California GAMA Special Study: Identifying Paleowater in California Drinking Water Wells

Jean E. Moran*, Menso de Jong*, Ate Visser, Michael J. Singleton, and Bradley K. Esser

Lawrence Livermore National Laboratory
*California State University, East Bay

Prepared in cooperation with the California State Water Resource Control Board
LLNL-TR- 661416

July, 2015
Suggested citation:
California GAMA Special Study:  
Identifying Paleowater in California Drinking Water Wells

J. E. Moran*, M. de Jong*, A. Visser, M. J. Singleton, and B. K. Esser:  
Lawrence Livermore National Laboratory, *California State University-East Bay  
Prepared in cooperation with the State Water Resources Control Board

KEYPOINTS

Paleowater, which herein includes pre-Holocene groundwater and also water that recharged well before the onset of significant human alteration of the hydrologic system in California, is identified using three key isotopic indicators of groundwater residence time: tritium, radiocarbon, and terrigenic helium-4. The goal of this GAMA Special Study is to examine the chemical and isotopic indicators of paleowater in the GAMA data set and identify wells, basins and regions where a portion of the produced water has a very long subsurface residence time.

Identification of paleowater in aquifers tapped for the public water supply is important both for sound water management, because paleowater is likely to be withdrawn unsustainably in high demand regions in the absence of large scale artificial recharge operations, and for contamination susceptibility assessment, as paleowater is typically isolated from contamination sources.

Although considerable uncertainty is associated with calculated apparent ages for individual samples, non-parametric statistical tests indicate that the composite set of isotopic indicators support classification of samples into categories that allow identification of wells most likely to produce paleowater. Approximately 7% of wells with a complete set of isotopic indicator data show strong evidence for producing paleowater.

Wells in desert basins of southeastern California and wells in the southwestern quadrant of the Central Valley frequently produce paleowater that is likely pre-Holocene in age. Very few wells in the northwestern portion of the state, the foothills and Sierra Nevada regions, and coastal basins with intensive artificial recharge activities are categorized as producing paleowater.

Climate is the primary control on paleowater occurrence, with arid portions of the state that were wetter in the distant past having the largest number of wells categorized as producing paleowater. Secondarily, paleowater can be found at the end of very long flow paths in confined aquifers e.g., in the center of the Northern Sacramento Valley. In contrast, paleowater may be masked in areas where unconfined or semi-confined conditions allow substantial mixing between modern recharge and paleowater. Modern, artificially recharged water has replaced very old water on a large scale in urban coastal basins.

Suggested citation:  
Table of Contents

1 Introduction .......................................................................................................................... 3
2 Methods ................................................................................................................................ 5
  2.1 Paleowater Markers ........................................................................................................ 6
    2.1.1 Tritium ..................................................................................................................... 6
    2.1.2 Radiogenic 4-Helium .............................................................................................. 7
    2.1.3 Radiocarbon ............................................................................................................. 7
    2.1.4 Ancillary Indicators ................................................................................................. 8
3 RESULTS and DISCUSSION .............................................................................................. 10
  3.1 Statistical Results ........................................................................................................... 10
  3.2 Paleowater Categorization ............................................................................................ 14
    3.2 Wells with strong evidence of paleowater ................................................................. 15
    3.3 Spatial distribution of wells with evidence of paleowater ......................................... 16
4 REFERENCES ...................................................................................................................... 18
California GAMA Special Study:
Identifying Paleowater in California Drinking Water Wells

J. E. Moran*, Menso De Jong*, A. Visser, M. J. Singleton, and B. K. Esser:
Lawrence Livermore National Laboratory, *California State University East Bay
Prepared in cooperation with the State Water Resources Control Board

1 INTRODUCTION

Inclusion of multiple analyses that allow estimation of groundwater residence time makes the GAMA program unique among groundwater monitoring programs. Groundwater samples collected under GAMA provide a large, spatially distributed, consistently sampled and analyzed set of data from which groundwater residence times can be determined. Groundwater residence times may vary from weeks or months to thousands of years and the long screened intervals typical of water supply wells sampled mean that GAMA wells almost invariably produce groundwater with a wide age distribution. Previous GAMA reports have focused on dividing groundwater age between modern and pre-modern, using presence or absence of tritium as a marker, in order to characterize wells according to susceptibility, given that anthropogenic contaminants are more likely to be found in modern (tritium-containing) groundwater (Visser et al., 2013a, Moran et al., 2004). In this report, we focus on the other end of the residence time spectrum, and examine the distribution of markers of much older groundwater. The primary reason for identifying paleowater is to provide information regarding the sustainability of groundwater extraction and consider situations in which the rate of discharge via wells is much greater than the present rate of recharge. Production of paleowater at drinking water wells may be sustainable, however, if recharge is managed to balance extraction over an acceptable time period.

The terms ‘paleowater’ or ‘fossil’ water are typically reserved for describing pre-Holocene-aged water. Pre-Holocene-aged water has been documented in isotopic investigations of aquifers in Northern Africa (Appelgren, 2004), the Arabian Peninsula (Abderraman, 2004), Maryland (Aeschbach-Hertig et al., 2002), Kansas (McMahon et al., 2004), the Dakota aquifer (Clark et al., 1998), southern Michigan (Ma et al., 2004), and the Rio Grande Basin (Plummer et al., 2004). In California, Izbicki and Michel (2004) and Kulongoski et al. (2005) discuss pre-Holocene water in the Mojave Desert region, and McMahon et al. (2011) include the Central Valley aquifer system in a study of pre-Holocene water in unconfined aquifers in the U.S. This type of groundwater is of interest because it recharged during pluvial periods under cooler and wetter conditions that ended about 10,000 years ago. Since the hydrologic conditions were quite different from those that are currently replenishing groundwater, pre-Holocene groundwater is usually considered a non-renewable groundwater resource.

In this study, we also include water that recharged more recently, but well before human activity significantly affected the natural hydrologic cycle and therefore under hydrologic conditions that may differ considerably from current conditions. This may include water recharged during extended climate oscillations (e.g., Pacific Decadal Oscillation; Kuss and Gurdak, 2014), when water from a former lake was trapped during sediment deposition, or when fluctuations in sea level allowed recharge of fresh water in a zone later inundated. Additionally, groundwater may have recharged under natural conditions which have been altered significantly over the last century by human activities. For example, the major, perennial rivers that drain the Sierra Nevada flooded and recharged portions of the Central Valley each spring, but are now controlled by dams and diversions. Widespread recharge by excess
irrigation water and large-scale managed aquifer recharge in urban areas are two more examples of anthropogenic alteration of natural recharge and discharge patterns. Thus, continued, rapid extraction of groundwater with ages not typically considered ‘paleo’ or ‘fossil’, but still recharged before significantly alteration by human activity in California (>~200 years old), may not be sustainable. Identification of paleowater in resources that are currently exploited for drinking water or for irrigation is therefore critical for predicting the future potential of groundwater as a reliable water supply. Furthermore, identification of paleowater is important for understanding aquifer contamination vulnerability, and for examining geochemical reactions involved in the occurrence of naturally-occurring contaminants.

In previous GAMA studies, GAMA Priority Basin Project (PBP) samples from wells have been categorized as ‘young’ or ‘old’ based in part on tritium activity, with either 1 tritium unit (3.2 pCi/L) or 1 pCi/L serving as the benchmark (Bennett, V et al., 2010; Landon et al., 2010). In those and other GAMA PBP studies, age is considered as one possible contributing factor to explain the distribution of various natural and non-natural constituents. Similarly, Kulongoski et al. (2010) use a combination of tritium, radiogenic helium-4, and carbon-14 to discriminate between ‘young’, ‘old’, and ‘mixed’ aged groundwater (Belitz et al., 2011), and to find statistically significant correlations between groundwater age and several individual water quality indicators.

The focus of this study is to examine the proportion of wells in the entire (state-wide) data set producing very old water and the spatial distribution of those wells. Hence, tritium, radiocarbon, and radiogenic helium-4 results are scrutinized individually and in combination, with emphasis on concentrations that may indicate the presence of paleowater. In addition, statistical methods are employed in order to compare sample populations for paleowater indicators and to examine the significance of co-occurrence of indicators. Other parameters measured under GAMA may be indirect markers of paleowater or show correlations with the three main paleowater indicators. For instance, a cooler climate, as existed during the Pleistocene epoch, has been shown to depress $\delta^{18}$O and noble gas recharge temperatures (Aeschbach-Hertig et al., 1998; Plummer et al., 2012). On the other hand, $\delta^{13}$C and specific conductance may be expected to show cumulative effects of water-rock interaction over time. These parameters were therefore examined in wells that were categorized according to the likelihood that they produce paleowater.
2 METHODS

The data analyzed for signatures of very old water were collected as part of a number of larger and smaller research projects. The majority of the samples were collected by USGS staff as part of the GAMA PBP monitoring program (Belitz et al., 2003). The PBP applied well selection on an irregular grid, to provide a spatially unbiased data set. The PBP data set was supplemented with local and regional studies of groundwater systems, with samples collected by LLNL and local water district staff, as part of the California Aquifer Susceptibility (CAS) program (Moran et al., 2002a; Moran et al., 2002b; Eaton et al., 2003; Moran and Halliwell, 2003; Moran et al., 2004a; Moran et al., 2004b; Moran et al., 2005b; Moran et al., 2005a; Carle et al., 2006) and GAMA Special Studies (McNab Jr et al., 2007; McNab Jr et al., 2010; Singleton et al., 2010; Singleton and Moran, 2010; Moran et al., 2011; Singleton et al., 2012; Visser et al., 2012; Moran et al., 2013; Visser et al., 2013c; Visser et al., 2014).

No single isotopic signature can be used to date water from less than one to tens of thousands of years old or to characterize the age distribution of water produced at long screened wells. Fortunately, several isotope tracers with a wide range in half lives or production rates are introduced to groundwater either in the atmosphere or in porous media. In this report, the emphasis is on tracers of pre-modern and very old water, in particular, tritium (half-life 12.34 years), radiocarbon (half-life 5730 years), and radiogenic helium-4 (accumulates in groundwater over hundreds to thousands of years). As discussed in several previous reports (Moran et al., 2004b; Visser et al., 2013c), long-screened wells invariably produce water of mixed age. Individual wells may exhibit evidence for both recent recharge and ancient water, making quantification of a mean age of limited relevance. Because of mixing processes, markers of young water may mask the presence of fossil water in a sample drawn from a long-screened wellbore.

Tritium, a tracer of modern recharge, is used to identify mixing between modern and pre-modern water (with pre-modern water defined as having <1 pCi/L tritium) and as the first step in categorizing samples as most likely to contain paleowater. Filtering out “modern” samples that contain greater than 1 pCi/L tritium likely leads to ‘false negatives’ (wells where paleowater is present but not called out), but few ‘false positives’ (wells identified as producing paleowater that in fact are producing a significant proportion of modern water). An intermediate designation is applied to samples containing some tritium, which also show evidence for the presence of paleowater. Furthermore, since the focus is on identifying wells producing very old water, the analysis does not explicitly address the volume of such water produced or identification of aquifers from which the water is produced.

Samples were collected for tritium in 1 L glass bottles. Tritium analyses were performed at LLNL by helium-3 accumulation (Clarke et al., 1976; Surano et al., 1992; Visser et al., 2013b) and at the USGS in Menlo Park by enriched liquid scintillation counting (Thatcher et al., 1977). Both techniques are capable of achieving a detection limit of 1 pCi/L (0.3 TU) and have a typical accuracy of 5%.

Noble gas and helium isotope samples were collected in 10 mL copper tubes and were analyzed at the LLNL noble gas mass spectrometry facility (Cey et al., 2008; Visser et al., 2013a; Visser et al., 2013c). Measurement uncertainty of the helium isotope ratio and dissolved concentrations of helium, neon, and argon is 2%; and 3% for krypton and xenon concentrations. Stable isotope samples were collected in 40 mL glass vials and analyzed by IRMS at either LLNL and or by the USGS (Belitz et al., 2003; Fram and Belitz, 2007).
Samples for $^{14}$C analysis were collected by the USGS under the GAMA PBP using protocols described in Fram et al. (2007), and analyzed by accelerator mass spectrometry. Ancillary data such as alkalinity, $\delta^{13}$C, specific conductance, pH, groundwater discharge temperature were measured by the USGS either in the field or in the laboratory as described in Fram et al. (2007).

As described further below, data were analyzed with Analyse-It statistical software add-on for Microsoft Excel. The samples were binned by $^3$H value and then the non-parametric Steel-Dwass-Critchlow-Fligner all-pairs comparison was used to test for significant shifts of the distributions of $^{14}$C, $^4$He$_{terr}$, and depth to screen bottom values between the $^3$H subsets. Where significant differences were found in an order indicating a progression from older to younger water, median values of those indicators from appropriate $^3$H subsets were used as low-cutoff ($^4$He$_{terr}$, depth to screen bottom) and high-cutoff ($^{14}$C) filters on the dataset to categorize samples. Resultant categories were imported into ESRI ArcGIS 10.2 for cartographic presentation.

### 2.1 Paleowater Markers

#### 2.1.1 Tritium

The tritium concentration of a groundwater sample is a good indicator of the presence/absence of water that entered the saturated zone in the last 55 years (Nir, 1964; Vogel et al., 1974). The time period covered by the presence of tritium roughly coincides with the time period during which human activity has had a significant effect on water quality and on groundwater use. Both industrial and agricultural activity (including the use of synthetic nitrogen fertilizer and the expansion of animal populations) significantly increased in California after World War II (Burow et al., 2007; Puckett et al., 2011).

Groundwater recharged before 1950 contains no detectable tritium in 2012, due to decay (62 years is approximately five tritium half-lives). However, there are two conditions that could lead to modern water not being detected by the tritium content of a well water sample: mixing of modern groundwater with pre-modern groundwater (either in the aquifer or in the wellbore) resulting in dilution to a tritium concentration below 1 pCi/L, or recycling of tritium dead water by groundwater pumping and irrigation return flow.

As noted previously, samples collected from long screened production wells – including most public supply wells – contain a mixture of pre-modern and modern groundwater. If the produced water is mostly recent recharge or contains a large fraction of water recharged during the “bomb pulse” of tritium in precipitation during and after the peak of atmospheric nuclear weapons testing in the early 1960s, tritium will be present at concentrations well above the detection limit. On the other hand, produced water may contain no detectable tritium if the fraction of modern water is small.

Present day tritium concentrations in groundwater that recharged between 1982 and 1994 in California are below 4 pCi/L. With a detection limit of 1 pCi/L (consistently achieved for the analytical techniques used by the USGS and LLNL) a mixed groundwater sample can therefore contain as much as 25% of “modern” groundwater without being detected. Hydrogeologically, such a binary mixture of pre-modern groundwater and groundwater that recharged between 1982 and 1994 is not likely a common occurrence, yet it cannot be ruled out.
2.1.2 Radiogenic 4-Helium
Accumulation of helium-4 (\(^{4}\text{He}\)) in groundwater, from the decay of naturally occurring uranium and thorium, provides a qualitative age tracer on the ~500 year to million year time scale (Marine, 1979; Schlosser et al., 1989; Solomon et al., 1996) Groundwater residence time indicators that are useful over the range between those covered by tritium and radiogenic \(^{4}\text{He}\) (apprx. 55 to 500 years) are lacking, though \(^{39}\text{Ar}\) (half-life 269 yrs) shows some promise for bridging the gap. In contrast to cosmogenic and anthropogenic-sourced radionuclides used for dating groundwater (like tritium and radiocarbon), the build-up of radiogenic \(^{4}\text{He}\) is not complicated by a time-varying source, or by decay and dispersion in the vadose zone and saturated zone. Radiogenic \(^{4}\text{He}\) is not measured directly, but rather is calculated from the total \(^{4}\text{He}\) concentration and the \(^{3}\text{He}/^{4}\text{He}\) ratio measured in a dissolved gas sample.

Dating of groundwater using radiogenic \(^{4}\text{He}\) is complicated by heterogeneous rock and sediment compositions with varying concentrations of U and Th, leading to different production rates for \(^{3}\text{He}\) and \(^{4}\text{He}\) and to spatially variable production rates. Furthermore, preferential release or retention of \(^{4}\text{He}\) into fluids is caused by differences in the mineral composition of host rocks, variations in porosity, or diagenetic processes. In addition, variation in the \(^{3}\text{He}/^{4}\text{He}\) in circulating fluids may be caused by mixing with atmospheric, mantle, or other crustal sources of helium.

For the data examined here, terrigenic \(^{4}\text{He}\) (including radiogenic and possible mantle components) and non-atmospheric \(^{3}\text{He}\) (terrigenic + tritiogenic) were calculated from the measured helium isotope ratio, the measured helium concentration and the modeled equilibrium and excess air helium components as described in Visser et al. (2014). Associated uncertainties with measured and modeled concentrations were propagated accordingly.

The radiogenic \(^{4}\text{He}\) age of samples can be calculated using literature values of typical crustal U and Th contents, rock density and aquifer porosity (Larsen and Gottfried, 1960; Solomon et al., 1996; Wagner, 1998). The radiogenic helium age is then determined simply by multiplying the \(^{4}\text{He}\) radiogenic concentration by the accumulation rate. This simplified calculation of groundwater age does not take into account the variable rates of \(^{4}\text{He}\) production or release into water, including a basal flux of helium from the underlying basement rocks that are mentioned above.

2.1.3 Radiocarbon
Groundwater dating based on radiocarbon is possible up to approximately 30,000 years, given the \(^{14}\text{C}\) half-life of 5730 years. \(^{14}\text{C}\) contents (expressed as “percent modern carbon”, pmC) are measured on dissolved inorganic carbon (DIC) in groundwater samples. A single apparent age can be calculated for a sample based on the measured radiocarbon activity relative to an assumed initial activity at the water table of 100 pmC and the \(^{14}\text{C}\) half-life. However, this simple formulation ignores the geochemical evolution of carbon in the vadose zone and in aquifers. Reactions between aquifer materials and dissolved constituents may add C that does not contain \(^{14}\text{C}\) to DIC or may remove C containing \(^{14}\text{C}\) from DIC.

Complex reactions between atmospheric CO\(_2\) dissolved in groundwater, soil organic matter, and C-containing minerals, dictate the DIC concentration and C isotopic composition of groundwater at the water table. Both the radioactive decay of radiocarbon derived from the atmosphere at the time of recharge and water-rock interactions, especially dissolution of and exchange with much older carbonates during transport, will determine the \(^{14}\text{C}\) activity of DIC in groundwater. \(^{14}\text{C}\) ages of groundwater are generally considered highly uncertain because these chemical reactions are poorly understood in most settings and are site specific. In particular, determination of the ‘initial’ activity at
Identification of Paleowater in California Drinking Water Wells

recharge is complex. An uncorrected age assumes an initial $^{14}$C activity of 100 pmC, but much lower activities have been observed near recharge areas, especially beneath thick vadose zones or in strongly reducing conditions (Izbicki and Martin, 1997; McMahon et al., 2011). For a collection of $^{14}$C measurements from a basin, the chemical and isotopic composition of samples from the inferred recharge area may be used to estimate the initial composition and $^{14}$C activity. However, for pre-Holocene groundwater, recharge took place under different recharge conditions so samples revealing corresponding initial conditions do not exist. On the other hand, a significant component of post-1950 recharge would result in pmC values >100% (indicating the presence of ‘bomb pulse’ or other anthropogenic $^{14}$C).

Schemes developed for adjusting the measured radiocarbon activity for the thermodynamically favorable mass transfer reactions, such as those applied in NETPATH (Plummer et al., 1994), require information regarding initial DIC and radiocarbon content. However, several studies have shown that although the $\delta^{13}$C composition is sensitive to chemical reactions along the flow path, $^{14}$C ages are not as sensitive (Davis and Bentley, 1982; Izbicki and Michel, 2004). Furthermore, since this study crosses many different hydrologic settings, estimating initial chemical and isotopic compositions for each sample is not practical.

In samples from California for which $^{14}$C was <50 pmC, differences between uncorrected $^{14}$C ages and ages adjusted for mass transfer reactions were less than 30% (adjusted ages are generally lower than uncorrected ages) (Izbicki and Michel, 2004; McMahon et al., 2010). Samples with higher dissolved solids are more likely to have been strongly affected by water-rock interaction (carbonate dissolution) and show larger differences between uncorrected and adjusted ages. Apparent groundwater ages based on $^{14}$C may therefore overestimate ages by thousands of years. However, by excluding samples containing > 1 pCi/L tritium and having low radiogenic $^{4}$He concentrations in the analysis, the probability of including samples with erroneous $^{14}$C ages is minimized.

2.1.4 Ancillary Indicators

Several parameters that are measured or calculated for GAMA well data are expected to show predictable patterns with respect to groundwater age. Well depth (depth to deepest perforation), screen interval length, noble gas recharge temperature, $\delta^{18}$O of H$_2$O, $\delta^{13}$C of DIC, and specific conductance, were examined in sample groups categorized by the three main paleowater indicators.

In most hydrogeologic settings, the probability of finding paleowater increases with increasing depth, owing to longer flow paths and greater likelihood of confined conditions. Wells screened exclusively in deep, confined aquifers may therefore be expected to be categorized as likely to produce paleowater. Many of the wells in the GAMA dataset are screened over multiple aquifers and draw water with a wide age distribution. Wells with longer well screens are more likely to produce mixed-age water than wells with shorter screened intervals, drawing exclusively from either shallow or deep aquifers.

A robust estimate of the temperature at which recharge took place is determined by measuring the concentrations of all of the dissolved noble gases, and comparing the results to established solubility vs. temperature curves. Typically, recharging groundwater equilibrates with the mean annual air temperature within the vadose zone and recharge temperatures do not vary considerably within basins, except in mountainous regions (Manning and Solomon, 2003). Noble gas recharge temperatures have been applied in studies of paleoclimate, and groundwater recharged during the cooler Pleistocene epoch is identified and typified by a difference of about 5° C (Andrews and Lee, 1979; Stute et al., 1992; Stute et al., 1995; Clark et al., 1997; Aeschbach-Hertig et al., 2002). Similarly, fractionation of the stable
isotopes of the water molecule is dependent upon temperature, and a clear pattern of lower $\delta^{18}$O and $\delta$D is observed for water that precipitates at lower temperature. In California, mean annual air temperature (MAAT) and stable isotope ratios are strongly affected by the physiographic gradient from the Pacific Ocean maritime climate (relatively warm and constant temperatures) to the Sierra Nevada (cold temperatures with wider fluctuations). The ‘continental effect’, whereby water vapor becomes isotopically lighter as it moves inland because the heavier isotope rains out, also controls the stable isotope pattern in precipitation in California.

Chemical evolution of groundwater along flowpaths typically involves water-rock interaction that results in increasing total dissolved solids (or specific conductance) and an approach toward alkaline equilibrium conditions. In aquifers where methane production is not important, $\delta^{13}$C trends toward sedimentary carbonate values near 0‰. Groundwater, which typically acquires an initial $\delta^{13}$C value from atmospheric CO2 at -7.7‰ and soil organic matter at -23‰, may thus be expected to increase in specific conductance and $\delta^{13}$C over time. Most water-rock reactions, however, occur relatively quickly, leaving a stable geochemical groundwater composition after only months to years (Mazor, 2004). However, some of the important alluvial basins in California, dominated by sediments weathered from granitic terrain, are comparatively unreactive compared to sedimentary systems with extensive carbonate deposits.
3 RESULTS AND DISCUSSION

3.1 Statistical Results

Initially, all 6006 samples were examined using basic statistics and categorized according to the measured and calculated residence time indicators. Individual parameters were examined to determine whether the sample population followed a normal distribution.

The underlying hypothesis in the treatment of the indicator data is that the indicators will coincide in expected patterns (absence of tritium, low $^{14}$C activity and high $^{4}$He$_{rad}$ concentrations) in samples with paleowater present. Because of the complexity and expanse of the hydrogeologic conditions covered by the very large data set, linear correlations between indicators are weak or insignificant (near zero). In addition, distributions of $^3$H, $^{14}$C and $^4$He$_{rad}$ are non-normal. For these reasons, non-parametric statistical methods were employed.

A statistical software add-on for Microsoft Excel was used to examine the distributions and basic statistical properties of the measured and calculated residence time indicators within each subset. Pairwise comparisons of significant differences between median values were quantified using an extension of the non-parametric Wilcoxon method, the Steel-Dwass-Critchlow-Fligner test. The Steel-Dwass-Critchlow-Fligner all pairs comparison test was used to test for significant shifts of the distributions of $^{14}$C, $^{4}$He$_{rad}$, and depth to screen bottom values between the $^3$H subsets. Significance at the $p<.05$ level was determined using comparisons of the population median values. Where significant differences were found in an order indicating a progression from older to younger water, median values of those indicators from appropriate $^3$H subsets were used as low-cut ($^{4}$He$_{rad}$, depth to screen bottom) and high-cut ($^{14}$C) filters on the dataset to categorize samples, as described below. Results of the statistical analysis are shown in Figures 1-3.
Figure 1: Depth to bottom of screen binned by tritium activity
Identification of Paleowater in California Drinking Water Wells

Distribution of radiocarbon against tritium

Radiocarbon (as percent modern carbon): means (blue line and triangle), and medians (line and box) binned by tritium activity (as pCi/L $^3$H) with 95% certainty.

Radiocarbon mean and median statistics

<table>
<thead>
<tr>
<th>C14_USGS by H3 Bin</th>
<th>N</th>
<th>Mean</th>
<th>Mean SE (based on pooled SD)</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;H3=&lt;1</td>
<td>263</td>
<td>42.346</td>
<td>1.4534</td>
<td>738 533</td>
<td>27.176</td>
</tr>
<tr>
<td>1&lt;H3=&lt;5</td>
<td>294</td>
<td>70.612</td>
<td>1.3747</td>
<td>617 493</td>
<td>24.849</td>
</tr>
<tr>
<td>H3=0</td>
<td>123</td>
<td>33.913</td>
<td>2.1253</td>
<td>792 289</td>
<td>28.148</td>
</tr>
<tr>
<td>H3&gt;5</td>
<td>515</td>
<td>92.756</td>
<td>1.0386</td>
<td>370 825</td>
<td>19.257</td>
</tr>
<tr>
<td>Pooled</td>
<td>1195</td>
<td>55.570</td>
<td>23.571</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Radiocarbon all pairs $^3$H subset comparisons by Steel-Dwass-Critchlow-Fligner method

Steel-Dwass-Critchlow-Fligner all pairs comparisons

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Hodges-Lehmann location shift</th>
<th>Simultaneous 95% CI</th>
<th>0</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;H3=&lt;1 - 1&lt;H3=&lt;5</td>
<td>-29.490</td>
<td>-36.120 to -22.970</td>
<td></td>
<td>&lt;0.00011</td>
</tr>
<tr>
<td>0&lt;H3=&lt;1 - H3=0</td>
<td>8.220</td>
<td>1.030 to 16.290</td>
<td></td>
<td>0.01539</td>
</tr>
<tr>
<td>0&lt;H3=&lt;1 - H3&gt;5</td>
<td>-51.830</td>
<td>-57.570 to -46.130</td>
<td></td>
<td>&lt;0.00011</td>
</tr>
<tr>
<td>1&lt;H3=&lt;5 - H3=0</td>
<td>39.735</td>
<td>30.230 to 49.410</td>
<td></td>
<td>&lt;0.00011</td>
</tr>
<tr>
<td>1&lt;H3=&lt;5 - H3&gt;5</td>
<td>-20.050</td>
<td>-23.470 to -16.630</td>
<td></td>
<td>&lt;0.00011</td>
</tr>
<tr>
<td>H3=0 - H3&gt;5</td>
<td>-64.530</td>
<td>-72.380 to -54.840</td>
<td></td>
<td>&lt;0.00011</td>
</tr>
</tbody>
</table>

H0: B = 0
The shift in location between the distributions of the populations is equal to 0.
H1: B ≠ 0
The shift in location between the distributions of the populations is not equal to 0.
1 Reject the null hypothesis in favour of the alternative hypothesis at the 5% significance level.

Figure 2: Radiocarbon activity binned by tritium activity
**Terrigenic $^4$He mean and median statistics**

<table>
<thead>
<tr>
<th>He$_4$ by H$_3$ Bin</th>
<th>N</th>
<th>Mean</th>
<th>Mean SE (based on pooled SD)</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;H$_3$$\leq$1</td>
<td>586</td>
<td>0.000000668</td>
<td>0.000000049</td>
<td>0.0000000000</td>
<td>0.000001769</td>
</tr>
<tr>
<td>1&lt;H$_3$$\leq$5</td>
<td>696</td>
<td>0.000000504</td>
<td>0.000000045</td>
<td>0.0000000000</td>
<td>0.000001575</td>
</tr>
<tr>
<td>H$_3$$=$0</td>
<td>169</td>
<td>0.000000536</td>
<td>0.000000092</td>
<td>0.0000000000</td>
<td>0.000001258</td>
</tr>
<tr>
<td>H$_3$$&gt;$5</td>
<td>1599</td>
<td>0.000000102</td>
<td>0.000000030</td>
<td>0.0000000000</td>
<td>0.000000565</td>
</tr>
<tr>
<td>Pooled</td>
<td>3050</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>He$_4$ by H$_3$ Bin</th>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Maximum</th>
<th>Inter-quartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;H$_3$$\leq$1</td>
<td>0.0000000000</td>
<td>0.0000000005</td>
<td>0.0000000073</td>
<td>0.0000000523</td>
<td>0.000001674</td>
<td>0.0000000518</td>
</tr>
<tr>
<td>1&lt;H$_3$$\leq$5</td>
<td>0.0000000000</td>
<td>0.0000000000</td>
<td>0.0000000028</td>
<td>0.0000000273</td>
<td>0.000002414</td>
<td>0.0000000273</td>
</tr>
<tr>
<td>H$_3$$=$0</td>
<td>0.0000000000</td>
<td>0.0000000007</td>
<td>0.0000000097</td>
<td>0.0000000717</td>
<td>0.0000000700</td>
<td>0.0000000710</td>
</tr>
<tr>
<td>H$_3$$&gt;$5</td>
<td>0.0000000000</td>
<td>0.0000000000</td>
<td>0.0000000001</td>
<td>0.000000013</td>
<td>0.0000001265</td>
<td>0.0000000013</td>
</tr>
</tbody>
</table>

**Terrigenic $^4$He all pairs $^3$H subset comparisons by Steel-Dwass-Critchlow-Fligner method**

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Hodges-Lehmann location shift</th>
<th>Simultaneous 95% CI</th>
<th>$0$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0&lt;H$_3$$\leq$1 - 1&lt;H$_3$$\leq$5</td>
<td>0.000000001</td>
<td>0.0000000000 to 0.0000000025</td>
<td>1</td>
<td>&lt;0.0001&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>0&lt;H$_3$$\leq$1 - H$_3$$=$0</td>
<td>0.0000000000</td>
<td>-0.0000000023 to 0.0000000009</td>
<td>1</td>
<td>0.9215&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>0&lt;H$_3$$\leq$1 - H$_3$$&gt;$5</td>
<td>0.0000000001</td>
<td>0.0000000003 to 0.0000000074</td>
<td>1</td>
<td>&lt;0.0001&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>1&lt;H$_3$$\leq$5 - H$_3$$=$0</td>
<td>-0.000000001</td>
<td>-0.0000000060 to 0.0000000000</td>
<td>1</td>
<td>0.0019&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>1&lt;H$_3$$\leq$5 - H$_3$$&gt;$5</td>
<td>0.0000000006</td>
<td>0.0000000010 to 0.0000000023</td>
<td>1</td>
<td>&lt;0.0001&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
<tr>
<td>H$_3$$=$0 - H$_3$$&gt;$5</td>
<td>0.0000000007</td>
<td>0.0000000034 to 0.00000000145</td>
<td>1</td>
<td>&lt;0.0001&lt;sup&gt;1&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>1</sup> The shift in location between the distributions of the populations is equal to 0.

<sup>2</sup> The shift in location between the distributions of the populations is not equal to 0.

<sup>1</sup> Do not reject the null hypothesis at the 5% significance level.

<sup>2</sup> Reject the null hypothesis at the 5% significance level.

**Figure 3: Terrigenic $^4$He concentration binned by tritium activity**
3.2 Paleowater Categorization

Of the complete dataset of 6006 samples, 3666 had results for $^3$H concentration. These samples were assigned to four $^3$H subsets representing possible paleowater ($^3$H = 0, 0 < $^3$H < 1), mixed origin water (1 < $^3$H < 5), and modern recharge water ($^3$H > 5). Only samples with values for at least three of the four indicators (screen bottom depth, $^3$H, $^4$He$_{rad}$, and $^{14}$C) were eligible for categorization as paleowater or mixed water in this analysis. As other studies have concluded (e.g., Belitz et al., 2010; Visser et al., 2014), individual indicators cannot be relied upon to accurately characterize an individual sample, so a minimum of three values was used to make the categorization more robust. For those indicators, the respective median values for the 0 < $^3$H <= 1 subset were used as filters. The exception was the depth to screen bottom—because there was no significant difference between $^3$H subsets less than 5 pCi/L, the median of the 1-5 pCi/L was used, as it was still significantly different from the $^3$H > 5 subset. All filters were inclusive. For example, a sample was ineligible for paleowater characterization if any one or more of its indicators met or exceeded the values shown in Table 1. Once excluded from the paleowater category, a second filter set was applied to the excluded samples in the same manner to identify older/mixed water, using median values of the $^3$H > 5 subset indicators. This category possibly includes wells drawing from mixed new and paleowater zones. If any one or more indicators met or exceeded those values, the sample was categorized as modern water. If a sample had values for two or fewer of the four indicators and those indicators present did not meet or exceed the older water filter, the sample was judged to have insufficient data for categorization. Missing $^{14}$C data was the most common reason for putting samples into the insufficient data category.

Table 1: Cutoffs used to classify samples as indicated.

<table>
<thead>
<tr>
<th>Insufficient Data</th>
<th>Modern</th>
<th>Older/Mixed</th>
<th>Paleo</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3$H: High-cut</td>
<td>None</td>
<td>5</td>
<td>1</td>
<td>pCi/L</td>
</tr>
<tr>
<td>Screen Bottom Depth: Low-cut</td>
<td>None</td>
<td>310</td>
<td>479.5</td>
<td>Feet</td>
</tr>
<tr>
<td>$^{14}$C: High-cut</td>
<td>None</td>
<td>95.91</td>
<td>40.89</td>
<td>pmC</td>
</tr>
<tr>
<td>$^4$He: Low-cut</td>
<td>None</td>
<td>0</td>
<td>7.35E-08</td>
<td>cm$^3$STP/g</td>
</tr>
<tr>
<td>n</td>
<td>2154</td>
<td>3179</td>
<td>511</td>
<td>162</td>
</tr>
</tbody>
</table>

The other paleowater indicators were found to statistically differ in the expected pattern between groups defined by tritium concentrations. For example, samples with tritium concentrations below 1 pCi/L showed significantly lower $^{14}$C activity and higher $^4$He$_{rad}$ concentrations than samples with 1-5 pCi/L or >5 pCi/L (fig 2 and fig 3).

Interestingly, the pattern of significantly lower $^{14}$C held for even lower tritium concentrations (<0 pCi/L; p-value .0003). Although tritium values of <0 pCi/L do not have physical meaning, the detection limit for tritium analyzed by the $^3$He accumulation method depends upon the accumulation time and the reported uncertainty. The uncertainty of the measured tritium activity is calculated by error propagation of the standard sensitivity uncertainty, multiplier signal uncertainty, air correction uncertainty and blank uncertainty. Analyses are accepted if the propagated uncertainty at the time of sampling is less than the quadratic sum of 1 pCi/L plus 5% of the tritium activity at the time of sampling. Correction to the tritiogenic $^3$He for $^3$He from an air leak can result in reported values of <0 pCi/L. The cutoff at 0 pCi/L is arbitrary, but it is clear from the analysis that in this large data set where statistically significant
differences exist between the <0 pCi/L and the 0-1 pCi/L groups, some values below the reported 1 pCi/L (+/- 1 pCi/L) are determined with a lower detection limit and lower uncertainty, and contain some information about the relationship between tritium and other age indicators.

The $\delta^{13}$C distribution of the entire data set is normally distributed with a mean of about -15% and with a range expected for groundwater that has incorporated atmospheric CO$_2$ (-7.7%) and remineralized organic C (-23%). In groundwater that has undergone extensive water-rock interaction, $\delta^{13}$C is expected to approach chemical equilibrium with low $\delta^{13}$C carbonates, as described above. This pattern appears as hypothesized, with higher median $\delta^{13}$C values observed in lower tritium sample subsets. Although ‘initial’ values for the parameters discussed so far vary spatially due to diverse geologic and hydrogeologic conditions, the spatial patterns are difficult to quantify on the scale of the data set.

On the other hand, $\delta^{18}$O and recharge temperature both vary spatially according to ambient air temperature and follow predictable patterns in California, with cold temperatures and low $\delta^{18}$O at high elevations and high latitudes. Recharge temperatures determined from dissolved noble gas concentrations were therefore adjusted according to the mean annual air temperature of the sample location (Visser et al., 2014). The resulting NGRT-MAAT difference is expected to be about 5°C lower in pre-Holocene water (Stute et al., 1992; Stute et al., 1995; Aeschbach-Hertig et al., 1998). However, this relationship was not observed in any of the data subsets. Likewise, $\delta^{18}$O values in the subsets do not differ statistically, even for wells within the same geographic provinces. Possible reasons include mixing effects in long-screened wells that mask paleo recharge temperatures, smaller paleotemperature differences in southern and maritime climates of California, or identification of fossil water not recharged during the Pleistocene in this analysis. An approximate 5°C difference in paleo recharge temperature (accompanied by low $^{14}$C and high $^4$He$_{rad}$) was observed in deep, short-screened monitor wells in the northern Sacramento Valley (Moran et al., 2005b), suggesting that the effects of wellbore mixing are a likely factor in masking depressed paleo recharge temperatures.

### 3.3 Wells with strong evidence of paleowater

The subset of wells with results for at least three of the four main indicators and defined by strict cutoffs is most likely to include samples of pre-Holocene groundwater. The subset is made up with wells with screens extending greater than 479.5 feet below ground surface, in which $^3$H is less than 1 pCi/L, $^{14}$C activity is less than 40.89 pmC, and $^4$He$_{rad}$ exceeds $7.35 \times 10^{-6}$ cm$^3$STP/g$_{water}$. As noted under the description of each of these tracers above, determining precise groundwater residence times from measured quantities is associated with a high degree of uncertainty. However, the cutoffs for this subset of 162 wells (of 2330 wells for which at least three of the four indicators were reported), and the combination of three indicators, make it very likely that these wells do produce fossil water with a residence time of at least a few thousand years. That is, the $^3$H cutoff of 1 pCi/L makes it very likely that the water has a residence time of >60 years, the level of <40.89 pmC (uncorrected for geochemical evolution) indicates a residence time of at least 7,400 years, and the radiogenic $^4$He concentration of $>7.35 \times 10^{-6}$ cm$^3$STP/g indicates a residence time of at least 2900 years, based on a production rate of $2.5 \times 10^{-11}$ cm$^3$STP/g-yr. For the subset of 162 wells, indicated $^4$He$_{rad}$ ages range from 2,900 to 616,000 yr and $^{14}$C ages range from 7800 to 44,000 yr. In addition to the cutoff value for $^{14}$C, the application of cutoffs for the other indicators has produced a dataset with a median $^{14}$C age of 15,500 years and mean of age of 16,700 years.
3.4 Spatial distribution of wells with strong evidence of paleowater

Some areas of the state are devoid of wells with strong evidence of paleowater (Figure 4). As noted above, one reason for this may be that mixing with younger water in the wellbore or in the aquifer masks the ‘pure’ paleowater signal required by the cutoffs. However, availability of the very large data set makes it likely that geographic groupings of wells are reliably categorized.

Notable areas without wells in this category include northern volcanic aquifers and alluvial aquifers such as the Scott River Valley, Smith River Valley, and Klamath River Valley where precipitation rates are the highest in the state. The foothill areas of the Sierra Nevada are likewise nearly devoid of wells in this category. The relatively high recharge rates due to snowmelt runoff in the foothills and relatively shallow aquifers tapped by supply wells having relatively low storage volumes do not create conditions conducive to the preservation or sampling of paleowater. The intensively managed coastal, urban basins such as Santa Clara Valley, Orange County, and Los Angeles have few wells with paleowater indicators, owing to rapid turnover caused by high rates of extraction and replenishment. However, in the most distal areas of those basins, upgradient of seawater barrier injection wells, and between the Los Angeles and Orange County recharge flow fields, some wells do produce “archaic” groundwater that falls into the paleowater category, as previously reported in Hudson et al. (2002).

Very few wells from the eastern side of the Central Valley fall into this category, in spite of the very large number of wells from this area included in the data set. The consistent, perennial recharge from rivers draining the Sierra Nevada and perennial irrigation return flow of the same river water, along with the semi-confined nature of the eastern Central Valley aquifer system are likely important explanatory factors.

Paleowater indicators are found in several wells in the Western San Joaquin Valley subbasins, especially from Los Banos to Kettleman City. Of the 162 wells in this subset, 19 are located in the Western San Joaquin Valley USGS PBP study unit. These wells are distant from the main recharge sources of the Sierra Nevada and are in an area where the Corcoran clay is an effective confining unit (Faunt, 2009).

Wells in the desert areas of southeastern California are most likely to produce paleowater as categorized here. Sixty-two of the 162 wells are from desert basins such as Owens Valley, the Upper Mojave Valley, Antelope Valley and other relatively low use basins. The presence of pre-Holocene water in the Mojave region has been reported previously (Izbicki and Michel, 2004; Kulungoski et al., 2005). Here, its prevalence in drinking water wells has additional significance since it indicates reliance upon the resource by the local population.

Deep wells in the center of the northern Sacramento Valley produce paleowater, as reported previously in Moran et al. (2005b). As noted above, in addition to the key residence time indicators used to build this subset, these wells have noble gas recharge temperatures that are about 5°C cooler than MAATs, offering further evidence that Pleistocene-aged water is present in the deep portion of this aquifer system occupied by the Tuscan Formation.
Figure 4: GAMA priority basin study units with wells identified as producing paleowater. Wells producing paleowater are shown as circles with red outline and fill. Other wells sampled are shown as circles with black outline and no fill. GAMA Study provinces by color as shown in legend.
Identification of Paleowater in California Drinking Water Wells

Wells categorized as producing paleowater are also found in the confined portion of the aquifer system of the Salinas Valley, where groundwater extraction rates are very high and managed recharge replenishes only the shallow aquifer (Moran et al., 2011). A few wells on the San Lorenzo plain likewise produce paleowater (also noted for the lower aquifer system in Izbicki et al, 2003) in an area with little groundwater management. Similarly, a few wells in the Sonoma Valley and Napa Valley, where groundwater extraction rates are high and no water is imported into the area, reveal very old water even in relatively shallow aquifer systems (Moran et al., 2010).
4 REFERENCES

GAMA Project reports are available on the State Water Board GAMA website:


Moran, de Jong, Visser Singleton and Esser (2015) LLNL-661416


Moran J. E., Esser B. K., Singleton M. J. and Visser A. (2013) Intrinsic and Extrinsic Tracers for Tracking Water Quality Changes During Managed Aquifer Recharge. In Managed Aquifer Recharge in the...
Identification of Paleowater in California Drinking Water Wells

Urban Environment (Groundwater Resources Association of California; May 22-23, 2013; Burlingame, CA). Available at: http://www.grac.org/aquiferrecharge.asp.


