Title

Genetic Algorithm-Based Embedded Networked Sensing Design Coupled to an Environmental Simulator

Authors

Juyoul Kim
Thomas C. Harmon

Publication Date

2005

CENS Technical Report

Genetic Algorithm-Based Embedded Networked Sensing Design Coupled to an Environmental Simulator

by

Juyoul Kim[a] and Thomas C. Harmon[b]

[a]*Department of Civil & Environmental Engineering,*
University of California, Los Angeles
5732 Boelter Hall, Los Angeles, CA 90095-1593
[b] *School of Engineering*
University of California, Merced
P.O. Box 2039, Merced, CA 95344

April 10, 2005

[b]corresponding author:
Email: tharmon@ucmerced.edu
Phone: +1-209-724-4337
Fax: +1-209-724-2912

Abstract. Embedded networked sensing (ENS) technology is rapidly expanding into environmental application domains, where network coverage issues are tightly coupled to the environmental media and observational objectives. The goal of this work is to develop and test an automated, real-time ENS coverage design algorithm in the context of an environmental simulation model. The algorithm combines the application of a genetic algorithm (GA) with a deterministic inverse modeling approach, and is demonstrated in the context of a bench-scale groundwater test bed in which the ENS objective is to identify the location of a heat source. More specifically, optimal sensor locations are determined in real-time using a GA-based evolution algorithm whose objective function is the trace minimization of the model-prediction covariance with respect to potential sensor locations. Next, measured temperature sensor data and a descent-based inverse technique are used to update the source location estimate. The procedure is repeated (2 monitoring sensors per design cycle) until a pre-determined sensor supply is exhausted. Two transient heat transport experiments are undertaken in which sources placed upstream of a manually configurable ENS comprising thermocouples for mapping spatiotemporal temperature distributions. The ENS approach successfully corrected an erroneous initial source location estimate and incrementally improved upon this estimate with the addition of new sensors. A

point of diminishing improvement was eventually achieved at an imperfect source location estimate. This result was most likely the result of discrepancies between the mathematical model and the experimental system. For the dual-source experiment, the real-time source locator converged on a single source between the two sources, indicating the need for more sophisticated logic for increasingly complex cases.

Keywords: embedded networked sensing, sensor network design, genetic algorithm, inverse modeling, real-time parameter identification

1. Introduction

Science and engineering researchers working in earth, environmental and ecological systems understand that the key to answering many of their questions lies in better observation and decoding of the overlapping, multiscale spatiotemporal patterns that arise in the real world (NAP 2001, NSF 2003, Culler and Mulder 2004). This understanding is evidenced by the broad array of on-going large-scale observatory planning efforts (CUAHSI 2003, NRC 2003, AIB 2004). The ultimate goal of embedded networked sensing (ENS) technology research is to create a new way for applications domain scientists and engineers to observe their systems by creating the cyberinfrastructure enabling the network of specific environmental sensors to self-configure as a sophisticated, interactive virtual sensor for observing an environmental problem. The ENS approach extends beyond distributed sensors with data-logging capabilities, which are commonly deployed, to an integrated system in which the sensor readings are visualized and modeled in real-time, and which can be remotely re-tasked (e.g., to change sampling granularity) to continuously optimize performance. This work describes the development and testing of an embedded networked sensing (ENS) design algorithm for environmental media.

Environmental ENS systems currently need to be designed collaboratively by technology and applications domain researchers. An interdisciplinary approach is needed because it is difficult to conceive of an effective spatiotemporal sampling plan without domain-specific knowledge and network programming tools are not yet user-friendly enough to see widespread use amongst application domain experts. For example, ecologists are challenged with resolving issues of biodiversity and habitat restoration in the face of human population growth, land use change and climate change (NRC 2003).

The environmental ENS problem addressed here is the design of sensor network supporting groundwater observations, such as may be associated with optimal observation and management of water supplies or subsurface pollutants in the groundwater. The groundwater monitoring network design problem is a good potential application domain for

ENS because it represents a trade-off between the availability of data and their reliability. A common groundwater monitoring problem is concerned with the location of contaminant sources when transport model parameters are unknown or known with some uncertainty, and where monitoring points (which are costly) are to be located to minimize model prediction uncertainty for a given budget and collected data. Inverse modeling strategies have been developed for estimating model transport parameters and reconstructing unknown source information (Knowles and Yan, 2004; Sun, 1995; Sun and Sun, 2002; Wagner, 1992; Woodbury et al., 1998; Yeh, 1986; Yeh and Yoon, 1981), but have generally been applied either to synthetic data or in *post priori* exercises using real observations.

The illustrative example used in this work is the groundwater monitoring network design problem, a combinatorial problem in which the objective is to identify the optimal set of sensor locations from a set of potential locations. Even for a simple example of a sampling design of 5 sensor nodes to be deployed among 100 possible locations, there are C_{100}^{5} ($\approx 10^8$) possible sampling designs. This means that complete enumeration can only be used for small numbers of sensors. To solve the above large-scale combinatorial problem, a non-deterministic search approach is required.

The large-scale combinatorial optimization problem has been studied by many researchers and various numerical approaches have been proposed including simulated annealing (SA), tabu search (TS), genetic algorithm (GA), and others (Cieniawski et al., 1995; De Schaetzen et al., 2000; Giacobbo et al., 2002; Glover, 1986; Glover and Laguna, 1997; Hudak et al., 1995; Jaramillo et al., 2002; Kirkpatrick et al., 1983; Sciortino et al., 2002; Youssef et al., 2001; Zheng and Wang, 1996). SA, TS, and GA are all local search methods with particular rules for acceptance of a solution. In contrast to TS, SA and GA have rules for accepting solutions which are worse than the previous one (i.e., uphill moves) in order to escape local minima. In contrast to GA, SA and TS are based on the concept of neighboring designs which constitute the neighborhood. The experimental design problem in this work is addressed using GA.

The concept of GA was first proposed by Holland (1975), and has since been further developed by Goldberg (1989) and subsequently by many others. GA differs from many other types of optimization algorithms in that it searches the space from a population of points, not just from a single point. This is why GA can be applied to large, complex problems that are non-linear with multiple local optima. An advantage of GA over traditional gradient-based optimization searching approaches is that it can find global or near-global optimum of non-linear multi-objective function with multiple local optima (Carroll, 1996; Goldberg, 1989; Holland, 1975; Michalewicz, 1999; Michalewicz and Fogel, 2002). Another advantage of GA, particularly in the context of sensor network design, is its capacity to address combinatorial (0-1) integer variables that reflect the placement or absence of sensor at each potential sampling location. This integer combinatorial programming optimization problem is computationally intense. In this study, a GA-based combinatorial optimization technique is employed to solve the integer programming optimization problem of identifying optimal sensor locations for the purpose of delineating a heat source.

The GA approach has been applied previously to groundwater monitoring network design problems, typically for employing synthetic data generated by numerical simulators. For example, Cieniawski et al. (1995) investigated a method of optimization using GA to consider the two objectives of maximizing reliability expressed as percent of contaminant plumes detected and minimizing contaminated area at the time of first detection. They demonstrated GA utility in solving multiple-objective optimization problems, where they offered a distinct advantage over conventional optimization techniques because they were not affected by the complexity, convexity, or linearity of the objective function. More recently, and in a problem similar to the present one, Sciortino et al. (2002) examined spatial solute distributions generated in a controlled chemical release experiment, employing a GA to select optimal observation points from which to invert a model for the purpose of identifying the contaminant solute source location.

The objective of this paper is to develop and test an intelligent system supporting real-time ENS design in the context of the environment domain known as the groundwater monitoring network design problem. The analytical procedure combining a heuristic GA-based evolutionary algorithm is with a gradient-based inverse modeling technique is presented along with a brief description of the simulation model assumed to govern the physical processes in the environmental system. This is followed by a description of the test bed and three-dimensional heat transport physical aquifer model experiments used to test the proposed real-time algorithm.

2. Real-time Embedded Networked Sensing Design Algorithm

2.1. Experimental design algorithm

The goal of ENS design is to identify the optimal sampling sets from among many potential sensor locations. Given a model for the physical system in which the ENS is to be deployed, the ENS design problem can be expressed as an optimization problem employing an integer programming formulation. More specifically, the ENS design problem is to identify sensor locations which minimize the trace of the model-prediction covariance (Sciortino et al., 2002):

$$\min \operatorname{trace} \operatorname{Cov}(\mathbf{p}) \tag{1}$$

$$\text{subject to } \sum x_i C_i \leq B \tag{2}$$

where $\mathbf{p}$ is model parameter vector to be estimated

x_i is the indicator variable associated with sampling i

$x_i = 1$ if sampling i is selected; 0 otherwise

B is the budget, and C_i is the cost of sampling i

Here the covariance matrix provides a quantitative measure of the reliability of model parameters and can be used to evaluate and compare alternative ENS deployment strategies. The reliability of the estimated parameters is characterized by a norm of the covariance matrix. Covariance matrix of the estimated parameters is given as (Yeh and Yoon, 1981; Cleveland and Yeh, 1990):

$$Cov(\mathbf{p}) = \frac{E(\mathbf{p})}{M - L}\left(\mathbf{J}_D^T \mathbf{J}_D\right)^{-1} \tag{3}$$

where E is the least squares error, M is the total number of observations, L is the parameter dimension, and $\mathbf{J}_D$ is the Jacobian matrix. The proposed ENS design method identifies the optimal set of sampling points by minimizing a trace of covariance matrix (A optimality criterion). Other optimality criteria such as D optimality and E optimality have been used successfully in the source identification inverse model, but the A criterion has been proven to be equal or superior to them in terms of average percentage of decrease in parameter uncertainty with the increase of the number of observation points in this type of system (Sciortino et al., 2002).

2.2. Heat transport model for ENS test bed

An analogy between heat and mass transport is exploited in the test bed employed in this study to take advantage of the small form factor and low cost of temperature sensors relative to those for dissolved mass. The heat-mass transfer analogy is valid if the following conditions apply (Bird et al. 2002; Eckert and Drake 1987; Kim et al. 2005): physical properties are constant in time, no heat or mass is produced in the system (e.g., no chemical reaction), no radiant energy is emitted or absorbed, no viscous energy is dissipated, and the velocity is not affected by heat or mass transfer. The governing equation of transient heat transport for a continuous point heat source in a three-dimensional, homogeneous porous medium under steady, uniform flow conditions is as follows.

$$R_T \frac{\partial T(x,y,z,t)}{\partial t} = \kappa_x \frac{\partial^2 T}{\partial x^2} + \kappa_y \frac{\partial^2 T}{\partial y^2} + \kappa_z \frac{\partial^2 T}{\partial z^2} - u_x \frac{\partial T}{\partial x} + Q(x,y,z,t) \qquad (4)$$

$$\text{I.C.} \quad T(x,y,z,0) = T_o \qquad (5)$$

$$\text{B.C.} \quad T(\pm\infty,y,z,t) = T_o, \quad T(x,\pm\infty,z,t) = T_o, \quad T(x,y,\pm\infty,t) = T_o \qquad (6)$$

$$\text{where} \quad Q(x,y,z,t) = q(t)\delta(x-x_o)\delta(y-y_o)\delta(z-z_o) \qquad (7)$$

If the heat is liberated at a rate q(t) from t = 0 to t = t_f at the point (x_o, y_o, z_o), an analytical solution for the continuous point heat source is then (Carslaw and Jaeger, 1986):

$$T(x,y,z,t) = T_o + \frac{R_T^{1/2}}{8\left(\pi^3 \kappa_x \kappa_y \kappa_z\right)^{1/2}} \int_0^{t_f} \frac{q(t)}{(t_f - t)^{3/2}} \cdot \exp\left\{ \frac{-R_T}{4(t_f - t)} \left[\frac{(x - x_o - u_x/R_T \cdot (t_f - t))^2}{\kappa_x} + \frac{(y - y_o)^2}{\kappa_y} + \frac{(z - z_o)^2}{\kappa_z} \right] \right\} dt \tag{8}$$

2.3. Inverse Modeling Algorithm

For the above heat transport model, we formulated algorithms to identify the best-estimate of the model heat source parameters (i.e., source location coordinates). The minimization problem is solved using the L-M method, which is a descent method that has been used to identify source terms in contaminant transport models (Levenberg 1944; Marquardt 1963; Sciortino et al. 2000; Sun 1995). The objective function $\mathbf{\Phi}$ which minimizes sum of the squared errors between the calculated and observed temperature values and inverse algorithm used for parameter estimation of source location are:

$$\min \ \mathbf{\Phi}(\mathbf{p}) = \sum_{m=1}^{M} \omega_m^2 \left(T_m(\mathbf{p}) - T_m^{obs}\right)^2 \tag{9}$$

$$\mathbf{p}_{n+1} = \mathbf{p}_n - \left(\mathbf{A}_n^{\mathrm{T}} \mathbf{W} \mathbf{A}_n + \lambda \mathbf{I}\right)^{-1} \mathbf{A}_n^{\mathrm{T}} \mathbf{W} \mathbf{f}_n \tag{10}$$

where $T_m(\mathbf{p})$ is the calculated temperature, T_m^{obs} is the observed temperature, M is the number of observations in space, ω_m is a weight associated to measurement m, $\mathbf{p} = (x_o, y_o, z_o)$ is the source location to be identified, n is the iteration number, the term inside the parentheses is an approximation of the Hessian matrix, $\mathbf{A}_n$ is the matrix of partial derivatives (the Jacobian matrix) of the temperature function, and $\mathbf{f}_n$ is the vector of residual values for the current iteration:

$$\mathbf{f}_n = \begin{bmatrix} T_1(p) - T_1^{obs} \\ \vdots \\ T_M(p) - T_M^{obs} \end{bmatrix} \tag{11}$$

In (10), $\lambda\mathbf{I}$ is a correction term, where λ is an adjustable constant and $\mathbf{I}$ is the identity matrix. $\mathbf{W}$ is a diagonal matrix whose elements represent the weights ω_m associated with each element of the residual vector $\mathbf{f}_n$. The purpose of this term is to guarantee that the estimated objective function decreases from one iteration to the next, and that the parameter vector is within the range of admissible values.

3. Experimental ENS Test Bed

The experiment was performed in a three-dimensional intermediate-scale physical groundwater test bed that has been detailed elsewhere (Dela Barre et al., 2002; Kim et al., 2005). The test bed consists of a 1.5×0.5×0.4 m glass tank containing a water-saturated sandy porous medium as shown in Figure 1. Framed stainless steel screening is used to fabricate constant head boundaries at the influent and effluent ends of the tank. Steady, unidirectional flow through the sand is achieved by constant peristaltic pumping into the influent clear well (Masterflex® Model 7420, Cole-Parmer, Vernon Hills, IL), while maintaining constant head conditions in the effluent clear well using a weir. The model groundwater system is packed with homogeneous, clean sand (nominal grain diameter 0.33 mm, Lonestar Sand, Monterey, CA). The sandy medium is saturated with water to an average depth of 12 cm. The final porosity and bulk density of the model aquifer are determined to be 0.38 and 1.60 g/cm^3. The heat source is introduced at a fixed location as a continuous 31.6 mL/min source of warm water via the same peristaltic pump equipped with small precision tube (1.6 mm i.d., Masterflex® L/S™ 14). Sixteen thermocouples (J type, 1.5 mm o.d.) are deployable in positions indicated in Figure 1 to monitor three-dimensional temperature distributions resulting from point source injection. One thermocouple is fixed to the outlet of the warm water injection tube to monitor the source temperature, which was relatively constant (see inset plot in Figure 1).

The automated monitoring system of National Instruments data acquisition system (NI-DAQ) is connected to the distributed and networked temperature sensors. The thermocouple network is connected to signal conditioning and analog-to-digital switching modules (NI SCXI-1303, -1326) mounted on a 12-slot chassis (NI SCXI-1001). This chassis serves to power the SCXI modules while handling timing, trigger, and signal routing between the digitizer and SCXI modules. The chassis is connected to 16-bit data acquisition (DAQ) card (NI DAQCard™-AI-16XE-50), which is connected to a PC via a PCMCIA card. The DAQ system is controlled using LabVIEW (v6.1, National Instruments, 2001) software, which employs an object-oriented programming language called Virtual Instrument (VI) to create graphical users interface (GUI) for creating input and displaying output. MATLAB (v6.5, The MathWorks, Inc., 2002) routines are embedded into the LabVIEW program using the MATLAB ActiveX automation server to support the mathematical modeling and parameter estimation approaches discussed above.

The ENS design algorithm for source location identification described above was developed and integrated into the LabVIEW-based DAQ system using Matlab. The design algorithm is executed as follows: (1) an initial guess of the source location along with the number of sensors and monitoring period are selected prior to the test, (2) the GA-based strategy is used to identify the optimum set of sampling locations, (3) the sensors are manually deployed, (4) the L-M inverse modeling algorithm is used to update the estimate of the source location based on sensor readings, (5) the updated source location is then implemented in subsequent iterations to improve accurate source identification (with expanded budget for additional sensors). A flow diagram summarizing the ENS design algorithm is shown in Figure 2. To facilitate convergence of the GA, the searchable source region was restricted to include the upstream portions of the test bed and potential sensor locations to areas to integer number coordinates downstream of the searchable heat source region (see Figure 2 for defined regions).

The proposed ENS design algorithm was tested for two transient heat transport experiments comprising of single source (Figure 1) and dual sources (Figure 8). For each case, an estimate for the source location(s) was provided to the GA for determination of the initial two sensor locations. In the single-source case, the initial source estimate (x = 10 cm, y = 40, z = 10) was grossly incorrect compared to the actual source location (25, 25, 6). The investigators immediately installed the sensors at locations dictated by the GA and the observed temperature histories were then used to update the source location estimate per the L-M algorithm. Given the updated source location, the GA then proposed locations for two additional sensors. This process was repeated until an 8-sensor network was achieved.

4. Results and Discussion

The release of the warm water stream results in the propagation of a 3D quasi-Gaussian 3D temperature distribution in the groundwater flow direction. Sensors placed downstream exhibit the arrival of a dispersed temperature front which builds to a steady response. In order to use the mathematical heat transport model in the context of the ENS design problem, key parameters which cannot be independently estimated must be determined through model calibration. It is important to emphasize that this calibration step is a necessary and important step in most environmental modeling efforts. The reason for this is that in spite of the fact that the physical transport of mass and energy in environmental systems are reasonably well-understood, environmental media (and the associated model parameters) are typically comprised of distributed properties which are difficult to independently determine. In a previous investigation using this test bed (Kim et al. 2005), a nonlinear least squares regression algorithm was employed to estimate the thermal dispersion coefficients for the test bed sand-water system as K_x = 0.6 cm^2/min and K_y =K_z = 0.48 cm^2/min by fitting the analytical solution (8) to the transient temperature values observed at the prefixed location of x=65 cm, y=25 cm, and z=8 cm for the flow velocity of 7.8 cm/h. Figure 3 demonstrates the agreement between the experimental data and the best-fitting model temperature history for these thermal dispersion coefficient estimates.

The real-time sequential ENS design procedure is demonstrated in Figures 4 through 7 for the cases of 2, 4, 6, and 8 sensors, respectively. For each of these figures, the plots exhibit the GA-determined sensor locations, the L-M-based source prediction (from the perspective of the x-y and x-z planes), and a comparison of the simulated and observed temperature histories at the determined sensor locations. The resulting progression of the ENS design alternates between locations directly downstream of the source and symmetrically arranged off-center sensors. As more sensors are added, the same pattern propagates downstream. As expected, this behavior suggests that, for a steady-state temperature distribution, the best sensor network characterizes the gradient some optimal distance downstream of the source.

For the specific experimental conditions applied, this distance was roughly 10 to 15 cm downstream from the source. As these locations become occupied, subsequent choices are forced downstream where the gradient is less sharp and therefore contributes less to the source delineation effort.

The sequential approach is advantageous relative to a single-step design in that inaccurate initial estimates can be corrected using data from the first sensor deployments, and then refined by additional deployments. For the single-source case, the initial source estimate is errant by 15, 15 and 4 cm in the x-, y- and z-coordinates. In Figure 4, both longitudinally and with respect to depth, the source location estimate error has been reduced to 2 to 3 cm by the addition of the 2 sensors. With respect to the y-coordinate, the estimate error has been reduced to between 5 and 6 cm. Considering the error of the initial source estimate, these results are encouraging. With two additional sensors, the estimates improve greatly in the x-direction, closing to within 1 cm. The z-coordinate is slightly improved, but the y-coordinate estimate is less accurate than before. For the next two sensors (6 in total), the x- and z-coordinates are highly accurate, while the y-coordinate remains errant by 5 to 6 cm. For 8 total sensors, the accuracy of all coordinates decreases slightly.

Regardless of the trends in accuracy of the source location, Figures 4d through 7d demonstrate that model agreement with sensor observations continuously improves with the increasing spatial granularity of sensor network, as does the certainty in the model simulations (as determined by the 95% confidence interval). These results demonstrate that it is possible for the ENS results to be biased so as to report less accurate results with more certainty. This undesirably results may be an artifact of sensor errors, environmental model error, or both. While the sensors employed in this work are both accurate and precise, our ability to install them at a particular location is less so and clearly a source of random error. However, our prior work with this test bed (Kim et al. 2005) suggests that the systematic inability to correctly locate the heat source is more likely a result of the discrepancy between the uniform flow model employed and the real system in which a flow field perturbation was caused by the source injection. This perturbation could be simulated using a more complex numerical groundwater flow model, but such a model would be more difficult to invert in a computationally efficient manner. Numerical model inversion will undoubtedly need to be addressed as ENS extends to more realistic systems.

The dual source experiment illustrates a shortcoming in this ENS design strategy: the inability to distinguish between individual and multiple sources. The sources, which were identical in strength, created the temperature distributions depicted by the contour plots in Figure 9. As the ENS design sequence logged in Figures 10 through 13 indicates, the design algorithm initially distinguishes between the two sources in the y-coordinate. However, as the experiment continues, and the heat plumes from the dual sources intermix, the ENS design algorithm loses the capacity to make this distinction, concluding that the dual sources a aligning in the direction of the groundwater flow. This result is an indication that the relatively simplistic descent method employed for updating the source location will not adequately address the problem of multiple sources, and may result in a non-optimal ENS design. One possible way to eliminate such a non-optimal design from consideration is to monitor the evolution of the model fit with the data. In contrast to the single-source cases discussed above, significant discrepancies between the observed and simulated temperature histories arise, most obviously in Figures 12d and 13d. For a well-defined system, such as the test bed employed here, discrepancies on this order are indicative of a design progression gone awry. In more realistic systems, disagreement between simulations and observations are more common, however, and these types of errors will be more difficult to detect and remedy. This aspect of environmental ENS design warrants further research.

5. Summary and Conclusions

This work develops the concept of creating real-time decision-making algorithms in support of automating embedded networked sensing (ENS) design in an environment-specific context. The design algorithm used to demonstrate the concept combines a GA to solve the combinatorial optimization problem associated with identifying the best combination of limited number of sensor locations for identifying a heat source location, and a descent-based inverse modeling algorithm for updating the location estimate. While such algorithms have been used in computational experiments and for *post priori* analyses, their integration into a real-time ENS design algorithm based on the environmental medium in question is the key to successful design in such systems. The algorithm was validated for the case of a single heat source. However, the optimal inversion of sensor readings to identify the source location maintained a significant bias, suggesting that the phenomenological inverse model was overly simplistic relative to the experimental system. The algorithm failed to correctly identify the presence of dual sources, pointing to the need for incorporation of additional intelligence in the ENS design algorithm. Overall, the results of this demonstration point to the need for collaborative research between artificial intelligence, distributed networking, and environmental systems investigators to better couple learning and decision-making algorithms with realistic distributed parameter environmental simulators in the context of creating robust ENS designs for a variety of applications.

Acknowledgements

This work was supported by the Post-doctoral Fellowship Program of Korea Science & Engineering Foundation (KOSEF) and UCLA's Center for Embedded Networked Sensing (CENS) under cooperative agreement #CCR-0120778 with the National Science Foundation.

References

AIB, 2004. A Plan for Developing and Governing of the National Ecological Observatory Network, Report from the NEON Coordination and Implementation Conference, American Institute of Biological Sciences.

Bird, R.B., Stewart, W.E., Lightfoot, E.N., 2002. Transport Phenomena, 2nd ed. John Wiley & Sons, New York.

Carroll, D.L., 1996. Chemical laser modeling with genetic algorithms. American Institute of Aeronautics and Astronautics Journal 34 (2), 338-346.

Carslaw, H.S., Jaeger, J.C., 1986. Conduction of Heat in Solids, 2nd ed. Oxford University Press, New York.

Cieniawski, S.E., Eheart, J.W., Ranjithan, S., 1995. Using genetic algorithms to solve a multiobjective groundwater monitoring problem. Water Resources Research 31 (2), 399-409.

Cleveland, T.G., Yeh, W.W.G., 1990. Sampling network design for transport parameter identification, Journal of Water Resources Planning and Management, 116 (6), 764-783.

CUAHSI, 2003. A National Center for Hydrologic Synthesis: Scientific Objectives, Structure and Implementation, Report from a CUAHSI workshop, Santa Barbara, CA, July 10-12. [http://www.cuahsi.org/docs/NCHS_Workshop_Report.pdf]

Culler, D.E., Mulder, H., 2004. Smart Sensors to Network the World. Scientific American, June issue.

De Schaetzen, W.B.F., Walters, G.A., Savic, D.A., 2000. Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms. Urban Water 2 (2), 141-152.

Dela Barre, B.K., Harmon, T.C., Chrysikopoulos, C.V., 2002. Measuring and modeling the dissolution of nonideally shaped dense nonaqueous phase liquid pools in saturated porous media. Water Resources Research 38 (8), 1133, doi:10.1029/2001WR000444.

Eckert, E.R.G., Drake, R.M. Jr., 1987. Analysis of Heat and Mass Transfer, Hemisphere Publishing Corporation, Washington.

Giacobbo, F., Marseguerra, M., Zio, E., 2002. Solving the inverse problem of parameter estimation by genetic algorithms: the case of a groundwater contaminant transport model. Annals of Nuclear Energy 29 (8), 967-981.

Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Computers and Operations Research 13 (5), 533–549.

Glover, F., Laguna. M., 1997. Tabu Search, Kluwer Academic Publishers, Dordrecht, Netherlands.

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning, Addison

Wesley Longman, Reading, Massachusetts.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor, Michigan.

Hudak, P.F., Loaiciga, H.A., Marino, M.A., 1995. Regional-scale ground water quality monitoring via integer programming. Journal of Hydrology 164 (1-4), 153-170.

Jaramillo, J.H., Bhadury, J., Batta, R., 2002. On the use of genetic algorithms to solve location problems. Computers & Operations Research 29 (6), 761-779.

Kim, J.Y., Park, Y.J., Harmon, T.C., 2005. Real-Time Model Parameter Estimation for Analyzing Transport in Porous Media. Ground Water Monitoring and Remediation, in press.

Kirkpatrick, S., Gelat, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680.

Knowles, I., Yan, A., 2004. On the recovery of transport parameters in groundwater modeling. Journal of Computational and Applied Mathematics 171 (1-2), 277-290.

Levenberg, K., 1944. A method for the solution of certain problems in least squares. Quarterly of Applied Mathematics 2, 164-168.

Marquardt, D., 1963. An algorithm for least-squares estimation of nonlinear parameters. SIAM Journal on Applied Mathematics 11, 431-441.

Mayer, A.S., Huang, C., 1999. Development and application of a coupled-process parameter inversion model based on the maximum likelihood estimation method. Advances in Water Resources 22 (8), 841-853.

Michalewicz, Z., 1996. Genetic Algorithms + Data Structures = Evolution Programs. third, revised and extended ed. Springer, Berlin, German.

Michalewicz, Z., Fogel, D.B., 2002. How to Solve It: Modern Heuristics. Springer, Berlin.

NAP, 2001. Embedded, Everywhere: A Research Agenda for Networked Systems of Embedded Computers. National Academies Press, Washington D.C.

National Instruments, 2001. LabVIEW v6.1. Available from http://www.ni.com/

NRC, 2003. NEON: Addressing the Nation's Environmental Challenges, Committee on the National Ecological Observatory Network, National Research Council, Washington, D.C.

NSF, 2003. Environmental Cyberinfrastructure Needs for Distributed Sensor Networks, Report from National Science Foundation sponsored workshop, Scripps Oceanographic Institute, La Jolla, CA, Aug 12-14.

Sciortino, A., Harmon, T.C., Yeh, W.W.G., 2000. Inverse modeling for locating dense nonaqueous pools in groundwater under steady flow conditions. Water Resources Research 36 (7), 1723-1735.

Sciortino, A., Harmon, T.C., Yeh, W.W.G., 2002. Experimental design and model parameter estimation for locating a dissolving dense nonaqueous phase liquid pool in groundwater. Water

Resources Research 38 (5), 1057, doi:10.1029/2000WR000134.

Sun, N.Z., 1995. Inverse Problems in Groundwater Modeling, Kluwer Academic Publishers, Dordrecht, Netherlands.

Sun, N.Z., Sun, A.Y., 2002. Parameter identification of environmental systems, In: Shen, H.H. et al. (Eds.), Environmental Fluid Mechanics: Theories and Applications. American Society of Civil Engineers, Reston, Virginia, pp. 297-337.

The MathWorks, Inc., 2002. MATLAB v6.5. Available from http://www.mathworks.com/

Wagner, B.J., 1992. Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modeling. Journal of Hydrology 135 (1-4), 275-303.

Woodbury, A., Sudicky, E., Ulrych, T.J., Ludwig, R., 1998. Three-dimensional plume source reconstruction using minimum relative entropy inversion. Journal of Contaminant Hydrology 32 (1-2), 131-158.

Yeh, W.W.G., 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. Water Resources Research 22 (2), 95-108.

Yeh, W.W.G., Yoon, Y.S., 1981. Aquifer parameter identification with optimum dimension in parameterization. Water Resources Research 17 (3), 664-672.

Youssef, H., Sait, S.M., Adiche, H., 2001. Evolutionary algorithms, simulated annealing and tabu search: a comparative study. Engineering Applications of Artificial Intelligence 14 (2), 167-181.

Zheng, C., Wang, P., 1996. Parameter structure identification using tabu search and simulated annealing. Advances in Water Resources 19 (4), 215-224.

Biographies

Juyoul Kim received a Ph.D. in radioactive waste disposal from Seoul National University in 2002 and currently holds a position as postdoctoral researcher in Department of Civil & Environmental Engineering at University of California, Los Angeles. His research interests include numerical simulation of water flow, solute, and heat transport in unsaturated and saturated media, parameter estimation of nonlinear transport model using inverse technique and heuristic approach, geostatistical application to site characterization and environmental monitoring of waste disposal. He may be reached at Department of Civil & Environmental Engineering, University of California, Los Angeles, 5732 Boelter Hall, Los Angeles, CA 90095-1593.

Thomas C. Harmon is an Associate Professor in the School of Engineering and Founding Faculty member at the University of California, Merced. Prior to this, he was a faculty member in the Department of Civil & Environmental Engineering at UCLA from 1992-2003. He received a B.S. in Civil Engineering from the Johns Hopkins University in 1985, and M.S. and Ph.D. degrees from the Environmental Engineering program at Stanford University in 1986 and 1992, respectively. He currently directs the contaminant transport monitoring thrust of the Center for Embedded Networked Sensing at UCLA (http://cens.ucla.edu). His research and teaching is concerned with the fate and transport of chemicals in the subsurface environment (https://ucmeng.net/people/tharmon).

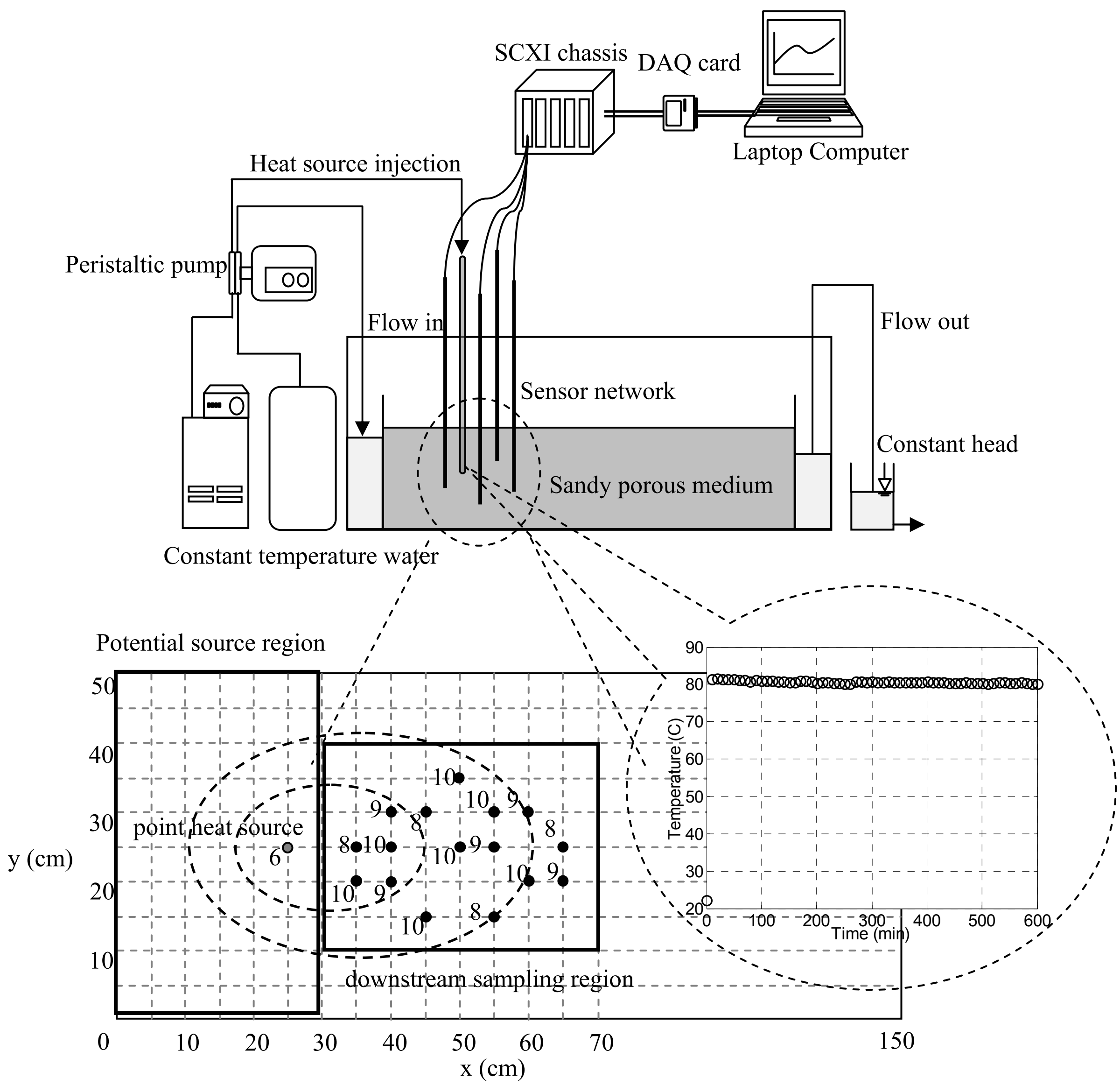

Figure 1. Plan view of the experimental layout for exercising the real-time monitoring network design algorithm, and temperature history from a point heat source. The numbers denote the elevation (z coordinate) of the potential sensor location above the aquifer bottom.

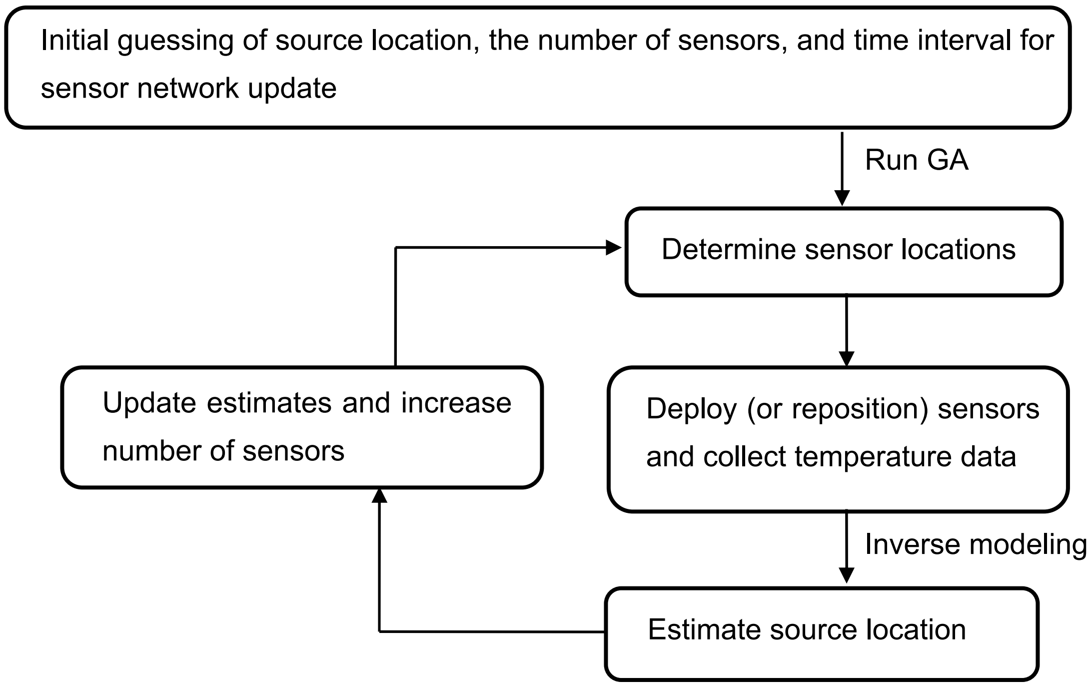

Figure 2. The proposed experimental sequences for real-time monitoring network design using GA and L-M algorithm

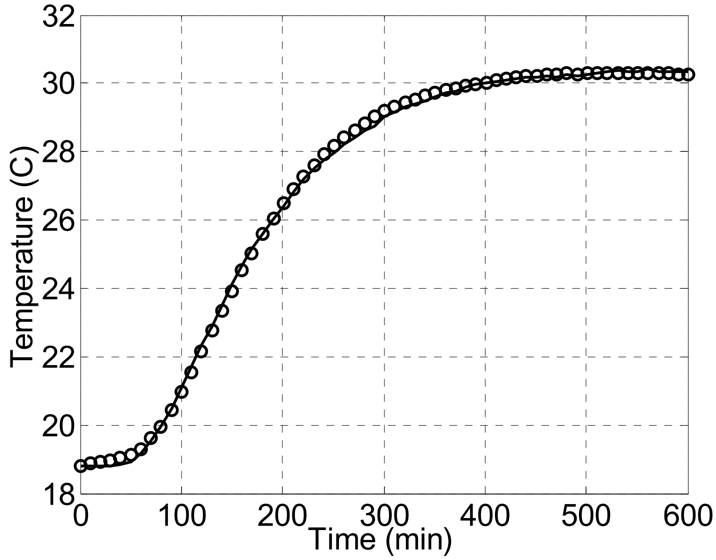

Figure 3. Observed temperature data (open circles) and simulated temperature history (solid curve) using estimated thermal dispersion coefficients

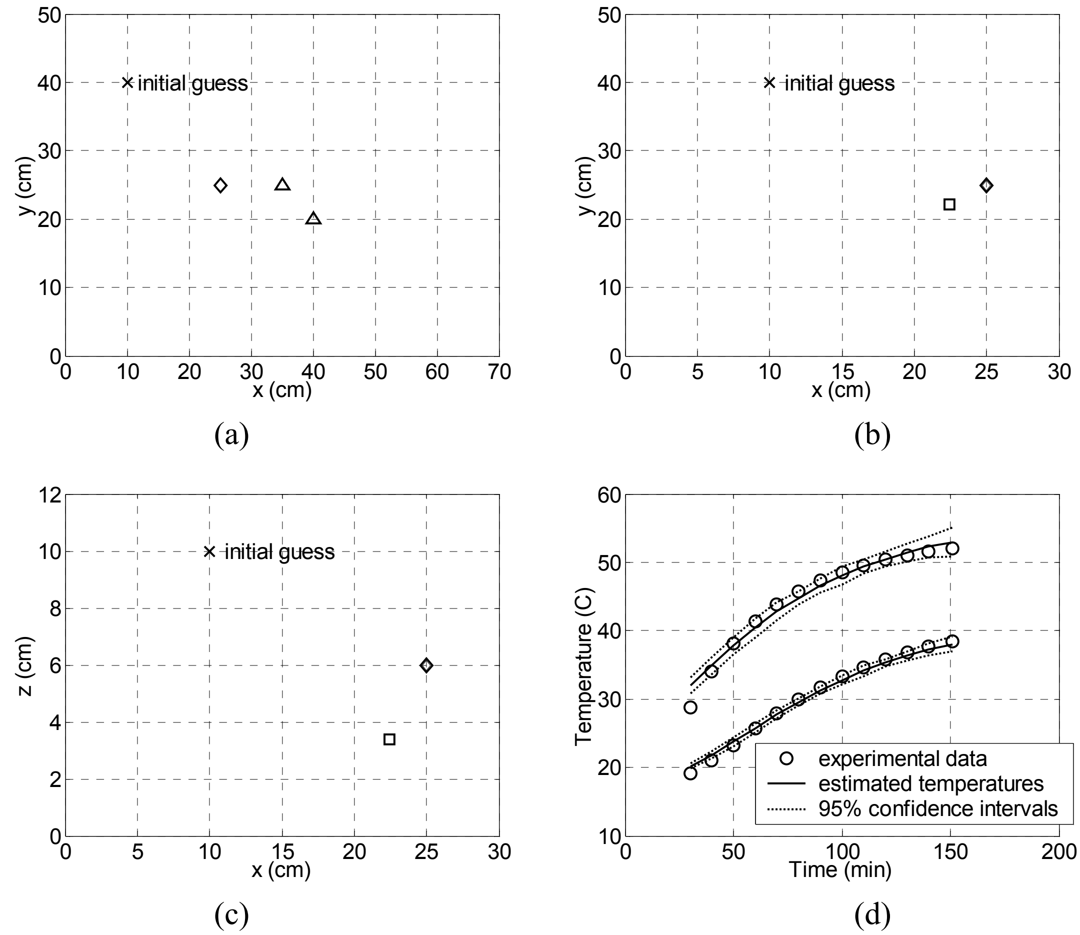

Figure 4. 2-sensor design with an initial guess of (10, 40, 10) for 0.5 ~ 2.5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

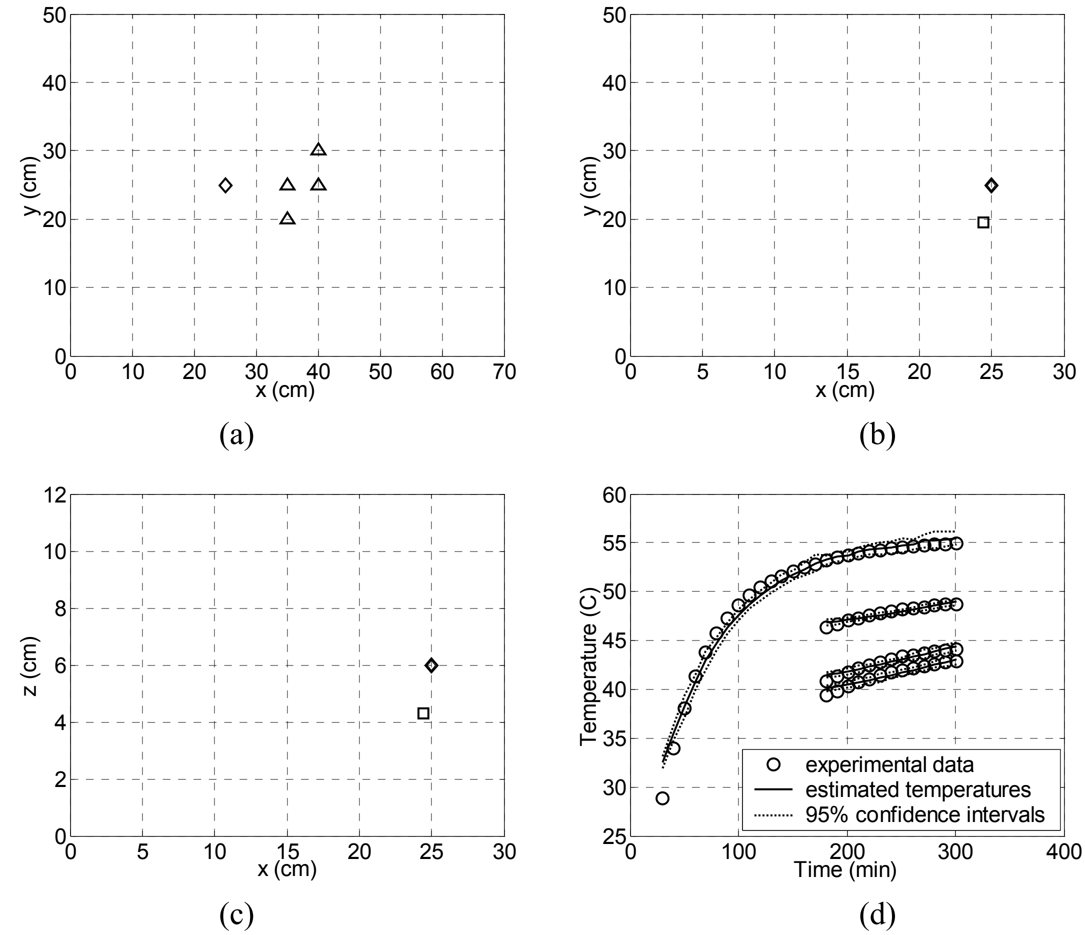

Figure 5. 4-sensor design with an initial guess of (10,40,10) for 3 ~ 5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

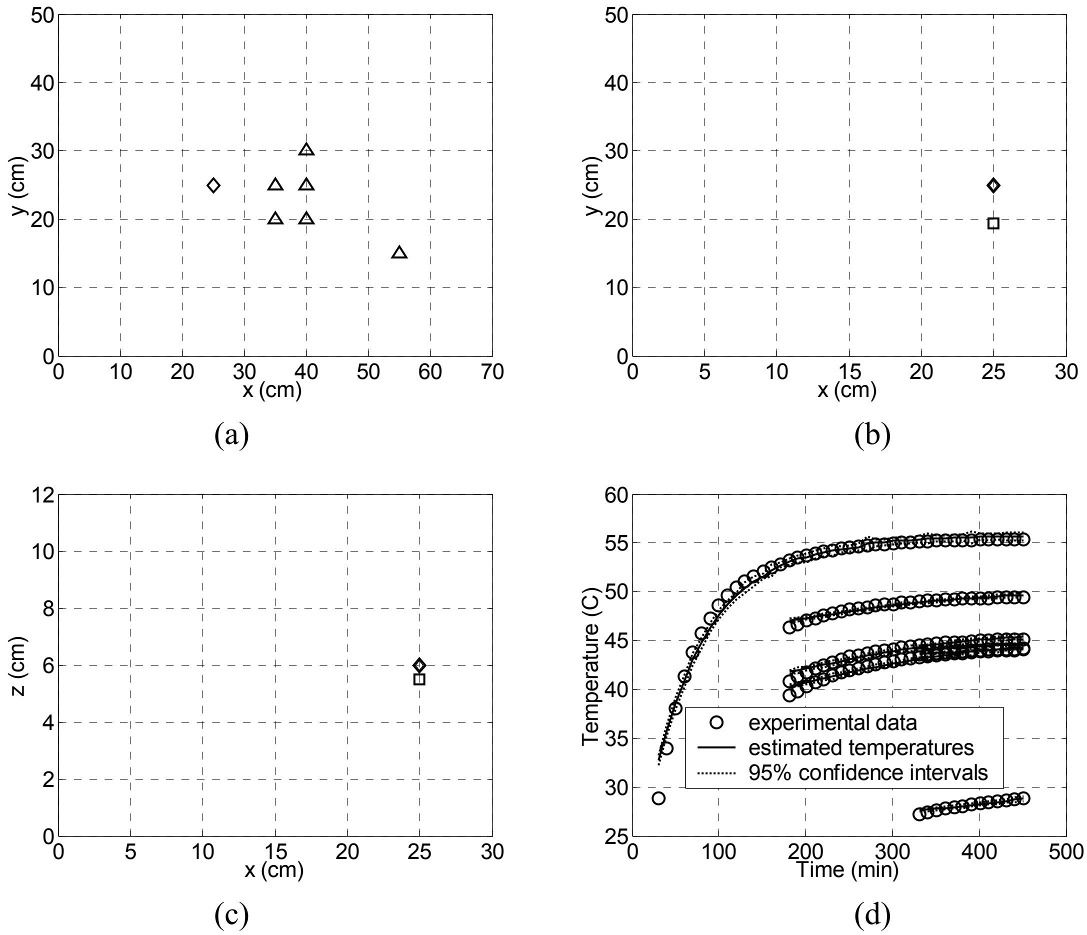

Figure 6. 6-sensor design with an initial guess of (10,40,10) for 5.5 ~ 7.5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

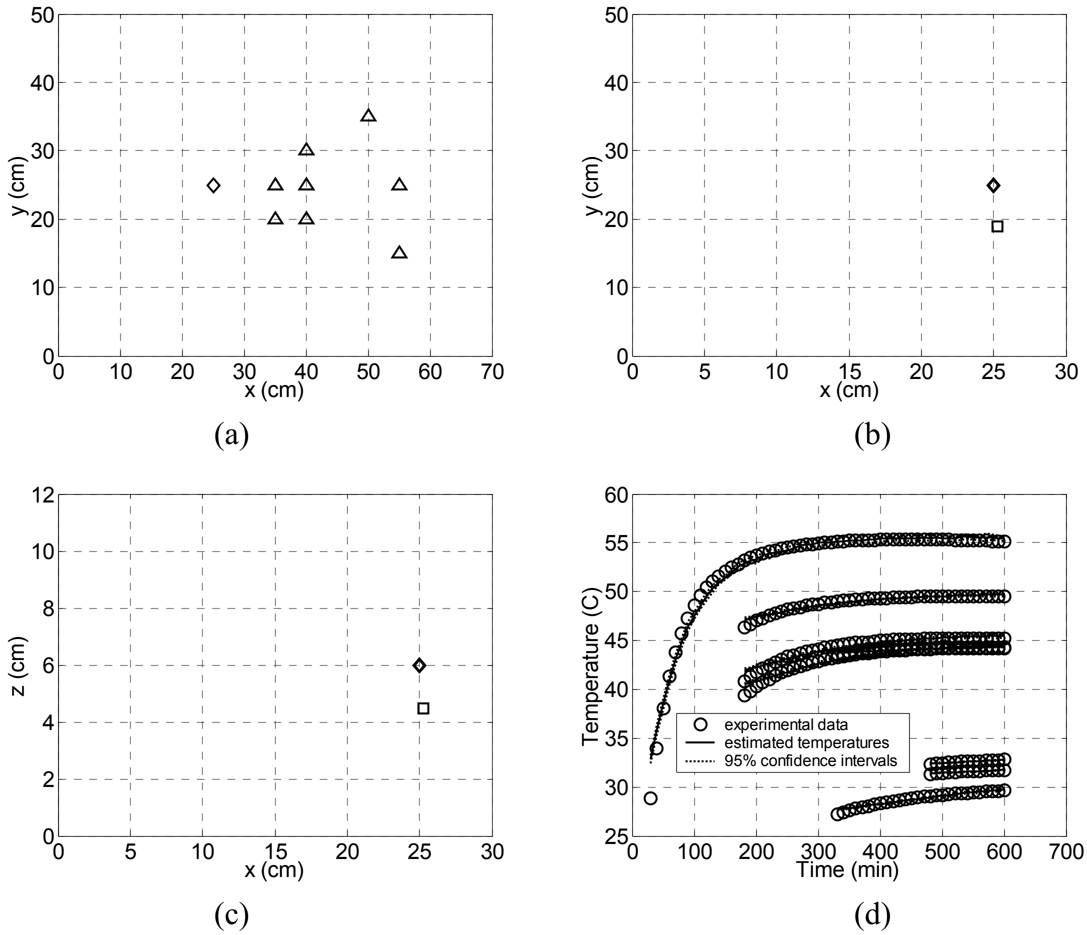

Figure 7. 8-sensor design with an initial guess of (10,40,10) for 8 ~ 10 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

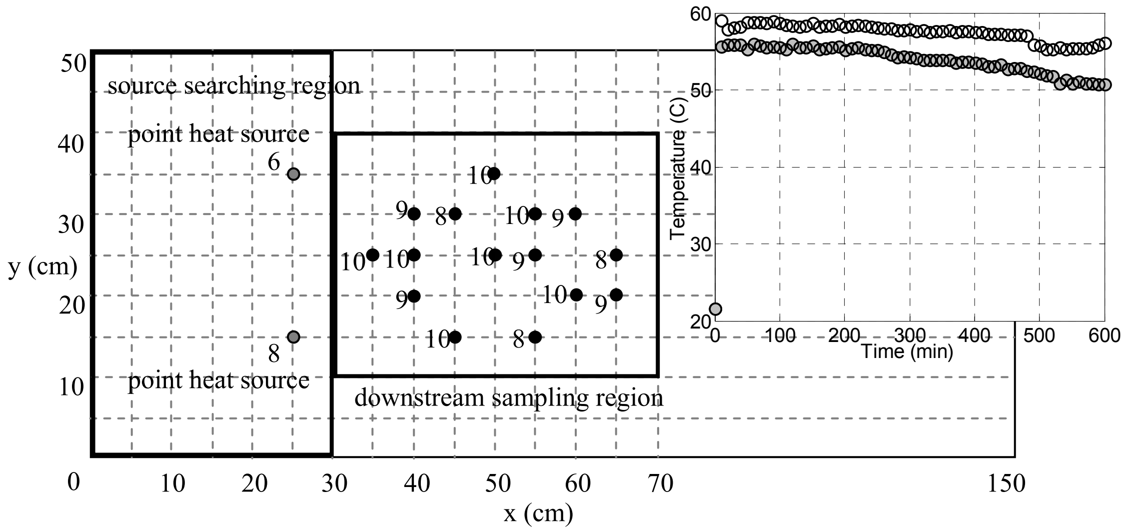

Figure 8. Plan view of the multiple heat experimental layout and potential sensor locations for exercising the real-time monitoring network design algorithm, and temperature history from two point heat sources

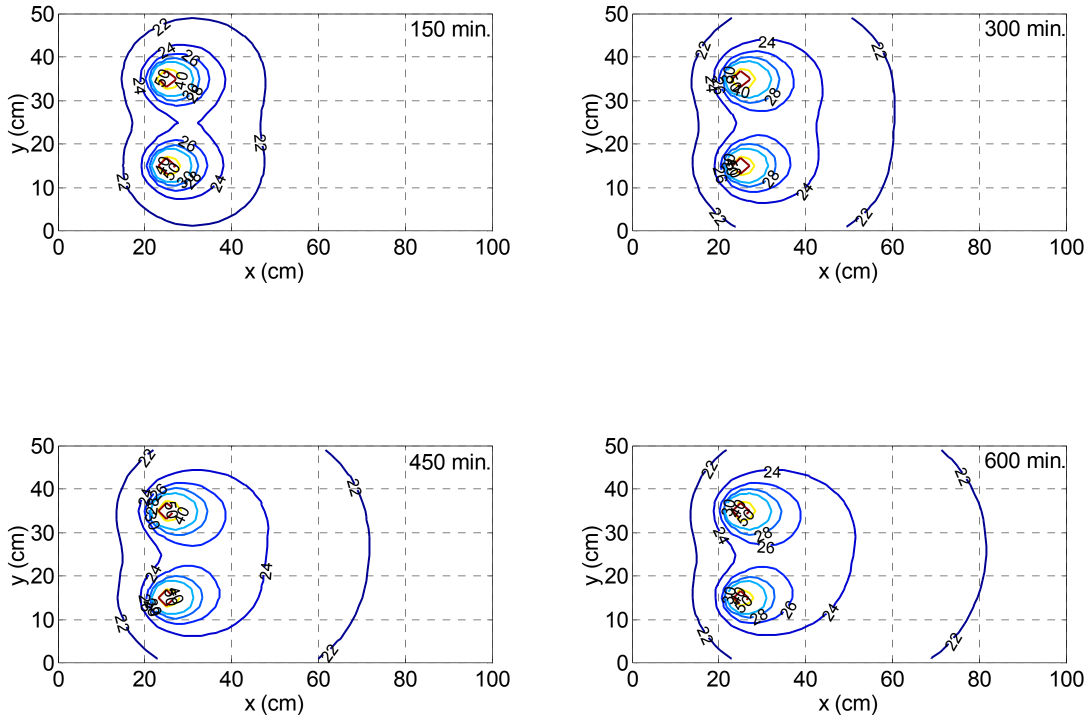

Figure 9. Snapshots of computed temperature contours from two point sources over a time horizon of 10 hours, with time intervals of 2.5 hours. Results are shown only in the middle of aquifer (z=6 cm) for convenience.

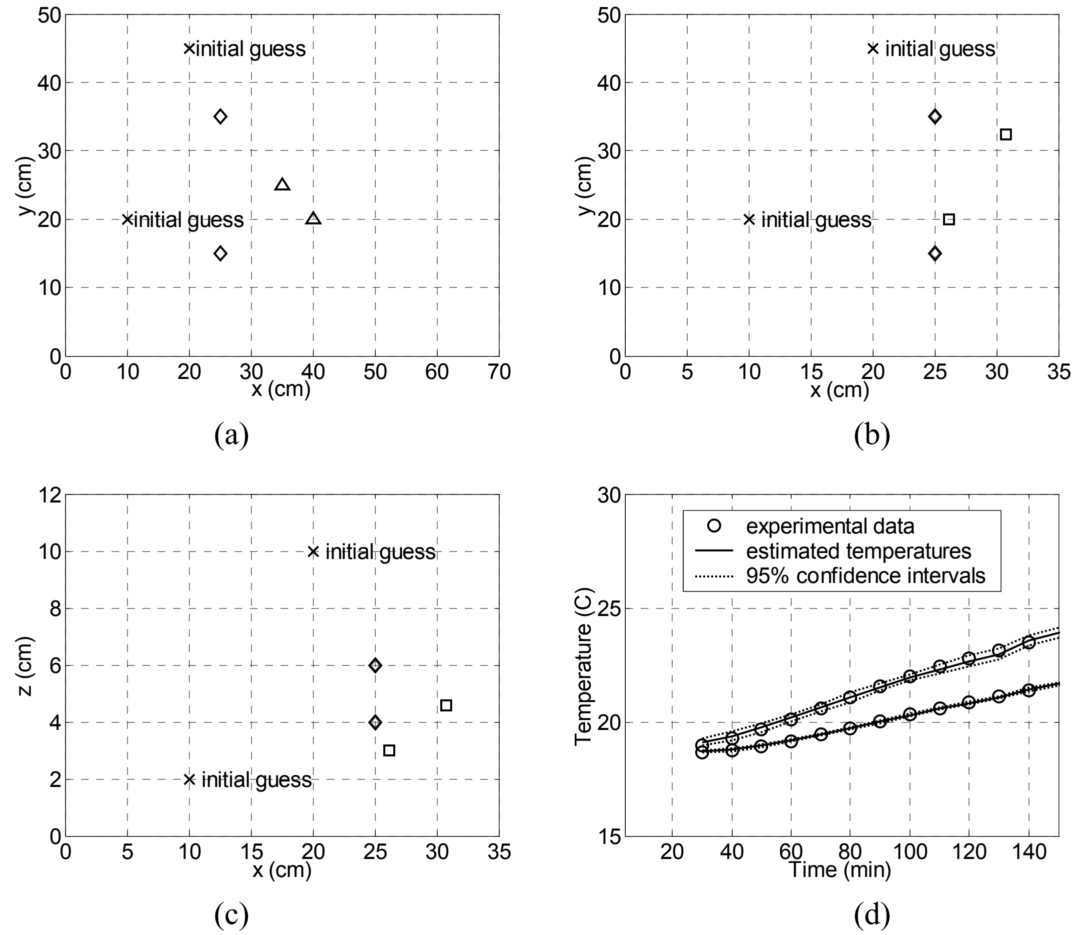

Figure 10. 2-sensor design with two initial guesses of (10,20,2) and (20,45,10) for 0.5 ~ 2.5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

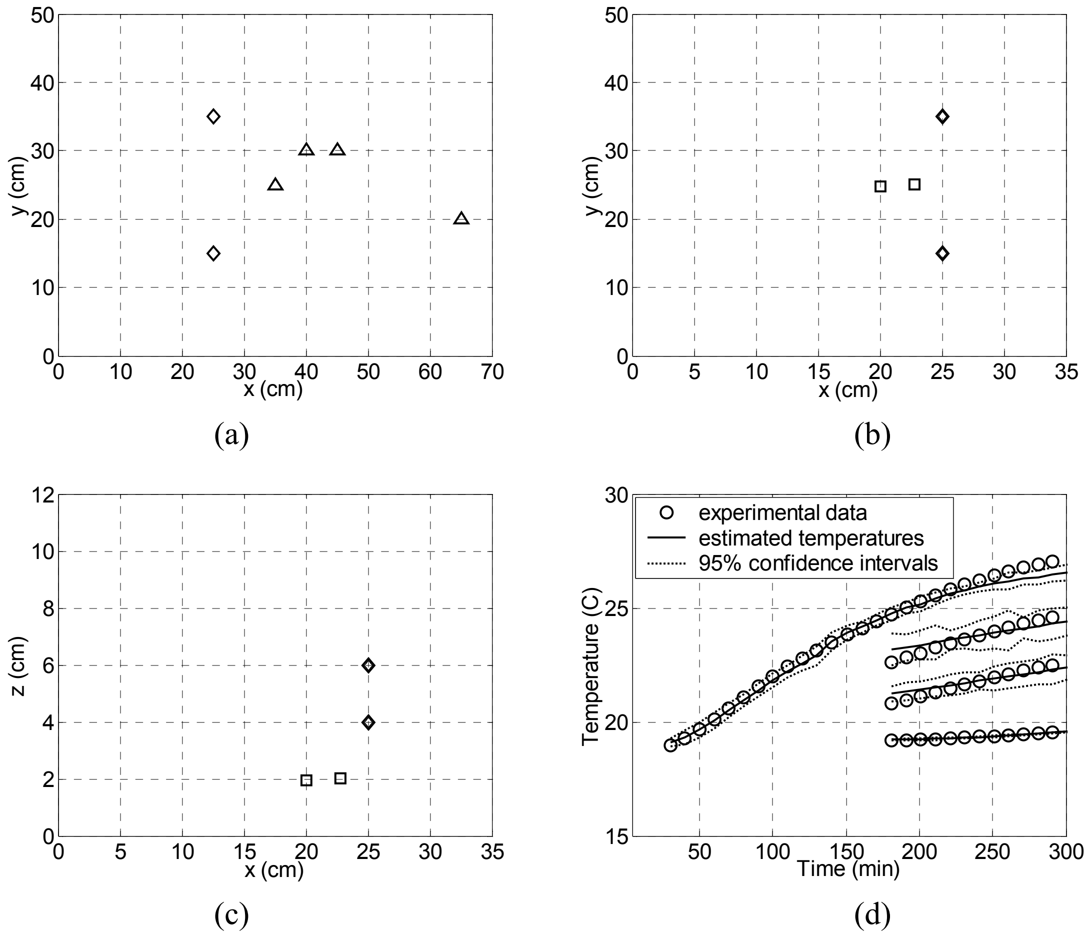

Figure 11. 4-sensor design with two initial guesses of (10,20,2) and (20,45,10) for 3 ~ 5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

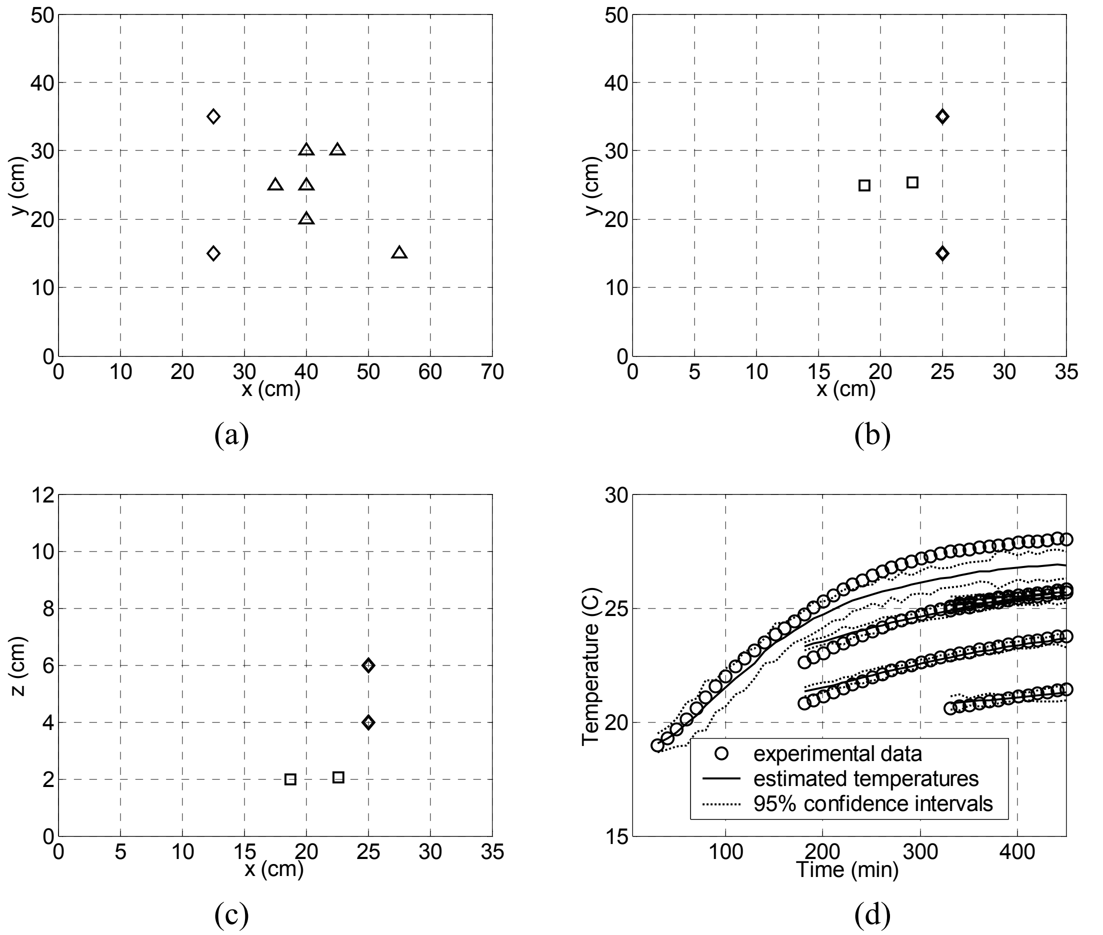

Figure 12. 6-sensor design with two initial guesses of (10,20,2) and (20,45,10) for 5.5 ~ 7.5 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).

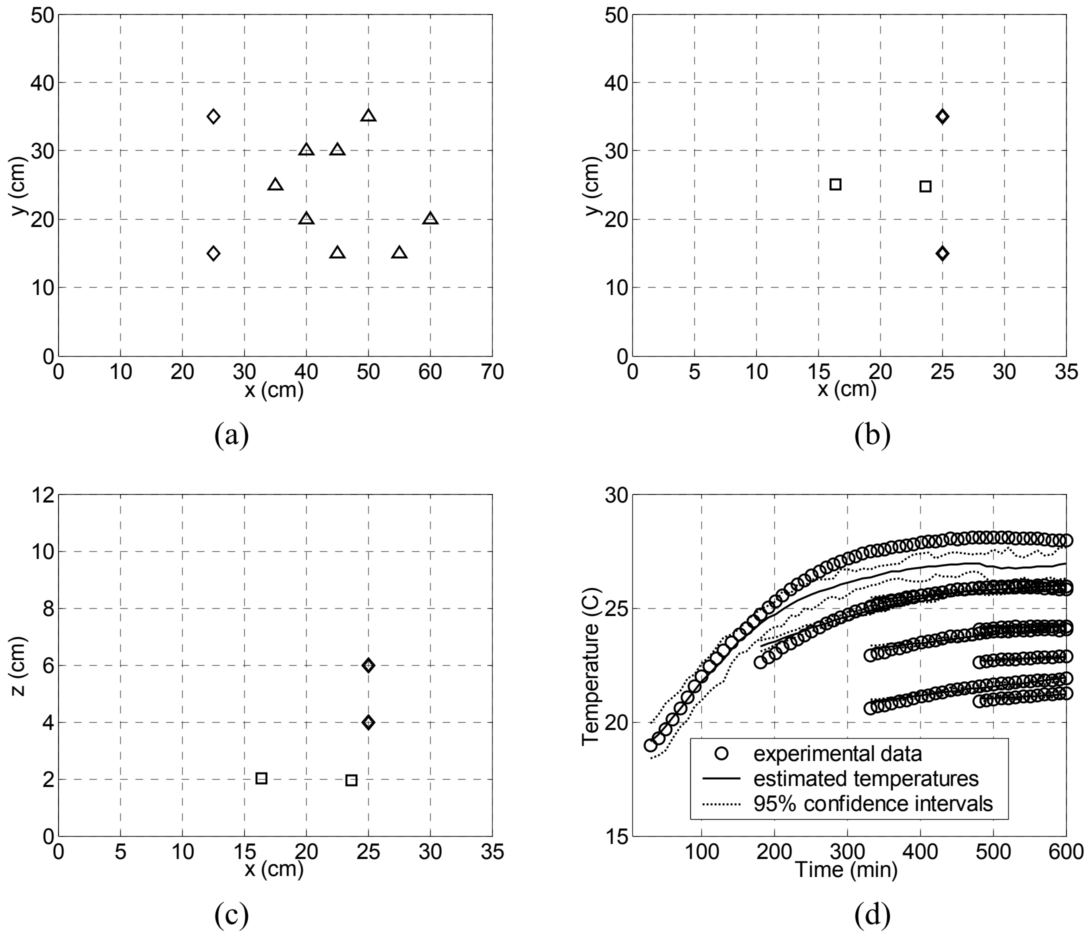

Figure 13. 8-sensor design with two initial guesses of (10,20,2) and (20,45,10) for 8 ~ 10 hours: (a) monitoring network design with GA, (b and c) source identification with L-M, and (d) comparison between experimental data and estimated temperatures for the optimal sensor locations selected (◊ is true source location, Δ are optimal sensor locations, □ are estimated source locations, and ○ are sensor readings).