Environmental Modelling & Software 25 (2010) 1769–1780

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

Flow guided interpolation – A GIS-based method to represent contaminant concentration distributions in groundwater

M. Morio [a,*], M. Finkel [a], E. Martac [b,c]

[a] Universität Tübingen, Center for Applied Geoscience, Sigwartstr. 10, D-72076 Tübingen, Germany
[b] tgag–Tübinger Gesellschaft für Angewandte Geowissenschaften e.V., Sigwartstr. 10, D-72076 Tübingen, Germany
[c] geo-log GmbH, Bahnhofstraße 9, D-72147 Nehren, Germany

ARTICLE INFO

Article history:
Received 25 November 2009
Received in revised form
11 May 2010
Accepted 24 May 2010
Available online 7 July 2010

Keywords:
Interpolation
Groundwater contamination
Contaminant transport
Chlorinated hydrocarbons
TCE
GIS
Particle tracking
Ordinary kriging
Numerical groundwater flow modeling

ABSTRACT

This paper introduces a new interpolation method to estimate the spatial distribution of contaminant concentrations in groundwater. The method is intended to identify areas of risks in early investigation stages when groundwater sampling data is typically scarce and available interpolation methods fail to provide reasonable results. As a consequence, the method does not only incorporate available sampling data, but also makes use of information about the groundwater flow field, in order to "guide" the interpolation with e.g. ordinary kriging or inverse distance method. The guidance includes the augmentation of available data by auxiliary point data and the segmentation of the estimated plume area into a series of sectors. The method is evaluated for several settings and different sampling data sets. Each data set reflects a specific level of field investigations at the model site, an abandoned military base in Potsdam near Berlin, Germany. The results reveal that flow guidance improves the representation of contaminant distribution for all cases examined in this study compared to "unguided" interpolation. These findings are underpinned by the results of the method's application to real sampling data. The method especially shows its strength when data of only a few sampling points are available.

1. Introduction

Different approaches exist to estimate the spatial distribution of contaminants dissolved in groundwater. Proposed approaches range from experience-based practitioners' methods, such as best guess delineation of source zones and plumes based on small amounts of sample data and experiences from similar sites and cases, to advanced methods involving geostatistics (e.g. Isaaks and Srivastasa, 1989; Deutsch and Journel, 1997; Michalak and Kitanidis, 2004a, 2005), modeling (e.g. Shlomi and Michalak, 2007; Prommer et al., 2002; D'Affonseca et al., 2008) and data assimilation techniques e.g. using Bayesian filtering techniques (e.g. Kalman, 1960; Chen, 2003). The appropriateness of individual methods is mostly dependent on the type, amount and quality of available data as well as on the particular objectives. Plume delineation focusing on detecting the plume's extent (e.g. Meyer et al., 1994; Storck et al., 1997; McGrath and Pinder, 2003) may call for methods other than the estimation of the concentration distribution (e.g. Boufassa and Armstrong, 1989; MacKay, 1990; Kerry and Oliver, 2007).

Data assimilation techniques such as Particle Filter and Ensemble Kalman Filtering have gained considerable interest in the last decade for utilization of available measurement data to update mathematical model predictions of groundwater flow and plume propagation (e.g. Evensen, 1994; Eigbe et al., 1998; Porter et al., 2000; Chang and Jin, 2005; Chang and Latif, 2009). Filtering methods seem to be best suited for transient problems of groundwater system state estimation when time series of measurements are to be repeatedly (i.e. sequentially) assimilated into mathematical models (e.g. Liu et al., 2008; Huang et al., 2009; Chang and Latif, 2010).

Geostatistical interpolation methods have been widely applied in the past decades. Mehrjardi et al. (2008) proposed the use of ordinary kriging and cokriging for interpolation of contaminants in groundwater, while Reed et al. (2004) evaluated different interpolation methods to estimate the distribution of perchloroethylene (PCE) in three heterogeneous test cases of different size and complexity for a groundwater plume with differing amounts of non-gridded sampling data. They recommend quantile kriging and

* Corresponding author.
E-mail addresses: max.morio@uni-tuebingen.de (M. Morio), michael.finkel@uni-tuebingen.de (M. Finkel), e.martac@geo-log.com (E. Martac).

1364-8152/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2010.05.018

multigaussian kriging to be most robust and least biased compared to ordinary kriging, intrinsic kriging and inverse distance weighted methods. Journel and Rossi (1989) showed that universal kriging (also called "kriging with trend") yields similar results to ordinary kriging on data sets with trend when the trend component is unknown and kriging is conducted in local neighborhoods for non-stationary data sets.

Cooper and Istok (1988) discusses the requisite of data preparation and analysis, including additivity, stationarity and amount of samples for the estimation using geostatistical interpolation methods. Sufficient data is required, e.g. to make use of an empirical semivariogram (e.g. Deutsch and Journel, 1997; Fuest et al., 1998; Kitanidis and Shen, 1996; Reed et al., 2004; Kistemann et al., 2008). A practical rule for the minimum amount of samples is given by Journel and Huijbregts (1978):

$$N(h) > 30 - 50 \quad \text{with} : |h| < (L/2), \tag{1}$$

where $|h|$ denotes the magnitude of separation vector h for N sample pairs and L stands for the longest dimension of the contaminant plume in the direction of h. For a complete list of symbols used in this article, please refer to Table A.5

Interpolation may utilize concentration measurements either in terms of point observation data from distributed monitoring networks (e.g. Sudicky et al., 1983; Warrick et al., 1998) or in terms of data from so-called control planes or monitoring fences (e.g. Schwarz et al., 1998; King and Barker, 1999; Bockelmann et al., 2001; Basu et al., 2006; Kübert and Finkel, 2006; Bayer-Raich et al., 2009).

Several suggestions were made to improve interpolation by including additional information in the estimation process, e.g. hydraulic gradient or head data from sampling campaigns (e.g. Burger and Schafmeister, 2000; Shlomi and Michalak, 2007). Other approaches are based on coupling numerical transport models to interpolation methods (Michalak and Kitanidis, 2004a, 2004b). Neupauer and Wilson (2004) used a probabilistic numerical flow and transport model to relate concentration measurements to possible upgradient source locations. Rautman and Istok (1996) and Istok and Rautman (1996) proposed stochastic geostatistical modeling of contaminant plumes as an approach to derive probabilities of having a contamination at a certain point with respect to a specific concentration threshold and probability cutoff.

A common problem in practice is that available contaminant concentration measurement data is not sufficient to make sensible use of the abovementioned geostatistical interpolation methods. Numerical transport modeling (e.g. Anderson and Cherry, 1979; Chu et al., 1987), which further requires comprehensive information about aquifer and transport properties, is also inappropriate for this reason e.g. (Batu, 2006). Limited availability of contaminant data is characteristic of early project stages in tiered decision-making procedures, when information on subsurface contamination only stems from historical data and some initial site investigation. This is particularly true, e.g. when extent and complexity of a site require an early identification and prioritization of focal areas and origins of risks in order to drive further decisions on detailed investigation programmes and remediation measures (compare Triad approach, e.g. in Crumbling et al., 2001; Mack et al., 2004; Critto et al., 2007; O'Reilly and Brink, 2006). Thus, especially for early site investigation stages, there is a need for enhancement and appropriate processing of sparse amounts of available data in order to produce the premise for a reasonable application of interpolation methods such as kriging.

In this paper we present a flow guided interpolation (FGI) method that has been specifically adapted to the type, scale, and information basis that is typically available in early stages of revitalization projects at contaminated sites. The method proposes to add extra sampling points in a standardized way through coupling of flow data and existing samples plus information about possible source zone extents to enable the application of kriging methods. The purpose of the proposed method is to bridge the gap between elaborate, data intensive approaches and subjective and often non-reproducible methods, which are frequently applied in practice if available data is scarce. The FGI method builds upon a groundwater flow model, assuming that basic information about the groundwater flow regime can be made available at relatively low cost. This model is not supposed to provide a highly sophisticated representation of the flow situation but is supposed to show the major i.e. characteristic features. Guidance by groundwater flow is intended to improve the interpolation especially if contaminant concentration data is scarce. The idea is to incorporate upgradient information in a sequential downgradient-moving interpolation procedure. The relevance of the quantity of available data is addressed through analyzing the FGI method's performance for different knowledge states, i.e. different sample amounts.

The remainder of this paper is organized as follows: the FGI method is described in section 2 by use of pseudo-code algorithms to explain the sequential procedure execution; section 3 gives a description of the model site in which the FGI method was applied; the evaluation of the FGI method for different parameter settings is discussed in section 4; results of the case study are presented in section 5; conclusions are made in section 6.

2. Methodology

The FGI method is implemented into the geographic information system ArcGIS (ArcMap, Version 9.1 or higher, (ESRI, 1992–2005) with VBA and ArcObjects (Razavi, 2004; Burke and Arana, 2003). The method is comprised of a sequence of procedures with major portions that are controlled by VBA modules.

In short, the following steps are performed: (1) delineation of the known or expected source zone of contamination using particles, which are equidistantly distributed along the source zone edges; (2) delineation of the plume fringe by tracing these particles advectively downgradient; (3) segregation of the plume into several sectors according to the plume's tortuosity and curvature; (4) employment of sector-wise flow guided interpolation, utilizing sampling data in the sector, as well as auxiliary sampling point data along the plume fringe and along the boundary to the previously processed sector; (5) merging of the results of all sectors into a single grid by mosaicking.

A series of three pseudo-code algorithms describe the process of (i) plume fringe delineation based on the previously defined source zone (Algorithm 1, see also Fig. 1a), of (ii) segmentation of the derived plume shape dependent on groundwater flow direction (Algorithm 2, see also Fig. 1b) and of (iii) sector-wise interpolation to estimate the spatial concentration distribution within the plume (Algorithm 3, see also Fig. 1c–g). The procedures are described in more detail below. Please note that multiple source zones and corresponding plumes can be considered. For the sake of clarity we limit the description to one source zone and one plume. The delineation of the plume fringe requires polygon data of the source zone, a numerical groundwater flow model, and a particle tracking module. The flow model is required to calculate the groundwater flow field, which in turn is required to calculate pathlines of groundwater flow by particle tracking. We used MODFLOW 96 (Harbough and McDonald, 1996) and MODPATH 3.0 (Pollock, 1994), respectively. Note that other models, e.g. a version of ESRI's Groundwater Modeling application (ESRI, 2009), could be used as well after some adaption work (see also the concluding discussion of this section further below). The identification of the source zone is based on desk work examining information on former use of the site, possible locations of contaminant spills, the geological and hydrological situation, and subsurface sampling information. Source zone polygons are then created manually as polygon shape files in GIS, based on the assembly of given data. As shown in Algorithm 1, the source zone polygons are converted into a set of equidistantly distributed points, which are then used as water particle starting locations for particle tracking with MODPATH. The pathline data returned by MODPATH is automatically converted to a polyline shape file consisting of polylines for each particle starting point and a corresponding point shape file bearing information on the pathline time steps. The pair of outermost polylines in the shape file represents the plume fringe and is converted to a polygon shape file. The resulting polygon has to be cut in order to account for the expected plume age. Using travel time information from the point shape file, a polygon is created that represents the plume extent corresponding to the given time after spill.