

A web-based platform for integrated groundwater data management

^{1,2}El Sawah, S., ^{1,2}Hicks, A., ^{1,2}Manger, P., ³Athanasiadis, I.N., ^{1,2,4}Croke, B.F.W., ^{1,2}Jakeman, A.J.

¹*Fenner School of Environment and Society, The Australian National University, Canberra, Australia*

²*National Centre for Groundwater Research and Training, The Australian National University, Canberra, Australia*

³*Democritus University of Thrace, Xanthi, Greece*

⁴*Department of Mathematics, The Australian National University, Canberra, Australia*
Email: sondoss.elsawah@anu.edu.au

Abstract:

Integrated management of groundwater involves implementing policies that can achieve economic and ecologically sustainable outcomes. Developing integrated policies needs to be supported by long term reliable data of groundwater levels and factors that influence recharge and extraction rates. In Australia, hydrological and groundwater data are currently collected and managed by different agencies, and scattered in various formats and locations. The objective of the National Groundwater Information System (NGIS) project is to provide a unified framework for collecting and managing groundwater data. The implementation of the NGIS requires: (1) setting up mechanisms for acquiring and storing data, (2) developing methods and tools to examine and improve the quality of collected data, (3) developing analytical tools to support various research and management purposes, and (4) providing methods and protocols for data access and sharing among relevant stakeholders. In this paper, we present a web-based platform that aims to serve these requirements in six sites: Wellington (NSW), Namoi (NSW), Ti Tree (NT), Willunga (SA), North Stradbroke Island (QLD), and Ovens (Victoria). The platform presents an architecture (i.e. technologies and tools) and data model that can be transferred to other locations. The paper gives an overview of the system architecture and the implemented prototype version. It concludes by highlighting future directions to extend the capacity and use of the platform

Keywords: *Integrated groundwater management, National Groundwater Information System, data management, web-based technologies*

1 INTRODUCTION

Improved understanding and management of groundwater need to be backed by long term reliable data sets. This rationale has driven a worldwide growing attention towards investment in developing groundwater data warehouses and information systems (Refsgaard *et al.*, 2010). In Australia, hydrological and groundwater data are currently collected and managed by different agencies, and scattered in various formats and locations. This fragmented approach hampers, both technically and legally, the development of national-scale integrated groundwater management plans. Acknowledging the importance of a unified framework for collecting and managing groundwater data, the Australian Water Commission launched a project to implement a National Groundwater Information System (NGIS) in 2009. The Water Division of the Australia Bureau of Meteorology (BoM) takes the leading role in managing the NGIS project.

As a part of the NGIS project, the National Centre for Groundwater Research and Training (NCGRT) has received \$15 million funding from the Education Investment Fund (EIF) through the Australian Super Science Initiative, to be spent on groundwater infrastructure (i.e. field and cyber) at six sites throughout Australia: Wellington (NSW), Namoi (NSW), Ti Tree (NT), Willunga (SA), North Stradbroke Island (QLD), and Ovens (Victoria). The development of the cyber infrastructure includes building a web-based platform to support integrated groundwater data management by:

- (1) bringing together spatial and non-spatial data types required to support groundwater research and decision making (e.g. runoff data, climate data, groundwater levels);
- (2) providing policies, processes and procedures to enable effective management (i.e. both technically and legally) across data lifecycle phases: creation and receipt, distribution, use, and maintenance;
- (3) providing a wide range of functions and tools to support users' various requirements for data storage and access, quality assurance, data analysis and visualization; and
- (4) facilitating information sharing among stakeholders (i.e. public, managers, scientists).

The development of the platform follows an iterative process which encompasses the design and implementation of a prototype system including: a high-level architecture of the platform, schematic database structure, a prototype single-site version, and a documentation of the protocols and practices for database population. The prototype version is used to demonstrate key system functions and capabilities, and facilitate communication among the project stakeholders about the final product (i.e. funders, software developers, project managers, site managers).

The main objective of this paper is to share the outputs of the prototype phase. Specifically, we cover the following:

- Present an overview of the platform architecture: functions, components and technologies (Section 2);
- Describe the implemented prototype version (Section 3); and
- Highlight future directions to extend the capacity and use of the platform (Section 4)

2 PLATFORM ARCHITECTURE

The platform is built using an open-source software freely available from the internet to allow for a wider access of data with lower costs and greater flexibility in maintaining and upgrading the system. The system architecture consists of *four* main components: (1) relational database, (2) SISS, (3) analysis and visualization tools via a web-portal, and (4) quality assurance tools (See Figure 1). This architecture represents a framework of technologies and tools that can be transferred to other locations. In this section, we describe each of the system components.

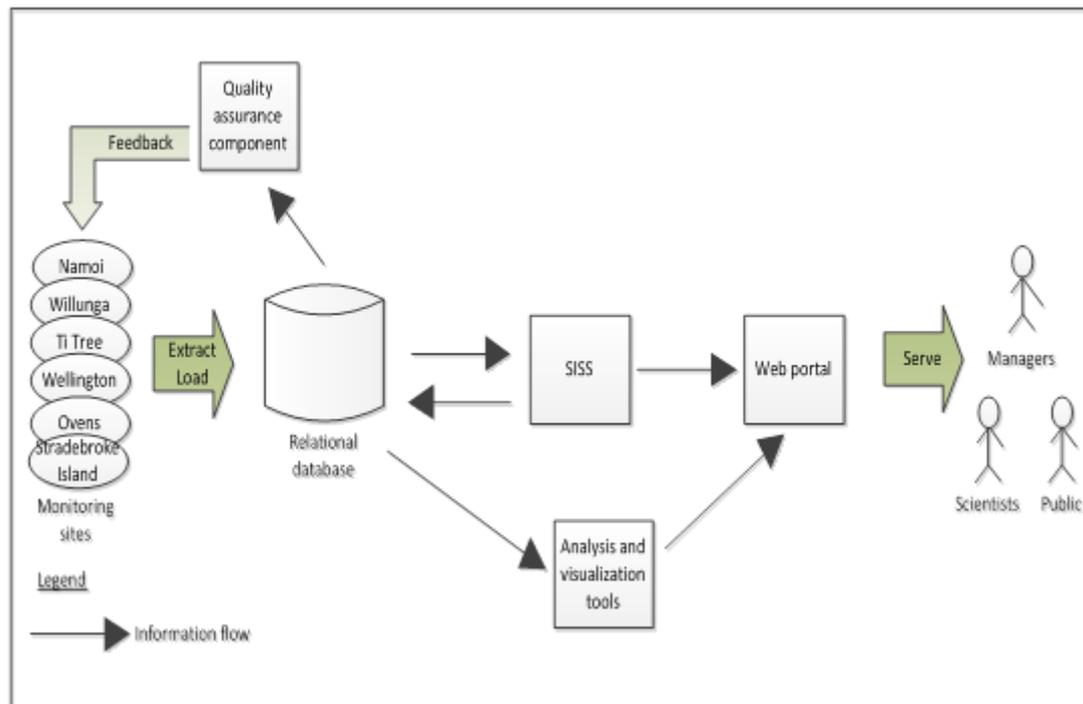


Figure 1: A high level architecture of the web-based groundwater data management platform

2.1 Relational database

At the core of the system, a relational database is built to store data sets and associated metadata, along with user access rules and policies. For each geographic site, a dataset is created and linked to the recording sensor, to allow for tracking of site-specific information, such as: attribution references, ownership, copyright, information access rights and rules. Data sets include time series data organized into five types: climate data, stream flow data, bore data, soil water data, and injection point data. Figure 2 shows the preliminary conceptual data model used in creating the database and validating acquired datasets. Individual time-series can be associated with quality indicator information, linking the value with a quality indicator, allowing errors and anomalies to be flagged during multiple stages of the quality assurance process (See section 2.4). The database uses the PostgreSQL object-relational database with the PostGIS extension for storing and querying spatial data.

2.2 Data analysis and visualization tools via the web portal

The web front provides visualization of the data in graph or table form. It is powered by an Apache Web Server running PHP. GeoServer software provides the spatial information in various formats, which are displayed in the user's web browser using the OpenLayers library. This spatial information provided by GeoServer can be used directly by common GIS software, such as ArcGIS. The time-series data will be provided for download in a number of formats: common text formats (e.g. CSV), and in binary data formats (e.g. NetCDF).

2.3 SISS

The Spatial Information Services Stack (SISS) is a suite of tools for spatial data interoperability. It uses tools that implement the standards set out by the Open Geospatial Consortium (OGC). The stack extends the normal capabilities of GeoServer, allowing it to provide not only the spatial information in Geography Markup Language (GML), but also richer, structured metadata and the underlying time-series data attached to the spatial data, in formats such as BoM's Water Data Transfer Format (WDTF), GroundwaterML, and the standard currently under development to bring the various water-domain formats together, i.e. WaterML 2.0. This effort aligns with the work being done by AuScope to provide data interoperability to the mineral sciences using GeoSciML, and is similar to other international efforts such as CUAHSI's Hydrological Information System.

2.4 Quality assurance tools

The quality assurance process is essential for supporting both the scientific work and the dissemination of data to a wider audience. Local experts in each site will check data quality, using statistical indexes, boundary rules, and more importantly their experience and knowledge of local conditions. Data submitted to the central cyber infrastructure will include both raw and quality-assured data, and a log of maintenance operations in each site. Based on this information, we aim to develop a semi-automated system, to perform a second round of quality assurance checks to all data stored in the central database. The QA sub-system will ensure that a common vocabulary is used for encoding data and errors, and that a common process will be applied to identify erroneous data that have not been identified by the checks at site level. The latter will be performed by exploiting data from neighboring sites.

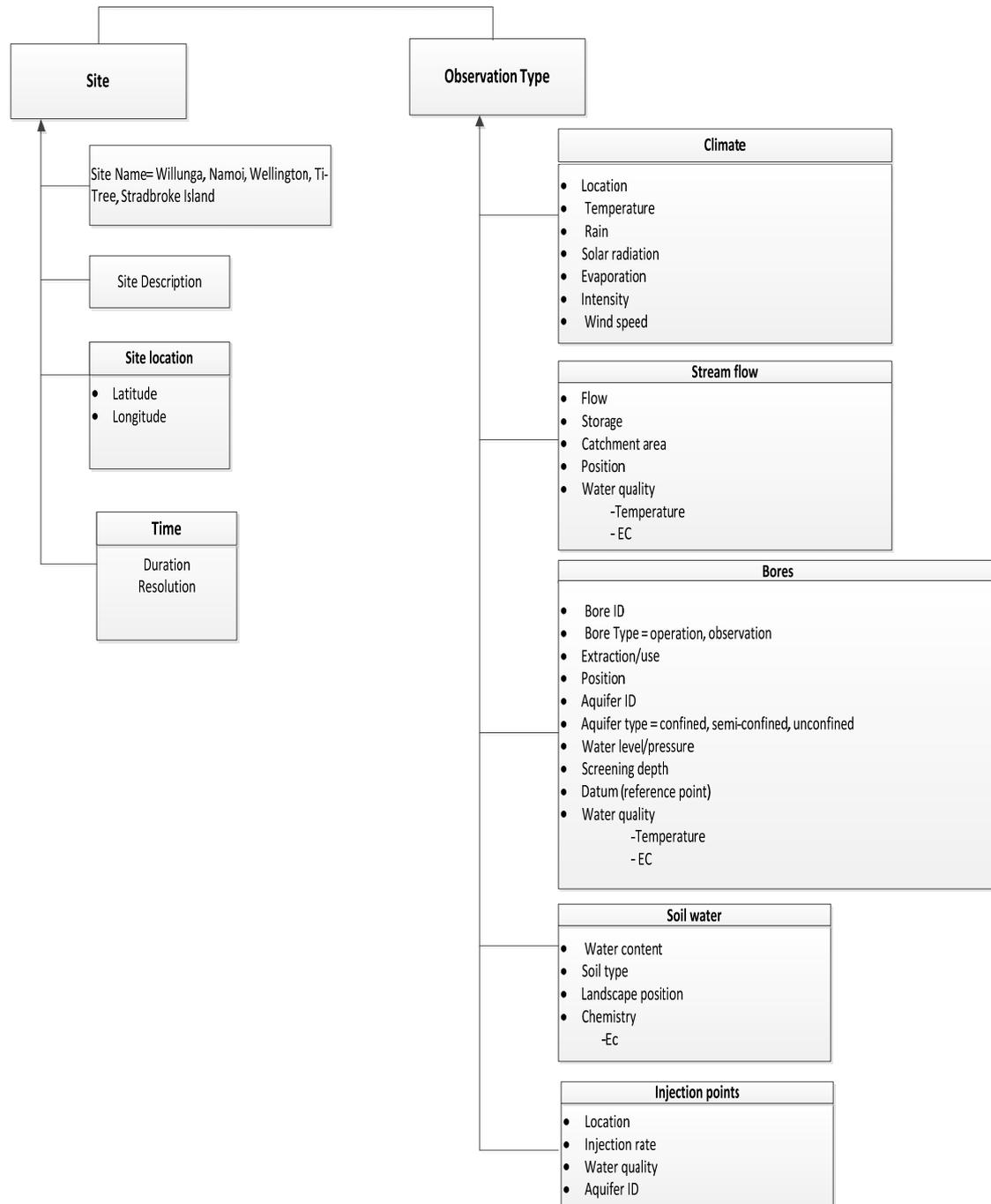


Figure 2: The conceptual data model underlying the database

3 PROTOTYPE VERSION: WELLINGTON SITE

3.1 Description of the Wellington site

The Wellington Site is located at a 400ha research farm owned by the UNSW and is being developed as a major teaching laboratory for groundwater field methods. The study site is underlain by folded and fractured Ordovician metamorphic sequences which will form one of the fractured-rock type sequences being investigated. The adjacent limestone at Wellington and the Devonian granite sequences close by, will form two others. Existing monitored catchments operated by NSW DPI will be incorporated into this LTERM to supplement the regional monitoring activities.

Long term monitoring and investigations at this site will allow firm predictions of useful groundwater resources in a variety of fractured rock types and also to monitor the response of these systems to climate change. The systems are particularly important for understanding and monitoring dryland salinity and for supporting the increasing demands for groundwater from fractured rock environments. Changed vegetation management will also be trialled at the site so that the impacts on the surface and subsurface hydrology can be ascertained.

The LTERM facility at Wellington has heritage in a number of related research efforts (driven predominantly in the United States) which seek to monitor aspects of the Earth system in an integrated manner. Some of these activities include the Long Term Ecological Research (LTER) network of sites (<http://www.lternet.edu/>), the Consortium of Universities for the Advancement of Hydrologic Science (CUASHSI), Hydrologic Observatories (<http://www.cuahsi.org/>), as well as other process focused investigations such as FluxNET (<http://www.fluxnet.ornl.gov/>) for land surface fluxes; the Aerosol Robotic Network (AERONET) (<http://aeronet.gsfc.nasa.gov/>) for aerosol quantification and distribution; and Oklahoma's meso-scale network of environmental monitoring stations (<http://www.mesonet.org/>). Related Australian efforts include the University of Melbourne OZNet soil moisture sites (<http://www.oznet.unimelb.edu.au/>) and the CSIRO led multi-institution collaboration of flux observations (<http://www.cmar.csiro.au/>), which has a more recent incarnation in the Terrestrial Ecosystem Research Network (TERN).

The Wellington LTERM differs quite significantly from these activities in both the scope and intent of observations being proposed. While most existing networks focus on particular processes within the Earth system (i.e. soil moisture, surface fluxes, surface meteorology), the Wellington facility is driven by the realization that improved understanding of groundwater relies on the accurate characterisation and quantification of the numerous interconnected hydrometeorological processes and mechanisms that influence behavior and response. To improve this understanding requires accurate monitoring of the principal groundwater-environmental interfaces. These include surface-water/groundwater (SW/GW) and groundwater-vegetation-atmosphere (GW/V/A) interactions.

3.2 Data description (including metadata)

Schematics of LTERM elements are shown in Figures 4 and 4 below.

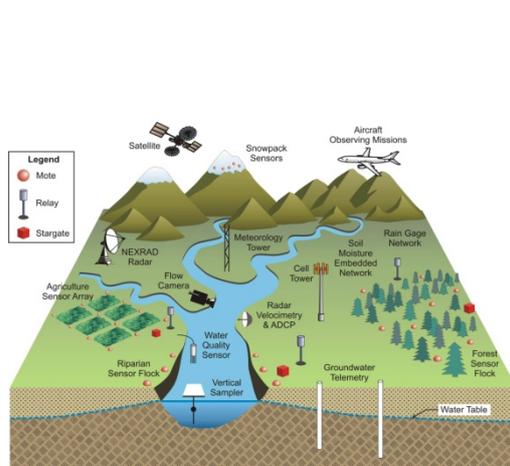


Fig. 3 Elements of an LTER

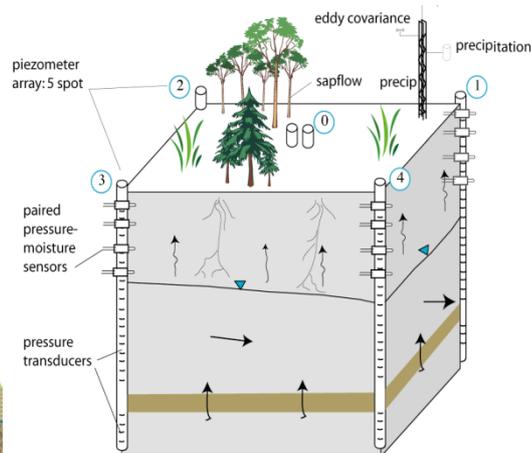


Fig. 4 Detailed site monitoring

The interconnectivity of surface water and groundwater is recognised as a priority in the NWI Groundwater Action Plan. To this end, a variety of terrestrial (surface and sub-surface) and atmosphere measurement programmes will be established, using the EIF Superscience funds, so that research into the following aspects can be carried out:

- Rainfall characterisation: an X-band Doppler-radar facility will monitor rainfall across a 30km radius focused around Wellington, providing accurate measurement of the primary driver of groundwater and surface hydrological response.
- Groundwater response to rainfall: a network of monitoring bores, rain gauges, water level monitoring and associated soil moisture monitoring sensors will provide a clear assessment of point-scale discharge/recharge relationships. These bores will be sited both on the farm but also throughout the surrounding area. An average density of 1 per 25 km² is anticipated with a total of 50 bores in the area around Wellington (20km radius). These measurement sites will provide an integrated approach to understanding water-table dynamics across local-to-regional scales, allowing real-time assessment of groundwater response in fractured rocks. Long-term measurement of these sites will provide information on fundamental drying and wetting behaviour across spatial and temporal scales. River gauging and bore-hole transects will be installed at the study site to monitor surface water-groundwater exchanges as a response to rainfall;
- Groundwater-vegetation interactions: the exchange of moisture between the terrestrial surface (through different vegetation cover types) and the atmosphere will be determined using a combination of flux towers and in-tree transpiration devices, as well as through undertaking detailed classification and characterisation of vegetation type and structure and radiation based forcing of vegetation response. Instruments based on eddy-covariance and scintillometry will be used to monitor surface moisture and heat flux exchanges between the terrestrial surface and the atmosphere, providing accurate characterisation of diffuse discharge relationships;
- Geophysical characterisation: groundwater moves through fractures in hard rock. The number and degree of interconnectivity of fractures is of fundamental importance in determining the quantity of groundwater storage that can be developed - yet little is known of the degree and nature of fracturing. It is envisaged in current resource planning that groundwater abstraction from fractured rock aquifers can provide a significant increment of water to rural communities. The LTERM will be used to test this hypothesis. International studies in fractured rocks have focussed on high-level radioactive waste disposal and have concentrated on detecting flow through significant individual fractures. At this site, the significance of interconnected fracture systems can be evaluated as an important resource. To establish statistically meaningful estimates of fracture flow, many hundreds of measurements are required. High density networks (10 to 15 bores per hectare) are being installed in the research farm as well as the adjacent limestone and granite systems.
- In addition, a suite of high-precision, cutting-edge instrumentation will be sought to provide greater detail on the linking mechanisms and enhanced process representation between these water-cycle components and stores. One of these approaches will include analysis of stable water isotope measurement in both the liquid (rainfall) and vapor (evapotranspiration) phases, which will enable constraints on sources and sinks of hydrological exchange to be better represented and identified.

All data will be transmitted wirelessly (where appropriate) directly back-to-base for quality checking and processing before making it freely available on the web. Much of the data will be uploaded in real-time, providing an unprecedented level of information to the science community and also to the general public. Efforts to engage and link with national and international research activities will be a priority of the LTERM effort, so as to maximise the scientific exposure and utility of collected data. The data sets that will be collected as part of the LTERM will be of interest to multi-disciplinary scientific research and activities. They will be of particular interest to the international remote sensing community, in that there will be a large-scale well calibrated facility that can be incorporated into a number of upcoming National Aeronautics and Space Administration (NASA) and European Space Agency (ESA) missions and calibration/validation campaigns.

3.3 Existing functions

The initial version of the site focuses on groundwater level data for the Wellington Super Science site, collected over a period of six years for a previous research project done by UNSW. This data will be used as a proof of concept for the system while the data collection process is refined and the various sensor types are installed and configured.

The locations of the monitoring bores for the Baldry node of the Wellington site are presented on a map view. Individual bores can then be selected to view their information, including the time-series water level information, which can be viewed as a graph or downloaded for analysis.

Work is currently being done to select and compare the levels of multiple bores, as well as improve the interface and responsiveness of the system. Support for other types of data is being added to the system, beginning with weather and stream level data.

4 FUTURE DEVELOPMENTS

- Beyond physical data types: Including socio-economic data

Groundwater systems do not only include hydro-ecological processes but social and economic aspects as well. Therefore, monitoring the physical system state (e.g. groundwater levels) and analyzing related variables are not sufficient to improve understanding and management of the whole system. Socio-economic data is an integral part of the groundwater database. Examples of socio-economic data include population size, formation, water consumption rates and crops grown (e.g. yield and type). To include socio-economic data, the existing conceptual data model needs to be extended and updated. What are the types of socio-economic observations to be collected and stored? What is the appropriate level of aggregation? What are the legal issues to be considered including data confidentiality? These questions and more need to be discussed with the project stakeholders.

- Beyond the NGIS: Towards the development of a national Groundwater-Decision Support System (GW-DSS)

DSS are defined as information systems that support decision making activities. DSS can be broadly classified into: data-driven DSS and model driven DSS. The primary objective of data-driven DSS is to help analyse large pools of data accumulated over long periods. This enables users to induce information about the historical and present system behaviour (e.g. how did groundwater level, in a particular area, change over the last 10 years?). Model-driven DSS are distinguished by a model base to help users answer policy-oriented and future-visioning questions (e.g. what are the predicted outcomes of changes in groundwater policies in future?).

In the next development phase, our objective is to extend the platform to provide model-based decision support tools (i.e. GW-DSS). Towards this aim, the NCGRT funds two major projects for developing integrated groundwater models in two Super science sites: Namoi and Willunga. These projects are complementary to the NGIS initiative. Collected data will be used in building and validating groundwater management models.

- Intelligent data analysis toolbox

On intelligent data analysis, we aim to incorporate data-driven intelligence for improving decision making, especially for rudimentary and repetitive tasks. To achieve this, an interactive software system could be deployed for eliciting expert knowledge, which is hard to capture in a straightforward manner. We envision a tool where scientists identify interesting cases that need to be considered in the future, and to reject trivial or false alarms discovered by the data driven system. Such an approach will put the foundations for a win-win cycle, where experts deposit their expertise along with their data, while at the same time review the results of data-driven models.

REFERENCES

- Refsgaard, J. C., Højberg, A. L., Møller, I., Hansen, M. and Søndergaard V. (2010) “Groundwater Modeling in Integrated Water Resources Management—Visions for 2020”, *Groundwater*, 48(5), pp 633–648