

Integration of Groundwater Transport Models with Wireless Sensor Networks

Kevin Barnhart¹, Iñigo Urteaga², Qi Han², Anura Jayasumana³, Tissa Illangasekare¹

¹*Division of Environmental Sciences and Engineering, Colorado School of Mines, kbarnhar@mines.edu, tissa@mines.edu, Golden, CO, USA*

²*Department of Mathematical and Computer Sciences, Colorado School of Mines, urteaga@gmail.com, qhan@mines.edu, Golden, CO, USA*

³*Department of Electrical and Computer Engineering, Colorado State University, anura.jayasumana@colostate.edu, Fort Collins, CO, USA*

ABSTRACT

Groundwater transport modeling is intended to aid in remediation processes by providing prediction of plume behavior. Wireless Sensor Networks (WSNs), an emerging technology, allows integration of data from multiple sensors to form an integrated, distributed network that monitors dynamic hydrological and environmental processes. As the wireless nodes are resource constrained, mass transport predictions may be used to increase the efficiency of the application-driven WSN protocols. A symbiotic relationship then exists between the WSN, which supplies data to calibrate the transport model in real-time, and the model, which optimizes the WSN monitoring performance. As current model calibration methods often require manual adjustment of calibration parameters, real-time calibration procedures pose significant challenges. Based on insights from a previous intermediate-scale, proof-of-concept study, we are developing the models, tools, and protocols necessary for a closed-loop simulation. The results presented here address the setup of a WSN simulator which cooperates with transport models, development of fault detection techniques into the network protocols, and a preliminary discussion of how transport models may be conceptualized in the WSN context.

INTRODUCTION

As groundwater contamination is an established problem with many environmental and health consequences, substantial research has focused on the science underlying the relevant processes and cleanup methodologies. Increased availability of computer power and resources has been instrumental in the almost ubiquitous use of computer-based transport models to calculate health risks, determine cleanup strategies, guide environmental regulatory policy, and to determine culpable parties in lawsuits. This culminates in large investments of effort and funds, and important environmental decisions being made based on the outcome of transport models (Eggleston and Rojstaczer 2000).

Given the frequent use of these transport models it may be surprising that very few studies exist designed to establish model credibility (cf. Anderson and Woessner 1992 and Hassan 2004). There is a growing belief that transport models will prove to be a far more effective tool if the parameters of the model are periodically updated or re-calibrated as new data becomes available, cf. Anderson and Woessner (1992) and Kim et al. (2005). Still, as pointed out by Anderson and Woessner (1992), obtaining such data is largely impeded by lack of time and monetary resources.

A promising solution is the use of Wireless Sensor Networks (WSNs) where in-situ measurements of contaminant concentration and possibly other aquifer parameters (such as flow rate and direction, head, and temperature) could be gathered in real-time at less cost than traditional sampling techniques. A WSN is an ad hoc network of sensor nodes. Each sensor node (mote) contains a microprocessor and a wireless receiver/transmitter, and thus is capable of sensing, processing, and broadcasting data via wireless (radio) transmission. Encouragingly, there have been many recent advances in sensor network technology (cf. Culler et al. 2004) and in the development of miniature sensors which sample contaminants (e.g. Ho 2005).

Motes are resource-constrained; that is, they have limited computational, memory, and power capacities. Because radio hardware is by far the greatest power consumer, applications and protocols are required to

minimize transmissions through various means; for example, by compressing or aggregating data locally. Such optimizations can be very dependent on the needs of the application as applications require data reliability assurances. Thus, application specific information can intuitively optimize these protocols for better performance. In this case, predictions of mass transport by a transport model may be used to tune WSN network parameters. A symbiotic relationship then exists between the WSN, which is supplying the data to calibrate the transport model, and the transport model, which supplies predictive information to the WSN for improved monitoring performance.

This paper presents preliminary research on the coupling of a WSN with a groundwater transport model. Based on a preliminary proof-of-concept laboratory study, important challenges are elucidated. The methodologies that we are currently employing to overcome these obstacles are then given, followed by preliminary simulative results. We conclude with the direction of future research.

PROBLEM DESCRIPTION

At the Center for Experimental Study of Subsurface Environmental Processes (CESEP), Porta (2007) conducted sodium bromide transport experiments in an intermediate-scale test bed using a WSN. Sodium bromide concentration was measured using Electrical Conductivity (EC) probes (Decagon ECH₂O-TE) connected to Crossbow TelosB motes. Porta attempted to use EC data to calibrate a 2D transport model in real-time. Based on this study, two conclusions that may be gathered are: (1) sensor noise, calibration drift, and network faults are important concerns which require a robust solution, and (2) successful Real-time Automatic Calibration (RAC) of transport models is hindered by the presence of anomalous data, prohibitive calibration times, and the need to manually adjust calibration parameters.

Though no other study has attempted to incorporate transport models with WSNs, there are several documented cases of using WSNs to monitor subsurface hydrologic properties, e.g., Lundquist et al. (2003), Ramanathan et al. (2006), and Musaloiu-E. et al. (2006). Invariably, these studies report the collection of considerable amounts of anomalous data due to drifting sensor calibration, faulty electronics, and varying environmental conditions. Hence, though a WSN is capable of monitoring and measuring the environment at scales and resolutions not before possible, it poses enormous data analysis and management challenges. Approaches such as Virtual Sensor Networks may help manage the monitoring of plumes spread over large areas (Jayasumana et al. 2007), but a fault detection scheme is still needed.

Next, despite coping with anomalous data, a major barrier to implementing real-time automatic inversion is that available tools often require manual adjustment of subjective calibration parameters (e.g., regularization parameters, weighting factors, objective function threshold) and modeling techniques (e.g., selection of objective function, number of parameters, and parameter placement) before convergence is achieved. Thus, automatic calibration may also be considered a trial-and-error procedure. Furthermore, currently available automatic inversion codes frequently require hours or days to invert the governing partial differential equations, as they utilize general, brute-force inversion methodologies.

When considering the incorporation of a transport model into the WSN, frequent manual adjustment of the automatic inversion procedure would negate the advantages of having a real-time, closed-loop monitoring system. There is no evidence to suggest that existing codes can be used in a RAC procedure, and experience from modelers such as Porta (2007) suggests otherwise. The choice of model, parametrization, and inverse technique will likely have significant impacts on the success of coupling transport model with a WSN.

METHODOLOGY

Fault Detection Service

To manage network and data faults, we developed a distributed fault-detection WSN service. The software, called REDFLAG, is written in NesC for TinyOS, a mote operating system, and serves as a *RE*al-time, *D*istributed, *F*lexible, *D*etector of *F*aults, that is also *L*ightweight And *G*enerous. It provides a fault detection service to the Application Layer on each node, enabling the application to make distributed

fault management decisions without burdening it with low-level detection algorithms. The service is also independent of Routing and MAC Protocol Layers, so that maximum control of packet delivery, duty-cycling, etc. are still available to the network developer. Beyond layer independence, REDFLAG: operates with both event-driven and periodic data collection, and consumes trim resources in dense and sparse networks. REDFLAG is a generous service application which is suitable in a myriad of scenarios.

In REDFLAG, two different types of faults are considered: (a) abnormal sensor readings, and (b) unresponsive nodes. *Abnormal sensor readings* refer to both systematic (e.g., dealing with offset, scale ranges, sensitivity variations, nonlinearity, calibration drift, etc. of sensor readings) or random errors (e.g., noisy readings due to changing environmental conditions) which cause reported values to be erroneous. In contrast, *unresponsive nodes* may fail to report readings entirely because of faulty mote hardware, power depletion, and inconsistent radio transmission due to varying environmental conditions.

Each node is responsible for detecting abnormal sensor readings and managing these faults appropriately. To do this, the calibration parameters, detection ranges, and expected noise standard deviations can be programmed into a mote for the specific characteristics of the attached sensor. REDFLAG detects: noisy data, i.e., the standard deviation of readings is larger than expected; NLDR (Non-Linear Detection Range) data, i.e. the reading is out of the linear calibration range of the sensor; out of range data, i.e. the reading is outside the total detection range of the sensor; stuck data, i.e. the values reported seem unusually steady; and abruptly changed data, the reading is drastically different than the last read value.

Due to changing environmental conditions, some intermittent link failure should be expected between motes. 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. With power consumption in mind, the REDFLAG service allows some packet loss, but, when multiple neighboring nodes fail to communicate with a particular mote, a corroborated warning is triggered and may be sent to the base station by the Application Layer.

A more complete description of REDFLAG is provided by Barnhart et al. (2008). The REDFLAG service helps to ensure a reasonable level of data reliability for the end application. Here, only concentration data that has been validated by REDFLAG are used in RAC.

Real-time Automatic Calibration

Before considering custom transport models and inversion procedures, RAC was attempted using available tools. The finite difference codes MODFLOW (Harbaugh et al. 2000) and MT3DMS (Zheng and Wang 1999) were coupled for 2D mass transport prediction. Zoned hydraulic conductivity parameters were calibrated to periodic WSN concentration data by using the Gauss-Marquardt-Levenberg non-linear optimization algorithm in PEST (Doherty 2000).

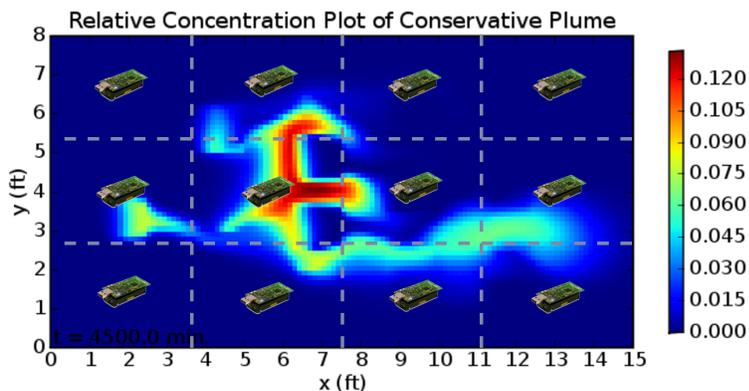


Figure 1: Simulation Setup

case in practice, since small-scale heterogeneities tend to dominate mass transport (Eggleston and

The “true” concentration data came from a MODFLOW/MT3DMS simulation motivated by the design of the next CESEP WSN experimental test bed. As shown in Figure 1, 12 motes are evenly distributed within the domain. These mark the center of the 12 hydraulic conductivity zones used in model calibration. Note that the plume, shown after $t=3.125$ days, is quite complex, yet we attempt to predict mass transport with only 12 crudely placed zones. We contend that this “under-parametrization” is often the

Rojstaczer 1998) but (1) the location and hydraulic properties of these small-scale features are usually unknown, and (2) parametrizing the model at such a small scale is typically computationally prohibitive for model inversion.

The RAC procedure was connected to TOSSIM, a realistic TinyOS WSN simulator (it supports MAC and Radio Layers, and uses sophisticated noise models to mimic actual WSNs). Following the flowchart in Figure 2, once the WSN simulation has begun, the transport model begins with an initial “non-informative” parameter set. Calibration is first performed using head data from the 12 node locations. Mass transport predictions are passed to the WSN protocols to improve their performance. Periodically thereafter (once some re-calibration criterion is met), the transport model is re-calibrated using concentration observations from the WSN, which consist of the “true” data subjected to WSN link failures and truncated precision.

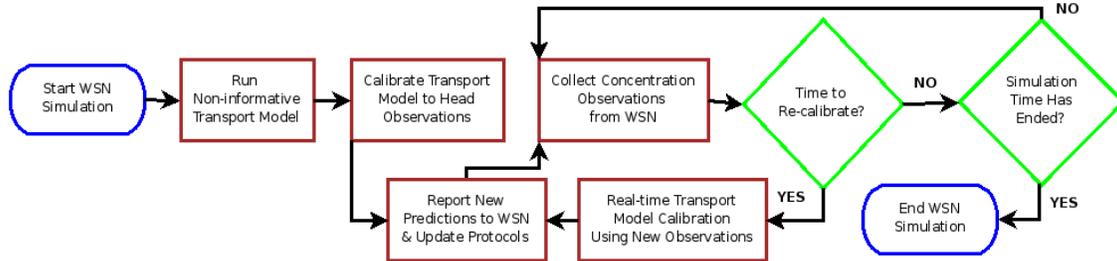


Figure 2: Basic RAC Procedure Combining WSN Simulator and Mass Transport Model

RESULTS

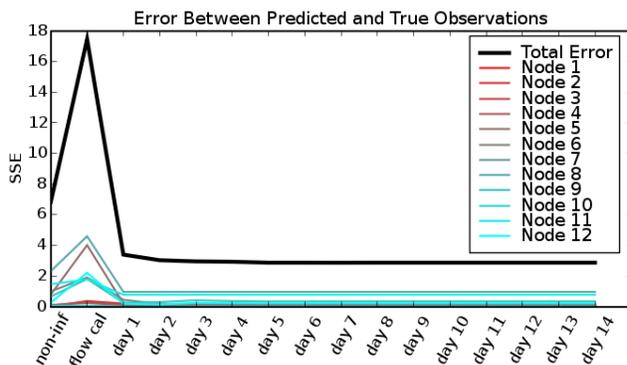


Figure 3: Prediction Error After Each Calibration

The simulation lasted 14 (virtual) days, with real-time re-calibration occurring after each day. The Sum of Squared Errors (SSE) between the predicted and observed breakthrough curves was calculated at each node after each calibration. These and the total SSE are shown in Figure 3. The most striking point in Figure 3, is that the calibration of the flow model to head data yielded poor prediction of mass transport. Perhaps if the initial inversion also calibrated budget data, the predictions would improve. Next, observe that the minimum total SSE is nearly reached after calibration using concentrations from the first day. Minor improvements are seen thereafter through

day 4, after which the prediction error appears to be asymptotic despite complex plume development. It seems that calibrating to concentration data does improve predictions of mass transport, but the efficacy of continued re-calibration is not apparent from this simulation. Further study is needed.

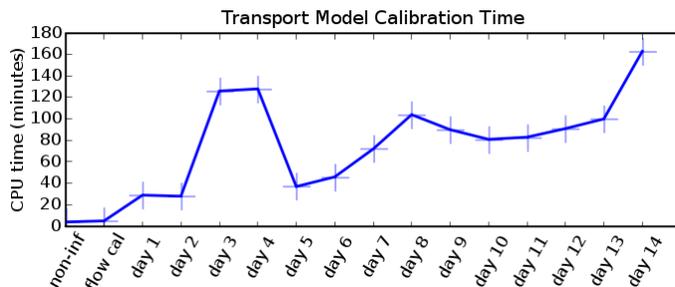


Figure 4: Time for Transport Model Calibration

The asymptotic prediction error may indicate some stability in the parameter set. While this is indeed the case (results not shown here), the calibration times of PEST’s Gauss-Marquardt-Levenberg algorithm continue to grow as simulation time increases (Figure 4). The magnitude of calibration times is worrisome since the given simulation is quite simple in comparison to models used in practice; i.e., one would expect to use far more than 12 parameters and run the simulation over a period longer than 14 days.

SUMMARY AND FUTURE WORK

Presented preliminary results suggest that available transport modeling codes and inverse techniques may be employed to perform RAC in the context of a WSN. This is stated under the following provisions: (1) the available observations contain no noise or faults, (2) there is sufficient computational resources available to complete the inversions, (3) initial, boundary, and source conditions are known, and (4) the contaminant behaves conservatively. Appropriately, future research will: (a) create network and sensor faults, add noise to the data, simulate sensor calibration drift, and employ regularization techniques, (b) develop more efficient ways of performing RAC by considering other transport models and inversion approaches, and (c) relax assumptions. Finally, though the results indicate that early re-calibration of the transport parameters with WSN observations leads to better predictions, it is not known whether these forecasts, and, moreover, the transport model, are useful in practice. Further work will suggest how WSNs can be used to quantify such model quality and credibility.

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.

REFERENCES

- Anderson, M. P. and W. W. Woessner, 1992. The role of the postaudit in model validation. *Adv. in Water Res.* 15, 167-173.
- Barnhart, K., I. Urteaga, and Q. Han, 2008. REDFLAG: A REal-time, Distributed, Flexible, Lightweight, And Generous fault detection service. *IEEE MASS 2008 Conference Proceedings*, in submission.
- Culler, D., D. Estrin, and M. Sirivastava, 2004. Overview of sensor networks. *IEEE Computer* , 41-49.
- Doherty, J. (2000). PEST: Model-independent parameter estimation. *Watermark Numerical Computing*.
- Eggleston, J. and S. Rojstaczer, 1998. Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport. *Water Resources Research* 34 (9), 2155-2168.
- Eggleston, J., and S. Rojstaczer, 2000. Can we predict subsurface mass transport? *Environ. Sci. Technol.*, 34(18), pp. 4010-4017.
- Harbaugh, A., E. Banta, M. Hill, and M. McDonald, 2000. MODFLOW-2000, the USGS modular groundwater model user guide to modularization concepts and the groundwater flow process. USGS. USGS Open File Report 00-92.
- Hassan, A. E., 2004. Validation of numerical ground water models used to guide decision making. *Ground Water* 42 (2), 277-290.
- Ho, C., 2005. Micro-chemical sensors for in-situ monitoring and characterization of volatile contaminants. <http://www.sandia.gov/sensor/>.
- Jayasumana, A. P., Q. Han and T. Illangasekare, 2007. Virtual Sensor Networks - A Resource Efficient Approach for Concurrent Applications. *Proc. Info. Techn.: New Generations (ITNG 2007)*
- 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(2), 78-86.
- 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. *IPSN 2003*, pp. 553. .
- Musaloiu-E., R., A. Terzis, K. Szlavetz, 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., 2007. Continuous plume monitoring and model calibration using a wireless sensor network: Proof of concept in intermediate-scale tank test, Masters thesis, Colorado School of Mines.
- 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 condense: Calibration and fault detection in environmental sensor networks, Tech. rpt. for Lab. for Emb. Collaborative Sys., UCLA.
- 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.