The model is wrong get over it – uncomfortable truths and decision making with mathematical models

Jason Morrissey, Yellow Sub Hydro Ltd., Canada
Mary Jeddere-Fisher, Yellow Sub Hydro Ltd., UK
Jeremy (Jez) Lofts, Yellow Sub Hydro Ltd., UK
Martin Lacroix, Wood PLC, Canada
Chris Sonntag, BHP Ltd., USA

Abstract

In engineering, mathematical modelling is ubiquitous and often an inter-disciplinary activity for many design processes. In the mining sector, it is frequently the case that the mathematical modelling output from one discipline becomes the input to another. This is particularly true with designs that are affected by surface water and groundwater systems, where water can be viewed simultaneously as a valuable resource, a nuisance, a hazard, and a critical component of our environment that must be protected. With this in mind, it is commonly thought that the purpose of a mathematical model is to help understand how a system works. While this perspective on mathematical modelling is largely true, for an engineer who is responsible for a design, the model's utility must be taken one step further by using the model results to make a decision. This additional step creates a significant challenge for the decision maker because a mathematical model is not a complete description of the system, in fact, it would be more accurate to describe a model as a reflection of what we know and don't know about the system. This inherent limitation to mathematical modelling is much more pronounced in surface and groundwater studies since unlike most engineering disciplines, modelling of natural water systems is exceptionally challenging due to extreme aleatory and epistemic uncertainties. Nonetheless, design decisions must still be made for a project to progress.

While science and engineering literature provides some useful guidance on creating and testing mathematical models, guidance on how to use a model to make a decision is difficult to find. This paper discusses the application of a decision framework for use with mathematical models of surface and groundwater systems through an iterative approach as opposed to the more traditional linear planning process. Following this framework can dramatically improve the transparency and clarity of any mathematical model with a specific focus on the more challenging applications such as in surface water and groundwater systems. Applying the framework will help determine the reliability of a model, highlight

what is known from what is not (of great importance for climate resilience), assist in preparing to procure modelling services and importantly, to facilitate communication of the decision-making process to all stakeholders.

Introduction

In engineering, mathematical modelling is ubiquitous and often an inter-disciplinary activity for many design processes. In the mining sector, it is frequently the case that the mathematical modelling output from one discipline becomes the input to another. This is particularly true with designs that are affected by surface water and groundwater systems, where water can be viewed simultaneously as a valuable resource, a nuisance, a hazard, and a critical component of our environment that must be protected. It is commonly thought that the purpose of a mathematical model is to help understand how a system works. While this perspective is largely true, for an engineer who is responsible for a design, the model's utility must be taken one step further by using the model results to make a decision. However, it is typically the case that the decision maker for a project is not the same person who produced the model. This creates a communication gap where the modeller's interpretation of the real-world system is usually perceived as a 'black box' by the decision maker. This communication gap exists because there is a depth and nuance to mathematical modelling not conveyed in the conventional framework widely used in mathematical modelling. In this paper we review the 'Conventional Framework' to modelling and provide additional steps intended to better bridge the communication gap between the modeller and the decision maker. We refer to this enhanced modelling framework as the Decision Framework.

The Conventional Framework

Figure 1 depicts the Conventional Framework for mathematical modelling promoted in literature. This framework has been widely adopted and discussed in surface and groundwater literature (which we will collectively referred to as "hydrology" in this paper). Although it is worth noting that in hydrology this Conventional Framework is often subdivided into additional steps which can be broadly listed as: 1) define the purpose, 2) create the conceptual model, 3) design the mathematical model, 4) calibrate model, 5) conduct a sensitivity analysis, 6) model verification, 7) create prediction model, and 8) presentation of results. Note that there is no consensus on the number of sub-steps in hydrology literature, but they all follow the same general pattern (e.g. Anderson and Woessner, 2002; Kresic, 1997; Spitz and Moreno, 1996). It is interesting to note that in hydrology text this framework is usually presented as part of the numerical modelling discussion, whereas, in mathematical text this framework is presented as a fundamental to all mathematical models and so is not linked to complexity (Meyer, 2004). We would suggest that this is because while many relatively simple models are regularly used in hydrology (e.g., Theis equation) they are usually not subject to calibration, so the Conventional Framework is perhaps being applied unconsciously in simple cases.

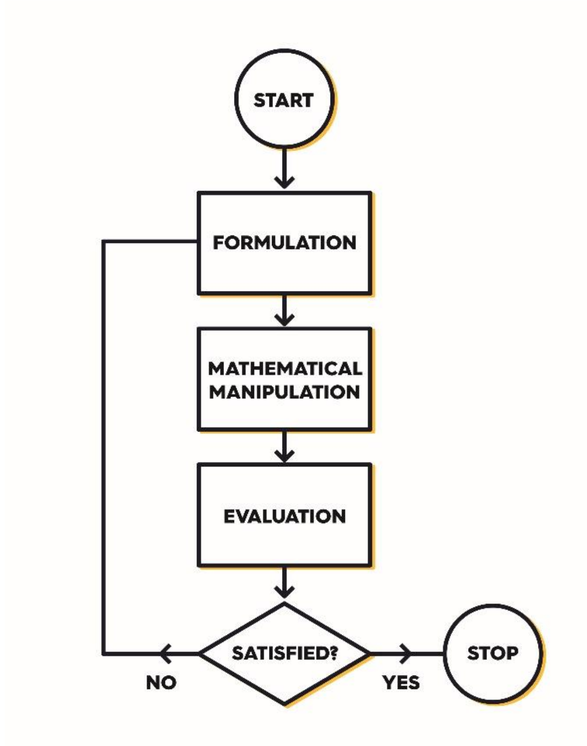

Figure 1: The Conventional Framework for Modelling (after Meyer, 2004)

Although the Conventional Framework is widely published and accepted, it is not clear where it came from, and there is no clear rationale provided as to why we should use it. Dym (2004) offers a brief and plain explanation in that mathematical modelling is a principled activity and that the Conventional Framework allows one to simply determine the "intentions and purpose" of the mathematical model. While Meyer (2004) suggests that mathematical modelling cannot be done mechanically so having some guidelines on how to proceed is merely helpful. From an application point of view, we could say that the Conventional Framework provides steps to modelling which are listed in the order that are normally undertaken because they simply make sense. Suffice to say that this conventional framework has proven useful as it has been adopted by practicing engineers and hydrologists all around the world. However, given that the Conventional Framework is based on common sense rather than a rigorous algorithm, it would be amenable to reasonable changes.

Perhaps the most useful change to the Conventional Framework would be an improvement to the overall communication of the model. Note from Figure 1 that there is an iterative loop in the framework until a "satisfactory" model has been achieved. There are two communication issues here: 1) qualifying a model as satisfactory is subjective to the modeller, and 2) the results of the model are normally reported only after it is deemed satisfactory. In effect, the Conventional Framework promotes a “best explanation” approach to a mathematical model in a linear fashion from the data, through the model, to a result. It is important to note that in this context, the best explanation does not necessarily mean the best possible explanation since it is impossible to prove that any single model is “correct”. Rather the best explanation model is the one selected by the modeller because it seems to be the best-fit to the available observations. This approach may be considered adequate when the modeller is also the decision maker as they would have the benefit of the insight developed through trial and error of the iterative loop. However, when the modeller needs to communicate the findings of the model to another person, the linear nature of the best explanation approach tends not to be very convincing.

The best explanation approach of the Conventional Framework is usually not very convincing because it is prediction focused, and from outside of the process, appears as a “black box” shielding the model from inspection. Engineers and scientists generally understand that a mathematical model is not a complete description of the system, but rather as a reflection of what we know and don’t know about the system and that models, mathematical or otherwise, usually provide only one possible explanation for the real-world observations. This nuance and depth of a model cannot be communicated through a single prediction. In fact, it is widely recognized that most of the insight and confidence in a model is not found in the numerical prediction but rather developed as part of the modelling process (Kay and King, 2020; Tetlock and Gardner, 2015; Barbour and Krahn, 2004; Dym, 2004; Watson and Burnett, 1995). In fact, Pilkey and Pilkey-Jarvis (2007) document several real-life applications of mathematical models which fail as numerical predictions but were still qualitatively useful through the insights developed in the modelling process. In this way, the modelling process, or in our metaphor, the ‘stuff’ hidden in the black box, is more important than the prediction because in the words of psychologist and economist Daniel Kahneman, “No one ever made a decision because of a number. They need a story.”

This evolves into a communication dilemma for the modeller who follows the Conventional Framework, and it is a dilemma which compounds as the number people who need to understand the model increases. In the mining sector, it is frequently the case that the mathematical modelling output from one discipline becomes the input to another and consequently there are often many decision makers involved. Additionally, mining projects need to address outside stakeholders which usually include regulators and the public. Therefore, mining companies have a clear need for better communication than can be offered through the Conventional Framework.

The Decision Framework

The Decision Framework modifies the Conventional Framework by assuming that the modeller and the decision maker are not the same person. The Decision Framework also contends that, for decision making, the qualitative sense of the model is more insightful than the numerical output. Therefore, the Decision Framework seeks to provide a more complete picture of the model by describing the process taken to reach the numerical output. Following this Decision Framework can dramatically improve the transparency and clarity of any mathematical model as well as help determine the reliability of a model, highlight what is known from what is not, and to facilitate communication of the decision-making process to all stakeholders. In effect, the Decision Framework seeks to make the modelling process transparent and interactive, which is beneficial to decision making because in the words of Bender (2000), "When any of us approaches a problem, we do so in a limited, biased fashion. The more open-mined, communicative, and creative we can be, the better our model is likely to be.". Figure 2 depicts the Decision Framework which comprises a loop with six stages. The rationale for the loop and a description of the stages are provided in the following paragraphs.

The looping nature of the Decision Framework indicates that modelling is truly an ongoing process. This looping nature of the framework speaks to the timeless advice that we need to update our understanding of the system as we learn more about it (Kay and King, 2020; Tetlock and Gardner, 2015; Meyer, 2004; Bear, 1979; and Peck, 1969). This is increasingly important when we are modelling under conditions of extreme aleatory and epistemic uncertainty such as hydrology and particularly when we need to reuse the model continuously over very long periods (e.g., decades) such as mining projects. The loop can be applied in two general ways: 1) chronologically as new data becomes available through time, and 2) by evaluating the results of an initial model by running an alternative data set through the process. It should be obvious from the looping nature of the Decision Framework that we as modellers we need to consider the decision maker's longer-term need for revisiting and updating the models. This means that we need to think about how to put more control over the legacy of our mathematical models into the hands of the decision maker such that they own and can pass historical modelling efforts into the future.

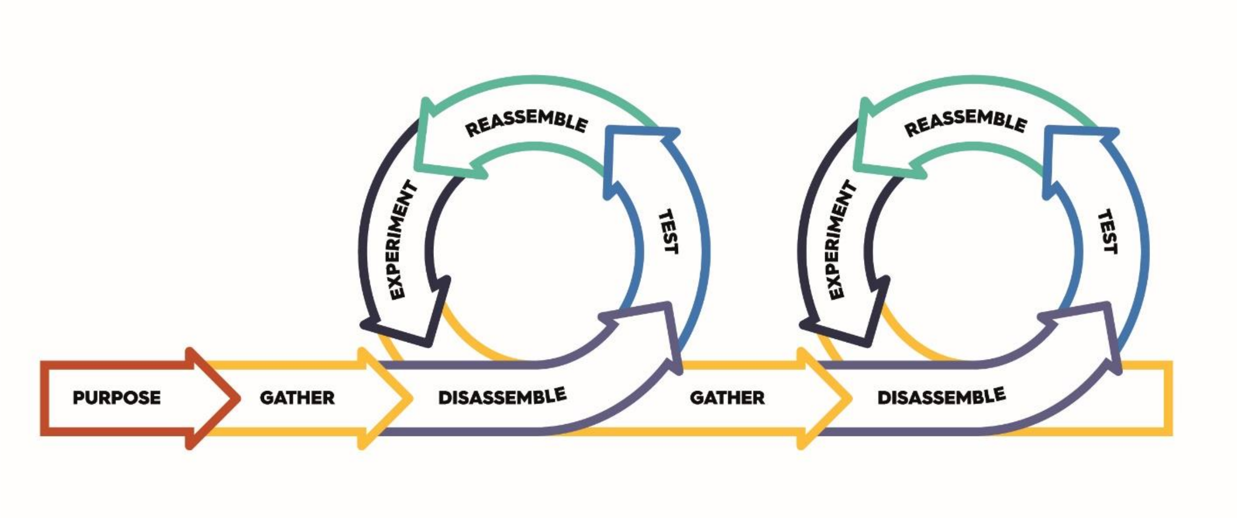

Figure 2: The Decision Framework for Modelling

Purpose

The first stage is defining the purpose for the mathematical model. This Stage is common to all mathematical models and as such it is also the first step in the Conventional Framework. Simply put, at this Stage we must clearly and concisely identify what it is that we want to know. Having a clear purpose allows for the modeller to design the appropriate model. It is interesting to note that uncertainty can and often does enter the process at this Stage. This uncertainty is created by starting with a vague purpose which can lead to modelling the wrong process.

Gather

During the Gather Stage we collate all the available information and make it spatially visible. However, at this Stage it is equally important to highlight any critical data which is missing (e.g., precipitation data from the site). The Gather Stage allows us to critic the starting quality of the data, helps clearly articulate the current limits of our knowledge and helps inform the decision maker on which data should be targeted for future collection. Note that under the Conventional Framework it is typically assumed that the available data is of sufficient quality and quantity. However, communicating the Gather Stage as a distinct step in the process allows the decision maker to develop an initial qualitative sense of the model's reliability. It is also noted that the Gather Stage initiates subsequent loops in the Decision Framework.

Disassemble

The objective of this Stage is to clearly identify and communicate what is objectively true from what needs to be assumed in the mathematical model. This stage comprises 1) data inspection, 2) identification of the important features and dynamics of the system, and 3) establishing a conceptual model of the problem. This stage has much in common with the conceptual model creation step in the Conventional Framework; however, this Stage puts more emphasis on highlighting the clear distinction between objective truths and assumptions necessary to create the model. We believe this distinction to be important as it forms an important part of the discussion around aleatory and epistemic uncertainty. For example, in hydrology modelling, the quality and quantity of stream flow data (including missing data) are very important to the assumptions made, which can have a strong affect on the model results. If the objective truths and necessary assumptions are mixed, then the decision maker does not receive a fully transparent model.

Test

The objective of the test stage is to break the large system model, which is often intractable, into smaller and more tractable components. In the application of mathematical models breaking an intractable model into many tractable ones is considered best practice (Kay and King, 2020, Dym, 2004; Meyer, 1984, and Bender, 1978). In effect the Test Stage aims to demonstrate how the metaphorical black box of the mathematical model works by breaking it into smaller and more comprehensible components. This stage comprises: 1) a first principles deconstruction of the system, 2) application of simple models and hand calculations, and 3) benchmarking against similar problems. This stage allows the modeller to provide the decision maker with multiple perspectives on the problem and allows both a deeper exploration of aleatory and epistemic uncertainties and some ability to quantify the effects of these uncertainties. Under the Conventional Framework this step is not formally undertaken. This is regrettable because in our experience this testing stage provides the most convincing lines of evidence that the reinforce the choices made by the decision maker and it is relatively easy to do.

Reassemble

The Reassemble stage of the Decision Framework comprises building a mathematical model that represents the whole system. Essentially, the Reassemble Stage is the Conventional Framework for modelling, and it is assumed that the reader is familiar with this general process. Within the Decision Framework the benefit of the completing the previous four Stages is that the decision maker is given the opportunity to see how the model was created. This transparency in the model development is not typically provided in the Conventional Framework and it is this lack of transparency which understandably makes the decision maker skeptical about the model results. During this Stage multiple models should be created to reflect alternative possible scenarios and each model should include parameter uncertainty analysis.

Under the Conventional Framework there is no universally accepted means of creating a hydrology model; however, the common default is to create a very complex model to simulate the system. This can be problematic and unnecessary. It is problematic because increasing model complexity results in more time to create the model, increasing simulation runtimes, and it becomes more difficult to understand how the model parameters are interrelated. This in turn typically means that fewer models can be created to test alternative concepts of the hydrological system. It is unnecessary because increasing the complexity of the model does not ensure greater accuracy or a better predictive output. This is particularly true when the available data is sparse, which is usually the case in hydrology. In general, and regardless of the framework being used, the simplest model that captures the important features of the system is usually the best choice.

Experiment

With the acceptable mathematical model of the entire system completed during the Reassemble Stage, the Experiment Stage allows the model to become useful proxy of the real-world system through an interactive sequence of 'what-if' analysis. This 'what-if' analysis, is arguably the most useful aspect of mathematical models when used for engineering design. Under the Conventional Framework there is no specification for 'what-if' type analysis; however, in hydrology 'what-if' questions can be posed at the outset of a well-planned modelling project. Unfortunately, even a well-planned modelling project will result in additional 'what-if' questions being asked at the end of the project. However, in hydrology, most conventional models are too complex to run quickly enough for near real-time 'what-if' analysis and they are typically too rigid to allow significant changes to the model domain. This limits the utility of many conventional hydrology models. One possible work around to this complexity and rigidity problem is to complete 'what-if' analysis during the Test Stage, where rapid simple models can be easily developed for this purpose. Another option is to choose a modelling method during the Reassemble Stage which is both quick enough and flexible to allow for rapid 'what-if' analysis. We recommend the latter approach.

Summary

Hydrological systems can have a profound effect on human endeavours including mining. As such there is an ongoing need to understand how the hydrological system might impact the man-made system and vise-versa. However, hydrological systems are complex and difficult to model. Yet modelling is still our best available tool to understand that system and out interaction with it. Unfortunately, models are merely approximations both in terms of how we think a system works and how we construct mathematical formulae to represent it. Therefore, our ability to accurately simulate hydrological systems is severely limited. Nonetheless, design decisions must still be made for a project to progress.

With respect to decision making, what mathematical models lack in predictive accuracy they to make up through qualitative insights derived through the modelling process. However, this presents a challenge when the modeller and decision maker are not the same person. This paper discussed the application of a

Decision Framework for use with mathematical models of hydrology. This Decision Framework assumes that the modeller must communicate all the insights derived through the modelling process to the decision maker. Following this framework can dramatically improve the transparency and clarity of any mathematical model and to facilitate communication of the decision-making process to all stakeholders. It is our hope that by following the Decision Framework the need to put more control over the legacy of our mathematical models into the hands of the decision maker becomes a serious consideration. Doing so in the mining sector will allow the modeller and decision maker to effectively pass historical modelling efforts and the insights gained to their successors.

References

Anderson, M., W. Woessner. 2002. Applied Groundwater Modeling. Elsevier Academic Press. London, UK.

Barbour, L., J. Krahn. 2004. Numerical Model: Prediction or Process? Geotechnical News Quarterly. Canadian Geotechnical Society. December Issue.

Bear, J. 1979. Hydraulics of Groundwater. McGraw-Hill Inc. New York, USA.

Bender, E. 2000. An Introduction to Mathematical Modeling. Dover Publications Inc. New York, USA.

Dym, C. 2004. Principles of Mathematical Modeling. Elsevier Academic Press. MA, USA.

Kay, J., M. King. 2020. Radical Uncertainty: Decision-Making Beyond the Numbers. W.W. Norton & Co. Inc. New York, USA.

Kresic, N. 1997. Quantitative Solutions in Hydrogeology and Groundwater Modeling. CRC Press LLC. FL, USA.

Meyer, W. 2004. Concepts of Mathematical Modeling. Dover Publications Inc. New York, USA.

Peck, R. 1969. Advantages and Limitations of the Observational Method in Applied Soil Mechanics. Geotechnique, 19(2), 171–187.

Pilkey, O., L. Pilkey-Jarvis. 2007. Useless Arithmetic: Why Environmental Scientists Can’t Predict the Future. Columbia University Press. New York, USA.

Spitz, K., J. Moreno. 1996. A Practical Guide to Groundwater and Solute Transport Modeling. John Wiley & Sons. NY, USA.

Tetlock, P. and D. Gardner. 2015. Super Forecasting: The Art and Science of Prediction. Penguin Random House LLC. New York, USA.

Watson, I., A. Burnett. 1995. Hydrology: An Environmental Approach. CRC Press LLC. Fl, USA.