Stratigraphic and structural controls on groundwater salinity variations in the Poso Creek Oil Field, Kern County, California, USA

Michael J. Stephens¹ · David H. Shimabukuro² · Will Chang³ · Janice M. Gillespie¹ · Zack Levinson²

Received: 22 July 2020 / Accepted: 6 July 2021 / Published online: 18 September 2021

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2021, corrected publication 2021

Abstract

Groundwater total dissolved solids (TDS) distribution was mapped with a three-dimensional (3D) model, and it was found that TDS variability is largely controlled by stratigraphy and geologic structure. General TDS patterns in the San Joaquin Valley of California (USA) are attributed to predominantly connate water composition and large-scale recharge from the adjacent Sierra Nevada. However, in smaller areas, stratigraphy and faulting play an important role in controlling TDS. Here, the relationship of stratigraphy and structure to TDS concentration was examined at Poso Creek Oil Field, Kern County, California. The TDS model was constructed using produced water TDS samples and borehole geophysics. The model was used to predict TDS concentration at discrete locations in 3D space and used a Gaussian process to interpolate TDS over a volume. In the overlying aquifer, TDS is typically <1,000 mg/L and increases with depth to ~1,200–3,500 mg/L in the hydrocarbon zone below the Macoma claystone—a regionally extensive, fine-grained unit—and reaches ~7,000 mg/L in isolated places. The Macoma claystone creates a vertical TDS gradient in the west where it is thickest, but control decreases to the east where it pinches out and allows freshwater recharge. Previously mapped normal faults were found to exhibit inconsistent control on TDS. In one case, high-density faulting appears to prevent recharge from flushing higher-TDS connate water. Elsewhere, the high-throw segments of a normal fault exhibit variable behavior, in places blocking lower-TDS recharge and in other cases allowing flushing. Importantly, faults apparently have differential control on oil and groundwater.

Keywords Groundwater salinity · Geophysical methods · Oil and gas · Stratigraphy · USA

Introduction

While there has long been interest in the distribution of shallower fresh—<1,000 mg/L total dissolved solids (TDS)—groundwater in California, USA (Page 1973; Hansen et al. 2018), there has recently been interest in mapping deeper brackish (1,000–10,000 mg/L TDS) groundwater throughout California (Kang and Jackson 2016; Gillespie et al. 2017; Kang et al. 2020). There is a focus in mapping TDS near San Joaquin Valley (SJV) oil fields, where oil

production and agricultural water usage exist in proximity (Taylor et al. 2014; Davis et al. 2018). These efforts, which determine TDS concentration using a mix of geochemical and geophysical approaches, have focused both at the basin scale (Kang and Jackson 2016; Gillespie et al. 2017; Metzger and Landon 2018; McMahon et al. 2018; Kang et al. 2020) and in specific oil fields (Gillespie et al. 2019; Stephens et al. 2019; Ball et al. 2020).

Processes creating the regional pattern of groundwater TDS concentration in the SJV are well understood. Unlike many sedimentary basins, the SJV does not have buried evaporites, making processes like the initial composition of pore fluids, meteoric water recharge, and water–rock interactions the primary factors in developing the basin wide vertical and lateral TDS structure (Ramseyer and Boles 1986; Wilson et al. 1999; Fisher and Boles 1990; Kharaka and Hanor 2004; Gillespie et al. 2017; McMahon et al. 2018). These processes are modified by stratigraphy containing extensive low-permeability clays and shales (Faunt 2009; Scheirer and Magoon 2007) that can compartmentalize TDS by inhibiting

✉ Michael J. Stephens
mjstephens@usgs.gov

¹ US Geological Survey, California Water Science Center, 6000 J Street, Sacramento, CA 95819, USA

² Department of Geology, California State University, Sacramento, 6000 J Street, Sacramento, CA 95819, USA

³ Hypergradient LLC, Berkeley, CA 94703, USA

meteoric water recharge pathways. All these processes combine to produce a TDS depth profile that generally increases with depth and that is fresher toward the eastern margin of the SJV (Kharaka and Hanor 2004; Gillespie et al. 2017).

At the scale of specific oil fields within the basin, the same factors that create regional TDS distributions occur, but the influence of stratigraphy and structure becomes more important. Confining stratigraphy can separate volumes of groundwater with different TDS concentrations, leading to vertical TDS gradients across clays (Kharaka and Berry 1974; Gillespie et al. 2017, 2019; Stephens et al. 2019). Lateral thinning of confining clays at the edge of the basin or adjacent to syndepositional folded uplifts are important in enabling recharge pathways that modify TDS (Gillespie et al. 2017, 2019).

It is not as clear how structural elements such as faults affect oil-field scale TDS. Regulatory proposals for subsurface injection in California often assume faults seal lateral and vertical fluid flows (Schloss 2017a, b), whereas the fault zone hydrogeology literature describes the wide variability of fault behavior in impeding or enhancing fluid flow (Caine et al. 1996; Evans et al. 1997; Bense and Person 2006; Fossen et al. 2007; Faulkner et al. 2010; Bense et al. 2013). Presumably faults exhibit a similar behavior with respect to TDS and can act as impermeable structures that confine TDS (Stephens et al. 2019) or as permeable structures that allow downward migration of meteoric water or upward migration of deeper fluids.

In the present study, a three-dimensional (3D) TDS model of the Poso Creek Oil Field located within the SJV (Fig. 1) was used to explore the oil-field scale stratigraphic and structural controls on TDS. First, a TDS model was constructed from historical produced water (water extracted with oil) sample TDS and borehole geophysics using a novel technique that estimates TDS using Archie's Law (Archie 1942) and models TDS spatially at once. This TDS volume was then compared to a new structural and stratigraphic model constructed from borehole geophysics to examine the role that clay layers and faults have in controlling the TDS distribution in the field.

The Poso Creek Oil Field was chosen for this study because of the proximity between oil and gas production and potentially useable groundwater resources. The mature oil fields of California often have high water-to-oil ratios (Clark and Veil 2009), resulting in up to 17 barrels of produced water extracted for every barrel of oil produced. Produced water is often high in TDS and contains inorganic and organic compounds associated with hydrocarbons that often limit its practical use (Clark and Veil 2009). Therefore, produced water is regularly injected back into hydrocarbon reservoirs during water/steam flood projects, or it is injected into the subsurface for disposal purposes (CalGEM 2019a). Understanding the factors that control groundwater TDS in oil fields can help

explain issues regarding potential migration of this injected water, displacement of oil field formation water during enhanced oil recovery (EOR), containment of water disposal, and the movement of water from oil-producing formations resulting from natural geologic processes (Kharaka and Hanor 2004; Taylor et al. 2014; Jordan and Gillespie 2016; Gillespie et al. 2019; McMahon et al. 2018; Wright et al. 2019).

The Poso Creek Oil Field was among fields selected for regional groundwater mapping and monitoring by the California State Water Resources Control Board (Water Board) as part of the Oil and Gas Regional Monitoring Program (RMP), conducted in cooperation with the US Geological Survey (USGS; California State Water Resources Control Board 2015, 2019; Davis et al. 2018; US Geological Survey 2020). The US Bureau of Land Management also identified a need to delineate the distribution of brackish groundwater in the Poso Creek Oil Field and is working in cooperation with the USGS on the mapping effort. More generally, there is also an increased interest in use of brackish and produced water for irrigation and water supply in California (Cooley et al. 2006; Faunt 2009; Leitz and Boegli 2011; EMWD 2013). In Kern County, some produced water is treated (separated from oil and filtered), blended with groundwater and used for irrigating nearby agriculture in several water districts such as in the Cawelo Water District (Kondash et al. 2020).

Geologic background

The Poso Creek Oil Field is located on the eastern side of the southern SJV in the Central Valley of California (Fig. 1). The SJV is bounded to the west by the Coast Ranges and the Sierra Nevada to the east. Sedimentary rocks in the SJV reach >7,500 m (24,600 ft) to basement and range from Cretaceous to Quaternary in age (Fig. 2). Sedimentation is controlled by sea level changes, regional tectonics, and sediment supply (Callaway 1990; Scheirer and Magoon 2007). The tectonic setting in the SJV transitioned from a forearc setting above a subduction zone to a transform setting with the initiation of the San Andreas Fault ~30 ma (Atwater 1970; Fig. 1b). During Miocene time, sedimentation along the west side of the SJV was from the predominantly granitoid Salinian block, while sedimentation along the east side was sourced from the Sierra Nevada (Reid 1995). Throughout the Miocene, transgressive–regressive sequences, controlled by intermittent basin-sea communication, laid down sediments throughout the SJV including diatomaceous sediments (primarily on the west side). Sedimentation in the axis of the southern SJV is characterized by a change in depositional environment from deep water turbidites and pelagic sediments in Miocene time to fluvial and lacustrine deposition during the

Fig. 1 a–c The Poso Creek Oil Field is in the southeastern San Joaquin Valley and is divided into three administrative areas: Premier, Enas, and McVan. Fault locations are from Weddle (1959), well locations and field boundary data are from CalGEM (2019c)

Pleistocene (Bartow and Pittman 1983; Bartow 1991; Reid 1995). These sediments were carried into the basin from the Sierra Nevada to the south and east.

The stratigraphy of the Poso Creek Oil Field consists of sands, clays, and conglomerates that range from Eocene to Pleistocene in age (Fig. 2; Fig. S1 of the electronic supplementary material (ESM; Weddle 1959; Scheirer and Magoon 2007). The deepest formation involved in oilfield activities is the upper Miocene Santa Margarita Sandstone (Fig. S2 of ESM). This unit, interpreted as a shoreline sand, is not a major oil producer in the Poso Creek Oil Field and is typically used for produced water disposal. Conformably overlying the Santa Margarita is the nonmarine Chanac Formation. The Chanac Formation is composed of interbedded sands and shales and is an important reservoir for oil production in the field. The Etchegoin Formation unconformably overlies the Chanac Formation and is split into three units: the basal Etchegoin sands, the Macoma claystone, and the upper Etchegoin. The basal Etchegoin sands are ~30 m (100 ft) thick, thin to the east, and are a significant oil reservoir. The basal Etchegoin sands are overlain by the regionally extensive, transgressive marine Macoma claystone. The Macoma claystone is ~35 m (115 ft) thick inside the field boundary and plays an important role restricting vertical fluid flow and separates the overlying aquifer system from the hydrocarbon zone below. The marine sands and shales of the upper Etchegoin Formation grade upward into the Miocene-Pliocene (to possibly Pleistocene)

Kern River Formation (Bartow and Pittman 1983). The Kern River Formation also grades basinward into the San Joaquin and Tulare formations. The Kern River Formation is composed of sands, clays, and conglomerates. There is little lateral lithologic continuity within the Kern River Formation and is thought to be deposited in a braided stream environment (Bartow and Pittman 1983).

The structure in the Poso Creek Oil Field is a west-dipping homocline modified by numerous faults. Many of these faults have a normal sense of displacement, trend roughly north–south, and are down to the west-southwest; however, there are normal faults down to the east, and strike-slip faults mapped in the area (i.e., East McVan fault; Weddle 1959; Saleeby and Saleeby 2019). The faulting in this region is thought to be linked to stresses from the Garlock fault (Bartow 1991) or related to extension of overthickened Sierra Nevada crust (Saleeby et al. 2013; Sousa et al. 2016).

In the Poso Creek Oil Field area, the Pond–Poso Creek fault strikes northwest and is southwest-down. Bartow (1991), plate 1 separates this fault into two segments: (1) the Poso Creek fault to the east can be mapped at the surface through Tertiary rock outcrops and trends northwest-southeast (Fig. 1), and (2) the Pond fault to the west in the subsurface, which trends more north outside the oil field boundary (not shown on Fig. 1). This study is concerned with the Poso Creek fault segment. Another important normal fault within the field is the Premier fault, which trends north and is

Fig. 2 Stratigraphic column of the southern San Joaquin Valley, modified from Scheirer and Magoon (2007)

down to the west. Across all but the southern part of the oil field, this fault bifurcates into two segments. Weddle (1959) labels the western fault segment the Premier fault, whereas plate 1 from Castle et al. (1983) labels the eastern fault segment the Premier fault. To avoid confusion, this study refers to the western and eastern fault segments as the West Premier fault and East Premier fault, respectively (Fig. 1c).

The oil field is split into three administrative areas (Fig. 1c; Weddle 1959; Welge 1973). The Premier area (main area) is the largest in the field and is bounded to the north by the Poso

Creek fault. The Enas area lies just north of the Poso Creek fault and the McVan area occupies the northeast portion of the field. The McVan area is not considered in this study because of limited data availability.

The Sierra Nevada recharge provides a widespread fresh-water source for the eastern margin of the SJV (Faunt 2009; Gillespie et al. 2017). As a result, groundwater in the eastern SJV is commonly significantly lower in TDS compared to the western SJV, which only receives small amounts of recharge from ephemeral streams draining metamorphic and marine rocks of the Coast Ranges (Laudon and Belitz 1991; Gillespie et al. 2017, 2019). The major aquifer in the study area is the Kern River Formation and, potentially, the upper Etchegoin Formation. It receives most of its recharge from rainfall and snowmelt originating from the Sierra Nevada to the east, and minor recharge from irrigation returns.

Methods

This section describes the data and TDS/clay/fault mapping techniques used in this study. First, the data used to create the TDS model are described. The TDS model incorporates formation resistivity derived from borehole geophysics, output values from porosity and temperature models, and laboratory measurements of TDS from produced water samples. Archie's relationship (Archie 1942) was used to convert the observed formation resistivity to water resistivity and then TDS predictions were made using an empirical power law relationship between laboratory-measured resistivity and TDS from the produced water samples. After TDS predictions were made at the locations of collected formation resistivity, TDS was interpolated over a volume using a Gaussian process. Lastly, the geophysical and stratigraphic data available to map the Macoma claystone, which also were used to define faults in the Poso Creek and Premier fault systems, are described.

Data for TDS model

The TDS model relies on the following five datasets which are available from Stephens et al. (2021):

Dataset A. TDS measurements from produced water samples from 51 oil wells

Dataset B. Formation resistivity from 144 oil well geophysical logs

Dataset C. Porosity readings with depth from 14 oil well geophysical logs

Dataset D. Temperature measurements with depth from 83 oil wells

Dataset E. Borehole logs for clay/fault mapping from 551 oil wells

The TDS and borehole resistivity data used in this study are from before 1980, primarily from the early 1960s, and represent natural TDS conditions, not modern, post-injection conditions. Borehole resistivity and TDS data collected after this time may be affected by steam injection activities that would necessitate other methods to consider the effects of steam injection on borehole log interpretation, which is not the subject of this paper. Thus, the TDS model represents an estimation of pre-injection baseline conditions.

Dataset A: produced water TDS

TDS measurements taken from produced water samples were used to map TDS and in setting the TDS model parameters. These data were compiled from scanned copies found on the California Geologic Energy Management Division (CalGEM; formerly named California Division of Oil, Gas, and Geothermal Resources or DOGGR) online archive of historical produced water chemical analyses (CalGEM 2016). The historical produced water TDS data originated from laboratory analyses that were reported by the oil field operators.

This dataset consists of 51 TDS values, each attributed with x (longitude), y (latitude) denoting the well location, z (elevation relative to sea level), and a lab-measured water resistivity. The elevation (z) was calculated from the upper-most perforated depth found either on the lab analysis report or the well record from the CalGEM Well Search website (CalGEM 2019b). Only a subset of the TDS measurements available were considered appropriate to use in the TDS model for a variety of reasons discussed in section S1 of ESM. One of the reasons TDS measurements were rejected was because the sample TDS was too low. The TDS model systematically overpredicted produced water TDS values <650 mg/L. It is unclear exactly why the model overpredicted these samples; however, it is hypothesized that when TDS is this low, clay content plays a relatively larger role in the bulk resistivity response, leading to overprediction. However, TDS this low should not impact interpretations because TDS below 650 mg/L exists virtually only in the overlying aquifer and not below the Macoma claystone where this study focuses. This is also the reason TDS measurements from groundwater wells overlying and adjacent to the study area (all of which are <650 mg/L; Metzger et al. 2018) were not used in the model.

Dataset B: formation resistivity

CalGEM provides an online database that contains scans of borehole geophysical well logs, well records, and other information reported about oil and gas wells (CalGEM 2019b). The well logs typically have multiple data readouts with depth including spontaneous potential (SP), deep- and shallow-reading resistivity, and on some newer logs (post ~1980s)

neutron and density porosity readings. From the continuous deep-reading formation resistivity data, values that correspond to “clean” or, clay- and hydrocarbon-free sands, were carefully chosen to minimize the impact of electrical charges found in clays and resistivity effects from hydrocarbons. When the uninvaded (by drilling mud) zone formation resistivity (R_t) is 100% water-saturated (hydrocarbon-free), it is denoted R_o . Clean sand zones were found by analyzing the separation between short and deep resistivity (interpreted to correspond to permeable sands, see Asquith and Krygowski 2004; Ellis and Singer 2007), referring to deflections in the SP curve, and consulting core analyses and driller’s notes when available in the well record. After these considerations, the filtered formation resistivity data should reflect only the formation water and clay-free rock as best as possible. Virtually all the resistivity data collected are from the early 1960s to ensure the data reflect original conditions and were not affected by injection operations—particularly steam injection. Steamed zones were intentionally avoided for the few data collected from wells drilled after steaming operations started (~1964). The resistivity dataset consists of 2,257 points of R_o , each with x (longitude) and y (latitude) denoting the well location, and z (elevation). The locations of datasets A and B are visualized in the interactive 3D file in ESM.

Dataset C: porosity readings with depth

Formation resistivity is related to formation water resistivity (R_w) via Archie’s Equation (Archie 1942). This equation requires knowing porosity at the location of the formation resistivity; however, porosity borehole logs were not common until ~1980s, but the formation resistivity data are almost exclusively from before that time. Therefore, 14 newer well logs with neutron porosity, density porosity, and caliper (borehole diameter) data were digitized to create a model of porosity with depth in the study area. The model is discussed in section ‘Porosity model’.

Dataset D: temperature measurements with depth

The temperature at the location of the formation resistivity is needed to relate R_w from Archie’s Equation to TDS. Like porosity logs, older (~1960s) temperature logs are not readily available in the study area. However, a temperature measurement is recorded on the well header of the well logs that corresponds to the bottom depth of the well. These data were used to construct a general model of the relation between temperature and depth in the study area. This dataset consists of temperature and the associated depths from 83 wells throughout the oil field. This model is discussed in section ‘Temperature model’.

Dataset E: borehole logs for clay/fault mapping

From 551 geophysical well logs from the study area, the top and base of the Macoma claystone were identified by analyzing the SP, resistivity, porosity curves, and driller's notes when available. This dataset consists of the Macoma claystone depth values (measured from the kelly bushing or derrick floor), the elevation of the kelly bushing or derrick floor, x (longitude), and y (latitude). Similar data were collected to construct a map of the top of the Santa Margarita.

Modeling

This subsection describes how TDS was estimated and interpolated. The TDS model created here aims to improve upon established TDS mapping methods (Hamlin and de la Rocha 2015; Kang and Jackson 2016; Gillespie et al. 2017, 2019; Stephens et al. 2019; Kang et al. 2020) by modeling measurements of TDS and geophysical observations simultaneously to create a 3D continuous TDS model. The clay layer/fault mapping is also described.

Porosity model

The porosity model predicts porosity, given depth within the oil field (Fig. 3a). The porosity data were filtered before fitting the model, so only sand units free of hydrocarbons were considered. First, bulk density readings were converted to density porosity using methods described in Asquith and Krygowski (2004):

$$\phi_D = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \quad (1)$$

where the variables are as follows:

ϕ_D	Density derived porosity (dimensionless)
ρ_{ma}	Rock matrix density, assumed to be 2.65 (g/cm ³)
ρ_b	Measured bulk formation density (g/cm ³)
ρ_f	Fluid density, assumed to be 1 (g/cm ³)

Next, the caliper data were used to filter the continuous porosity readings by removing intervals where the caliper was more than 1.27 cm (0.5 in) larger than the bit size. When the caliper reads much larger than the bit size, it is assumed the hole collapsed and the logging tools were not in contact with the borehole wall, causing the porosity readings in these intervals to be inaccurate (Asquith and Krygowski 2004; Ellis and Singer 2007). The presence of hydrocarbons, especially in the gas phase, will cause the density porosity tool to read higher than the neutron tool; this condition is known as “crossover” and these intervals were removed because they do not represent rock properties. Porosity data over 0.42 were

also removed. Even in ideal cubic sandstone grain packing, porosities cannot exceed 0.48 (Fetter 2001) and in practice, the upper bound is 0.42 (Jin et al. 2009). Porosities over 0.42 are likely influenced by some combination of clay-bound water, hydrocarbons, or non-sand lithologies. Next, depths were identified where sands were most likely to be clean where the neutron and density readings were within 2% of one another, yielding 685 density porosity points (Stephens et al. 2019). Lastly, density and neutron porosity readings were combined into a composite porosity ϕ_{N-D} using the root mean square formula (Asquith and Krygowski 2004):

$$\phi_{N-D} = \sqrt{\frac{\phi_N^2 + \phi_D^2}{2}} \quad (2)$$

To ensure that uncertainty in porosity observations is reflected in the TDS model, porosity was modeled as a random variable using a Bayesian linear regression:

$$\mathbf{p}_C = G_C^T \mathbf{q} + \boldsymbol{\varepsilon}$$

where vector $\mathbf{p}_C$ contains N_C porosity observations in the form of $\log(\phi_{N-D})$, and G_C is a $2 \times N_C$ matrix that featurizes porosity data locations, with ones in the first row and the depths in the second row. The subscripts of C are meant to evoke the connection between the variable and Dataset C. Random variable $\mathbf{q}$ is a two-element vector of coefficients with an uninformative, improper prior distribution $\mathbf{q} \sim \mathcal{N}(0, \infty \cdot I)$; and $\boldsymbol{\varepsilon}$ is a vector of uncorrelated noise, of the same length as $\mathbf{p}_C$ with prior distribution $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^{(\varepsilon)2} I)$, where I is the identity matrix and $\sigma^{(\varepsilon)2}$ is a parameter that is fitted by maximizing likelihood $p(\mathbf{p}_C)$. This model has a posterior distribution

$$\mathbf{q} \mid \mathbf{p}_C \sim \mathcal{N}(\boldsymbol{\mu}^{(q)}, \Sigma^{(q)}) \quad (3)$$

where

$$\begin{aligned} \boldsymbol{\mu}^{(q)} &= (G_C G_C^T)^{-1} G_C \mathbf{p}_C \\ \Sigma^{(q)} &= \sigma^{(\varepsilon)2} (G_C G_C^T)^{-1} \end{aligned}$$

These coefficients are used to model unobserved porosity values $\mathbf{p}$ at new locations, featurized by G . The posterior predictive distribution for $\mathbf{p}$ is

$$\mathbf{p} \mid \mathbf{p}_C \sim \mathcal{N}(\boldsymbol{\mu}^{(p)}, \Sigma^{(p)}) \quad (4)$$

where

$$\begin{aligned} \boldsymbol{\mu}^{(p)} &= G^T \boldsymbol{\mu}^{(q)} \\ \Sigma^{(p)} &= G^T \Sigma^{(q)} G + \sigma^{(\varepsilon)2} I \end{aligned}$$

Fig. 3 **a** Porosity is needed to make total dissolved solids (TDS) predictions from well logs; however, most logs in the study area do not have porosity curves so a depth-dependent porosity model is used instead. The porosity data from well logs (blue) exist only in the shallower zones. However, some of the resistivity data are derived from deeper zones (extent shown with the red regression line). Porosity was modeled using a Bayesian linear regression to incorporate uncertainty (red-shaded area) into the TDS model. **b** Similarly, historical temperature curves are not common so temperatures from well log headers were fit with a linear model to use throughout the field

This posterior predictive distribution is used to propagate what is known about porosity via the resistivity model (section ‘Resistivity model’) into the TDS model (section ‘TDS model’).

Temperature model

The temperature model provides temperature predictions with depth in the oil field (Fig. 3b). Dataset D ($n = 95$) was fit with a linear regression to create the temperature model that was used to provide a temperature estimate to accompany the formation resistivity data in dataset B (Gillespie et al. 2019; Stephens et al. 2019). This model represents the pre-injection, natural geothermal gradient by only using data not affected by water and steam injection.

Resistivity model

This model relates the TDS of groundwater to formation resistivity by chaining together three equations which relate:

- TDS to R_{w75} , which is the resistivity of groundwater at 75 °F (24 °C)
- R_{w75} to R_w , which is the resistivity of groundwater at the temperature at which it occurs
- R_w to R_o , which is formation resistivity

The first relationship depends on the ionic composition of groundwater, which tends to vary with TDS. Using the lab-measured water resistivity and TDS in dataset A, an empirical power law relationship between TDS and resistivity that is specific to Poso Creek Oil Field was derived:

$$\log R_{w75} = c_0 + c_1 \log \text{TDS} \quad (5)$$

The relationship between lab-measured $\log R_{w75}$ and $\log \text{TDS}$ was fit with linear regression (Fig. 4). Note, $\log$ represents the natural logarithm (base e) throughout this work.

The second relationship is between R_{w75} and R_w (Asquith and Krygowski 2004):

$$\log R_w = (\log R_{w75}) + \log(75 + 6.77) - \log(T + 6.77) \quad (6)$$

where T is the temperature of the groundwater as given by the temperature model.

The third relationship is Archie’s Law (Archie 1942; Winsauer et al. 1952):

$$\log R_o = (\log R_w) + (\log a) - (m \log \phi_{N-D}) \quad (7)$$

where tortuosity factor a and cementation factor m are assumed to be constant throughout the oil field, and ϕ_{N-D} is estimated porosity at the location of the groundwater. The three equations were combined to derive an affine relationship between $\log R_o$ and $\log \text{TDS}$:

$$\log R_o = c_0 + \log(75 + 6.77) - \log(T + 6.77) + (\log a) - (m \log \phi_{N-D}) + c_1 \log \text{TDS} \quad (8)$$

Note that c_0 , c_1 , a , and m are field-wide parameters, whereas R_o , T , ϕ_{N-D} , and TDS are variables for which there is one value per data location. For notational clarity each location-specific variable is transformed and a subscript indexing the location is applied to each:

$$\begin{aligned} r_i &= \log(R_o)_i \\ t_i &= \log(75 + 6.77) - \log(T_i + 6.77) \\ p_i &= \log(\phi_{N-D})_i \\ u_i &= \log \text{TDS}_i \end{aligned}$$

Then,

$$r_i = c_0 + t_i + (\log a) - mp_i + c_1 u_i \quad (9)$$

TDS model

Underlying the TDS model is a Gaussian process, which models spatially correlated random quantities (Rasmussen and Williams 2006). In this case the Gaussian process jointly models the N_A produced water TDS measurements of dataset A, and the N_B clean sand resistivity readings of dataset B, thus accounting for $N = N_A + N_B$ observations. After the model parameters were fitted, the model was used to produce conditional probability distributions of TDS at grid locations in the volume of analysis. These predictions were then visualized as maps.

For convenience, the observations are ordered so that $1, \dots, N_A$ refer to TDS measurements and $N_A + 1, \dots, N$ refer to resistivity readings. Let $\mathbf{u}$, $\mathbf{r}$, $\mathbf{v}$ be vectors of N elements each that denote log TDS, log resistivity, and actual observations, respectively, at the N observation locations, so that

$$v_i = \begin{cases} u_i & \text{for } i \in \{1, \dots, N_A\} \\ r_i & \text{for } i \in \{N_A + 1, \dots, N\} \end{cases}$$

Now probability distributions will be constructed for $\mathbf{u}$, $\mathbf{r}$, and $\mathbf{v}$, each in turn.

Modeling TDS values TDS observations $\mathbf{u}$ are modeled by a Gaussian process with a non-stationary mean to account for large-scale TDS trends. The joint distribution for log TDS is

$$\mathbf{u} | \mathbf{w} \sim \mathcal{N}(\boldsymbol{\mu}^{(u|w)}, \Sigma^{(u|w)}) \quad (10)$$

Fig. 4 The power law relationship between measured water resistivity and total dissolved solids (TDS) from produced water samples. The linear model here (red line) is used to relate equation-based estimates of formation water resistivity to TDS

where the trend

$$\boldsymbol{\mu}^{(u|w)} = F^T \mathbf{w} \quad (11)$$

is derived from a vector of parameters $\mathbf{w}$ and a $6 \times N$ feature matrix F where each column i is a featurization of the location (x_i, y_i, z_i) of observation i . In particular, each column i is $(1, x_i, y_i, z_i, z_i^2, z_i^3)^T$, which allows for the modeling of spatial trends using these polynomial basis functions. To avoid overfitting to the data, $\mathbf{w}$ was assigned a diffuse prior and was marginalized over:

$$\mathbf{w} \sim \mathcal{N}(0, \alpha I) \quad (12)$$

where I is the identity matrix, and α is set to 10,000, which is large enough, since x , y , and z distances are in units of kilometers and log TDS has a full range of around 5.

In Eq. (10), the covariance matrix $\Sigma^{(u|w)}$ is constructed using a gamma-exponential covariance function. Each entry is

$$\sigma_{ij}^{(u|w)} = s \cdot \exp \left[- \left(\frac{(x_i - x_j) + (y_i - y_j) + k(z_i - z_j)}{r} \right)^\gamma \right] + \begin{cases} 0 & \text{if } i \neq j, \\ n_A & \text{if } i = j, i \leq N_A \\ n_B & \text{if } i = j, i > N_A \end{cases}$$

where the parameters are (to use terms from kriging) sill s , range r , anisotropy k , shape γ , and two nuggets: n_A for produced water measurements, and n_B for resistivity readings. Shape γ has bounds $0 < \gamma \leq 2$, range must be positive, and the other parameters must be non-negative.

Combining Eqs. (10)–(12) and marginalizing over $\mathbf{w}$ yields

$$\mathbf{u} \sim \mathcal{N}(0, \Sigma^{(u)}) \quad (13)$$

where

$$\Sigma^{(u)} = \alpha F^T F + \Sigma^{(u|w)}.$$

Modeling log resistivity A probability distribution for log resistivity $\mathbf{r}$ can be derived using Eq. (9), which is restated here using vector notation:

$$\mathbf{r} = c_0 + \mathbf{t} + (\log a) - m\mathbf{p} + c_1 \mathbf{u} \quad (14)$$

Each vector has N elements, corresponding to N observations. For $\mathbf{r}$, $\mathbf{t}$, and $\mathbf{p}$ (derived from resistivity, temperature, and porosity), the first N_A elements are actually irrelevant but are included for notational convenience. Pulling in Eqs. (4) and (13) for $\mathbf{p}$ and $\mathbf{u}$, one obtains

$$\mathbf{r} | \mathbf{p}_C \sim \mathcal{N}(\boldsymbol{\mu}^{(r)}, \Sigma^{(r)}) \quad (15)$$

where

$$\mu^{(r)} = c_0 + (\log a) + t - m\mu^{(p)}$$

$$\text{and } \Sigma^{(r)} = m^2 \Sigma^{(p)} + c_1^2 \Sigma^{(u)}$$

Modeling mixed observations A probability distribution for a mixed set of observations $\mathbf{v}$ can be derived by noting that mean and covariance are $\mu^{(u)}$ and $\Sigma^{(u)}$ for the first N_A observations, and $\mu^{(r)}$ and $\Sigma^{(r)}$ for the next N_B observations. To complete the probability distribution, one must derive the covariance between u_i and r_j when $i \leq N_A$ and $j > N_A$, which is $\text{Cov}(u_i, r_j) = c_1 \sigma_{ij}^{(u)}$. Thus,

$$\mathbf{v} | \mathbf{p}_C \sim \mathcal{N}(\mu^{(v)}, \Sigma^{(v)}) \quad (16)$$

where

$$\mu_i^{(v)} = \begin{cases} \mu_i^{(u)} & \text{for } i \leq N_A \\ \mu_i^{(r)} & \text{for } i > N_A \end{cases}$$

$$\sigma_{ij}^{(v)} = \begin{cases} \sigma_{ij}^{(u)} & \text{for } i \leq N_A, j \leq N_A \\ \sigma_{ij}^{(r)} & \text{for } i > N_A, j > N_A \\ c_1 \sigma_{ij}^{(u)} & \text{otherwise} \end{cases}$$

Estimation of parameters Equation (16) yields a likelihood function in terms of the model's eight scalar parameters: Archie parameter a is embedded in $\mu^{(v)}$; geostatistical parameters s , r , n_A , n_B , k , and γ , are in $\Sigma^{(v)}$; and Archie parameter m is in both. The likelihood function was written in TensorFlow (Abadi et al. 2015) and maximized using the Adam optimizer (Kingma and Ba 2014). Results are shown in Table 1. TDS model performance was characterized by using only the TDS estimations at the clean sand resistivity locations to predict TDS at the produced water TDS locations (dataset A). The average root-mean-squared error for log TDS is 0.24 (Fig. 5); therefore, one-sigma relative prediction error for TDS is $e^{0.24} - 1 \approx 27\%$.

Table 1 Results of total dissolved solids (TDS) model parameter optimization. The units of the range are in kilometers

Parameter	Value
Archie a	0.936
Archie m	1.47
Shape γ	0.762
Range r	1.44
Sill s	0.205
Resistivity nugget n_B	0.0010
Produced water nugget n_A	0.0290
Anisotropy k	4.40

Model discussion This TDS model contains advances over some previous work in the modeling of environmental variables. Modeling petrophysical and geochemical data jointly has many advantages over a typical multistep process such as the following:

1. Work out the relationship between resistivity and TDS by estimating petrophysical parameters
2. Compute TDS at locations where there are resistivity measurements
3. Estimate field-wide TDS trends
4. Detrend the data and then estimate geostatistical parameters by creating experimental variograms

Jointly fitting petrophysical and geostatistical parameters avoids parameterizations that are suboptimal with respect to the complete data. Not converting each resistivity measurement into a TDS value avoids the possibility that the conversion denoises the data in the process, which would cause the TDS model to be more certain about TDS than warranted. The trend model developed here avoids overfitting by integrating over trend parameters (Eq. 12), which supersedes traditional universal kriging models that may overfit when trend models are complex. Finally, incorporating the sill s , range r , shape γ , and nuggets n_A , n_B into a likelihood function avoids the complications of fitting an experimental variogram, which entails first detrending the data, while any discernible trend might be quite equivocal (Gringarten and Deutsch 2001).

Clay and fault mapping

Using dataset E, maps of the Macoma claystone and faults were constructed. Maps of the elevation of the base and thickness are used to describe the structure and stratigraphy of the Macoma claystone. More detailed maps highlighting the offsets of the base of the Macoma claystone were used to identify faults. The discontinuities in elevation of the base Macoma claystone are interpreted to represent fault planes in the Poso Creek and Premier fault systems. The depths derived from the well logs (which are measured from the kelly bushing or derrick floor) were converted to elevation by subtracting the elevation of the kelly bushing or derrick floor.

The previously mapped faults were georeferenced from structural contour maps on the base of the Macoma claystone, which is also the top of the basal Etchegoin Formation, from Weddle (1959, plate II). Despite fewer wells being available at the time these faults were mapped, the general fault locations agree with data collected in this study.

The elevations of the top and base of the Macoma claystone were also superimposed onto the TDS cross sections described in the next section 'Results and discussion'. A universal kriging model with a linear variogram and linear trend model was used to produce estimations of the top and base

Fig. 5 To quantify model performance, the total dissolved solids (TDS) model was used to predict TDS (y-axis) at the locations of the measurements of TDS from produced water samples (x-axis). Average root-mean-square error (RMSE) is 0.24 for log TDS

Macoma claystone at the locations of the TDS cross sections using dataset E. The Macoma claystone top and base are shown on the TDS cross sections for visualizing the location and thickness of the Macoma claystone relative to TDS distributions and not for detailed faulting interpretation. The smoothing effect from the kriging model on the abrupt offsets from the faults should be noted. The fault dip of the West Premier fault was calculated to be 49° using a well that intersected the fault in the original hole as well as two subsequent redrills (API: 040291489) and is the lowest angle fault in the field. A fault dip of 85° was assumed for all other faults, which is consistent with Weddle (1959).

Results and discussion

The TDS model was used to interpolate over the volume of analysis to visualize groundwater TDS. Maps of the regional Macoma claystone and fault maps are shown to characterize zones where groundwater flow may be restricted and where flow paths may exist in the Poso Creek and Premier faults systems.

Deep groundwater TDS trends are interpreted as a proxy for long-term groundwater movement pathways. Large amounts of freshwater recharge originating from the Sierra Nevada over geologic time accounts for the generally low TDS ($<10,000$ mg/L) in the Poso Creek area (Faunt 2009; Gillespie et al. 2017; McMahon et al. 2018). The presence of biodegraded oil along the eastern margin of the SJV coupled with evidence of shallow isotopic signatures in deep ($\sim 1,500$ m) groundwater indicate that freshwater recharge has

replaced and/or mixed with higher TDS formation water (Fisher and Boles 1990). The dominant groundwater flow paths from higher elevation recharge areas contain lower-TDS water than where flow of fresh groundwater is inhibited. These flow paths are partially controlled by the Macoma claystone, which can restrict vertical groundwater flow, and the fault systems within the oil field which may control lateral variability. It is shown that groundwater TDS variations are collocated with changes in lithology and fault system geometry—consistent with previous conceptual models of hydrogeology (Fetter 2001; Faulkner et al. 2010; Bense et al. 2013).

Fig. 6 **a** Map showing the elevation of the base of the Macoma claystone in the Poso Creek Oil Field area. The Macoma claystone is a significant regional clay layer that restricts vertical groundwater flow and separates the overlying aquifer system from the hydrocarbon-bearing zone below. **b** Isochore map showing the thickness of the Macoma claystone. The Macoma claystone is thickest to the west and thins and becomes sandier to the east. Faults shown are from Weddle (1959), well locations and field boundary data from CalGEM (2019c), background from ESRI

The Macoma claystone

The Macoma claystone separates the overlying aquifer zone from the hydrocarbon-bearing zone. The Macoma claystone, along with faulting, plays an important role in trapping hydrocarbons in the field (Weddle 1959). The Macoma claystone dips to the west-southwest and is mapped in the producing zones at approximately -300 to -600 m (-980 to $-2,000$ ft; elevation relative to sea level) and to approximately $-1,400$ m ($-4,600$ ft) west of the oil field in the groundwater basin (Fig. 6a). The Macoma claystone also thins and becomes sandier to the northeast, eventually pinching out at elevations near sea level in the McVan area (Fig. 6b). The Macoma claystone thickens to the west to over 140 m (460 ft) outside the oil field boundary. Thickness is rather uniform in the oil producing zones inside the field at approximately 20 – 40 m (65 – 130 ft), except in the northwest portion of the field near the Poso Creek fault where it is >40 m (130 ft) thick. Note that contoured maps (Fig. 6) at this scale do not show the offsets in the Macoma claystone along the faults.

Faulting

A detailed view of the elevation of the base of the Macoma claystone is shown in Fig. 7. The abrupt discontinuities in

the elevation of the base of the Macoma claystone are interpreted to correspond to normal-sense fault throw along a fault plane. In the Poso Creek fault zone (Fig. 7a), the previously mapped faults generally agree with inferred fault locations in the newly compiled data except where the West Premier fault intersects the Poso Creek fault. Offsets in the data indicate this intersection is likely ~ 200 m (660 ft) west compared to those shown in Weddle (1959). This suggests the northern most continuation of the West Premier fault may also lie farther to the west. Other faults in the eastern portion of Fig. 7a, especially where dashed, are more poorly constrained. The en-echelon faulting in the northwest part of Fig. 7a just north of the Poso Creek fault has ~ 67 m (220 ft) of total offset and contains the highest density of faulting in the area.

The newly compiled base Macoma claystone data in the Premier fault zone (Fig. 7b) demonstrate a more complex fault architecture than depicted by earlier maps. The two strands of the Premier fault form separate en-echelon fault segments with variable amounts of throw. Maximum fault throw along the West Premier fault is ~ 60 m (197 ft), with displacement decreasing to the south to become negligible approaching the southern fault tip. Along the East Premier fault, throw is greatest toward the center at ~ 40 m (131 ft) decreasing both south and north toward the fault tips.

Fig. 7 **a** Map showing a detailed view of the elevation of the base of the Macoma claystone in the Poso Creek fault area. Abrupt discontinuities in the elevations of base Macoma claystone are interpreted to represent displacement along fault planes. There is ~ 67 m (220 ft) of total offset in the zone of en-echelon faulting in the northwest. **b** Map detailing the offset along the Premier fault zone. Offset is highest along the northern

section of the West Premier fault then decreases along-strike to the south. Offset along the East Premier fault is greatest toward the center and decreases to the south and north. This pattern follows a relay ramp model where displacement decreases toward the fault tips of two main faults connected by a ramp structure. Faults from Weddle (1959), well locations and field boundary data from CalGEM (2019c), background from ESRI

This pattern may be explained with a relay ramp, or stepover, model where extension is accommodated by two main faults connected by a relay ramp feature linking the fault tips of separate faults (Larsen 1988; Peacock and Sanderson 1991; Trudgill and Cartwright 1994; Fossen and Rotevatn 2016). A cross-sectional view of the Macoma claystone base elevation data in the Premier area is shown in Fig. 8a alongside a conceptual model of the relay ramp system in Fig. 8b. The comparison shown in Fig. 8 indicates the Macoma claystone data fit the relay ramp fault zone architectural model.

Groundwater TDS cross sections

The TDS model was used to visualize TDS in vertical cross sections throughout the volume of analysis. The sections extend from near the ground surface to −900 m of elevation. The Macoma claystone top and base, and the faults, were superimposed onto the sections using dataset E and the techniques described in section ‘Clay and fault mapping’. These sections focus on areas with relatively dense data that are near the Poso Creek and Premier fault zones. Since the TDS model infers a Gaussian distribution for each interpolated point, more reliable predictions are represented with a lower standard deviation which are shown in Figs. S3 and S4 of [ESM](#). Where produced water TDS and resistivity data are nearby, interpolated TDS values have lower associated uncertainty. Whereas areas farther away from model input data have higher associated uncertainty—for example, the northeastern portion of TDS cross sections 1, 2, and 3 on Fig. 9 have higher uncertainty (Fig. S3 of [ESM](#)) because of the lack of well data in this area.

Poso Creek fault area

Groundwater TDS near the Poso Creek fault area is shown in Fig. 9. TDS is typically <1,000 mg/L in the shallower aquifer and increases with depth to ~1,200 to 4,500 mg/L, with a higher-TDS zone reaching ~7,000 mg/L. On cross section 1 of Fig. 9, TDS increases rapidly with depth in the western part of the section, whereas TDS increases more slowly with depth elsewhere. The TDS distribution is largely controlled by lithology and further modified by faulting. In general, the Macoma claystone separates overlying lower-TDS from underlying higher-TDS zones. This separation is stronger in the western and southern areas of the cross sections where the Macoma claystone is thicker and more competent. Toward the northeast, the Macoma claystone thins (also see Fig. 6b), becomes sandier, and appears to exert less control on vertical TDS distributions. On cross sections 1–3, the TDS model uncertainty is highest toward the east approximately >3,000 m along section (Fig. S3 of [ESM](#)). However, there

are four wells with resistivity data that indicate little to no TDS gradients exist across the Macoma claystone in the east showing that stratigraphic control diminishes toward the McVan area.

The strongest vertical discontinuities of TDS, as well as the highest TDS (~7,000, mg/L) in the oil field, are found on cross section 1 at ~700 to ~1,200 m along the section and are collocated with the thickest Macoma claystone inside the oil field. This is also the location with the highest density of faulting in the field with a total offset of ~67 m (220 ft). While it is not clear why the higher zone of TDS exists here, it is hypothesized that the thickness of the Macoma claystone coupled with the high density of sealing faults has created a low-permeability structure that has been isolated from fresh-water recharge leaving behind relatively higher-TDS formation water. Development of fine-grained material (fault gouge, clay smear) and/or sediment mixing in the fault planes in areas with thick clays and higher displacement is a likely mechanism for fault seal in this area (Yielding et al. 1997). The presence of thick clay layers in this zone provides an opportunity for clay smear to develop over a large area of the fault planes (Yielding et al. 1997). Furthermore, in zones of closely spaced, higher displacement faults it is possible for fault damage zones to link or nearly link causing low-permeable zones to effectively connect, rendering a larger structure impeding groundwater flow (Faulkner et al. 2010; Bense et al. 2013). Conversely, outside the dense en-echelon fault zone to the south (Fig. 9, sections 2 and 3), TDS gradients decrease, further suggesting that fault density plays a role in controlling TDS distributions. Although this area contains elevated clay content (thicker Macoma claystone) that could contribute to higher TDS predictions, it is not thought that overall TDS patterns are dramatically affected by clay content within the sands. Also, the higher zone of TDS in section 1 of Fig. 9 is strongly supported by resistivity data, evidenced by lower model uncertainty in Fig. S3 of [ESM](#). However, higher model uncertainty below the Macoma claystone away from the high zone of TDS suggests that alternative interpretations could exist if more resistivity and/or produced water TDS data become available. Currently, the en-echelon faulting and the thick Macoma claystone are hypothesized to play a role in this higher-TDS zone given the available data.

Premier fault area

Groundwater TDS cross sections in the Premier area are shown in Fig. 10. Here, TDS also generally increases with depth, although at variable rates. In the shallower aquifer zone above the Macoma claystone, TDS is generally <1,000 mg/L and increases to ~1,500 to 3,500 mg/L in the hydrocarbon zone, reaching ~5,700 mg/L in isolated areas below the Macoma claystone. In the west, TDS rapidly increases with depth across the thicker part of the Macoma claystone (Figs. 6

Fig. 8 **a** Cross section of the base Macoma claystone data from Fig. 7b. This view shows the offset along the West and East Premier faults at the margins and the relay ramp feature between. **b** Conceptual model of the relay ramp system in the Premier fault zone. Modified from Uzel (2016)

and 10), indicating a TDS discontinuity at this level, and implying the Macoma claystone largely restricts vertical groundwater flow between the overlying fresher and underlying brackish groundwater zones. In contrast, in the east, lower-TDS water extends below the thinner and sandier Macoma claystone, indicating reduced control of the stratigraphy on TDS likely by allowing meteoric recharge from higher elevation to the east. This shows that lateral stratigraphic facies changes play an important role in creating the TDS structure by confining water of different TDS, or by controlling recharge of meteoric water.

Stratigraphic controls do not explain all TDS features seen in the cross sections in the Premier area. The lower-TDS water that infiltrates below the pinched-out Macoma claystone from the east has an irregular lateral extent toward the west. In the north, a thin layer of lower-TDS water penetrates across the East Premier fault (Fig. 10, cross section 4). Southward along the East Premier fault, the lower-TDS volume does not extend as far west, stopping coincidentally with the East Premier fault (Fig. 10, sections 5 and 6). Continuing southward, a ~130-m (426 ft) layer of lower-TDS water can be seen beneath the Macoma claystone starting at ~1,500 m (Fig. 10, section 7) and ~2,100 m (Fig. 10, section 8) along the section. South of that, the lower-TDS region alternately stops at the fault (Fig. 10, section 9) and continues across the faults (Fig. 10, section 10).

When comparing the behavior of lower-TDS recharge to the fault geometry, it becomes apparent that there is uneven structural control by the faults on groundwater. As noted before, the East Premier fault has maximum offset toward the center of the fault, decreasing to zero offset at the fault tips (Figs. 7 and 8). The higher-displacement segment of the East Premier fault (Fig. 8, cross section lines 5–8) appears to prevent passage of lower-TDS waters in some places (Fig. 10, sections 5, 6, and 9) but allow it in others (Fig. 10, sections 7 and 8). While the lower-offset fault tips on the East Premier fault (near cross sections 4 and 10) are mapped as faults in the previously published CalGEM fault map (Weddle 1959, plate II), the displacement decreases to near zero and the fault does not seem to impede groundwater here.

This variable control from fault offset on groundwater contrasts with the trapping effect the faults have on oil. The East and West Premier faults act as structural traps where oil production is centered (Fig. S5 of ESM) (Weddle 1959). Where displacement on the West Premier fault is largest (Fig. 8, cross sections 4 and 5), the fault forms an up-dip trapping structure for oil on the west side of the fault (Fig. S5 of ESM). The East Premier fault forms an up-dip structural trap centered on the site of maximum throw directly west of the New Hope fault's southern terminus (between cross sections 6 and 7), resulting in hydrocarbon production in the middle ramp area. In both cases, the locus of petroleum production is at the point of

Fig. 9 Total dissolved solids (TDS) cross sections in the Poso Creek fault area (see Fig. 1 for cross section locations). The gray lines are the top and base Macoma claystone and the black dashed lines are faults. Cross sections are numbered in the lower right corner. There is no vertical exaggeration

maximum displacement, consistent with views of hydrocarbon accumulations in normal fault systems (Harding and Tuminas 1989).

Two conclusions can be drawn from comparing mapped TDS to the geometry of the Premier fault system. First, the ability of these faults to seal out lower-TDS water varies along the strike of the fault and is not directly related to the throw of the fault. This pattern highlights the complicated mix of factors that govern fault zone hydrogeology (Caine et al. 1996; Evans et al. 1997; Bense and Person 2006; Fossen et al. 2007; Faulkner et al. 2010; Bense et al. 2013). Second, the faults are apparently different in sealing nature with respect to oil and water, as shown by the fact that segments of the faults that trap oil are permeable to lower-TDS water. Berg (1975) noted that capillary pressures between oil and water in rock pores are responsible for trapping oil in stratigraphic traps and Watts (1987) noted that membrane seals prevent hydrocarbon leakage when capillary forces exclude hydrocarbon intrusion through the smaller pore throats of water wet seals. This has been demonstrated in the laboratory by Teige et al. (2005) who showed that water may flow through a low-

permeability water wet membrane but, at the same pressure, the flow of oil was impeded by capillary forces. Therefore, it is likely that the faults are sealing to oil but less so to water due to the presence of smaller, water-wet pore throats in the clays and fault gouge. It is expected that the distributions of these zones with differential flow properties are likely influenced by the fault zone geometry and the zones are therefore spatially variable. Furthermore, the hydrocarbons exist in the upper part of the reservoir, and the sealing mechanisms that apply to that portion of the fault may not be operative down-dip, also resulting in spatial variability of fault control.

Conclusions

Groundwater total dissolved solids (TDS) distribution in the Poso Creek Oil Field is controlled by stratigraphy and geologic structure. The controlling factors were discovered using deep TDS patterns to identify long-term groundwater flow paths. TDS concentration is generally low (<3,000 mg/L) in most of the field and reaches its highest value in the northwest

Fig. 10 Total dissolved solids (TDS) cross sections in the Premier fault area. See Fig. 1 for cross section locations. The gray lines are the top and base Macoma claystone and the black dashed lines are faults. Cross sections are numbered in the lower right corner. There is no vertical exaggeration

at ~7,000 mg/L at ~700 m (~2,460 ft) elevation. Relatively higher- and lower-TDS zones are largely divided by the regionally extensive Macoma claystone. Stratigraphic control on TDS from the Macoma claystone diminishes to the east as it thins and becomes sandier, demonstrating how facies changes affect TDS patterns. Inside the oil field boundary, the Macoma claystone base elevation reaches approximately –300 to –600 m (–980 to –2,000 ft) with a thickness of ~20–40 m (65 to 130 ft). TDS concentration in the aquifer overlying the Macoma claystone is typically <1,000 mg/L and increases with depth to ~1,200–3,500 mg/L in the hydrocarbon-bearing zone below the Macoma claystone. There is an isolated zone of higher TDS reaching ~7000 mg/L in the northwest near the Poso Creek fault, which may be the result of formation water being isolated by a low-permeability structure created by a high density of faults and the Macoma claystone. There is also a lower-TDS feature below the Macoma claystone that differs in its western extent indicating that the fault control on TDS is variable along-strike. The lower TDS concentration below the Macoma claystone also suggests the faulting has differential control on groundwater and oil flow dynamics in the field.

This study also provides distinct advancements in modeling TDS in three-dimensions. The TDS model developed here used a marginal likelihood function to find the optimal petrophysical and geostatistical parameters concurrently, which improves model uncertainty estimates. Also, the TDS model can account for complex spatial trends, like universal kriging, but without overfitting by marginalizing over trend model parameters. Lastly, fitting the geostatistical parameters while modeling the spatial trends jointly avoids the complications of fitting an experimental variogram. Together, these methods improve TDS modeling and uncertainty estimates, and could be incorporated in future spatial prediction models of inferred environmental variables.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10040-021-02381-5>.

Acknowledgements Thank you to Matt Landon, Lyndsay Ball, Rhett Everett, Loren Metzger, Emily Haugen, Kim Taylor, Pete McMahon, and Ayman Alzraiee of the USGS for providing advice and reviews. Thank you to Theron Sowers and the Water Resources Group at California State University, Sacramento, for countless hours of data compilation, digitizing, and quality checks. This work was also improved by reviews from Neil Terry of the USGS and two anonymous reviewers.

Funding This work was funded by the California State Water Resources Control Board's Oil and Gas Regional Monitoring Program, the US Bureau of Land Management, US Geological Survey Cooperative Matching Funds, and USGS/California State University Cooperative Agreement G16AC00042.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as

you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado GS, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, Isard M, Jia Y, Jozefowicz R, Kaiser L, Kudlur M, Levenberg J, Mané D, Monga R, Moore S, Murray D, Olah C, Schuster M, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) TensorFlow: large-scale machine learning on heterogeneous distributed systems. <https://www.tensorflow.org/about/bib>. Accessed December 2019
- Archie GE (1942) The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans Am Inst Min Metal Eng* 146: 54–62. <https://doi.org/10.2118/942054-G>
- Asquith GB, Krygowski D (2004) Basic well log analysis, 2nd edn. AAPG, Tulsa, OK
- Atwater T (1970) Implications of plate tectonics for the Cenozoic tectonic evolution of western North America. *Geol Soc Am Bull* 81:3513–3536. [https://doi.org/10.1130/0016-7606\(1970\)81\[3513:IOPTFT\]2.0.CO;2](https://doi.org/10.1130/0016-7606(1970)81[3513:IOPTFT]2.0.CO;2)
- Ball LB, Davis TA, Minsley BJ, Gillespie JM, Landon MK (2020) Probabilistic groundwater salinity mapping from airborne electromagnetic data adjacent to the Lost Hills and Belridge oil fields, southwestern San Joaquin Valley, California. *Water Resour Res* 56(6), 20 pp. <https://doi.org/10.1029/2019WR026273>
- Bartow JA (1991) The Cenozoic evolution of the San Joaquin Valley. *California US Geol Surv Prof Pap* 1501, 40 pp. <https://doi.org/10.3133/pp1501>
- Bartow JA, Pittman GM (1983) The Kern River formation, southeastern San Joaquin Valley, California. *US Geol Surv Bull* 1529-D, 17 pp. <https://doi.org/10.3133/b1529D>
- Bense VF, Person MA (2006) Faults as conduit-barrier systems to fluid flow in siliciclastic sedimentary aquifers. *Water Resour Res* 42(5), 18 pp. <https://doi.org/10.1029/2005WR004480>
- Bense VF, Gleeson T, Loveless SE, Bour O, Scibek J (2013) Fault zone hydrogeology. *Earth-Sci Rev* 127:171–192. <https://doi.org/10.1016/j.earscirev.2013.09.008>
- Berg RR (1975) Capillary pressures in stratigraphic traps. *AAPG Bull* 59(6):939–956. <https://doi.org/10.1306/83D91EF7-16C7-11D7-8645000102C1865D>
- Caine JS, Evans JP, Forster CB (1996) Fault zone architecture and permeability structure. *Geol* 24:1025–1028. [https://doi.org/10.1130/0091-7613\(1996\)024<1025:FZAAPS>2.3.CO;2](https://doi.org/10.1130/0091-7613(1996)024<1025:FZAAPS>2.3.CO;2)
- CalGEM (2016) Chemical analysis database. CalGEM, Sacramento, CA. ftp://ftp.consrv.ca.gov/pub/oil/chemical_analysis. Accessed February 2019
- CalGEM (2019a) Annual reports. CalGEM, Sacramento, CA. https://www.conservation.ca.gov/calgem/pubs_stats/annual_reports/Pages/annual_reports.aspx. Accessed December 2019
- CalGEM (2019b) Well search. CalGEM, Sacramento, CA. <https://secure.conservation.ca.gov/WellSearch>. Accessed December 2019
- CalGEM (2019c) GIS mapping. CalGEM, Sacramento, CA. <http://www.conservation.ca.gov/calgem/maps/Pages/GISMapping2.aspx>. Accessed February 2019

- California State Water Resources Control Board (2015) Model criteria for groundwater monitoring in areas of oil and gas well stimulation. California State Water Resources Control Board, 39 pp. http://www.waterboards.ca.gov/water_issues/programs/groundwater/sb4/additional_resources/index.shtml. Accessed December 2019
- California State Water Resources Control Board (2019) Regional monitoring plan. https://www.waterboards.ca.gov/water_issues/programs/groundwater/sb4/regional_monitoring/index.shtml. Accessed December 2019
- Callaway DC (1990) Organization of stratigraphic nomenclature for the San Joaquin basin, California. AAPG Pac sect 5-20, AAPG, Tulsa, OK
- Castle RO, Church JP, Yerkes RF, Manning JC (1983) Historical surface deformation near Oildale, California. US Geol Surv Prof Pap 1245, 42 pp. <https://pubs.er.usgs.gov/publication/pp1245>. Accessed June 2018
- Clark CE, Veil JA (2009) Produced water volumes and management practices in the United States. ANL/EVS/R-09-1 Argonne Natl Lab. <http://www.ipd.anl.gov/anlpubs/2009/07/64622.pdf>. Accessed December 2015
- Cooley H, Gleick PH, Wolff G (2006) Desalination, with a grain of salt: a California Perspective. Pacific Institute. https://pacinst.org/wp-content/uploads/2013/02/desalination_report1.pdf. Accessed December 2019
- Davis T, Landon MK, Bennett GL (2018) Prioritization of oil and gas fields for regional groundwater monitoring based on a preliminary assessment of petroleum resource development and proximity to California's groundwater resources. US Geol Surv Sci Invest Rep 2018-5065. <https://doi.org/10.3133/sir20185065>
- Ellis DV, Singer JM (2007) Well logging for earth scientists. Springer, Dordrecht, The Netherlands
- EMWD (2013) Eastern municipal Water District. Desalination Program. https://www.emwd.org/sites/main/files/file-attachments/desalinationprogram_englis.pdf?1537294374. Accessed December 2019
- Evans JP, Forster CB, Goddard JV (1997) Permeability of fault-related rocks, and implications for hydraulic structure of fault zones. J Struct Geol 19(11):1393–1404. [https://doi.org/10.1016/S0191-8141\(97\)00057-6](https://doi.org/10.1016/S0191-8141(97)00057-6)
- Faulkner DR, Jackson CAL, Lunn RJ, Schlische RW, Shipton ZK, Wibberley CAJ, Withjack MO (2010) A review of recent developments concerning the structure, mechanics and fluid flow properties of fault zones. J Struct Geol 32(11):1557–1575. <https://doi.org/10.1016/j.jsg.2010.06.009>
- Faunt CC (ed) (2009) Groundwater availability of the Central Valley aquifer, California. US Geol Surv Prof Pap 1766, 225 pp. <https://pubs.usgs.gov/pp/1766/>. Accessed June 2018
- Fetter CW (2001) Applied hydrogeology. Prentice Hall, Upper Saddle River, NJ
- Fisher JB, Boles JR (1990) Water–rock interaction in tertiary sandstones, San Joaquin basin, California, USA: diagenetic controls on water composition. Chem Geol 82:83–101. [https://doi.org/10.1016/0009-2541\(90\)90076-J](https://doi.org/10.1016/0009-2541(90)90076-J)
- Fossen H, Rotevatn A (2016) Fault linkage and relay structures in extensional settings: a review. Earth-Sci Rev 154:14–28. <https://doi.org/10.1016/j.earscirev.2015.11.014>
- Fossen H, Schultz RA, Shipton ZK, Mair K (2007) Deformation bands in sandstone: a review. J Geol Soc 164(4):755–769. <https://doi.org/10.1144/0016-76492006-036>
- Gillespie JM, Kong D, Anderson SD (2017) Groundwater salinity in the southern San Joaquin Valley. AAPG Bull 101:1239–1261. <https://doi.org/10.1306/09021616043>
- Gillespie JM, Davis TA, Stephens MJ, Ball LB, Landon MK (2019) Groundwater salinity and the effects of produced water disposal in the Lost Hills-Belridge Oil Fields, Kern County, California. Environ Geosci 26(3):73–96. <https://doi.org/10.1306/eg.02271918009>
- Gringarten E, Deutsch CV (2001) Teacher's aide variogram interpretation and modeling. Math Geol 33(4):507–534. <https://doi.org/10.1023/A:1011093014141>
- Hamlin HS, de la Rocha L (2015) Using electric logs to estimate groundwater salinity and map brackish groundwater resources in the Carrizo-Wilcox aquifer in South Texas, Gulf Coast. Geol Soc J 4: 109–131
- Hansen JA, Jurgens BC, Fram MS (2018) Quantifying anthropogenic contributions to century-scale groundwater salinity changes, San Joaquin Valley, California, USA. Sci Total Environ 642:125–136. <https://doi.org/10.1016/j.scitotenv.2018.05.333>
- Harding TP, Tuminas AC (1989) Structural interpretation of hydrocarbon traps sealed by basement normal block faults at stable flank of foredeep basins and at rift basins. AAPG Bull 73(7):812–840. <https://doi.org/10.1306/44B4A276-170A-11D7-8645000102C1865D>
- Jin G, Torres-Verdín C, Lan C (2009) Pore-level study of grain-shape effects on petrophysical properties of porous media. SPWLA 50th annual logging Symp, SPWLA. <https://www.onepetro.org/conference-paper/SPWLA-2009-62136>. Accessed June 2018
- Jordan PD, Gillespie J (2016) Produced water disposal injection in the southern San Joaquin Valley: no evidence of groundwater quality effects due to upward leakage. Environ Geosci 23(3):141–177
- Kang M, Jackson RB (2016) Salinity of deep groundwater in California: water quantity, quality, and protection. Proc Natl Acad Sci 113(28): 7768–7773. <https://doi.org/10.1073/pnas.1600400113>
- Kang M, Perrone D, Wang Z, Jasechko S, Rohde MM (2020) Base of fresh water, groundwater salinity, and well distribution across California. Proc Natl Acad Sci 117(51):32302–32307. <https://doi.org/10.1073/pnas.2015784117>
- Kharaka YK, Berry FA (1974) The influence of geological membranes on the geochemistry of subsurface waters from Miocene sediments at Kettleman North Dome in California. Water Resour Res 10(2): 313–327. <https://doi.org/10.1029/WR010i002p00313>
- Kharaka YK, Hanor JS (2004) Deep fluids in the continents: I. sedimentary basins. Treatise Geochem 5:1–48. <https://doi.org/10.1016/B0-08-043751-6/05085-4>
- Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. <https://arxiv.org/abs/1412.6980>. Accessed June 2019
- Kondash A, Hoponick Redmon J, Lambertini E, Feinstein L, Weinthal E, Cabrales L, Vengosh A (2020) The impact of using low-saline oilfield produced water for irrigation on water and soil quality in California. Sci Total Environ 733. <https://doi.org/10.1016/j.scitotenv.2020.139392>
- Larsen PH (1988) Relay structures in a lower Permian basement-involved extension system, East Greenland. J Struct Geol 10(1):3–8. [https://doi.org/10.1016/0191-8141\(88\)90122-8](https://doi.org/10.1016/0191-8141(88)90122-8)
- Laudon J, Belitz K (1991) Texture and depositional history of late Pleistocene-Holocene alluvium in the central part of the western San Joaquin Valley, California. Bull Assoc Eng Geol 28(1):73–88. <https://doi.org/10.2113/gsegeosci.xxviii.1.73>
- Leitz F, Boegli W (2011) Evaluation of the Port Hueneme demonstration plant: an analysis of 1 MGD reverse osmosis, nanofiltration and electrodialysis reversal plants under nearly identical conditions. Rep no. 65, US Bur Reclam Desalination and Water Purification Res and Dev Progr. <https://www.usbr.gov/research/dwpr/reportpdfs/report065.pdf>. Accessed December 2019
- McMahon PB, Kulongoski JT, Vengosh A, Cozzarelli IM, Landon MK, Kharaka YK, Davis TA (2018) Regional patterns in the geochemistry of oil field water, southern San Joaquin Valley, California, USA. Appl Geochem 98:127–140. <https://doi.org/10.1016/j.apgeochem.2018.09.015>
- Metzger LF, Landon MK (2018) Preliminary groundwater salinity mapping near selected oil fields using historical water-sample data, central and southern California. US Geol Surv Sci Invest Rep, 2018-5082. <https://doi.org/10.3133/sir20185082>

- Metzger LF, Davis TA, Peterson MF, Brilmyer CA, Johnson JC (2018) Water and petroleum well data used for preliminary regional groundwater salinity mapping near selected oil fields in central and southern California. US Geol Surv Data Release. <https://doi.org/10.5066/F7RN373C>
- Page RW (1973) Base of fresh ground water approximately 3,000 micromhos, in the San Joaquin Valley, California. US Geol Surv Hydrol Atlas Series 489. <https://doi.org/10.3133/ha489>
- Peacock DCP, Sanderson DJ (1991) Displacements, segment linkage and relay ramps in normal fault zones. *J Struct Geol* 13(6):721–733. [https://doi.org/10.1016/0191-8141\(91\)90033-F](https://doi.org/10.1016/0191-8141(91)90033-F)
- Ramseyer K, Boles JR (1986) Mixed-layer illite/smectite minerals in tertiary sandstones and shales, San Joaquin Basin, California. *Clays Clay Min* 34(2):115–124. <https://doi.org/10.1346/CCMN.1986.0340202>
- Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. MIT Press, Cambridge, MA
- Reid SA (1995) Miocene and Pliocene depositional systems of the southern San Joaquin basin and formation of sandstone reservoirs in the Elk Hills area, California. Cenozoic paleogeography of the western United States II. In: Fritsche AE (ed) Cenozoic paleogeography of the western United States: II, Pacific section. Society of Economic Paleontologists and Mineralogists, Broken Arrow, OK, pp 131–150
- Saleeby J, Saleeby Z (2019) Late Cenozoic structure and tectonics of the southern Sierra Nevada-San Joaquin Basin transition, California. *Geosphere* 15:1164–1205. <https://doi.org/10.1130/GES02052.1>
- Saleeby J, Saleeby Z, Le Pourhiet L (2013) Epeirogenic transients related to mantle lithosphere removal in the southern Sierra Nevada region, California: part II, implications of rock uplift and basin subsidence relations. *Geosphere* 9:394–425. <https://doi.org/10.1130/GES00816.1>
- Scheirer AH, Magoon LB (2007) Age, distribution, and stratigraphic relationship of rock units in the San Joaquin Basin province, California, chap 5. In: Petroleum systems and geologic assessment of oil and gas in the San Joaquin Basin province, California. US Geol Surv Prof Pap 1713:1–107. <https://doi.org/10.3133/pp17135>
- Schloss J (2017a) Aquifer exemption application Basal Etchegoin, Etchegoin Formation Poso Creek Oil Field Kern County, California. Hathaway, LLC Bakersfield, CA. ftp://ftp.consrv.ca.gov/pub/oil/Aquifer_Exemptions/County/Kern/Kern-Front-Chanac/Application/Aquifer_Exemption_Application.pdf. Accessed December 2019
- Schloss J (2017b) Aquifer exemption application Upper Chanac Formation Oil Sand Kern Front Oil Field Kern County, California. Hathaway, Bakersfield, CA. ftp://ftp.consrv.ca.gov/pub/oil/Aquifer_Exemptions/County/Kern/Kern-Front-Chanac/Application/Aquifer_Exemption_Application.pdf. Accessed December 2019
- Sousa FJ, Farley KA, Saleeby J, Clark M (2016) Eocene activity on the Western Sierra fault system and its role incising Kings Canyon, California. *Earth Planet Sci Lett* 439:29–38. <https://doi.org/10.1016/j.epsl.2016.01.020>
- Stephens MJ, Shimabukuro DH, Gillespie JM, Chang W (2019) Groundwater salinity mapping using geophysical log analysis with in the Fruitvale and Rosedale Ranch oil fields, Kern County, California, USA. *Hydrogeol J* 27(2):731–746. <https://doi.org/10.1007/s10040-018-1872-5>
- Stephens MJ, Haugen EA, Shimabukuro DH, Gillespie JM, Sowers TA, Ducart A, Levinson Z, Medrano V (2021) Geochemical, geological, and geophysical data for wells in the Poso Creek Oil and Gas field, Kern County, California (ver. 2.0, 2020). US Geol Surv Data Release. <https://doi.org/10.5066/P9RR9UYN>
- Taylor KA, Fram MS, Landon MK, Kulongoski JT, Faunt CC (2014) Oil, gas, and groundwater in California: a discussion of issues relevant to monitoring the effects of well stimulation at regional scales. US Geological Survey, CA Water Sci. Center, Sacramento, CA. ftp://ftp.consrv.ca.gov/pub/oil/SB4EIR/docs/Taylor_et_al_2014.pdf. Accessed December 2019
- Teige GM, Hermanrud C, Thomas WH, Wilson OB, Bolås HMN (2005) Capillary resistance and trapping of hydrocarbons: a laboratory experiment. *Pet Geosci* 11(2):125–129. <https://doi.org/10.1144/1354-079304-609>
- Trudgill B, Cartwright J (1994) Relay-ramp forms and normal-fault linkages, Canyonlands National Park, Utah. *Geol Soc Am Bull* 106(9):1143–1157. [https://doi.org/10.1130/0016-7606\(1994\)106<1143:RRFANF>2.3.CO;2](https://doi.org/10.1130/0016-7606(1994)106<1143:RRFANF>2.3.CO;2)
- US Geological Survey (2020) California oil, gas, and groundwater (COGG) Program. <https://webapps.usgs.gov/cogg/>. Accessed May 2020
- Uzel B (2016) Field evidence for normal fault linkage and relay ramp evolution: the Kırkağaç fault zone, western Anatolia (Turkey). *Geodin Acta* 28(4):311–327. <https://doi.org/10.1080/09853111.2016.1184778>
- Watts NL (1987) Theoretical aspects of cap-rock and fault seals for single-and two-phase hydrocarbon columns. *Mar Pet Geol* 4(4):274–307. [https://doi.org/10.1016/0264-8172\(87\)90008-0](https://doi.org/10.1016/0264-8172(87)90008-0)
- Weddle JR (1959) Premier and Enas areas of Poso Creek Oil Field: summary of operations 45. California Division of Oil and Gas, Sacramento, pp 41–50
- Welge EA (1973) Poso Creek oil field, McVan area: summary of operations 59. California Division of Oil and Gas, Sacramento, CA, pp 59–68
- Wilson AM, Garven G, Boles JR (1999) Paleohydrogeology of the San Joaquin basin, California. *Geol Soc Am Bull* 111(3):432–449. [https://doi.org/10.1130/0016-7606\(1999\)111<0432:POTSJB>2.3.CO;2](https://doi.org/10.1130/0016-7606(1999)111<0432:POTSJB>2.3.CO;2)
- Winsauer WO, Shearin HM Jr, Masson PH, Williams M (1952) Resistivity of brine-saturated sands in relation to pore geometry. *AAPG Bull* 36:253–277. <http://archives.datapages.com/data/bulletns/1949-52/data/pg/0036/0002/0250/0253.htm?doi=10.1306%2F3D9343F4-16B1-11D7-8645000102C1865D>. Accessed July 2021
- Wright MT, McMahon PB, Landon MK, Kulongoski JT (2019) Groundwater quality of a public supply aquifer in proximity to oil development, Fruitvale oil field, Bakersfield, California. *Appl Geochem* 106:82–95. <https://doi.org/10.1016/j.apgeochem.2019.05.003>
- Yielding G, Freeman B, Needham DT (1997) Quantitative fault seal prediction. *AAPG Bull* 81(6):897–917. <http://archives.datapages.com/data/bulletns/1997/06jun/0897/0897.htm?doi=10.1306%2F522B498D-1727-11D7-8645000102C1865D>. Accessed July 2021

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.