Environmental Modelling & Software 53 (2014) 35–52

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Stochastic reconstruction of paleovalley bedrock morphology from sparse datasets

J.C. Castilla-Rho [a,c,d,*], G. Mariethoz [a,c,d], B.F.J. Kelly [b,c,d], M.S. Andersen [a,c,d]

[a] *School of Civil and Environmental Engineering, University of New South Wales, Sydney, NSW 2052, Australia*
[b] *School Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia*
[c] *Connected Waters Initiative Research Centre, University of New South Wales, Sydney, NSW 2052, Australia*
[d] *Affiliated with the National Centre for Groundwater Research and Training (NCGRT), Australia*

A R T I C L E I N F O

Article history:
Received 19 March 2013
Received in revised form
25 August 2013
Accepted 29 October 2013
Available online 7 December 2013

Keywords:
Geological uncertainty
Stochastic hydrogeology
Geological analog
Geostatistics
Spatial interpolation

A B S T R A C T

Stochastic groundwater models enable the characterization of geological uncertainty. Often the major source of uncertainty is not related to aquifer heterogeneity, but to the general shape of the aquifer. This is especially the case in paleovalley-type alluvial aquifers where the bedrock surface limits the extent of easily extractable groundwater. Determining the shape of a bedrock surface is not straightforward, because it is typically non-stationary and defined by few data points that are generally far apart. This paper presents a new workflow for the stochastic reconstruction of bedrock surfaces using limited datasets that are typically available for aquifer characterization. The method is based on a lateral propagation of basement cross-sections interpreted from geophysical surveys, and conditions the reconstructed surface to existing well-log data and digital elevation model. To alleviate the typical limitations of sparse data, we use an analog approach to incorporate prior geological knowledge. We test the methodology on a synthetic example and a dataset from an alluvial aquifer in Northern Chile. Results of these case studies show that the algorithm is capable of enforcing the general notion of structural continuity, with the aquifer shape being conceptualized as an elongated, continuous and connected valley-shaped body. Our method captures the large-scale topographic features of fluvial incision into bedrock and the uncertainty in the positioning of the surface. Small-scale spatial variability is incorporated using Sequential Gaussian Simulation informed by geological analogs. Being stochastic, the methodology allows characterization of the uncertainty associated with positioning of the bedrock surface, by generating an ensemble of models via a Monte-Carlo analysis. This makes it possible to quantify the uncertainty associated with estimating the aquifer volume. We also discuss how this methodology may be used to better quantify the influence of uncertainty associated with defining the aquifer geometry on water resource assessment and management.

Software availability

Name of software: stochastic paleovalley interpolation V 1.0
Availability and cost: freeware downloadable from: https://github.com/juancastilla/Paleovalley-Modelling.git including documentation and demo datasets. The program is available as a set of Matlab functions and scripts. SGEMS and mgstat are open-source software and can be downloaded at no cost from the developers' website.
Developers: Juan Carlos Castilla-Rho, Gregoire Mariethoz
Contact address: School of Civil and Environmental Engineering, University of New South Wales, Kensington Campus, Sydney, NSW 2052, Australia
Phone: +61 (0) 478074291
Email: j.castilla@unsw.edu.au
Hardware required: 32- or 64-bit PC with Windows or Mac OS. We recommend a high-speed processor and at least 4 GB of RAM.
Software required: Matlab R2012b, mGstat geostatistical toolbox (http://mgstat.sourceforge.net), SGEMS geostatistical modelling software (http://sgems.sourceforge.net)

* Corresponding author. School of Civil and Environmental Engineering, University of New South Wales, Kensington Campus, Sydney, NSW 2052, Australia. Tel.: +61 (0) 478074291.

E-mail addresses: j.castilla@unsw.edu.au (J.C. Castilla-Rho), gregoire.mariethoz@unsw.edu.au (G. Mariethoz), bryce.kelly@unsw.edu.au (B.F.J. Kelly), m.andersen@unsw.edu.au (M.S. Andersen).

1364-8152/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envsoft.2013.10.025

1. Introduction

The lack of spatial information is a prevalent issue for scientists and practitioners needing to assemble 3D models of geological regions. While carefully designed studies should rely on dense data collection campaigns, in practice economic and technical constraints often result in poor control of the quality and spatial arrangement of the data. Hydrogeology is particularly affected by this problem, as information mainly originates from localized data that sample a very small portion of the geological region of interest. In addition, subsurface morphology is usually heterogeneous, anisotropic and non-stationary. Furthermore, geoscientists can produce a range of interpretations from a single dataset, hence introducing knowledge bias and interpretational uncertainties (Bond et al., 2007). For all these reasons, geological uncertainty is often large, and significant research efforts are aimed at quantifying and capturing it for 3D structural, volumetric, facies and flow modeling applications.

Alluvial aquifers or basins are often modeled as a filling sequence overlying a bedrock surface (Hoyos et al., 2012; Jaireth et al., 2010; Whiteley, 2005). Their conceptualization is therefore controlled by the shape and geomorphology of the underlying structure, typically defined by a paleovalley bedrock surface. The aquifer then constitutes the permeable units of the unconsolidated sediments that overlie low permeability consolidated deposits or bedrock. However, the inherent limitations of geological datasets can make the task of characterizing bedrock surfaces challenging. This has significant implications for defining the transmissivity distribution in the aquifer or even for simpler characterization such as estimating the total volume of porous material in the reservoir.

Geostatistics is a widely used framework for making predictions at unmeasured locations from limited and sparsely arranged data. Geosciences and water resources rely heavily on geostatistics for spatial interpolation, with frequent use of kriging in its various forms (Li and Heap, 2011). A limitation of kriging is that it relies on assumptions of stationarity and smoothness, and limited representation of anisotropy (such as zonal anisotropy). Although traditional geostatistical methods can include data sources such as digital elevation models (DEMs), wells and geophysics, it is often found that representing complex geometries of geological systems is difficult (Neuweiler and Vogel, 2007; Zinn and Harvey, 2003). These issues are general and arise when working on structures having characteristics of non-stationary and meandering geometries, such as 3D groundwater models demanding realistic bedrock surfaces. Some of the approaches that have been proposed to model geological complexity include multiple-point statistics (Guardiano and Srivastava, 1993; Hu and Chugunova, 2008; Mariethoz and Kelly, 2011), object-based methods (Haldorsen and Chang, 1986) or complex forms of indicator simulation such as the pluriGaussian method (Le Loc'h et al., 1994; Mariethoz et al., 2008). Although these methods are appropriate for 3D gridded facies models, they are not always applicable to modeling geological surfaces, which are continuous and strongly non-stationary.

The literature on spatial interpolation of topographic datasets in meandering, non-stationary valley-shaped landscapes presents a series of solutions and customizations to improve the performance of simple interpolators when applied to such data (Nordfjord et al., 2005). One successful approach has been to transform the datasets to a channel-oriented coordinate system prior to interpolation (Legleiter and Kyriakidis, 2008; Merwade, 2009; Merwade et al., 2006, 2005). This coordinate conversion, however, is more appropriate for dense datasets and requires accurate definition of the channel centerline, often a time-intensive task (Goff and Nordfjord, 2004). Unfortunately, in hydrogeology the data is frequently too sparse for this method to be successful. Another general principle often used to improve topographic reconstructions is the concept of separating trend from the data using empirical functions. This idea has been the focus of recent studies related to river bathymetry modeling (Legleiter and Kyriakidis, 2008; Merwade, 2009), allowing subsequent application of isotropic interpolation. These methods still present challenges, particularly in addressing non-stationarity (Merwade, 2009) and their suitability to site-specific data. Even though some studies derive generic trend functions using modern analogs (Allan James, 1996), their applicability to different settings should be dealt with caution. For instance, the morphological differences between fluvial and glacial environments may invalidate the applicability of a method derived for a specific environment (Anderson et al., 2006; Graf, 1970; Li et al., 2001). Therefore, a general methodology is needed to fully extract trends from available data, avoiding loss of information typically occurring during curve-fitting procedures. Also relevant to this work, is the use of random fields through Sequential Gaussian Simulation (SGS). Gringarten et al. (2005) use this approach to simultaneously confer realistic geological features to topographic models, while conditioning to well-log data. As a by-product of this workflow, several possible realizations are obtained which can be used to convey quantitative measures of uncertainty.

In recent years, many authors have demonstrated that aquifer conceptualization is responsible for a significant proportion of the uncertainty in groundwater model predictions (Bond et al., 2007; Bredehoeft, 2005; Neuman, 2004; Poeter, 2007; Rojas et al., 2010; Zeng et al., 2013). To date however, the specific problem of generating 3D reconstructions of paleovalley geomorphology with sparse data has received little attention. The goal of this paper is to model uncertainty related to the alluvium/bedrock interface caused by limited subsurface information, which is often an important source of predictive uncertainty (Poeter, 2007; Refsgaard et al., 2012). In this regard, multi-realization methodologies have a proven track-record in the quantification of uncertainty due to data limitation and other factors (Refsgaard et al., 2012; Troldborg et al., 2007). This last point highlights the need for specialized geostatistical algorithms capable of generating multiple realizations of alluvium–bedrock interfaces.

In this paper, we address the challenge of integrating, in an automated workflow, the various pieces of information available for the characterization of alluvial-aquifer bedrock surfaces using a stochastic multi-realization approach. We adopt a hybrid modeling framework (Bertoncello et al., 2013; Dubrule, 1993; Michael et al., 2010), which combines different approaches and data types to tailor the modeling workflow to a specific geologic environment. Fig. 1 illustrates the general problem of a bedrock surface overlain by valley-filling unconsolidated sediments, along with the different types of data typically available for a geological modeling. The data sources considered here are:

1. Lithological data from boreholes that intersect the bedrock, as well as those that do not intersect it but provide an estimate of the upper bound of the alluvium–bedrock interface,
2. Geophysical surveys providing cross-sections of the bedrock surface,
3. Digital elevation models (DEMs), and
4. Conceptualization of bedrock geometry as elongated, continuous and V-shaped, typical of fluvial valleys.

Each data source brings information of a different nature. Lithological well-logs inform the depth to bedrock quite accurately; however they are typically not numerous enough to