High-resolution simulation of basin-scale nitrate transport considering aquifer system heterogeneity

Steven F. Carle  
Atmospheric, Earth, and Energy Department, Lawrence Livermore National Laboratory, Livermore, California 94551, USA

Brad K. Esser  
Jean E. Moran  
Chemical Biology and Nuclear Science Division, Lawrence Livermore National Laboratory, Livermore, California 94551, USA

ABSTRACT

Nitrate contamination presents a growing threat to many groundwater basins relied upon for drinking water. This study combines geostatistical techniques, parallel computing of flow simulation, and particle tracking to develop realistic nitrate loading and transport scenarios. The simulation scenarios are patterned after the rapidly urbanizing Llagas groundwater subbasin in the south San Francisco Bay area of California. In the Llagas subbasin, groundwater is the sole municipal water supply. A key component of this study is the development of a highly resolved model of the heterogeneity in the aquifer system using a new geostatistical technique for simulating hydrofacies architecture that can incorporate uncertain or “soft” data, such as well driller logs. Numerical simulations of nitrate transport indicate the degree to which a heterogeneous conceptual model can account for dispersion relative to a conventional homogeneous model assuming typical dispersivity coefficients. The heterogeneous model transport results are found to be consistent with observed nitrate contamination patterns and depth distribution as well as groundwater-age trends with depth. The model provides a realistic test-bed for prediction of future nitrate concentrations, including the time frame for potential nitrate impacts to deep wells, given that geochemical data indicate that denitrification is not likely to occur.

Keywords: geostatistics, simulation, nitrate, heterogeneity, geochemistry.

INTRODUCTION

Nitrate contamination of surface and groundwater is a pervasive and growing problem in the nation and in the world. High levels of nitrate affect both human and ecosystem health. High nitrate in drinking water is regulated because it has been associated with methemoglobinemia or “blue baby syndrome” (Johnson and Kross, 1990). High nitrate in surface and coastal waters leads to eutrophication. The human activities that deliver nitrate to groundwater—animal operations, crop fertilization, wastewater treatment discharge, and septic systems—are a legacy of commerce and growth over the last half-century yet are likely to continue into the future. Nitrate loading from these activities has increased significantly in the past 60 yr. Since 1945, the population of the United States, which is correlated to wastewater treatment and septic system use, has more than doubled; synthetic nitrate fertilizer use has increased 20-fold; commercial cattle production has doubled; and poultry production has increased ~100-fold. More specialized and efficient animal production has created more and larger confined animal feeding operations on the majority of livestock or poultry farms with sales greater than $100,000 (Natural Resources Conservation Service, 2003).

In California, >40% of the state’s population uses groundwater for at least a portion of their public-supply needs. About 10% of California’s public drinking-water supply wells produce water that exceeds the nitrate regulatory drinking-water standard (10 mg/L N), and a much larger fraction produces water that approaches the limit (Fig. 1). Since 1984, ~8600 of 25,000 drinking-water supply wells have been shut down because of excessive nitrate. As the population of California increases by 50% over the next 20 yr, largely by urbanization of agricultural areas, water resources will be further stressed, and the loss of drinking-water supply to nitrate contamination will become an increasingly serious water supply issue (Anton et al., 1988; Esser et al., 2002). As California moves increasingly to using groundwater basins to store water in lieu of above-ground reservoirs, nitrate contamination may become a severe water-storage capacity issue (California Department of Water Resources [CDWR], 2003).

Groundwater-quality problems, particularly nitrate, are increasingly being viewed from a basin-scale perspective. Prediction of contamination trends in groundwater is particularly problematic because aquifer systems have long response times (years to centuries) and are heterogeneous and difficult to characterize. Even with precise knowledge of land use and source loading history, prediction of impacts to groundwater is highly uncertain (Fogg et al., 1998b; McLay et al., 2001; Thorsen et al., 2001) and would likely require site-controlled field experiments to decipher site-specific impacts to wells (Harter et al., 2002). These uncertainties magnify when modeling transport of reactive contaminants. The biogeochemistry of nitrate has a significant role on its transport: under oxic conditions, nitrate does not appreciably sorb and generally behaves conservatively; under anoxic conditions, however, nitrate can undergo denitrification, the biogeochemical process of converting nitrate to harmless molecular nitrogen. The effect of denitrification on nitrate transport can be difficult to characterize or predict over field to basin scales.

Numerical models offer means to predict long-term transport and future occurrence of nitrate contamination in aquifer systems, but realistic model implementation remains a challenge. In addition to source uncertainties, the impacts of subsurface heterogeneity, hydrogeological processes (e.g., pumping and recharge patterns, groundwater flow, vadose-zone processes), and denitrification need to be addressed. Some two- and three-dimensional models have addressed denitrification in large-scale scenarios (Uffink and Romkens, 2001; Molenat and...
Gascuel-Odoux, 2002), but without explicit representation of aquifer system heterogeneity.

“Macrodispersion” or large-scale dispersion attributed to hydraulic heterogeneity is a major component to contaminant spreading in aquifer systems. Transport models typically use dispersivity coefficients or Gaussian random field representations to address macrodispersivity (Gelhar and Axness, 1983; Gelhar, 1986; Rubin, 1990; Tompson 1993). However, at the basin scale, heterogeneity associated with the spatial distribution of aquifers and aquitard units or “hydrofacies” within an alluvial basin will likely dominate basin-scale macrodispersion. In most basins, all heterogeneity associated with the “hydrofacies architecture” or hydrostratigraphy of the aquifer system cannot be represented deterministically. Yet, the heterogeneity will profoundly affect predictions of long-term impact to drinking-water wells. These ideas are not new; the problem is how to account for heterogeneity and dispersion at the basin scale using commonly available characterization data and prior knowledge of the hydrogeology. As Bouwer (1991, p. 43) states on the subject of macrodispersion, “As always, the main difficulty in dealing with heterogeneous media is characterization of underground conditions and simplification into manageable systems.” Furthermore, once a realistic high-resolution multimillion-cell hydrogeologic model is constructed, the problem remains to simulate flow and transport on a multimillion-cell numerical grid.

This paper addresses impacts of basin-scale heterogeneity as follows: (1) Abundant but uncertain well driller log data are incorporated as “soft data” in conjunction with accurate but less available lithologic data. (2) Heterogeneity is explicitly accounted for by generating a geostatistical “realization” of aquifer system heterogeneity or “hydrofacies architecture.” (3) High-resolution simulations of the flow field are computed using a robust and parallelized numerical simulation code. (4) Nonreactive transport in the highly heterogeneous system is simulated using particle tracking. (5) The high-resolution transport simulations are compared to lower-resolution simulations assuming homogeneity with effective dispersivity coefficients.

Rather than attempting to reproduce all historical nitrate sources and impact to wells in a basin, this study simply poses realistic nitrate transport scenarios related to historical agricultural use. The nitrate simulation scenarios assume nonreactive transport or unfavorable conditions for denitrification (e.g., oxic groundwater and/or lack of electron donor). Recently acquired geochemical and isotopic data show no evidence for saturated-zone denitrification in the Llagas subbasin (Lawrence Livermore National Laboratory, 2005). Two scales of heterogeneity are considered: (1) hydrofacies related to aquitard, interbedded, and aquifer units, and (2) intermediate scale (2–20 m) hydraulic conductivity variability within the hydrofacies. Obviously, considerable heterogeneity occurs on smaller scales of meters or less. Many previous studies have focused on determining effective dispersivity coefficients for longitudinal scales of hundreds of meters or less. Our basin-scale conceptual model assumes that hydraulic conductivity variability associated with hydrofacies architecture will dominate basin-scale dispersion on scales of kilometers or more. Advantages of the high-resolution simulation approach over use of dispersion coefficients are: (1) Fast pathways and fingering of plumes are resolved; (2) the impact of large-scale “matrix diffusion” or trapping of contaminants in low-hydraulic conductivity zones is accounted for, for example, tailing of breakthrough curves and recalcitrance of residual contamination in source areas; and (3) the model of heterogeneity is consistent with field observations and prior knowledge of the hydrogeologic system.

The main goals of the high-resolution nitrate simulation are to better understand the long-term impact of nitrate loading to address basic questions such as: (1) Although nitrate does not appear to impact deeper municipal wells today, is there a future threat? (2) How much do location and depth affect vulnerability to nitrate contamination? (3) If current levels of nitrate loading continue, will nitrate concentrations continue to increase? (4) If major nitrate sources are eliminated, how will nitrate concentrations change? (5) Is a simplified, low-resolution model sufficient to address long-term basin-scale nitrate issues?

In addition to nitrate transport, the scenarios presented could be modified to address other relatively nonreactive contaminants, such as methyl tertiary butyl ether (MTBE) or perchlorate.

LLAGAS SUBBASIN STUDY AREA

The study area is the Llagas subbasin within the Gilroy-Hollister groundwater basin in California, located ~30 km south of San Jose, California (see Fig. 1). The 25-km-long northwest-trending Llagas subbasin is drained southward by tributaries of the Pajaro River, including Uvas and Llagas creeks, which enter the alluvial basin near Gilroy and Morgan Hill, respectively. The water-bearing materials in the subbasin consist of Pliocene- to Holocene-age unconsolidated to semiconsolidated alluvial deposits (CDWR, 2002).
Aquifer materials associated with fluvial channels include three principal hydrofacies: (1) coarse-grained and laterally continuous lenses composed of gravel, such as overbank and floodplain deposits, and (3) interbedded discontinuous layers of gravel and silt or clay (CDWR, 1981; Brown and Caldwell Geotechnical Consultants, Inc., 1981, personal commun.; Hoose, 1986). These hydrofacies are related to fluvial depositional processes and, thus, geologic knowledge of fluvial depositional architecture can be used to advantage to develop realistic models of hydrofacies architecture. For example, ancient channel belt systems will be relatively anisotropic in shape with elongation in the ancient direction of flow, whereas floodplain deposits are more extensive and isotropic in lateral directions.

The richest source of data for constructing the hydrofacies architecture is found in “well driller logs,” which are understandably of variable quality. Landon and Belliz (1991) and Burow et al. (2004) provided examples of the usefulness of well driller logs for large-scale hydrogeologic characterization in California with recognition of the limitations of these data. Over 300 well driller logs (Fig. 2) are used in this study to provide “soft conditioning” to geostatistical realizations as described in the following. High-quality lithologic descriptions are also available for several boreholes. For this study, well driller logs and other borehole lithologic descriptions were categorized according to hydrofacies with a degree of uncertainty.

The uncertainty in hydrofacies classification using well driller log data was quantified by assigning a correlation value to each log. The correlation value reflects an estimation of the correlation between the well driller log and reality. Thus, the correlation value could range between 0.0 (no correlation to reality) and 1.0 (perfect correlation to reality). Ideally, one could calculate the correlation value by comparing a well driller log to a high-quality lithologic log located a short distance away, then assign that same correlation value to other well driller logs of similar quality. Unfortunately, no such direct comparisons of well driller logs and high-quality lithologic logs were available. For this study, hydrofacies interpretations from the well driller logs were assigned correlation values between 0.3 and 0.7 based on expert opinion of the quality of lithologic description. These correlation values are used by the geostatistical simulation algorithms (summarized in a following section) to treat well driller log data as uncertain or “soft” data. A few high-quality lithologic logs were included in the data set, and these data were assigned correlation values of 1.0.

Hydrology

Water-level elevations indicate an overall southeastern groundwater-flow direction. Figure 3 shows generalized water-level contours for spring 2001 water-level data (Santa Clara Valley Water District, 2002). Figure 3 also shows that groundwater conditions are regarded as confined in the northern and upland areas of the subbasin and confined in the central southern portion of the basin. Groundwater recharge occurs primarily in the unconfined portions of the subbasin near and south of Morgan Hill and immediately south and west of Gilroy.

CDWR conducted a detailed study of the Llagas subbasin (CDWR, 1981). In that study, CDWR developed a simple one-layer groundwater model with 38 cells in the Llagas subbasin. In this study, we used the CDWR model’s average annual cell-by-cell stream infiltration values and average precipitation values to

Figure 2. Posting of well driller log data used for this study. Data are categorized into aquifer, aquitard, and interbedded hydrofacies.

Geosphere, June 2006

197
Historically, nitrate contamination is greatest in wells east of the valley axis, in the central and southern portions of the subbasin, where a large fraction of the wells have concentrations above the MCL of 45 mg/L as NO₃⁻ (Fig. 1). In a study carried out by the local water agency, about half of over 600 private domestic wells tested for nitrate concentration exceeded the MCL (Santa Clara Valley Water District, 1998). Deep, high-capacity public drinking-water wells have lower, but slowly rising nitrate concentrations. This is a serious water supply problem because the region relies exclusively on groundwater for its drinking water, and at least 19 public supply wells are in the contaminated portion of this basin, although none has as yet exceeded the MCL.

Geochemical analyses of well-water samples for major anions and cations, nitrogen and oxygen isotopes of nitrate, dissolved excess nitrogen, tritium and groundwater age, and trace organic compounds indicate that synthetic fertilizer is the most likely source of nitrate in highly contaminated wells (Lawrence Livermore National Laboratory, 2005). In wells with nitrate concentrations greater than 40 mg/L N, nitrate-δ¹⁵N values fall between 3.8‰ and 6.6‰. This range is covered by reported ranges for mineralized inorganic fertilizers and soil organic N, but not by animal wastes or precipitation. Current and historical data consistently show nitrate concentrations that are higher in shallow monitoring and domestic wells than in deep production wells. Several geochemical parameters (dissolved oxygen, tritium, and some major ions) also exhibit vertical stratification. Chemically stratified groundwater could be the result of a transition from oxidizing to reducing conditions at depth, in which case denitrification could potentially account for the observed drop in nitrate concentrations. However, nitrate-δ¹⁵N and nitrate-δ¹⁸O isotopic compositions and absence of excess nitrogen are not consistent with denitrification as a significant process in the fate of nitrate in the subbasin (except in an area of recycled water application).

Tritium concentrations and tritium-helium age data reveal a very dynamic shallow aquifer flow system, with significant recharge and relatively rapid groundwater flow over a large part of the subbasin. Wells screened exclusively in deeper aquifers are devoid of tritium, indicating that the groundwater produced from those zones recharged more than fifty years ago (Lawrence Livermore National Laboratory, 2005).

METHODS

The area encompassed by our geostatistical and numerical flow and transport models is shown in Figure 3. The model area includes most of the Llagas subbasin. The most northerly portion of the subbasin was excluded as a simplification to better approximate confined flow conditions prevalent in the south. The version of the ParFlow code used for this high-resolution multimillion-cell flow simulation is restricted to confined and steady-state flow conditions. The geostatistical and flow and transport simulation methods are summarized in the following, along with necessary details on integration of soft data (e.g., well driller logs) within the geostatistical algorithm.

Transition Probability–Markov Geostatistics

The key step in model development is to generate a highly resolved and realistic simulation of the alluvial heterogeneity in the subbasin. This study uses the transition probability–Markov geostatistical approach described by Carle and Fogg (1996), Carle (1997), and Carle and Fogg (1997), and applied by Carle et al. (1998), Fogg et al. (1998a), Weissmann et al. (1999), Tompson et al. (1999), Weissmann et al. (2002), Maxwell et al. (2003), and Jones et al. (2005). This study adds implementation of a new approach to incorporate uncertain or soft data, such as well driller logs. Carle (2003) contains more details on derivation, application, and validation of the soft data geostatistical approach.
To summarize, the transition probability–Markov geostatistical approach involves quantification of the spatial variability of categorical variables (e.g., hydrofacies) using the transition probability as a bivariate spatial statistic in lieu of indicator cross covariances or variograms as proposed by Deutsch and Journel (1992). Markov chain models of spatial variability are developed for each depositional direction (e.g., dip, strike, and upward) and then combined into a three-dimensional (3-D) Markov chain model (Carle and Fogg, 1997). The 3-D Markov chain model is used in a two-stage conditional simulation algorithm that (1) generates an initial hydrofacies configuration using the sequential indicator simulation algorithm (Deutsch and Journel, 1992) but reformulates it with transition-probability–based cokriging (Carle and Fogg, 1996), and (2) simulates quenching (zero-temperature simulated annealing) with an objective function to minimize the mean squared difference between simulated and modeled transition probabilities (Carle, 1997).

The simulation algorithm can also account for variations in depositional direction using predetermined dip and strike angles, as shown by Carle et al. (1998), Tompson et al. (1999), and Weissmann et al. (1999). Variable dip and strike angles are applied in this study’s application to the Llagas subbasin. Figure 4 shows a geostatistical realization generated for this study. Steep dip angles are present toward the east. Variable directions of depositional strike are evident on the top surfaces.

Transition Probability
A transition probability matrix provides a relatively simple quantitative model to describe a pattern of spatial heterogeneity. An entry \( t_{jk}(h) \) in a transition probability matrix \( T(h) \) is defined by

\[
t_{jk}(h) = \Pr \{ k \text{ occurs at } x + h | j \text{ occurs at } x \}, \tag{1}
\]

where \( j \) and \( k \) are categories (e.g., facies, classifications), \( x \) is a location, and \( h \) is a lag or separation vector.

The transition probability has long been used by geologists to analyze vertical successions (Vistelius, 1949; Krumberin and Dacey, 1969; Doveton, 1971; Miall, 1973). Figure 5 shows vertical transition probability measurements obtained for hydrofacies data with assumed correlation values of 0.7–1.0.

Markov Chain Model
The Markov chain is a stochastic (probabilistic) model that can be applied to transition probability statistics as in equation 1. Conceptually, a one-dimensional Markov chain model assumes that an outcome depends entirely on the closest datum. For example, as applied to a vertical succession of facies, a Markov chain model assumes the probability of the next higher facies occurrence depends entirely on what facies occurs immediately below. Mathematically, the one-dimensional spatial Markov chain model formulates a transition probability matrix \( \hat{T}(h_v) \) as a matrix exponential

\[
\hat{T}(h_v) = \exp \mathbf{R} h_v, \tag{2}
\]

where \( h_v \) is the lag in the \( v \) direction, and \( \mathbf{R} \) is the transition rate matrix in units of length\(^{-1}\).

The matrix exponential is computed by eigenvalue decomposition, such that each transition probability model \( f_{jk}(h) \) for hydrofacies \( j \) and \( k \) consists of a linear combination of exponential functions and the assumed proportions (Agterberg, 1974).

The transition probability–Markov concepts can be extended laterally (e.g., depositional dip and strike) to formulate 3-D Markov chain models \( \hat{T}_{jk}(h) \) for any lag vector \( h \) in any direction (Switzer, 1965; Carle and Fogg, 1997). The 3-D Markov chain is used in both the cokriging and simulated quenching steps of the geostatistical simulation algorithm to produce a realization of hydrofacies architecture (Carle et al., 1998).

Typically, data are abundant for characterizing vertical spatial variability of lithology. The transition probability measurements guide fitting of the upward or \( z \)-direction Markov chain model shown in Figure 5. Hydrofacies categories are ordered in the matrices as (1) aquitard, (2) interbedded, and (3) aquifer. The fitted model assumes that the upward successions of hydrofacies are statistically random:

\[
\mathbf{R} = \begin{bmatrix}
-1/4.0 & R & * \\
* & * & * \\
R & * & -1/3.0
\end{bmatrix} \text{ m}^{-1}, \tag{3}
\]

where proportions are: aquitards = 0.391; aquifers = 0.329; and interbeds = 0.280; 4.0 m and 3.0 m denote mean \( z \)-direction lengths, \( R \) denotes a random juxtapositional tendency, and * invokes application of row and column summing constraints according to probability theory (Agterberg, 1974; Carle and Fogg, 1996) to compute the interbedded category model parameters.

Infusion of Geologic Interpretation
Lateral spatial variability is more difficult to characterize directly from data because of insufficiently close data spacing and uncertainties in vertical and angular control (e.g., depositional dip and strike). However, Markov chain models are readily developed from prior knowledge of facies architecture. Diagonal transition rates are directly related to facies mean lengths, and off-diagonal transition rates are directly related to juxtapositional tendencies—the frequency at which facies occur adjacent to each other. In this study, the

![Figure 4. Geostatistical realization of hydrofacies architecture for Llagas subbasin.](image-url)
transition rate matrices for the $x$ (strike) and $y$ (dip) directions were developed as follows:

$$ \mathbf{R}_x = \begin{bmatrix} -1 / 2500 & 0.5 / 2500 \\ 0 & * & * \\ * & * & * \end{bmatrix} \text{m}^{-1}, \quad (4) $$

where proportions are: aquitards = 0.391; aquifers = 0.280; 2500 m and 300 m denote the mean $x$-direction lengths of aquitards and aquifers, 0.5 indicates the probability of an aquifer being juxtaposed to an aquitard in the $+x$ direction; and $S$ indicates an assumption of symmetry, and

$$ \mathbf{R}_y = \begin{bmatrix} -1 / 5000 & 0.0 \\ 0.0 & * & * \\ * & * & * \end{bmatrix} \text{m}^{-1}, \quad (5) $$

where proportions are: aquitards = 0.391; aquifers = 0.280; 5000 m and 1000 m denote the mean $y$-direction lengths of aquitards and aquifers, and 0.0 off-diagonal entries are placed to prevent aquifers and aquitards from linking together in the downstream direction.

Note that the $x$- and $y$-direction mean length ratios indicate the degree of lateral anisotropy in the hydrofacies geometry. For example, the aquifer hydrofacies $y:x$ mean length ratio of 10:3 is related to the assumption of aquifers derived from channel belt deposition with the $y$ direction aligned in the major direction of flow.

Integration of Soft Data

This study utilizes a novel approach to integrate soft data into the two-step stochastic simulation method. The integration of soft data is implemented by modifying both (1) the cokriging equations for the sequential indicator simulation step and (2) the decision-making process for changing categories at individual grid-cell locations as performed within the simulated quenching step.

Cokriging

The cokriging system of equations is used to estimate the probability that a facies occurs at a certain location given the occurrences of facies at surrounding locations (e.g., data or previously simulated grid cells). In application to stochastic simulation, the “sequential indicator simulation” (SIS) algorithm (Gomez-Hernandez and Srivastava, 1990) incorporates conditional probability estimates to select the simulated facies at cell locations along a random path. Carle and Fogg (1996) showed that indicator (categorical) cokriging equations can be formulated with transition probabilities, such as a Markov chain model, to approximate the conditional probability estimates. To account for soft data, the transition probabilities entered into the cokriging equations should be modified to reflect the uncertainty of the data. This can be accomplished by substituting $I(h)$ in place of $I(h)$ into the cokriging equations, with $I(h)$ defined as

$$ I(h) = \alpha(x)\alpha(x+h)I(h) + [1-\alpha(x)\alpha(x+h)]p_j, \quad (6) $$

where $\alpha(x)$ is the correlation value of data located at $x$, $I(h)$ is the entry in the Markov chain model transition probability matrix for lag $h$, and $p_j$ is the probability that $k$ occurs in the domain of interest (stationary or global probability).

In formulating the cokriging equations with $I(h)$ for the SIS algorithm, only soft data locations were assigned correlation values less than unity; previously simulated locations or hard data locations assumed correlation values of unity. Throughout the cokriging step, the soft data categories remained unchanged to allow the cokriging step to impart structure conditional to the soft data and associated correlation values. Soft data could be changed in the simulated quenching step as described next.

Simulated Quenching

Because a cokriging (or indicator kriging) estimate is not an exact solution to the conditional probability that a facies occurs at a location given surrounding data, a second simulated quenching step is recommended. The simulated quenching step attempts to minimize an objective function defined as the squared difference between calculated and modeled transition probabilities (Carle, 1997). The simulated quenching step is also used at soft data locations to address the uncertainty of the data. As the correlation value of the soft data decreases, the simulated quenching step enables the simulation algorithm to weigh more toward using the model of spatial variability and less toward using the soft data value to select a category at a soft data grid location.

During implementation of the simulated quenching algorithm, the objective function is reduced through change of categories at individual grid-cell locations along a random path. The change of category that reduces the objective function the most is accepted. If no change
reduces the objective function, no swap is performed. To consider soft data, the correlation value is used to prescribe a probability that the category change will be rejected. If the correlation value is 1.0, a category change will always be accepted. If the correlation value is 0.0, a category change that reduces the objective function will always be accepted. If the correlation value is 0.7, for example, a category change will be rejected with 30% probability. In this manner, the correlation value is used to permit facies changes in the realizations proportionately to the degree of certainty in the data. Through comparison of simulation results at and near data and simulated and modeled transition probabilities, Carle (1997) and Carle et al. (1998) provided validation of the transition probability–Markov stochastic simulation approach, while Carle (2003) provided additional testing and validation for consideration of soft data.

**Flow and Transport Simulation**

Once a geostatistical model is developed, it is not difficult to generate multimillion-cell geostatistical realizations and assign hydraulic properties to hydrofacies. Performing high-resolution flow and transport simulation presents a greater computational challenge.

**Hydraulic Property Variation and Calibration**

In the heterogeneous model, each hydrofacies is expected to exhibit spatial variability of hydraulic properties. On the heterogeneous model’s grid spacing of \( dx = 20 \text{ m}, dy = 20 \text{ m} \), however, we assume that hydrofacies spatial variability between grid blocks is uncorrelated. Each hydrofacies is assigned constant porosity and log-normally distributed hydraulic conductivity \((K)\) as shown in Table 1.

<table>
<thead>
<tr>
<th>Hydrofacies</th>
<th>Porosity</th>
<th>Mean ( K ) (m/day)</th>
<th>Std. Dev. ( \log_2(K) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquitard</td>
<td>0.3</td>
<td>0.001</td>
<td>1.0</td>
</tr>
<tr>
<td>Interbedded</td>
<td>0.3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Aquifer</td>
<td>0.3</td>
<td>100.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 1. Hydraulic Properties Assigned to Hydrofacies

Figure 6. Geostatistical realization of hydraulic conductivity for Llagas subbasin.

...
in 46 shallow wells, 47 intermediate wells, and 9 deep wells. Historical pumping rates were used where available. However, pumping rates at many wells, particularly shallow- and intermediate-depth wells, were increased by a factor of 10 or more to compensate for the model’s known under-representation of the total number of active wells in the Llagas subbasin. All pumping rates are constant under the steady-state flow assumption.

**Particle Tracking**

To simulate nonreactive transport in a mass-conservative, accurate, and computationally efficient manner, a Lagrangian particle-tracking algorithm was used (Maxwell et al., 1999). Advection transport is simulated via a quasi-analytical streamline algorithm (Schafer-Perini et al., 1991), while dispersive and diffusive transport is simulated by a random walk algorithm using bilinear velocity interpolation (Labolle et al., 1996). Particles are moved independently with optimized time steps for particle movements. Particles are bifurcated in regions of low particle density.

The homogeneous model assumes relatively large dispersivity coefficients of 10 m longitudinal and 1 m transverse, whereas the heterogeneous model assumes zero dispersivity. Assignment of fixed dispersivity coefficients is problematic because dispersivity increases with scale, although a 10 m longitudinal dispersivity coefficient is believed to be representative for some aquifers up to scales of multiple kilometers (Gelhar, 1986). Dispersion in the heterogeneous model is caused entirely by heterogeneity built into the hydraulic conductivity field.

**Source Loading**

The nitrate transport simulations are conducted as realistic hypothetical scenarios for a few likely sources based on past land use in the Llagas subbasin. However, the entire nitrate source loading history for the subbasin, including septic systems, is by no means accounted for. Suspected nitrate loading sources (with number of source locations) considered by this study are dairies (3), poultry farms (2), greenhouse operations (2), and irrigated agriculture (2) (see Fig. 9). Based on estimates of nitrate produced by these types of farming (Santa Clara Valley Water District, 1994) and assuming increased nitrate loading over time between 1945 and 1985, a source loading curve was developed (Fig. 10). The source loading curves arbitrarily do not consider any nitrate management practices and, therefore, are largely hypothetical. The most concentrated source of nitrate is assumed to be poultry farms. For practical purposes, however, the dairy, poultry farm, and greenhouse sources can be viewed as point sources covering a range of nitrate source loading rates, while the irrigated agriculture sources represent patch-shaped nonpoint sources. The source loading function assumes that the agricultural nitrate sources remain constant after 1985, except for the case of poultry farms. In this modeling scenario, the poultry farm sources impacting groundwater are assumed to cease in 2045.

**Nitrate Transport Results**

Two nitrate transport simulation scenarios are presented and compared: (1) a homogeneous case, where the hydraulic conductivity field is assumed to be homogeneous and effective dispersion coefficients are implemented for transport simulation, and (2) a heterogeneous case, where the geostatistical model defines a spatially variable hydraulic conductivity field. In both cases, the flow models are calibrated to head,
well pumping is identical, and nitrate loading is identical. Visualizations of plume behavior and breakthrough curves are used to compare results from these two conceptual models of basin-scale nitrate transport.

Given that the nitrate loading scenarios are not comprehensive (particularly for irrigated agriculture), simulated breakthrough at the wells is expected to under-represent the total impact of nitrate loading in the Llagas basin. The nitrate transport simulation results focus on the character of well impacts and breakthrough curves under the assumption of a limited number of nitrate sources. These results aid understanding of long-term nitrate impacts and relationships to other observations, such as geochemical and groundwater-age data.

**Flow**

Figures 7 and 8 show simulated head for the homogeneous and heterogeneous cases, respectively, assuming steady-state flow. Hydraulic head is matched reasonably well in both cases to typical head values in the basin (e.g., Fig. 2). The models differ particularly in the west and south.

In the homogeneous model, heads are too high (above 100 m) in the western recharge area because the relatively large proportion of coarse-grained deposits is not accounted for. The heterogeneous model, however, incorporates the known abundance of coarse-grained deposits in the western recharge area to achieve a more realistic calibration to local head. In support of this result, isotopic and groundwater-age results delineate an area of active recharge and relatively rapid flow in the southwestern portion of the study area. In the south, heads are relatively uniform and higher in the homogeneous model because the increased proportion of fine-grained deposits is not accounted for. The heterogeneous model adapts the proportion of coarse- and fine-grained deposits according to local lithologic data. Thus, it is possible to obtain a more realistic flow calibration with a geostatistical model by incorporating well driller log data as soft data to condition local spatial distribution of hydrofacies.

**Dispersion**

Although dispersion can be quantified through moment calculations, the basin-scale dispersive effects from multiple contaminating loading sites and pumping locations can be better appreciated from visual inspection. Figure 11 shows the simulated evolution of nitrate concentration at 1965, 2045, and 2145 yr (20, 100, and 200 yr after the beginning of nitrate loading) for both the homogeneous and heterogeneous conceptual models.
Character of Breakthrough

The character of the breakthrough curves is useful to predict future trends, such as the impact of eliminating a nitrate source (e.g., the hypothetical poultry farms), delay in impact relative to timing of source loading, and vulnerability of wells with respect to depth.

Figures 12A and 12B show breakthrough curves for shallow wells in the heterogeneous and homogeneous model. Wells 18 and 43 show similar shapes in breakthrough for both the heterogeneous and homogeneous models, suggesting that a homogeneous model could account for heterogeneity-related impacts of dispersion in some cases. Upon closer inspection of Figure 11, wells 18 and 43 are located along the lateral periphery of the nitrate plumes.

Of the shallow wells, wells 32, 34, and 39 in the heterogeneous model, and 33, 47, and 61 in the homogeneous model exhibit peaked breakthrough, suggesting impact from the poultry farm sources (which terminate in year 2045). Wells 32, 34, and 39 are all located very close to the poultry farm sources. The breakthrough curves in Figure 12A suggest that removal of the poultry farm source would result in a near-immediate nitrate concentration decrease in these shallow wells over a time frame of 40–100 yr. Similarly, well 40 in the homogeneous case shown in Figure 12B exhibits a rapid nitrate concentration decrease within a time frame of 25 yr. Shallow well 50, however, shows a delayed response to the poultry farm nitrate source, with peak breakthrough ~90 yr after termination of the poultry farm source. Interestingly, none of the shallow wells in the heterogeneous model exhibit a comparable delayed response. The difference in response can be attributed to the stratigraphic dip angles incorporated into the heterogeneous model. These dip angles cause the poultry farm source nitrate contamination to dive downward below apparently downgradient shallow wells.

Wells 18 and 43 in both models and wells 15, 16, and 46 in the heterogeneous model exhibit relatively stable long-term nitrate concentrations. This stabilization indicates that if loading sources remain constant, many shallow wells can be expected to exhibit a relatively constant long-term nitrate concentration. However, the breakthrough responses from wells 15, 16, 18, and 43 indicate that stabilization in nitrate breakthrough for shallow wells could lag by several decades behind the stabilization of nitrate loading.

Figures 12C and 12D show breakthrough in intermediate-depth wells. Breakthrough occurs in both the heterogeneous and homogeneous models for wells 27, 33, 41, and 65. Breakthrough for wells 27 and 65 is similar for both models, but quite different for wells 33 and 41. In the homogeneous model, wells 33 and 41 show similar peak breakthrough related to separate poultry sources. For well 33, the long-term stabilization indicates contribution from another source, most likely dairy. Well 51 exhibits delayed breakthrough from a distant poultry source in the homogeneous model, and no response in the heterogeneous model.

In the heterogeneous model, multiple sources may be impacting well 33, while well 41 exhibits a double peak, indicating an initial impact from the closest poultry farm source and a later impact from a more distant poultry farm source, an impact not seen in the homogeneous model. Breakthrough in wells 78, 82, and 85 for the heterogeneous model is caused by the western irrigated agriculture source. Breakthrough in wells 48 and 65 for the homogeneous model is caused by the central irrigated agriculture source.

Figure 12E shows breakthrough for two deep wells, 79 and 86, in the heterogeneous model. The delayed, relatively low-magnitude breakthrough indicates that the deep, municipal wells could expect impacts from nitrate in the future given that denitrification is not likely to occur. However, only two of the nine deep wells show any breakthrough over the course of the 500-yr-long simulation. This, combined with the long gradual rise in nitrate concentration for the two impacted deep wells, indicates that mean groundwater travel times to deep well screens would likely be on the order of centuries or more.

The animation (Animation 1) compares simulated nitrate transport over one century—between 1945 and 2045—for the homogeneous and heterogeneous conceptual models with the same hypothetical nitrate loading scenario. The differences in plume evolution highlight impacts of preferential pathways and macrodispersion on prediction of breakthrough behavior, particularly for deeper wells.

DISCUSSION

The transport simulation results for both the heterogeneous and homogeneous models are comparable to the general trends in nitrate contamination in the Llagas subbasin. Based on hundreds of nitrate samples in Llagas subbasin, nitrate concentrations at individual wells have shown consistent upward trends since the 1960s (Santa Clara Valley Water District, 1994, 1998; Lawrence Livermore National Laboratory, 2005). These upward trending nitrate concentrations are prevalent at depths of 60–120 m, which corresponds to the shallow and upper-immediate zones of this modeling study. Low tritium concentration and low nitrate are prevalent below depths of 110–120 m, which would correspond to deeper intermediate wells and portions of deep wells simulated in this study (Lawrence Livermore National Laboratory, 2005). The transition from shallow high-nitrate...
Figure 11. Three-dimensional view of simulated nitrate concentrations at years 1965, 2045, and 2145 (20, 100, and 200 yr since assumed beginning of nitrate loading).
Figure 12. Simulated nitrate breakthrough curves (of anthropogenic nitrate concentration against time) for wells that exhibit peak anthropogenic nitrate concentrations of greater than 0.1 mg/L as nitrate.
that suggest that nitrate does indeed behave conservatively in the Llagas subbasin. However, these modeling results have many limitations. Significantly, the assumption was made that the hydrogeologic conceptual model is correct, which involves several simplifying assumptions, such as: (1) Facies-related heterogeneity patterns are statistically stationary throughout the basin (except for trends in proportions). (2) Only one realization was examined. (3) Structural assumptions, including depositional direction, deformation, and lack of fault displacement, are correct. (4) Flow is steady state and thus ignores seasonal and annual variation of recharge and pumping. (5) Only a small portion of the shallow- and intermediate-depth wells is accounted for. (6) The hydrogeologic complexities of the northern portion of the Llagas subbasin, including the city of Morgan Hill and the northern boundary between Llagas subbasin and the Coyote subbasin north, are not addressed. (7) Only a small portion of irrigated agriculture is accounted for, and geochemical data indicate that irrigated agriculture may be the primary source of nitrate (Lawrence Livermore National Laboratory, 2005). (8) Irrigation return flow is not accounted for. (9) Only a small portion of total nitrate loading is accounted for, and that portion is highly simplified. (10) Changing land use is not accounted for, except for changing nitrate loading rates that either stabilize or terminate. (11) Vadose-zone processes are not accounted for.

The high-resolution, heterogeneous model was primarily constructed to demonstrate the capability to perform basin-scale high-resolution groundwater flow and nitrate transport simulations under a realistic scenario. However, by no means does this paper conclude that the use of heterogeneous conceptual model to account for dispersivity is in all cases more appropriate than using dispersivity coefficients applied to a homogeneous conceptual model. Choice of conceptual model is certainly discretionary. In either the heterogeneous or homogeneous case, the modeling results suggest that calibration to both groundwater-age and nitrate concentration trends will vary improve long-term predictions of basin-scale flow and transport.

In the past, geostatistical techniques and computational capabilities were not available to perform high-resolution basin-scale simulation. Now, lithologic data of variable quality, including well driller logs, can be used through geostatistical techniques to help constrain and characterize 3-D heterogeneous flow and transport models. The practitioner now can choose between conceptual models incorporating explicitly heterogeneous assumptions or effective properties based on homogeneous assumptions. In some respects, a heterogeneous model can be more intuitive to construct because it can be made directly consistent with lithologic data and other geologic knowledge. A homogeneous model may require more effort in determining appropriate effective properties, which may or may not be realistic.

The heterogeneous hydrogeologic model presented in this paper could be improved by adding more lithologic data, particularly high-quality lithologic descriptions obtained by continuous or semicontinuous core, geophysical logs, or cone-penetrometer data. Hydrostratigraphic interpretations could be used to define packages of alluvial deposition instead of lumping the entire subbasin’s alluvial deposits into one set of geostatistical parameters, as was done in this study. Pump test data would be useful to refine the estimates of the hydraulic conductivity distributions for different hydrofacies. A transient flow model could be used, particularly a Richards Equation–based model (Jones and Woodward, 2001), to better account for vadose-zone, recharge, or irrigation return processes. The particle model could also be applied in a backward mode in time to reproduce groundwater-age distributions at wells (Tompson et al., 1999). This would be particularly useful to calibrate a model to groundwater-age data.

**CONCLUSIONS**

Given commonly available data, such as well driller logs, well locations, and pumping rates, combined with hydrogeologic interpretation, geostatistical techniques, parallelized flow, and particle-tracking modeling codes, it is becoming computationally feasible to model basin-scale flow and transport at high

---

**Animation 1.** Simulated nitrate transport at 5-year increments between 1945 and 2045 for the hypothetical nitrate loading scenario in Llagas subbasin with homogeneous and heterogeneous conceptual models. If you are viewing the PDF, or if you are reading this offline, please visit www.gsa-journals.org or http://dx.doi.org/10.1130/GES00032.S1 to view the animation.
resolution with multimillion-cell numerical grids. It is possible to account for uncertainty in lithologic data to enable abundant well driller logs to influence, but not absolutely constrain, the distribution of hydrofacies in a geostatistical realization. Given conditioning data, the geostatistical model can account for differences in proportions of hydrofacies, such as more coarse-grained materials in proximal areas and more fine-grained materials in distal areas. Prior structural knowledge, such as dip and strike and depositional or deformation directions, can be incorporated into the basin-scale geostatistical realizations. Such structural information can influence prediction of transport behavior by, for example, causing contamination migration to dive downward along a dipping bedding plane and miss an apparent downgradient receptor.

A highly resolved heterogeneous numerical model can produce complex transport behavior that produces fingering and greater lateral and transverse dispersion compared to a conventional homogeneous model with typical dispersivity coefficients. Compared to a simplified homogeneous model, the heterogeneous model exhibits broader impact to wells from a fixed number of nitrate sources, including prediction of potential long-term impacts to deeper wells. A highly resolved basin-scale model can be used to evaluate impacts of past and future nitrate loading scenarios and to predict concentration trends and timing of peak breakthrough. As for all modeling of flow and transport, such a basin-scale model can greatly benefit from calibration or consistency checking to groundwater-age data. The highly resolved groundwater flow and transport model developed in this study yields nitrate transport results that are consistent with existing geochemical and groundwater-age data in the Llagas subbasin, which suggest that denitrification is not significant. Future modeling work is needed to assess a more complete inventory of nitrate source loading in Llagas subbasin, particularly for irrigated agriculture, because recent geochemical data indicate synthetic fertilizer is the most likely nitrate source in highly contaminated wells.

ACKNOWLEDGMENTS

The authors thank the Santa Clara Valley Water District for providing lithologic data, cross sections, pumping data, nitrate concentrations, and water-level data in addition to valuable geologic insight. This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract no. W-7405-ENG-48.

REFERENCES CITED


Santa Clara Valley Water District, 1994, Llagas groundwater basin nitrate study, nitrate source area identification: Santa Clara, California, Santa Clara Valley Water District, 39 p.
High-resolution simulation of basin-scale nitrate transport


Geosphere, June 2006