Environmental Modelling & Software 77 (2016) 95–107

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Evaluating the sources of water to wells: Three techniques for metamodeling of a groundwater flow model

Michael N. Fienen [a, *], Bernard T. Nolan [b], Daniel T. Feinstein [c]

[a] U.S. Geological Survey Wisconsin Water Science Center, 8505 Research Way, Middleton, WI 53562, USA
[b] U.S. Geological Survey Office of Water Quality, 12201 Sunrise Valley Drive, Reston, VA 20192, USA
[c] U.S. Geological Survey Wisconsin Water Science Center, University of Wisconsin-Milwaukee Lapham Hall, Geosciences Dept., Room 338 3209 North Maryland Avenue, Milwaukee, WI 53211, USA

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form
23 November 2015
Accepted 25 November 2015
Available online 24 December 2015

Keywords:
Metamodeling
Groundwater
Bayesian networks
Artificial neural networks
Gradient boosted regression trees
Prediction

ABSTRACT

For decision support, the insights and predictive power of numerical process models can be hampered by insufficient expertise and computational resources required to evaluate system response to new stresses. An alternative is to emulate the process model with a statistical "metamodel." Built on a dataset of collocated numerical model input and output, a groundwater flow model was emulated using a Bayesian Network, an Artificial neural network, and a Gradient Boosted Regression Tree. The response of interest was surface water depletion expressed as the source of water-to-wells. The results have application for managing allocation of groundwater. Each technique was tuned using cross validation and further evaluated using a held-out dataset. A numerical MODFLOW-USG model of the Lake Michigan Basin, USA, was used for the evaluation. The performance and interpretability of each technique was compared pointing to advantages of each technique. The metamodel can extend to unmodeled areas.

Published by Elsevier Ltd.

1. Introduction

Numerical models are powerful tools for decision-making, allowing managers to evaluate potential outcomes of new stresses (for example, new high-capacity groundwater extraction well impacts on headwater stream ecosystems). Several challenges make it difficult for numerical models to be used in some assessments. In particular, the trade-off between resolution and computational effort often means that a model to evaluate a new stress covering a large enough area with sufficient detail requires more computational effort than is practically available. This challenge can be met by making available smaller-scale models that can be adjusted for new stresses, particularly if served online (e.g. Jones, 2012). Another approach is to simplify the physics of the problem using simplified models or analytical solutions that can be solved quickly (e.g. Hamilton and Seelbach, 2011). We seek an intermediate approach in which the insights from a fine-scaled but regional numerical model can be summarized as a statistical model—a "metamodel" (Blanning, 1975)—which can make predictions nearly instantly, albeit with less precision and certainty than the numerical model. The concept is to generate model outputs simulating a variety of conditions, treating those model results as data, and then training a statistical model to those data such that predictions can be made with only the statistical model. The dataset of model outputs can be generated using a Monte Carlo approach to systematically vary parameter values (e.g. Nolan et al., 2012) and run the model many times with those varied parameter sets. Alternatively, a sufficiently large sample set may be generated from few model runs if enough variability among input parameters is represented over time and space in a single model parameterization (e.g. Fienen et al., 2013). The latter approach is adopted in this work where the source of water to groundwater wells makes up the data of interest, as evaluated in a numerical groundwater flow model. By simulating several wells (and insuring that they are far enough from one another as to not interact in a single model run), a sample of several hundred data points can be obtained from a single run of the numerical model. This allows variability of the system to be sampled from natural system variability (as implemented in the numerical model) rather than through varying model input parameters.

In temperate regions such as the upper Midwest of the USA,

* Corresponding author.
E-mail addresses: mnfienen@usgs.gov (M.N. Fienen), btnolan@usgs.gov (B.T. Nolan), dtfeinst@usgs.gov (D.T. Feinstein).

http://dx.doi.org/10.1016/j.envsoft.2015.11.023
1364-8152/Published by Elsevier Ltd.

shallow groundwater is closely connected with lakes, wetlands, and streams (Winter et al., 1998). Groundwater discharge is important both in supplying water and moderating temperatures and geochemical conditions that maintain ecosystem functions. Groundwater pumped for human use typically is supplied at the expense of discharge that, under natural conditions, would supply surface water. The fact that pumping results in a deficit in discharge rather than a connection with recharge is sometimes misunderstood resulting in the famous “myth of safe yield” (Bredehoeft, 1997). In most cases, increased pumping does not change recharge, so for mass balance to be honored, the decrease in discharge is the main impact of pumping. This renders the recharge rate largely irrelevant when answering the question of whether a groundwater well will impact surface water or not (Barlow and Leake, 2012).

Natural resource managers need reliable models to predict surface water impacts due to installation of new groundwater pumping wells. Our modeling efforts are focused on this outcome, consistent with the best practice outlined by Jakeman et al. (2006).

In the Lake Michigan Basin (LMB) in the upper Midwest, USA, a large regional, steady-state, groundwater model (Feinstein et al., 2010) covers 215,000 km^2 mostly in the states of Wisconsin and Michigan, and minor portions of surrounding states (Fig. 1.1). Surface water features are represented in the groundwater model using the MODFLOW River (RIV) package. The spatial coverage of this model is well-suited to the needs of the upper Midwest and explicit simulations using the model can answer many important management questions. The challenges outlined above of long runtimes and the need to consider uncertainty remain. First, the original scale of the regional model is too coarse for the level of detail needed to evaluate local-scale impacts. We focused on the nearfield region of the model, illustrated in Fig. 1.1. Taking advantage of a new unstructured-grid version of the MODFLOW code used for the regional model (Panday et al., 2013), Feinstein et al. (2015) created a “semi-structured” version that collapses multiple deep layers into one—ultimately reducing from 19 to 4 layers—and laterally refines the model grid in the shallow system. This allows local surface-water impacts to be evaluated with what is still a regional model. Another major challenge, however, is not mitigated—the run time for a single simulation remains high. In order to evaluate the impact of a single well, one model run is performed - thus, evaluating the impact of many potential well locations requires many model runs, each at a potentially large computational expense. One strategy to overcome the issue of computational expense is through metamodeling (e.g. Fienen et al., 2015).

Razavi et al. (2012) and Asher et al. (2015) provide detailed reviews of metamodeling (a.k.a. surrogate modeling) techniques in various hydrologic applications. The goal of our work is to evaluate three techniques for creating metamodels and to compare their characteristics and performance. We chose these three techniques to evaluate, each with different strengths. One propagates uncertainty through to predictions (Bayesian Networks, BNs), another is a black box (Artificial neural networks, ANNs), and the third is an efficient technique with potentially better performance (Gradient Boosted Regression Trees, GBRTs). BNs and ANNs have recently been used as metamodels for groundwater applications (e.g. Fienen et al., 2013; Nolan et al., 2012) although GBRT has not, to our knowledge, previously been applied as a metamodeling technique for groundwater modeling. ANN and GBRT require a separate metamodel for each prediction of interest while BN is able to predict multiple outcomes using a single metamodel.

In the next section, we discuss the generation of the datasets on which to train the metamodels. Following that, we describe the three techniques evaluated and then their relative performance in cross validation, hold out prediction, and filling in unsampled areas. Filling in unsampled areas shows the power of these techniques in providing resource managers with information about susceptibility of streams to stress by nearby groundwater extraction without needing to run a numerical model. Managers can use this information in screening applications for water extraction to rule out many clear cases and focus their effort on borderline situations for proposed supply wells, and have a way to focus on those which potentially could have a substantial effect on surface water flows.

2. Causal relationships and variables

The application and evaluation of the three techniques described in this work focused on the source of water-to-wells application described in the introduction section. A sample of sources of water for 4911 hypothetical wells drawn from the MODFLOW-USG model makes up the dataset on which metamodels were built. Each sample includes an instance of each input variable and each output variable, as defined in Table 1 and calculated using the methodology outlined in Feinstein et al. (2015). The input variables were chosen as variables that are expected to have predictive power for the output of interest. Surface water features were intersected with a uniform grid with cells 500 ft on a side in the shallowest model layer. For each sample ”seeded” well, input distances were calculated from the hypothetical well location to the nearest grid cell centroid containing a surface water feature of the type of interest (e.g. first order, second order, or third and higher orders). Surface water density (e.g. percent surface water) in the local area were calculated as the number of grid cells containing the relevant type of surface water feature divided by the total number of grid cells in the local area. It is important, at this stage, to be precise in how we characterize the output of interest. All output variables involve changes in water budget due to simulation of a hypothetical ”seeded” well. These budgets are calculated as the change in flux of surface water features with or without simulation of seeded wells. More detail on these processes is provided in Feinstein et al. (2015). Pre-existing wells are present in the simulation, but the goal of our analysis is to quantify potential surface water impact due to the introduction of a new well. As a result, the superposition of a base model run (including existing wells) and a new run (the only changes to the system being the introduction of new hypothetical well locations) allows for calculation of the incremental impact of the new wells. In this way, we are supporting a decision-making strategy for managing additional groundwater extraction, not managing existing extraction.

The general issue laid out above is clear—namely, in temperate regions, managers acknowledge that groundwater pumping impacts on surface water features is important. However, a range of terminology can cloud the precise analysis being made. We seek to quantify “the source of water-to-wells” as the amount of water pumped by a well that is intercepted discharge or induced (reversed) flow from a surface water body. We parse the sources into categories of first, second, or third and higher order streams. This parsing is important to make the distinction between fragile (e.g. first order streams) and robust (e.g. third and higher order streams) ecosystems. This concept is consistent with previous work on “surface water depletion” where the mass balance is centered on the well (e.g. Barlow and Leake, 2012; Leake et al., 2013). Another perspective would be to evaluate impacts with mass balance centered on a surface water feature. In that case, we would refer to “baseflow reduction” which we define as the decrease in baseflow supplied to a surface water feature due to pumping in a well nearby. This subtle distinction has important ramifications on how a model is constructed and sampled, and how metamodels are generated. It is for these reasons, that in this work we focus on surface water sources supplying wells (e.g. depletion). This definition must be