Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-436-SC4, 2019

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "On the assimilation of environmental tracer observations for model-based decision support" *by* Matthew J. Knowling et al.

Matthew Knowling

m.knowling@gns.cri.nz

Received and published: 15 October 2019

We thank Mr Chris Turnadge for his detailed and constructive comments. We feel that his comments can be addressed where appropriate through some minor yet important additions and modifications to the manuscript text. His comments also provide us with an opportunity to reiterate and expand on some of our current decision-support modeling perspectives and philosophies. We respond to each of his comments below.

We agree that the guidance provided in the manuscript is timely for the reasons Mr Turnadge describes (and for reasons described in the Introduction and Discussion sections of the manuscript). It was one of our primary motivations to rigorously investigate

the calls for increased use of diverse data (e.g., Schilling et al., 2019) in the context of decision-support modeling. We thank Mr Turnadge for his positive sentiments.

We agree that our conclusions are "indicative", and we are pleased Mr Turnadge believes our carefully formulated conclusions are appropriate. We feel that indicative conclusions are really all that can be drawn on the basis of two (or even more) real-world case study example demonstrations. We also consider empirical demonstrations to provide an important adjunct to theoretical demonstrations; we feel empiricism in the presence of inevitable site specifics are important to accompany theoretical investigation.

Response to major comments:

1. First, our results do in fact already show the worth of MRT observations in the absence of spring discharge observations (albeit in the absence of other observations too—therefore representing the case where the maximum worth of MRT observations is apparent with an otherwise "empty" observation dataset). Please see the blue bars in the MRT column of Figure 2. Second, the Heretaunga Plains case study presented reflects a real-world investigation, i.e., whereby tritium concentrations were measured after and in combination with discharge measurements. It is not the experience of the authors that tracer concentrations are typically only sampled where flux observations are lacking. This would imply that flux and tracer data contain the same information and can therefore be substituted for one another. Contrast this with much literature suggesting the benefits of as many data types and as much data as possible. For these reasons, we consider the first case study to represent a fair and useful test case.

2. We regret that this comment appears to reflect a misunderstanding (and therefore a need to improve the communication of the manuscript) more than anything else. While we agree with Mr Turnadge's intuition regarding the integration of tracer-derived information along flow paths, we do not show that tritium concentration observations in the second case study are of limited value or "worth" (regardless of where the observations

occurred within the basin). In fact, despite that the second case study does not actually explore the worth or information content of tritium observations, the results suggest that the tritium observations contain "too much" information (or, rather, misinformation when considered through the lens of an imperfect model that lacks parameter receptacles for the information contained in the tritium observations). That is, the sensitivity of the forecast of interest to uncertain parameters that were conditioned by tritium concentration observation led to forecast bias. This occurs due to two factors: 0. the "Firth" nitrate-load forecast aggregates flow paths across the entire domain (i.e., this forecast represents the only nitrate flow sink of the system) and in time; and 1. the tritium observations provide insight into spatially and temporally averaged recharge and lateral flux rates in the upgradient portion of the domain, where most of the surface-water/groundwater exchange occurs. In other words, the bias reflects the information content of the upgradient tritium observations related to averaged upgradient model parameters on which the forecast is sensitive. To address this comment, we will add text regarding the nature of the forecast to the Second case study section, and this explanation to the Discussion and Conclusions section.

3. We agree that exploration of the conservativeness or otherwise of simplified model forecast PDFs (the former of which can be viewed as a metric for accepting such a simplified model) is an important undertaking, especially when performed with respect to a specific management decision threshold. We are pleased to inform Mr Turnadge that we have two papers that explicitly tackle this question—the first in terms of model parameterization (Knowling et al., 2019) and the second in terms of model vertical discretization (White et al., forthcoming; we will send this manuscript to Mr Turnadge). We will therefore address this comment by adding an explicit reference to these manuscripts in the Discussion and Conclusions section. We also refer Mr Turnadge to Prof. Ferre's related question in his review of the manuscript regarding how consideration of a management decision threshold may provide a more appropriate basis for assessing data worth from a decision maker's perspective; we will respond to this comment shortly.

Response to minor comments:

1. We agree with Mr Turnadge regarding the simplicity and lack of physical basis of LPMs, and the potential benefits of full advective-dispersive (and reactive) transport numerical models for simulating tracer concentrations (as described comprehensively by Turnadge and Smerdon, 2014), notwithstanding their practical limitations, which are amplified in formal decision-support modeling contexts. Importantly, we reiterate here that LPM-derived MRT observations are only used (in combination with advective-transport simulations) in the first case study; the second study employs full advective-dispersive transport modelling together with a first-order reaction rate to simulate radioactive decay of tritium. Our intention here was to employ "standard practice" tracer modeling techniques as a basis for exploring the ramifications of model tracer-data assimilation, such that the findings are as useful as possible to industry. The literature reflects the common use of both advective-only particle-tracking simulations (combined with LPM-based "age" observations) and advective-dispersive simulations (combined with tracer concentrations) (e.g., Gusyev et al., 2014). We will address this comment by presenting the above explanation and justification in the revised manuscript, and also by making explicit mention of reactive-transport modeling approaches as one means of increased model complexity that may facilitate improved imperfect model-data assimilation.

2. The MRT observations in the first case study were derived using a combination of exponential piston flow models (EPMs) and BMMs (comprising two "parallel" EPMs), as described in detail in Morgenstern et al. (2018). For the EPMs, Morgenstern et al. (2018) states "For wells with a long well screen interval in unconfined conditions, a high fraction of exponential (mixed) flow of 80–95% was applied. For wells with a narrow screen interval in confined conditions, a low fraction of exponential flow of 50–60% was used". BMMs were employed for most of the drinking-water wells where long time-series data are available for multiple tracers and where an adequate LPM fit (to different tracer signals) could not be obtained on the basis of a single EPM (e.g., due to

complex geological features). We will add these details to the First case study section of the manuscript.

3. We agree with the need to be more specific regarding atmospheric tritium concentrations. We thank Mr Turnadge for his suggested revision to the text. We will revise the manuscript directly following his suggestion.

4. The second case study does in fact use a first-order decay rate to simulate the process of denitrification reactions. We consider this approach to be “standard” modeling practice. We address this comment by making this point clear in the Second case study section of the revised manuscript. Note that we do not simulate the full range of nitrate reactive processes explicitly simply due to computational resource constraints; these constraints become more limiting where model deployment is undertaken stochastically for decision-support purposes.

5. We agree with Mr Turnadge regarding the importance of the well screen length for tracer interpretation—especially when using LPMs for interpretation. As described in response to minor comment (2), we will indicate the role that well screen lengths had on the deployment of LPMs to infer MRT in the First case study section.

6. We will add a description to both case study sections of the revised manuscript listing the parameterization device (e.g., spatially uniform, zones, pilot points, grid-based) employed for each parameter type. Briefly, for the Heretaunga Plains case study, pilot points were used to parameterize hydraulic conductivity (horizontal and horizontal-vertical anisotropy ratio), effective porosity, specific storage and specific yield, while river-bed and boundary conductance parameters are defined on a reach and zone basis. For the Hauraki Plains case study, the parameterization approach has already been described in detail in White (2018) and Knowling et al. (2019); we will therefore add only a brief summary to the revised manuscript along with a reference to these papers.

7. We will address this comment through indicating which spatially distributed parameters are represented by which variograms. Briefly, for the Heretaunga Plains case study, the same variogram (variogram parameters already defined in manuscript) is used for pilot-point based distributed parameters (only subsurface property-related parameters); no spatial correlation is assumed otherwise. For the Hauraki Plains case study, variogram details regarding the various different parameter types in the second case study have already been described in White (2018) and Knowling et al. (2019); we will therefore add only a brief summary to the revised manuscript along with a reference to these papers. On a more general note, we agree that variograms could theoretically be defined on a parameter type-specific basis from a physically-based (or perhaps more appropriately a “physically-motivated”) parameter standpoint. However, given our recent experience and findings regarding the significant potential for ill-effects (e.g., forecast bias) in real-world decision-support modeling (e.g., Knowling et al., 2019; White et al., forthcoming), we tend to consider spatially distributed parameters employed by regional-scale models to be significant “abstractions” (i.e., from their intended property representation; e.g., Watson et al., 2013). It follows that questions such as “what is the variogram for spatially distributed recharge bias-correction parameters?” and “how do we represent uncertainty in the variogram model used to describe prior parameter correlation and heterogeneity (i.e., a “hyper-parameter”)?” arise when trying to rigorously deal with real-world model error.

8. On Line 55, we are making only a general statement that data assimilation through history matching an imperfect model can result in forecast bias and uncertainty underestimation; these ill effects occur as a result of both model simplification and history matching. We cite literature that demonstrate these phenomena. This statement does not relate to tracers or tritium specifically, or any other data in particular. Therefore, no comments can be made at this point as to the nature of bias or variance underestimation. The “direction” of tritium assimilation-induced bias (i.e., under- or over-estimation) and is covered in the Results section of the second case study (although we note that the direction of bias may not be very generalizable between different forecasts and between different sites). Nevertheless, we will revise this sentence to unpack this

sentence by defining the terms “first moment” and “second moment” explicitly here.

9. The ensemble size was in fact selected on the basis of an approximation of the solution space dimensionality. This approximation was obtained through a subspace analysis of predictive error variance (Moore and Doherty, 2005). We refer Mr Turnadge to the Supplementary Material of Knowling et al. (2019) for more information on this, including a plot of the singular value spectrum.

10. Fewer relatively long flow paths occur when vertically coarsening the model grid simply due to the aggregation of numerical discretization effects—the flow paths of a coarser-layer model will be a smoother and averaged representation of those derived from a finer-layer model. We agree that this is an important explanation to support the current findings. As described above, the ramifications of model simplification in terms of reduced vertical discretization in the uncertainty quantification and data assimilation context more generally is covered by the separate manuscript White et al. (forthcoming). We will nevertheless build on the brief explanation provided in the current manuscript for completeness.

Many thanks again to Mr Turnadge for his helpful comments.

Matt Knowling

References:

1. Gusyev, M., Abrams, D., Toews, M., Morgenstern, U., and Stewart, M.: A comparison of particle-tracking and solute transport methods for simulation of tritium concentrations and groundwater transit times in river water, Hydrology and Earth System Sciences, 18, 3109, 2014

2. Knowling, M. J., White, J. T., and Moore, C. R.: Role of model parameterization in risk-based decision support: An empirical exploration, Advances in Water Resources, 128, 59 – 73, https://doi.org/10.1016/j.advwatres.2019.04.010, 2019

3. Moore, C., and Doherty, J.: Role of the calibration process in reducing model pre-

dictive error, Water Resources Research 41(5), 1-14, 2005

4. Morgenstern, U., Begg, J., van der Raaij, R., Moreau, M., Martindale, H., Daughney, C., Franzblau, R., Stewart, M., Knowling, M., Toews, M., Trompetter, V., Kaiser, J., and Gordon, D.: Heretaunga Plains aquifers : groundwater dynamics, source and hydrochemical processes as inferred from age, chemistry, and stable isotope tracer data, https://doi.org/10.21420/G2Q92G, 2018

5. Schilling, O. S., Cook, P. G., and Brunner, P.: Beyond Classical Observations in Hydrogeology: The Advantages of Including Exchange Flux, Temperature, Tracer Concentration, Residence Time, and Soil Moisture Observations in Groundwater Model Calibration, Reviews of Geophysics, 57, 146–182, https://doi.org/10.1029/2018RG000619, 2019

6. Turnadge, C., and Smerdon, B.D.: A review of methods for modelling environmental tracers in groundwater: Advantages of tracer concentration simulation, Journal of Hydrology, 519, 3674-3689, https://doi.org/10.1016/j.jhydrol.2014.10.056, 2014

7. Watson, T. A., Doherty, J. E. and Christensen, S.: Parameter and predictive outcomes of model simplification, Water Resources Research, http://dx.doi.org/10.1002/wrcr.20145, 2013

8. White, J.T.: A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions, Environmental Modelling and Software, 109, 191-201, https://doi.org/10.1016/j.envsoft.2018.06.009, 2018

9. White, J. T., Knowling, M. J., and Moore, C. R.: Consequences of model simplification in risk-based decision making: An analysis of groundwater-model vertical discretization, Groundwater, accepted subject to minor revisions, forthcoming

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-436, 2019.