Editorial: Rapid, reproducible, and robust environmental modeling for decision support: worked examples and open-source software tools

Jeremy T. White[1]*, Michael N. Fienen[2], Catherine R. Moore[3] and Anneli Guthke[4]

[1]Intera Geosciences, Perth, WA, Australia, [2]U.S. Geological Survey, Upper Midwest Water Science Center, Madison, WI, United States, [3]GNS Science, Wellington, New Zealand, [4]Stuttgart Center for Simulation Science, Cluster of Excellence EXC 2075, University of Stuttgart, Stuttgart, Germany

KEYWORDS

modeling, decision support, reproducability, uncertainty analysis (UA), open-source

Editorial on the Research Topic
Rapid, reproducible, and robust environmental modeling for decision support: worked examples and open-source software tools

To provide support for resource management decision making, computational modeling workflows in environmental simulations need to be efficient, reproducible, and robust with regard to informing assessments of the risk of unwanted outcomes. Each of these three attributes is difficult to achieve in practice; aspirations to simultaneously achieve all of them are truly lofty. Too often, modeling analyses are inefficient, the workflow is largely opaque and unknown, and the important simulated outcomes lack the context of uncertainty and/or risk. This Research Topic called for papers that demonstrate rapid, reproducible and/or robust modeling through worked examples and software tools (a preference for open source). The worked examples should demonstrate how the researcher aspired to be rapid, reproducible, and robust; we were interested in the process and approach as much as the results. We aim to stimulate discussion based on lessons learned and results presented, for other researchers and practitioners to build on. We particularly welcomed descriptions of trials and tribulations: What was difficult? What did not work? How were these issues overcome?[1]

Generally, we identified three categories of contributions:

- New open-source software tools designed to facilitate aspects of environmental, hydrological and geophysical modeling;
- New approaches to enable better decision support with modeling;
- Demonstrations/case studies of rapid, reproducible, and robust modeling.

---

1 Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

These contributions came from a wide range of author backgrounds and institutions. This diversity shows that there is broad interest from academia, industry, and government agencies in rapid, reproducible, and robust modeling workflows. We continue to help promote such methods through convening dedicated sessions at international conferences.

Open-source software to support modeling

Leaf and Fienen present Modflow-setup, a workflow toolset to automate the construction of numerical groundwater models for the MODFLOW platform from original geospatial and tabular datasets. The open-source, online code base is extensible through collaborative version control.

Moges et al. call for reproducible model benchmarking and diagnostics, which will find wide acceptance in modeling communities only through standardized methods and ready-to-use toolkits. Using the Jupyter platform, they have introduced HydroBench: an open-source toolset for objectively benchmarking hydrological models that can further be developed by the hydrological community.

Larsen et al. present pyGSFLOW, a Python toolset to transparently and reproducibly prepare input for and postprocess output of the integrated surface-water/groundwater model GSFLOW.

James et al. provide a new standard for geophysical data formats, termed GS Convention, to improve the interoperability, transferability, and long-term archival of such data. Their open-source toolset GSPy provides methods and workflows to build the respective standardized files.

Morvillo et al. present VisU-HydRA, a Python toolbox to compute exceedance probabilities and resilience measures as a basis for assessing the risk of groundwater contamination. It comes with a step-by-step tutorial to ensure reproducibility of the workflow.

Schorpp et al. introduce ArchPy, a python toolset for automating the construction of Quaternary geological models. This is an important step toward including these uncertainties in subsurface modeling workflows in a transparent and reproducible way, because the traditional approach required multiple manual steps using different software, which rendered updates with new data or automation almost intractable.

Pryet et al. present a scripted workflow that facilitates the use of reverse particle tracking in applied groundwater modeling as an efficient surrogate to more computationally demanding advection-diffusion transport modeling for well susceptibility evaluation.

Mudunuru et al. present an approach to improve the calibration of large-scale integrated hydrological models such as SWAT via deep-learning techniques. Compared to more traditional approaches, the proposed routine is more efficient and achieves higher skill scores in calibration.

New approaches to support model-based decision making

Hugman and Doherty discuss the challenge of choosing the right amount of model complexity for decision-making and propose a methodology that allows expert knowledge of system properties to inform the parameters of a structurally simple model. They demonstrate navigating the conflicting and competing objectives of simple and complex model designs on a case study of predictive modeling to support the management of a stressed coastal aquifer.

Elshall et al. present a method for prescreening-based subset selection with decision relevant metrics to exclude non-representative model runs from the prediction ensemble. Following the FAIR (Findability, Accessibility, Interoperability, and Reuse) Guiding Principles for scientific data management and stewardship, they developed and shared interactive Colab notebooks for data analysis.

Moore et al. present a sequential conditioning approach to account for geostatistical model uncertainty, which is shown to have a decisive impact on representing the connectivity of high permeability pathways in contaminant transport assessment.

Manewell et al. investigate spatial averaging functions to infer aquifer properties from aquifer test drawdowns under heterogeneity and feature boundaries. This helps to characterize and robustly estimate aquifer property heterogeneity in hydrogeological site investigation.

Case studies of rapid, reproducible, and robust workflows

Kitlasten et al. present a scripted, reproducible workflow to analyze the impact of ensemble size and vertical resolution on groundwater age predictions for New Zealand.

Standen et al. demonstrate a scripted and open-source application of decision-support modeling for managed aquifer recharge scenarios to mitigate aquifer contamination from saltwater intrusion in the Algarve region of Portugal.

Chambers et al. present a decision-support modeling analysis of the potential for increased groundwater flooding as a result of projected sea-level rise in the low-lying South Dunedin region of New Zealand. They incorporate risk into the analysis proving valuable new information to decision makers.

Brakenhoff et al. present a fully repeatable demonstration of large-scale transfer-function-noise modeling to differentiate contributions to observed groundwater level variations in a region of the Netherlands. Differentiating pumping and climate sources on water level impacts has important implications in how to manage water resources.

De Sousa et al. present a surface-water/groundwater modeling analysis of a semi-arid closed-basin in southwest Australia, and demonstrate efficient, at-scale application of several advanced analyses.

Höglund et al. report a fully-scripted decision-support modeling analysis within the context of contaminated groundwater discharging to surface-water. Innovative techniques are used to assimilate thermal measurements to better resolve patterns of surface-water/groundwater exchange, leading to improved modeling predictions, and ultimately decision support.

The editors are grateful to all of the authors for providing valuable contributions in the space of rapid, reproducible, and robust modeling. We hope that you enjoy reading these contributions. We also hope that in reading these contributions some of you may feel inspired to engage with the open-source community of your modeling field.

Author contributions

JW: Writing–original draft. MF: Writing–review and editing. CM: Writing–review and editing. AG: Writing–review and editing.

Conflict of interest

Author JW was employed by Intera, INC., a for-profit environmental consulting organization with no mechanism of direct financial benefit arising from JW's involvement in this Research Topic.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author Disclaimer

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.