Journal of Geochemical Exploration 144 (2014) 63–73

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

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

# Reactive transport modeling of thorium in a cloud computing environment

Xiang Huang [a], Guoliang Cao [a], Jie Liu [a], Henning Prommer [b,c,d], Chunmiao Zheng [a,e,*]

[a] Center for Water Research, College of Engineering, Peking University, Beijing, China
[b] School of Earth and Environment, University of Western Australia, Crawley, WA, Australia
[c] CSIRO Land and Water, Floreat, WA, Australia
[d] National Centre for Groundwater Research and Training, Adelaide, SA, Australia
[e] Department of Geological Sciences, University of Alabama, Tuscaloosa, AL, United States

ARTICLE INFO

*Article history:*
Received 21 September 2013
Received in revised form 6 February 2014
Accepted 4 March 2014
Available online 19 March 2014

*Keywords:*
Thorium
Cloud computing
Tailings impoundment
Bayan Obo
Reactive transport modeling
Yellow River
Baotou

ABSTRACT

Despite the long standing interest in modeling the fate and environmental impacts of radionuclides, simulations addressing the fate and transport of rare earth elements (REEs) and thorium (Th) have received comparably little attention. This study presents an architecture that enables reactive transport modeling and parameter sensitivity analysis on cloud computing platforms. We adapted an existing groundwater modeling framework to perform some of the computationally most expensive steps within a cloud environment based on Microsoft Windows Azure. The cloud computing architecture was evaluated and validated through the development of a schematic, cross sectional model along a transect across a tailings impoundment at a REE mine tailings site in northwest China. The model framework employs a suite of flow, solute transport and reactive transport simulation tools, i.e., MODFLOW, MT3DMS, and PHT3D. On the basis of our model simulations, the collection-trench for the impoundment constructed above the ground surface appears to collect a substantial portion of the leachate fluxes, but the remainder will bypass the trench and migrate downstream. Those bypassed leachate fluxes will subsequently interact with downstream fluviolacustrine aquifers and eventually discharge into the Yellow River south of the study site under the idealized simulation environment. Further investigations of the hydraulic parameters of the aquifer system and the impoundment dam, and other geochemical characteristics are needed to elucidate the fate and transport of thorium and improve the reliability of the numerical model. Although the discussion and analysis of this study is tailored to thorium reactive transport modeling of a REE tailings impoundment, such a framework can also be applied to deploy different types of scientific modeling applications on Azure Cloud.

## 1. Introduction

Groundwater pollution due to tailings impoundment leakage has become and persisted as a serious worldwide environmental problem, especially where it is associated with acid mine drainage (Miekeley et al., 1992; Pan, 2010; Xu et al., 2005). Numerical modeling of such problems can provide an indispensible supplement to elucidate, quantify and predict the migration of contaminants in complex groundwater systems. Understanding and quantifying the mechanisms controlling radionuclide fate and transport in groundwater from tailings impoundments is vital for effective management of REE mine wastes, selection of potential disposal sites to store nuclear fuel wastes (Ashley et al., 2012; Ewing, 1999; Ma et al., 2012), remediation of accidental contamination problems, and for the protection of water resources (He et al., 2010; Xu et al., 2005). Although the specific toxicity of thorium is greater than uranium, its environmental fate is seldom reported, possibly because of its lower solubility. Thorium in natural groundwater prevails predominantly in a tetravalent redox state and is not redox sensitive. It is usually present in ultra-trace concentration and transported as a complexed species (LaFlamme and Murray, 1987; Langmuir and Herman, 1980). Like uranium, the fate of Th is significantly affected by hydrolysis and subsequent oligomerization reactions, which increase its solubility in water (Anderson et al., 1982; Kim et al., 2010). Therefore, it is possible that Th can be leached into groundwater and migrate towards nearby ditches or rivers from Th-rich REE tailings impoundments.

Human exposure and severe Th-pollution risks may occur through bioaccumulation in foods grown adjacent to the disposal sites or via direct consumption of water with elevated Th concentrations. Transport models incorporating a comprehensive suite of geochemical reactions (e.g., Ma et al., 2010; Prommer et al., 2003; Steefel and Lasaga, 1994) are required (i) for predicting the future long-term fate of the contaminants and (ii) to provide more detailed insights into the factors that control the contaminant behavior under the site-specific spatially and temporally varying hydrogeological and geochemical conditions.

---

* Corresponding author at: Center for Water Research, College of Engineering, Peking University, Beijing, China. Tel./fax: +86 10 62767687.
*E-mail address:* czheng@pku.edu.cn (C. Zheng).

http://dx.doi.org/10.1016/j.gexplo.2014.03.006
0375-6742/

Clearly, such modeling efforts have to be underpinned by detailed monitoring efforts and geostatistical analysis of the concentration distribution to characterize the extent of the contamination and to provide model calibration constraints. On the other hand, models can provide invaluable guidance for data collection and assist with the optimization of monitoring locations, sampling times and data types. Ideally, data collection and modeling are carried out as an integrated effort with a repeated iteration between both tasks.

While an extensive number of previous reactive transport modeling studies were dedicated to simulate the complex fate and transport characteristics of uranium (e.g., Ma et al., 2010, 2012; Olin and Lehikoinen, 1997), relatively little attention has been given to thorium. Thus this study represents one of the first attempts at simulating the reactive transport of aqueous Th and how it might be influenced by the physical and chemical conditions generated by REE tailings impoundments.

Based on the underlying models that describe the flow and conservative transport processes, reactive transport processes are successively invoked until a satisfactory description of the entire geochemical environment is achieved. Two important steps in the development of reliable reactive transport models are the model calibration step and a comprehensive parameter sensitivity analysis. The latter often plays an important role for the former as the effects of parameters on model outputs are evaluated through sensitivity analysis and used to update estimated parameters and to obtain optimized values. These steps require a large number of model runs, depending on the number of parameters to be evaluated and/or estimated. During the parameter estimation process many of the model runs can be performed completely independently (Hunt et al., 2010), while overall incurring substantial computational cost. When relying on traditional computing infrastructure, i.e., personal computers (PCs) the model calibration step and its associated computing times can often become a bottleneck.

The recent emergence of cloud computing has provided a unique opportunity to eliminate such bottlenecks by harnessing the unprecedented storage and high-performance computing resources for the groundwater community (Hunt et al., 2010; Langevin and Panday, 2012; Liu et al., 2012). To date the potential provided by parallel processing is rarely applied for the assessment of groundwater contamination by reactive transport modeling, despite the considerable computational burden (Arnett and Greenwade, 2000) and the fact that parameter estimation can easily be parallelized (Hunt et al., 2010). Cloud computing provides an ideal opportunity for such applications. It offers model users a suitable infrastructure and services from private computing environments to third-party data centers through the Internet (Armbrust et al., 2010; Yang et al., 2013). These various services are referred to as (1) software as a service, (2) platform as a service (PaaS), and (3) infrastructure as a service (IaaS) (Liu et al., 2012; Luchette et al., 2009). Specific examples of cloud computing services for the scientific community are scientific computing as a service (SCaaS) (Saripalli et al., 2011) and scientific modeling as a service (SMaaS) (Liu et al., 2012). Sun (2013), for example, described how service-oriented cloud computing can enable collaborative decision-making in watershed management by alleviating the technical burdens of Environmental Decision Support Systems. Cloud computing can also provide scientific software in connection with powerful computing resources and data storages (Kollet et al., 2011) by simultaneously supporting the massively distributed and millions of requests with elasticity, on-demand, and pay-as-you-go features (Fox, 2011; Yang et al., 2013). While numerous studies have already explored the feasibility of utilizing cloud computing for deploying geoscientific applications and first experiences on how to work efficiently and effectively on this new computing paradigm, to our knowledge no reactive transport modeling studies have benefited from cloud computing.

The present study is intended to demonstrate the applicability of cloud computing as an effective and promising tool to perform reactive transport simulations and sensitivity analysis, using thorium migration from a tailings impoundment as a representative application example. The remainder of this paper is structured as follows. First, we provide an overview of the emerging scientific cloud computing literature and specifically highlight recent research efforts for utilizing cloud computing in support of geoscientific applications. Then, we outline an architecture adapted from Subramanian et al. (2010) and ModflowOnAzure (Liu et al., 2011b, 2012) for using the Microsoft Windows Azure Cloud computing platform to implement the case study and evaluate its performance and reliability across desktop and cloud platforms. The case study involves a Th reactive transport modeling along a transect across a tailings impoundment in Baotou, China for which the fate and transport characteristics of leaching Th are not yet well understood. Next, we present a sensitivity analysis that was performed on the Azure platform in which the sensitivity of the net thorium mass-fluxes into the collection-trench to various hydrogeological parameters was investigated. Finally, we discuss some of the lessons and insights in developing our cloud computing framework for reactive transport modeling applications and point out possible future directions.

## 2. Scientific cloud computing

Like nebulous ‘cloud’ itself, the definition of “cloud computing” in the literature is still incomplete and confusing (Foster et al., 2008; Vaquero et al., 2009). From a highly cited article by Armbrust et al. (2010), “cloud computing refers to both the applications delivered as services over the internet and the hardware and systems software in the data centers that provide those services”. Cloud computing represents a new computing paradigm and more reliability–flexibility–scalability than previous grid computing and high-performance computing (HPC), but still inherits many characteristics of traditional computing. For more detailed discussions on differences among them, see Foster et al. (2008), Bias (2010) and Ebejer et al. (2013). The recent surge in cloud computing is leading to a transition towards a new computing paradigm that will profoundly affect every aspect of our social life and economy (Armbrust et al., 2010; Yang et al., 2013). Concurrently, research efforts, including geoscientific applications, are increasingly relying on the advantages of cloud computing. Examples of these applications include CloudClustering (Dave et al., 2011), CloudGIS (Yang et al., 2013), CloudPEST (Fienen et al., 2011), synthetic seismograms and seismic source inversion (Subramanian et al., 2010), computational-intensive molecular modeling (Ebejer et al., 2013), comparative genomics in bioinformatics (Kim et al., 2012), calibration of watershed modeling (Humphrey et al., 2012) and uncertainty analysis of flood modeling in hydraulic studies (Quiroga et al., 2013).

Commercial cloud platforms such as Amazon EC2/AWS, GoGrid, Windows Azure, Google AppEngine/GCE and Rackspace Mosso are now available to hydrogeological and environmental researchers and practitioners. Luchette et al. (2009) solved two Parallel PEST parameter estimation problems by using virtual machines (VMs) on GoGrid. Kollet et al. (2011) implemented the integrated hydrologic simulation platform ParFlow using a cloud computing web service, which included a web interface, optimization capabilities (e.g., PEST, EnKF) and basic visualization utilities. Saripalli et al. (2011) presented a numerical simulator for the subsurface fluid and heat transport model (TOUGH2) on the VM-based Amazon AWS, which helps users to access data-intensive visualizations via a web-browser.

Compared with other cloud computing platforms, Microsoft Windows Azure has been demonstrated to be one of the most effective cloud platforms (Gannon and Vargas, 2012; Kim et al., 2012; Liu et al., 2012). There are numerous data- and computationally intensive scientific applications on Azure such as AzureBlast (Lu et al., 2010) and Twister4Azure (Gunarathne et al., 2013). Liu et al. (2011b) developed an easy-to-use cyber-enabled MultiModel computing framework for groundwater risk analysis on Azure. MODISAzure is a large-scale satellite image processing system, which implements the MODIS data-intensive reprojection and reduction pipeline in Azure (Humphrey et al., 2012). Liu et al. (2012) developed and utilized ModflowOnAzure for groundwater modeling uncertainty analysis and showed that the