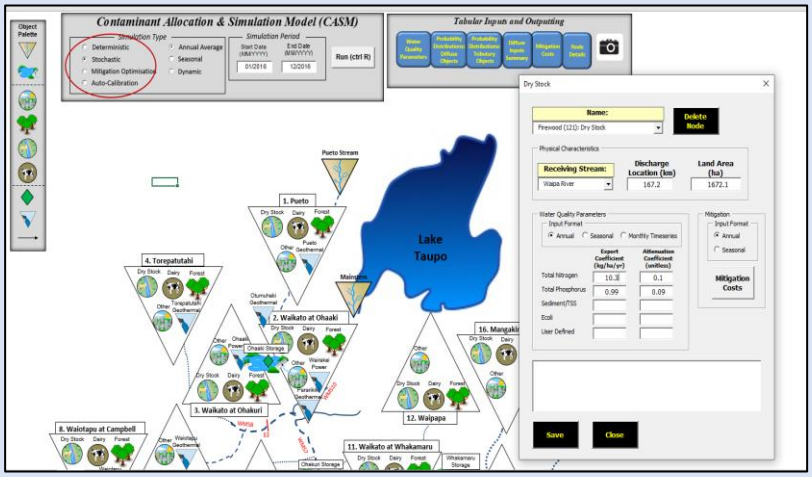
On the other hand, a certain level of model complexity is often required to answer the fundamental questions of a given study. Explicit mechanistic representation of complex processes may be required if those processes directly feature in predictive simulations. In such cases, lumped parameterization, or implicit representation, won't cut it. For example, investigating lake nutrient internal loading response to a change in external loading can't be achieved without a rigorous representation of lake sediment diagenesis processes and a mechanistic link between internal and external loads. The mechanistic link might be provided with separate inputs of water column settling rates, sediment adsorption/desorption and mineralization rate constants, and vertical diffusion coefficients. A lumped model might only include user-prescribed net sediment nutrient flux rates, independent of water column conditions, which would show no response to a long-term change in catchment loads. Further, overall model accuracy might be enhanced with increased complexity if that complexity is supported by adequate data to achieve confidence in model parameterization. In such cases, supporting data must allow for the isolation of key model parameters in the calibration process. For example, a mechanistic lake sediment diagenesis model might require sufficient sediment nutrient data to adequately quantify lake sediment adsorption, desorption, and mineralization rates in the model.

Catchment and lake water quality models, specifically...

Catchment water quality models can span a broad range of complexity levels. At the simplified end of the spectrum, export coefficient models, like our own CASM software[5] (Box 2), are a popular option. The focus of export coefficient models is on the delivery of contaminant mass to receiving waters at a coarse timescale (e.g. seasonal or annual) and, typically, at a broad spatial scale (e.g. sub-catchment). Although these models can accurately depict such aggregated mass delivery, the specific mechanisms involved in this delivery are not explicitly represented. Hydrology is only implicitly represented in the export coefficients, and mass fate and transport processes are aggregated into lumped attenuation coefficients. However, because of the small number of input parameters in export coefficient models, data requirements are relatively low, and the model construction and calibration processes are streamlined. Further, the models easily lend themselves to uncertainty and sensitivity analyses. The primary sources of parameter uncertainty in such models are the export and attenuation coefficients.

Box 2 – CASM: A simplified approach to simulating complex catchment water quality

Streamlined Environmental's Contaminant Allocation & Simulation Model (CASM) was developed to address an identified need for a flexible, and usable, generalised modeling tool for simulating diffuse and point source contamination at a catchment scale. Key model inputs are export coefficients (kg^{-1} ha^{-1} $year^{-1}$), which can vary by both sub-catchment and by landuse category, and two types of lumped attenuation coefficients (diffuse pathway and instream). Export and attenuation coefficients can be specified as annual average values or can vary seasonally or as a monthly timeseries. A fully stochastic simulation mode is available in CASM, whereby uncertain parameters can be defined with distributions of values, and output provided probabilistically. CASM also allows for the simulation of mitigation options and offers an optimisation simulation mode to identify optimal catchment mitigation strategies.



In contrast, more complex catchment models, like the USEPA's HSPF[6] or SWAT[7] models, generally include a combination of mechanistic rainfall-runoff hydrologic calculations and mass balance algorithms. In these models, input parameters might include, for example, precipitation interception rates, surface and sub-surface seepage rates, evapotranspiration rates sub-surface advective rates, contaminant surface application and deposition rates, and surface sorption and loss rates. Depending on supporting data and the calibration process, the uncertainty associated with each individual parameter can be high and quantifying overall model uncertainty can be extremely challenging.

Similarly, lake water quality models vary widely on the complexity spectrum. This variation can relate to, for example, lake hydrodynamics, the number of chemical and biological species included in the model, or the model's representation of sediment contaminant dynamics. We might, for example, consider a model like CE-QUAL-W2[8] to reside at the complex end of the spectrum. This software includes detailed lake bathymetry in two dimensions, sophisticated hydrodynamic calculations, and the simulation of up to six different floating algae groups (as well as multiple zooplankton and macrophyte options). At the simplest end of the spectrum, in terms of publicly available software, might be USACE's Bathtub[9] model. Bathtub simulates steady-state eutrophication dynamics in a single well-mixed lake zone with user-prescribed sediment nutrient fluxes.

As with catchment models, increasing lake model complexity generally equates to increasing input parameter and supporting data requirements. Each input parameter can be a source of significant model uncertainty, depending on supporting data. Further, depending on the end use of the model, sophisticated representation of lake processes may or may not be required. For example, model predictions of lumped phytoplankton response to a change in nutrient loading, without changes in lake bathymetry or hydrology, likely don't require the mechanistic detail offered by a model like CE-QUAL-W2. Conversely, a simplified tool like Bathtub would be insufficient for simulating the response of internal (sediment) nutrient loads to long-term changes in external loads. CDM Smith's Simplified Lake Analysis Model (SLAM)[10] (Box 3) likely resides somewhere between these two options on the complexity spectrum.

Concluding thoughts...

Usable and practical models, that are transparent in what they do and don't do well, are required in the New Zealand water resources sector. This is particularly true within a high-level regulatory and planning context. As modelers, we need to develop tools that are fit for purpose and are as accurate and defensible as possible. We also need to work within the practical constraints of any given study and not try to do too much with our models. We may, at times, need to consider the usability of our models for a potentially wide range of end users. Lastly, we need to fully understand our model's limitations and better communicate concepts like uncertainty, likelihood, and risk. Assessing the merits of complexity vs. simplicity in modeling is critical to this process.

Box 3 – SLAM: A practical lake water quality model

CDM Smith's Simplified Lake Analysis Model (SLAM) was developed to address an identified need for a practical water quality model, focused on lake eutrophication, that could be easily and simply applied in planning studies by a wide range of end-users. SLAM is designed to be intuitive in its use and streamlined in functionality and data requirements, while still providing for a robust simulation of lake nutrient and phytoplankton dynamics. SLAM calculates lake mass and flow balances on a daily timestep assuming one or more well-mixed lake zones.

SLAM is simplified in its handling of both lake hydrodynamics and phytoplankton growth. With respect to the former, we argue that, in many cases, detailed and explicit hydrodynamic calculations aren't necessary to simulate water quality dynamics. For example, lake hydraulics might be roughly steady, on an annual or seasonal basis, and/or predictive scenarios may not require changes to lake hydraulics. Phytoplankton dynamics, on the other hand, are always complex. It might be argued here that we don't understand enough about the underlying processes, in any given lake, to represent confidently with process-based equations. Therefore, in SLAM, we use empirical regression relationships to simulate phytoplankton, with the underlying complex processes only implicitly represented. There are multiple options in the software for defining different types of empirical regression equations, which are intended to be site-specific.

SLAM does include a fairly sophisticated dynamic sediment nutrient flux module. This module calculates internal nutrient loads from the sediments to the water column as a function of shallow sediment nutrient dynamics and diffusive exchanges between sediment porewater and the overlying water column. In this way, the model provides for a mechanistic link between external and internal nutrient loads, critical to many lake studies. The inclusion of dynamic and rigorous sediment nutrient calculations within a practical planning level water quality model distinguishes SLAM from most other published lake water quality models.

Lastly, SLAM includes a stochastic simulation option, whereby uncertain input parameters can be defined with ranges, or distributions, of values, and output predictions are provided probabilistically. As has been discussed, it is critically important to understand where we may be wrong in our predictions. This feature helps with that. SLAM also includes an "auto-calibration" option. This allows for easy identification of, potentially, multiple, equally plausible, calibration parameter sets that achieve performance targets within a prescribed error threshold. The parameter sets can then be used within larger uncertainty analyses.

Acknowledgement:

This paper benefited greatly from the review and input of Dr. Mal Green. Thanks mate!

Dr. Tim Cox is a water resources engineer, scientist, and modeler, with over 20 years of experience in the consulting industry. If you haven't already figured it out, Tim splits his time between CDM Smith (USA) and Streamlined Environmental (NZ). If you haven't also already figured out, he is the principal developer of the CASM and SLAM modeling platforms. For more information, contact Tim at tim@streamlined.co.nz.

References

(1) Parliamentary Commissioner for the Environment (PCE). 2018. Overseer and regulatory oversight: models, uncertainty, and cleaning up our waterways. December.

(2) Ministry for the Environment (NZ). 2016. A draft guide to communicating and managing uncertainty when implementing the National Policy Statement for Freshwater Management 2014.

(3) Saltelli, A.; Ksenia, A.; Becker, W.; Fennell, P.; Ferretti, F.; Holst, N.; Li, S.; and Wu, Q. 2019. Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. Environmental Modelling and Software 114. pp. 29 – 39.

(4) Cox, T.J. 2020. On water allocation modelling and the context within New Zealand. Streamlined Environmental Ltd. (hot link)

(5) Contaminant Allocation and Simulation Model (CASM). Version 2.0. User's Manual. 2019. Streamlined Environmental Ltd. (https://streamlined.co.nz/wp-content/uploads/2019/05/CASM-May-2019.pdf)

(6) https://www.epa.gov/ceam/hydrological-simulation-program-fortran-hspf

(7) https://swat.tamu.edu/

(8) http://www.ce.pdx.edu/w2/

(9) http://www.wwwalker.net/bathtub/help/bathtubWebMain.html

(10) Simplified Lake Analysis Model (SLAM). Version 2.0. User's Manual. 2017. CDM Smith.