ground water

On Integrating Groundwater Transport Models with Wireless Sensor Networks

by Kevin Barnhart[1], Iñigo Urteaga[2], Qi Han[2], Anura Jayasumana[3], and Tissa Illangasekare[4]

Abstract

The emerging technology of wireless sensor networks (WSNs) is an integrated, distributed, wireless network of sensing devices. It has the potential to monitor dynamic hydrological and environmental processes more effectively than traditional monitoring and data acquisition techniques by providing environmental information at greater spatial and temporal resolutions. Furthermore, due to continuing high-performance computing development, these data may be introduced into increasingly robust and complex numerical models; for instance, the parameters of subsurface transport simulators may be automatically updated. Early field deployments and laboratory experiments conducted using in situ sensor technology and WSNs indicated significant fundamental issues concerning sensor and network hardware reliability—suggesting that investigations should first be conducted in controlled environments before field deployment. A first step in this validation process involves evaluating the predictive capability of a computational advection-dispersion transport model when incorporating concentration data from a WSN simulation. Data quality is a major concern, especially when sensor readings are automatically fed into data assimilation procedures. The appropriate employment of an independent WSN fault detection service can ensure that erroneous data (e.g., missing or anomalous values) do not mislead the model. Parameter estimation regularization techniques may then deal with remaining data noise. The primary purpose of this study is to determine the suitability of WSNs (and other in situ data delivery technologies) for use in contaminant transport modeling applications by conducting research in a realistic simulative environment.

Introduction

Very recently, the hydrology community has become increasingly interested in sensor networks that involve collecting hydrological data in situ without manual sampling and analysis. Although complex phenomenon such as subsurface contaminant transport may arguably still be undersampled, data are being made available as never before and it is anticipated that practitioners will soon be burdened with doing data-rich rather than data-poor research.

Wireless sensor networks (WSNs) are ad hoc networks of sensor nodes where in situ measurements (e.g., contaminant concentration, flow rate and direction, head, and temperature) may be gathered in real time at less cost than traditional sampling techniques. Each sensor node (aka, mote) contains a microprocessor and a wireless receiver/transmitter and is thus capable of sensing, processing, and broadcasting data via wireless transmission (usually radio frequencies). Encouragingly, there have been numerous advances in sensor network technology (Culler et al. 2004) along with the development of miniature sensors suitable for this application (e.g., in situ chemical sensors, Ho 2005).

[1]Corresponding author: Center for Experimental Study of Subsurface Environmental Processes, Division of Environmental Sciences and Engineering, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401; (303) 273-3483; fax: (303) 273-3413; kbarnhar@mines.edu

[2]Department of Mathematical and Computer Sciences, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401.

[3]Department of Electrical and Computer Engineering, Colorado State University, Engineering Room B104, 1373 Campus Delivery, Fort Collins, CO 80523.

[4]Center for Experimental Study of Subsurface Environmental Processes, Division of Environmental Sciences and Engineering, Colorado School of Mines, 1500 Illinois St., Golden, CO 80401.

Received December 2008, accepted January 2010.

doi: 10.1111/j.1745-6584.2010.00684.x

Though enticing, initial investigations into using WSNs for surface and subsurface monitoring suggest (sometimes implicitly) that sensor and network technology maturation is compulsory before water quality and quantity can be accurately monitored and predicted (Lundquist et al. 2003; Musăloiu-E. et al. 2006; Ramanathan et al. 2006; Porta et al. 2009; Trubilowicz et al. 2009). These studies invariably report the collection of considerable amounts of anomalous data due to drifting sensor calibration, faulty electronics, and unforeseen transient environmental conditions. Hence, although a WSN is capable of monitoring and measuring the environment at scales, locations and resolutions not before possible, it poses enormous data analysis and management challenges. Further investigations should be prudently conducted in controlled simulative and laboratory environments prior to field deployment where poorly quantified uncertainties often dominate results.

Increased availability of computer resources has been instrumental in the almost ubiquitous use of computational models to calculate health risks, design cleanup strategies, guide environmental regulatory policy, and determine culpable parties in lawsuits. Such use culminates in computational models dramatically influencing environmental decisions involving large investments of effort and funds (Eggleston and Rojstaczer 2000). Given this, it may be surprising that few studies exist that are designed to establish model credibility (Anderson and Woessner 1992; Hassan 2004). There is a growing belief that models will become more effective if model parameters are periodically updated or recalibrated as new data become available (Anderson and Woessner 1992; Kim et al. 2005). However, the effect that erroneous data may have on simulation results is largely overlooked—especially in light of the above pioneering WSN studies.

This paper focuses on the periodic inclusion of concentration data into a computational advection-dispersion transport model. The data are synthetically generated and a WSN simulator is used to inject authentic anomalies into the data set. Data assimilation is considered using four scenarios: (1) infrequently collected data with minimal noise; (2) daily WSN data containing only theoretical noise levels according to sensor specifications; (3) data containing WSN faults (including noise); and (4) data with fewer WSN faults due to the use of a WSN fault detection application. The presented methodologies employ modeling tools commonly accepted in industry (i.e., MT3DMS, Zheng and Wang 1999; MODFLOW-2000, Harbaugh et al. 2000; and PEST, Doherty 2005). Thus, the aim is to provide practical insight into the directions of future modeling efforts in a new data context.

The next section outlines specific challenges associated with assimilating noisy and faulty data into a computational transport model. A methodology is then proposed, implemented, and evaluated in the sections that follow. Finally, conclusions are drawn with respect to this study and its broader impacts.

Problem Description

The use of concentration data to estimate numerical transport model parameters can be traced back to the mid-80s. Chu et al. (1987) calibrated transmissivity in a two-dimensional (2D) finite difference model of transport using sparsely sampled synthetic data (with additive noise). Their results focused on model predictions that only proved valid for short periods of time. The work by Medina et al. (1990) is another noteworthy study based on a maximum likelihood approach. Their investigation showed that better model calibration may be achieved by including synthetic concentration values but performed no predictive analysis. "Various conceptual and numerical difficulties" were cited as the reason most previous research neglected concentration data during parameter estimation.

Among others, Harvey and Gorelick (1995) noted that parameter estimation is greatly simplified by using mean arrival time of the solute as it seemed to represent the critical information of the breakthrough curve (BC). Their conclusions focused on the accuracy of the estimated hydraulic conductivity fields rather than on the model's ability to forecast transport location. Ezzedine and Rubin (1996) concurred that measured concentration is inferior to travel time for efficient conditioning. Although using temporal moments and travel time may help stabilize parameter estimation methods, these studies inherently assume that sufficient concentration data are available to estimate the zeroth and first moments of the BC. Because of this, model predictions can only benefit from such statistics after the plume has already (mostly) passed borehole locations. This may be why most of these studies concentrate on conditioning quality and parameter/dataset reproducibility instead of model forecast capacity.

There is little evidence that transport model predictions have been accurate in the field (Hassan 2004) or even in fair numerical studies (Zimmerman et al. 1998). However, the advent of new sensing technologies may give modelers some hope of quality short-term forecasts. It appears that researchers may return to the problem originally presented by Chu et al. (1987) but now armed with: (1) significantly greater amounts of data; (2) at least a qualitative idea of the information that may be gathered from a BC; (3) many new parametrization, parameter estimation/inversion, and forward modeling techniques; and (4) drastically improved computing resources. Hendricks Franssen and Gòmez-Hernàndez (2003) is an example of such recent research.

Emerging sensing and network technology is also motivating researchers to develop data assimilation techniques, that is, readjusting model parameters (usually automatically) as new data become available. Kim et al. (2005) performed real-time model parameter estimation of simplified (diffusive-only) transport in porous media; Liu et al. (2008) assimilated concentration data from the MADE site to validate tritium concentrations. Graham and McLaughlin (1991) and McLaughlin et al. (1993) gave early examples of concentration data assimilation. Although the first study uses elementary models and the others do not involve sensor networks, such methodologies may prove important in the data context of WSNs.

As already mentioned, data collection technologies like WSNs are prone to network and sensor faults. Porta et al. (2009) and Ramanathan et al. (2006) give examples of the sorts of errors that may occur while monitoring subsurface transport. Herein, it is demonstrated that data assimilation and the possibility of useful model forecasts may be hindered by the presence of erroneous data but that some of these difficulties may be overcome by employing simple data handling techniques.

Methodology

This study highlights challenges that exist with assimilating WSN concentration data into an accepted model of subsurface transport. In this section, an accessible approach to data assimilation is outlined. A distributed WSN software algorithm is also summarized, which aids in identifying and removing erroneous data during the data assimilation process. The subsequent section actualizes the presented methodology in a synthetic example.

Transport Model Data Assimilation

Transport data assimilation methods such as that described by Liu et al. (2008) are not commonly used in practice where modelers usually: (1) manually update the parameter estimation process to recognize the new data and (2) constrain numerical model selection to those that have been "accepted." Questions pertaining to data quality/weighting may also be difficult, especially when the information content and rate of anomalies in the dataset are not known. The approach taken here attempts to address both of these issues: a practicable data assimilation methodology is given using accepted tools and which takes into account data that are subject to noise and other abnormalities.

Presently, the popular flow and transport numerical codes MODFLOW-2000 and MT3DMS are employed as the coupled forward (and predictive) transport model. The parameters of this model are estimated with PEST using a hybrid regularization approach similar to that presented in Tonkin and Doherty (2005). These codes are freely available, are regularly used in practice, and offer many state-of-the-art techniques.

The data assimilation methodology utilized for this study is outlined in Figure 1 and is now described in detail:

Step 1: Develop a conceptual model and calculate a best-guess or average value(s) for each model parameter using initial site knowledge. Setup the predictive model, herein also referred to as the forward model, using these estimates.

Step 2: Collect data from the field for some period of time and analyze the data for anomalies, discarding those data that are believed to be erroneous. Detecting anomalies in a dataset is a problem-specific and often subjective task. However, using a rule-based methodology founded on problem intuition and known collection errors, many of these abnormalities may be automatically detected using software. Such a service is described in the following section.

Step 3: Employ a data reduction algorithm to identify and remove those data that are not, or are no longer, informative to the parameter estimation process. An approximate data reduction method is now suggested: Stummer et al. (2004) present an algorithm for designing subsurface electrical resistivity experiments. Because it is becoming economically impractical to test all experimental combinations, they propose a heuristic approach which will determine which combinations will give the most information to the model. A basic data reduction procedure may be derived from this idea. Select a base dataset to be all "new" data that have just been collected and not yet assimilated into the model. Rank the remaining observations (or groups of observations) according to a goodness function, Equation 1, and eliminate those which are not above a certain threshold (or, alternately, by only accepting a given number). By reducing the set of data, the parameter estimation process is less likely to be overwhelmed by the sheer number of data, much of which may be redundant or insignificant.

The goodness function is defined as:

$$GF(i) = \sum_{j-1}^{M} \frac{|J_{ij}|}{J_j^{\text{sum}}} \left(1 - \frac{R_{jj}^{\text{base}}}{R_{jj}^{\text{compr}}}\right) \quad i = 1 \ldots (N - N_{\text{base}}) \tag{1}$$

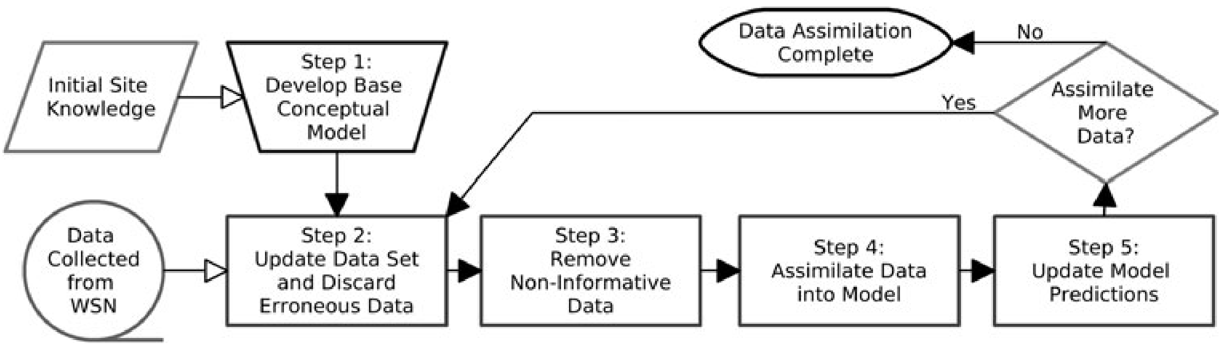

Figure 1. Proposed data assimilation methodology.

where J_{ij} is the ith row and jth column of the Jacobian ($J_{ij} = \partial d_i / \partial p_j$ in which d_i is the ith data point and p_j is the jth parameter out of N and M data points and parameters, respectively). Here,

$$J_j^{\text{sum}} = \frac{1}{N} \sum_{l=1}^{N} |J_{lj}| \tag{2}$$

is a normalization factor of the composite observation sensitivities. From linearizing the nonlinear parameter estimation problem, let $\mathbf{R} = \mathbf{J}^{-j}\mathbf{J}$ where $\mathbf{J}^{-j}$ is the generalized or pseudo-inverse of $\mathbf{J} \cdot \mathbf{R}$ is the resolution matrix; $\mathbf{R}^{\mathbf{compr}}$ refers to the resolution matrix calculated using all available data (comprehensive) and R^{base} is calculated using only the base dataset. The right-hand side of Equation 1 estimates the additional resolution that d_i would give to the parameter estimation algorithm.

Step 4: Update forward model parameters so that model output better fits the filtered and reduced dataset utilizing a parameter estimation technique. Assimilating the filtered and reduced dataset will require parameter estimation expertise since there are many sensitive and subjective parameters used in most robust parameter estimation algorithms. Fienen et al. (2009) and Moore and Doherty (2005) give examples of this assertion. However, as experienced in this study, once a model parameter estimation procedure has been customized to the initial dataset, the assimilation of new data may be achieved using little or no manual intervention.

Step 5: Execute the predictive model with the updated parameters to provide predictions for the purpose of decision making. Predictive error analysis may be carried out if it lends credibility to the predictions (Moore and Doherty 2005). Data assimilation can continue as long as desired by returning to Step 2.

As stated in Introduction, this study avoids new research methods (such as the Ensemble Kalman Filter) in favor of commonly accepted modeling tools with the aim of evaluating how easily accessible and existing techniques handle this new data context. Still, several aspects of the above procedure have not been widely considered by the hydrological modeling community. First, Steps 2 through 5 are performed automatically. Although many modeling studies check for data anomalies and redundancies prior to model calibration, none automatically perform these tasks as new data arrive. Other fields (notably, meteorology) have already made this transition. Secondly, the detection and alert of abnormal data are done using algorithms distributed within the WSN itself; that is, important data processing is performed in each wireless mote which reduces the communication load in such networks. Few other examples of this exist even in the WSN literature (see Urteaga et al. 2009). Lastly, the data reduction method in Step 3 is novel and easily integrated with available tools.

Fault Detection Service

Motivated by the problem of monitoring subsurface transport, Urteaga et al. (2009) developed a fault-detection WSN application to manage network and data faults. The software is implemented in the WSN-optimized programing language of NesC for the TinyOS mote operating system, which is specifically designed for resource constrained wireless sensor applications. The fault detection service comprised a sensor reading validity (SRV) subservice, which detects erroneous sensor readings, and a network status report (NSR) subservice, whose task is to abate data loss by identifying unresponsive nodes. Below, the SRV and NSR subservices are summarized. The interested reader will find a more complete description of the service in Urteaga et al. (2009).

Sensors translate a physical magnitude of interest into human or machine readable signals. This translation process is subject to many nonideal factors: sensitivity variations, scale and offset dynamic error, calibration drift, hysteresis, noise, etc. Systematic errors (influenced by offset, scale ranges, sensitivity variations, nonlinearity, etc.) may be handled by calibration, whereas signal processing techniques may compensate for random errors (primarily noise).

The basis of the SRV subservice is to (in)validate sensor readings. Basic signal processing is first applied to sampled sensor outputs in order to minimize the impact of noise on provided readings. Then the signal processed value is subjected to a set of rules that are designed to elucidate the existence of an erroneous reading (based on faults classified from previous sensor deployments, Szewczyk et al. 2004; Ramanathan et al. 2006; Porta et al. 2009):

1. Abnormally noisy data: the standard deviation of readings is larger than expected.
2. Nonlinear detection range (NLDR) data: the reading is out of the linear calibration range of the sensor.
3. Out of range data: the reading is outside the total detection range of the sensor.
4. Stuck data: the values reported seem unusually steady.
5. Abruptly changed data: the reading is drastically different than the last read value.

By applying the rules above, both systematic and random sensor errors are detected and reported by the SRV subservice, thus reducing data quality uncertainty.

Typical WSN deployments are subject to unforgiving environments at remote locations where node and link failures are likely to happen. Nodes may appear unresponsive if they run out of battery power, are broken, or if connectivity with their neighbors is lost. Some packet loss may be tolerated since the loss of a few data still yields a significant improvement over current monitoring methods; however, consistent link and/or node failure in a sparse network leads to loss of important information. Hence, the primary goal of the NSR subservice is to detect unresponsive nodes and diagnose their root causes. When multiple neighboring nodes fail to communicate with a particular mote, a corroborated warning is triggered.

Intelligent decisions can be made by processing warning messages either within the motes themselves or at a base data collection station. For instance, a mote may do its own filtering by ignoring all readings outside the linear detection range. The mote might be alternately designed to simply report all reading and warning messages—letting a software application or technician at the base station filter data as needed. For this paper, the latter was chosen; a network technician browses the warning messages every few days and corrects network and sensor problems as they arise.

Computational Performance Evaluation

This section provides an example of the practicable data assimilation method described above. A fine-scale synthetic dataset was generated based on an actual field site. A WSN simulator added abnormalities to this data and the fault detection service detected some of these anomalous readings. Using these tools, four datasets were produced: an infrequently collected dataset containing minimal noise, a second from a WSN containing theoretical noise levels according to sensor specifications, a third including realistic sensor and network faults (including noise), and a fourth with fewer WSN faults due to the employment of the fault detection service. The three resulting datasets are assimilated into a forward model and the subsequent model predictions are compared to the reference dataset.

Synthetic Data Model

The above methods may initially be best evaluated using a synthetic dataset since: (1) any number of arbitrary sensor locations may be chosen in the domain; (2) predictions can be evaluated at all locations and at each time step; and (3) the level of model complexity is easily controlled. The challenge then exists to create synthetic data to be realistic, that is to say, governed by many of the same processes and properties that exist in a natural setting; interpolating/extrapolating field data with geostatistical tools is the conventional approach. A synthetic data model of conservative solute transport was inferred from the MADE field site (Boggs et al. 1992) for this research.

A sequential Gaussian simulator produced a stochastic field based on flowmeter measurements (Boggs et al. 1992) and variogram data (Rehfeldt et al. 1992) from the MADE site. Nonstationarity was achieved by adding a second-order polynomial trend to the logarithmic hydraulic conductivity field, ln(**K**) (Rehfeldt et al. 1992), and the final result is depicted in Figure 2.

MODFLOW-2000 solved the steady-state flow equation ($\nabla\cdot(\mathbf{K}\nabla h) = 0$) using the previously generated hydraulic conductivity fields. No flow conditions were placed on the left, right, and bottom boundaries. A constant head of 67 and 64 m was forced on the inlet and outlet boundaries, respectively, to create a head drop similar to that reported in Boggs et al. (1992).

The MT3DMS software computed conservative solute transport based on the typical governing equation:

$$\frac{\partial(\phi C)}{\partial t} = -\mathbf{v}\cdot\nabla(\phi C) + \nabla\cdot(\mathbf{D}^H\nabla(\phi C)). \qquad (3)$$

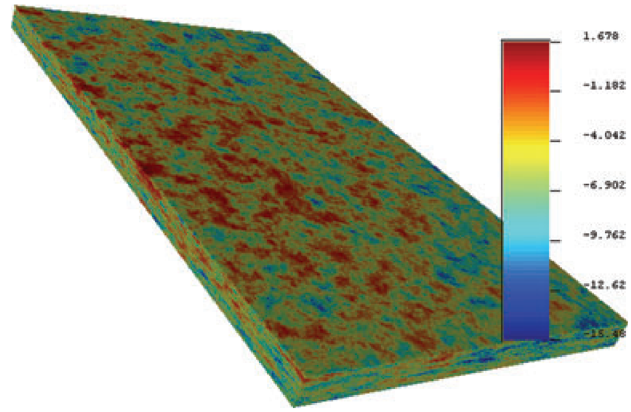

Figure 2. Synthetic ln(K) field used to generate reference data.

Porosity, ϕ, varies throughout the domain based on a Kozeny-like relationship to permeability (Bear 1972) and field porosity measurements (Boggs et al. 1992). Dispersion is treated locally such that $D^H = 0.1$ m in the x and y directions and $D^H = 0.001$ m in the z direction. Two source zones were created, 40 m$^2\times$ 5 m deep. Each zone is 10 m from the upstream boundary, 50 m from each side, and centered in the z direction. Initial concentrations in the source zone are randomly correlated to permeability by using a stochastic process. The randomly generated plume was allowed to develop for 1095 d (approximately 3 years) in the three-dimensional (3D) domain.

The predictive transport model used here is 2D, thus predicted concentrations were compared to the depth-averaged 3D concentration field. This 2D field is referred to as the reference dataset, and a snapshot is shown in Figure 3.

Noisy, Faulty, Filtered, and Sparse Datasets

One hundred random "sensor" locations were chosen in the synthetic model domain. Noise and errors which reflect measurement and environmental conditions were added to the concentration values at these points. For this example, sensor noise is based on electrical conductivity (EC) sensors used in WSN laboratory experiments (Porta et al. 2009). Let κ be the actual EC without noise then, based on empirical data and sensor documentation, EC reading noise is modeled as:

$$\kappa_\eta = \kappa + \eta \text{ where } \eta \sim N(0, \sigma^2_{\text{base}} + \varepsilon\kappa) \qquad (4)$$

That is, a base amount of Gaussian noise is present in all readings and additional Gaussian noise is included as κ increases ($\varepsilon \in (0, 1)$). κ is found through an empirical linear calibration curve custom to each sensor; the reference data at each node location was translated into EC values (using the calibration curve), random noise was then added according to Equation 4, and the result is translated back to concentrations. Such a dataset will be referred to as the noisy dataset.

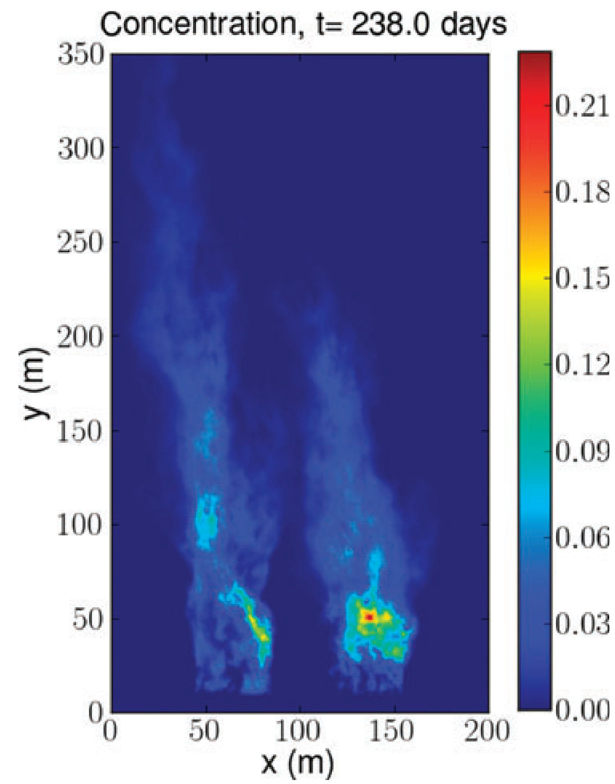

Figure 3. Depth-averaged snapshot of conservative plume at 238 d.

Based on reported errors in Ramanathan et al. (2006), Porta et al. (2009), and Szewczyk et al. (2004), Bernoulli Processes generated the following abnormal sensor readings:

1. *Stuck readings*: the sensor continues to report approximately the same value despite changing conditions.
2. *Out of range*: the sensor will report a value out of range for a contiguous period of time.
3. *Abrupt shifts*: the linear calibration curve will shift randomly and abruptly.
4. *Abnormal noise*: the variance in random noise for each reading is increased.

In the absence of random faults, κ is still perturbed by: (1) adding random Gaussian noise as before and (2) slightly drifting the calibration. A fault persists for a random period of time after which the faulty sensor is "fixed" (presumably by a network technician) and the sensor again reports sensible values. The dataset containing all of the above random faults, random noise, and calibration drifts is called the *faulty* dataset—created by a WSN simulator called TOSSIM (Levis et al. 2003) which realistically reproduces WSN physical and link layer features (Lee et al. 2007). In addition, packet loss rates follow the empirical data from Zhao and Govindan (2003).

The above fault detection service was performance tested inside TOSSIM (Urteaga et al. 2009). For this study, a *filtered* dataset was created by passing the faulty dataset through the fault detection service. If a sensor reading was flagged as faulty, that data was not included in the filtered dataset—effectively filtering out likely erroneous data. In this case, the fault may be "fixed" more quickly than in the faulty dataset because a warning message was issued to a network technician. Thus, data will be missing from the filtered dataset for a shorter random period of time.

Those datasets already presented include concentration data on a daily basis. This temporal resolution is achievable with a WSN but is not possible using traditional data acquisition techniques. Mimicking manual data collection, a *sparse dataset* was obtained by including reference data at each node location every three months.

For illustration, the BCs for one of the nodes is given in Figure 4. Although the sparse dataset contains significantly less data than the others, the shape of the BC may still be easily reconstructed from the available values. Several types of errors are apparent in the faulty BC: lost packets from poor network link quality, longer periods of missing data caused by mote failure, abnormally noisy data, and abrupt shifts in readings. The fault detection service was able to identify most of these at the cost of filtering out additional data.

Forward (Predictive) Transport Model

There is always a significant discrepancy between the complexity of the processes and properties that govern the evolution of transport in a natural setting and the complexity of the computational model used to capture the main features of this phenomenon. The reference dataset is taken from a 3D model with 2.1 million grid cells, trended stochastic K-field, random porosity and random source zones. In contrast, the forward model uses 17,500 grid cells in 2D with transmissivity for all cells estimated by kriging 646 uniformly placed pilot points. Porosity is constant, the mean of the measured field porosity given in Boggs et al. (1992). The location of the source zones are assumed known but they contain only the average concentration (taken over all cells in the zone) of the random fields in the synthetic data model. Dispersivity is adjusted to have the same grid-peclet number and the boundary conditions are identical to those used in the synthetic data model.

Again, MODFLOW-2000 and MT3DMS form the computational model. The pilot points were adjusted using PEST's regularization mode. Tikhonov and truncated singular value decomposition (TSVD) regularization methods were used together to stabilize the Gauss-Newton solver. PEST is unique in its ability to simultaneously utilize both of these forms of regularization in a hybrid scheme and the interested reader may wish to read the paper by Tonkin and Doherty (2005) for algorithmic details.

Initially, all pilot points were set to the average transmissivity of the field flowmeter data (Rehfeldt et al. 1992). The forward model made an initial forecast; data were assimilated every 92 d thereafter (up through day 552) based on the methodology described in Transport

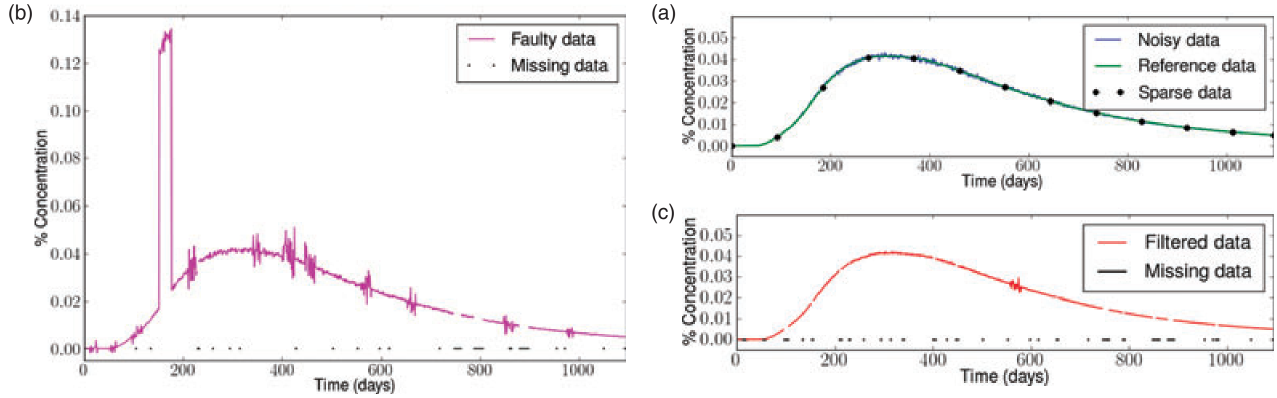

Figure 4. Sample BCs for node 44. Reference and noisy data are shown in (a), faulty data in (b), and filtered in (c).

Model Data Assimilation section. In other words, PEST calibrated the initial homogeneous parameter set to the first 92 d of BC data. The resulting parameter set was then used as the initial parameter set for calibrating to the next 92 d worth of data, etc. A special version of PEST, called BeoPEST (http://www.prinmath.com/pest/), was used to execute concurrent forward model runs on a high-performance cluster belonging to the Golden Energy Computing Organization (http://geco.mines.edu/).

Results and Discussion

For a given time, t, the following formula was used to measure the *congruence* between predicted and true observations:

$$v(t) = 1 - \int_V \frac{|C^{\mathrm{t}} - C^{\mathrm{p}}|}{C^{\mathrm{t}} + C^{\mathrm{p}}} \mathrm{d}V \tag{5}$$

C^{t} is the reference (relative) concentration given from the synthetic dataset and C^{p} is the predicted (relative) concentration provided by the predictive model. In practice, of course, the discrete analog is used.

Figure 5 displays v (V is the entire 2D domain) for each data assimilation experiment and for each iteration therein. Dashed lines indicate the congruence of model forecasts, that is, time periods from which no data had yet been collected. The quantities of data included for each parameter estimation run are shown in Figure 6. In addition, the contour plots in Figure 7 qualitatively show how well the predictive model is capturing large-scale features of the plume at day 672, which is 120 d (or about 4 months) after the last data assimilation.

Transport model calibration using sparse data gives poor results. From Figure 7d, data assimilation was ineffectual until 368 d of data was amassed. Even then, congruence to the reference model never reaches 0.4. The inability to capture early plume development causes the second source zone to be washed out much sooner than the first—this is not a feature in the reference data (cf. Figures 7a and 7e).

From Figure 5, parameter estimation using the noisy dataset clearly has the best congruence with the reference model. The abrupt drops in the congruence curves might be influenced by the selection of Tikhonov regularization and kriging parameters. Model congruence continues to improve as new data are assimilated despite the removal of most previous data using the data reduction methodology, Figure 6. This suggests that old data have already been "memorized" by the parameter set. The forecast in Figure 7b indicates that the predictive model locates the remaining solute in roughly the same areas as the reference model, Figure 7a. Still, the noisy dataset represents an idealized version of reality—one in which sensors and networks perform flawlessly.

Prima facie, no significant difference in congruence is shown between the faulty and filtered datasets (cf. Figures 5b and 5c, respectively); the reason for this will be discussed below. Yet, the forecasts for these datasets show two important distinctions. First, in Figure 7c, more source has moved from the first source zone than the second, whereas Figure 7d indicates a more rapid depletion of the second source. Secondly, the forecast from the filtered dataset shows 100% of the source mass still within the model domain but, in contrast, the model forecast from the faulty dataset more closely agrees with the fact that mass has already exited the top boundary (cf. Figure 7a). These observations suggest that model calibration with the unfiltered, faulty data was more successful than when using the filtered data. However, both forecasts underestimated permeability, leaving considerably more mass in the source areas than the other plots in Figure 7.

Over 90% of the erroneous data from the faulty dataset were discovered and corrected using fault detection. Despite this dramatic decrease, the final filtered dataset results indicate how parameter estimation is sensitive to just a small number of abnormal data (cf. Figures 5c and 7d). Furthermore, because only a few anomalous data exist, the parameter estimation procedure was capable of fitting these data too closely. In this case, the optimal data misfit (PHIMLIM in PEST) of Tikhonov regularization should be adjusted to help correct this problem but this proves difficult to estimate since statistics of data anomalies are not usually known a priori.

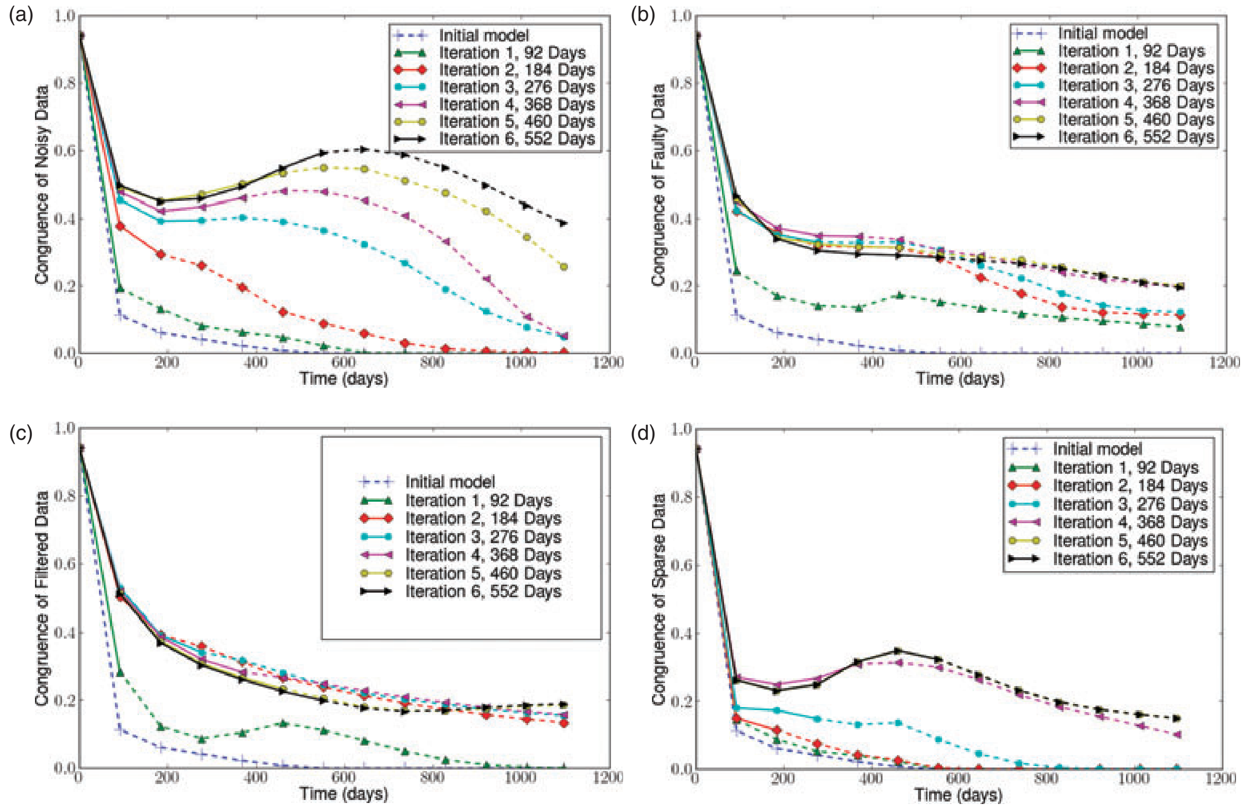

Figure 5. The congruence, Equation 5, of the predictive model to the reference model when calibrating with each dataset, noisy (a), faulty (b), filtered (c), and sparse (d). Solid lines indicate predictions when data were already collected, whereas dashed lines are true forecasts.

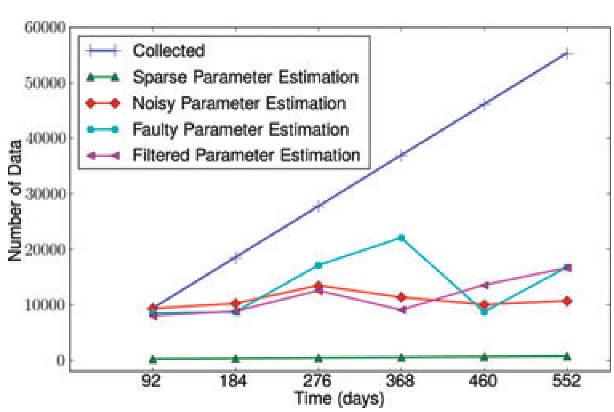

Figure 6. Amounts of data used for each data assimilation iteration.

In all cases, when using filtered data, the parameter estimation software obtained a value of the measurement objective function considerably smaller than with the faulty data. As seen in Figure 8, ϕ was reduced by almost three orders of magnitude due to the removal of a large majority of faulty data, bringing the ϕ values for the filtered dataset quite close to the dataset containing only theoretical noise levels. Prior to computing model congruence and plotting the final results, one may expect this improvement in calibration to translate into an enhancement in final model results. This did not occur. As just stated, the satisfactory decrease in the value of the measurement objective function for the filtered dataset can be explained by successful calibration to relatively few erroneous data. In the case of the faulty dataset, the quantity of erroneous data overwhelmed PEST even during the first data assimilation iteration. The value for PHIMLIM was then automatically chosen to be slightly larger than the best value of the measurement objection function to account for model misfit (a standard technique). PEST output results subsequently show that parameter smoothing (from Tikhonov regularization) dominated the parameter estimation for the faulty dataset. Again, reflecting on Figures 5 and 7c, even though PEST was unable to calibrate well to the faulty data, model congruence is at least as good as when using filtered data by exploiting statistical correlations in the reference model.

Summary and Conclusions

This study considered the fate of accepted numerical subsurface contaminant models in light of a new data context. A complex 3D synthetic model was developed using field site measurements and, from this, four datasets were constructed to mimic different scenarios. A simple data assimilation methodology was employed to calibrate a 2D numerical transport model and subsequent model predictions were compared to the reference dataset.

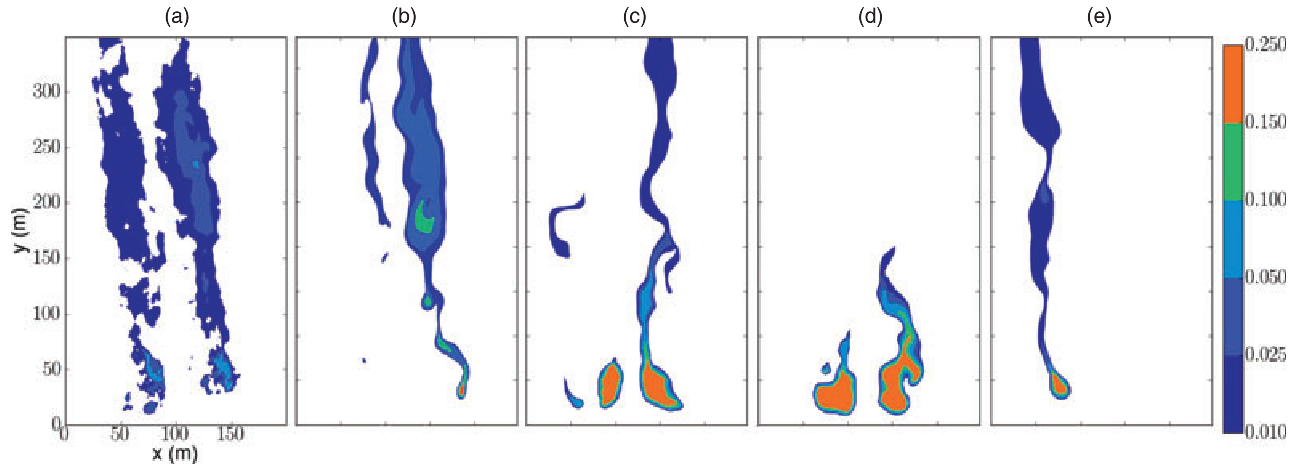

Figure 7. Reference (a) versus predictive snapshots of relative concentration using the noisy (b), faulty (c), filtered (d), and sparse datasets (e), respectively, at day 672 of the simulation.

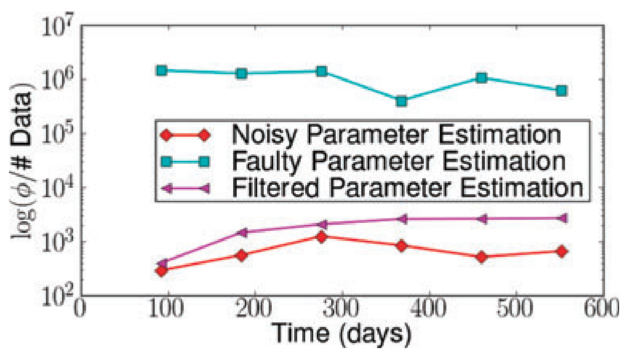

Figure 8. Objective function (ϕ) reduction due to REFLAG.

This investigation is concluded with the following suggestions and remarks:

1. Even small amounts of erroneous data may significantly affect the outcome of the calibrated model while a fault detection service may remove a majority of these abnormalities, manual inspection of data may still be advisable. Furthermore, satisfactory parameter estimation results do not guarantee reasonable model predictions if faults persist. When statistics of data faults, such as random noise, can be estimated a priori then regularization techniques may be used to ensure robust solutions.

2. Transport model forecasts are highly sensitive to parameter estimation input variables, particularly regularization parameters. Although PEST includes many features to induce robust parameter estimation and model inversion (Doherty 2005, 2008), expertise is needed to achieve reliable results (Fienen et al. (2009)). Parameter estimation software should be viewed as another component in the modeling process which contributes to the overall model error and for which credibility should be sought (Moore and Doherty 2005).

3. Model agreement generally improves with each iteration of data assimilation even when most previous data has been eliminated through data reduction. This iterative approach to parameter estimation may be appropriate whenever transient data is available.

4. In agreement with Chu et al. (1987), predictions of transport fate only remain valid for short time periods after model calibration. The level of heterogeneity at the actual site may impact this length of time. Still, this work advocates that data assimilation and wireless networks be left in place during the entire field study.

The results suggest that contaminant transport models will benefit as in situ sensing and network technologies mature. It is imperative that existing modeling tools must be adapted and new approaches should be developed in order to efficaciously assimilate WSN data. Finally, though the presented case study is anticipatory of technological advancement, further modeling studies are likely to help guide technology development and deployment.

Acknowledgments

This project is supported in part by the Army Research Office, Terrestrial Sciences, Environmental Science Division, and also by NSF CSR grant CNS-0720875. Special thanks to Dr. John Doherty for insights into the parameter estimation process. Dr. Willem Schreuder is credited for the development and guidance in using BeoPEST. The Golden Energy Computing Organization is acknowledged for providing computer resources on their high-performance cluster.

References

Anderson, M.P., and W.W. Woessner. 1992. The role of the postaudit in model validation. *Advances in Water Resources* 15, no. 3: 167–173.

Bear, J. 1972. *Dynamics of Fluids in Porous Media*. New York: American Elsevier Publishing Company, Inc.

Boggs, J., S. Young, L. Beard, L. Gelhar, K. Rehfeldt, and E. Adams. 1992. Field study of dispersion in heterogeneous aquifer, 1, over view and site description. *Water Resources Research* 28, no. 12: 3281–3291.

Chu, W., E. Strecker, and D. Lettenmaier. 1987. An evaluation of data requirements for groundwater contaminant transport modeling. *Water Resources Research* 23, no. 3: 408–424.

Culler, D., D. Estrin, and M. Sirivastava. 2004. Overview of sensor networks. *IEEE Computer* 13, no. 88: 41–49.

Doherty, J. 2008. *Addendum to the PEST Manual*. Brisbane, Queensland, Australia: Watermark Numerical Computing.

Doherty, J. 2005. *PEST Version 9 Users Guide.* Brisbane, Queensland, Australia: Watermark Numerical Computing.

Eggleston, J., and S. Rojstaczer. 2000. Can we predict subsurface mass transport? *Environmental Science Technology* 34, no. 18: 4010–4017.

Ezzedine, S., and Y. Rubin. 1996. A geostatistical approach to the conditional estimation of spatially distributed solute concentration and notes on the use of tracer data in the inverse problem. *Water Resources Research* 32, no. 4: 853–861.

Fienen, M., C. Muffels, and R. Hunt. 2009. On constraining pilot point calibration with regularization in PEST. *Ground Water* 47, no. 6: 835–844.

Graham, W., and D. McLaughlin. 1991. A stochastic-model of solute transport in groundwater—Applications to the Borden, Ontario, tracer test. *Water Resources Research* 27, no. 6: 1345–1359.

Harbaugh, A., E. Banta, M. Hill, and M. McDonald. 2000. MODFLOW-2000, the USGS modular ground-water model user guide to modularization concepts and the groundwater flow process. USGS Open File Report 00-92. Reston, Virginia: USGS.

Harvey, C., and S. Gorelick. 1995. Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity. *Water Resources Research* 31, no. 7: 1615–1626.

Hassan, A.E. 2004. Validation of numerical ground water models used to guide decision making. *Ground Water* 42, no. 2: 277–290.

Hendricks Franssen, H., and J. Gòmez-Hernàndez. 2003. Coupled inverse modelling of groundwater flow and mass transport and the worth of concentration data. *Journal of Hydrology* 281, no. 4: 281–295.

Ho, C. 2005. Micro-chemical sensors for in situ monitoring and characterization of volatile contaminants. http://www.sandia.gov/sensor/.

Kim, J., Y. Park, and T.C. Harmon. 2005. Real-time model parameter estimation for analyzing transport in porous media. *Ground Water Monitoring & Remediation* 25, no. 2: 78–86.

Lee, H., A. Cerpa, and P. Levis. 2007. Improving wireless simulation through noise modeling. In *The Sixth International Conference on Information Processing in Wireless Sensor Networks (IPSN)*, 21–30.

Levis, P., N. Lee, M. Welsh, and D. Culler. 2003. TOSSIM: Accurate and scalable simulation of entire tinyOS applications. In *Proceedings of the First ACM Conference on Embedded Networked Sensor Systems (SenSys 2003)*, 126–137.

Liu, G., Y. Chen, and D. Zhang. 2008. Investigation of flow and transport processes at the MADE site using ensemble Kalman filter. *Advances in Water Resources* 31, no. 7: 975–986.

Lundquist, J.D., D.R. Cayan, and M.D. Dettinger. 2003. Information processing in sensor networks, chapter meteorology and hydrology in Yosemite National Park: A sensor network application. In *The Second International Conference on Information Processing in Wireless Sensor Networks* (*IPSN*) 2003, 553.

McLaughlin, D. 2002. An integrated approach to hydrologic data assimilation: Interpolation, smoothing, and filtering. *Advances in Water Resources* 25, no. 8–12: 1275–1286.

McLaughlin, D., L. Reid, S. Li, and J. Hyman. 1993. A stochastic method for characterizing groundwater contamination. *Ground Water* 31, no. 2: 237–249.

Medina, A., J. Carrera, and G. Galarza. 1990. Inverse modelling of coupled flow and solute transport problems. In *ModelCARE 90: Calibration and Reliability in Groundwater Modelling.* IAHS Publication no. 195: 185–194.

Moore, C., and J. Doherty. 2005. Role of the calibration process in reducing model predictive error. *Water Resources Research* 41, no. 5: W05020, doi: 10.1029/2004WR003501.

Musăloiu-E., R., A. Terzis, K. Szlavecz, A. Szalay, J. Cogan, and J. Gray. 2006. Life under your feet: A wireless soil ecology sensor network. In *Third Workshop on Embedded Networked Sensors (EmNets)*.

Porta, L., T.H. Illangasekare, P. Loden, Q. Han, and A.P. Jayasumana. 2009. Continuous plume monitoring using wireless sensors: Proof of concept in intermediate scale tank. *Journal of Environmental Engineering* 135, no. 9: 831–839.

Ramanathan, N., L. Balzano, M. Burt, D. Estrin, T. Harmon, C. Harvey, J. Jay, E. Kohler, S. Rothenberg, and M. Srivastava. 2006. Rapid deployment with confidence: Calibration and fault detection in environmental sensor networks. Technical report for the Center for Embedded Network Sensing, UCLA.

Rehfeldt, K.R., J.M. Boggs, and L.W. Gelhar. 1992. Field study of dispersion in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resources Research* 28, no. 12: 3309–3324.

Stummer, P., H. Maurer, and A. Green. 2004. Experimental design: Electrical resistivity data sets that provide optimum subsurface information. *Geophysics* 69, no. 1: 120–139.

Szewczyk, R., J. Polastre, A. Mainwaring, and D. Culler. 2004. Lessons from a sensor network expedition. In *European conference on Wireless Sensor Networks* (*EWSN)* 2001, 307–322.

Tonkin, M., and J. Doherty. 2005. A hybrid regularized inversion methodology for highly parameterized environmental models. *Water Resources Research* 41, no. 10: W10412.

Trubilowicz, J., K. Cai, and M. Weiler. 2009. Viability of motes for hydrological measurement. *Water Resources Research*, 45, no. 6: 16.

Urteaga, I., K. Barnhart, and Q. Han, 2009. REDFLAG: A RuntimE, Distributed, Flexible, Lightweight, And Generic fault detection service for data-driven wireless sensor applications. *Pervasive and Mobile Computing* 5, 432–446.

Zhao, J., and R. Govindan. 2003. Understanding packet delivery performance in dense wireless sensor networks. In *ACM Sensys,* 1–13.

Zheng, C., and P. Wang. 1999. MT3DMS: A modular 3D multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems. US Army Engineer Research and Development Center Contract Report SERDP-99-1.

Zimmerman, D., G. de Marsily, C. Gotway, M. Marietta, C. Axness, R. Beauheim, R. Bras, J. Carrer, G. Dagan, P. Davies, D. Gallegos, A. Galli, J. Gòmez-Hernàndez, P. Grindrod, Gutjahr, P. Kitanidis, A. Lavenue, D. McLaughlin, S. Neuman, B. RamaRao, C. Ravenne, and Y. Rubin. 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport to groundwater flow. *Water Resources Research* 34, no. 6: 1373–1413.