XXXVIII IAH Congress

Groundwater Quality Sustainability
Krakow, 12–17 September 2010

Extended Abstracts

Editors:
Andrzej Zuber
Jarosław Kania
Ewa Kmiecik

University of Silesia Press 2010

abstract id: **260**

title: **Well bore cross-aquifer contamination**

author(s): **Alan L. Mayo**
Brigham Young University, United States, `alan_mayo@live.com`

keywords: groundwater monitoring, cross-aquifer contamination, well purging, ambient well bore flow

ABSTRACT

Water quality tests were performed on two long-screened alluvial aquifer wells (15 to 30 m of screen) that are completed in a heterogeneous aquifer that exhibits extreme temporal water quality variability. When stressed the total dissolved solids (TDS) in one well decreased from 10,600 to 3,500 mg/L and in another well the TDS increased from 136 to 2,255 mg/L. Nested short screened monitoring wells were constructed in chemically distinct horizons affecting each well. Water level measurements and solute and isotopic samples were obtained from the production wells and the monitoring wells during a water quality test. Results of a time drawdown tests demonstrate transmissivity differences between horizons. Ambient water quality in the production wells and aquifer cross-contamination are controlled by well bore mixing due to head differences of as little as 0.01 m between chemically distinct horizons which are linked by the production well screen. During non-stress periods the ambient well bore chemistry is controlled by the horizon with the greatest hydraulic head, whereas during stressed conditions, horizon transmissivity controls the well bore chemistry. In one well aquifer cross-contamination, driven by an ambient head differential of 1.2 m, persisted until about 1,600 well bore volumes were purged.

INTRODUCTION

Production wells are often the primary source of temporal and spatial water quality and hydraulic head data. Such data are used to evaluate the origin, mixing patterns, stratification, and movement of groundwater; to describe the location, geometry, and migration of groundwater contamination; and to calibrate and verify groundwater flow and solute transport models. Because confidence in water quality data is critical, numerous purging and sampling protocols have been developed to help ensure representative water quality data, and considerable attention has been paid to monitoring well design.

Well purging is often needed because stagnant casing storage may not be representative of aquifer water quality. Non-representative water may result from chemical and microbiological induced changes in borehole water quality, internal ambient well bore flow and mixing, the affects of pumping rates and time, and the pump location (Barber and Davis, 1987; Martin-Hayden, 2000a, b; Martin-Hayden and Wolfe, 2000). If the aquifer is chemically homogeneous, well purging should result in a representative water quality sample. Sampling protocols include pumping until field parameters stabilize, evacuation of three or more well-bore volumes, low-flow purging, and calculations of purging times and volumes based on theoretical considerations (Barcelona et al., 1994; Barber and Davis, 1987; Boylan, 2004; Capel et al., 2002; Gibs et al., 1990; Hardy et al, 1989; Knobel, 2006; Robbins et al., 2005; Varljen et al., 2006).

Some aquifers have vertical and/or spatial chemical heterogeneity, thus attention has been paid to monitoring well design including long vs. short screen monitoring wells and discrete and multiport sampling devices (Britt, Tunks, 2003; Einarson, Cherry, 2002; Gibs et al., 1993). Long well screens (i.e., > ~ 2 to 3 m) can bias the sample by diluting water drawn from contaminated horizon(s) with water from non-contaminated horizons. Short screen wells and discrete sampling may bias the sample by either missing contaminated or non-contaminated horizons.

Internal and external factors can bias water quality samples obtained from heterogeneous aquifers. Well bores can induce cross-aquifer contamination in multilayered aquifers (Church and

Granato, 1996; Henrich, 1998; Meiri, 1989; Santi et al., 2006; Sloto, 1996; Sloto et al., 1992). Cross-contamination may occur with small hydraulic head or thermal differentials under ambient conditions (Elci et al., 2001, 2003; Reilly et al., 1989). Temporal solute variability may also result from physical and chemical heterogeneity in the aquifer and from skin effects (Church and Granato, 1996; Reilly and LeBlanc, 1998).

Although many processes which can bias well water quality have been investigated, previous work has focused on: 1) theoretical and modeling approaches (Barber and Davis, 1987; Elci et al., 2001, 2003; Lacombe et al., 1995; Reilly et al., 1989), and 2) laboratory and field experiments generally using low pumping rates (< 5 L s^{-1}), short time intervals (< 1 day), and low solute concentrations (< 100 mg L^{-1}) (Church and Granato, 1996; Hutchins and Acree, 2000; Martin-Hayden 2000b; Reilly and LeBlanc, 1998). Low pumping rates and short pumping intervals mean the travel distance for water drawn into the well bore is small and only a limited aerial extent of the aquifer is involved. Additionally low solute concentrations can make it difficult to chemically distinguish water bearing horizons, to fully evaluate well skin effects, and to identify cross-aquifer contamination.

Short well screens and multiport sampling designs are generally preferred over long well screens for water quality sampling and water level measurements. However, long-screened, production wells with high pumping rates, which may sample chemically heterogeneous aquifers, are often the only data source. Despite the numerous investigations of water quality bias associated with monitoring wells, the potential for water quality bias due to long-term pumping of high output, long-screen production wells is poorly understood.

Because long-screened production wells commonly draw water from heterogeneous aquifers, a critical question is what does water quality data from long-screened production wells in heterogeneous aquifers tell us about the aquifer? To examine this question two unconfined aquifer production wells, which are located in the San Luis Valley, Colorado (Figure 1) and which exhibit large temporal TDS variability, have been investigated.

GEOLOGIC AND HYDROLOGIC CONDITIONS

The San Luis Valley, a major agricultural area located in south-central Colorado (Fig. 1), contains approximately 2.5 x 10^{12} m^3 of groundwater within 1,800 m of the land surface (Romero and Fawcett, 1978). The 95 km wide and 170 km long valley was a closed basin from about 4.5 Ma until a few hundred thousand years ago, when the ancestral Rio Grande overflowed the basin and cut a channel through the San Luis Hills in the southern portion of the valley (Machette, 2004). The northern portion of the valley, known as the Closed Basin, remained a region of internal drainage. Mayo et al. (2007) designated the central portion of the Closed Basin as the ancestral sump due to methane (CH$_4$) evolving, high TDS groundwater in the unconfined and upper portion of the confined aquifer.

In the Closed Basin the ~ 30 m thick unconfined aquifer occurs in the upper part of the Pliocene-Pleistocene Alamosa Formation (Mayo et al., 2007). The Alamosa Formation, which consists of a series of discontinuous, lakebed clay, and other interbeds that are up to several hundred meters thick, also supports underlying confined groundwater systems (Emery et al., 1973; Hanna and Harmon, 1989; Huntley, 1976; Powell, 1958; Romero and Fawcett, 1978). Interbeds in the ance-

stral sump area include well-sorted fluvial deposits, fine-grained lake sediments, organic sediments, and evaporite minerals deposited in ancestral Lake Sipapu (Mayo et al., 2007).

In ancestral sump area, the U.S. Bureau of Reclamation constructed 170 long-screened (15-30 m of screen) production wells and 35 monitoring wells, known as SW and EW wells, respectively (Figure 1). The SW wells have a casing diameter of 0.28 m, a mean screen length of 16.6 m, and a mean well depth of 29.9 m. EW wells have a well casing diameter of 0.1 m, a mean screen length of 4.5 m, and a mean depth of 39.8 m.

Using SW and EW data maximum TDS concentrations in the ancestral sump are contoured on Figure 2. The waters have a maximum concentration of more than 44,000 mg L^{-1} and evolve from Ca^{2+} - HCO$_3^-$ type water outside the ancestral sump (mean TDS 247 mg L^{-1}) to sump area Na$^+$ - HCO$_3^-$ - SO$_4^{2-}$ - Cl$^-$ rich water (mean TDS 2,619 mg L^{-1}). The chemical evolution of the waters is described by Mayo et al. (2007).

METHODS OF INVESTIGATION

Thirty-five SW wells, with a maximum TDS greater than 300 mg L^{-1}, have exhibited TDS variability greater than 25 percent in response to pumping stress. In some wells the TDS increases during pumping and in other wells the TDS decreases during pumping. Absolute differences between minimum and maximum TDS for individual wells range from 52 to 14,394 mg L^{-1}. Wells exhibiting TDS variations are located along a linear trend (Figure 3), which corresponds to the locations of playa and organic rich environments in ancestral Lake Sipapu (Mayo et al., 2007). The combined effects of chemical stratification well bore mixing due to hydraulic head driven ambient well bore flow, and differential well bore inflows during pumping stress were suspected as the cause of temporal TDS variability.

Two production wells, SW-67 and SW-89 (Figure 1), were selected for study because they exhibit large temporal water quality changes in response to pumping stress (Figure 4). Low TDS in SW-67 corresponds to pumping periods, whereas low TDS in SW-89 corresponds to non-pumping periods (Figure 5). After periods of pumping concentrations of Na$^+$, HCO$_3^-$, SO$_4^{2-}$, and Cl$^-$ decreased in SW-67, whereas concentrations of Na$^+$, Ca^{2+}, SO$_4^{2-}$, Cl$^-$, and HCO$_3^-$ increased in SW-89 (Figure 4). The recorded TDS range in SW-67 is 1,030 to 15,427 mg L^{-1} and the recorded TDS range in SW-89 is 123 to 3,250 mg L^{-1}.

SW-67 and SW-89 are located about 10.5 km apart (Figure 1). The unconfined aquifer lithology is not continuous between the two wells. SW-67 lithology is dominated by finer grained sediments, whereas coarser grained sands are more common at SW-89 (Figures 6a and 6b). In both wells numerous thin clay and interbedded clay and sand horizons are encountered. Lithologic logs of wells located between SW-67 and SW-89 suggest that horizons commonly pinch over distances of a kilometer or more.

Using SW and deep boring lithologic logs, water levels, and geophysical logs several water-bearing horizons were identified at SW-67 and SW-89. At each well a series of nested monitoring wells were constructed in three distinct aquifer horizons traditionally considered and legally defined as part of the unconfined aquifer. At SW-67 a monitoring well was also completed in the upper part of the underlying unconfined aquifer. Seven monitoring wells were constructed near SW-67 and six were constructed near SW-89 (Table 1).

At SW-67 a combined 8.5-day time-drawdown pumping and water quality test, using observation well responses, was performed. During the test the pumping rate declined from 16.7 to 11 L s^{-1}. At SW-89 a 23-day water quality test was performed with a pumping rate that declined from 12.7 to 9.1 L s^{-1}. Pumping rates declined in response to falling water levels in the pumping wells (Figure 7). Because the wells contained permanently installed electric pumps, the impeller speed could not be increased to compensate for the declining pumping rates. Water levels were measured using an electrical sounder. Discharge water was conveyed 300 m from the pumping wells via pipeline and discharged into small ponds.

Solute, gas, and isotopic samples were collected prior to and at the end of each test. Pumping well water samples were collected from permanently attached sampling faucets. Each SW well has such a faucet because production wells are sampled at least twice a year. Prior to and at the end of the water quality tests samples were collected from the monitoring wells using a low volume pump. Samples were collected after at least 3 well bore volumes had been removed and the field parameters pH, temperature and conductivity stabilized. The purging protocol is considered adequate because the monitoring wells were designed to sample discrete water bearing horizons.

In addition to field parameters samples were collected for major ion and isotopic analysis (Table 2). Isotopic analysis included δ^2H and $\delta^{18}O$, $\delta^{13}C$, and $\delta^{34}S$. A wide range of solutes and isotopes were collected because it was uncertain which parameters would be most useful. Major ions and stable isotopes were collected to help distinguish geochemical horizons, ambient well bore mixing, and aquifer cross-contamination. The stable isotopes δ^2H, $\delta^{18}O$, $\delta^{13}C$ and $\delta^{34}S$ were collected because they sometimes provide insight into geochemical horizons independently rock-water interactions that effect solute compositions. Major ion charge balance errors for SW-67 data are <3% and errors for SW-89 data are typically <5%. Because many confined and some unconfined aquifer waters exsolve gas, the gas was analyzed for C1-C5 (i.e., methane, ethane, propane, butane, pentane), CO_2, N_2, O_2, Ar, He, and H_2 content (Table 3). $\delta^{13}C$ and δ^2H were determined on CH_4, and $\delta^{13}C$ was determined on CO_2 to help determine the origin of the gas.

RESULTS

SW-67 and SW-89 Water Quality Tests

Water quality tests, which included observation well water quality sampling, were conducted at SW-67 and SW-89 in an attempt to better understand the relationship between TDS variability, pumping stress, and aquifer heterogeneity. SW-67 was pumped for 8.5 days and SW-89 was pumped for 23 days. Prior to the tests SW-67 and SW-89 were not pumped for 123 and 71 days, respectively. Beginning and end of test solute compositions are illustrated as Stiff diagrams in Figures 6a and 6b. During the tests SW-67 TDS declined from 10,600 to 3,530 mg L^{-1} and the TDS in SW-89 increased from 136 to 2,282 mg L^{-1} (Table 2). Water quality stabilization in SW-89 was not achieved until about 15 days of pumping (Figure 8). Similar temporal water quality data are not available for the 8.5 day test at SW-67.

At each pumping well the water bearing horizons are chemically stratified. Three distinct water types were identified: low to moderate TDS Na^+ - HCO_3^- type water, elevated TDS Na^+ - HCO_3^- - SO_4^{2-} type water, and elevated TDS Na^+ - Cl^- - Ca^{2+} - SO_4^{2-} type water. Na^+ - HCO_3^- type water

occurs in horizons 1, 3, and 4 at SW-67 and in horizons 1 and 3 at SW-89. Na^+ - HCO_3^- -SO_4^{2-} type water occurs in horizon 2 at SW-67, and Na^+ - Cl^- - Ca^{2+} - SO_4^{2-} type water occurs in horizon 2 at SW-89 (Figures 6a and 6b).

Only the upper and lower horizons at each pumping well maintained stable chemical compositions during the tests. Changes in chemical compositions in horizons 2 and 3 at SW-67 and horizon 2 at SW-89 resulted from either pumping induce vertical leakage from an adjacent horizon or from the removal of water which invaded the horizon via the pumping well bore during non-pumping periods. Where no pumping induced vertical leakage occured, end of test monitoring well chemistries are assumed to be representative of the background solute compositions of the horizon.

Although an in depth analysis of the chemical evolution is beyond the scope of this investigation, a brief discussion of the chemical evolution will help to provide context for the observed chemical stratification. Diverse carbon histories are evidenced by HCO_3^- and $\delta^{13}C$ contents which vary from 0 to -12 ‰ (Table 2). End of test SW-89 horizons 1 and 3 have HCO_3^- concentrations ~ 1.3 to 3.7 meq L^{-1} and $\delta^{13}C$ ~ -12 to -9 ‰, that are typical of carbon acquired from soil zone CO_2 gas and the dissolution of soil zone carbonate minerals. Most end of test waters have $\delta^{13}C$ compositions of ~ -7 to -5 ‰, however SW-67 horizon 4 has a $\delta^{13}C$ of about 0 ‰. Possible explanations for the less negative $\delta^{13}C$ compositions include: 1) the acquisition of dissolved carbon during a different climatic time, or 2) the acquisition of CO_2 gas or H^+ from additional sources such as methanogenic reactions. δ^2H and $\delta^{18}O$ isotopic composition of SW-89 horizons 2 and 3, and SW-67 horizon 2 may suggest recharge during different climatic conditions than during the recharge of other waters or evaporation at the time of aquifer deposition (Figure 9). Both of these mechanisms could affect the $\delta^{13}C$ composition; however, neither mechanism would increase HCO_3^- to > 20 meq L^{-1}.

At SW-67, $\delta^{13}C$ of -7.4 ‰ or less combined with elevated HCO_3^- concentrations are accompanied by in situ production of methanogenic carbon. Methanogenic processes are described by Doelle (1969), Hunt (1979), Whiticar et al. (1986), and Wolin and Miller (1987). Evolving methane gas (CH_4) was evident during the SW-67 test from several monitoring wells (Table 3). The slight odor of HS^- gas occurs at SW-67, WS-89, and other unconfined aquifer wells. HS^- is a product of sulfate reduction and often associated with methanogenesis. Although the $\delta^{13}C$ composition of horizon 2 may suggest methanogenesis, the SO_4^{2-} content of this water is too elevated for appreciable anaerobic methanogenesis to have occurred. The apparent contradiction between the $\delta^{13}C$ and SO_4^{2-} may indicate that different chemical processes are occurring in different horizons. In summary elevated Na^+ and HCO_3^- contents are attributed to in situ H^+ driven - methanogenic - cation exchange – carbonate mineral dissolution mechanism, whereas low Na^+ and HCO_3^- contents are attributed to cation exchange.

The elevated SO_4^{2-} content of horizon 2 at both wells is attributed to gypsum dissolution. Gypsum dissolution as evidenced by the positive $\delta^{34}S$ compositions (Table 2). Sulfate from reduced sulfur sources (e.g., pyrite) typically have $\delta^{34}S$ values ~ 0 ‰, whereas sulfate from oxidized sources (e.g., gypsum) typically have $\delta^{34}S$ compositions >10 ‰ (Clark and Fritz, 1997). The idea of gypsum dissolution in SW-89 horizon 2 is also supported by gypsum saturation (gypsum SI = 0.01). The low SO_4^{2-} concentrations and very positive $\delta^{34}S$ compositions of SW-89 horizon 1 and 3 waters are attributed to the original sulfate content of the recharge waters. Most stream wa-

ters entering the Valley from both the Sangre de Cristo Range and the San Juan Mountains have very low SO_4^{2-} concentrations, however $\delta^{34}S$ data are only available for one stream entering the valley, thus correlation of $\delta^{34}S$ contents with closed basin groundwaters is problematic (Mayo et al., 2007). The elevated Cl^- at SW-89 horizon 2 most likely results from the dissolution of halite in the aquifer matrix.

δ^2H and $\delta^{18}O$ compositions suggest that some pre-test waters have been subjected to evaporation (Figure 9). Well bore evaporation may be responsible for some pre-test data, although all of the wells were capped. The data also suggest that end of test SW-67 horizon 2 and possibly SW-89 horizon 2 have been evaporated, whereas groundwaters in horizons above and below have not. The significance of the elevated TDS contents of horizons 2 at both wells and the potential evaporation water in SW-67 horizon 2 is that, under natural conditions, upward vertical flow from horizon 2 to horizon 1 has been limited. Under current conditions the vertical gradient is slightly downward from horizon 1 to horizon 2. When the idea of limited vertical flow is considered in light of the closed TDS contours (Figure 2), it is apparent that a zone or zones of stagnate or nearly stagnant groundwater exists in the subsurface. The lateral continuity of the elevated TDS horizons beyond SW-67 and SW-89 is unknown.

SW-67 Time-Drawdown Pumping Test

Unconfined aquifer vertical gradients are downward near SW-67, but the total head difference is generally less than 0.15 m (Table 1; Figure 10). However, in horizon 3 well 3A had a static head about 0.3 m less than the corresponding well 3B, possibly due a facies change between the two wells or vertical leakage from horizon 1. The confined aquifer antecedent water level (horizon 4) was about 1 m above unconfined aquifer levels indicating upward potential between the two aquifers. No discernable trends in antecedent water levels were measured in the 20 days prior to the test.

Results of the 8.5 day aquifer test are shown in Figures 11 and 12. Drawdown data for unconfined aquifer wells (horizons 1, 2, and 3) initially decreased in a straight-line fashion on a semi-log plot (Figure 11). During the test the rate of decline decreased after 200 minutes or less and water levels subsequently increased. The non-linear semi-log slopes are attributed to the declining pumping rate during the test, although facies changes and vertical leakage may also have been factors.

The rise above the static level in the confined aquifer monitoring well 4A (Figure 12) is a common response in Closed Basin upper confined aquifer wells that exsolve (CH_4) methane gas (Huss, 2004). The rise in water level is attributed to the Noordbergum effect which is the result of three-dimensional deformation of adjacent aquifer materials induced by pumping (Hsieh, 1996; Wolff, 1970).

The combined factors that drawdown occurred in each well in each horizon in response to pumping SW-67, horizon 1 drawdown was <0.3 m whereas horizons 2 and 3 drawdowns were several meters, drawdowns in the distant B-series wells were less than in the nearby A-series wells, and SW-67 is only screened opposite horizons 2 and 3 suggests that:

■ horizon 1 is unconfined,

■ horizons 2 and 3 are confined or semi-confined,

■ water from horizon 1 enters the SW-67 well bore by vertical leakage into horizon 2,

■ lateral hydrodynamic communication occurs in each horizon.

Understanding the potential contribution from each horizon to SW-67 is important for evaluating the time-drawdown data. The fact that the horizons are chemically stratified means that it should be possible to use the chemical compositions to help understand the contribution of each horizon to each other and to SW-67.

Assuming the SW-67 well completion isolates horizon 1 from the well bore, drawdown in horizon 1 results from SW-67 pumping induced vertical leakage. Examination of the SW-67 as built documentation and discussions with Rio Grande Water Conversancy personnel (Huss, 2004) supports this assumption. The fact that the post-test TDS of horizon 2 water is less than the end of test TDS and the TDS of both pre-test and end of test horizon 1 water is greater than the TDS of horizon 1 water suggests that ambient leakage from horizon 1 to horizon 2 is minimal. Head differentials between the two horizons of 0.01 to 0.03 m support this idea. Pumping SW-67 created downward head differentials of ~2 to 6 m which could readily induce vertical leakage from horizon 1 to horizon 2. The potential contribution of horizon 1 to horizon 2 during pumping was calculated assuming the final horizon 2 chemistry is a mixture of horizon 1 and pre-test horizon 2 waters. Calculations using SO_4^{2-}, Cl^- , and TDS suggest about 25% of end of test horizon 2 water is from horizon 1.

Mixing ratios were calculated in an attempt to evaluate contributions from horizons 2 and 3 to SW-67 during pumping (Table 4). Mixing calculations only used the A-series end of test results for conservative solute species (SO_4^{2-} and Cl^-) and the isotopic compositions ($\delta^{18}O$, δ^2H, and $\delta^{13}C$). End of test compositions accommodates the effect of vertical leakage from horizon 1 to horizon 2. The A-series wells were chosen because they are close to SW-67 and their water chemistry can not have been impacted by water flowing near the B-series wells. For SO_4^{2-} and Cl^- the calculations suggest that horizon 2 could contribute ~22 to 34% of SW-67 discharge and that horizon 3 contributes most of the water discharging from SW-67.

Mixing calculations for the stable isotopes of water (i.e., $\delta^{18}O$ and δ^2H) suggest that both horizon 2 and 3 could contribute ~ 50% of the water to SW-67. However, the value of the final $\delta^{13}C$ of SW-67 is less than the $\delta^{13}C$ value for water near A-series horizon 3 water, thus the calculated contribution from this horizon is listed as 100 % in Table 4. The $\delta^{13}C$ of water near the B-series horizon 3 well is -5.0 ‰ , which means the combined contributions of water from near the A- and B-series wells is consistent with the end of test SW-67 $\delta^{13}C$ composition. Therefore, based on $\delta^{13}C$ both horizons 2 and 3 could contribute to SW-67 discharge.

Although the chemical data suggest that during pumping horizon 2 contributes 30 to 40 %, horizon 3 contributes 60 to 70% of SW-67 discharge, the data also suggest that ~ 25% of horizon 2 water is from horizon 1. The uncertainty in contributions of from each horizon combined with the non-linear drawdown responses complicate the analysis of the SW-67 observation well time-drawdown data. Because of the uncertainty both analytical and numerical methods were used to analyze aquifer parameters using the time-drawdown data.

Analytical analysis involved curve matching with respect to delayed yield, vertical leakage, and boundary conditions by the methods described in Lohman (1978) and Batu (1998). Only horizons 2 and 3, which have direct hydraulic communication with SW-67, were analyzed by curve

matching methods. The SW-67 pumping rate was apportioned between the horizons based on the calculated mixing ratios. Assigned Q contributions from horizons 2 and 3 are 27% and 73% of SW-67 discharge, respectively, and only the first 300 minutes of data were used. The 27 and 73% values were used to account for some water from the vicinity of the B-series wells. Aquifer parameters were calculated using leaky without storage log-log type curves. The drawdown data were also evaluated relative to other type curves. Using the Q apportionment method, calculated storativity (S) of horizons 2 and 3 are similar, ~3 to 6×10^{-4} and ~2 to 4×10^{-4}, respectively and calculated transmissivity (T) for horizons 2 and 3 were ~ 15 and 45 m^2 day^{-1}. The rate of vertical leakage between horizons 1 and 2 was not quantified.

Numerical analysis, using a radial flow model similar to Hoffmann et al. (1996), was performed on the fist 360 minutes of data by Halford (2009). The numerical model has the advantage that flow rates do not need to be assigned to specific horizons and all horizons can be analyzed simultaneously. Assumptions included vertical to horizontal anisotropy = 0.2, specific storage = 9.9 $x10^{-6}$, and Sy = 0.15. Calculated T values for horizons 1, 2, 3, and 4 were 10, 8, 84, and 67 m^2 day^{-1}, respectively.

Calculated T results using both analytical and numerical methods show similar patterns and are relativity consistent with each other. The results should only be viewed as 1st order approximations due to the numerous assumption used in the analysis. Results of the numerical analyses confirm that the transmissivity of horizon 3 is appreciably greater than the overlying horizons and that most groundwater discharging from SW-67 originates in horizon 3.

DISCUSSION

SW-67 and SW-89, which are open to heterogeneous aquifers, exhibit extreme water quality variability in response to pumping stress. Each well encounters water bearing horizons that are chemically distinct from each other. At each well site chemical differences between the horizons include concentration and in some instances chemical composition. Because pumping well water quality variability is associated with TDS differences between the horizons as well as temporal solute variability within some horizons, the water quality relationships between the pumping wells and the horizons are complex. Understanding this complexity is complicated by the fact that hours to days of well purging are required to stabilize water quality parameters in the production wells. Such purging times greatly exceed typical sampling protocols.

In order to sort out the factors responsible for the temporal water quality variability in the production wells several factors need to be evaluated: 1) the natural water quality in each horizon and the spatial distribution of this water quality, 2) water quality changes within horizons due to pumping induced head changes, and 3) water quality changes induced by ambient (i.e., non-pumping) cross-aquifer contamination via the pumping well bore. The following evaluation will only include analysis of SW-67 data, because only one horizon in SW-89 contains two monitoring wells. Two monitoring wells are needed to evaluate spatial relationships.

At SW-67 only horizon 1 monitoring wells have end of test compositions that are essentially unchanged from pre-test conditions and that are similar in both the A and B series wells. These waters are Na^+ -HCO_3^- type with a TDS of ~2,200 mg L^{-1}. The spatial and temporal consistency of pre-test and end of test compositions suggest chemical homogeneity in horizon 1 that has not influenced by pumping stress.

Horizons 2 and 3 monitoring wells exhibit both spatial and temporal chemical heterogeneity. Horizon 2 pre-test compositions in the A- and B-series wells are chemically similar (Figure 6a). The pre-test composition (Na^+ -HCO_3^- -SO_4^{2-} type water, TDS of ~13,000 mg L^{-1}) likely represents background conditions because: 1) significant natural vertical leakage from horizon 1 into horizon 2 is unlikely do to the small head differential between horizons 1 and 2, 2) horizon 3 water can not invade horizon 2 via SW-67 well bore due to the downward gradient, and 3) horizon 2 waters have the highest TDS. Mixing calculations, discussed above, suggest that the horizon 2 end of test compositions are diluted by a ~25% contribution from horizon 1 via pumping induced vertical leakage. The cause of the large TDS difference between end of test well 2A and well 2B waters may be the result of greater pumping induced vertical leakage from horizon 1 in the vicinity of the B-series wells than in the vicinity of the A-series wells. Lower TDS water beyond well 3B is unlikely for several reasons. The gradient from well 2A to well 2B is 0.007 suggesting horizon 3 ambient water occurring west of well 3B should be similar or more saline than water in encountered in well 2A. The pre-test horizon 2 compositions support this idea. Because prior to pumping the natural groundwater flow is toward well 3B, the TDS beyond the well prior to pumping should also be elevated.

Pre-test and end of test horizon 3 water in well 3A exhibits large compositional and concentration differences, whereas both the composition and concentration in well 3B were relatively stable during the test. In well 3A the end of test TDS declined from ~6,500 to ~700 mg L^{-1} indicating the invasion of substantial amounts of high TDS water into the horizon prior to pumping. Elevated TDS horizon 2 water is the most reasonable source of this water. Because the head differential between well 2A and 3A is very small, only 0.01 m, and only horizon 3 water in the vicinity of the A-series wells was affected, vertical flow in the SW-67 well bore is the most likely avenue for fluid migration between horizon 2 and horizon 3. Mixing calculations using pre-test well 2A and post test well 3A as end members suggest that ~ 50% of the water encountered in well 3A originated in horizon 2. The ~ 45 m thick clay zone separating horizon 4 from horizon 3 would limit vertical leakage from the confined to the unconfined aquifer.

The lateral extent horizon 2 water invasion into horizon 3 was evaluated by two methods: 1) calculation of the volume of water removed from horizon 3 during the water quality test, and 2) calculation of the radius of water invasion into horizon 3 from horizon 2 during a specified time. The volume of water removed calculation utilizes several simplifying assumptions: piston flow, aquifer porosity = 0.2, the pumping time required to remove mixed water from horizon 3 = 8.5 days, average pumping rate = 14 L s^{-1}, saturated thickness in horizon 2 = 7 m, horizon 3 = 10 m. Based on these assumptions the radius impacted in horizon 3 is ~ 35 m. The 8.5 day purging time was selected because this was the duration of the drawdown test. Using a shorter purging time would result in a small radial impact. It should be noted however, that a the pre-test solute composition in well 3B, located 22.5 m from SW-67, was slightly impacted water from horizon 2.

The water invasion estimate involved calculating the ambient flow rate from horizon 2 to horizon 3 via SW-67 well bore. The flow rate was then used to estimate the radius of water invasion. The flow rate calculation assumed horizons 2 and 3 satisfy the Theis assumptions. Assigned aquifer parameters, based on the results of the time-drawdown aquifer test were T= 18 m^2 d^{-1} and S = 0.2 for horizon 2, and T= 84 m^2 d^{-1} and S = 10^{-4} for horizon 3. Calculated ambient flow rates ranged from ~ 0.01 to 0.04 L s^{-1} depending on the assigned head differential.

Using the range of calculated ambient flow rates, the radius of horizon 3 affected and the minimum purging time require to remove the water invaded from horizon 2 were calculated. Assumptions included horizon 3 saturated thickness = 10 m, porosity = 0.2, inflow rate = 0.01 to 0.04 L s^{-1}, and pumping rate = 14 L s^{-1}. The calculated radius impacted after 1 year of ambient well bore flow is ~7 to 12 m, and the purging time required to remove horizon 2 water from horizon 3 ranged from ~ 4 to 24 hours. In addition to the uncertainty in the ambient flow rate calculations and the porosity of horizon 3, the calculations include the simplifying assumptions of piston flow from the well bore into horizon 3, and the regional gradient does not impact the shape of the plume (i.e., radial flow from the well).

The purging times calculated using the ambient flow rate method appears to be low based on field observations, and the 8.5 day purging assumed in method 1 is likely too long. Short term synoptic conductivity data are not available to better fix the necessary purging time for SW-67. SW-67 had only been idle for 71 days prior to the test. The time between pumping events would also greatly impact the purging time as the volume invasion water via ambient flow is proportional to time between pumping events. Synoptic TDS data collected during the 23 day SW-89 test (Figure 8) provides insight in potential purging times. At SW-89, which had only idle for 71 days prior to the test. During the SW-89 water quality test conductivity data were measured frequently (Figure 13). At SW-89, which had been idle for 123 days prior to the test, ~15 days of pumping at ~11 L s-1 were required to remove all invaded water.

CONCLUSIONS

Vertical stratification may affect water quality in sampled wells in the following manner. When the well is not under stress, groundwater from the horizon with the greatest hydraulic head flows into the well bore and displaces water from horizon(s) with lesser hydraulic heads. Water can move up or down the well bore, depending upon the direction of head differentials. Under non-stressed conditions, where the aquifer contains hydrochemically distinct or contaminated horizons, the water chemistry or contaminant concentrations in the well is dominated by the chemistry of the horizon with the greatest hydraulic head. If the well has not been pumped for some time, water from the horizon with the greatest hydraulic head will also move into and mix with groundwater in horizons with lower hydraulic heads.

Pumping removes the mixed groundwater from the well bore and contaminated horizons. Thus, after a well has been pumped for sufficient time the water chemistry in the well will represent the chemistry or contamination of each horizon mass weighted for its transmissivity. Numerous schemes have been developed to determine the pumped volume of water necessary to help ensure samples that are representative of the aquifer (Barcelona et al., 1994; Barber and Davis, 1987; Gibs et al, 1990; Hardy et al, 1989). The general idea in most schemes is that a limited amount of groundwater extraction is required to obtain representative water quality data. Inherent in this is the assumption that the influence of well-bore cross-contamination does not extend for a great distance into the aquifer. In most situations water quality differences between horizons are not great and the true extent of cross-contamination is difficult to quantify.

Because substantial water quality differences exist between aquifer horizons in this study, several observations regarding the potential meaning of water quality samples are possible from the San Luis Valley testing. At SW-89, where chemical stratification is not subtle and thus purging effects can be readily measured, the necessary purging volume greatly exceeded most

protocols. The daily effects of pumping on solute compositions of SW-89 waters are illustrated in Figure 13. Similar data are not available for SW-67. Solute compositions increased steadily until about day 14-16 when compositions stabilized. This is equivalent to vacating about 1,600 well bore volumes before representative water quality was obtained. After chemical stabilization the representative water quality carries the caveat that is does not represent a single horizon, but it represents mixed water quality.

Observation well data from SW-67 and SW-89 demonstrate that well bore mixing in long-screened wells can result in appreciable aquifer water mixing away from the well bore under small head differentials. Thus, mixing can influence the solute, isotopic, and contaminant concentrations in nearby short-screened monitoring wells. Mixing of aquifer waters by invasion of water via the well bore is observed in well 3B (SW-67), at a distance of 22.5 m from the long-screened well pumping under a head differential of only 0.11 m.

Another critical issue is what does the water quality from a sampled well tell about the aquifer system? Water quality sampled from SW-67 and SW-89 provided only limited insight into subsurface conditions. The temporal water quality data did suggest well bore mixing from two or more horizons. However, the pumping well data did not suggest the existence of four hydrogeochemical horizons and it did not suggest chemical facies changes over short distances. For example, the $\delta^{13}C$ of HCO_3^- and the major ion compositions of end test SW-67 horizon 2 waters (i.e., 2A and 2B; Table 2) and the beginning and end of test TDS differences in SW-67 wells 3A and 3B are fundamentally different from each other indicating facies changes over a distance of less than 30 horizontal m.

Some monitoring well data also provided misleading information. For example, OW-3 (SW-89) has only 1.5 m of screen, yet the water quality and isotopic composition of samples at the beginning and ending of the 23-day test are fundamentally different (Figure 6B, Table 2). A similar conditions occurs in well 3A at SW-67 (Figure 8A) Daily sampling from SW-89 (Figure 13) indicate that a considerable volume of water, 1,600 well bores in this case, must be removed to eliminate cross-contaminated water for the aquifer system. Even in cases where the TDS did not appreciable change between the beginning and end of the test, such as well 3B (SW-67) fundamental isotopic and isotopic compositional changes occurred (Table 2).

Knowledge of aquifer lithology and aquifer parameters may provide little comfort in assessing the meaning of water quality data. The extent of chemical stratification and aquifer cross-contamination was not apparent from the borehole lithologic and geophysical data at SW-67 and SW-89.

Because chemical stratification is pronounced over relatively short vertical distances, the water quality variability in long- and short-screened wells in the San Luis Valley provides valuable insights into the impact of chemical stratification on water quality samples and on knowledge of the groundwater system gained from samples. Such insight is not readily apparent in groundwater systems where chemical stratification is not pronounced yet subtle differences occur. This is particularly true for groundwater contamination investigations where concentration differences as small as 0.001 to 0.01 mg L^{-1} may be critical. Such small values may be the difference between meeting or exceeding a water quality standard. As an example of how this critical difference may be important Capel et al. (2002) found that purging 3 well-bore casing may be adequate for major ion analysis but not for some inorganic constituents. In this case it is likely

that different hydrostratigraphic horizons had similar overall water chemistry but not all horizons were nitrate and atrazine contaminated. Thus conventional purging stabilized field parameters, but either cross-contamination or organic constituents remained or other factors affecting the reliability of purged samples remained.

Results of this investigation suggests that: 1) in relatively low TDS groundwater cross-aquifer contamination may persist appreciably longer that previously thought and that typical well purging techniques may not result in representative water quality sample, 2) in long-screened production wells cross-aquifer contamination is common, although it may not be readily apparent when the solute concentrations of the various horizons are similar, and 3) it may not be possible to obtain a non-biased water quality sample. Therefore, the question remains. You've sampled the well, so what do you now know about the aquifer? Clearly thoughtful consideration is required when collecting, interpreting, and evaluating water quality results from long-screened wells in heterogeneous aquifers.

ACKNOWLEDGMENTS

REFERENCES

Barber, C., and Davis, G.B., 1987, Representative sampling of ground water from short screened boreholes: Ground Water v. 25, n. 5, p. 581-587.

Barcelona, M.J., Wehrmann, H.A., and Varljen, M.D., 1994, Reproducible well-purging procedures and VOC stabilization criteria for ground water sampling: Ground Water, v. 32, n. 1, p. 12-22.

Boylan, J.A., 2004, Chemical artifacts of sampling methods in groundwater: Geological Society of America, Abstracts w/Programs, v. 36, n. 5, p. 242.

Britt, S.L., and Tunks, J., 2003, Thorough mixing of contaminant stringers demonstrated in sand tank ground water model: results support expanded use of no-purge sampling techniques: Proceedings of the NGWA conference on Remediation: Site closure and total cost of clean-up, November 13-14, New Orleans, Louisiana, 14 p.

Capel, P.D., Kruglov, V., Landon, M.K., Delin, G.N., and Certain, H.J., 2002, Variations in concentrations of atrazine and nitrate during well purging in relation to field water-quality parameters *in* Hydrologic Science; Challenges in the 21st Century: American Institute of Hydrology, Hydrological Science and Technology Journal, v. 18 n. 1-4, p. 217-230.

Clark, I.D., and Fritz, P., 1997, Environmental isotopes in hydrology: Lewis Publishers, 328 p.

Church, P.E., and Granato, G.E., 1996, Bias in ground-water data caused by well-bore flow in long-screened wells: Groundwater, v. 34, n. 2, p. 262-273.

Batu, V., 1998, Aquifer Hydraulics: A comprehensive guide to hydrogeologic data analysis: John Wiley & Sons, New York, 727 p.

Doelle, H.W., 1969, Bacteria metabolism: New York, Academic Press.

Elci, A., Molz, F.J., III, and Waddrop, W.R., 2001, Implications of observed and simulated ambient flow in monitoring wells: Ground Water, v. 39, n. 6.

Elci, A., Flach, G.P., and Molz, F.J., 2003, Detrimental effects of natural vertical head gradients on chemical and water level measurements in observation wells: identification and control: Journal of Hydrology, v. 28, p. 70-81

Einarson, M.D., and Cherry, J.A., 2002, A new multilevel ground water monitoring system using multichannel tubing: Ground Water Monitoring and Review, v. 22, n. 4, p. 52-65.

Emery, P.A., Snipes, R.J., Dumeyer, J.M., and Klein, J.M., 1973, Water in the San Luis Valley, south central Colorado: Colorado Water Conservation Board, Water Resource Circular 18, 26 p.

Gibs, J., Imbrigiotta, T.E., and Turner, K., 1990, Bibliography on sampling ground water for organic compounds: U.S. Geological Survey Open-file report 90-564.

Gibs, J., Brown, G.A., Turner, K.S., MacLeod, C.L., Jelinski, J.C., and Koehnlein, S.A., 1993, Effects of small-scale vertical variations in well-screen inflow rates and concentrations of organic compounds on the collection of representative ground-water-quality samples: Ground Water, v. 31, n. 2, p. 201-208.

Halford, K.J., 2009, Personal communication: U.S. Geological Survey.

Hanna, T.M., and Harmon, E.J., 1989, An overview of the historical, stratigraphic, and structural setting of the aquifer systems of the San Luis Valley, Colorado in Water in the Valley - A 1989 perspective of water supplies, issues, and solutions in the San Luis Valley, Colorado: Colorado Groundwater Association, p. 1-34.

Hardy, M.A., Leahy P.P., and Alley, W.M., 1989, Well installation and documentation, and ground-water sampling protocols for the pilot National Water-Quality Assessment Program: U.S. Geological Survey Open-File Report 89-396.

Henrich, W.J., 1998, Evaluation of a decommissioned water production well as a potential conduit for cross contamination of confined aquifers using heat-pulse flowmeter, temperature and fluid conductivity geophysical logging techniques: Association of Engineering Geologists, Annual Meeting Abstracts v. 41 p. 96.

Hsieh, P.A., 1996, Deformation-induced changes in hydraulic head during ground-water withdrawal: Ground Water, v. 34, n. 6, p. 1082-1089.

Hunt, J.M., 1979, Petroleum geochemistry and geology: Freeman and Company.

Huntley, D., 1976, Groundwater recharge to the aquifers of the northern San Luis Valley, Colorado: A remote sensing investigation, Unpublished Doctoral Thesis, Colorado School of Mines, 249 p.

Hutchins, S.R., and Acree, S.D., 2000, Ground water sampling bias observed in shallow, conventional wells,; Ground Water Monitoring and Remediation, v. 20, n. 1, p. 86-93.

Huss, F., 2004, Personal Communication: Rio Grande Water Conservation District, Alamosa, Colorado.

Knobel, R. L., 2006, Evaluation of well-purging effects on water –quality results for samples collected from the eastern Snake River Plain Aquifer underlying the Idaho National laboratory, Idaho: U.S. Geological Survey, Scientific Investigations Report 2006-5234, 52 p.

Lacombe S., Sudicky, E.A., Frape, S.K., and Unger, A.J.A., 1995, Influence of leaky boreholes on cross-formational groundwater flow and contaminant transort: Water Resources Research, v. 31, n. 8,p. 1871-1882.

Lohman, S.W., 1972, Ground-water hydraulics: U. S. Geological Survey Professional Paper 708, 70 p.

Machette, M.N., 2004, Personal Communication: U.S. Geological Survey, Denver, Colorado.

Martin-Hayden J.M., 2000a, Sample concentration response to laminar wellbore flow: implications to ground water data variability: Groundwater, v. 38, n. 1, p. 12-19.

Martin-Hayden J.M., 2000b, controlled laboratory investigations of wellbore concentration response to pumping: Groundwater, v. 38, n. 1, p. 121-128.

Martin-Hayden, J.M., and Wolfe, N., 2000, A novel view of wellbore flow and partial mixing: digital image analysis: Ground Water Monitoring Review, WMR, Fall 2000, p. 96-103.

Mayo, A.L., Davey, A., and Christiansen, D., 2007, Groundwater flow patterns in the San Luis Valley, Colorado, USA revisited: An evaluation of solute and isotopic data: Hydrogeology Journal, p. 383-408.

Meiri, D., 1989, A tracer test for detecting cross contamination along a monitoring well column: Ground Water Monitoring Review, v. 9, n. 2, p. 78-81.

Powell, W.J., 1958, Ground-water resources of the San Luis Valley, Colorado: U.S. Geological Survey Water Supply Paper 1379, 284 p.

Reilly, T.E. and LeBlanc, D.R., 1998, Experimental evaluation of factors affecting temporal variability of water samples obtained from long-screened wells: Ground Water, v. 36, n. 4, p. 566-576.

Reilly T.E., Franke, O.L., and Bennett, G.D., 1989, Bias in groundwater samples caused by wellbore flow: journal of Hydraulic Engineering, v. 115, n. 2, p. 270-276.

Robbins, G.A., Metcalf, M., and Budaj, R., 2005, Evaluation of purging and sampling bias in conducting remediation compliance monitoring: Geological Society of America, Abstracts w/Programs, v. 31, n. 1, p. 28.

Romero, J., and Fawcett, D., 1978, Geothermal resources of south central Colorado and their relationship to ground and surface water; State of Colorado, Department Natural Resources, Division Water Resources, 127 p.

Santi, P.M., McCray, J.E., and Martens, J.L., 2006, Investigating cross-contaminated aquifers: Hydrogeology Journal, n. 14, n.1-2, p. 51-68.

Sloto, R.A., 1996, Sample borehole flow to quantify aquifer cross-contamination by volatile organic compounds: U.S. Geological Survey Toxic Substances Hydrology Program; proceedings of the technical meeting, Water-Resources Investigations 94-4015, p. 361-368.

Sloto, R.A., Machiaroli, P., and Towle, M.T., 1992, Identification of a multiaquifer ground-water cross-contamination problem in the Stockton Formation by use of borehole geophysical methods, Hatboro, Pennsylvania *in* Proceedings to Engineering and Environmental Problems: Engineering and Environmental Geophysical Society, SEGEEP Proceedings, 1992, p. 21-35.

Varljen, N.D., Barcelona, M.J., Obereiner, J., and Kaminski, D., 2006, Numerical simulations to assess the monitoring zone achieved during low-flow purging and sampling: Ground Water Monitoring and Remediation, v. 26, v. 1 p. 44-52.

Whiticar, M.J., Faber, E., and Schoell, M., 1986, Biogenic methane formation in marine and freshwater: CO_2 reduction vs. acetate fermentation - Isotopic evidence: Geochim. Cosmochim. Acta, v. 50, p. 693-709.

Wolf, R.G., 1970, Relationship between horizontal strain near a well and reverse water level fluctuation: Water Resources Research, v. 6, n. 4, 1721-1728.

Wolin, M.J., and Miller, T.L., 1987, Bioconversion of organic carbon to CH_4 and CO_2: Geomicrobiology Journal, v. 5, n. 3/4, p. 239-259.

International Association of Hydrogeologists

AGH University of Science and Technology

AGH