Journal of Water Resource and Protection, 2022, 14, 542-570
https://www.scirp.org/journal/jwarp
ISSN Online: 1945-3108
ISSN Print: 1945-3094

Integrated Sequential Groundwater Contaminant Source Characterization and Pareto-Optimal Monitoring Network Design Application for a Contaminated Aquifer Site

Hamed K. Esfahani[1*], Adrian Heggie[2], Bithin Datta[3*]

[1]Urban Development Department, Stantec, Sydney, Australia
[2]Technical Executive, Contaminated Land Management, WSP, Sydney, Australia
[3]College of Science and Engineering, James Cook University, Townsville, Australia
Email: *Hamed.ehesfahani@Stantec.com, *bithin.datta@jcu.edu.au

How to cite this paper: Esfahani, H.K., Heggie, A. and Datta, B. (2022) Integrated Sequential Groundwater Contaminant Source Characterization and Pareto-Optimal Monitoring Network Design Application for a Contaminated Aquifer Site. *Journal of Water Resource and Protection*, **14**, 542-570.
https://doi.org/10.4236/jwarp.2022.147029

Received: June 30, 2022
Accepted: July 24, 2022
Published: July 27, 2022

Abstract

Accurate and reliable groundwater contaminant source characterization with limited contaminant concentration monitoring measurement data remains a challenging problem. This study presents an illustrative application of developed methodologies to a real-life contaminated aquifer. The source characterization and optimal monitoring network design methodologies are used sequentially for a contaminated aquifer site located in New South Wales, Australia. Performance of the integrated optimal source characterization methodology combining linked simulation-optimization, fractal singularity mapping technique (FSMT) and Pareto optimal solutions is evaluated. This study presents an integrated application of optimal source characterization with spatiotemporal concentration measurement data obtained from sequentially designed monitoring networks. The proposed sequential source characterization and monitoring network design methodology shows efficiency in identifying the unknown source characteristics. The designed monitoring network achieves comparable efficiency and accuracy utilizing much smaller number of monitoring locations as compared to a more ideal scenario where concentration measurements from a very large number of widespread monitoring wells are available. The proposed methodology is potentially useful for efficient characterization of unknown contaminant sources in a complex contaminated aquifer site, where very little initial concentration measurement data are available. The illustrative application of the methodology to a real-life contaminated aquifer site demonstrates the capability and efficiency of the proposed methodology.

Keywords

Groundwater Source Characterization, Optimal Monitoring Network Design, Fractal Singularity Mapping Technique

1. Introduction

Remediation of contaminated aquifers poses many challenges. The most important one is the accurate and reliable identification of the unknown contaminant sources in terms of location, magnitude, and duration of activity. This process of characterizing unknown sources of groundwater contamination is a complex one due to uncertainties in modelling and predicting the flow and transport processes in a contaminated aquifer. Also, often the contaminants are reactive multiple species. This requires the modelling of multiple species reactive transport processes, which also includes accurate description of the geochemical processes. In a regional scale, the unknown contaminant source characterization involves solution of the inverse problem with inherent nonunique responses of the aquifer to sources, ill-posed characteristics due to lack of adequate field measurement data, and also the possibility of the contaminant sources being spatially distributed in nature. Therefore, accurate contaminant source(s) identification being a first essential step for contamination remediation, reliable methodologies are necessary to ensure computational feasibility of solving such inverse problems, and to increase the reliability of source estimation by designing and applying efficient contaminant monitoring networks. This study utilizes a linked simulation optimization approach [1] [2] to solve the optimal contamination source identification problem, and also proposes and evaluates a methodology for sequential design of a mentoring network along with a linked simulation optimization model solution to sequentially improve the accuracy of source identification in aquifers.

The local fractal singularity mapping technique (FSMT) is used in this study to delineate the contaminant plume occurring in the aquifer. This information is then utilized to design an optimal monitoring network to better identify the contaminant sources. Because, often, the existing monitoring networks are arbitrary and may only result in erroneous identification of potential sources, the monitoring network design and implementation are followed by solution of the optimal source identification model. The identified sources at that stage are then utilized to design a new optimal monitoring network for implementation. The monitoring data obtained from the new network is then utilized to solve the source identification inverse model. This integrated sequential process of source identification and monitoring improves the accuracy of source identification results iteratively. This approach is particularly useful for accurate and reliable identification of unknown contaminant sources, especially when the preliminary contaminant concentration data are sparse, and obtained from an arbitrary set of measurement locations resulting in inaccurate initial estimation of source locations, magnitudes, and time history.

As an effective step for determining reliable groundwater management and remediation strategies, linked simulation-optimization models are increasingly being used for identification of unknown groundwater pollution source. [3]-[9]

A suitable monitoring network can be utilized to increase the reliability of source characterization in terms of location, magnitude, and duration of activity. Potential well locations to obtain the optimal monitoring network design should be selected carefully after considering likely scenarios of contamination. One of the effective methods to select the potential well locations can be to first determine the contaminant plume boundary and then select potential monitoring well locations within the boundary.

This study utilizes s the fractal singularity mapping index methodology [3] [10] together with sequential use of source identification as well as improved monitoring network design and implementation. The main aim is to improve the accurate estimation of the sources, by a sequential process of source identification and monitoring network design. The fractal singularity mapping index approach helps to better delineate the contaminant transport plume boundary in a contaminated aquifer to increase the utility and efficiency of the optimal monitoring network design. The sequential process helps utilization of enhanced monitoring data based on updated to characterize unknown distributed pollution sources in a complex contaminated aquifer.

Optimal design and implementation of an aquifer contamination monitoring network is complicated due to the presence of uncertainties in predicting the plume movement, and is based on specified objectives of design, single or multiple. Reference [3], lists some of the relevant prior studies related to this issue. These include: for detection of contamination; [11] [12] related to the reducing the cost of monitoring; while multiple objectives groundwater monitoring network design [13]. Reference [14] [15] proposed using monitoring network design for source identification and redundancy reduction with feedback information. Sampling strategy in space and time using Kalman filter was proposed by Kollat and Reed [16] [17]. Long long-term monitoring network design using multi-objective simulation–optimization model under uncertainties was proposed by [9] [18].

Application of the fractals concept was introduced by [19] [20]. The applying of fractal and multifractal based methodologies in recent studies include: flooding [21], geoscience [22] [23] [24] [25] and groundwater contamination plume delineation [26]. Fractal models such as Number-Size model (N-S), Concentration-Area model (C-A) [27], Spectrum-Area model (S-A) [28], singularity index [29], and Concentration-Volume model (C-V) [30] were used for geochemical data analysis. In this study the local fractal singularity mapping technique (FSMT) as presented in [3] is utilized for improved monitoring network design utilizing the plume boundary delineated using the FSMT. However, this concept is applied and evaluated for a new real life urban contaminated aquifer study area in

Australia, incorporating sequential integrated source characterization and monitoring network design. This sequential iterative approach is shown to improve the source characterization accuracy.

This study presents an application of developed methodologies to a real-life contaminated aquifer. The linked source characterization and optimal monitoring network design methodologies are used sequentially for a contaminated aquifer site located in New South Wales, Australia. The location details of this site are not disclosed because of confidentiality requirements. In this study, the developed methodology for unknown contaminant sources characterization and FSMT are applied. The performance of the optimal source characterization methodology using the FSMT is applied to establish the potential applicability of this approach. In the following sections, first, a review of the problem and the site properties is presented. Then, the implemented flow and transport simulation models are explained briefly. Finally, the application results of the integrated sequential source identification and monitoring network design which are discussed. The primary goal of this paper is the integrated source characterization with an optimal number of spatiotemporal concentration measurement data obtained by utilizing Pareto-optimal monitoring network design in sequence.

2. Materials and Methods

The methodology presented here has two components: As the first step, the singularity indices are estimated utilizing FSMT [3] [10]. The FSMT is used to construct plausible contaminant plume boundary later used as an input for optimal contaminant monitoring network design. The second step consists of utilizing two objectives linked simulation-optimization model to be solved for optimal monitoring network design. This monitoring network design model also specifies the constraint on maximum number of monitoring locations to be selected at that stage, based on constraints on the maximum number of monitoring wells to be implemented. The Pareto-optimal solutions obtained from the two-objective model are used to design a set of Pareto-optimal monitoring networks. The objectives considered are: 1) Maximize the weighted sum of the product of estimated concentration gradients, and the simulated concentration at selected monitoring locations, and 2) minimize the maximum normalized errors between actual concentration and those estimated with the concentration interpolation models, based on monitoring data from designed monitoring locations. Adoptive Simulated Annealing (ASA) algorithm is used as the optimization algorithm to solve the two-objective monitoring networks design model.

The linked simulation-optimization model for optimal source identification is solved using the ASA algorithm. The ASA is used as the optimization algorithm for solving the optimization problem that minimizes the difference between the simulated and measured pollutant concentrations at the optimally chosen monitoring locations obtained as solution. The source-identification model is solved using concentration measurements from a chosen Pareto-optimal monitoring network.

2.1. The Singularity Mapping Technique Methodology

Generalized fractal self-similarity is often characterized by a power-law relationship in the spatial or frequency domain [29]. In the singularity mapping technique, the C-A (Concentration-Area) model is used. In this study, the FSMT in 2D map data is describe as a power-law relationship between area A in a sampled region, and the total amount of a certain physical quantity $\mu(A)$ as Equation (1) [5].

$$\langle \mu(A) \rangle = cA^{\alpha/2} \tag{1}$$

Here $\langle \rangle$ denotes the statistical expectation, α is the Holder exponent or singularity index, and c is a constant. The areal density value of $\mu(A)$ in the area A is defined by concentration $\rho(A)$ as Equation (2).

$$\langle \rho(A) \rangle = \frac{\mu(A)}{A} = cA^{\frac{\alpha}{2}-1} \tag{2}$$

Singularity is an index representation of the scaling dependency from a multifractal point of view, and it characterizes how statistical behaviors change as the scale of geochemical values changes. In the singularity mapping technique, the indices are estimated using the window-based procedure. The improved window-based procedure [29] is conducted as per the following steps.

This approach as described below has been developed in this study to detect the boundary of the contamination plume and determine the effective potential well locations relevant to source characterization. The contamination plume boundary is estimated utilizing the characteristics that almost horizontal gradient generally prevails near the plume boundary which corresponds to the inflection point of the anomaly.

2.2. Multi-Objective Optimization Algorithm for Monitoring Network Design

A multi-objective optimization model is formulated for the design of an optimal monitoring network with the conflicting objectives. The two-objective optimization model is solved by optimizing one of the objectives subject to the other objective defined as an implicit constraint. The number of monitoring wells to be selected is essentially governed by budgetary constraints. The two objectives of the multi-objective optimization model for optimal monitoring network design for accurate identification of unknown pollution sources are defined by Equations (3) and (5), respectively [31]. The multi-objective optimization model can be mathematically expressed as:

$$\begin{aligned} \text{F1} = \text{Maximize}\Bigg(\sum f_{i,j} C^*_{i,j} \Bigg\{ & \frac{\left|C^*_{i-1,j} - C^*_{i,j}\right| + \left|C^*_{i,j+1} - C^*_{i,j}\right|}{dx} \\ & + \frac{\left|C^*_{i,j-1} - C^*_{i,j}\right| + \left|C^*_{i,j+1} - C^*_{i,j}\right|}{dy} \Bigg\} \Bigg) \end{aligned} \tag{3}$$

$$\varepsilon_{\max} \geq C_{i,j}^{*} \geq \varepsilon_{\min} \quad \forall i, j \tag{4}$$

$$\text{F2} = \text{Minimize}\left(\text{Maximum}\left(\frac{C_{sim} - C_{int}}{C_{sim} + N}\right)\right) \tag{5}$$

where $C_{i,j}^{*}$ is simulated concentration in cell *i, j*. $f_{i,j}$ represents the binary decision variable to place or not to place a monitoring well at grid location *i, j*. $f_{i,j} \equiv$ such that when $f_{i,j}$ value equal to 1 representing monitoring well to be placed at grid *i, j*, and zero otherwise. $\varepsilon_{\max}$ and $\varepsilon_{\min}$ are the high and low value of concentration, respectively. C_{int} is interpolated (kriged) concentration in cell *i, j*. C_{sim} is simulated concentration in cell *i, j*.

The two-objective optimization model is solved using the constrained method [32]. In the constrained method, one of the objective functions (F1) is maximized, constraining the minimum level of satisfaction of the second objective function (F2) as shown in Equation (6)

$$\sum_{i,j} C_{i,j}^{*} f \left\{ \frac{\left|C_{i-1,j}^{*} - C_{i,j}^{*}\right| + \left|C_{i,j+1}^{*} - C_{i,j}^{*}\right|}{dx} + \frac{\left|C_{i,j-1}^{*} - C_{i,j}^{*}\right| + \left|C_{i,j+1}^{*} - C_{i,j}^{*}\right|}{dy} \right\} - \gamma \geq 0 \tag{6}$$

where γ is the specified minimum level of satisfaction of the second objective function F2, also termed as the trade-off constant. Therefore, the resulting model can be solved iteratively as a single objective optimization model for different satisfaction levels of γ, thus a Pareto-optimal solution set is generated. The second objective function can be specified as a new implicit constraint. The upper limit of γ is defined by the new constraint the maximum value of the second objective function F2 when solved as a single-objective optimization [Equation (7)]. The lower limit of γ is the value of the second objective function F2 corresponding to the maximum value of the first objective function F1, when the optimization model is solved as a single objective model with F1 as the only objective [Equation (8)].

$$\text{MaxF2} \geq \gamma \tag{7}$$

$$\text{F1}_{\text{MaxF}_2} \leq \gamma \tag{8}$$

where $\text{F1}_{\text{MaxF}_2}$ is the value of the objective function F2 corresponding to the maximum value of the first objective function F1 when solved as a single objective model. All solutions obtained on a Pareto-optimal front correspond to a different Pareto optimal monitoring network.

2.3. Developing Source Identification Model

The flow and transport models and the source characterization optimization model algorithm utilizing the Adoptive Simulated Annealing algorithm (ASA) is combined to develop the source characterization model. The simulation models utilize the unknown source concentration candidate solutions generated by the optimization algorithm to obtain the estimated contaminant concentration at monitoring locations. Then the optimization algorithm evaluates the objective function. The objective function value is a function of the differences between models estimated and measured concentration values. The optimal source characterization is obtained by solving the optimization model to minimize the objective function (9). Reference [33] defined the objective function of a simulation-optimization model for unknown concentration source characterization as follows:

$$\text{Minimize F1} = \sum\nolimits_{k=1}^{n_k} \sum\nolimits_{job=1}^{n_{ob}} \left(C_{estiob}^{\ \ k} - C_{obsiob}^{\ \ k} \right)^2 \cdot w_{iob}^{k} \tag{9}$$

$$C_{estiob}^{\ \ k} = f\left(x, y, z, C_{si}\right) \tag{10}$$

$$w_{iob}^{k} = \frac{1}{C_{obsiob}^{\ \ k} + n} \tag{11}$$

where n is an appropriate constant chosen to ensure errors at low concentration values do not dominate the solution, $C_{obsiob}^{\ \ k}$ is the concentration measured data at observation monitoring location iob and at the end of time period k ($M \cdot L^{-3}$), $C_{estiob}^{\ \ k}$ is the concentration estimated by the simulation models at observation monitoring location *iob* and at the end of time period k ($M \cdot L^{-3}$), n_t is the total number of monitoring time steps, n_{ob} is the total number of observation wells, n_k is total number of concentration observation time periods, n_{ob} is total number of observation wells, obs_{iob}^{k} is observed concentration at well *iob* and at the end of time period *k*, $f\left(x, y, z, C_{si}\right)$ is the simulated concentration obtained from the transport simulation model at an observation location and source concentrations C_{si}. w_{iob}^{k} is the weight corresponding to observation location *iob*, and the time period *k*.

The constraint set Equation (10) represents the linked simulation model for flow and transport process simulation. This simulation model can be a numerical model, or it can be replaced by an approximate simulator such as a trained and tested surrogate model.

The present study incorporates the Adaptive Simulated Annealing as the optimization algorithm for optimal source characterization model. This algorithm is chosen based on its comparative efficiency in reaching a global optimal solution.

2.4. Linked Simulation-Optimization Model for Optimal Contaminant Source Identification Using Pareto-Optimal Monitoring Networks and Arbitrary Monitoring Network

Source identification in terms of magnitude of an unknown pollution source is often solved using a linked simulation-optimization approach. The linked simulation-optimization model simulates the physical and chemical processes of flow and solute transport within the optimization algorithm. The flow and solute transport simulation models are treated as an important binding constraint for the optimization model. Therefore, any feasible solution of the optimization model needs to satisfy the flow and the transport simulation model. The advantage of this approach is that it is possible to link any complex numerical model to the optimization model. However, running the simulation model for several thousand times to obtain the optimal source characterization is time consuming and may affect computational feasibility and efficiency of the methodology. To address these issues, trained and tested Genetic Programming (GP) based surrogate models are utilized as approximate simulators of the physical processes in the linked optimization algorithm to obtain the reasonable and acceptable results with enormous saving of CPU time. In this study, linked simulation optimization based methodology for characterization of unknown pollution sources utilizes GP-based surrogate models as approximate simulators of the flow and transport processes in the aquifer study area contaminated by multiple reactive chemical species [4].

A linked simulation-optimization model [objective function Equation (12) and constraint set Equation (13)] is solved to estimate the pollution sources concentration to evaluate the performance of the proposed methodology. The two sets of 10 Pareto-optimal monitoring networks are designed from the potential well locations using FSMT (MNSI1 to MNSI10), and without utilizing FSMT (MN1 to MN10), respectively. Concentration measurements from each of the monitoring networks corresponding to each of the Pareto-optimal solution are used to estimate the pollution sources concentration. These evaluation results using concentration observations from the two sets of 10 Pareto-optimal monitoring networks are compared to find the efficiency of using FSMT based on the two objectives. For this evaluation purpose, the observed aquifer responses are simulated by solving HYDROGEOCHEM [34] [35], along with appropriate initial and boundary conditions. In order to evaluate the performance for erroneous concentration measurement data, numerically simulated concentration measurements are perturbed to represent the effect of random measurement errors. The observed pollutant concentration data is perturbed with random measurement error with maximum specified deviation of 10% as shown in Equation (12).

$$pert_{C^k_{est_{iob}}} = C^k_{est_{iob}}\left(i + err\right) \tag{12}$$

$$err = \mu_{per} \times rand \tag{13}$$

where $pert_{C^k_{est_{iob}}}$ is the perturbed numerically simulated concentration values.

$C^k_{est_{iob}}$ is the numerically simulated concentration value; err is the error term; μ_{per} is the maximum deviation expressed as a percentage; and $rand$ is a random fraction between −1 and +1 generated using Latin hypercube distribution.

To illustrate the efficiency of singularity index technique on the source identification model, the linked optimization algorithm is used with three different monitoring networks. These three monitoring networks design scenarios evaluated are: are FSMT-arbitrary networks, FSMT-designed optimal monitoring networks, and optimized monitoring networks without utilizing FSMT for potential well locations. Also, the contaminant source estimates obtained using the totally arbitrary (without any FSMT information) monitoring network is compared with those obtained using other monitoring networks.

The methodology for source identification utilizing sequential design of monitoring network includes two main steps which are as follows:

1) Implementation of a numerical flow and transport simulation model based on the estimated hydrological and geochemical properties of the contaminated aquifer study area. In this study, MODFLOW and MT3DMS are used.

2) The transport simulation models are linked to the Adaptive Simulated Annealing (ASA) optimization algorithm within a linked simulation-optimization model to obtain the optimal characterization of the unknown contaminant sources in terms of location, magnitude, and time of activity.

These two steps are iteratively followed in sequence so that a monitoring network can be designed with updated source estimates. Then the new monitoring network can be implemented to obtain fresh concentration data, which can be used to refine the source estimates. Therefore, the source characterization can improve in accuracy as additional data are sequentially obtained from monitoring networks designed with improved source information. The details and advantages of using the linked simulation-optimization model, the utilization of the FSMT for designing a monitoring network based on FSMT and the resulting efficiency are discussed in details in [3] [10]. **Figure 1** shows a schematic representation of the proposed methodology.

3. Background and the Study Area Description

The contaminated aquifer is located in Macquarie Groundwater Area in New South Wales (exact location not provided), Australia. Some reports mentioned the sign of BTEX in the subsurface water in this area, and the polluted region was investigated using the observation data from the monitoring groundwater wells. However, it did not indicate the first record of the BTEX pollution in this area. The highest BTEX concentration of 320 mg/l was reported in October 2009 in one monitoring well in this area. The affected area was implicitly estimated as over 1 Km^2. Between October 2006 and July 2011, the groundwater level and the contaminant concentration were measured and recorded with seventy-four monitoring wells. These wells were installed and utilized at various times during the monitoring period to realize the pollutant plume and pollutant transport process in the study area.

It was suspected and later determined that the contamination source was an underground leaking petrol tank(s) at a service station. The aim of the methodology evaluation for this illustrative contaminated aquifer site is to test the feasibility of accurately characterizing the contaminant source in terms of location and release history, utilizing a sequential and integrated source characterization and monitoring network design for concentration measurements. For evaluation purposes, the potential source locations included a dummy potential source location, so that it is also determined if the actual location is identified using this methodology. The performance evaluation was based on the assumption that the best estimate of the source characteristics are those obtained using the extensive

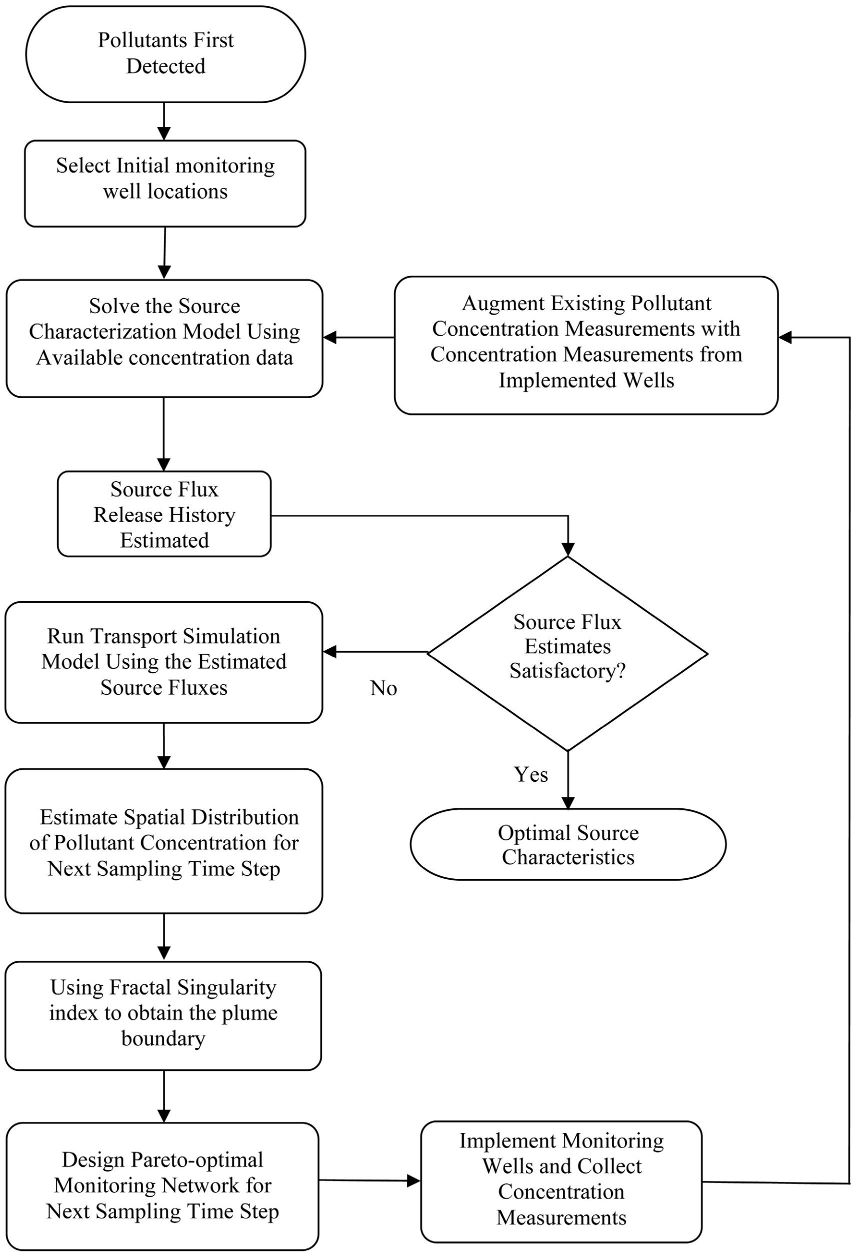

Figure 1. Schematic diagram of sequential source characterization and monitoring network design methodology.

concentration measurements available for this site. The deviation between these best source characterizations estimated and the source characteristics obtained using a much smaller number of monitoring locations based on the sequential optimal design procedure is considered as a measure of the accuracy of the results. This approach was necessary, as the source was detected many years after it become active, and initially the location was also unknown. Therefore, comparison of the source characterization solutions results, and actual source characteristics was not possible for a real-life site like this, and the indirect method of evaluation was adopted.

In this site, most of the existing wells were installed close to the potential source, and only a few were located to regions further away from the source. However, the time and location of measured observation data were selected arbitrarily. It seems based on the previous investigations, the leaking underground storage tank at a gas station was the potential source of the pollution. However, source characteristics regarding starting time of leakage, the magnitude of flux and the time history of flux releasing out of the source were not specified in the investigation reports.

The aim of this study is to characterize the unknown contaminant source(s) characteristics in terms of source location, starting time, and flux release history of any potential source in the polluted aquifer. The contaminant measured data are available at almost all the seventy-four monitoring wells every three months during the investigation period (October 2006 to July 2011). However, the aim of the source characterization study is using the optimal number of spatiotemporal observed measurement data. There are some limitations in collected observation data including: concentration data are not available for all wells and all times, some wells installed later during the investigation period, some wells installed far from the sources and these data are less relevant. The polluted aquifer site is a part of the Upper Macquarie Groundwater Management Area. 292 m AHD is the starting ground elevation at East, decreasing until elevation reaches the 251 m AHD at West. The aquifer is mainly recharged from the rainfall and the river. The majority recharge comes from the Macquarie River, and minority comes from precipitation (average 583 mm/year) in the wet season, from November to February.

The aquifer water in the study area is mainly extracted for potable water and irrigation usage through pumping wells. There have been substantial variations in the range of pumping rates due to changes in groundwater usage policy from the available aquifer, and also a voluntary limitation of extraction pumping rate from 2010 [36]. Another source of losing the water from the aquifer is through evapotranspiration, which peaks to 260 mm/month during the dry season [37].

3.1. Groundwater Flow and Contaminant Transport Simulation Models

The numerical groundwater flow and transport simulation models utilized for flow and transport processes are MODFLOW [38] and MT3DMS [39], respectively. The aquifer parameters calibrated earlier for the flow process and the flow model developed by [40] are utilized in this study. The utilized aquifer properties and other details in the flow and transport simulation models are described briefly in the next two sections. More details about the simulation models used are described by [40].

3.2. Groundwater Flow Modelling of the Contaminated Aquifer

There were no easily detectable natural flow boundary conditions around the contaminated study area. Therefore, a bigger site area with natural boundary conditions was used for flow simulation and area was designated as the extended study area. This study area dimensions are 2.187 Km by 2.426 Km, while the dimensions of the modelled smaller contaminated area are 608 m by 864 m, which is called the specified area. The groundwater flow in the aquifer study area is modeled as an unconfined aquifer. The natural boundary condition of the study area at West is the specified constant head condition based on average stage at the Macquarie River (**Figure 2**). Similarly, the ground topography is laid down toward the river in the West. Other boundaries are specified as constant head boundary conditions. The hydrogeologic properties and the boundary conditions used in [40] for the flow model are the same as those used in modeling groundwater flow in the entire Upper Macquarie Groundwater Management Area, developed by [37].

The groundwater flow process is modeled for 18 years with 18 one-year time steps from 1 January 1995 until 31 December 2012. In the three-dimensional simulation models, the study area is discretized into small grids of size 21.87 m by 21.08 m in the x and y directions respectively. The soil layers thickness are variable at different places; therefore, the grid size in the z direction is different and matches with the layer thickness. The hydrogeological properties, such as hydraulic conductivity, porosity, specific storage and specific yield, were obtained from previous studies conducted in this study area by [40]. These hydrogeological properties are listed in **Table 1**.

Table 1. Hydrogeological Properties used in Flow Modelling of the extended Study Area [10].

Parameter	Unit	Value
Maximum length of study area	m	2187.1
Maximum width of study area	m	2425.6
Saturated thickness, b	m	Variable
Number of layers in z-direction		3
Grid spacing in x-direction, Δx	m	21.87
Grid spacing in y-direction, Δy	m	21.08
Grid spacing in z-direction, Δz	m	Variable
Kxx (Layer 1, Layer 2, Layer 3)	m/d	12.37, 16.24, 0.001
Kyy (All Layers)	m/d	0.2
θ (All Layers)	dimensionless	0.27
Longitudinal Dispersivity, aL	m/d	12
Transverse Dispersivity, aT	m/d	6
Horizontal Anisotropy	dimensionless	1.5
Specific Yield Sy (All Layers)	dimensionless	0.1
Specific Storage Ss (All Layers)	dimensionless	0.000006
Initial pollutant concentration	g/l	0.00

Figure 2. Plan views of the study area and the impacted area.

The calibration of the flow model was achieved using observed measurement data from all well locations. The five years calibration period started from October 2006 to July 2011. One metre tolerance from the observed head data with 90 percent confidence was the calibration targets. The satisfaction level of calibration process was obtained with boundary conditions adjustment.

Once the extended study area is modeled and calibrated, the flow model parameters used for the specified area are derived from the calibrated extended model. The GMS7.0 feature, Regional to Local, is used to interpolate the starting head and layer thickness values for the specified area from the extended study area model. Figure 3 shows the specified area where the pollutant is estimated to be present. The grid sizes are refined further in the flow model for the specified area, and the area is discretized into 75 rows, 50 columns, and three layers. All of the boundaries are considered as time-varying specified head boundary conditions. The value of the time varying specified heads at the edge of the specified area are extrapolated from the calibrated model for the extended study area. All of the other hydrogeological flow parameters are kept the same as in Table 1.

3.3. Pollutant Transport Simulation in the Impacted Area

MT3DMS code (USGS) was utilized as a three-dimensional numerical transient transport simulation model for the study area model the fate and transport of the petrochemical pollutant BTEX originating from potential point sources. For the purpose of implementation, the pollutant is assumed to be conservative in nature, and the pollutant plume boundary is assumed to be contained within the boundary of the impacted area. The transport simulation model predicts the movement of the contaminants using the results of the flow simulation model in the impacted area of the aquifer, over time. The initial concentration of BTEX is assumed to be zero at the first-time step in the study area. All the other relevant transport parameters used in the transport model are shown in Table 1.

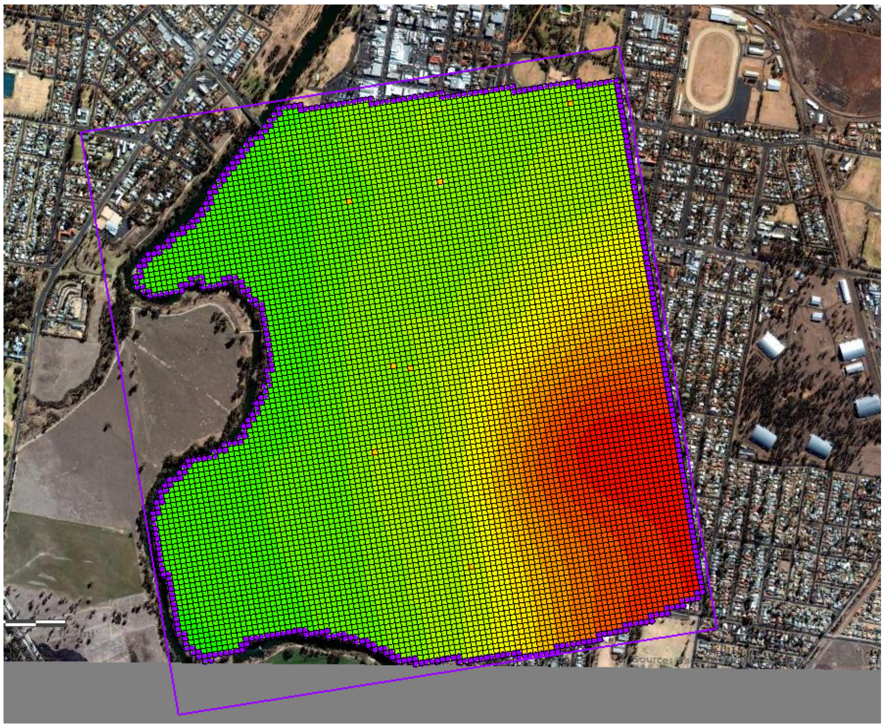

Figure 3. Plan view of the extended study area [40].

3.4. Sequential Integrated Model

The proposed methodology integrates the source identification model [4] and the multi-objectives monitoring network design model [3] by sequentially solving for the source characterization followed by monitoring network design. This sequence may be followed for several iterations. The monitoring network design is to be implemented for new concentration measurement to be utilized for another sequence of source characterization. This methodology evaluation is performed for the specified study area for contamination transport modelling. Although no actual measurements are conducted at the designed optimal monitoring locations, the available set of extensive concentration measurement data at already existing monitoring locations are used as the actual measurements. This is possible only because the extensive concentration measurement data is collected over several time periods, and also, the potential monitoring locations specified in the monitoring network design model coincides with some of the existing monitoring locations.

The unknown groundwater source characterization model uses the observed measurement data at arbitrary well locations to estimate the initial source characteristics. Next, the optimal monitoring network design utilizes the source identification results from the previous step and determines the new monitoring well locations. The linked optimization source characterization methodology is solved to identify the source characteristics using the observation data from the new monitoring wells in addition to the measured data from the existing monitoring wells. The developed integrated methodology needs to be used iteratively until reasonable accuracy of source characterization is obtained. The results of the source characterization model are utilized for obtaining new monitoring well locations using the Pareto-optimal monitoring network design for each sequence.

The following steps outline the present the new optimal monitoring network design process. The first step involves the estimation of spatial concentration values utilizing available concentration data and spatial extrapolation methods. In the illustrative application of the methodology, the initial spatial concentration measurements were simulated using a contaminant transport simulation model. The transport simulation model provides the estimates of contaminant concentration throughout the aquifer using the initial source characterization results. Second, the Fractal Singularity Mapping Technique (FSMT) provides the plume boundary as a guideline for selecting the potential monitoring wells. Third, the Pareto-optimal two-objectives monitoring network design is implemented for collecting new concentration measurement data. In the next sampling time step, pollutant concentration measurements from these newly implemented monitoring wells and data at already existing monitoring wells are obtained. Subsequently, the pollutant concentration measurements from the current sampling time step and previous sampling time steps are utilized for source identification. **Figure 1** shows the schematic diagram of the proposed methodology.

3.5. Performance Evaluation of the Applied Methodologies

The groundwater unknown source characterization uses the calibrated contaminated transport simulation model. The unknown source identification methodology developed by [3] [5] [10] is utilized for recreating the source flux release history and the source activity initiation time. To increase the accuracy and efficiency of source identification, the Pareto-optimal monitoring network design methodology to identify the source characteristics is developed using fractal singularity mapping technique (FSMT) [3]. The integrated monitoring network design and source identification sequentially, increases the efficiency of source characterization in real life scenarios with sparse spatiotemporal measurement data. The aim of this study is to identify the unknown source characteristics in terms of flux magnitude, contaminant flux release history, and activity starting time. The location of the contaminant source is explicitly known for evaluating the methodology, but unknown to the source characterization model. For evaluating the efficiency of developed source characterization methodology, two possible source locations are assumed as unknown sources. One of these should be identified as a non-active or non-existent source. The two potential sources in the study area are shown in **Figure 4**. The points marked in red circles are the grid locations containing the possible sources, and the yellow points are the observation wells where the concentration of BTEX is observed. A total of seventy-four concentration measurement locations are present in the study area.

Figure 4. Plan view of the specified area. (Red circle: Potential sources; Yellow Square: Monitoring wells)

3.6. Simultaneous Source Flux Release History and Source Activity Initiation Time Identification

ASA optimization linked source characterization methodology is used to identify the source flux magnitude. The potential contaminant sources are named source 1 and source 2 which are located at cell (1, 17, 29), and (1, 16, 24), respectively. The simulation model starts from 1 January 1995. However, the unknown starting time of the source activity can be anywhere between 1 January 1995 and 31 December 2011. Ten equal stress periods (1 year each) cover the ten years activity duration of the sources. The pollutant flux from each of the sources is represented as $S_{i,j}$ ($i = 1, 2$, $j = 1, 2, \ldots, 10$) where i represents the source number, and j accounts for the stress period number. In this case, S1 is the actual source and S2 is the unreal source. Other assumptions of source identification include the contaminant source flux is constant over each stress period, and the contaminant releases of both sources S1 and S2 starts at the same time. To realize the starting time of the sources an additional time lag variable ΔT is introduced in the optimization linked program [40]. In source characterization model, observed measurement data are used from 22 January 2009 and continued every three months.

3.7. Sequential Optimal Monitoring Network for Efficient Source Characterization

To increase the efficiency of source identification, the Pareto-optimal monitoring network design is integrated with the linked simulation-optimization source identification model. The source identification methodology regarding flux magnitude and source activity starting time is started using the observed measurement data from three arbitrary wells. These wells are randomly selected as a subset of all available monitoring wells within the study area. At each sequence of model running, three new monitoring wells are chosen for the next monitoring time step. The available installed observation wells in the specified area are considered as the potential monitoring wells. The preliminary source characterization model is solved using concentration data measured on 22 January 2009 from the first three arbitrary wells.

The source characterization solution results are utilized in the transport simulation model as inputs to predict the BTEX concentration at the next sampling time step (30 April 2009) in the specified area. The plume concentration in the specified area on 30 April 2009 is utilized to obtain the next monitoring wells locations. The two-steps optimal monitoring network design procedure (Esfahani and Datta 2018) is applied to locate the next three monitoring well locations.

Firstly, FSMT is used to compute the singularity indices based on the initially estimated source fluxes at sources and corresponding aquifer response at the specified area. FSMT is applied to estimate the likely contamination plume boundary which is then used as one of the guidelines for locating potential monitoring well locations. All installed monitoring wells close to the plume boundary are suitable potential monitoring well locations for the monitoring network. These wells are selected out of all available monitoring wells for the next step. In the second phase, a multi-objective optimization methodology is used for Pareto-optimal monitoring network design, with constraints on the total number of new monitoring wells to be installed. In this performance evaluation scenario, the total number of new monitoring wells to be optimally chosen was restricted to three locations S out of the potential monitoring well locations specified. The Pareto-optimal solutions obtained from the two-objective model are used to design a set of Pareto-optimal monitoring networks. The monitoring network is chosen based on the following two objectives: 1) maximize the summation of the product of estimated concentration gradients, and the simulated intensity at that location, and 2) minimize the maximum normalized error between actual concentration and those estimated with the kriging interpolation models, based on monitoring data from designed monitoring locations (Esfahani and Datta 2018).

Once a new monitoring network is designed and implemented, the concentration measurements from all wells in the monitoring network obtained on 30 April 2009, together with measurements obtained from the pre-existing three arbitrarily chosen wells on 22 January 2009, are utilized in the source identification model. This approach of sequential source characterization and monitoring network design and implementation is repeated for the subsequent sampling time steps, until changes in the source flux and starting time estimates are negligible.

4. Results and Discussion

Well M03, M14 and M20 are selected as the three arbitrary initial monitoring wells, located at (1, 21, 25), (1, 24, 25), and (1, 20, 29), respectively. After every sequence of monitoring wells implementation, the source characterization results are presented. To reflect the real-life conditions, measurement errors are incorporated with the concentration measurement data. The concentration measurement data are perturbed using random measurement error with a maximum specified deviation of 10 percent of the actually measured value.

4.1. First Sequence

In the beginning, as the initial step, the linked simulation-optimization model is solved to obtain the source characteristics regarding contaminant release fluxes and the source activity starting times. The objective function of the linked optimization methodology for source characterization is formulated using the observed measurement data for January 2009 at the three arbitrary well locations. Adaptive Simulated Annealing optimization algorithm is used to link the MT3DMS transport simulation model to obtain the best solution results based on the provided information. Figure 5 shows the source identification results in this step, using three arbitrary monitoring wells. The x-axis is marked by the source flux variables for two sources at different time steps (Sij i = 1, 2, j = 1, 2, …, 10), and also time lag variable ΔT. The source flux magnitude (mg/s) is shown on the primary y axis, and the secondary y-axis shows the lag time (day).

The lag time estimated by this methodology indicates that source activity started in the year 2001. The first sequence of source characterization indicates that both sources were active over the time steps. These preliminary solution results based on arbitrary observation measured data does not appear to be reasonable. Therefore, to determine more accurate source characteristics, the next sequence of the optimal monitoring network is essential. To obtain the next three optimal monitoring well locations, the two-steps FSMT multi-objectives monitoring network design model is solved for the next time step, with concentration measurement data for dated 30 April 2009.

4.2. Second Sequence

The first sequence results of source characterization indicate that the monitoring

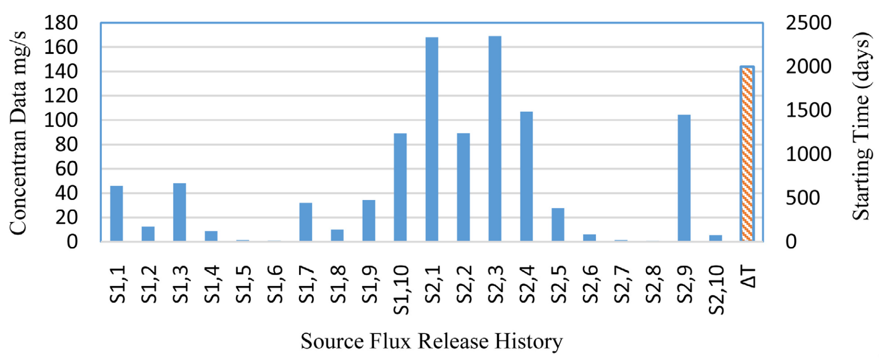

Figure 5. Source identification result using initial observed concentration measurements.

network for the next monitoring time step (April 2009) should be designed. To increase the feasibility and efficiency of monitoring network design, FSMT is used. Singularity Index contours indicate the plume boundary. The source characterization is more efficient using the wells close to the plume boundary. The wells which are close to the singularity index contours are used as candidates in the two objectives monitoring network design model, and the wells which are far from the plume boundary is eliminated. **Figure 6** and **Figure 7** show the singularity analysis results and singularity index contours for a value of 2, respectively.

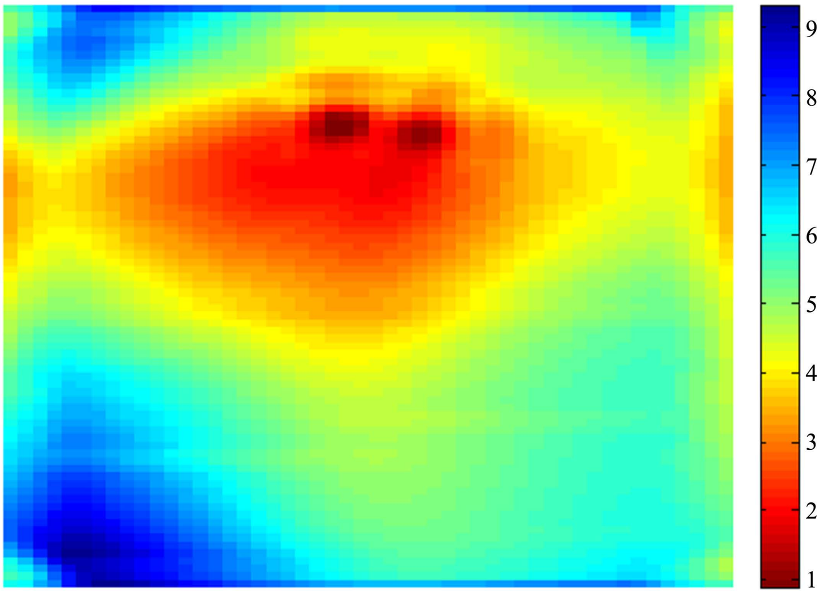

Figure 6. Singularity analysis result using the first source characterization results.

Figure 7. Singularity index using the results of first source characterization.

The next stage monitoring well locations are chosen using the Pareto-optimal monitoring network design based on the following two objectives: 1) maximize the summation of the product of estimated concentration gradients, and the simulated concentration at that location, and 2) minimize the maximum normalized error between actual concentration and those estimated with the kriging interpolation models, based on monitoring data from designed monitoring locations. In the monitoring network process, three well locations out of 18 potential monitoring wells are selected as observation wells for April 2009. The selected wells from the Pareto-optimal monitoring network design model are M17, M19 and M16 located at (1, 21, 23), (1, 18, 24), and (1, 17, 26), respectively.

The observed measurement data from all six monitoring wells both the first sequence and the second sequence (M03, M14, M20, M17, M19, and M16) are recorded for April 2009. These concentration measurement data, in addition to the concentration data collected in January 2009 (already available from the previous sequence) are utilized in the linked simulation-optimization model to improve the solution for optimal source characterisation in the study area. **Figure 8** illustrates the source flux release history results and the lag time estimation from the source characterization results using concentration measurements from all wells in the monitoring network, obtained on 30 April 2009, and measurements obtained on 22 January 2009 from the three pre-existing arbitrary wells.

It can be noted that there is an improvement in the estimate of the source flux magnitude for source two (S2). However, the source characterization model could not identify the actual and the dummy (not actually a source) sources completely. The same integrated monitoring network design and source characterization model solution sequence is repeated until the reasonable accurate results are obtained.

4.3. Third Sequence

In the third sequence, again new monitoring wells need to be selected for the next monitoring time (July 2009) to further improve the source characterization. Similar to the previous sequence, plume boundary contours are obtained using the FSMT. **Figure 9** and **Figure 10** show the results of FSMT methodology

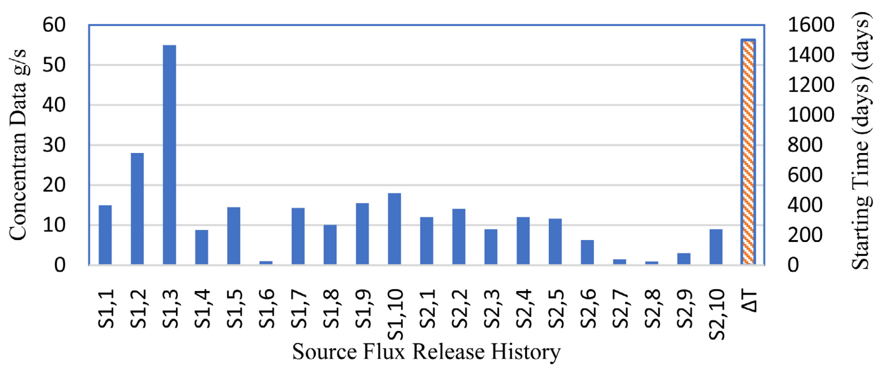

Figure 8. Source identification results from the second sequence.

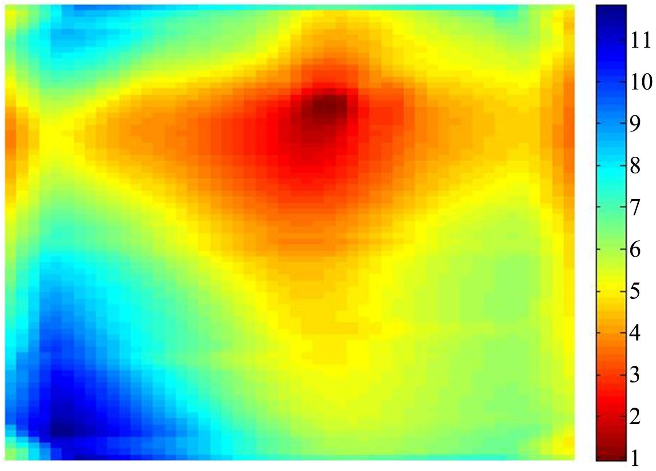

Figure 9. Singularity analysis result using the second source characterization results.

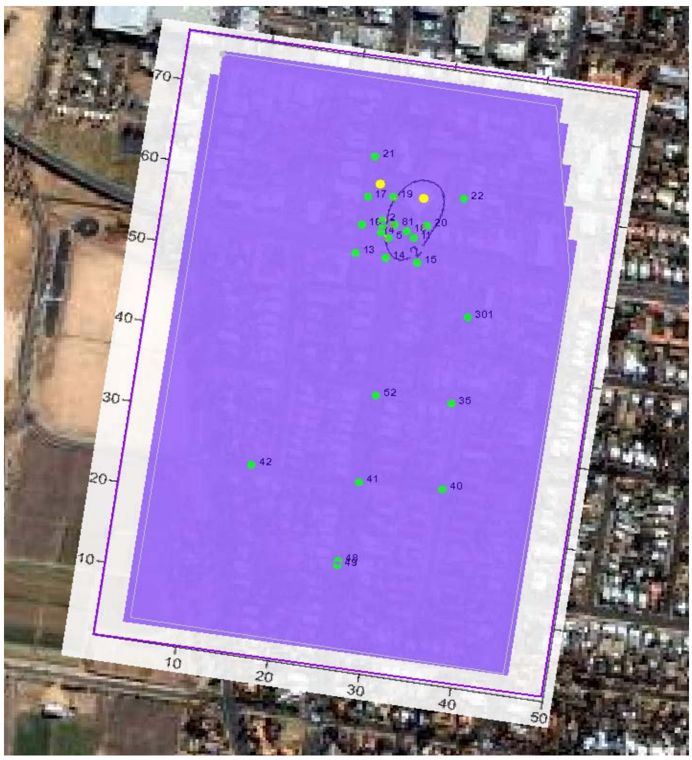

Figure 10. Singularity index using the results of second source characterization.

solution in the third sequence. The new potential wells are identified, and those wells which are far from the plume boundary are eliminated. Then the two objectives Pareto-optimal monitoring network design is utilized to obtain the location of next three monitoring wells on July 2009. Well locations M02 (1, 20, 25), M05 (1, 22, 25) and M18 (1, 21, 27) are selected as the new monitoring wells for next monitoring time step.

The concentration data at existing monitoring wells (M03, M14, M20, M17, M19 and M16) and the newly selected monitoring wells (M02, M05, and M18) are recorded for July 2009. The source identification model is solved using the concentration observation data in July 2009 in addition to the existing available concentration data (in January and April 2009). Figure 11 shows the source flux estimation, and the lag time estimation for two sources for the third sequence of source characterization.

The starting time estimates do not seem to change from the previous design sequence to this design sequence. It is also evident that the dummy source is identified correctly in this sequence; therefore, the methodology is terminated. Subsequent to this all the available concentration measurement data from all the 74 well locations are utilized for accurate source characterization. These solution results are utilized as the benchmark for evaluating or validating the source characterization solution results obtained with a very small number of designed monitoring locations. These solution results of source flux characterization and the selected optimal monitoring locations are presented in Table 2.

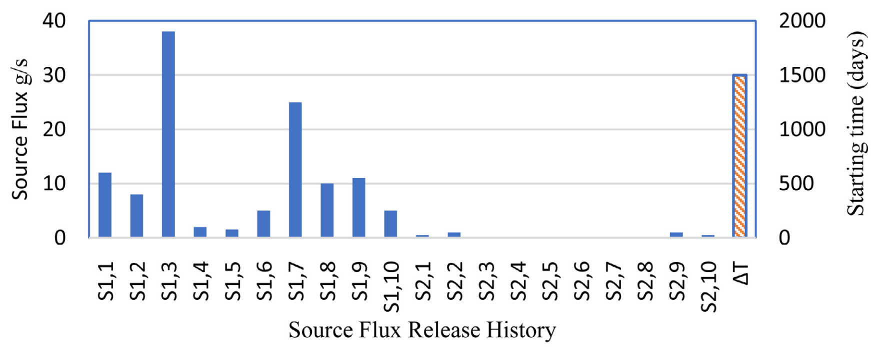

Figure 11. Third source identification results.

Table 2. Results of source characterization and designed monitoring well locations.

	Source & Source Location (i, j, k)	Flux Values (g/s) at different time steps									
		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Evaluation Model	Source 1 (Actual) (17, 29, 1)	14.10	11.05	35.08	1.15	2.61	7.90	27.40	7.32	8.14	15.95
	Source 2 (dummy) (16, 24, 1)	0	0	0	0	0	0	0	0	0	0
Third Sequence	Source 1 (Actual) (17, 29, 1)	12	8	38	2	1.5	6	27	6	9	5
	Source 2 (dummy) (16, 24, 1)	0.2	0.1	0	0	0	0	0	0	0.1	0.1
Second Sequence	Source 1 (Actual) (17, 29, 1)	15	28	55	8.8	14.5	1	14	10	15.5	18
	Source 2 (dummy) (16, 24, 1)	12	14	9	12	11.6	6	1.4	0.9	3	9
First Sequence	Source 1 (Actual) (17, 29, 1)	46	12..6	48	8.7	1.4	1	32	10	34	89
	Source 2 (dummy) (16, 24, 1)	168	201.2	169	107	27.6	6.3	1.4	0.9	164	5.4

4.4. Evaluation

In order to evaluate the efficiency and accuracy of the proposed sequential methodology in estimating the source flux magnitude, release history and source activity starting time, accurate values of the actual fluxes, release history in terms of time including the activity initiation time are necessary. Performance evaluation of the proposed methodology requires these benchmark values for comparison. As in almost all such real-life scenarios, actual source flux magnitudes, possibly location, and the release history or the source(s) activity initiation time are not known. Therefore, for evaluation purpose, the performance evaluation is based on the comparison of the estimates obtained using the proposed sequential methodology with the more accurate values obtained by utilizing the extensive concentration measurement network for this polluted urban aquifer site, consisting of 74 monitoring wells covering part of a small urban city in New South Wales, Australia. In fact, the motivation behind choosing this real-life site was the availability of extensive concentration and head measurement data, as well as the available information regarding the hydrogeologic parameters. The bench-mark values of the source characteristics in terms of location, magnitude, and release history were established using the source characterization based on this extensive monitoring information for this site. Subsequently, these benchmark values were compared with the solution results obtained using the proposed methodology.

Therefore, the linked simulation-optimization model for optimal source characterization, without any monitoring network design component is solved using all available contaminant concentration measurement data for all three observed time steps (January, April, and July 2009). Figure 12 presents the flux release history for the evaluation model. Table 2 shows the optimal source flux estimation using all available measured concentration data.

Figure 12 shows the source characterization results in the evaluation model using the extensive concentration measurement network for this polluted urban aquifer site, consisting of seventy-four monitoring wells in the study area. The estimated source flux magnitude for S10 shows a steep jump in source flux value. As the lag time ΔT estimate indicates that the source activity started in the year

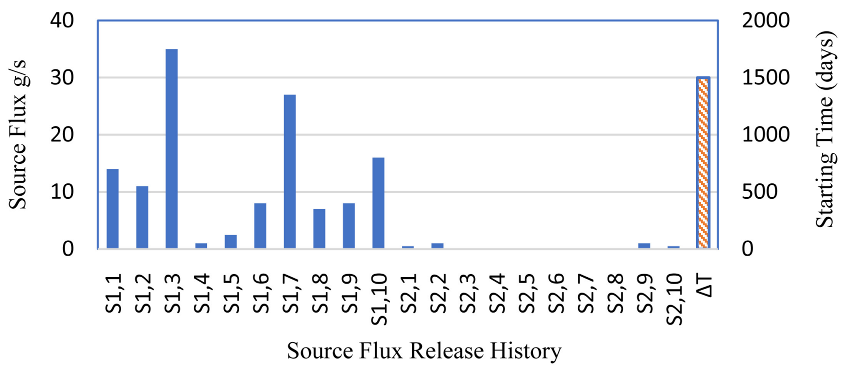

Figure 12. Flux release history and lag time in the evaluation source identification.

1999, source flux S10 represents the source flux magnitude for 2009. All the concentration measurements used in the identification of source characteristics are from the beginning of the year 2009 (22 January 2009). The observation monitoring wells always obtain the source activity with time delay. Therefore, it seems that the source flux magnitude S10 may not have impacted the concentration measurements taken in January, April, and July 2009. Figure 13 shows the concentration breakthrough curves at the selected monitoring locations in the specified study area. These breakthrough curves show the relevance of the selected monitoring locations, as these are impacted in a time varying manner by the source flux. The locations M17 and M19 appear to be minimally affected by the source and may not have proved very effective in the source characterization process. However, these wells are selected using the Pareto-optimal monitoring network design for identifying the source characteristics and therefore, it is possible that some of these locations do not show large concentrations. However as discussed in Prakash and Datta (2015), a few monitoring locations with very small concentrations may also help in distinguishing multiple overlapping plumes.

To evaluate the efficiency of the proposed sequential methodology of monitoring network design and subsequent source characterization, the normalized absolute errors between the estimated temporal release history of source one at different sequences, and corresponding benchmark source fluxes (obtained using extensive information and 74 monitoring location data) are calculated. Figure 14 shows the normalized absolute errors for various sequences of the integrated source characterization and monitoring network design methodology.

These limited performance evaluation results for a real-life contaminated aquifer site results show that the developed methodology of sequential source characterization and FSMT based multi-objectives monitoring network design can

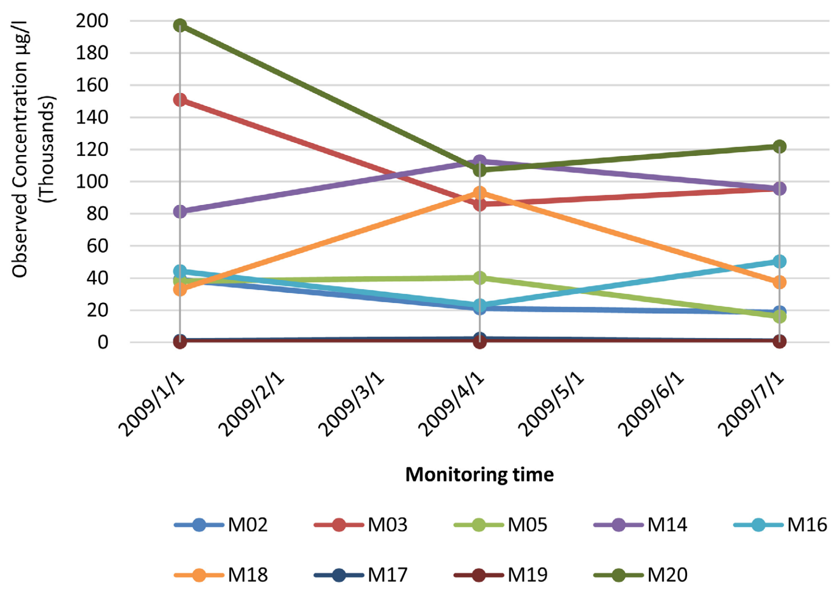

Figure 13. Breakthrough curves at monitoring well locations.

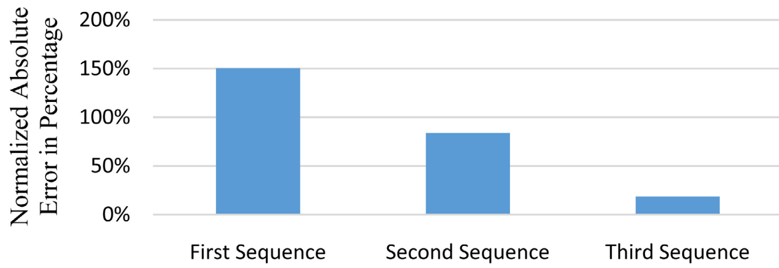

Figure 14. Normalized absolute errors for potential source location one.

successfully identify the source characteristics of the unknown contaminant source, and also correctly identify the dummy (not an actual) source in terms of locations, flux release history, and source activity starting time. These reasonably accurate solution results are obtained using a limited number of concentration measurement data, at only a few selected monitoring locations (18 concentration measurement data) in the third sequence of the iterative procedure. Also, these solution results are comparable to those obtained using a much more expensive and comprehensive concentration monitoring network. These results show the potential application of the methodology to design an economically efficient and effective monitoring network utilizing FSMT for initial determination of the potential monitoring locations.

5. Conclusions

In this study, the performance of an integrated sequential source characterization methodology based on the solution of an optimal linked simulation-optimization based source characterization model and a sequential Pareto-optimal monitoring network design methodology is evaluated for a real-life contaminated aquifer in an urban area. The reasonably accurate solution results for source characterization in terms of flux release history, source activity starting time and accurate source locations demonstrate the potential applicability of the proposed methodology to real contaminated aquifer sites. The available hydraulic head data from the observation wells are utilized to obtain the calibrated flow model, and the transport simulation models for the study area. Three sequences of source identification and monitoring network design are applied to the study area to obtain the final solution results for the unknown source characteristics.

Initially, three arbitrary monitoring wells are utilized at the first round of source characterization. Then the estimated plume concentration data are used to choose the next three monitoring well locations using FSMT based two objectives monitoring network design model. The well locations which are close to the plume boundary, are suitable candidates for monitoring network design for source identification model. A Singularity Index guideline improves the optimal design of the monitoring network by effectively decreasing the number of potential monitoring well locations. The concentration measurement data from the new wells in the designed monitoring network, in addition to the previous monitoring well locations are utilized to identify the source characteristics. In this application, the sequential process is repeated three times to obtain the optimal accuracy in estimating release fluxes and source flux starting time.

The proposed methodology shows efficiency in identifying the unknown source characteristics as only nine monitoring wells are utilized in the final sequence. The designed monitoring network uses less number of well locations as compared to the source identification model using seventy-four available observation wells, with comparable results. Only three temporal readings (January, April, and July 2009) were utilized to estimate the source characteristics satisfactorily. Therefore, the proposed methodology is potentially useful for efficient characterization of unknown contaminant sources in a complex contaminated aquifer site, where very little initial concentration measurement data are available. The proposed sequential procedure helps in designing relevant and efficient monitoring networks, which when implemented, provides fresh concentration measurement data. The illustrative application of the methodology to a real-life contaminated aquifer site demonstrates the capability and efficiency of the proposed methodology.

Funding

Graduate Research School, James Cook University, and College of Science and Engineering provided financial support including a Graduate Research School Scholarship (JCUPRS) for this research. Stantec supported financially with the publication charge for this paper.

Authors' Contributions

Conceptualization: Bithin Datta and Hamed Esfahani; methodology, Bithin Datta and Hamed Esfahani; validation, Hamed Esfahani; formal analysis, Hamed Esfahani and Bithin Datta, providing the raw measurement data from the contaminated aquifer site: Adrian Heggie.

This study forms a part of the Ph.D. Thesis related work of Hamed Esfahani. The thesis was supervised and guided by Bithin Datta. This Ph.D. Thesis was submitted to James Cook University, Australia in 2016.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] Chandalavada, S., Datta, B. and Naidu, R. (2012) Optimal Identification of Groundwater Pollution Sources Using Feedback Monitoring Information: A Case Study. *Environmental Forensics*, **13**, 140-154. https://doi.org/10.1080/15275922.2012.676147

[2] Datta, B. and Kourakos, G. (2015) Preface: Optimization for Groundwater Characterization and Management. *Hydrogeology Journal*, **23**, 1043-1049.

https://doi.org/10.1007/s10040-015-1297-3

[3] Esfahani, H.K. and Datta, B. (2018) Fractal Singularity-Based Multiobjective Monitoring Networks for Reactive Species Contaminant Source Characterization. *Journal of Water Resources Planning and Management*, **144**, 1-12. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000880

[4] Esfahani, K.H. and Datta, B. (2016) Linked Optimal Reactive Contaminant Source Characterization in Contaminated Mine Sites: A Case Study. *Journal of Water Resource Planning and Management ASCE*, **142**, 1-14. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000707

[5] Datta, B., *et al.* (2016) Preliminary Hydrogeologic Modeling and Optimal Monitoring Network Design for a Contaminated Abandoned Mine Site Area: Application of a Developed Monitoring Net-work Design Software. *Journal of Water Resource and Protection*, **8**, 46-64. https://doi.org/10.4236/jwarp.2016.81005

[6] Esfahani, K.H. and Datta, B. (2015) Use of Genetic Programming Based Surrogate Models to Simulate Complex Geochemical Transport Processes in Contaminated Mine Sites. In: Gandomi, A.H., Alavi, A.H. and Ryan, C., Eds., *Handbook of Genetic Programming Application*, Springer International Publishing, Berlin, 359-379. https://doi.org/10.1007/978-3-319-20883-1_14

[7] Esfahani, K.H. and Datta, B. (2015) Simulation of Reactive Geochemical Transport Processes in Contaminated Aquifers Using Surrogate Models. *International Journal of GEOMATE*, **8**, 1190-1196. https://doi.org/10.21660/2015.15.4263

[8] Neupane, R. and Datta, B. (2021) Optimal Characterization of Unknown Multispecies Reactive Contamination Sources in an Aquifer. *Journal of Hydrologic Engineering*, **26**, Article ID: 04021035. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002134

[9] Hayford, M.S. and Datta, B. (2021) Source Characterization of Multiple Reactive Species at an Abandoned Mine Site Using a Groundwater Numerical Simulation Model and Optimization Models. *International Journal of Environmental Research and Public Health*, **18**, 4776. https://doi.org/10.3390/ijerph18094776

[10] Esfahani, H.K. (2016) Development of Integrated Methodologies for Optimal Characterization of Reactive Contaminant Sources and Monitoring Network Design in Polluted Aquifer Sites. PhD Thesis, James Cook University, Townsville.

[11] Fethi, B.J., Loaiciga, A.H. and Marino, A.M. (1994) Multivariate Geostatistical Design of Groundwater Monitoring Networks. *Journal of Water Resource Planning and Management ASCE*, **120**, 505-522. https://doi.org/10.1061/(ASCE)0733-9496(1994)120:4(505)

[12] Loaiciga, H.A., *et al.* (1992) Review of Ground-Water Quality Monitoring Network Design. *Journal of Hydraulic Engineering*, **118**, 11-38. https://doi.org/10.1061/(ASCE)0733-9429(1992)118:1(11)

[13] Reed, P.M. and Minsker, B.S. (2004) Striking the Balance: Long-Term Groundwater Monitoring Design for Conflicting Objectives. *Journal of Water Resources Planning and Management*, **130**, 140-149. https://doi.org/10.1061/(ASCE)0733-9496(2004)130:2(140)

[14] Datta, B., Chakrabarty, D. and Dhar, A. (2009) Optimal Dynamic Monitoring Network Design and Identification of Unknown Groundwater Pollution Sources. *Water Resources Management*, **23**, 2031-2049. https://doi.org/10.1007/s11269-008-9368-z

[15] Dhar, A. and Datta, B. (2007) Multi-Objective Design of Dynamic Monitoring Networks for Detection of Groundwater Pollution. *Journal of Water Resource Planning and Management*, **133**, 329-338. https://doi.org/10.1061/(ASCE)0733-9496(2007)133:4(329)

[16] Kollat, J.B., Reed, P.M. and Maxwell, R.M. (2011) Many-Objective Groundwater Monitoring Network Design Using Bias-Aware Ensemble Kalman Filtering, Evolutionary Optimization, and Visual Analytics. *Water Resources Research*, **47**, W02529. https://doi.org/10.1029/2010WR009194

[17] Reed, P.M. and Kollat, J.B. (2013) Visual Analytics Clarify the Scalability and Effectiveness of Massively Parallel Many-Objective Optimization: A Groundwater Monitoring Design Example. *Advances in Water Resources*, **56**, 1-13. https://doi.org/10.1016/j.advwatres.2013.01.011

[18] Luoa, Q., Wub, J., Yangb, Y., Qian, J. and Wu, J. (2016) Multi-Objective Optimization of Long-Term Groundwater Monitoring Network Design Using a Probabilistic Pareto Genetic Algorithm under Uncertainty. *Journal of Hydrology*, **534**, 352-363. https://doi.org/10.1016/j.jhydrol.2016.01.009

[19] Mandelbrot, B. (1972) Possible Refinement of the Lognormal Hypothesis Concerning the Distribution of Energy Dissipation in Intermittent Turbulence. In: Rosenblatt, M. and Atta, C., Eds., *Statistical Models and Turbulence*, Springer, Berlin, 333-351. https://doi.org/10.1007/3-540-05716-1_20

[20] Mandelbrot, B. (1967) How Long Is the Coast of Britain. *Science*, **156**, 636-638. https://doi.org/10.1126/science.156.3775.636

[21] Cheng, Q. (2008) A Combined Power-Law and Exponential Model for Streamflow Recessions. *Journal of Hydrology*, **352**, 157-167. https://doi.org/10.1016/j.jhydrol.2008.01.017

[22] Sadeghi, B., Madani, N. and Carranza, E.J.M. (2015) Combination of Geostatistical Simulation and Fractal Modeling for Mineral Resource Classification. *Journal of Geochemical Exploration*, **149**, 59-73. https://doi.org/10.1016/j.gexplo.2014.11.007

[23] Sadeghi, B., Moarefvand, P., Afzal, P., Yasrebi, A.B. and Saein, L.D. (2012) Application of Fractal Models to Outline Mineralized Zones in the Zaghia Iron Ore Deposit, Central Iran. *Journal of Geochemical Exploration*, **122**, 9-19. https://doi.org/10.1016/j.gexplo.2012.04.011

[24] Sadeghi, B. (2021) Concentration-Concentration Fractal Modelling: A Novel Insight for Correlation between Variables in Response to Changes in the Underlying Controlling Geological-Geochemical Processes. *Ore Geology Reviews*, **128**, Article ID: 103875. https://doi.org/10.1016/j.oregeorev.2020.103875

[25] Sadeghi, B. (2021) Simulated-Multifractal Models: A Futuristic Review of Multifractal Modeling in Geochemical Anomaly Classification. *Ore Geology Reviews*, **139**, Article ID: 104511. https://doi.org/10.1016/j.oregeorev.2021.104511

[26] Datta, B., Amirabdollahian, M., Zuo, R. and Prakash, O. (2016) Groundwater Contamination Plume Delineation Using Local Singularity Mapping Technique. *International Journal of GEOMATE*, **11**, 2435-2441. https://doi.org/10.21660/2016.25.5157

[27] Cheng, Q., Agterberg, F.P. and Ballantyne, S.B. (1994) The Separation of Geochemical Anomalies from Background by Fractal Methods. *Journal of Geochemical Exploration*, **51**, 109-130. https://doi.org/10.1016/0375-6742(94)90013-2

[28] Cheng, Q., Xu, Y. and Grunsky, E. (2000) Integrated Spatial and Spectrum Method for Geochemical Anomaly Separation. *Natural Resources Research*, **9**, 43-52. https://doi.org/10.1023/A:1010109829861

[29] Chen, G., Cheng, Q., Zuo, R., Liu, T. and Xi, A.Y. (2015) Identifying Gravity Anomalies Caused by Granitic Intrusions in Nanling Mineral District, China: A Multifractal Perspective. *Geophysical Prospecting*, **63**, 256-270. https://doi.org/10.1111/1365-2478.12187

[30] Afzal, P., Alghalandis, Y.F., Khakzad, A., Moarefvand, P. and Omran, N.R. (2011) Delineation of Mineralization Zones in Porphyry Cu Deposits by Fractal Concentration-Volume Modeling. *Journal of Geochemical Exploration*, **18**, 220-232. https://doi.org/10.1016/j.gexplo.2011.03.005

[31] Prakash, O. and Datta, B. (2014) Multi-Objective Monitoring Network Design for Efficient Identification of Unknown Groundwater Pollution Sources Incorporating Genetic Programming Based Monitoring. *Journal of Hydrologic Engineering*, **19**, 61-75. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000952

[32] Datta, B. and Peralta, R.C. (1986) Interactive Computer Graphics-Based Multiobjective Decision-Making for Regional Groundwater Management. *Agricultural Water Management*, **11**, 91-116. https://doi.org/10.1016/0378-3774(86)90023-5

[33] Mahar, P.S. and Datta, B. (2001) Optimal Identification of Ground-Water Pollution Sources and Parameter Estimation. *Journal of Water Resource Planning and Management*, **127**, 20-30. https://doi.org/10.1061/(ASCE)0733-9496(2001)127:1(20)

[34] Sun, J. (2004) A Three-Dimensional Model of Fluid Flow, Thermal Transport, and Hydrogeochemical Transport through Variably Saturated Conditions. M.S. Thesis, Department of Civil and Environmental Engineering, University of Central Florida, Orlando.

[35] Yeh, G.T., *et al.* (2004) HYDROGEOCHEM 5.0: A Three-Dimensional Model of Coupled Fluid Flow, Thermal Transport, and Hydrogeochemical Transport through Variably Saturated Conditions—Version 5.0. Oak Ridge National Laboratory, Oak Ridge.

[36] Manderson, J. (2011) Report for Strengthening Basin Community-Planning Study Business Case—Enhancing Dubbo's Irrigation System. Dubbo City Council, Dubbo.

[37] Puech, V. (2010) Upper Macquarie Groundwater Model. Technical Report VW04680, Office of Water, NSW Government and National Water Commission.

[38] Harbaugh, A.W., Banta, E.R., Hill, M.C. and McDonald, M.G. (2000) MODFLOW-2000, the U.S. Geological Survey Modular Ground-Water Model—User Guide to Modularization Concepts and the Ground-Water Flow Process. U.S. Geological Survey Open-File Report 121, 1-92. https://doi.org/10.3133/ofr200092

[39] Zheng, C. and Wang, P.P. (1999) MT3DMS: A Modular Three-Dimensional Multispecies Model for Simulation of Advection, Dispersion and Chemical Reactions of Contaminants in Groundwater Systems; Documentation and User's Guide. Contract Report SERDP-99-1, US Army Engineer Research and Development Center, Vicksburg.

[40] Prakash, O. and Datta, B. (2015) Optimal Characterization of Pollutant Sources in Contaminated Aquifers by Integrating Sequential-Monitoring-Network Design and Source Identification: Methodology and an Application in Australia. *Hydrogeology Journal*, **23**, 1089-1107. https://doi.org/10.1007/s10040-015-1292-8