Key Points:
- Data and models at multiple scales were integrated to assess the geochemical impact of managed aquifer recharge with recycled water
- Injection of aerobic water triggered pyrite oxidization, but native buffering capacity prevented groundwater acidification
- Injection of low-salinity water triggered fluoride release, but sufficient attenuation occurred

Supporting Information:
- Supporting Information S1

Correspondence to:
H. Prommer,
henning.prommer@csiro.au

Assessing and Managing Large-Scale Geochemical Impacts From Groundwater Replenishment With Highly Treated Reclaimed Wastewater

Jing Sun[1,2,3] iD, Michael J. Donn[2], Philippe Gerber[2,4], Simon Higginson[5], Adam J. Siade[2,3,6] iD, David Schafer[2,3,6] iD, Simone Seibert[7], and Henning Prommer[2,3,6] iD

[1]State Key Laboratory of Environmental Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang, China, [2]CSIRO Land and Water, Wembley, Western Australia, Australia, [3]School of Earth Sciences, University of Western Australia, Perth, Western Australia, Australia, [4]Hydrique Ingénieurs, Lausanne, Switzerland, [5]Water Corporation of Western Australia, Leederville, Western Australia, Australia, [6]National Centre for Groundwater Research and Training, Adelaide, South Australia, Australia, [7]Federal Institute for Geosciences and Natural Resources, Hannover, Germany

Abstract Reuse of wastewater through a combination of advanced wastewater treatment (AWT) and managed aquifer recharge (MAR) is an important water management option. As an integral part of any AWT-MAR system, the geochemical compatibility of the recharged water with the targeted aquifer must be assessed to avoid groundwater quality deterioration. Although short-term field experiments may uncover potentially concerning sediment-water disequilibria, an advanced analysis is often required to understand the long-term geochemical impacts of large-scale MAR. Here, we develop and apply a pragmatic approach to upscale and verify the previously derived local-scale geochemical understanding to the spatial and temporal scale required for assessing and managing large-scale and long-term impacts resulting from the recharge of AWT-processed water. We use Australia's largest MAR scheme, in which aerobic, purified water is injected into deep, anaerobic, pyritic aquifers as an illustrative example. In this case, the local-scale understanding was derived from a comprehensive field trial and the interpretation of the data through a trial-scale high-resolution reactive transport model (RTM). Based on (i) the trial-scale RTM, (ii) new data from the early phase of the full-scale MAR, and (iii) downscaling of an existing, regional-scale flow model, we developed an upscaled RTM to assess two critical issues for the large-scale groundwater replenishment of Perth deep aquifers, that is, the sustainability of native pH buffering processes and the dynamics of fluoride release and attenuation. In a final step we illustrate how the upscaled RTM can be applied to assess how changes in the AWT affect long-term groundwater pH changes.

1. Introduction

Managed aquifer recharge (MAR) is the intentional recharge of water to aquifers for later recovery or for environmental benefits. MAR can act as an environmental buffer and can be used to temporarily store various source water types, especially surface waters that are available in excess at times when water demand is low (Dillon et al., 2018). Where such excess water is unavailable, (re)use of treated wastewater is becoming an increasingly important water management option (Pavelic et al., 2007; Shah, 2014). This is particularly relevant where the availability of water resources is in decline, for example, due to a drying climate, while water demand, for example, from urbanization, is increasing (Loáiciga et al., 2000; Xanke et al., 2016). Advanced wastewater treatment (AWT) can be used to reclaim large volumes of wastewater that is otherwise lost, for example, via discharge to ocean outfalls, while achieving water quality standards such that the AWT-produced water is acceptable for drinking water supply.

Nevertheless, to date direct (re)use of this water as a source of drinking water has, with few exceptions such as the NEWater initiative in Singapore (e.g., Lee & Tan, 2016), not seen a significant uptake, due to various barriers including the psychological "yuck factor" as a key obstacle (Leong & Lebel, 2020; Mankad & Tapsuwan, 2011; Wester et al., 2016). Risk perception was reported as the most dominant social factor influencing the acceptance of alternative drinking water sources (Mankad & Tapsuwan, 2011; Marks et al., 2008).

An environmental buffer, such as an aquifer, can greatly aid in increasing social acceptance of recycled water by providing the disconnect between the wastewater source and the potable supply. As a result, indirect potable reuse, which combines AWT with a subsequent MAR step to facilitate water recycling for potable supply, has seen a steadily increasing number of small-scale and several large-scale implementations over the last >30 years (Corneille & Dawes, 2001; Houtte & Verbauwhede, 2008; Maliva, 2019).

Using aquifers for water storage offers several advantages including avoiding evaporative losses, minimizing the footprint of storage, replenishing depleted aquifers, protecting against land subsidence, and creating hydraulic barriers to protect against saline intrusion (Bekele et al., 2019; Dillon et al., 2018). However, the aquifer storage zone is not inert; therefore, MAR can result in improvements (e.g., inactivation of pathogens and degradation of trace organic chemicals) or deteriorations on the quality of the groundwater (e.g., Hellauer et al., 2018; Pavelic et al., 2007; Prommer et al., 2019; Sharma et al., 2012). Depending on the combination of (i) the injectant composition and (ii) the sediment composition and reactivity, the recharged water can induce a geochemical disequilibrium within the receiving aquifer matrix. The type and magnitude of this disequilibrium controls the risk of water quality deterioration. One of the typical risks is the mobilization of toxic metal(loid)s. For example, Oren et al. (2007) attributed in situ mobilization of geogenic manganese during a MAR site in Israel to reduction of sedimentary manganese oxides under suboxic conditions. Jones and Pichler (2007) and Wallis et al. (2011) linked the temporary release of arsenic at aquifer storage and recovery sites in southwest Florida to the oxidation of arsenic-bearing pyrite by the oxidized injectant. McNab et al. (2009) attributed elevated arsenic concentrations at a MAR site in the Central Valley, California, to the more alkaline character of the recharge water and arsenic desorption due to pH changes. Fakhreddine et al. (2015) explained arsenic mobilization at a MAR site in Orange County, California, to result from the lack of divalent cations in the purified injectant and subsequent arsenate desorption from negatively charged clays. These MAR studies have demonstrated that concerns of groundwater quality deterioration are even warranted in cases where neither the ambient groundwater nor the injectant contained any elevated levels of contaminants.

The aforementioned cases also highlight that the planning and implementation of MAR schemes requires a thorough, site-specific geochemical understanding, as well as the necessary tools for predicting the potential for groundwater quality risks and identifying technical options to mitigate these risks. Such a site-specific understanding is typically developed by a combination of hydrogeological and geochemical characterization activities (e.g., Rathi et al., 2017), laboratory tests (e.g., Fakhreddine et al., 2015; Schafer et al., 2018) and short-term field injection trials (e.g., Seibert et al., 2014, 2016), or push-pull tests (e.g., Fakhreddine et al., 2020; Prommer et al., 2018; Rathi et al., 2017). Short-term field experiments can often provide some early warning signs of geochemical disequilibrium, and the interpretation of these results can be used to identify the geochemical mechanisms that control the water quality evolution as the injectant migrates through the native aquifer sediments (Rathi et al., 2017). Additionally, the mechanistic understanding derived from the field experiment can be used to modify the AWT process to enhance the geochemical compatibility between the injectant and the aquifer sediments (Prommer et al., 2018). However, the nature of such short-term experiments precludes the discovery of any geochemical mechanisms that rely on the longer-term, sustained abundance of potentially beneficial aquifer constituents such as carbonate or iron oxide minerals that might be actively consumed, in preventing groundwater quality deterioration. To predict the impacts of large-scale groundwater replenishment and thus the long-term sustainability of a MAR operation, a more advanced analysis is required (Ganot et al., 2018).

In this study, we developed a pragmatic approach to integrate geochemical and hydrogeological data with numerical models, at multiple scales, to assess and manage large-scale and long-term impacts resulting from groundwater replenishment, by using Australia's largest AWT-MAR scheme as an illustrative example. This AWT-MAR project was planned and implemented over the last 15 years in Perth, Western Australia, during which a series of site-specific studies were conducted, including a 4-year-long groundwater replenishment trial (GWRT), as well as associated numerical models that analyze the trial data (Schafer et al., 2018, 2020; Seibert et al., 2014, 2016). With current extensions, this AWT-MAR project will soon replenish deep aquifers with over 28 GL year^{-1} of recycled water, rising to >100 GL year^{-1} over the next decades (Dillon et al., 2018). Given the large aquifer volumes that will be affected by these GWR operations, it is important to understand and predict future groundwater quality evolution. This prediction requires the development of a suitable upscaling approach that incorporates the quantitative geochemical understanding that was developed at

the GWRT scale, while also considering the hydrogeological features that impact the kilometer-scale reactive transport behavior of the injectant and its interaction with the Perth aquifer system over several decades of GWR operations. Accordingly, we developed a new and computationally feasible methodology to tackle this "upscaling problem." The key steps of the methodology included (i) the development and evaluation of a local-scale surrogate dual-domain

reactive transport model (RTM) with new parameters representative for the new scale and for the new monitoring protocol and (ii) the development of an injectant plume-scale RTM that integrates the downscaled groundwater flow information from the regional model with the geochemical reaction conceptualization and parameterization derived during the surrogate RTM development. This injectant plume-scale RTM was used to assess the two potentially most critical groundwater quality risks. First, an assessment of whether there is any risk that groundwater acidification could be triggered by the oxidation of pyrite, a ubiquitous mineral in Perth deep aquifers, and specifically, whether the previously identified in situ pH buffering mechanisms (Seibert et al., 2016) will persist over the long term, maintaining pH at levels that prevent the risk of metal mobilization. Second, the mobilization of fluoride occurred during GWRT, which was shown to be triggered by the low ionic strength of the injectant resulting from AWT (Schafer et al., 2020). While fluoride concentrations have not exceeded any health guideline levels to date, developing a process-based quantitative understanding is crucial for reliable predictions of the long-term and large-scale behavior of fluoride. The presented upscaling approach in this study will be applicable to support the planning, implementation, and operation of many other large-scale MAR schemes.

2. Background

2.1. Study Site

In Perth, due to reducing rainfall and surface runoff, combined with a continued population growth, a projected freshwater deficit of up to 250 GL year^{-1} by 2030 was predicted (CSIRO, 2009). To reduce this gap, a range of options were considered, including (i) increased water efficiency, (ii) development of deep groundwater resources away from the metropolitan area, (iii) seawater desalinization, and (iv) increased recycling of wastewater, including indirect potable reuse through a combination of AWT and MAR. As part of exploring the latter option, the technical feasibility and societal acceptance of indirect potable reuse was investigated. Initial investigations focused on targeting the Leederville aquifer, with injection taking place near the Beenyup wastewater treatment plant, a site located ~20 km north of the city center of Perth, which is referred to hereafter as the Beenyup recharge site (Figure 1). At this site the Leederville aquifer consists of siliciclastic, paralic sediments mostly encompassing the Wanneroo Member of the Cretaceous Leederville Formation (Descourvieres et al., 2010; Leyland, 2011). The upper section of the Leederville aquifer is characterized by tidally influenced distributary channel deposits consisting predominantly of sand, whereas the lower section contains predominantly intertidal flat deposits more enriched in silt and clay (Leyland, 2011). The currently operating (GWR Stage 1) full-scale implementation and its planned extension (GWR Stage 2) will involve multiple injection locations targeting both the Leederville and the deeper Yarragadee aquifer systems. In this study, we focus mainly on the methods developed and their application to the Leederville aquifer, while the same methods can be applied for GWR of the deeper Yarragadee aquifer, or other aquifer systems worldwide, without loss of generality.

2.2. GWRT

Prior to the full-scale implementation of GWR, a series of hydrogeological and geochemical characterization studies were conducted since 2005, culminating in a GWRT between November 2010 and September 2014. For the geochemical characterization, drill cores were collected through diamond core drilling using a triple tube coring method to minimize core loss and mud invasion of the cores. The barrels were split in a specifically designed anaerobic N$_2$-positive pressure chamber installed next to the drilling rig, which prevented oxidation of the sediments (Descourvieres, Prommer, et al., 2010). The GWRT was performed at the Beenyup recharge site through a single injection bore, *LRB1* (Figure 1) (Higginson & Martin, 2012). The injectant consists of highly treated recycled water from the Beenyup waste water treatment plant (Seibert et al., 2016; Water Corporation, 2009). The injectant is characterized by its very low salinity and oxic conditions, thus contrasting to the more saline and reducing conditions in the Leederville aquifer (Table 1). Over the 4-year GWRT period, the average injection rate was ~2.8 ML day^{-1}, and a total of 3,893 ML was injected.
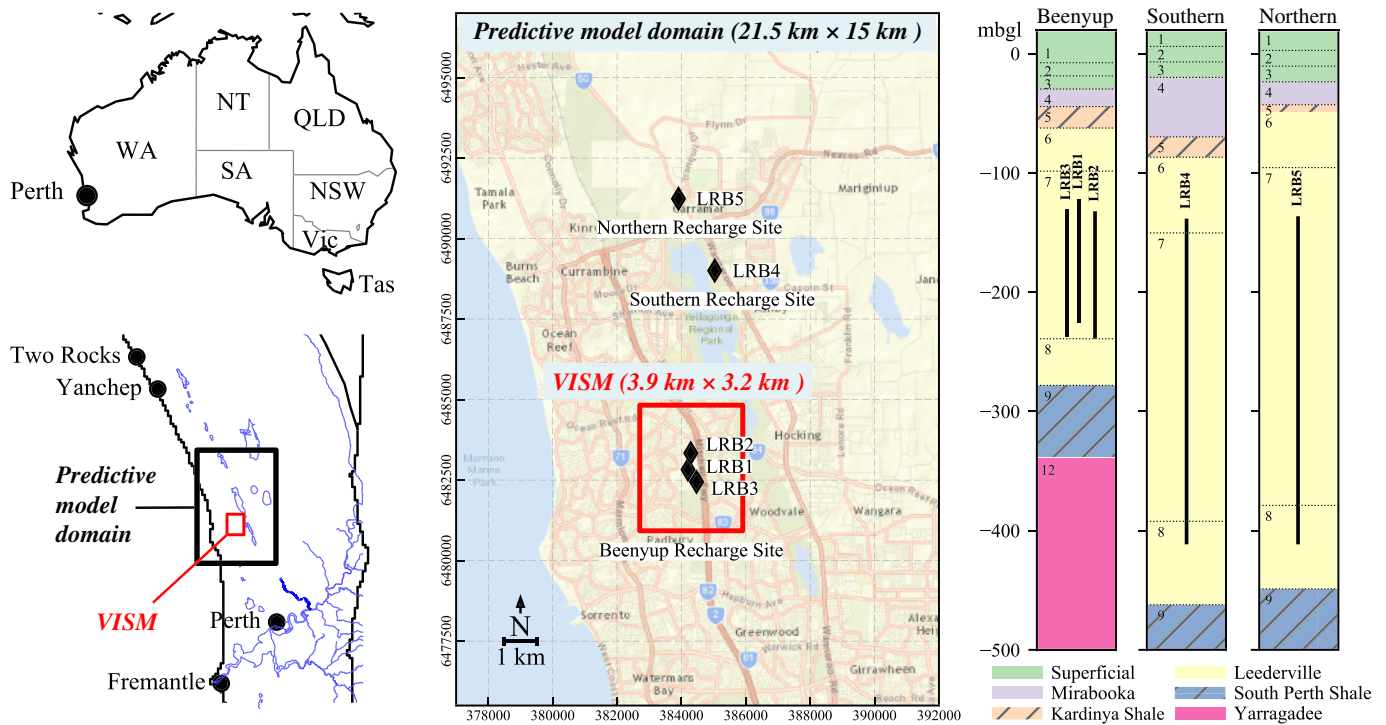
Figure 1. Groundwater replenishment sites in Perth, Western Australia, and the model domains for the VISM and the predictive model. Background map ©OpenStreetMap contributors and the GIS User Community. The right part of the figure shows the geological layout of the sites. The numbering shows the corresponding layer in the PRAMS model. The vertical black bars in the Leederville Formation represent the screened sections of the injection bores *LRB1–LRB5*.

Table 1
Typical Compositions of the Ambient Groundwater in the Leederville Aquifer and the Injectant

Species/Component	Unit	Ambient Leederville	Injectant
pH	—	6.62	7.04
Temperature	°C	25.0	26.8
TDS	mg L^{-1}	511	26
Al	mol L^{-1}	4.24×10^{-7}	<DL
Ba	mol L^{-1}	8.32×10^{-7}	<DL
C(-4)	mol L^{-1}	2.40×10^{-7}	<DL
C($+4$)	mol L^{-1}	1.21×10^{-3}	1.64×10^{-4}
Ca	mol L^{-1}	6.00×10^{-4}	<DL
Cl^a	mol L^{-1}	1.13×10^{-2}	1.66×10^{-4}
F	mol L^{-1}	8.44×10^{-6}	3.64×10^{-6}
Fe($+2$)	mol L^{-1}	1.39×10^{-4}	<DL
K	mol L^{-1}	3.55×10^{-4}	2.26×10^{-5}
Mg	mol L^{-1}	9.24×10^{-4}	<DL
Mn($+2$)	mol L^{-1}	1.04×10^{-6}	<DL
N(-3)	mol L^{-1}	1.19×10^{-5}	3.27×10^{-5}
N($+5$)	mol L^{-1}	<DL	8.29×10^{-5}
Na	mol L^{-1}	9.23×10^{-3}	3.34×10^{-4}
O(0)	mol L^{-1}	<DL	5.00×10^{-4}
P	mol L^{-1}	9.83×10^{-7}	<DL
S(-2)	mol L^{-1}	8.79×10^{-11}	<DL
S($+6$)	mol L^{-1}	4.00×10^{-4}	<DL
Si	mol L^{-1}	4.71×10^{-4}	8.49×10^{-6}
Sr	mol L^{-1}	1.59×10^{-6}	<DL

Note. Some of the measurements were below detection limits (DLs) and therefore reported as "<DL" in the table.
^{a}The reported numbers represent the adjusted values after charge balance.

Table 2
A Summary of All Previously and Newly Constructed GWR-Related Numerical Models

Model	Reference	Type	Lat. extent	Lat. grid size	Vert. extent	# of layers
GWRT-NR	Seibert et al. (2014)	3-D flow and solute/heat transport model	2 × 2 km	6 m around LRB1 and 160 m at the edge	Leederville aquifer at Beenyup recharge site	76 layers
GWRT-R	Seibert et al. (2016) and Schafer et al. (2020)	2-D radial-symmetric reactive transport model	1 km in radial direction	2 m around LRB1 and 100 m at the edge	Leederville aquifer at Beenyup recharge site	76 layers
VISM	This study	2-D plan view dual-domain reactive transport model	3.9 × 3.2 km	20 m around LRB1–LRB3 and 300 m at the edge	Leederville aquifer at Beenyup recharge site	1 layer
PRAMS	de Silva et al. (2013) and CyMod Systems Pty Ltd (2014)	3-D flow model	217 km in the north-south and up to 107 km in the east-west direction	500 × 500 m	Entire Perth aquifer system, thickness varies between 250 and 4,600 m	13 layers
PRAMS-USG	This study	As PRAMS	As PRAMS	As PRAMS	As PRAMS	As PRAMS
GWR-Predict-NR	This study	3-D dual-domain solute transport model	21.5 × 15 km	62.5 × 62.5 m	Entire Perth aquifer system	13 layers
GWR-Predict-R	This study	3-D dual-domain reactive transport model	21.5 × 15 km	62.5 × 62.5 m	Leederville aquifer	3 layers

Pyrite oxidation was shown to occur within the recharged aquifer, but no groundwater acidification was observed over the entire duration of the GWRT (Seibert et al., 2016). While no release of toxic metal(loid)s was observed, release of fluoride and phosphate was observed, with concentrations rising particularly during the injectant breakthrough, before generally receding thereafter (Schafer et al., 2018).

2.3. Local, GWRT-Scale RTM

The characterization data of the Leederville aquifer at the Beenyup recharge site, together with the geophysical and hydrochemical data collected during the GWRT, allowed for the development of a comprehensive understanding of groundwater flow, solute, heat, and reactive transport processes at the local, injectant plume scale (Schafer et al., 2018, 2020; Seibert et al., 2014, 2016). The conceptual hydrogeological model was based on the extensive geophysical and geological characterization (Descourvieres, 2011; Water Corporation, 2009). It was translated into a 3-D flow and conservative transport model with a high vertical and moderate lateral grid resolution and calibrated based on the observed heat and conservative transport behavior that was documented during the GWRT (Seibert et al., 2014) (GWRT-NR, Table 2). In the vertical direction, the 3-D flow and solute/heat transport model consisted of 76 layers to represent the strong vertical heterogeneity that characterizes the Leederville aquifer. The modeling results highlighted that the flow and transport of the injectant was mostly restricted to the sandy layers, whereas heat advected into and out of the underlying and overlying low permeability layers. Using the same vertical discretization, the model was subsequently (i) transformed to a 2-D radial-symmetric model geometry for computational efficiency and (ii) expanded to integrate a comprehensive set of geochemical processes (GWRT-R, Table 2). Based on the results of this RTM, pyrite oxidation during the GWRT showed to be largely driven by the dissolved oxygen contained in the injectant, with a less prominent role of nitrate also serving as electron acceptor; in response to the acidity generated by GWR-induced pyrite oxidation, pH buffering processes, in particular proton exchange, occurred (Seibert et al., 2016). Recently, this RTM was further extended to elucidate the fate of fluoride and phosphate during the GWRT, which attributed the occurrence of the fluoride and phosphate pulses to the low calcium concentration in the injectant and the induced dissolution of carbonate-rich fluorapatite in the aquifer (Schafer et al., 2020).

2.4. Large-Scale Implementation of Groundwater Replenishment

To date, large-scale implementation of GWR in Perth involves two stages, with Stage 1 being operational since August 2017 while planning and construction of Stage 2 operations is currently underway. Similar

to the GWRT, GWR Stage 1 has operated at the Beenyup recharge site (Figure 1). The total injection rate of 14 GL year^{-1} is distributed across four bores of which three (*LRB1*, *LRB2*, and *LRB3*) are targeting the Leederville aquifer and one (*YRB1*) the deeper Yarragadee aquifer. The inner diameters of the injection bores were between 40 and 50 cm. The injection had several flow interruptions that resulted from operational problems. During GWR Stage 1, groundwater samples were collected from the three short-screen monitoring bores (*BY151*, *BY162*, and *BY201*) that were installed for the GWRT 240 m from *LRB1* and from three newly installed long-screen monitoring bores *LMB1–LMB3* that were 60, 60, and 50 m away from the respective injection bores *LRB1–LRB3*. Prior to sample collection, each monitoring bore was purged with a minimum of three casing volumes. pH was recorded immediately following collection. Water quality samples were filtered as required using 0.45-μm syringe filters to polyethylene bottles. All water quality samples were stored on ice immediately after collection and submitted for analysis on the same day. The groundwater sampling campaign captured a broad range of water quality indicators, including physicochemical parameters, major ions, trace metals, and nutrients. In GWR Stage 2, an additional 14 GL year^{-1} will be injected through four new injection bores at two sites, the northern and southern recharge sites, which are located ~8 and 6 km north of the Beenyup recharge site, respectively (Figure 1). In agreement with various stakeholders such as the Department of Water and Environmental Regulation and the Department of Health, Western Australia, a recharge management zone concept was developed (Water Corporation, 2013). The recharge management zone extends over a radial distance of 250 m from each individual injection bore (supporting information Figure S1). The agreement requires that at the boundary of this recharge management zone the groundwater quality meets the federal health and environmental guidelines in Australia (ANZECC-ARMCANZ, 2000). Where dissolved concentrations of particular constituents (e.g., iron) in the ambient groundwater already exceed the guidelines, their concentrations are required not to exceed the pre-GWR concentrations at the boundary of the recharge management zone.

2.5. Regional-Scale Groundwater Flow

The understanding of regional-scale flow processes in the Perth Basin is manifested in the Perth Regional Aquifer Model (PRAMS) (CyMod Systems Pty Ltd, 2014; de Silva et al., 2013), which has been jointly developed by the key stakeholders of the groundwater system and is used to underpin a wide range of water management decisions such as groundwater allocations (Siade et al., 2020). The latest calibrated version, PRAMS v3.5.2, uses MODFLOW-2000 (Harbaugh et al., 2000) to simulate groundwater flow, and a groundwater recharge simulator for the vadose zone, the Vertical Flux Model (Dawes et al., 2012), is used to simulate the spatiotemporally varying groundwater recharge to the superficial aquifer. PRAMS consists of 13 layers with the layering being consistent with the hydrostratigraphic units, including three primary aquifers, the superficial, Leederville, and Yarragadee aquifers (Figure 2 and Table 2). The layers representing the superficial aquifer are bounded on the west side by the coastline, while the confined layers contain an offshore section. The flow boundaries consist of a constant-head boundary at the interface between the ocean and the superficial aquifer, general-head boundaries at the interface of the Leederville aquifer with the offshore fault system, and a no-flow boundary at the interface of the Yarragadee aquifer with the offshore fault system. The Vertical Flux Model, which computes the transient groundwater recharge to the superficial aquifer, takes into account climate variability, land use, vegetation density, soil hydraulic properties, and depth to the water table (Dawes et al., 2012). Groundwater abstraction is modeled using both the well and the recharge packages. The former is used to simulate the large abstraction schemes composed of the Water Corporation production bores and privately licensed bores, whereas the latter is used to simulate the private unlicensed abstraction. PRAMS consists of about 600,000 active grid cells with a constant cell size of 500×500 m, extending 217 km in the north-south, up to 107 km in the east-west, and between 250 and 4,600 m in the vertical direction. More details of the PRAMS setup and flow model simulations can be found in Siade et al. (2017).

3. Upscaling and Downscaling Procedures

3.1. Overview

The large-scale and long-term geochemical impacts of GWR could not be assessed with the RTM that was originally constructed to interpret the GWRT because (i) the flow and transport created by the large-scale injection is no longer radial as the injection is distributed across multiple injection bores, and increasingly
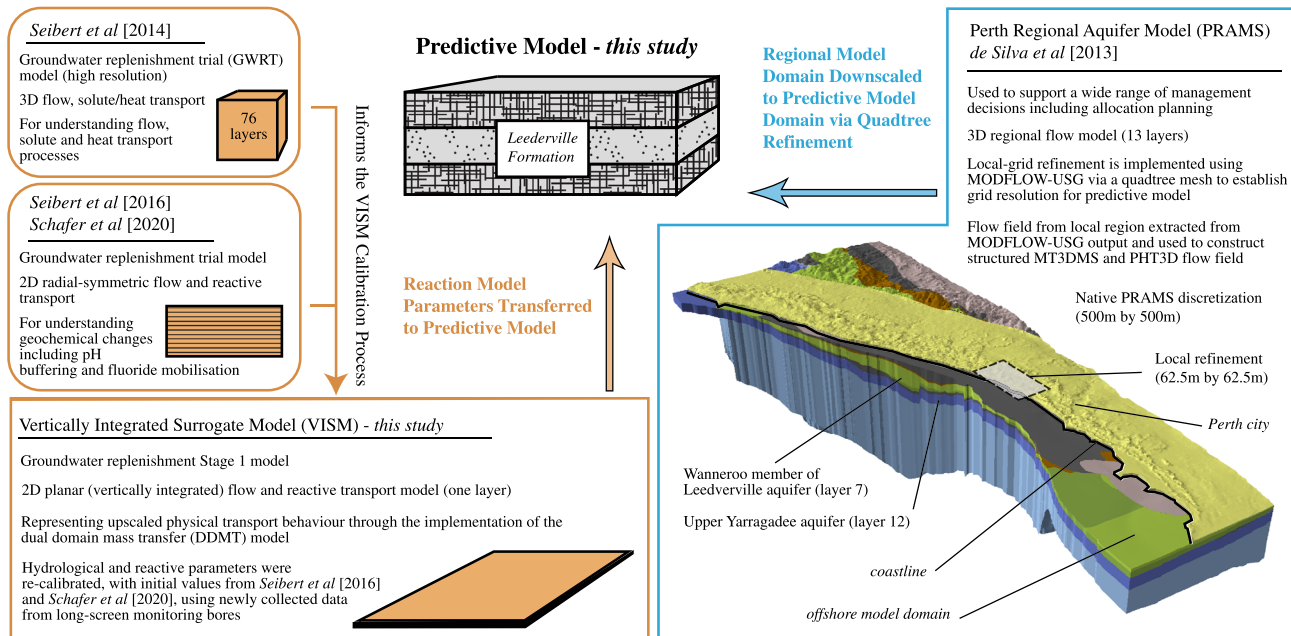
Figure 2. A schematic overview of previously existing and newly constructed models and how they were used in quantifying the large-scale impacts of GWR. The right part of the figure also provides an overview of PRAMS, which displays the model lithology and is reproduced from Siade et al. (2017).

impacted by hydrogeological features such as faults/barriers and groundwater abstractions, (ii) the assumed continuity of the vertical layering that was valid at the GWRT scale was unlikely to hold at the spatial scale of interest for long-term predictions, and (iii) the high vertical resolution of the GWRT model would have caused excessive model runtimes. Therefore, the prediction of long-term groundwater quality changes required the development of a suitable upscaling approach. Unfortunately, upscaling of flow and solute transport processes has been, and still is, a substantial challenge in hydrological sciences in general and an even bigger problem for reactive transport modeling studies. Despite significant research efforts, this problem has not been resolved satisfactorily (Sanchez-Vila & Fernàndez-Garcia, 2016). Where the "upscaling problem" has been addressed, it was generally tackled for synthetic problems (e.g., Hao et al., 2019; Jung & Navarre-Sitchler, 2018; Li et al., 2011), whereas practical applications to real-world problems are extremely limited, and therefore, there is a clear lack of workflows that provide pragmatic solutions for predicting large-scale groundwater quality evolution. Consequently, to assess the long-term impacts of GWR on the Perth aquifer system, we present a new and computationally feasible methodology, which

1. integrates the entirety of currently available quantitative hydrogeological and geochemical information that is available for the targeted groundwater system;
2. is based on more coarsely discretized and therefore computationally more efficient RTMs;
3. relies on upscaled hydrogeological parameters for these coarser models, which suitably represent the lateral and vertical solute spreading characteristics at the finer scale; and
4. relies on upscaled reaction parameters for these coarser models, which suitably represent the geochemical processes/rates that are triggered by GWR.

Figure 2 provides a schematic overview of previously existing and newly constructed models and how they were used in quantifying the long-term, large-scale impacts of GWR. The key steps were

1. the development and evaluation of a local-scale, vertically integrated surrogate RTM (VISM) and
2. the development of the injectant plume-scale RTM that integrates the downscaled groundwater flow information from the regional model with the geochemical reaction conceptualization and parameterization derived during VISM development.

These two key steps are described in more detail in the following sections.

3.2. VISM

3.2.1. Modeling Approach

Aquifer characterization data, groundwater monitoring, and the model-based interpretation that was performed for the GWRT were used to conceptually understand the significant impact of the (vertical) hydrogeological and geochemical heterogeneity within the Leederville aquifer. This provided the basis for developing a detailed local-scale modeling framework for the spreading behavior of the conservative and reactive constituents in the injectant within the Leederville aquifer, as discussed in section 2.3. The two main characteristics of the previously identified injectant transport behavior were (i) the overwhelmingly lateral solute migration that occurs at varying rates within the sandy aquifer layers, thus, over the thickness of the Leederville aquifer, causing a dispersing injectant front, and (ii) the presence of a high fraction of clayey or silty aquifer layers that effectively are inert to solute transmission. To represent the observed and simulated transport and reaction characteristics at the scale of the long-term GWR predictions, a vertically integrated and also laterally more coarsely discretized model was constructed and evaluated against new observation data in two steps (Figure 2):

1. The dual-domain mass transfer (DDMT) model (Feehley et al., 2000; Li et al., 1994; Zhang & Brusseau, 1999) was invoked to preserve the solute transport characteristics in the newly constructed VISM. Herein the mobile domain of the DDMT model is thought to represent the lateral transport that preferentially occurs in the sandy aquifer layers, while the immobile domain represents the silty and clayey aquifer layers. This method of upscaling solute transport model parameters using DDMT is consistent with the theoretical development and recent advances in the field (Fernàndez-Garcia et al., 2009; Guo et al., 2020; Li et al., 2011).
2. The conceptual and numerical geochemical reaction network and its earlier parameterization by Seibert et al. (2016) and Schafer et al. (2020) was evaluated against new field observations during GWR Stage 1. This step involved the (re)estimation of the hydrogeological and reaction model parameters of the VISM against data collected from long-screen monitoring bores. The samples from these long-screen bores represent vertically integrated and approximately flux-weighted groundwater compositions, that is, concentrations that are conceptually comparable to those simulated by the VISM. The parameters employed in the calibrated GWRT model were used as initial estimates.

3.2.2. VISM Domain and Boundary Conditions

VISM was constructed as a single confined layer (2-D plan view) model, which covered an area of 12.5 km^2 around the injection bores *LRB1–LRB3* with a gradually refined grid from the edge of the model to the injection bores (VISM, Table 2 and Figure 1). The outer edge of the VISM was >1 km from any injection bore. Similar to the GWRT models, the vertical extent of the VISM spanned across the Leederville aquifer at the Beenyup recharge site. All lateral model boundaries were defined as constant-head boundaries. Groundwater flow in the model domain was considered to be governed primarily by the injection fluxes, given that (i) at the Beenyup recharge site, the hydraulic gradient of the Leederville aquifer is small (~0.0006), and (ii) the closest bore accessing the Leederville aquifer for groundwater abstraction is >2 km away from the model boundaries. The VISM was run in transient mode for a total of 3,000 days, starting from November 2010 when GWRT commenced (Day 0) until January 2019 coinciding with the latest available monitoring data from GWR Stage 1 (Day 3,000). GWR Stage 1 commenced in August 2017, corresponding to Day 2,462. The logged injection rates for *LRB1–LRB3* were discretized into daily stress periods to capture the generally highly variable flow conditions.

3.2.3. VISM Reaction Network

Previous mineralogical analysis of the Leederville Formation indicated that, while the different main lithologies (sand, silt, and clay) have undergone different physical processing, the mineralogical composition of the silts and their geochemical response during laboratory incubation tests is very similar to the corresponding properties of sands (Descourvieres, Prommer, et al., 2010). Therefore, the mineral composition was set the same in the mobile and immobile domains. The network of water-sediment interactions defined in the VISM was based on the process understanding previously developed from the model-based interpretation of the GWRT. The defined reaction network consists of a mixture of equilibrium-based and kinetically controlled reactions. Here we provide a summary of the key processes that were found to be induced by GWR, that is, (i) acidity generation, (ii) pH buffering, and (iii) fluoride mobilization processes, while the

details of the original model development are available in Seibert et al. (2016) and Schafer et al. (2020). The details of the numerical implementation that was employed in this study are provided in Text S1 and Table S1.

Briefly, acidity generation during GWR at the Beenyup recharge site occurs mostly through the oxidation of pyrite (FeS_2), primarily by oxygen that is contained in the injectant and potentially also by nitrate. The pyrite oxidation reactions were included in the VISM based on a previously developed rate law (Eckert & Appelo, 2002; Prommer et al., 2018):

$$r_{pyr} = C_{H^+}^{-0.11} \left(C_{O_2}^{0.5} + f_2 C_{NO_3^-}^{0.5} \right) \left(10^{-10.19} \frac{SA_{pyr}}{V} \right) \left(\frac{C}{C_0} \right)_{pyr}^{0.67} \frac{f(T)}{f(T_{ref})} \tag{1}$$

where C_{H^+}, C_{O_2}, and $C_{NO_3^-}$ are the dissolved concentrations of proton, oxygen, and nitrate, respectively. f_2 represents the relative significance of nitrate (compared to dissolved oxygen) and was set to 0.005 based on previous results (Seibert et al., 2016). SA_{pyr}/V represents the initial ratio of pyrite reactive surface area to solution volume, 185 dm^{-1} mol^{-1}. $(C/C_0)^{0.67}$ considers the decrease in mineral concentration (C) relative to the initial concentration (C_0) with the progressing reaction, which is proportional to the decline in surface area assuming spherical particles. T refers to the groundwater temperature in °C, and Tref is a reference temperature at 25°C. $f(T)$ accounts for the temperature dependency, which was included following Prommer and Stuyfzand (2005):

$$f(T) = e^{-\left(\frac{1}{T + 273.15} a_1 + a_2 \right)} \tag{2}$$

where a_1 and a_2 are empirical constants that mimic the mechanism of the Arrhenius equation, which were −6758.1 and 16.1 K, respectively.

The two major processes that were assumed to compete with pyrite oxidation for the oxygen and nitrate were the oxidation of sediment organic matter and of ferrous iron that was present in the ambient groundwater as well as on cation exchange sites, while also being generated from dissolution of iron(II)-bearing minerals. The oxidation of sediment organic matter was represented in the VISM by a Monod-type reaction rate expression (Descourvieres, Hartog, et al., 2010; Parkhurst & Appelo, 1999), and a stoichiometric composition of $CH_2O(NH_4)_{0.1}H_{0.3}(PO_4)_{0.04}$ was assumed for sediment organic matter (Eiche, 2009; Wallis et al., 2020). The oxidation of ferrous to ferric iron was simulated following a previously established rate expression (Eckert & Appelo, 2002; Sun et al., 2018). The iron(III) produced from the reaction was allowed to precipitate as $Fe(OH)_3$, which was modeled as an equilibrium reaction (Prommer et al., 2018; Sun et al., 2018).

Seibert et al. (2016) suggested that in response to acidity generation, proton buffering was the major pH buffering process during the GWRT. Proton buffering was included in the VISM as one of the cation exchange reactions on an exchanger site (X):

$$H^+ + X^- = HX \tag{3}$$

The composition of the exchanger was calculated by assuming equilibrium with the surrounding groundwater. The employed thermodynamic constant (log K) for the proton exchange reaction (Reaction 3) was 5.08, which was adopted originally from Griffioen (2003) and refined in Seibert et al. (2016). Previous studies showed that the dissolution of aluminosilicates, particularly glauconite, and siderite ($FeCO_3$) might also contribute to native pH buffering capacity (Descourvieres, Prommer, et al., 2010; Seibert et al., 2016). The rate expressions for glauconite dissolution as proposed by Fernandez-Bastero et al. (2008) and for siderite dissolution as proposed by Duckworth and Martin (2004) were therefore implemented in the VISM. $Ca_{0.02}K_{0.85}Mg_{1.01}Fe^{II}_{0.05}Fe^{III}_{1.03}Al_{0.32}Si_{3.735}O_{10}(OH)_2$ was previously determined to be the formula for glauconite in the Leederville aquifer (Seibert et al., 2016) and adapted for this study.

Fluoride release was observed during the GWRT and Stage 1, which was accompanied by phosphate release. Carbonate-rich fluorapatite was previously identified to be the primary source of fluoride and phosphate (Schafer et al., 2018; Schafer et al., 2020). The composition of the carbonate-rich fluorapatite was previously determined to be $Ca_{9.75}Na_{0.25}(PO_4)_{5.37}(CO_3,F)_{0.55}F_{1.82}(OH)_{0.18}$ (Schafer et al., 2018). Fluorapatite

Table 3
Key Parameters Employed in the Calibrated VSIM Model, as Estimated by PEST++

Parameter	Estimated parameter value	Prior standard deviation	Posterior standard deviation	% variance reduction	Coefficient of variation	Composite scaled sensitivity
Physical transport parameters						
Longitudinal dispersivity (m)	$4.78E + 00$	$2.25E + 00$	$8.55E - 01$	62.0%	$1.79E - 01$	$3.68E - 06$
Mobile porosity	$9.67E - 02$	$5.00E - 02$	$2.76E - 03$	94.5%	$2.85E - 02$	$1.89E - 05$
Immobile porosity	$3.49E - 01$	$7.88E - 02$	$7.87E - 02$	0.02%	$2.26E - 01$	$2.27E - 08$
Mass transfer rate (day^{-1})	$2.26E - 06$	$2.48E - 05$	$2.71E - 07$	98.9%	$1.20E - 01$	$1.28E - 06$
Ambient sediment concentrations[a]						
Pyrite (mol L^{-1})	$4.98E - 01$	$1.80E - 01$	$6.77E - 02$	62.4%	$1.36E - 01$	$3.21E - 06$
Siderite (mol L^{-1})	$2.00E - 03$	$5.00E - 03$	$1.05E - 03$	79.0%	$5.26E - 01$	$5.81E - 07$
Glauconite (mol L^{-1})	$8.73E - 05$	$4.98E - 04$	$1.09E - 06$	99.8%	$1.25E - 02$	$2.49E - 05$
Cation exchange sites (mol L^{-1})	$1.99E - 02$	$1.13E - 02$	$9.74E - 04$	91.3%	$4.89E - 02$	$7.82E - 06$
Carbonate-rich fluorapatite (mol L^{-1})	$4.99E - 06$	$1.48E - 05$	$2.09E - 06$	85.8%	$4.19E - 01$	$2.21E - 06$
Hydrated dicalcium phosphate (mol L^{-1})	$3.36E - 06$	$7.25E - 06$	$1.81E - 07$	97.5%	$5.38E - 02$	$8.83E - 06$
Kinetic rate coefficients						
Sediment organic matter degradation k_{O_2} (mol L^{-1} s^{-1})	$5.14E - 10$	$5.99E - 09$	$9.02E - 10$	84.9%	$1.76E + 00$	$2.82E - 07$
Sediment organic matter degradation k_{NO_3} (mol L^{-1} s^{-1})	$8.94E - 11$	$3.00E - 09$	$1.13E - 11$	99.6%	$1.27E - 01$	$3.49E - 06$
Thermodynamic constants (log *K*s)						
Ammonium exchange	$7.20E - 01$	$7.50E - 02$	$3.52E - 02$	53.1%	$4.88E - 02$	$2.06E - 06$
Carbonate-rich fluorapatite dissolution	$1.26E + 00$	$7.50E - 01$	$4.35E - 01$	42.0%	$3.45E - 01$	$2.65E - 06$

Note. The "prior standard deviation" is calculated from the upper and lower bounds of a parameter during automatic calibration, which are selected based on literature data, sample characterization, and expert knowledge. The "posterior standard deviation" is the linearized estimate of the standard deviation of a parameter after the calibration process has been conducted (assuming the posterior is normally distributed). Ideally, a reduction is expected between the prior and posterior estimates. A small reduction indicates either the model-simulated equivalents of the observations are not sensitive to the parameter or the parameter is correlated with other parameters.
[a]Sediment concentrations are in the unit of mol L^{-1} of bulk aquifer volume.

dissolution often involves an initial rapid exchange reaction whereby calcium is exchanged by protons on the mineral surface and triggers a preferential release of calcium and fluoride over phosphate (Chaïrat, Oelkers, et al., 2007; Chaïrat, Schott, et al., 2007). The incongruent dissolution of carbonate-rich fluorapatite was included in the VISM as an equilibrium exchange reaction. This produces a surface layer with a composition equivalent to hydrated dicalcium phosphate ($\equiv Ca_{9.75}Na_{0.25}(PO_4)_{5.37} \cdot nH_2O$) (Chaïrat, Oelkers, et al., 2007; Chaïrat, Schott, et al., 2007), which affects the fate of calcium and phosphate and regulates further dissolution of the carbonate-rich fluorapatite. The hydrated dicalcium phosphate surface was modeled as an equilibrium mineral phase.

3.2.4. VISM Calibration

The model parameters were initially adopted from literature, where available, and then reestimated based on observations from long-screen monitoring bores *LMB1–LMB3* (Table 3). The adjustable (physical) flow and conservative transport parameters were initially adopted from Seibert et al. (2014) and Gerber (2016), which included

1. hydrodynamic longitudinal and transverse dispersivity;
2. mobile and immobile porosity—the partition between the mobile and immobile domains was determined by the proportion of the aquifer thickness with low hydraulic conductivity representing the immobile domain (~75%). Taking an estimation of ~0.4 as the total porosity, the mobile and immobile porosity was initially set to be 0.1 and 0.3, respectively; and
3. mass transfer rate coefficient.

The adjustable parameters controlling the geochemical reactions were initially adopted from Seibert et al. (2016) and Schafer et al. (2020), which included

1. reaction rate parameters controlling the rate of sediment organic matter oxidation—parameters were varied within the previously identified range;

2. reaction rate parameters for pyrite, siderite, and glauconite—the calibration allowed for adjustment of their initial concentrations within the range of plausible values; and

3. parameters affecting the cation exchange process, including cation exchange capacity, proton buffering, and the rapid incongruent dissolution of carbonate-rich fluorapatite—the calibration allowed for adjustment of the cation exchange capacity and of the log Ks for ammonium exchange and for carbonate-rich fluorapatite dissolution.

The parameters were first adjusted through a manual trial-and-error calibration. The resulting parameter estimates were subsequently used as initial values for the Gauss-Levenberg-Marquardt method contained in PEST++ for automatic calibration refinement and sensitivity analysis (Welter et al., 2015) (Table 3). The sum of squared residuals between measurements and their associated model-simulated results was used as the objective function and minimized during the calibration procedure. The observations used to constrain the calibration consisted of chloride, pH, sodium, potassium, calcium, magnesium, iron, bicarbonate, silicate, phosphate, sulfate, fluoride, nitrate, and ammonium concentrations. The procedure of observation weight assignment during automatic calibration was adopted from Sun et al. (2018), and based on the magnitude of each observation, the number of observations available, and the uncertainty inherent in each observation type. The resulting inverse problem is well determined as most parameters exhibit a significant reduction in variance upon estimation (Table 3). The only parameter exhibiting a small variance reduction is immobile porosity; however, this parameter has a very small sensitivity. The thermodynamic constants also have a somewhat small variance reduction, which may be due to how their prior ranges were chosen; however, their posterior standard deviation is relatively small compared to their estimated value.

3.2.5. Observed and VISM-Simulated Local-Scale Geochemical Evolution

3.2.5.1. Overview

The simulations with the calibrated VISM provide a reasonable representation of all the measured key groundwater constituents at the long-screen monitoring bores *LMB1–LMB3* (Figures 3 and S2). Although there was evidence for some weak lateral geochemical/mineralogical heterogeneity, which was not considered in the VISM, the simulations successfully captured the main characteristics of the hydrogeochemical observations. It should be noted that injection at *LRB1* as part of the GWRT altered the ambient condition for *LMB1* for the start of GWR Stage 1 injection and that *LMB1* was installed post-GWRT; hence, there are no data from long-screened monitoring bore to compare with for the initial ~1,000 days of the simulation period. The good agreement between the simulations and the observations indicates that the previously identified reaction network also remains valid to represent the coupled and now vertically integrated flow, transport, and reaction processes during large-scale GWR. The simulations show that the behavior of many monitored groundwater constituents such as chloride and heavy metals was predominantly governed by physical flow and transport processes, and therefore, these constituents displayed similar concentrations to those in the treated recycled water soon after injectant breakthrough. In contrast, the spatiotemporal evolution of some other species such as sulfate, pH, and fluoride were reaction driven. To focus on the validation of the VISM, the following discussion is centered on the analysis of the latter class of species and on the processes that cause these changes.

3.2.5.2. Pyrite Oxidation and pH Buffering Processes

Consistent with the results of the GWRT and the high-resolution model simulations, the observed and simulated breakthrough behavior for the long-screen monitoring bores *LMB1–LMB3* show that pyrite oxidation is the most prominent water-sediment interaction to occur during large-scale GWR. Even after injectant breakthrough, which is indicated by the sharp decline in chloride concentrations, groundwater sulfate concentrations remained at ~0.1 mM, substantially higher than in the injectant, where sulfate concentration was below the detection limit of ~0.001 mM (Figure 3). Pyrite oxidation could have resulted in a significant release of free protons (H^+) and therefore an acidification of the groundwater as illustrated by the simulations from a model variant (mv1, Table S2) in which all buffering processes remained deactivated (Figure 4). Nevertheless, pH measurements showed that circumneutral pH conditions were maintained in the recharged Leederville aquifer during GWR (Figure 3). In fact, upon injectant breakthrough, groundwater pH increased temporarily above the ambient pH before successively declining back to the ambient value. Previously, Seibert et al. (2016) suggested that proton buffering was the primary process to buffer the acidity generated by pyrite oxidation, which was considered in the VISM through a proton exchange reaction on cation exchange sites (mv2, Table 2). Solely considering this proton exchange reaction was
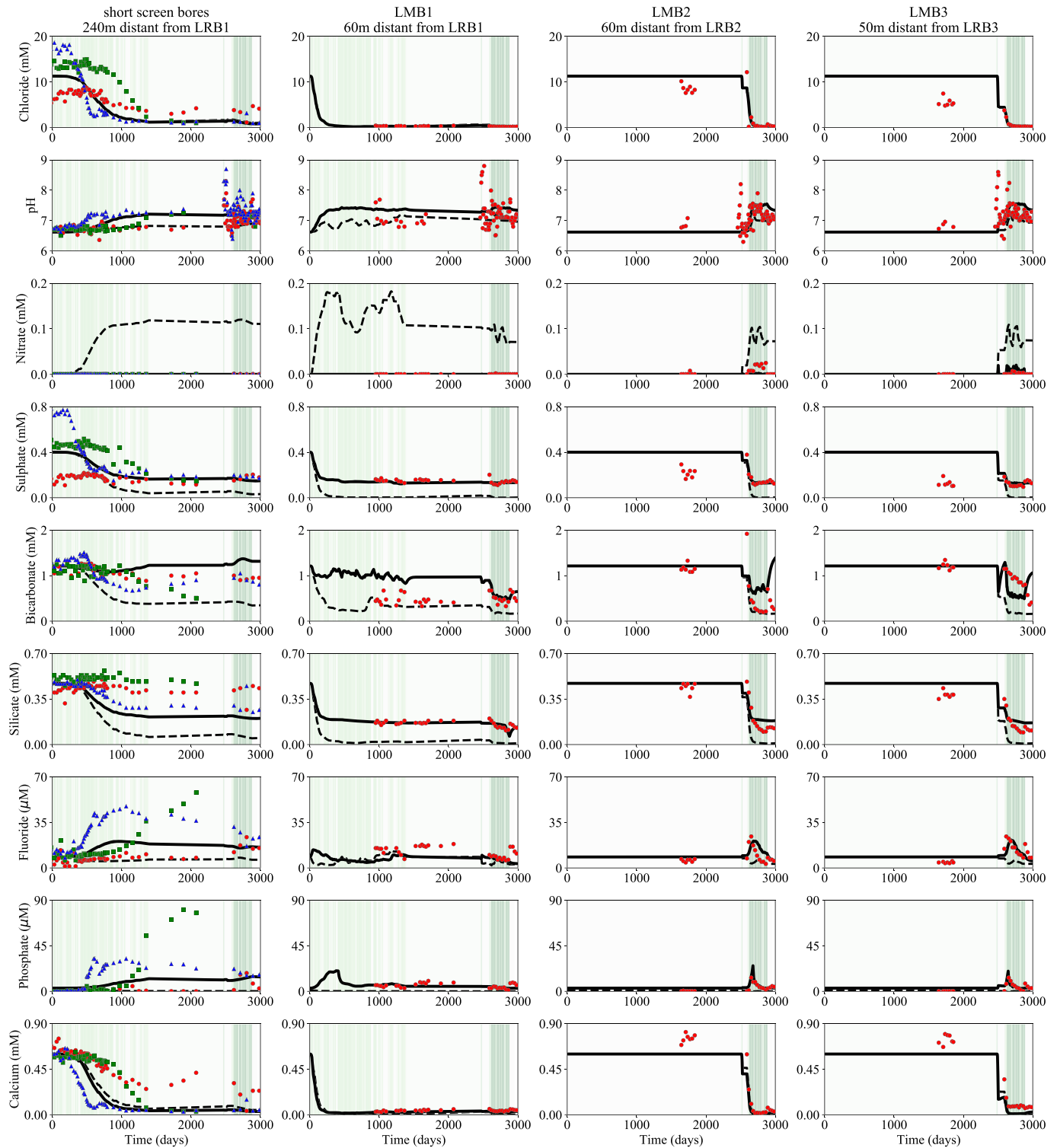
Figure 3. Observed (symbols) and simulated groundwater chloride, pH, nitrate, sulfate, bicarbonate, silicate, fluoride, phosphate, and calcium during GWRT and Stage 1. The results of iron, sodium, potassium, magnesium, and ammonium are shown in Figure S2. Time 0 represents November 2010 when the GWRT started. Injections are marked in each subplot in green background with more intense color representing higher injection rate. Reactive and conservative transport simulations are plotted as solid and dashed black lines, respectively. *LMB1–LMB3* are long-screen monitoring bores, which are ~60, 60, and 50 m distant from injection bores *LRB1–LRB3*, respectively. On short-screen bore subplots, red "o," green "□," and blue "Δ" symbols represent observations from *BY151*, *BY162*, and *BY201* bores, respectively, all of which are 240 m distant from injection bore *LRB1*.
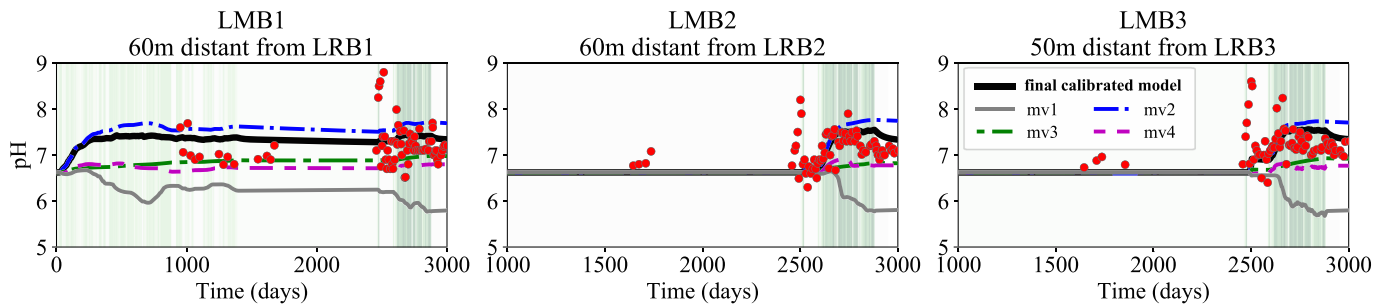
Figure 4. Simulated pH from final calibrated model and model variants mv1—no buffering reaction (no proton exchange, glauconite, or siderite dissolution), mv2—proton exchange alone (no glauconite or siderite dissolution), mv3—ambient Na concentration in the injectant, and mv4—iderite alone (no proton exchange or glauconite dissolution). Note the shorter timescales for *LMB2* and *LMB3* to show the details of the GWR Stage 1 injection.

found to yield a reasonable agreement between observed and simulated pH values, including the observed initial pH rise (Figure 4). This rise in pH is conceptually similar to observations from coastal aquifers (e.g., Appelo, 1994), where a freshening of brackish or salty aquifers and, associate with it, an elution of cations from cation exchange complexes occurs—here induced by the highly purified, dilute nature of the injectant that is employed by GWR (Table 1). Due to the higher selectivity of the sediment's exchange sites for protons (log K is 5.08) over many other cations (e.g., log K for divalent calcium is only 0.8), the dilution of porewater upon injectant breakthrough promotes a relative decrease in proton concentration in groundwater and therefore a pH rise. This phenomenon can be illustrated by the comparison with a separate model variant (mv3, Table S2) in which sodium concentrations in the injectant were set to the ambient sodium concentration and, accordingly, the initial pH rise vanished (Figure 4).

While the simulated proton exchange can match the pH measurements upon injectant breakthrough, the model variant that considered only this buffering reaction (mv2, Table S2) did not replicate the pH decline that was observed upon continued injection (Figure 4). Even when less exchange sites were assigned in the model, this resulted in an abrupt drop of the pH once the proton buffering capacity was exhausted, while a smooth decline was observed. This mismatch indicated the presence of at least one other simultaneously occurring buffering process. The most likely candidate for this process was thought to be the dissolution of aluminosilicate minerals such as glauconite (Descourvieres, Prommer, et al., 2010; Seibert et al., 2016). Including glauconite in the reaction network results in a stoichiometric release of the glauconite constituents ($Ca_{0.02}K_{0.85}Mg_{1.01}Fe^{II}_{0.05}Fe^{III}_{1.03}Al_{0.32}Si_{3.735}O_{10}(OH)_2$), many of which are subsequently removed from groundwater by cation exchange reactions (e.g., potassium, calcium, magnesium, and aluminum) and precipitation (e.g., iron), and therefore, concentration changes were not apparent from the monitoring data. Nevertheless, the observed increased concentrations of silicate at all monitoring locations, compared to the concentration in the injectant, are a strong indicator for the occurrence of glauconite dissolution (Figure 3). However, with the silicate-constrained low rate of glauconite dissolution, the simulated pH values were almost identical to those obtained by the proton buffering alone variant (mv2, Table S2). This suggests that glauconite dissolution, while important for explaining the observed silicate release, makes no significant contribution to the aquifer's pH buffering capacity.

The best agreement between the VISM simulations and the observed pH changes was achieved when the dissolution of siderite was also included. Seibert et al. (2016) suggested that dissolution of siderite and/or ankerite would only occur in the deeper, finer-grained sections of the Leederville aquifer. However, as the observations in this study were obtained from long-screen monitoring bores, the collected samples represent mixtures of groundwaters from different depths and possibly varying provenance of specific processes, including buffering reactions. While including siderite dissolution in the VISM can plausibly explain the pH decline that was observed at later times after complete injectant breakthrough, when simulating pH buffering solely through siderite dissolution, this model variant (mv4, Table S2) was unable to replicate the observed initial pH rise (Figure 4). Therefore, proton exchange and siderite dissolution are both necessary to explain the native pH buffering capacity.
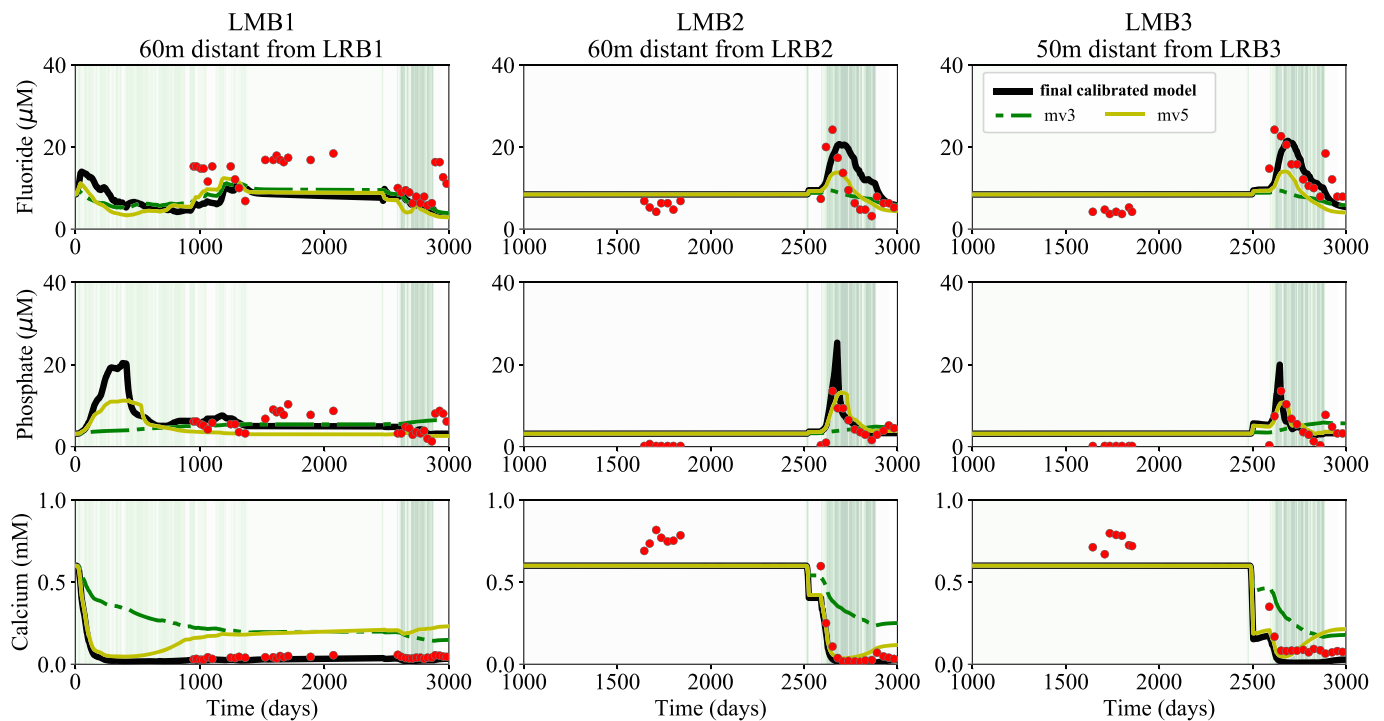
Figure 5. Simulated fluoride, phosphate, and calcium breakthrough curves from final calibrated model and model variants mv3—ambient sodium concentration in the injectant and mv5—ambient calcium concentration in the injectant. Note the shorter timescales for *LMB2* and *LMB3* to show details of GWR Stage 1 injection.

3.2.5.3. Fluoride Release and Attenuation

Consistent with the data collected during the GWRT, release of fluoride and phosphate was also observed during GWR Stage 1 (Figure 3). The occurrence of the fluoride and phosphate pulses is largely associated with the period of injectant breakthrough. Adopting the reactions proposed by Schafer et al. (2020) and including a low amount of carbonate-rich fluorapatite, the VISM could replicate the observed fluoride pulses (Figure 3). The observed phosphate pulses could also mainly be attributed to the dissolution of carbonate-rich fluorapatite, although sediment organic matter degradation also played a minor role on the phosphate release. The good agreement of the VISM simulation results with the observed data further underpins the previous hypothesis that dissolution of carbonate-rich fluorapatite caused the release of fluoride and phosphate in the Leederville aquifer during GWR.

Due to the highly purified nature of the injectant and the various associated cation exchange processes, the GWR caused particularly low groundwater calcium concentrations and thus perturbed the carbonate-rich fluorapatite mineral-groundwater equilibrium in the native aquifer (Chaïrat, Schott, et al., 2007; Schafer et al., 2018). The effect of cation exchange reactions on dissolution of carbonate-rich fluorapatite is demonstrated by two model variants in which cation concentrations in the injectant were explicitly changed (Figure 5). The model variant in which calcium concentrations were increased to ambient concentration level (mv5, Table S2) still shows fluoride and phosphate release to occur upon injectant breakthrough, as the initial breakthrough calcium concentration remains low due to the dynamic nature of the exchange reactions and the higher selectivity of the exchange sites for the divalent cation calcium (Appelo, 1994). In contrast, the model variant in which sodium concentrations were increased to ambient concentration level (mv3, Table S2) showed the initially high calcium concentrations during breakthrough to persist while fluoride and phosphate concentrations remained low.

3.3. Downscaling of PRAMS Groundwater Flow Model

Flow model simulations with PRAMS were performed to establish the flow conditions under consideration of the envisaged injection rates of full-scale GWR (i.e., GWR Stage 2), which were applied for a period of
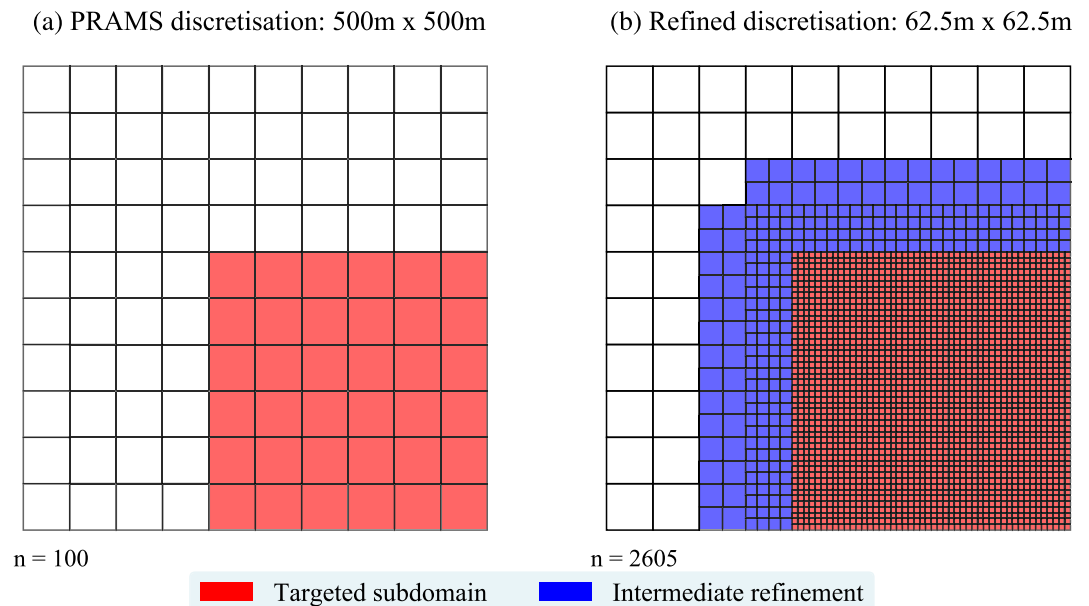
Figure 6. Illustration of the quadtree refinement procedure used on the upper left corner of a subdomain. The level of refinement is 8 on each side; therefore, cell size is reduced from 500 × 500 m to 62.5 × 62.5 m. The subdomain obtained is a structured grid, while the transition to the coarser initial grid is unstructured.

30 years and discretized into monthly stress periods. For the long-term predictions the impact of the GWRT and the short period for which only GWR Stage 1 operated were neglected. This is a reasonable simplification that does not significantly affect the long-term (30-year) predictions. Consistent with the planned GWR operations, a combined 14 GL year^{-1} was injected at the new northern and southern recharge sites, in addition to the 14 GL year^{-1} injected at the Beenyup recharge site. The injection bores in the Leederville aquifer were added to Layer 7 of the PRAMS model, while the bores in the Yarragadee aquifer were added to Layer 12 (Figure 2). Furthermore, a reduced groundwater recharge rate was considered in the flow model to incorporate the impacts of projected climate change in 30 years, and the groundwater abstraction rates used in the model also reflected estimates of future abstractions, as detailed in Water Corporation (2016). Based on the results of preliminary particle tracking simulations (data not shown) with the PRAMS model for the 30-year predictive period, a suitable subdomain size was selected, covering a model extent of 21.5 km in north-south direction and 15 km in east-west direction (Figures 1 and 2). The injection bores are approximately located at the center of the model domain.

While the GWRT-scale RTM employed a spatial discretization that was too fine to allow for large-scale model predictions, the grid resolution of PRAMS (500 × 500 m) is too coarse for conservative and reactive transport simulations at the spatial scale that is of interest for assessing the long-term geochemical impact of GWR. This concerns especially the geochemical changes between the injection bores and the boundaries of the recharge management zones (250-m radial distance), where a sufficiently high grid resolution is required to assess the evolution of the geochemical gradients and to provide robust support for any potential management decisions to be based on the modeling results. The regional flow model therefore needed to be locally refined to accommodate appropriately spatially discretized transport simulations. To facilitate local grid refinement within the region of interest, PRAMS was converted from a standard MODFLOW model into a MODFLOW-USG model to allow for the use of an unstructured model grid (Panday et al., 2013). To perform refinement of the subdomain within the MODFLOW-USG model, the "quadtree principle" method was recursively applied, which divides the "parent" cell into four equal cells first, and the neighboring four cells can then be further refined until the desired resolution is reached at the local study area (Figure 6). The applied refinement resulted in grid cell sizes of 62.5 × 62.5 m for the transport subdomain, while all other features remained consistent with PRAMS v3.5.2 (CyMod Systems Pty Ltd, 2014). Extraction of the necessary flux information for the subsequent transport simulations was automated using Python scripts in the subdomain from the results of the MODFLOW-USG simulations.

4. Large-Scale, Long-Term Impacts of Groundwater Replenishment

4.1. Conservative and Reactive Transport Simulations

Based on the flow field within the selected subdomain that was computed with MODFLOW-USG, conservative transport simulations were performed with MT3DMS (Zheng & Wang, 1999) to (i) assess the potential for any vertical leakage to occur toward overlying aquifers, specifically toward the superficial aquifer and (ii) determine the spatiotemporally evolving extent of injectant plume spreading and dilution within the Leederville aquifer (GWR-Predict-NR, Table 2). Consistent with the VISM, DDMT was invoked to consider the impact of some solute mass transfer between the "mobile" conductive aquifer zones formed by the sandy lithological units and the comparably "immobile" zones formed by the silty and clayey units. The transport properties of the Leederville aquifer were assigned based on the parameters that were estimated for VISM (Table 3).

Reactive transport simulations were then conducted using PHT3D (Prommer et al., 2003), to forecast and illustrate the future GWR-induced groundwater quality evolution in the Leederville aquifer (GWR-Predict-R, Table 2). Given the generally limited spreading of the injectant in vertical direction (as discussed in section 4.2), sufficiently accurate reactive transport simulations for the Leederville were performed with a three-layer model (PRAMS Layers 6–8) instead of the original 13-layer model to significantly reduce the computational effort that was required for the reactive transport simulations (Figure S3). The same reaction network and parameters that were employed in the VISM were used for the RTM as it was not expected that upscaling effects between the scale of the VISM and the large-scale model would have a significant impact on these reaction parameters (Table 3). The assumed ambient groundwater and sediment compositions were also adopted from those employed in the VISM. Based on limited information from analyses on sediment samples collected at the Beenyup site (Descourvieres, 2011; Descourvieres, Prommer, et al., 2010), the spatial variances in sediment mineralogy in the Leederville Formation are not significant. Therefore, similar to the VISM, the sediment composition was set homogeneous in the large-scale model. The composition of the injectant was assumed to remain constant over the 30-year simulation period and corresponds to the average water composition that was measured during GWR Stage 1 (Table 1).

4.2. Injectant Spreading Behavior

The simulated spreading of the injectant after a 30-year simulation time in both the vertical and lateral directions shows three distinct plumes to form in both the Leederville and the Yarragadee aquifers (Figure 7). The model simulations predict that at some isolated locations, a small fraction of the injectant could migrate upwards from the Leederville aquifer into the lowest (PRAMS Layer 3) of the three considered layers that represent the superficial aquifer in the PRAMS model. Minor leakage is predicted to occur at the Beenyup and the northern recharge sites, with the fraction of the injectant to occupy Layer 3 being estimated at ~1% after 30 years of continued injection (Figure 7). The predicted upward migration that occurs at the Beenyup recharge site is a result of the higher injection rate, as the injection into the Leederville aquifer occurs at ~30 ML day^{-1} via three closely spaced recharge bores, whereas at the other two sites, injection occurs only at ~10 ML day^{-1} via one recharge bore. The upward migration described in the model predictions at the northern recharge site can be attributed to the decreased thickness of the confining Kardinya Shale (Layer 5), compared to the thickness found at the other sites (Figure 1).

It is evident that near the recharge sites, lateral injectant spreading is strongly influenced by the presence of the considered geological faults (Figure 7). Generally, the assumed low hydraulic conductance of these faults largely prevents the GWR-induced plumes from traveling beyond the fault line. In contrast, the faults near the northern recharge site, especially in the northerly and easterly direction from the injection bore *LRB5*, had a smaller influence on the plume propagation. This can be attributed to the much higher hydraulic conductance compared to other segments of the fault system (CyMod Systems Pty Ltd, 2014; de Silva et al., 2013). However, it should be noted that the currently simulated impacts of these faults on the solute transport behavior remain somewhat unreliable due to the uncertainties in the more detailed understanding of their characteristics. Nevertheless, where required the modeling framework presented in this study can now also be used to quantify the impacts of this uncertainty on injectant behavior. The injectant propagation in the lateral direction is also influenced by the regional groundwater flow system, with a slow plume migration
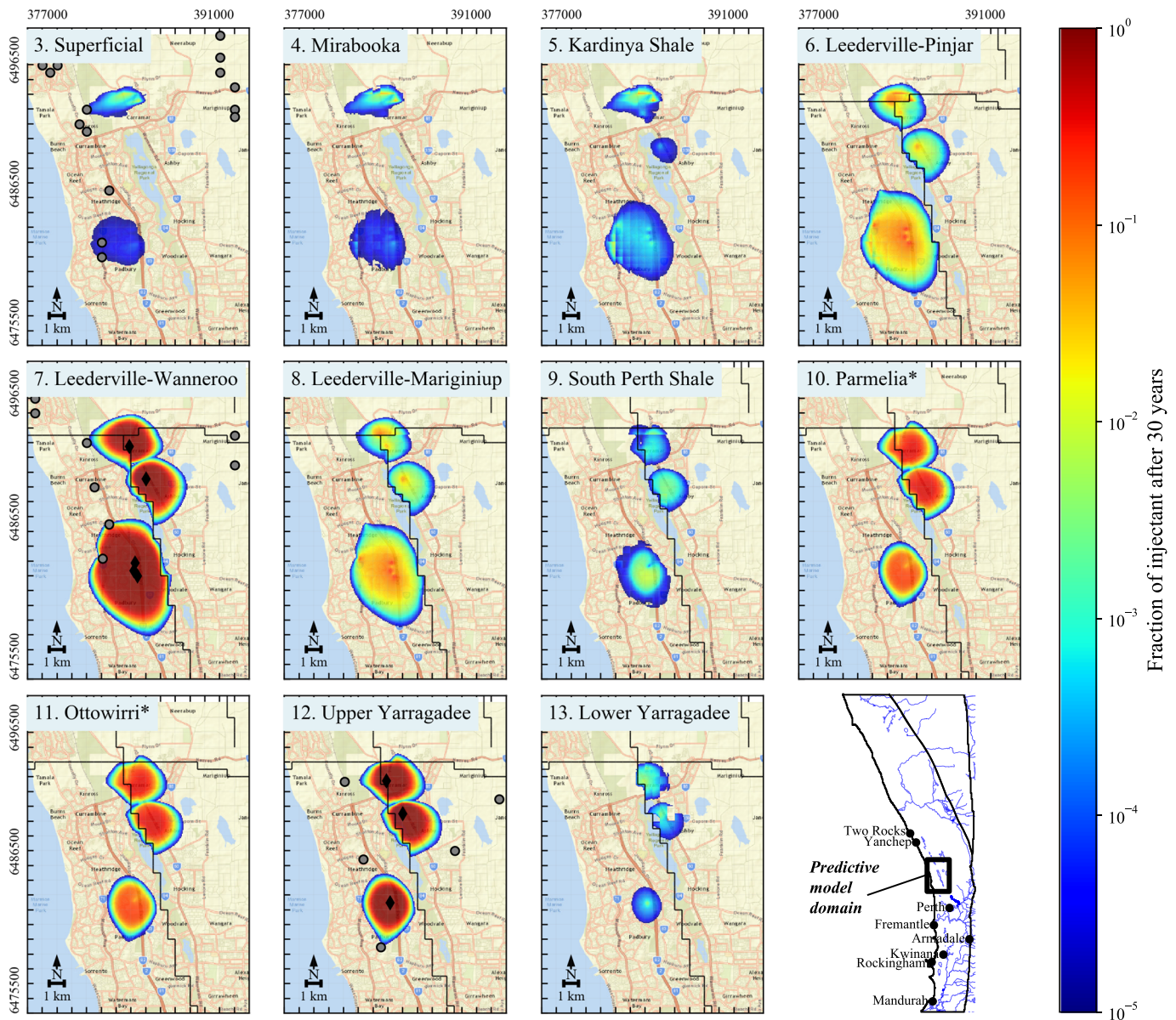
Figure 7. Lateral and vertical spreading of injectant over the model domain at the end of 30 years of GWR Stage 2. The intermediate spreading patterns are shown in Figure S4. The injection bores are depicted as black diamonds, while the abstraction bores are shown as gray circles. The black lines present in Layers 6–13 represent the geological faults. Layers 10 and 11 with the name marked with "*" are sudo 1-m-thick layers. Background map ©OpenStreetMap contributors and the GIS User Community.

occurring toward nearby abstraction bores (Layers 7 and 12, Figure 7). In the Leederville aquifer, the injectant is predicted to arrive at the Water Corporation production bores WT5/25, west of the Beenyup recharge site, between 5 and 15 years (Figure S4).

Due to the substantially lower salinity of the injectant compared to the ambient groundwater (Table 1), the predictive model results show growing zones of substantially lower chloride concentrations as the injectant spreads away from the injection bores (Figure 8). As indicated by the simulated chloride concentrations, the average arrival of the injectant plume at the boundary of the recharge management zone occurred after a relatively short time (Figure 9). Interestingly, the average arrival times differ noticeably among the different injection bores despite the injection rates being similar at ~10 ML day^{-1}. The slightly later arrival time at
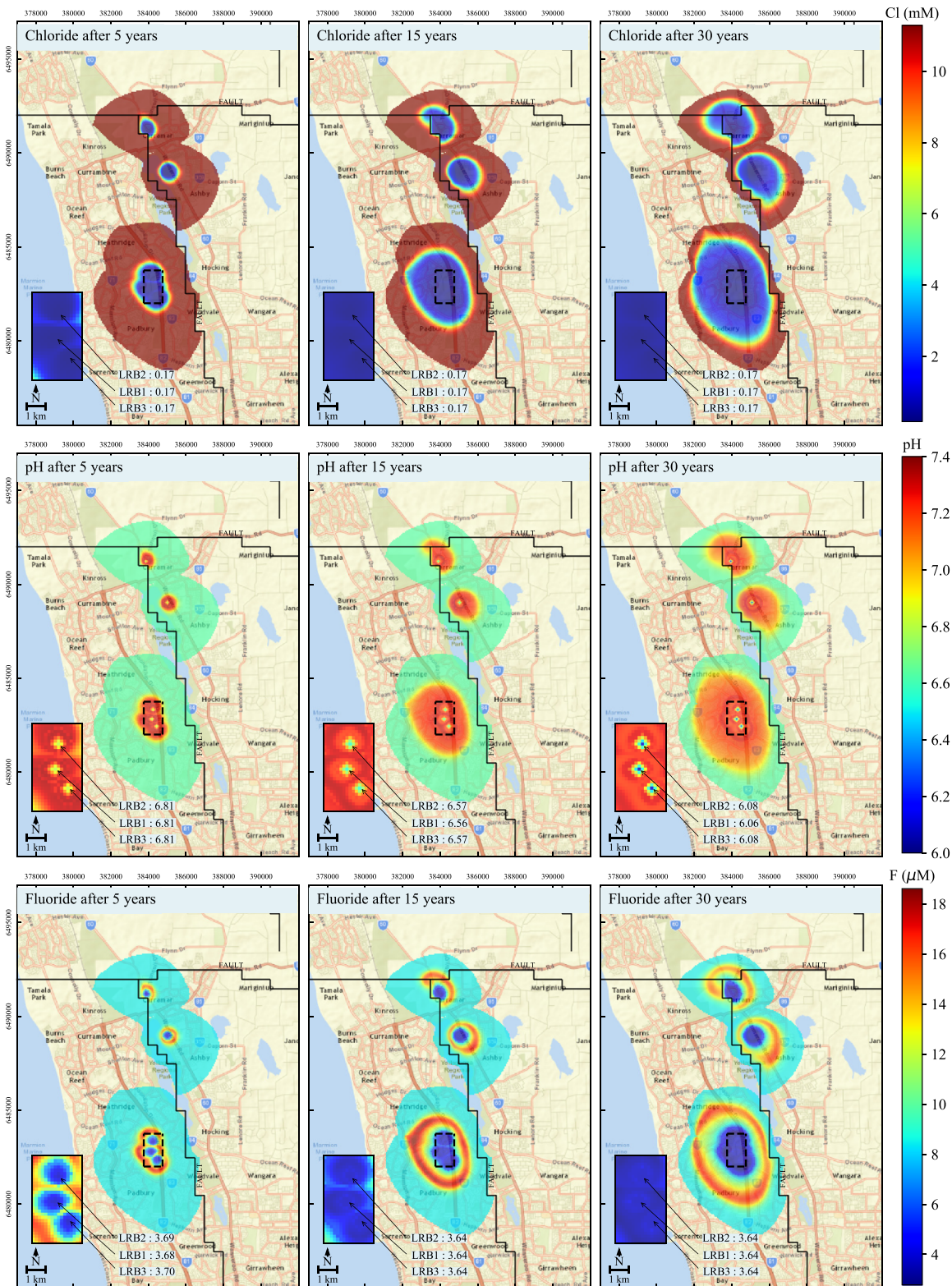
Figure 8. Predicted chloride concentrations, pH, and fluoride concentrations after 5, 15, and 30 years of GWR Stage 2. The results of other species are shown in Figure S5. Insert shows close-up views of the simulated values in the vicinity of the injection bores *LRB1–LRB3*, while the values around *LRB4* and *LRB5* are similar. The subplots only show the portion of the active model domain while not displaying the remainder of the model domain in which a fixed concentration boundary was applied, reflecting the fact that the ambient groundwater conditions would not change outside the active domain. The outlines of the active domain represent the maximum extent of lateral spreading after 30 years of GWR Stage 2. The black lines shown in each subplot represent the geological faults.
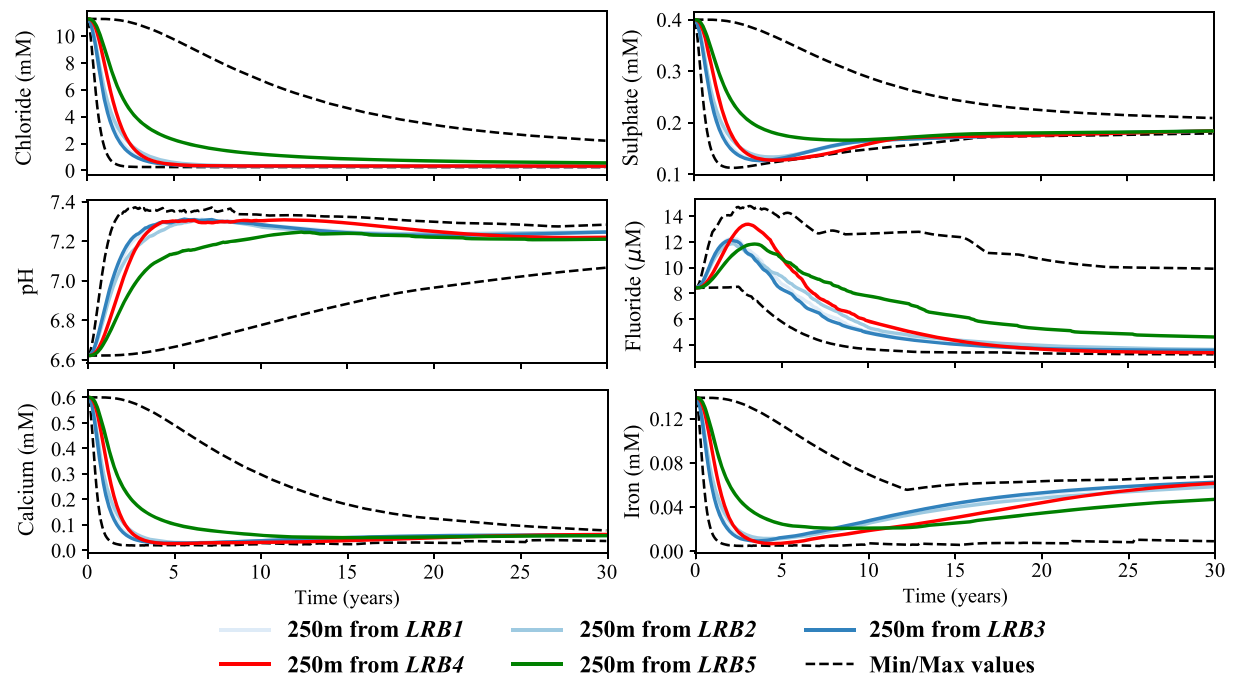
Figure 9. Predictive breakthrough curves of chloride, pH, sulfate, fluoride, calcium, and iron at the recharge management zone boundary of each injection bore in the Leederville aquifer. The values reported for individual recharge management zone are the average values at 250-m radius from the respective injection bore and shown as solid colored lines. The dashed black lines represent the minimum and maximum simulated values across all model cells at 250-m radius from *LRB1*–*LRB5*.

LRB4 relative to *LRB1*–*LRB3* is explained by the appreciably thicker Leederville aquifer at the southern recharge site relative to the Beenyup site (Figure 1). The even later average arrival time at *LRB5* is primarily a result of the presence of the low-conductance fault line at 250 m south of the injection bore.

4.3. Sustainability of pH Buffering by Native Sediments

Similar to the previously discussed models, the upscaled RTM simulations predict that groundwater sulfate concentrations after breakthrough are significantly higher than those in the injectant (Figures 9 and S5), mostly as a result of pyrite oxidation but also slightly due to some (slow) mass transfer from the "immobile" zones. The dissolution of pyrite also would lead to the release of ferrous iron. However, similar to the results from previous observations and models, this iron is subsequently oxidized and removed from groundwater by precipitation. Based on the predicted results (Figure 8), due to rapid proton exchange, an increase in groundwater pH occurs in the vicinity of the injection bores, similar to the behavior observed during the initial phases of the GWRT and GWR Stage 1. Also owing to this buffering process, the pH at the boundaries of the recharge management zone is predicted to stay above the initial groundwater pH of 6.6 over the 30-year simulation period (Figure 9). As formulated in the conceptual model, siderite and glauconite disso-lution occur in response to the acidity generated during pyrite oxidation, thereby releasing bicarbonate and silicate, respectively (Figure S5).

As the GWR operation keeps introducing exogenous oxidants (oxygen and nitrate) into the aquifer, pyrite oxidation will continue to occur due to the high abundance of pyrite in the Leederville sediments. This will cause a continuous consumption of the native buffering capacity. For the immediate vicinity of the injection bores, following the initial pH increases that were already observed, the pH is predicted to return to ~6.6, that is, the pre-GWR pH, after ~15 years of GWR, and then decreasing further to ~6.1 after 30 years (Figure 8). Due to the fact that the simulated value represents an average value over a cell size of 62.5×62.5 m, the local pH around the injection bores is expected to be even lower. The low pH fronts will successively migrate away from the injection bores, although the front velocity will be slowing down due to the radially decreasing flow velocities and due to the increasing sediment volume in contact with the injectant.

4.4. Predicted Fate of Fluoride

Consistent with the observed and simulated behavior of fluoride and phosphate during the GWRT and during GWR Stage 1, with continued injection, elevated concentrations of fluoride and phosphate are also predicted to occur at larger distances from the injection bores, with the front of the fluoride and phosphate plumes close to the injectant plume front (Figures 8 and S5). Fortunately, also similar to what was observed during the GWRT and the GWR Stage 1, peak fluoride and phosphate concentrations are not predicted to significantly increase over time or with larger travel distances. This is explained by a combination of dilution effects and the nature of the dissolution of carbonate-rich fluorapatite and (re) precipitation process. As already discussed in Schafer et al. (2020), where GWRT data were interpreted, although dissolution of carbonate-rich fluorapatite occurred to some extent due to the injection of low-calcium water, the groundwater that is being transported behind the plume front would quickly reach its new equilibrium with carbonate-rich fluorapatite locally within the aquifer. Consequently, such dissolution of carbonate-rich fluorapatite would not lead to uncontrollably high level of fluoride contamination in the recharged Leederville aquifer. While phosphate is not considered toxic, a high fluoride intake is undesired as it could cause dental and skeletal fluorosis. Therefore, the drinking water guideline for maximum fluoride concentrations is set 1.5 mg L^{-1} or 79 μM (WHO, 2017). At this stage, owing to the small amounts of carbonate-rich fluorapatite that are thought to prevail in the Leederville aquifer (as nodules), and as successfully replicated by the VISM (Table 3), the highest measured fluoride concentration is ~50 μM (0.95 mg L^{-1}) based on the extensive monitoring data from the GWRT (Schafer et al., 2018). It is predicted that fluoride concentration during long-term GWR will remain below the drinking water guideline (Figures 8 and 9).

5. Potential for Injectant Manipulation to Safeguard Water Quality

In a final step, the developed large-scale RTM was employed to illustrate how modifications of the AWT process through targeted amendments of the injectant could affect the long-term evolution of the groundwater quality. Even though the current model predictions suggest that the pH decrease resulting from GWR will remain restricted to the immediate vicinity of the injection bores over the 30-year predictive period (see inserts in Figure 8), it is prudent to investigate the possibility that acidification could proceed in a more pronounced manner. In this case, it will be important to understand whether changes of the treatment process could sufficiently mitigate such a potential groundwater quality deterioration. Two key scenarios were explored in this study (Table S3):

1. The first scenario involved the deoxygenation of the injectant to suppress pyrite oxidation, as recently proposed (e.g., Prommer et al., 2018) to prevent arsenic mobilization during (re)injection of coal seam gas coproduced water into the Precipice aquifer in Queensland, Australia. In this scenario, dissolved oxygen concentrations in the injectant were assumed to be zero, while all other water quality parameters were left similar to the current injectant composition.
2. The second scenario simulates the amendment of bicarbonate with the aim of enhancing the pH buffering capacity of the injectant (e.g., Robinson et al., 2009). In this scenario, bicarbonate concentrations were increased to 0.75 mM, which was the highest measured concentration in the treated recycled water (Seibert et al., 2016). For better comparison, the injectant pH in this scenario was maintained at 7.04.

Instead of starting with the manipulated injectant composition, it was assumed that a switch to the modified injectant composition would occur after 15 years of GWR, to consider the fact that there is currently no indication that an additional modification of the treatment process appears to be required. The simulations demonstrate that both the deoxygenation and the increased bicarbonate scenario were able to limit the magnitude of acidification in the immediate vicinity of the injection bores and attain <0.1 pH unit decrease compared to the ambient conditions (see inserts in Figure 10). This is an important result as it suggests that it is sufficient, as currently implemented, to monitor the groundwater quality at the 50/60 m bores (LMB1–LMB3) and to be able to still mitigate any acidification trends.

Given that the predicted fluoride concentrations remained below the required water quality guideline, no fluoride specific mitigation scenario was conducted as part of the present study. However, scenario simulations exploring the amendment of calcium chloride as described in Schafer et al. (2020) and ongoing
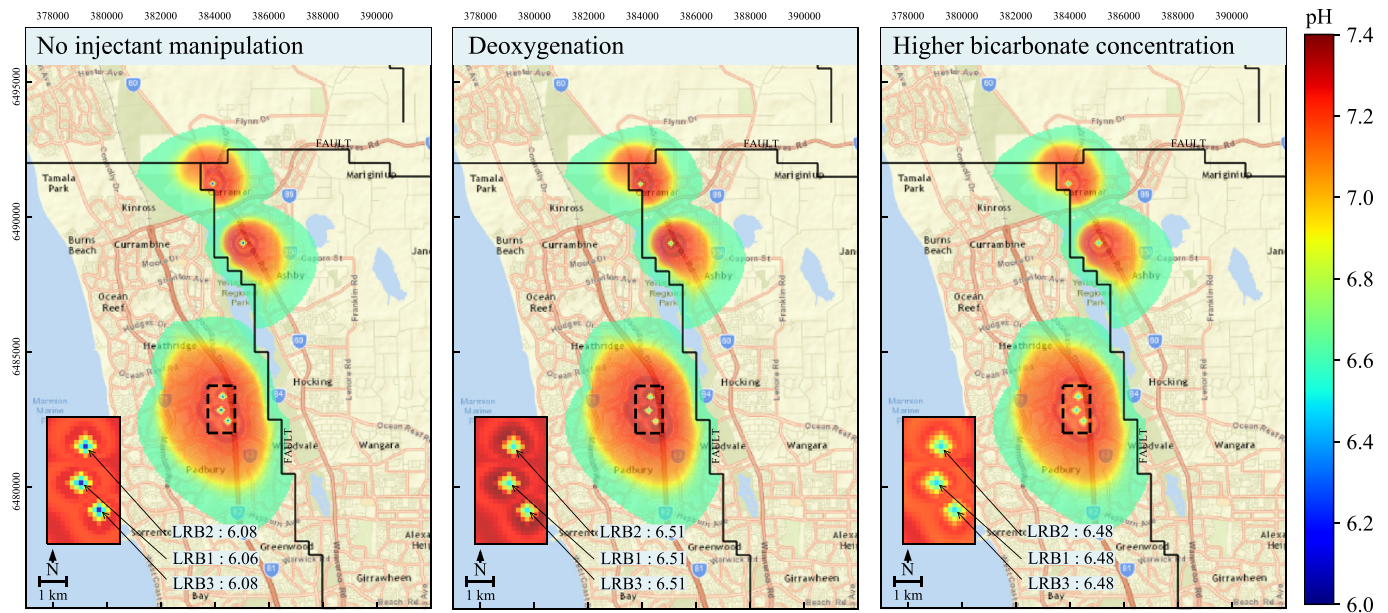
Figure 10. Predictive pH at the end of 30 years of GWR Stage 2 operation without and with injectant manipulation between 15 and 30 years, that is, (left) no injectant manipulation, (middle) deoxygenation, and (right) increased bicarbonate concentration. Insert shows close-up views of the simulated values in the vicinity of the injection bores *LRB1–LRB3*, while the values around *LRB4* and *LRB5* are similar. The subplots only show the portion of the active model domain while not displaying the remainder of the model domain in which a fixed concentration boundary was applied, reflecting the fact that the ambient groundwater conditions would not change outside the active domain. The outlines of the active domain represent the maximum extent of lateral spreading after 30 years of GWR Stage 2. The black lines shown in each subplot represent the geological faults.

modeling studies to explore an increase of sodium concentrations in the injectant to the ambient concentration level in the Leederville aquifer are both suggested as effective ways to suppress dissolution of carbonate-rich fluorapatite, if needed.

6. Conclusions

This study developed a pragmatic approach to integrate field observations with numerical models, at multiple scales, to assess and predict the solute movement and the key water-sediment interactions that are induced by the full-scale injection of aerobic, low salinity-treated recycled water into a deep, anaerobic aquifer system. New observation data collected during the initial phase of the full-scale GWR and the model-based assessment performed in this study confirmed and further refined the conceptual hydrogeological and geochemical models that were established during the interpretation of the earlier, local-scale field injection trial. A key element of this study was to develop a pragmatic upscaling procedure that allowed to perform computationally feasible large-scale reactive transport simulations for the assessment the long-term fate of GWR. An important aspect of this upscaling procedure was to invoke a DDMT approach at the larger scale to represent key transport characteristics invoked by the significant vertical aquifer heterogeneity. It should be noted that the model parameters that were employed by the upscaled model will not only be site specific but moreover depend, to some extent, even on the model discretization. Based on our new large-scale predictive simulations using this "fit-for-purpose" model framework and assuming the current injectant composition, it is anticipated that pyrite oxidation will continue to proceed in the vicinity of the injection bores and that groundwater acidification will be largely prevented by the buffering capacity of the native sediments. However, given that any significant acidification would imply an enhanced risk of metal mobilization, a regular careful analysis of the monitoring data will need to be performed to identify any "warning signs" that the aquifer behavior deviates from current predictions. Should this be the case it will be possible to mitigate the problem through a modified injectant (pre)treatment. The developed modeling framework can be used to assist with optimizing injectant treatment options and to address the associated predictive uncertainty for this AWT-MAR project if needed in the future. The developed modeling procedures will also be highly applicable to investigate other large-scale groundwater quality changes induced by MAR and a wide range of other water management options.

Data Availability Statement

All the data presented in this study are available for download from CSIRO Data Access Portal (https://doi.org/10.25919/5f98e14bbf0f7) (provided by CSIRO upon acceptance), and any additional information can be obtained by contacting the corresponding author.

References

ANZECC-ARMCANZ (2000). Australian and New Zealand guidelines for fresh and marine water quality. In *National Water Quality Management Strategy No. 4* (pp. 1–314). Canberra: Australian and New Zealand Environment and Conservation Council/Agriculture and Resource Management Council of Australia and New Zealand.

Appelo, C. (1994). Cation and proton exchange, pH variations, and carbonate reactions in a freshening aquifer. *Water Resources Research*, *30*(10), 2793–2805. https://doi.org/10.1029/94WR01048

Bekele, E., Zhang, Y., Donn, M., & McFarlane, D. (2019). Inferring groundwater dynamics in a coastal aquifer near wastewater infiltration ponds and shallow wetlands (Kwinana, Western Australia) using combined hydrochemical, isotopic and statistical approaches. *Journal of Hydrology*, *568*, 1055–1070. https://doi.org/10.1016/j.jhydrol.2018.11.059

Chaïrat, C., Oelkers, E. H., Schott, J., & Lartigue, J.-E. (2007). Fluorapatite surface composition in aqueous solution deduced from potentiometric, electrokinetic, and solubility measurements, and spectroscopic observations. *Geochimica et Cosmochimica Acta*, *71*(24), 5888–5900. https://doi.org/10.1016/j.gca.2007.09.026

Chaïrat, C., Schott, J., Oelkers, E. H., Lartigue, J.-E., & Harouiya, N. (2007). Kinetics and mechanism of natural fluorapatite dissolution at 25°C and pH from 3 to 12. *Geochimica et Cosmochimica Acta*, *71*(24), 5901–5912. https://doi.org/10.1016/j.gca.2007.08.031

Corneille, R., & Dawes, T. (2001). The groundwater replenishment system—A supplemental source of high quality water for Orange County, California. *Proceedings of the Water Environment Federation*, *2001*(6), 257–275. https://doi.org/10.2175/193864701784291839

CSIRO (2009). *Water yields and demands in south-west Western Australia: A report to the Australian Government from the CSIRO South-West Western Australia Sustainable Yields Project*. Perth: CSIRO.

CyMod Systems Pty Ltd (2014). *Construction and calibration of the Perth Regional Aquifer Model, PRAMS 3.5.2*, report prepared by CyMod Systems Pty Ltd. Rep. 1 74043 540 0. Perth, Western Australia: Departement of Water.

Dawes, W., Ali, R., Varma, S., Emelyanova, I., Hodgson, G., & McFarlane, D. (2012). Modelling the effects of climate and land cover change on groundwater recharge in south-west Western Australia. *Hydrology and Earth System Sciences*, *16*(8), 2709–2722. https://doi.org/10.5194/hess-16-2709-2012

de Silva, J., Wallace-Bell, P., Yesterner, C., & Ryan, S. (2013). *Perth Regional Aquifer Modelling System (PRAMS) v 3.5—Conceptual model (Rep. HR334), Western Australia*. Western Australia: Department of Water, Government of Western Australia.

Descourvieres, C. (2011). *Quantifying water quality changes during managed aquifer recharge in a physically and chemically heterogeneous aquifer* (p. 195). Perth, Western Australia: The University of Western Australia.

Descourvieres, C., Hartog, N., Patterson, B. M., Oldham, C., & Prommer, H. (2010). Geochemical controls on sediment reactivity and buffering processes in a heterogeneous aquifer. *Applied Geochemistry*, *25*(2), 261–275. https://doi.org/10.1016/j.apgeochem.2009.11.012

Descourvieres, C., Prommer, H., Oldham, C., Greskowiak, J., & Hartog, N. (2010). Kinetic reaction modeling framework for identifying and quantifying reductant reactivity in heterogeneous aquifer sediments. *Environmental Science & Technology*, *44*(17), 6698–6705. https://doi.org/10.1021/es101661u

Dillon, P., Stuyfzand, P., Grischek, T., Lluria, M., Pyne, R. D. G., Jain, R. C., et al. (2018). Sixty years of global progress in managed aquifer recharge. *Hydrogeology Journal*. https://doi.org/10.1007/s10040-018-1841-z

Duckworth, O. W., & Martin, S. T. (2004). Role of molecular oxygen in the dissolution of siderite and rhodochrosite. *Geochimica et Cosmochimica Acta*, *68*(3), 607–621. https://doi.org/10.1016/S0016-7037(03)00464-2

Eckert, P., & Appelo, C. (2002). Hydrogeochemical modeling of enhanced benzene, toluene, ethylbenzene, xylene (BTEX) remediation with nitrate. *Water Resources Research*, *38*(8), 38. https://doi.org/10.1029/2001WR000692

Eiche, E. (2009). *Arsenic mobilization processes in the Red River Delta, Vietnam: Towards a better understanding of the patchy distribution of dissolved arsenic in alluvial deposits*. Karlsruher Mineralogische und Geochemische Hefte 37: KIT Scientific Publishing.

Fakhreddine, S., Dittmar, J., Phipps, D., Dadakis, J., & Fendorf, S. (2015). Geochemical triggers of arsenic mobilization during managed aquifer recharge. *Environmental Science & Technology*, *49*(13), 7802–7809. https://doi.org/10.1021/acs.est.5b01140

Fakhreddine, S., Prommer, H., Gorelick, S. M., Dadakis, J., & Fendorf, S. (2020). Controlling arsenic mobilization during managed aquifer recharge: The role of sediment heterogeneity. *Environmental Science & Technology*, *54*(14), 8728–8738. https://doi.org/10.1021/acs.est.0c00794

Feehley, C. E., Zheng, C., & Molz, F. J. (2000). A dual-domain mass transfer approach for modeling solute transport in heterogeneous aquifers: Application to the Macrodispersion Experiment (MADE) site. *Water Resources Research*, *36*(9), 2501–2515. https://doi.org/10.1029/2000WR900148

Fernandez-Bastero, S., Gil-Lozano, C., Briones, M., & Gago-Duport, L. (2008). Kinetic and structural constraints during glauconite dissolution: Implications for mineral disposal of CO₂. *Mineralogical Magazine*, *72*(1), 27–31. https://doi.org/10.1180/minmag.2008.072.1.27

Fernàndez-Garcia, D., Llerar-Meza, G., & Gómez-Hernández, J. J. (2009). Upscaling transport with mass transfer models: Mean behavior and propagation of uncertainty. *Water Resources Research*, *45*, 26. https://doi.org/10.1029/2009WR007764

Ganot, Y., Holtzman, R., Weisbrod, N., Russak, A., Katz, Y., & Kurtzman, D. (2018). Geochemical processes during managed aquifer recharge with desalinated seawater. *Water Resources Research*, *54*, 978–994. https://doi.org/10.1002/2017WR021798

Gerber, P. (2016). Quantifying the long-term impacts of groundwater replenishment with highly treated wastewater. ETH Zurich

Griffioen, J. (2003). Kation-uitwisselingspatronen bij zoet/zout grondwaterverplaatsingen. In *Stromingen* (Vol. 9, pp. 35–45).

Guo, Z., Henri, C. V., Fogg, G. E., Zhang, Y., & Zheng, C. (2020). Adaptive Multirate Mass Transfer (aMMT) model: A new approach to upscale regional-scale transport under transient flow conditions. *Water Resources Research*, *56*, e2019WR026000. https://doi.org/10.1029/2019WR026000

Hao, Y., Smith, M. M., & Carroll, S. A. (2019). Multiscale modeling of CO₂-induced carbonate dissolution: From core to meter scale. *International Journal of Greenhouse Gas Control*, *88*, 272–289. https://doi.org/10.1016/j.ijggc.2019.06.007

Harbaugh, A. W., Banta, E. R., Hill, M. C., & McDonald, M. G. (2000). *MODFLOW-2000, the US Geological Survey modular ground-water model: User guide to modularization concepts and the ground-water flow process*. Reston, VA, USA: US Geological Survey.

Hellauer, K., Karakurt, S., Sperlich, A., Burke, V., Massmann, G., Hübner, U., & Drewes, J. E. (2018). Establishing sequential managed aquifer recharge technology (SMART) for enhanced removal of trace organic chemicals: Experiences from field studies in Berlin, Germany. *Journal of Hydrology*, *563*, 1161–1168. https://doi.org/10.1016/j.jhydrol.2017.09.044

Higginson, S., & Martin, M. (2012). *Groundwater replenishment trial—Groundwater report—2012* (p. 135). Western Australia: Water Corporation.

Houtte, E. V., & Verbauwhede, J. (2008). Operational experience with indirect potable reuse at the Flemish Coast. *Desalination*, *218*(1–3), 198–207. https://doi.org/10.1016/j.desal.2006.08.028

Jones, G. W., & Pichler, T. (2007). Relationship between pyrite stability and arsenic mobility during aquifer storage and recovery in southwest central Florida. *Environmental Science & Technology*, *41*(3), 723–730. https://doi.org/10.1021/es061901w

Jung, H., & Navarre-Sitchler, A. (2018). Physical heterogeneity control on effective mineral dissolution rates. *Geochimica et Cosmochimica Acta*, *227*, 246–263. https://doi.org/10.1016/j.gca.2018.02.028

Lee, H., & Tan, T. P. (2016). Singapore's experience with reclaimed water: NEWater. *International Journal of Water Resources Development*, *32*, 611–621. https://doi.org/10.1080/07900627.2015.1120188

Leong, C., & Lebel, L. (2020). Can conformity overcome the yuck factor? Explaining the choice for recycled drinking water. *Journal of Cleaner Production*, *242*, 118196. https://doi.org/10.1016/j.jclepro.2019.118196

Leyland, L. (2011). *Hydrogeology of the Leederville Aquifer, central Perth Basin, Western Australia* (p. 164). Western Australia: University of Western Australia.

Li, L., Barry, D., Cuiligan-Hensley, P., & Bajracharya, K. (1994). Mass transfer in soils with local stratification of hydraulic conductivity. *Water Resources Research*, *30*(11), 2891–2900. https://doi.org/10.1029/94WR01218

Li, L., Zhou, H., & Gómez-Hernández, J. J. (2011). Transport upscaling using multi-rate mass transfer in three-dimensional highly heterogeneous porous media. *Advances in Water Resources*, *34*(4), 478–489. https://doi.org/10.1016/j.advwatres.2011.01.001

Loáiciga, H., Maidment, D. R., & Valdes, J. B. (2000). Climate-change impacts in a regional karst aquifer, Texas, USA. *Journal of Hydrology*, *227*(1–4), 173–194. https://doi.org/10.1016/S0022-1694(99)00179-1

Maliva, R. G. (2019). Wastewater MAR and indirect potable reuse. In *Anthropogenic aquifer recharge* (pp. 717–763). Part of the Springer Hydrogeology book series (SPRINGERHYDRO).

Mankad, A., & Tapsuwan, S. (2011). Review of socio-economic drivers of community acceptance and adoption of decentralised water systems. *Journal of Environmental Management*, *92*(3), 380–391. https://doi.org/10.1016/j.jenvman.2010.10.037

Marks, J., Martin, B., & Zadoroznyj, M. (2008). How Australians order acceptance of recycled water: National baseline data. *Journal of Sociology*, *44*(1), 83–99. https://doi.org/10.1177/1440783307085844

McNab, W. W. Jr., Singleton, M. J., Moran, J. E., & Esser, B. K. (2009). Ion exchange and trace element surface complexation reactions associated with applied recharge of low-TDS water in the San Joaquin Valley, California. *Applied Geochemistry*, *24*(1), 129–137. https://doi.org/10.1016/j.apgeochem.2008.11.009

Oren, O., Gavrieli, I., Burg, A., Guttman, J., & Lazar, B. (2007). Manganese mobilization and enrichment during soil aquifer treatment (SAT) of effluents, the Dan Region Sewage Reclamation Project (Shafdan), Israel. *Environmental Science & Technology*, *41*(3), 766–772. https://doi.org/10.1021/es060576+

Panday, S., Langevin, C. D., Niswonger, R. G., Ibaraki, M., & Hughes, J. D. (2013). *MODFLOW–USG version 1: An unstructured grid version of MODFLOW for simulating groundwater flow and tightly coupled processes using a control volume finite-difference formulation* (Rep. 2328-7055, book 6, chap. A45, 66 p.). U.S. Geological Survey Techniques and Methods.

Parkhurst, D. L., & Appelo, C. (1999). User's guide to PHREEQC (Version 2): A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations.

Pavelic, P., Dillon, P. J., Barry, K. E., Vanderzalm, J. L., Correll, R. L., & Rinck-Pfeiffer, S. M. (2007). Water quality effects on clogging rates during reclaimed water ASR in a carbonate aquifer. *Journal of Hydrology*, *334*(1–2), 1–16. https://doi.org/10.1016/j.jhydrol.2006.08.009

Prommer, H., Barry, D., & Zheng, C. (2003). MODFLOW/MT3DMS-based reactive multicomponent transport modeling. *Ground Water*, *41*(2), 11. https://doi.org/10.1111/j.1745-6584.2003.tb02588.x

Prommer, H., & Stuyfzand, P. J. (2005). Identification of temperature-dependent water quality changes during a deep well injection experiment in a pyritic aquifer. *Environmental Science & Technology*, *39*(7), 2200–2209. https://doi.org/10.1021/es0486768

Prommer, H., Sun, J., Helm, L., Rathi, B., Siade, A. J., & Morris, R. (2018). Deoxygenation prevents arsenic mobilization during deepwell injection into sulfide-bearing aquifers. *Environmental Science & Technology*, *52*, 13,801–13,810. https://doi.org/10.1021/acs.est.8b05015

Prommer, H., Sun, J., & Kocar, B. D. (2019). Using reactive transport models to quantify and predict groundwater quality. *Elements*, *15*(2), 87–92. https://doi.org/10.2138/gselements.15.2.87

Rathi, B., Siade, A. J., Donn, M. J., Helm, L., Morris, R., Davis, J. A., et al. (2017). Multiscale characterization and quantification of arsenic mobilization and attenuation during injection of treated coal seam gas coproduced water into deep aquifers. *Water Resources Research*, *53*, 10,779–10,801. https://doi.org/10.1002/2017WR021240

Robinson, C., Barry, D., McCarty, P. L., Gerhard, J. I., & Kouznetsova, I. (2009). pH control for enhanced reductive bioremediation of chlorinated solvent source zones. *Science of the Total Environment*, *407*(16), 4560–4573. https://doi.org/10.1016/j.scitotenv.2009.03.029

Sanchez-Vila, X., & Fernàndez-Garcia, D. (2016). Debates—Stochastic subsurface hydrology from theory to practice: Why stochastic modeling has not yet permeated into practitioners? *Water Resources Research*, *52*, 9246–9258. https://doi.org/10.1002/2016WR019302

Schafer, D., Donn, M., Atteia, O., Sun, J., Macrae, C., Raven, M., et al. (2018). Fluoride and phosphate release from carbonate-rich fluor-apatite during managed aquifer recharge. *Journal of Hydrology*, *562*, 809–820. https://doi.org/10.1016/j.jhydrol.2018.05.043

Schafer, D., Sun, J., Jamieson, J., Siade, A., Atteia, O., & Prommer, H. (2020). Model-based analysis of reactive transport processes governing fluoride and phosphate release and attenuation during managed aquifer recharge. *Environmental Science & Technology*, *54*, 2800–2811. https://doi.org/10.1021/acs.est.9b06972

Seibert, S., Atteia, O., Ursula Salmon, S., Siade, A., Douglas, G., & Prommer, H. (2016). Identification and quantification of redox and pH buffering processes in a heterogeneous, low carbonate aquifer during managed aquifer recharge. *Water Resources Research*, *52*, 4003–4025. https://doi.org/10.1002/2015WR017802

Seibert, S., Prommer, H., Siade, A., Harris, B., Trefry, M., & Martin, M. (2014). Heat and mass transport during a groundwater replenishment trial in a highly heterogeneous aquifer. *Water Resources Research*, *50*, 9463–9483. https://doi.org/10.1002/2013WR015219

Shah, T. (2014). Towards a Managed Aquifer Recharge strategy for Gujarat, India: An economist's dialogue with hydro-geologists. *Journal of Hydrology*, *518*, 94–107. https://doi.org/10.1016/j.jhydrol.2013.12.022

Sharma, L., Greskowiak, J., Ray, C., Eckert, P., & Prommer, H. (2012). Elucidating temperature effects on seasonal variations of biogeochemical turnover rates during riverbank filtration. *Journal of Hydrology*, *428*, 104–115. https://doi.org/10.1016/j.jhydrol.2012.01.028

Siade, A. J., Cui, T., Karelse, R. N., & Hampton, C. (2020). Reduced-dimensional Gaussian process machine learning for groundwater allocation planning using swarm theory. *Water Resources Research*, *56*, e2019WR026061. https://doi.org/10.1029/2019WR026061

Siade, A. J., Hall, J., & Karelse, R. N. (2017). A practical, robust methodology for acquiring new observation data using computationally expensive groundwater models. *Water Resources Research*, *53*, 9860–9882. https://doi.org/10.1002/2017WR020814

Sun, J., Prommer, H., Siade, A. J., Chillrud, S. N., Mailloux, B. J., & Bostick, B. C. (2018). Model-based analysis of arsenic immobilization via iron mineral transformation under advective flows. *Environmental Science & Technology*, *52*, 9243–9253. https://doi.org/10.1021/acs.est.8b01762

Wallis, I., Prommer, H., Berg, M., Siade, A. J., Sun, J., & Kipfer, R. (2020). The river–groundwater interface as a hotspot for arsenic release. *Nature Geoscience*, *13*, 288–295. https://doi.org/10.1038/s41561-020-0557-6

Wallis, I., Prommer, H., Pichler, T., Post, V., Norton, S. B., Annable, M. D., & Simmons, C. T. (2011). Process-based reactive transport model to quantify arsenic mobility during aquifer storage and recovery of potable water. *Environmental Science & Technology*, *45*(16), 6924–6931. https://doi.org/10.1021/es201286c

Water Corporation (2009). *Site characterisation report—Groundwater replenishment trial*. Perth, Australia: Water Corporation.

Water Corporation (2013). *Groundwater replenishment trial—Final report*. Perth: Water Corporation.

Water Corporation (2016). *Perth groundwater replenishment scheme stage 2—Hydrogeological report*. Perth: Water Corporation.

Welter, D. E., White, J. T., Hunt, R. J., & Doherty, J. E. (2015). Approaches in highly parameterized inversion—PEST++ Version 3, a Parameter ESTimation and uncertainty analysis software suite optimized for large environmental models. *Rep*. 2328-7055, US Geological Survey.

Wester, J., Timpano, K. R., Çek, D., & Broad, K. (2016). The psychology of recycled water: Factors predicting disgust and willingness to use. *Water Resources Research*, *52*, 3212–3226. https://doi.org/10.1002/2015WR018340

WHO (2017). *Guidelines for drinking-water quality—Fourth edition incorporating the first addendum* (p. 541). World Health Organization.

Xanke, J., Jourde, H., Liesch, T., & Goldscheider, N. (2016). Numerical long-term assessment of managed aquifer recharge from a reservoir into a karst aquifer in Jordan. *Journal of Hydrology*, *540*, 603–614. https://doi.org/10.1016/j.jhydrol.2016.06.058

Zhang, Z., & Brusseau, M. L. (1999). Nonideal transport of reactive solutes in heterogeneous porous media: 5. Simulating regional-scale behavior of a trichloroethene plume during pump-and-treat remediation. *Water Resources Research*, *35*(10), 2921–2935. https://doi.org/10.1029/1999WR900162

Zheng, C., & Wang, P. P. (1999). MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems; documentation and user's guide, DTIC Document.