UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

LINKING IMAGE PROCESSING AND NUMERICAL MODELING TO IDENTIFY

POTENTIAL GEOHAZARDS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

RAFAEL AUGUSTO PIRES DE LIMA
Norman, Oklahoma
2017

LINKING IMAGE PROCESSING AND NUMERICAL MODELING TO IDENTIFY
POTENTIAL GEOHAZARDS

A THESIS APPROVED FOR THE
CONOCOPHILLIPS SCHOOL OF GEOLOGY AND GEOPHYSICS

BY

_____
Dr. Kurt Marfurt, Chair

_____
Dr. Shankar Mitra

_____
Dr. Zulfiquar Reza

Acknowledgements

environment, and Saurabh for clarifying so many petroleum engineering terms and concepts.

Lastly, I extend my thanks to friends that supported me during the period of this work; Ashley, Gabriel, Flavio, Priscila, Hernan, Laura, Paula, Javier, Henry, Daniela, Camila, Esther, Limeira, Samoes, Santos, Jussara, Andre, Rafael, Jairo, Ligia, Mae, Pai, Tio Ze, Silvio, Silvia, Marcos, and Ana. You all did a lot for me.

Table of Contents

List of Figures

Abstract

Faults, along with natural fractures, may enhance production when confined within the reservoir. However, if the fault is connected to an aquifer, it may cause early water breakthrough in the reservoir. Even if they are not conductive, they pose a significant geohazard during drilling as fault slippage can cause shearing of casing/tubing leading to either sidetracking, or complete abandonment of the well. In this thesis, I propose a simplistic approximation of dynamic conductivity of faults based on steady state flow equation. I use a geometric attribute; coherence, as a proxy for fault hydraulic conductivity and in a steady state flow equation to model dynamic flow.

This thesis was inspired by problems faced by several companies working the Eagle Ford shale reservoir of south Texas. Surveys often exceeds 1000 km^2 and exhibit hundreds of faults. Most faults are not problematic; however, some connect with the deeper Edwards limestone aquifer. Wells that complete near these faults produce water.

This algorithm can provide early water production warnings and can provide simple, easy to compute useful input in field development in the absence of the more complete datasets required more rigorous reservoir simulation implemented in commercial software. This simple tool is designed to be used in a statistical, rather than deterministic manner, identifying problematic faults by comparing their orientation and connectivity to those known to be bad by previous drilling history. The computational time is less than 17% of a more rigorous conventional reservoir simulation. The model can be updated easily as more and more dataset are available during various stages of

field development by ignoring important parameters for single well production such as facies, petrophysical and flow equations.

This algorithm is a fast and simple approximation that can be very useful in overall field management where one wishes to quickly identify problematic faults or fault sets.

Chapter 1: Introduction

3D seismic edge detecting algorithms such as coherence are routinely used to highlight geomorphological and structural features such as channels and faults. Gersztenkorn and Marfurt (1999) showed that coherence may be evaluated based on cross-correlation between the seismic traces, semblance or with an eigendecomposition of the seismic data covariance matrix. Höcker and Fehmers (2002), and Marfurt (2006) proposed alternative data conditioning algorithms to improve the signal-to-noise-ratio and generate a better coherence image.

While faults provide crucial geologic information that can be critical for reservoir modeling, fault picking and interpretation is a time-consuming activity. While significant effort has been devoted to fault enhancement (e.g. Pedersen et al., 2002; Machado et al., 2016), little effort has been devoted to estimating fault connectivity between reflectors of interest. In addition to being a potential conduit for water from a nearby aquifer in resource plays, faults, and collapse features may also give rise to seal risk for carbon sequestration (Holubnyak et al., 2014) and waste water injection (Machado et al., 2017).

Seismic attributes have been correlated with fluid flow in recent studies. Guo et al. (2012) generalize the idea of correlating seismic-derived volumetric curvature to fluid flow first applied by Nissen et al. (2009).

Nissen et al. (2009) studied the Dickman Field in Ness County, Kansas, and determined that fluid flow is associated with the proximity of the well to the nearest fracture lineament. Based on geologic analysis and production data, they interpret two lineaments sets. The first lineament set correspond to debris, clay, and silt-filled

fractures and serve as barriers to fluid flow. The second lineament set is associated with open fractures that channel water from the underlying aquifer. This determination is critical in the potential use of this depleted reservoir for carbon sequestration.

Nissen et al. (2009) method is labor intensive for the interpreter because it requires extensive manual picking. Adding to the weak points, the method only accounts for fluid flowing from the nearest lineament. Guo et al. (2012) generalize the idea and assume that the fracture density is proportional to the volumetric curvature and that the fluid production is reflected as the sum of fluid flow from all fracture lineaments rather than simply the closest one.

Guo et al. (2012) were able to a find correlation between the distance from sealed (NE) fractures and increasing oil production in contrast water production decreases with increasing distance from open (NW) trending fractures. They obtain their results generating a suite of azimuthally-limited volumetric curvature volumes that are sensitive to fracture orientation and intensity. These attributes were later directly correlated to production.

Holubnyak et al. (2014) attests that the Arbuckle in southwestern Kansas is a good candidate for carbon sequestration because of its lithological proprieties as well as its thickness and depth. The authors use volumetric curvature to improve identification and numerical characterization of karst features from a seismic 3D volume and use the lineament interpretation to place a well in the investigated area. The drilled well confirmed the presence of fractures and faults inferred from the volumetric curvature. The authors then create a reservoir model defining faults using volumetric curvature and porosity using seismic inversion. This model then is used as input to a commercial flow

simulation program to match production history. Holubnyak et al. (2014) show that the volumetric curvature derived from seismic amplitude data can be effectively used to improve fault interpretation as well as the characterization of karst features. They also find that additional methods may be necessary to better resolve karst features.

Reservoir simulation augmented by well tests can quantify the fault transmissibilites and can estimate the water influx using either discretized flow equations   or simple empirical correlations. Such analysis is a complex task and requires careful model building and detailed well control. My primary objective is to develop a "quick and dirty" image processing algorithm that will highlight faults that may be connected to nearby aquifers constituting potential geohazards. My secondary objective will be to match the results obtained with this image processing tool with statistical analysis of water production.

My hypothesis is that faults seen on seismic data will act as conductors between two sets of horizons – one of which is an aquifer, the second a producing or reservoir – that will act as sources or sinks. To do so, I initially assume the worst-case scenario, i.e. that all faults that can be resolved by the seismic coherence attribute are conductive. As more data are made available during field life cycle, the model can be updated with different conductivities assigned to fault sets having different orientations, and even individual faults themselves.

The initial goal is then to highlight faults that connect both horizons – one source and one sink, while unconnected faults should exhibit a weaker response. In this simple model, I ignore the fact that faults can be "dip sealing" or "dip leaking" and the same fault can have different sealing behavior depending on the lithology. Weber

(1987) studied growth fault structures in the Niger Delta and states that although there is controversy about fluid migration along faults, it is likely that such process occurs. The author also shows fault traps observed on listric faults on the Niger delta. Depending on the local lithology, faults can be sealed by clay smear. There are different approaches on how to calculate clay smear potential. Stuevold et al. (2003) use the Shale Gouge Ratio (SGR) and the shale smear factor (ratio to fault throw to clay layer thickness) for the Ormen Lange Field, offshore Norway. They find that the low values of SGR and faults with throws up to the juxtaposition (sealant layer on footwall and hanging wall side by side) threshold are unlikely to seal.

Reservoir simulator studies can integrate water influx models that simulate and predict aquifer influence on water-drive reservoirs. Commonly the aquifer parameters are estimated using either direct measurements, history matching or material balance calculations. Sverdrup et al. (2003) studied the sealing properties of faults in the Snorre oil field in the North Sea where oil was recovered using water-alternating-gas injection. By observing water and gas breakthrough, history matching, and tracers in the water-alternating-gas injection, Sverdrup et al. (2003) were able to evaluate the flow paths and communication patterns. The authors conclude that different seal properties for different fluids should be carefully assessed during the reservoir simulation process. Flow of water, oil and gas have different behavior on the studied field. Such detailed petrophysical proprieties are not incorporated in my algorithm at the present moment.

Fault permeability and porosity are related to lithology. Faults with small displacements may lie below seismic resolution and therefore may be invisible on seismic datasets. Fractured rocks may have enough porosity for fluid flow, and although

structural geologists can predict their occurrence, fractures in general will not be visible in edge detecting seismic attribute volumes.

Nevertheless, this simplistic assumption, that faults seen on seismic data have a higher hydraulic conductivity (or permeability) than the surrounding background geology, may prove itself useful in scenarios where the faults are not completely characterized. In my simulation, I assume a single-phase fluid system and no phase transformations, i.e. fluid above bubble point or dew point pressure, no solution gas converting to free gas during production and no composition information related to liquid/gas. The goal is to provide a simple, rapid analysis toll that can be run on any 3D seismic survey. If the solutions are shown to be statistically valid, they can then be used for subsequent well planning.

I believe this is particularly useful in resource plays where operators have the statistics of hundreds of wells and where the size of the reservoir may be too large to model with more accurate numerical simulator software. As more data from well testing, 4D seismic, chemical data, etc. are made available, the well placement and production profiles can be updated. Early information on water production can also help in early surface facility modifications to handle additional water which is frequently valuable in offshore applications.

By construction, reservoir simulator history matching of production data requires significant well control. In addition to the volumes of fluids produced over time, reservoir simulators require estimates of permeability and pressures from well tests as well as porosity from core and log analysis. My objective is to somewhat simpler – to evaluate the hypothesis that production of water from a nearby aquifer can

be correlated to faults and other discontinuities seen on 3D seismic data. Obviously, such problems can be handled by a simulator, however the geometry of each fault needs to be explicitly defined and the number of cells necessary to represent the aquifer will significantly increase the cost of the simulator.

As I conclude writing my thesis I encountered two problems, both of which are data issues. First, although I anticipated obtaining a license to a modern Eagle Ford survey, this has still not happened. Second, although I have access to several 3D surveys and oil production in the midcontinent of the USA, the corresponding water production is not available, since it was not required to report such volumes to the appropriate state regulatory agency (e.g. Trumbo, 2014). In general, the water-to-gas ratio is not reported either, with oublic information limited to water production for only a few points in well life, such as well tests, precluding correlation with my flux computation results. . For this reason, I am forced to use good quality seismic surveys that do not exhibit water problems as a proxy while I await the appropriate data.

I begin the thesis with a discussion about the methods and hypothesis used for the algorithm development. Next I apply the algorithm to synthetic and real datasets, for one of which I statistically compare the flux predictions and measured water production. Then I use the algorithm to predict head and link with pressure on synthetic datasets. Lastly, I conclude with a summary of the advantages and limitations of this work as well as suggestions for future technology development. Appendices summarize algorithm development, implementation and parallelization.

Chapter 2: Methods

I assume that the 3D volumetric result of an edge detecting attribute, such as coherence, is representative of hydraulic conductivity, allowing me to use these attributes as a proxy of conductivity to simulate the flow between two horizons. The two horizons on this conductivity scheme are: the aquifer layer (modeled as the source) and the reservoir layer (modeled as a sink).

The fluid head (h) for the three-dimensional saturated flow equation (Istok, 2013; Harbaugh, 2005) is

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + W = S\frac{\partial h}{\partial t} \qquad (1)$$

where K_x, K_y, and K_z are the conductivities of the media in the x, y, and z coordinate directions, W is a volumetric flux per unit volume (e.g. a pumping or injecting well), S is the specific storage of the media, and t is time. Equation 1 is commonly used for groundwater flow modeling.

The **K** and S properties are related to petrophysical properties common to an oil and gas field, where the hydraulic conductivity of a material is related to its permeability while the specific storage is associated with its porosity.

Next, the absolute flow **q** is defined as

$$\mathbf{q} = q_x\hat{\boldsymbol{x}} + q_y\hat{\boldsymbol{y}} + q_z\hat{\boldsymbol{z}} = -K_x\frac{\partial h}{\partial x}\hat{\boldsymbol{x}} - K_y\frac{\partial h}{\partial y}\hat{\boldsymbol{y}} - K_x\frac{\partial h}{\partial z}\hat{\boldsymbol{z}}, \qquad (2)$$

where $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ are the unit vectors in the x, y, and z directions respectively.

The finite-difference method used to compute the equations showed above are described in Appendix A, while the parallelization scheme used is described in the Appendix B.

Since my objective is to map areas with a higher conductivity that link or otherwise connect two different horizons, I can assume that the time derivative in Equation 1 is zero, i.e., the system is at equilibrium and the head does not change with time.

During the transient period (when the time derivative of the potential is not zero), the fluid may accumulate on different areas of the analyzed volume due to the specific storage properties. In the steady state, as all the fluid has completely saturated the media, the flow (or the fluid velocity) will be higher on the paths that link the two horizons. Assuming a fault with a higher hydraulic conductivity than the surrounding lithology linking two geological layers with different potential values, most of the fluid will flow through the higher conductivity area. For this reason, I hypothesize that those conductors connecting source and sink will exhibit a higher absolute flow value, $|\mathbf{q}|$, compared to the other areas.

Petroleum engineers use finite difference or finite element based (in case of coupled flow) flow simulators that compute the change in potential (or pressure) based on different fluid properties and saturation of these fluids in a reservoir. In a conventional simulator, to observe the effect of production and injection on the reservoir pressure, they need to model the transient pressure changes caused due to wells. Sinha et al. (2017) summarize the workflow as honoring the production data during the history matching process, requiring simultaneous matches of flow rates,

pressures and composition. A necessary dataset for such an analysis needs to include but is not limited to the production rate (all phases), lab composition data (modeled using equations of state) as well as casing and tubing information. Additional information such as the pressure profile in the tubing/casing using production logs at multiple times during the well life is required to assign the appropriate wellbore models and accurately represent the single phase or multi-phase flow regimes throughout the production life cycle of the well.

Realistic simulations rely on careful geological models and well data. Building an accurate geological model, such as the one shown by Senoglu (2017) representing an area of 72 square miles composed of 89x98x1525 cells (13,301,050 cells in total), might take up to two years (personal communication with the author). This "static" model then can be used as input to the simulation process. The history matching described by Sinha et al. (2017) used three wells and took around one and a half years to analyze an area of 36 mi^2 (personal communication with the main author).

The objective of the procedure described in this thesis is to develop an indicator of areas that might constitute geohazards based on 3D seismic data, and a statistical distribution of good wells and "problematic wells" that product excessive amounts of water on scarce geology and production data. In this workflow, the interpreter picks horizons if necessary and chooses a few algorithm parameters. Set up and run times using desktop multicore workstations (24 processors) are less than one day.

Chapter 3: Results and discussion

Synthetic model

To calibrate the algorithm and to better understand the results obtained using the proposed methodology in this work, I generated a 3D synthetic model and then computed the potential and flow.

Figure 1 represents a cross-section of a 3D volume showing several faults and two horizons - the reservoir and the overpressured aquifer. Faults that link the potential reservoir with the overpressured aquifer below with a higher hydraulic conductivity I denote as "connected faults". In contrast, faults that terminate between the aquifer and reservoir or are otherwise broke I donate as being "disjoint faults".

Faults F1, F4, and F5 are connected faults. Fault F3 terminates before reaching the reservoir and is a disjoint fault. Although fault F2 visually links the reservoir and aquifer, it is not connected due to the finite-difference stencil (Appendix A); therefore, it is also a disjoint fault in my scheme.

Figure 1a shows a vertical slice through the input, hydraulic conductivity volume. Figure 1b shows the resulting head computation. Figure 1c shows the corresponding flux.

The connected faults have a higher anomalous value for the absolute flow as expected. Note that the rightmost fault has a higher conductivity than the background and a much smaller conductivity compared with the other faults.

Even so, because it connects both horizons, the flow of the fault F5 is higher than the flow of the disjoint faults (gray arrow, and the gray circle). Occurrence of disjoint faults is not uncommon in the "reservoir aquifer" scenario, and provides partial

connectivity and weaker absolute flow (Stuevold et al., 2003). Such not fully connected faults and can be modelled effectively using my algorithm. There are no special computational challenges in this scenario as compared to conventional reservoir simulation. The latter involves significant computational effort to model flow across discontinuities with multiple fluid transmissibilities.

Reservoir simulation use is a standard petroleum industry predictive tool. The continuous media is approximated by a discrete finite difference or finite element mesh. At each grid point the partial differential equations that describe the relationship between different reservoir proprieties are approximated by simple algebraic equations. After defining appropriate initial and boundary conditions, one solves the algebraic equations to approximate flows, saturations, and pressure in the reservoir. Physical processes include the flow of fluids partitioned into as many as three phases (oil, water, and gas) and mass transfer between the various phases. Reservoirs simulators also account for the effects of viscous, capillary, and gravitational forces on fluid flow by making use of a generalized form of Darcy's Law. This approach makes possible to use the least number of simplifying assumptions for reservoir heterogeneity, mass transfer between phases, and the forces/mechanisms responsible for fluid flow (Ertekin et al., 2001).

While commercial reservoir simulator can be used to model the aquifer linkage problem, the computational overhead is high. In general, single-phase water flow is incompressible and inviscid, with (by construction) no nonlinear phase conversions. By restricting myself to simpler physics, I can address much larger sized problems, represented by seismic surveys on the order of 1000 km^2 or greater.

11

The intermediate step of my algorithm, head potential computation, is displayed in Figure 1b. This potential has the same nature as what is used for the initialization on reservoir simulator tools and can be viewed as continuous gravity column pressure potential. My algorithm at present does not consider competing capillary forces, but these can be included at negligible additional computational cost.
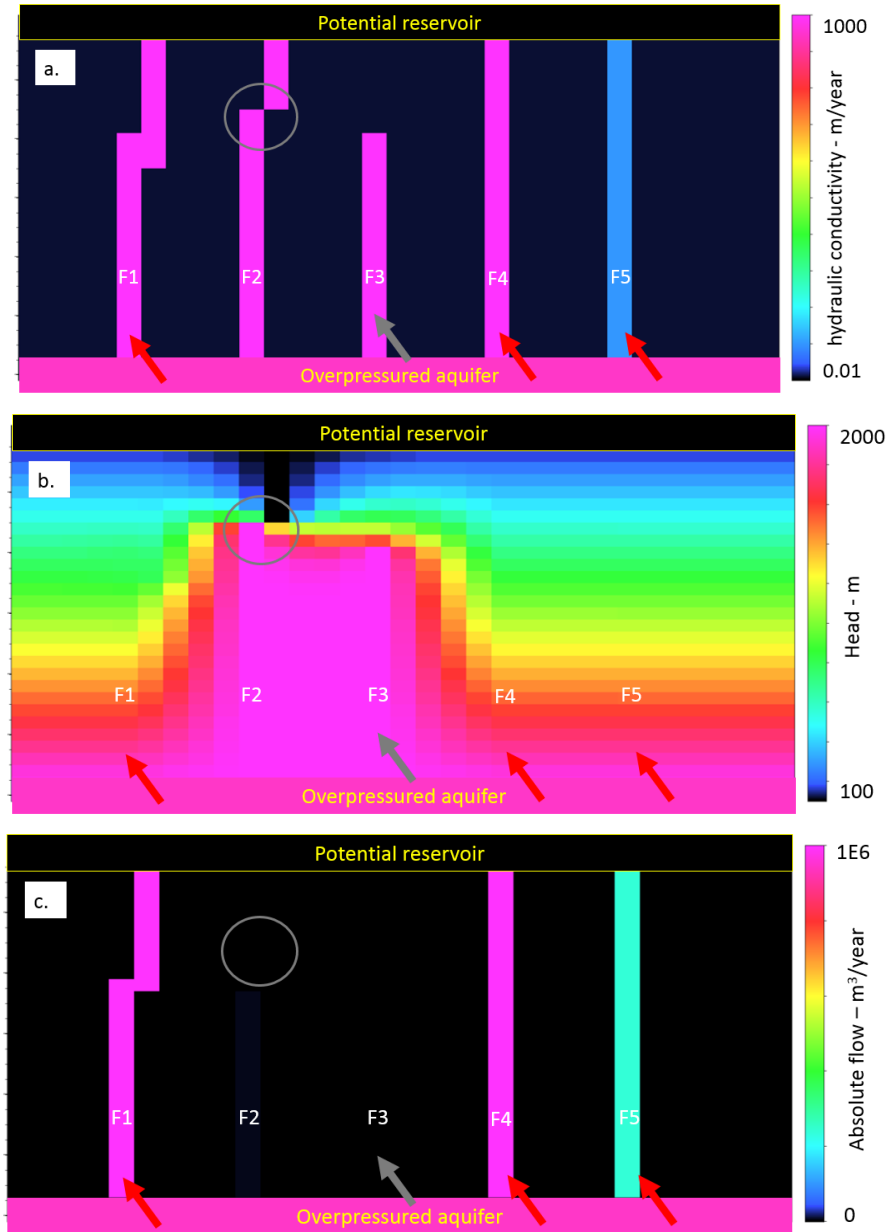
Figure 1: Cross-section of a 3D synthetic model with "connected" and "disjoint" faults. (a) The background hydraulic conductivity is 10 m/year while the stronger pink faults have a value of 1000 m/year. The weaker blue fault has a hydraulic conductivity of 100 m/year. (b) The head *h*. (c) The absolute flow |q|. On all panels, the red arrows indicate through-going faults, and gray arrows disjoint faults. The gray circle highlights the break point of a fault that is not connected due to the finite difference scheme. Even the rightmost fault, with a smaller hydraulic conductivity, is highlighted with the algorithm.

Application

I do not have licenses to an Eagle Ford example at this time. However, although the surveys presented here are not from resource plays, they are faulted, providing a useful proxy until I obtain rights to a faulted resource play. Given this current lack of data licenses, I demonstrate the value of the algorithm on two highly faulted data volumes from New Zealand. The first is represented by polygonal faulting, the second by a suite of radial faults.

Example 1: Modeled "flow" through polygonal faults

The Great South Basin (GSB) is highly faulted New Zealand Basin and covers an area of approximately 150,000 km^2. The GSB is one of the better-known New Zealand's deepwater basins (Uruski, 2010).

The proxy for hydraulic conductivity input (Figure 2 in grayscale and Figure 4 using blue-green-yellow-red color scale) was computed by applying a directional Laplacian of a Gaussian (dLoG) filter (Machado et al., 2016) to a coherence volume. The dLoG operator sharpens perpendicular and smooths parallel to the fault. Coherence and its dLoG filtered version exhibit values between 0 and 1. Inside the algorithm, I scale the dLoG to vary from 0.01 to 1000 m/year. These values were chosen to mimic shale (low hydraulic conductivity and permeability) and sand (high hydraulic conductivity and permeability) values. Heath (1983) finds hydraulic conductivity values to have a large. Bense et al. (2003) perform a study along fault zones in the Roer Valley Rift System (Netherlands) making use of recently developed digital-image-analysis techniques to estimate the spatial distribution of hydraulic conductivity at the

14

millimeter-scale. They use these images to describe the micromorphologic characteristics of the fault zone, and calibrate the results using laboratory measurements of hydraulic conductivity on core-plug samples. The authors explain that a sample taken from the fault core reveals that sand and clay are mixed together along the edge of the clay smear. They found values for the hydraulic conductivity in the fault vicinity ranging from ~2600 m/y to ~8700 m/y (mean) in the laminated sand layer. These values are comparable to the ones I used for high hydraulic conductivity.

To have a better geological significance, the top and bottom limits for the potential calculations should be based on geological horizons. However, for the initial testing, I limited the computations solely based on time slices. This procedure is also fundamentally important to observe the algorithm results if the interpreter has no time to pick horizons. Even with little geologic control, interesting features can still be quickly estimated with the flux computations.

Computing the absolute flow on the GSB (Figure 3 in grayscale and Figure 5 using blue-green-yellow-red color scale) balances the value of smaller and weaker faults, highlighting geological features that may otherwise be overlooked. The geometrical behavior of faulting displayed in Figure 3 associated with syneresis became clearer on the absolute flow results. I can observe one of the faults indicated in the image presents a higher flux value when compared with the dLoG filtered coherence. I accept this as an indication that this fault is connected from t=1.2 s to t=2.2 s. I also observe syneresis areas becoming anomalously smaller when compared with the dLoG input. This is an indication that the syneresis complex, although highly faulted, is vertically limited. Morley et al. (2017) describes the syneresis as exhibiting a

honeycomb-shaped morphology, with the structures approximately 200 m in diameter.

Morley et al. (2017) describe syneresis as the bulk contraction of the sediment

accompanied by fluid expulsion (a diagenetic origin). I use a different color scale so

that more information can be extracted from the same dataset.
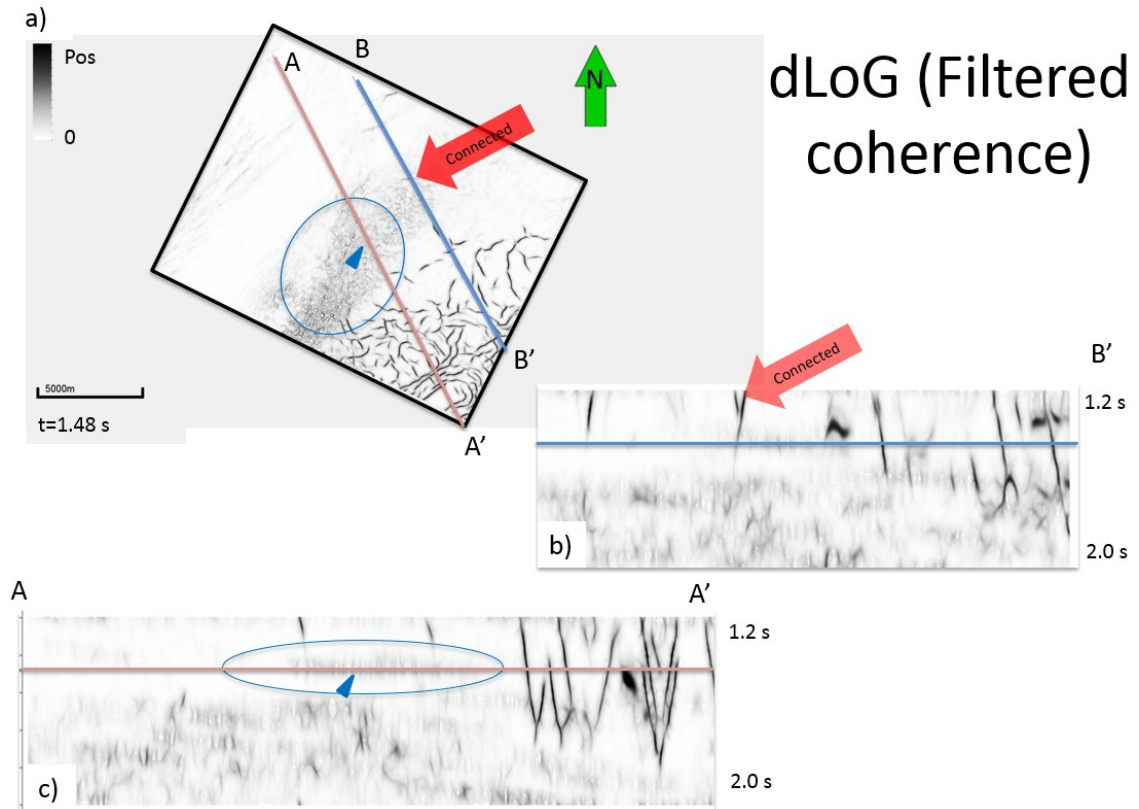
Figure 2: (a) Time-slice and (b, c) cross-section through a dLoG filtered coherence volume. The polygonal faults in the SE corner connect the hypothesized aquifer at t=2.2 s to the hypothesized reservoir at t=1.2 s. In contrast, the syneresis (commonly called "shale dewatering") features do not connect the two horizons and in general are disjoint. Two red arrows point to examples of connected faults.
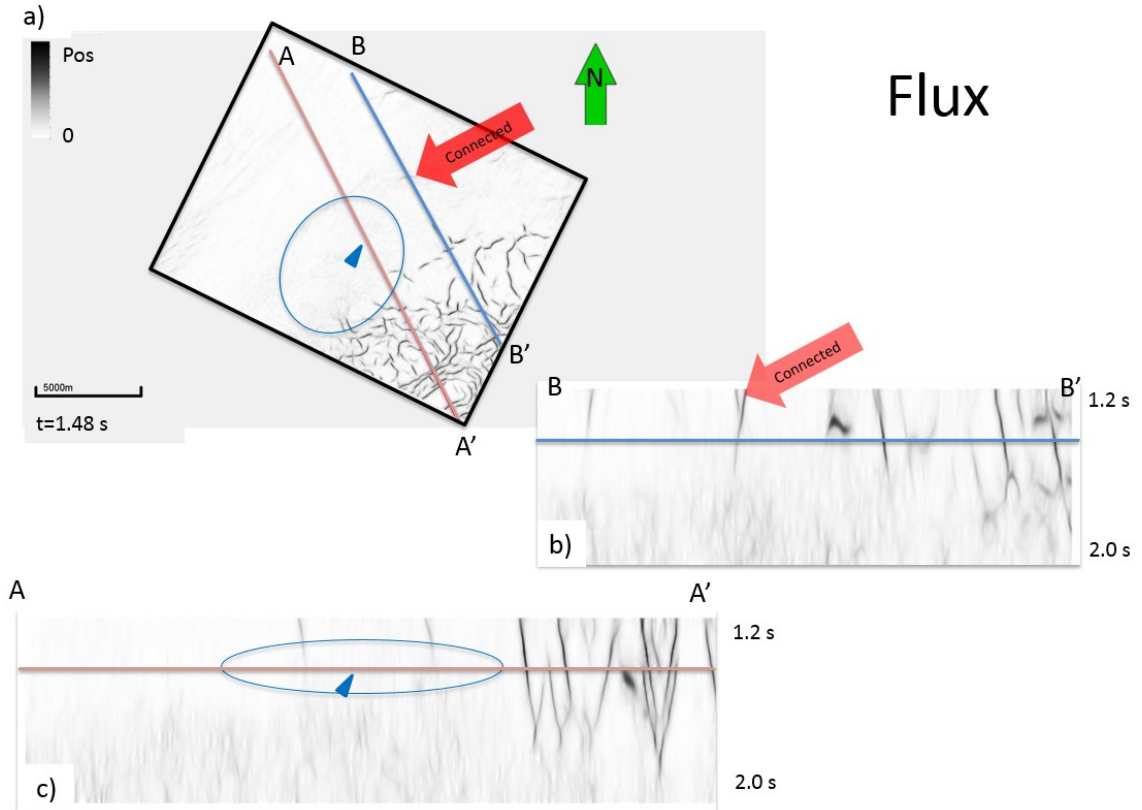
Figure 3: (a) Time-slice and (b, c) cross-section of the computed flux. Blue triangle point to areas that, although are stronger on dLoG filtered coherence and probably constitute syneresis complexes, are not connected from t=1.2 to t=2.2 s. The red triangle is used to point faults that became more visible. The cross-section shows a more amplitude balanced compared with the dLoG filtered coherence. Two red arrows point to examples of connected faults.
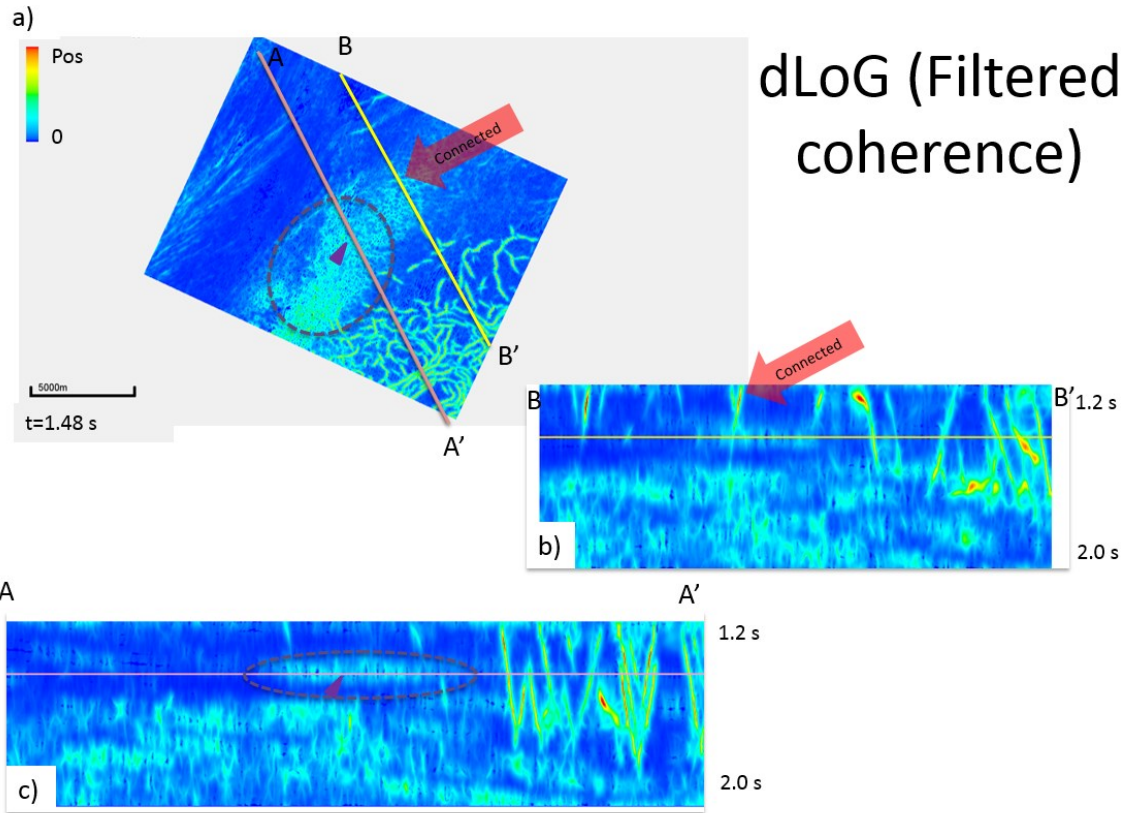
Figure 4: (a) Time-slice and (b, c) cross-section through a dLoG filtered coherence volume. Refer to Figure 2 for detailed description.
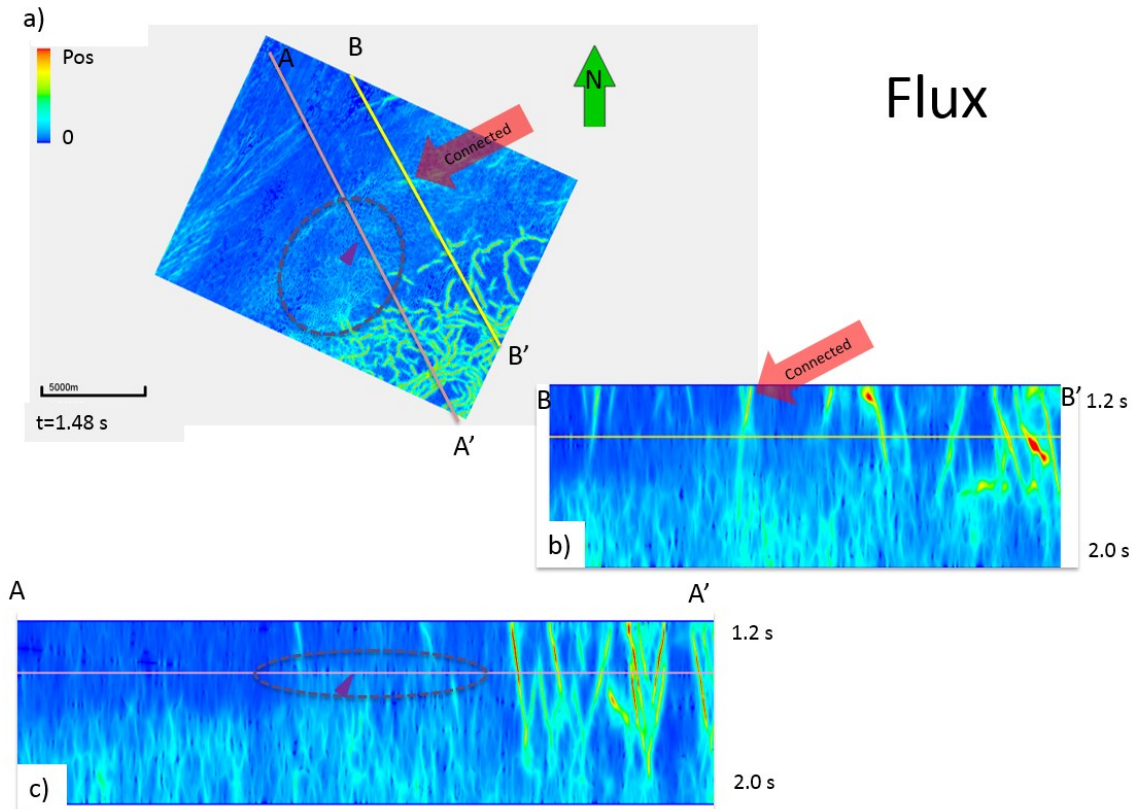
Figure 5: (a) Time-slice and (b, c) cross-section of the computed flux. Refer to Figure 3 for detailed description.

Example 2: Modeled "flow" through radial faults

Infante and Marfurt (2017) described the seismic expression of igneous rocks associated with the Miocene-age Kora volcano in the Taranaki Basin. Volcanic rocks have become particularly attractive for the oil industry in the last two decades, with significant quantities of hydrocarbons being produced from reservoirs drilled in volcanic rocks in China, New Zealand, and Argentina. Infante and Marfurt (2017) use a multi-disciplinary approach to map pyroclastic and lava flows within the Kora volcano and conclude that the cause of the dry well (Kora-4) was due to a lack of source charge and not to the absence of reservoir quality. The Kora-1 well drilled a thick sequence (>1000 m) of pyroclastic flows in this submarine volcano by chance and found high

19

peaks of gas in the mudlogs near the top 25m of this sequence. A long-term test in this upper volcanic section resulted in 32 API oil flow of 668 BOPD for 254 hours. (Infante and Marfurt, 2017).

As one possible cause for the dry well is insufficient charge, I will apply my fault connectivity algorithm to better understand the fault connections between the source and reservoir (the grayish transparent fault on Figure 6).

The magma conduit of the volcano appears incoherent on the 3D seismic data volume. However, I have no reason to believe the magma conduit will have a high hydraulic conductivity in this region. This doesn't mean that igneous rock will never have a high hydraulic conductivity compared with the background geology. As an example, the Serra Geral formation on the Parana Basin in Brazil is composed of igneous intrusions and still holds an important aquifer for the region. Bortolin et al. (2014) performed a hydrogeologic and hydrogeochemical study and concludes that the wells with higher water production are located in areas with a high fracture density.
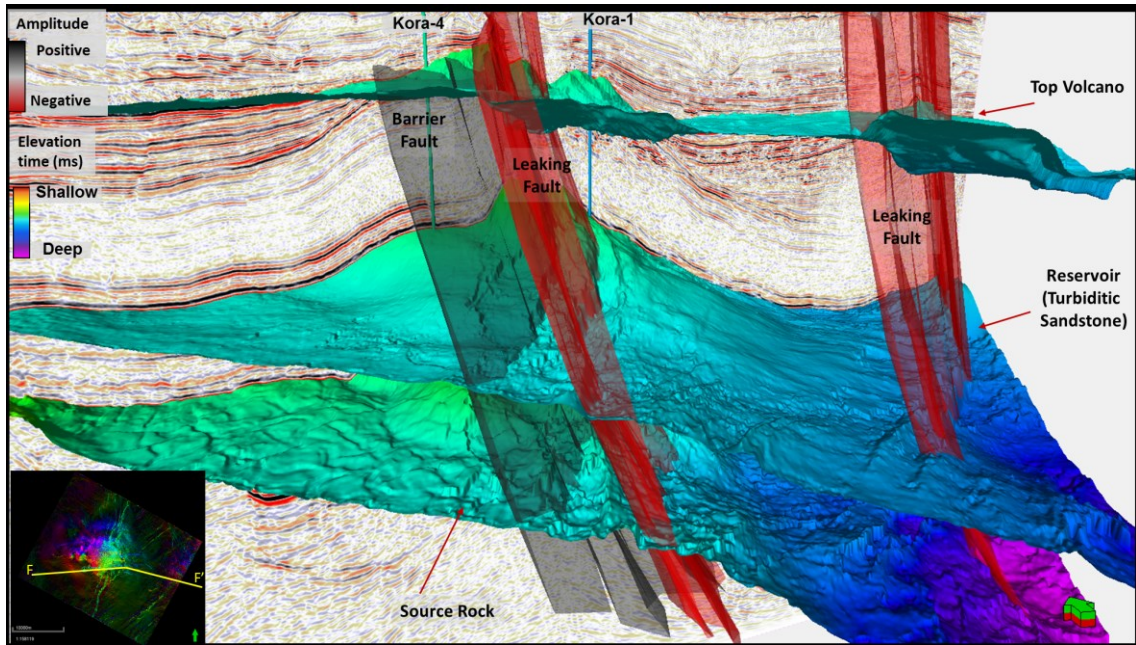
Figure 6: Seismic amplitude, horizons, interpreted faults, and wells for the Kora volcano in the Taranaki basin. Three faults are interpreted on this dataset. Infante and Marfurt (2017) do not believe that the grayish fault - barrier fault - constitute a migration path for the oil coming from the source rock below. (Infante and Marfurt, 2017)

To cope with the low coherence – low hydraulic conductivity (opposite of what I hypothesized for faults) I interpreted the base of the volcano and I set the volcano core to be a hydraulic insulator.

Figure 7 - Figure 9 show vertical and time slices through the seismic time migrated amplitude data, the dLoG filtered coherence (which is the input for the modeling computations) and the computed flux volumes respectively. The potential head was set to be low at the top volcanic reservoir and high on deep source rock horizon.

As in the GSB example, the flux computation balances the data and highlights areas that are connected (red triangles). Some laterally extensive but vertically limited geological features (blue triangles), including turbidites systems, are blurred by my

21

algorithm, losing their anomalous amplitude. Areas with anomalous values on the dLoG

filtered coherence volume that are not connected became weaker after computing the

flux results (purple triangles).
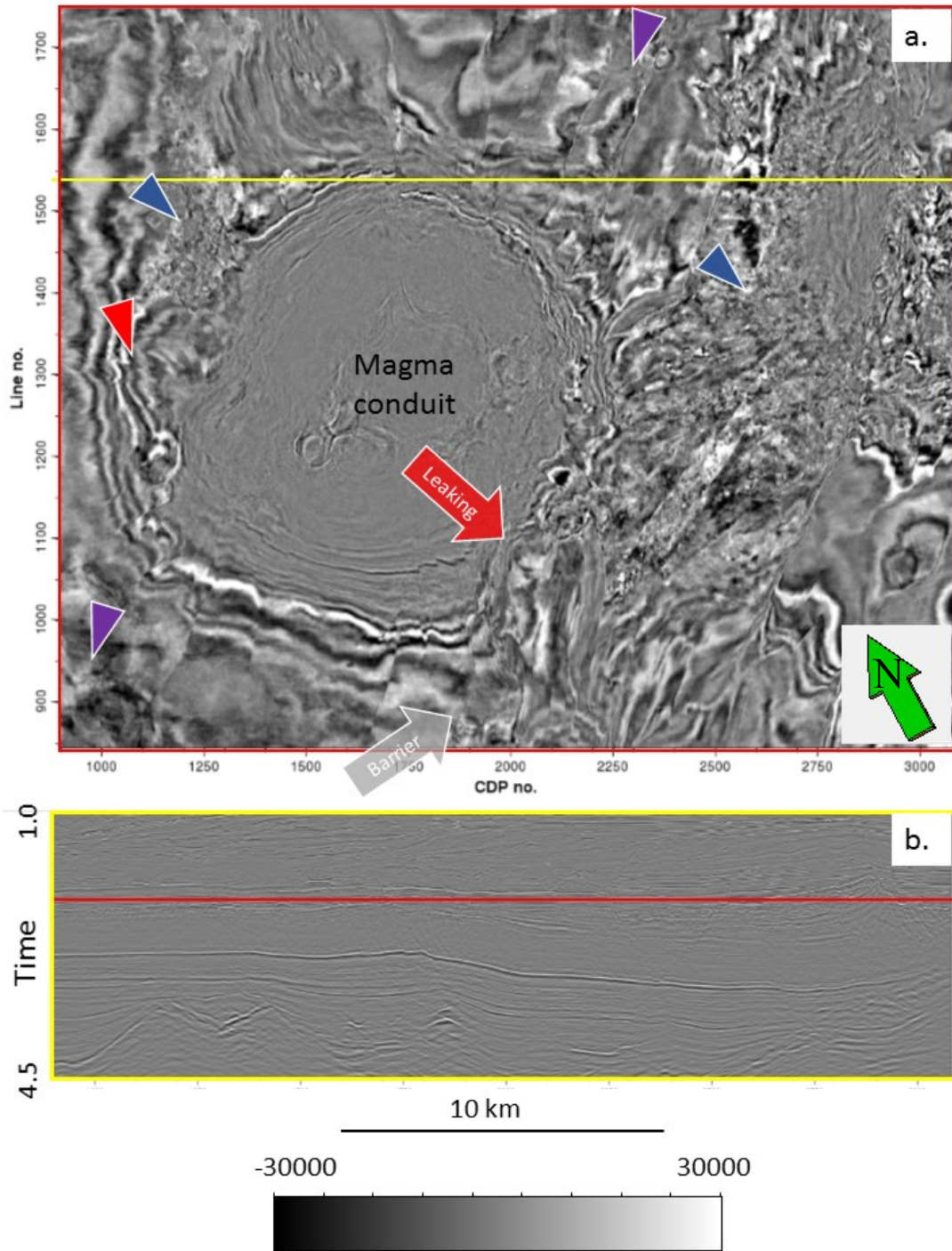
Figure 7: (a) Time-slice at t= 2.14 s and (b) vertical slice along AA' through the seismic amplitude volume. Arrows indicate the position of the leaking and barrier faults. Triangles are discussed in Figure 8 and Figure 9.
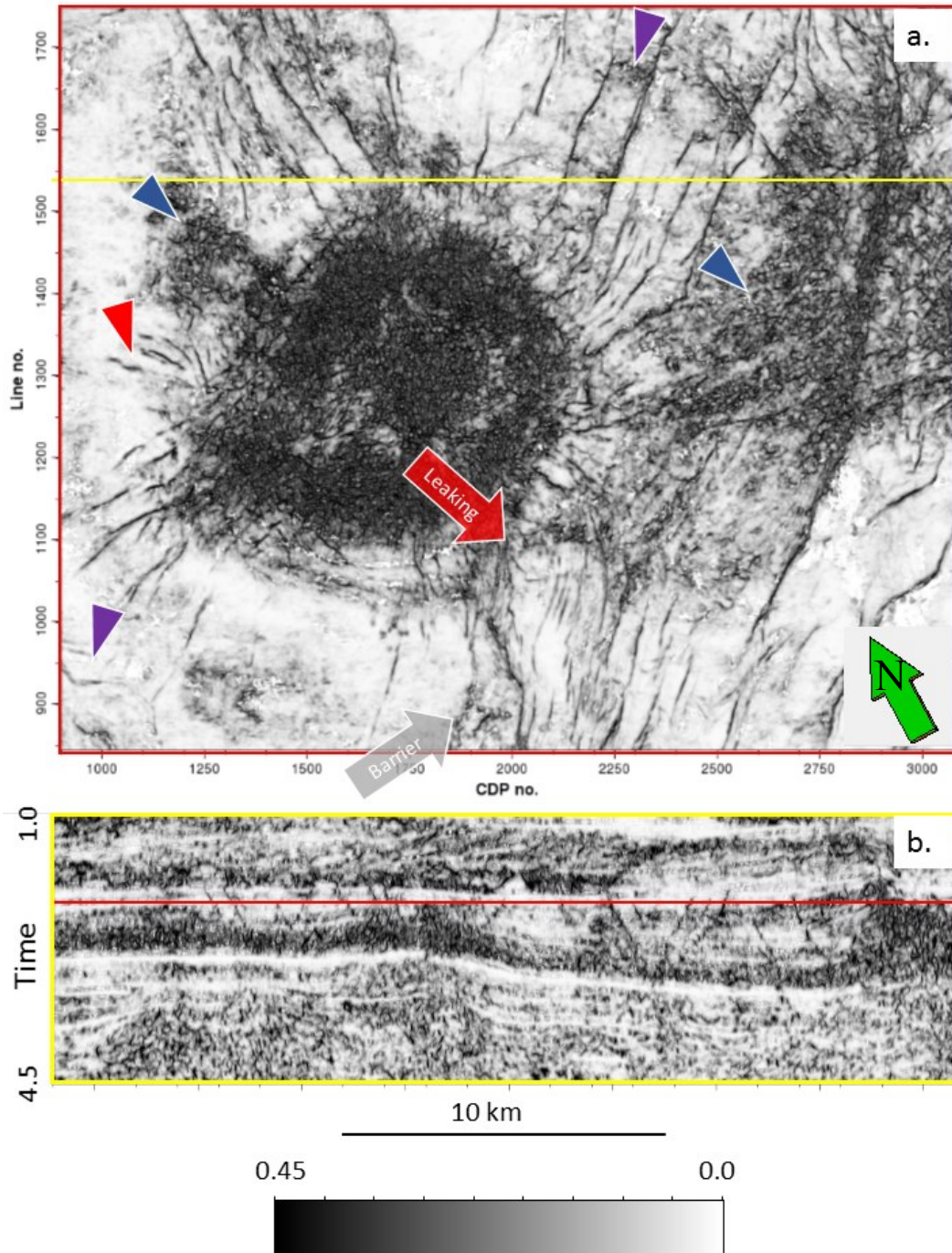
Figure 8: (a) Time-slice at t= 2.14 s and (b) vertical slice along AA' through the dLoG filtered coherence volume. The magma conduit has a high value of the dLoG, although I have no reasons to believe it is truly a higher hydraulic conductivity area. The blue triangles indicate to turbidites generated on the flanks of the volcano. The red triangle indicates a fault that becomes more continuous on the flux results. The purple triangles indicate to areas that became weaker on flux results because they are not connected.
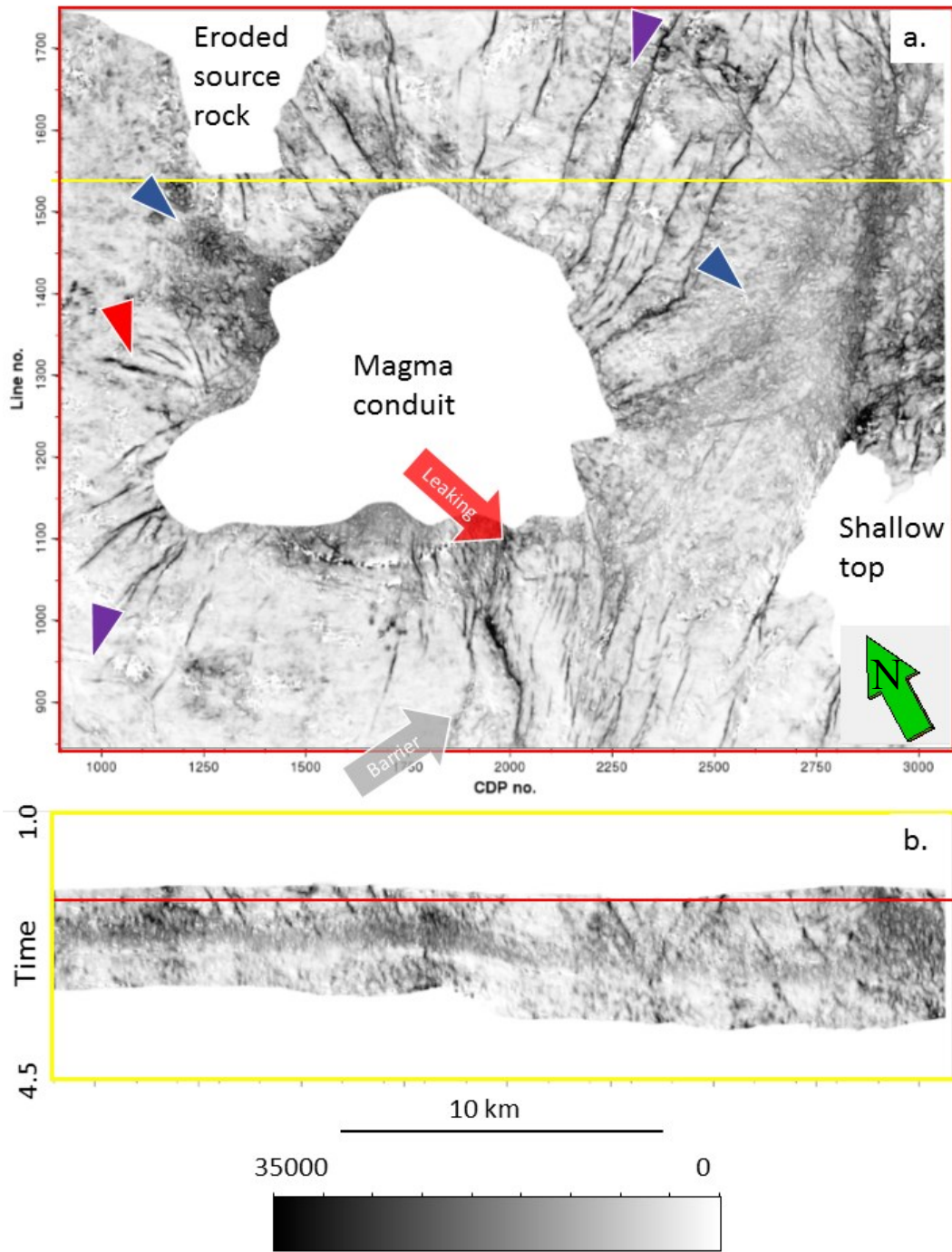
Figure 9: (a) Time-slice at t= 2.14 s and (b) vertical slice along AA' through the computed flux. The flux computation is limited to the zone between the top volcanic reservoir and the top of the source rock horizons. Also, the area interpreted as the base of the volcano has no flux (low conductivity).

Example 3: Modeled "flow" through karstic features

The Bemis-Shutts dataset is a depth migrated seismic acquired on the largest

field in Ellis county, Kansas, the latter of which has the highest reported oil production

in 2016 (Figure 10).
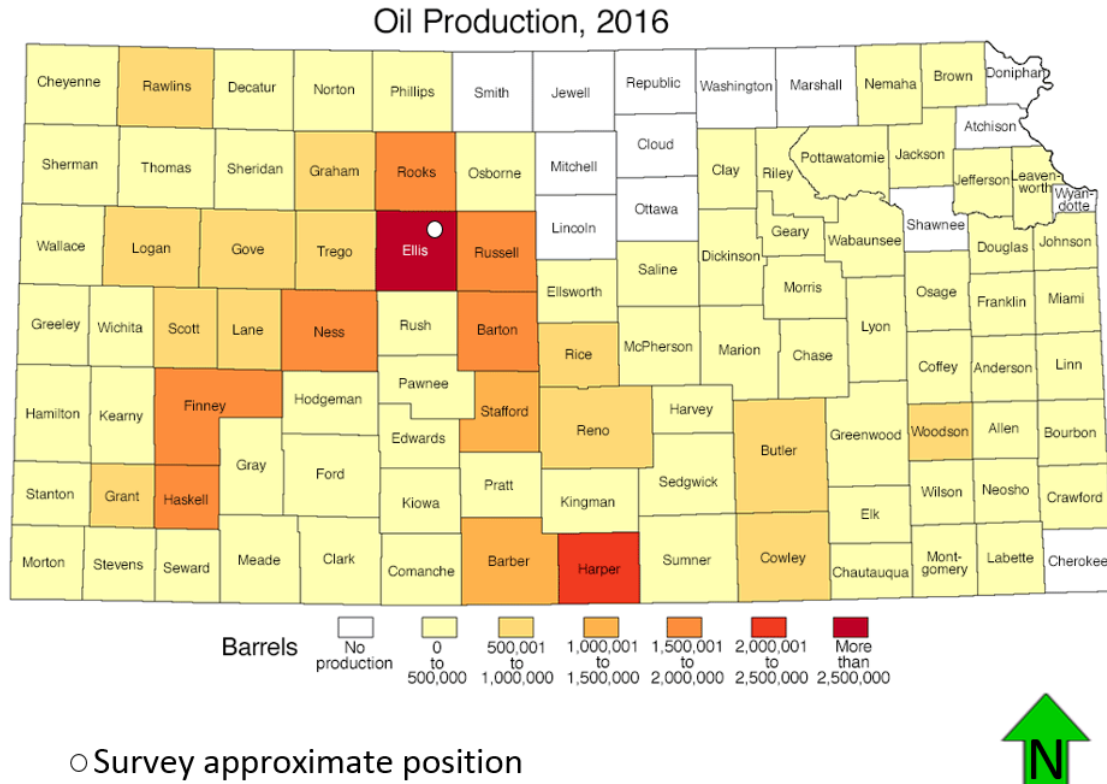
Figure 10: Map of Kansas with county oil production (adapted from Kansas geological survey, http://www.kgs.ku.edu/PRS/petro/interactive.html, accessed July 7, 2017)

The fracture-controlled karstic carbonate reservoirs within the Arbuckle group

are responsible for production of about 2.19 billion barrels of oil (BBO) and represent

approximately 36% of the total Kansas oil production. Up to 2003, production from the

Arbuckle in the Bemis-Shutts field alone accumulated over 240 million barrels

(Franseen et al., 2003).

Ferril and Morris (2008) describe the effect of clay content in faults and deformations in a carbonate controlled fault system in Texas. Faults with shallower dips (60° or less) are developed in clay-rich limestones while clay-poor massive limestones and dolomites tend to generate steep faults (70° or more).This lithology behavior may be helpful if extrapolated to other areas. Estimating the faults dip may be useful to infer carbonate or clastic lithology if no other information is available. However, the Bemis-Shutts seismic survey I present in this thesis has no significant faults and most geologic features highlighted by coherence images have karst characteristics.

Franseen et al. (2003) stated that studies indicate that the connectivity of the uppermost oil-productive Arbuckle strata with the underlying aquifer is extremely variable. Drill Stem Test (DST) data in numerous major mature fields show that reservoirs well connected to the underlying aquifer system will have pressures similar to the aquifer pressures (approximately 1100 psi). Poorly connected reservoirs exhibit low infill well DST pressures below aquifer pressures (generally around 900 psi). If the reservoir is not connected, the pressure can drop as low as 250 psi. The variability in connection of the upper Arbuckle strata to the underlying aquifer probably relates to a mixture of karst, structural, stratigraphic, and other diagenetic controls. Because most of the oil and gas zones in the Arbuckle are close to the top, wells are usually competed into the top of the Arbuckle with penetrations under 10 ft or 3 m (Franseen et al., 2003).

Holubnyak et al. (2014) find that fractures in the Arbuckle are correlated with volumetric curvature and use this relation to tune the geologic model to study carbon sequestration. Nonetheless, I will use the seismic coherence as a proxy for hydraulic conductivity. The curvature attribute is smoother when compared to the coherence. If

used without a thoughtfully pre-conditioning as input for my algorithm, the computed head would have no issues in flowing through all the survey. Coherence should provide a more separable petrophysical background. Unfortunately, at this moment, I do not have water production data information to correlate the results I obtain.

Using public well information from the Kansas Geological Survey, I pick the top of the Arbuckle in a depth migrated seismic data from the Bemis-Shutts field. I use this picked horizon as the top input horizon for my algorithm. I then shift the horizon 300 ft down to be the bottom horizon representing the overpressured aquifer.

Figure 11 shows the data I use as input for my algorithm, Figure 12 shows the computed flux normalized results. Because of the horizons used (the bottom is a shifted version of the top), the small number of samples between the horizons (around 30) and the geological setting of the area (semi-vertical features in the coherence images), the flux of the possible karst features resemble vertical lines.
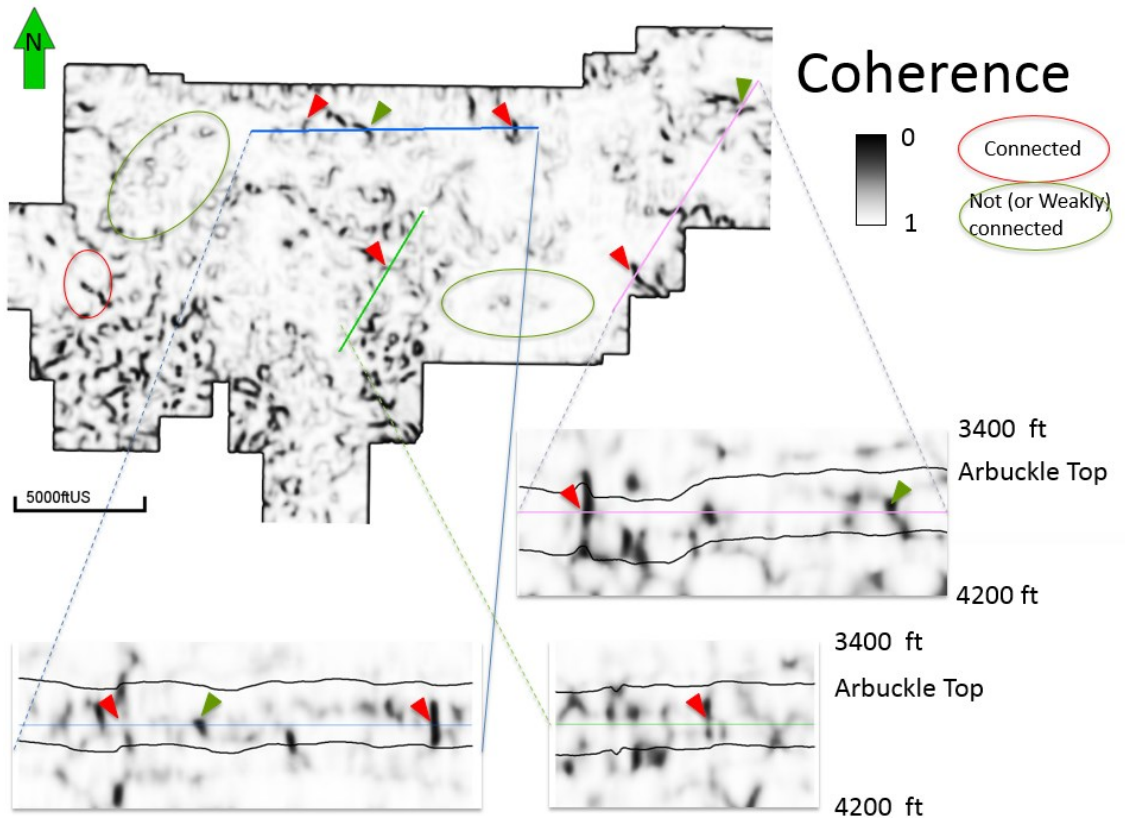
Figure 11: Depth-slice at 3780 ft and indicated vertical slices through the computed coherence. Red triangles and ellipses indicate connected features while green triangles and ellipses indicate disjoint or weakly connected geological features.
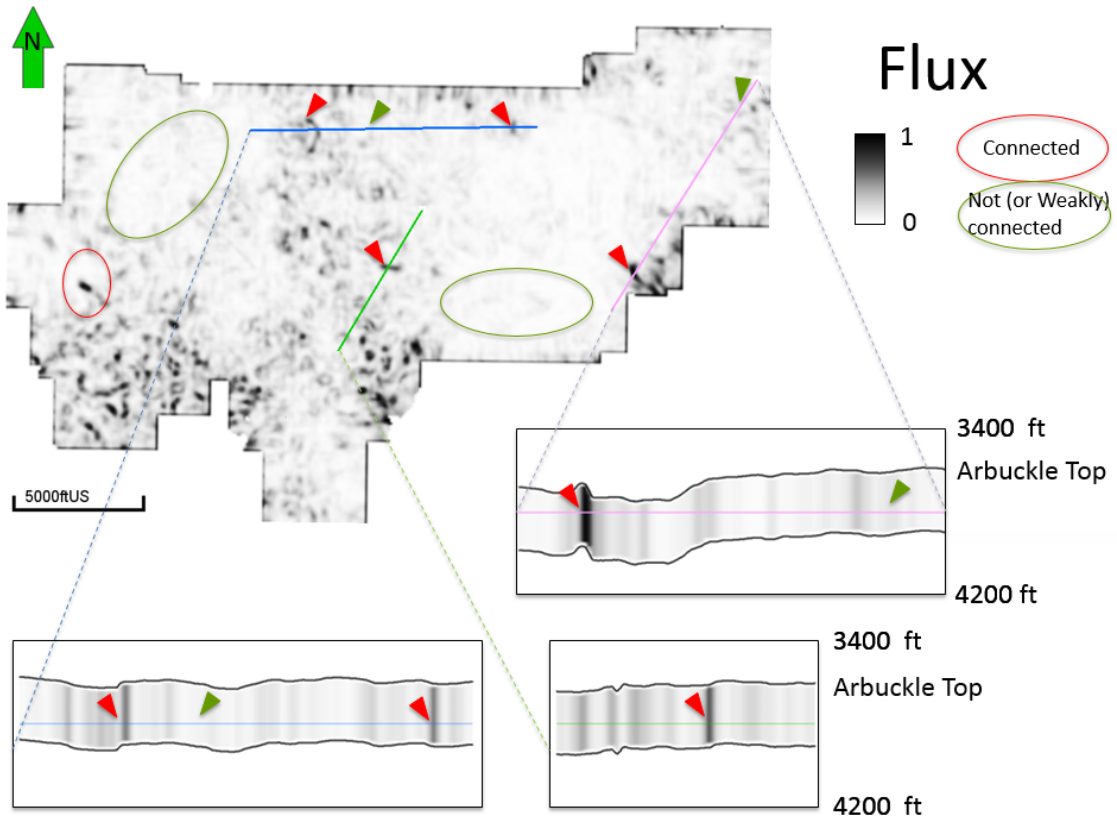
Figure 12: Depth-slice at 3780 ft and indicated vertical slices through the normalized flux. Red triangles and ellipses indicate connected features while green triangles and ellipses indicate disjoint or weakly connected geological features. The small space between top and bottom horizons (around 30 samples), the similarity between top and bottom horizon, and the input coherence make the flux looks almost vertical in the vertical slices.

Chapter 4: Water production application: Fairview

Unconventional shale gas now account for a significant part of the produced gas in the USA. One of the first unconventional shale gas fields in the United States, the Barnett Shale is the main source rock for oil and gas in the Fort Worth Basin (Jarvie et al.,2007).

The deeper Viola, Forestburg, and Marble Falls Limestones are water-bearing and act as fracture barriers (Perez-Altamar, 2013). Therefore, a fault connecting the petroleum producing Lower Barnett with the Ordovician Viola Limestone can conduct water from the deeper aquifer to the reservoir.

Figure 13Figure 16 show seismic attributes values extracted on the Lower Barnett horizon – amplitude, dip magnitude, coherence, and finally the dLoG filtered coherence respectively that I use as input for flux computations. The dip magnitude image (Figure 14) exhibits an important geometrical feature that can be inferred as a sink hole, or karst collapse. The coherence and dLoG filtered coherence don't show this important feature as clearly as the dip magnitude does.

I set the horizon 20 ms above the Lower Barnett as the top horizon with a low head value and the horizon 20 ms below the Viola with a high head value. I then compute the flux and normalize the results.

Using well data, I compare the relationship between the results I obtained and the water production for each well. To do that, I set a circle with 1000 ft diameter around the well position and I numerically integrate the horizon extracted values for the dLoG filtered coherence and the normalized flux. I expect to see a higher correlation

31

between the numerically integrated flux than the dLoG filtered coherence. However, the resulting correlation is disappointing.

Using the gas and water production data available, I generate a water-to-gas ratio so I can reduce the bias that could be produced by an increase in production in general. I then plot the median of the water-to-gas ratio on the same map as the normalized flux (Figure 18). Figure 19 shows the normalized flux extracted at the Lower Barnett horizon, the water to gas ratio median, and the dLoG filtered coherence numerical integration, while Figure 20 follows the same scheme, but presents the normalized flux numerical integration.

The possible sink hole in the middle of the image close to well ID 12 seems to have a great influence in the water inflow. Well 12 is by far the one with the highest peak of water production.

In Figure 21 the relationship between dLoG filtered coherence and flux numerical integrations becomes visible. Figure 22 shows the dLoG filtered coherence numerical integration and the median of the water to gas ratio while Figure 23 compares flux and water to gas ratio. There's a proportionality between both seismic attributes. I see this as a sign that the application of the algorithm described in this thesis was not completely adequate for this dataset.

Because the water-to-gas ratio may also be related to the well pressure, I show in Figure 24 and Figure 25 the water-to-gas ratio vs. the shut-in pressure for both the well with the higher dLoG/flux numerical integration and for the well with the higher water-to-gas ratio.
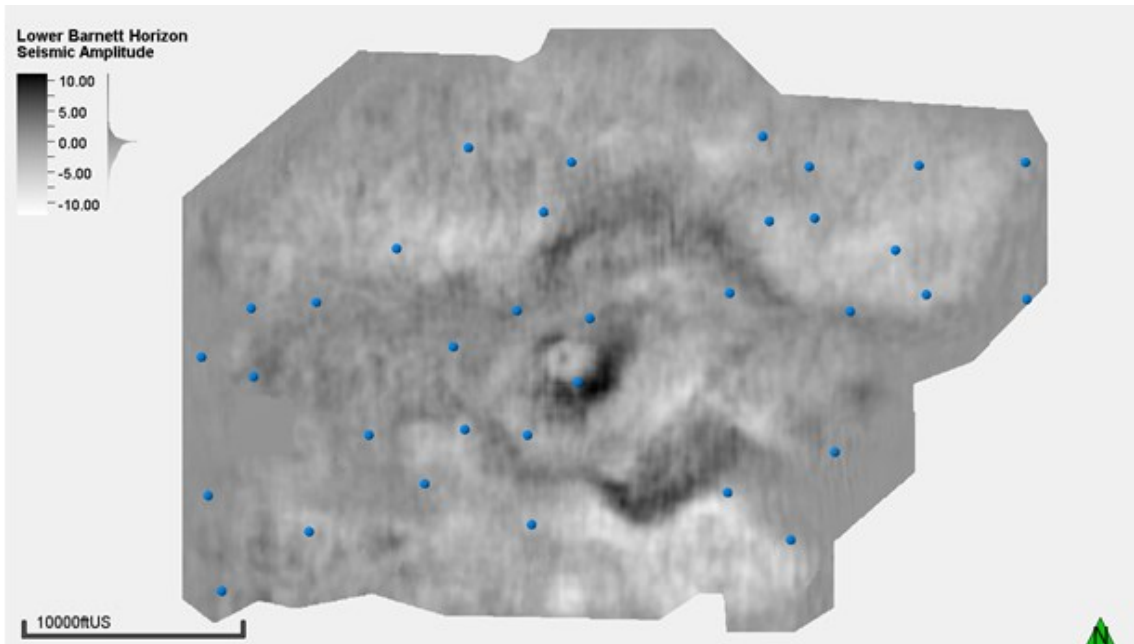
Figure 13: Seismic amplitude data of Lower Barnett and well location (blue points).
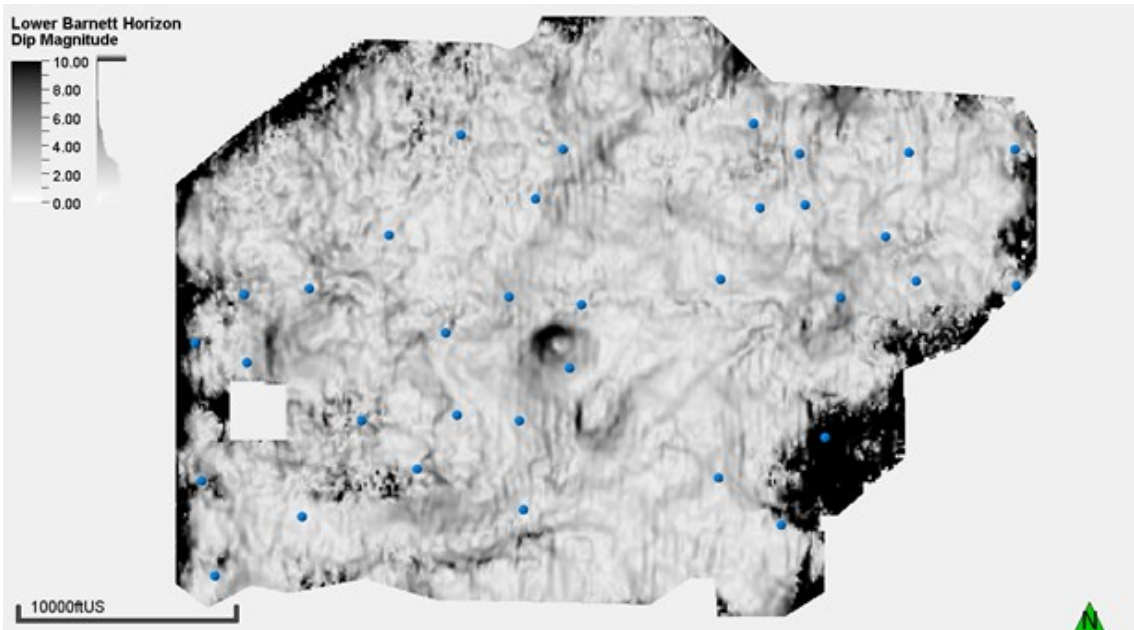
Figure 14: Dip magnitude of Lower Barnett horizon and well location (blue points). Stratigraphic geometrical features become more pronounced with this attribute extracted from the seismic amplitude data. Note the characteristically sink hole seismic dip magnitude signature on the center of the image.
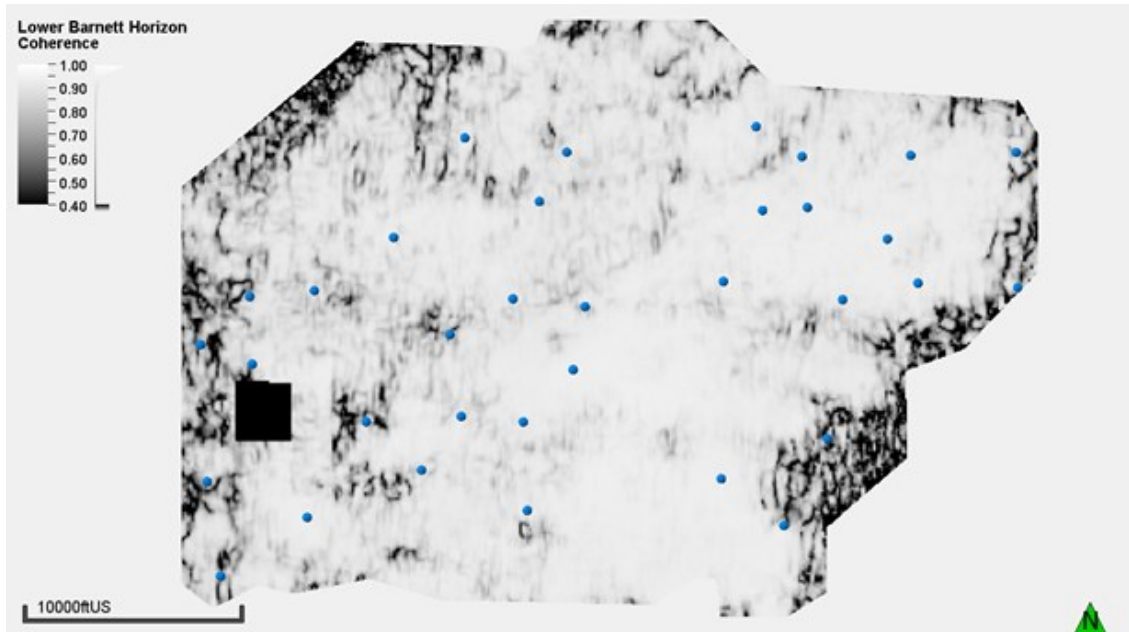
Figure 15: Coherence of Lower Barnett horizon and well location (blue points). This edge detection algorithm shows misalignment between seismic reflector. The possible sink hole feature on the center of the image is not so clear with this attribute.
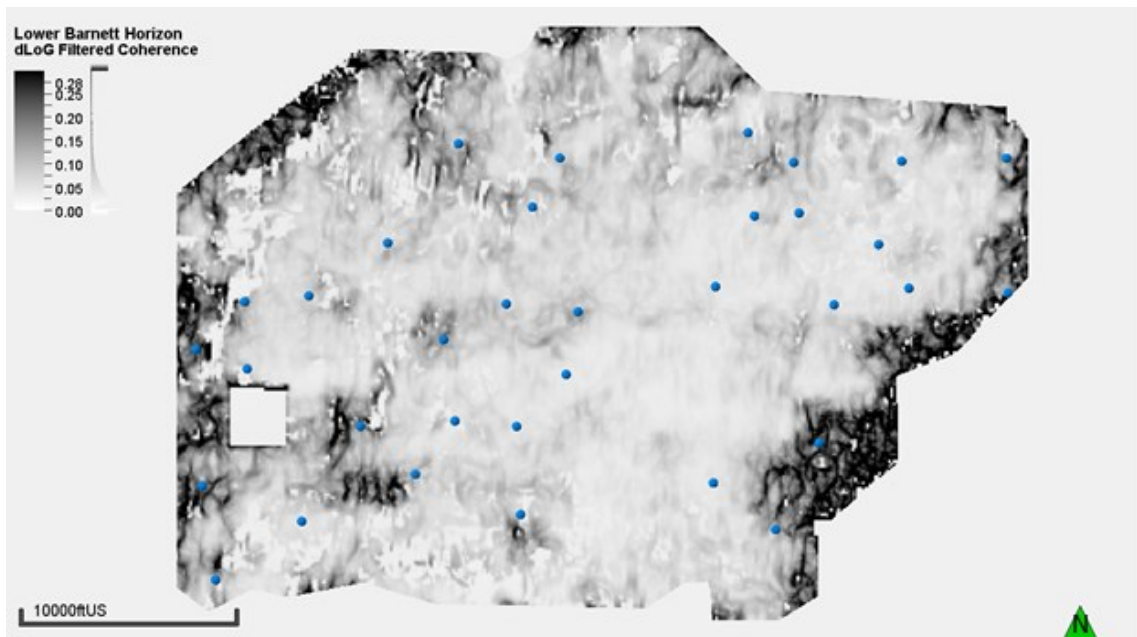
Figure 16: dLoG filtered coherence of Lower Barnett horizon and well location (blue points). This filter makes the coherence response smoother.
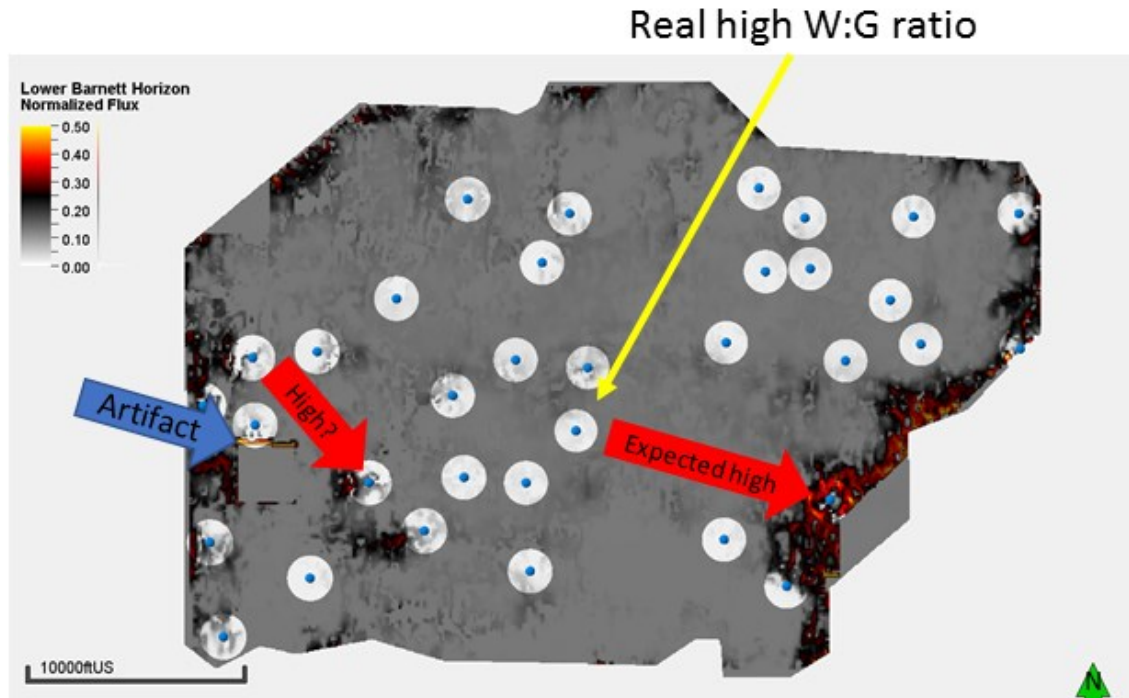
Figure 17: Normalized flux of Lower Barnett horizon and well location (blue points). The blue arrow points to an artifact created on the edge of the clipped area due to the acquisition parameters. The red arrows point to areas that I expect to have a higher water production because of my hypothesis and the yellow arrow points to the well drilled close to the sink hole-inferred feature, the well located there has the higher water to gas ratio of the survey. The image is clearer on a 1000 feet radius around the well. That's the area I use to compute the values from the seismic attributes.
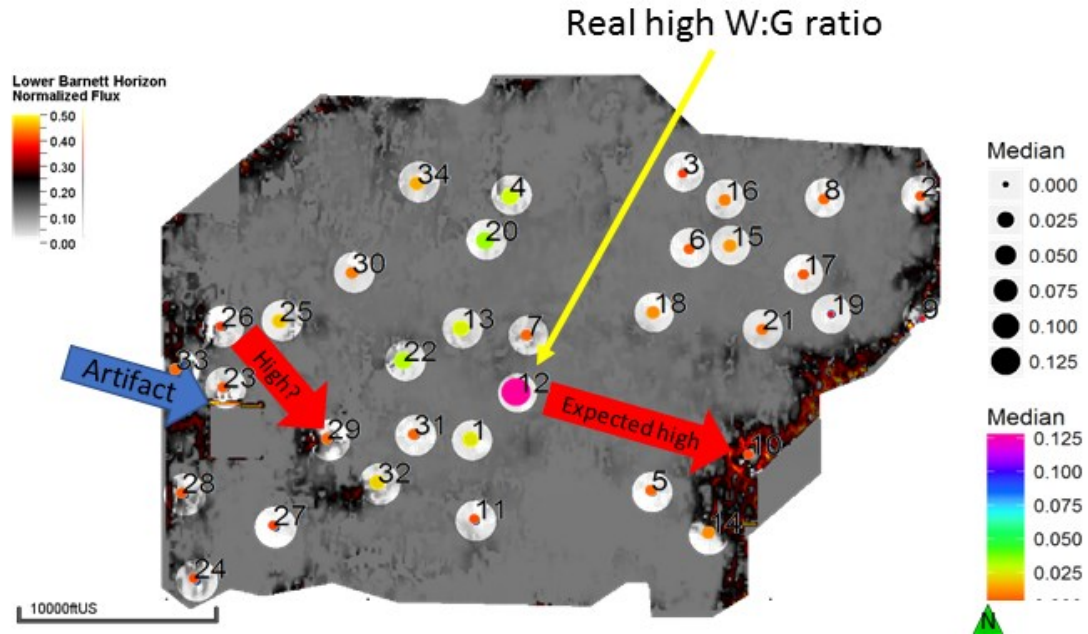
Figure 18: Normalized flux of Lower Barnett horizon, and the median of the water to gas ratio. The circle colors and circle sizes are based on the median of the water to gas ratio of the respective well. Note the well with the higher water production lies on the seismic-inferred sink hole.
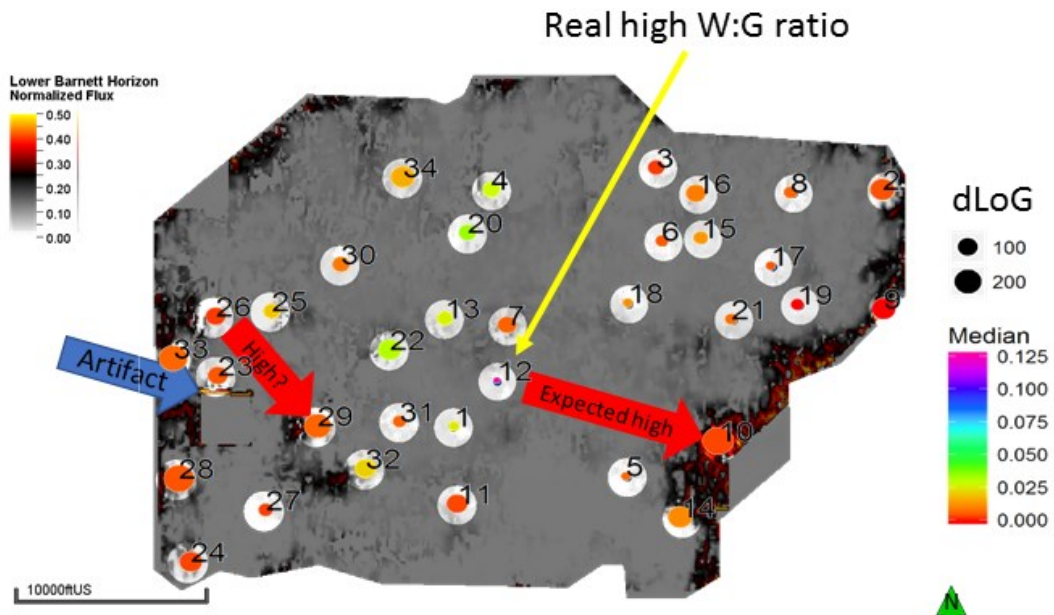
Figure 19: Normalized flux of Lower Barnett horizon, the dLoG filtered coherence area influence, and the median of the water to gas ratio. The circle colors are based on the median of the water to gas ratio of the respective well, the circle size is draw according to the value computed for the dLoG filtered coherence for a circle centered at the well with 1000 feet radius. It does not seem to be a correlation between water production and the dLoG filtered coherence.

36
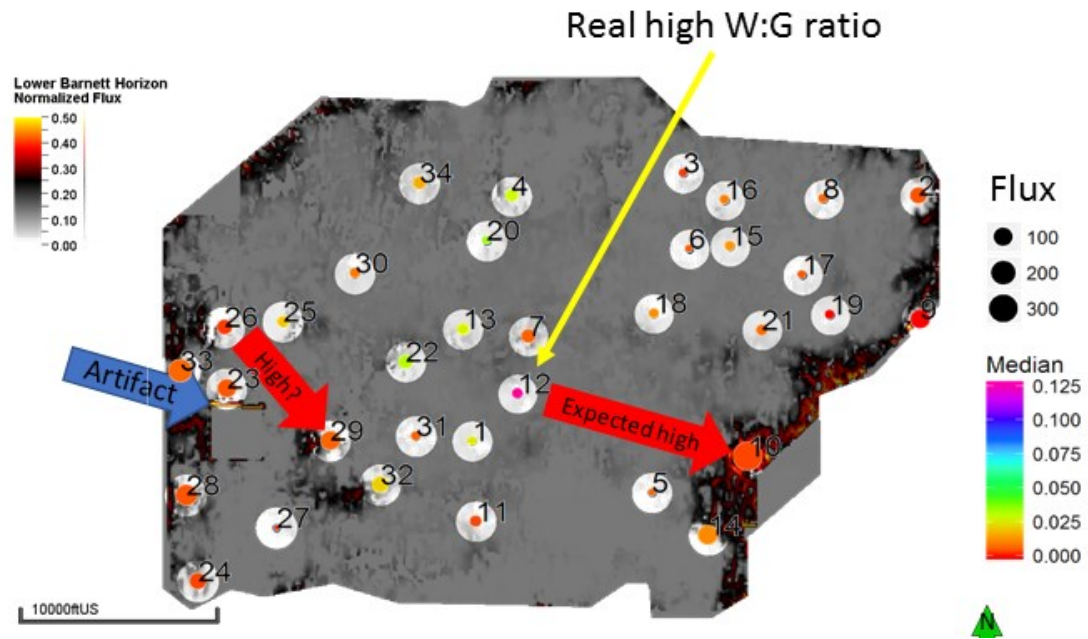
Figure 20: Normalized flux of Lower Barnett horizon, the normalized flux area influence, and the median of the water to gas ratio. The circle colors are based on the median of the water to gas ratio of the respective well, the circle size is draw per the value computed for the normalized flux for a circle centered at the well with 1000 feet radius. It does not seem to be a correlation between water production and the normalized flux. The areas with expected high (arrows) don't have a comparable higher water production.
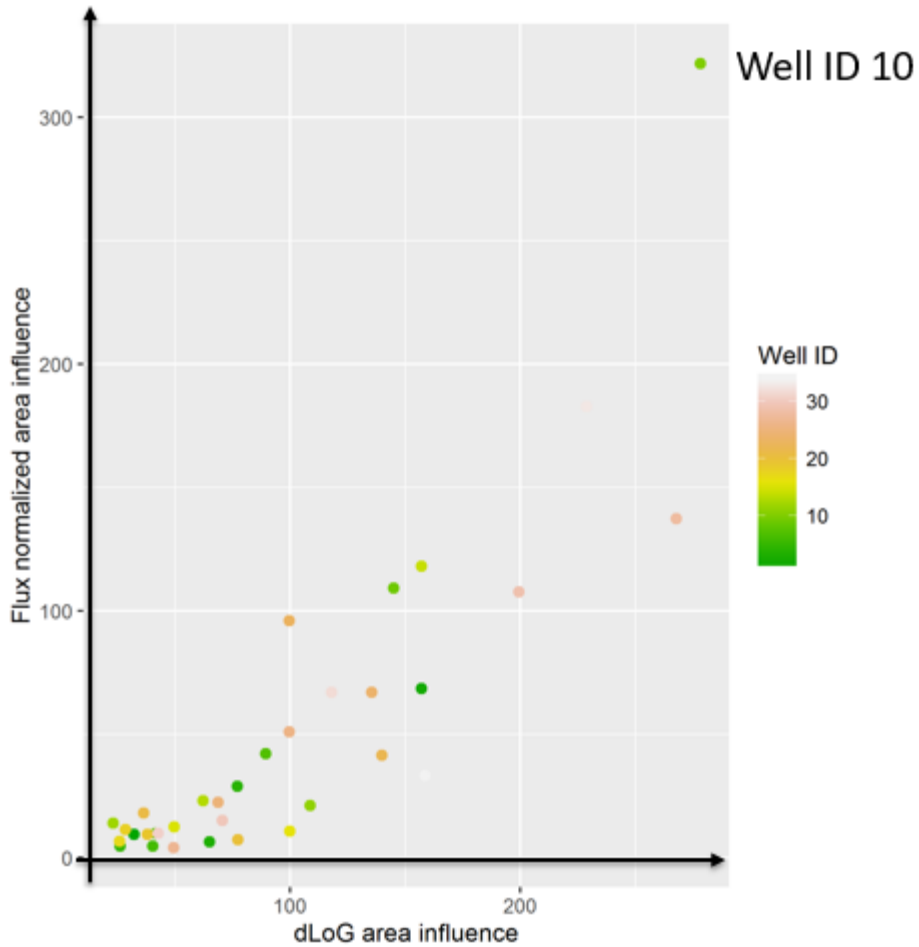
Figure 21: Comparison between flux and dLoG filtered coherence. The axis values are results from the computation of the integration of the seismic attribute value on a circle with radius of 1000 ft around the well. Well with ID 10 is the point with anomalous flux/dLoG values. The plot shows a possible correlation between dLoG filtered coherence and the flux.
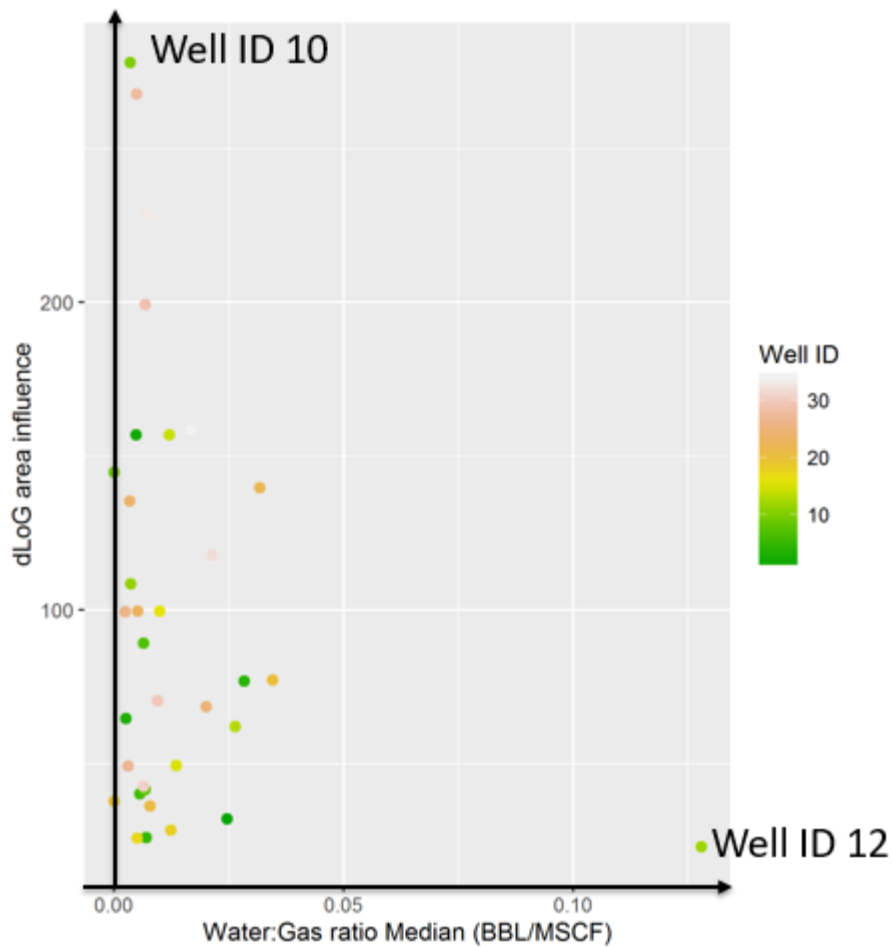
Figure 22: Relationship between dLoG filtered coherence numerical area integration and the median of water to gas ratio (water production indicator) for the wells. Well 12 has an anomalous high water production, while well 10 has an anomalous dLoG value.
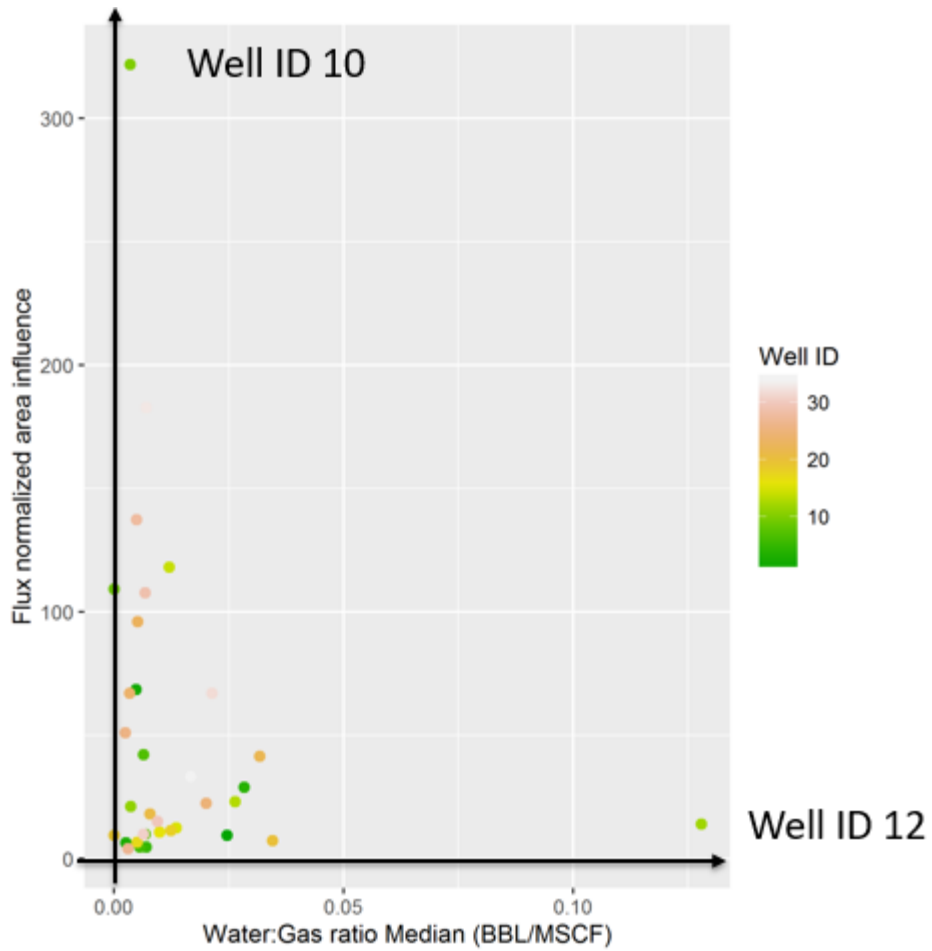
Figure 23: Relationship between normalized flux numerical area integration and the median of water to gas ratio (water production indicator) for the wells. Well 12 has an anomalous high water production, while well 10 has an anomalous normalized flux value.
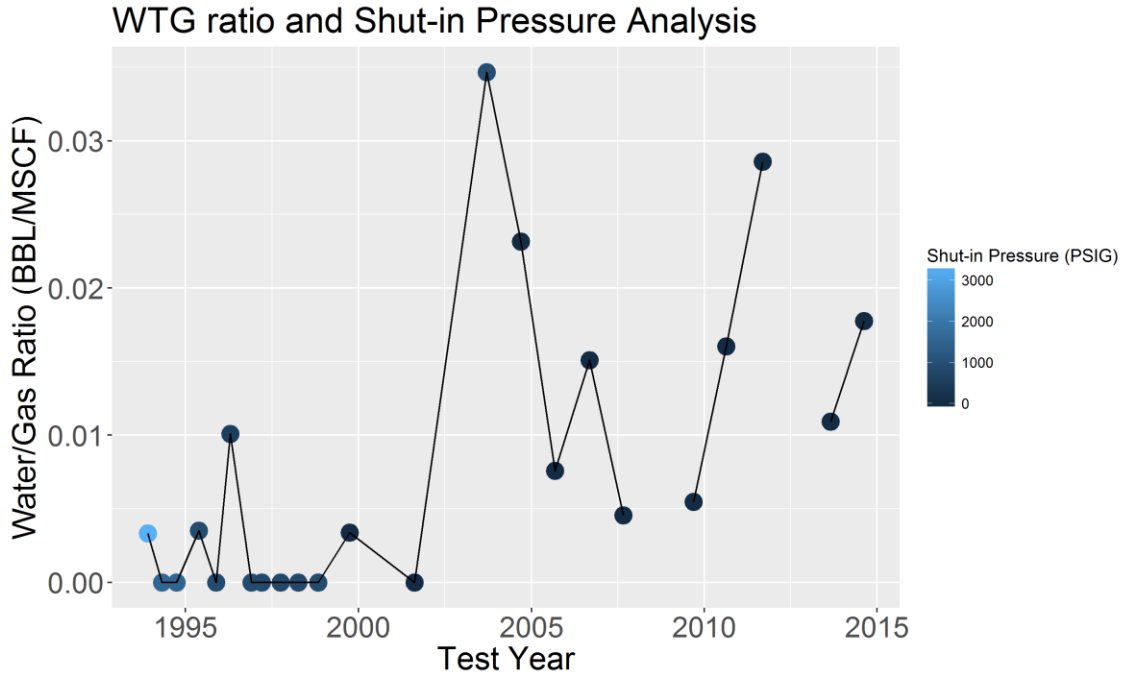
Figure 24: Water to gas ratio and shut-in pressure for the well ID 10. This is the well that had the higher values for the dLoG filtered coherence and flux numerical integration. However, the water to gas ratio is low.
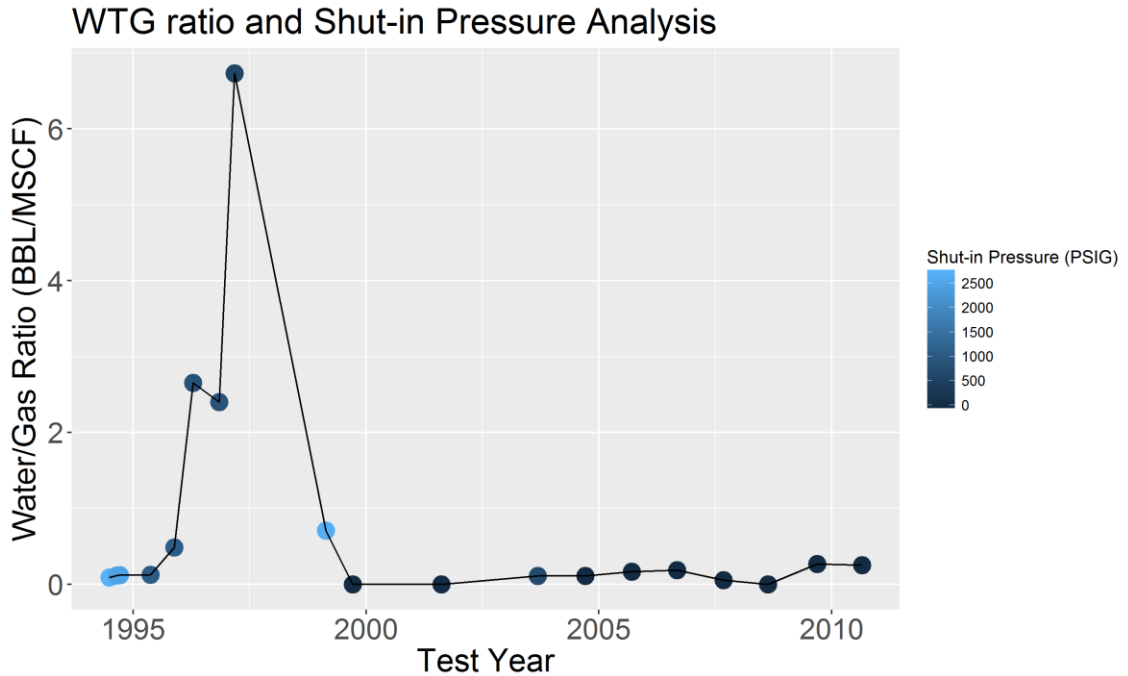
Figure 25: Water to gas ratio and shut-in pressure for the well ID 12. The peak water production between 1995 and 2000 is noticeable, however the median is robust regarding outliers. The shut-in pressure does not show a strong correlation with the water to gas ratio. A proportion relationship between water production and shut-in pressure may be related with a new fracking.

Chapter 5: Future work: using the algorithm to predict head and infer pressure

Synthetic model

Using a synthetic hydraulic model, I simulated a steady-state solution with a known pressure gradient. Then, using an initial guess, I iterate to recover the values used to create synthetic well pressure values using simulated annealing (Appendix C).

Figure 26 shows the hydraulic model where I assume a pressure gradient of 0.30 psi/f. Figure 27 shows the resulting steady-state solution.
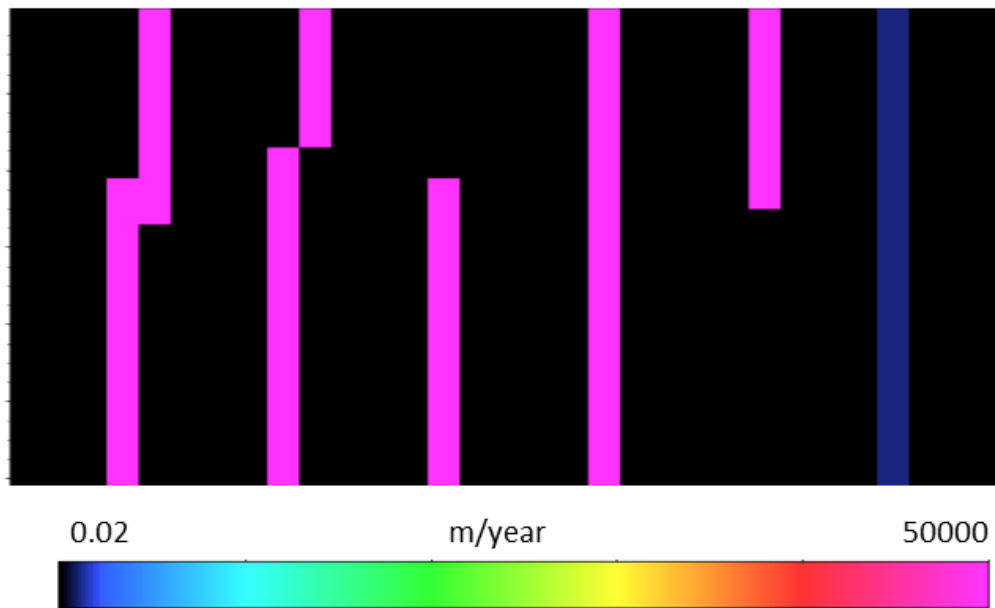
Figure 26: Hydraulic conductivity model used to generate synthetic pressure points (wells). The blue fault on the right has hydraulic conductivity of 2000 m/y.

The synthetic model consists of Nx=Ny=Nz=31 cells that measure Δx=12.5 m, Δy=25.0 m and Δz=8.0 m. The vertical extent of the model is 248 m (or 814 ft) therefore I set the bottom pressure to be 0.30x814=244 psi.

This psi value corresponds to 171 meters of head, which is applied as the boundary condition (the fluid source), at the bottom of the model.
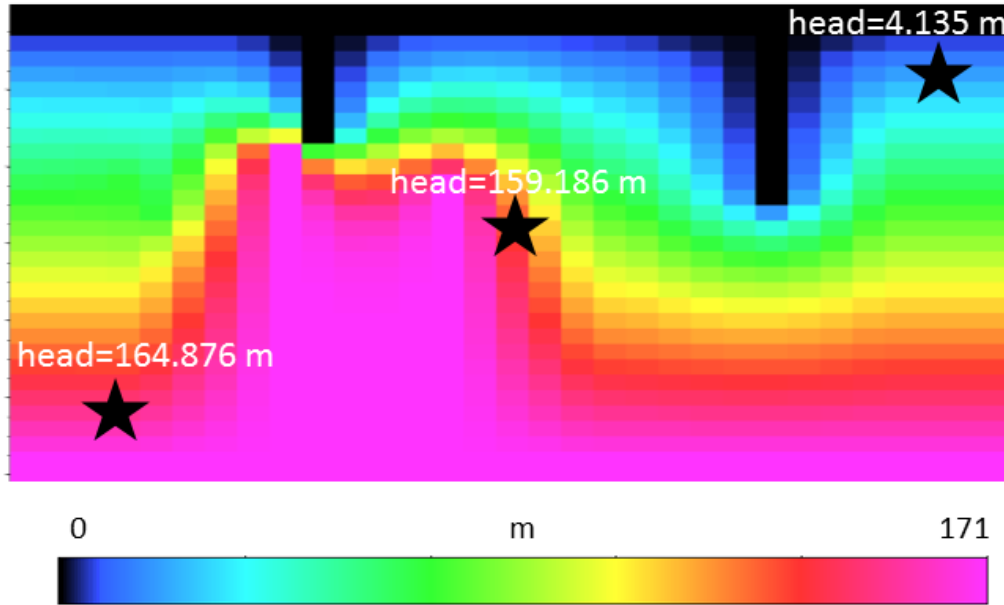
Figure 27: The computed head results. The black stars represent wells with identified pressure-head. In this synthetic simulation, I know the background hydraulic conductivity (Figure 26) and the pressure gradient of the area.

After computing the head on this synthetic model, I assume this is the correct head/pressure representation of this synthetic field. I then choose three locations to be my synthetic well head/pressure values (Figure 27). My goal is to recover the background hydraulic conductivity and pressure (head) using the algorithm described in this thesis and simulated annealing. To do that, I need a synthetic hydraulic conductivity model and an initial bottom head/pressure value.

To use the simulated annealing approach, I have to disturb my variables (hydraulic conductivity and the bottom head/pressure) looking for a combination that will give me the best answer to fit the well data. In summary, I wish to find a

combination of hydraulic conductivity (K) and bottom head/pressure for each well (h_{bw}) that minimizes the error (e):

$$e^2 = \Sigma_1^W (h_{bw} - h_{mw})^2 \qquad\qquad (3)$$

where h_{mw} is the modeled head for each well w and W is the total number of wells.

I initialize using the hydraulic conductivity to be the double that what was used in Figure 26 and a pressure gradient of 0.45 psi/ft. Therefore, the initial bottom pressure is 0.45x814=366 psi corresponding to a head of 257 m.

Because my hypothesis is that the seismic coherence attribute is a good proxy for the hydraulic conductivity, my objective is simply to find two weights or multipliers, one for the initial hydraulic conductivity and another for the initial bottom model pressure. Because I know the conditions that generated the head for my synthetic wells, I want to find the perfect solution to recover the same values. Therefore, for the input head I want to obtain a multiplier of 171/257=0.66, and 0.50 for the hydraulic conductivity (since my initial guess was the double of what was used to generate the modeled wells).

Perturbing my variables with random values, I obtained multiplier for the hydraulic conductivity (0.5071) and for the used input head (0.6649).

Using the multiplier I obtained with the simulated annealing, I obtain a bottom head of 263.12*0.6649=174.94 which gives an error of 3.95 feet or 2.31% when compared with the initial synthetic model. The error between the set pressure points input and modeled (the Euclidean distance from equation 3) is 5.52 feet.

Although I am in control of the synthetic parameters, I could not find the multipliers that would give the perfect solution. A possible explanation is related to the

parameters used for the simulated annealing and for the internal steady-state solution itself: as the cost of this solution is expensive, the number of iterations used in the simulated annealing were not capable of reaching the desirable results. I generated the synthetic wells using the double numbers of iterations that I used for the simulated annealing search. There is always the need to balance the weight of time constraints and desired accuracy.

Nonetheless, the results obtained show that the method is able to generate answers with a reasonable error for my prediction. However, this approach needs to be benchmarked against more established pore pressure procedures such as the ones revised by Dutta (2002).

Chapter 6: Conclusions

I have prototyped a very simple flow model that is built on the hypothesis that seismic attributes, such as coherence, delineate conductive faults. While such a simple flow model cannot replace more accurate (and interpreter intensive) models built using commercial flow simulators, it can be used to statistically correlate water production from a suite of horizontal wells to azimuthally limited fault families.

Such correlations may help earth scientists avoid problematic faults or target those that may enhance production. I envision that applying this procedure can be used to predict faults that can have higher water inflow from a neighboring aquifer.

I believe that the comparison between the results obtained by my method and a standard reservoir simulation studies on a well know field could create valuable insights for earth scientists working on new ventures fields where little or no well information is available.

Nevertheless, the tool developed here might be used for different objectives. As it constitutes a different balancing of well-established seismic attributes, such as coherence, I anticipate that the work described here might give rise to alternative geological interpretations. Faults are not the only geological feature highlighted by edge detecting algorithms applied on seismic amplitude data. Channels, for example, are easily highlighted as well, and can give rise to increased connectivity.

The GSB dataset highlighted several through-going faults previously masked by a complex network of syneresis features.

The results on the Taranaki basin – Kora volcano dataset showed indications that the method discussed in this paper may also be used as a filtering technique. If

smaller and weaker geological features may dominate some areas of the seismic data, calculating the flux may remove some of this possible geological noise.

The best separation between connected and disjoint geological features occurred in the application to the Bemis-Shutts field data volume. Images of local karst features in the carbonates in the Arbuckle group were suppressed while larger through-going karst linking top and bottom used horizons. Hopefully with a more accurate horizon interpretation of the aquifer top and water production data, a correlation between computed flux and connectivity with the underlying aquifer can be obtained.

Careful data preconditioning is of significant importance for this method. Removing the influence of a high hydraulic conductivity feature that is not geologically reasonable, such as the Kora volcano example, is a simple task that might help the interpreter obtain better results. Masking nonconductive features illuminated by coherence such as the Taranaki Basin magma conduit or of low coherence salt diapirs are part of the model building process.

The application of the algorithm on Fairview dataset shows that the hypothesis that hydraulic conductivity is proportional to coherence will not be correct all the times, but it also shows that other attributes could be related to (and used as input for the flux computations) permeability. The well with the higher water production lies close to a geometrical feature with seismic characteristics of a sink hole that can be the responsible for bringing water from deeper layers.

Applying the flux computation to estimate the pressure of a field might be a valuable new tool for oil fields in exploratory phase. However, a faster method needs to be implemented so that the algorithm can be used with these objectives.

References

Abbasi, B., and H. Mahlooji, 2012, Improving response surface methodology by using artificial neural network and simulated annealing: Expert Systems with Applications, **39** (3), 3461-3468.

Bense, V. F., E. H. Van den Berg, and R. T. Van Balen, 2003, Deformation mechanisms and hydraulic properties of fault zones in unconsolidated sediments; the Roer Valley Rift System, The Netherlands: Hydrogeology Journal, **11** (3), 319-332.

Bortolin, T. A., P. A. R. Reginato, M. I. Leao, and V. E. Schneider, 2014, Hidrogeologia e hidroquímica dos aquíferos fraturados associados às rochas vulcânicas ácidas no Município de Carlos Barbosa (RS): Rev. Ambient. Água, **9**(1), 55-67.

Dutta, N. C., 2002, Geopressure prediction using seismic data: Current status and the road ahead: Geophysics, **67**, 2012-2041.

Ertekin, T., J.H. Abou-Kassem, and G.R. King, 2001, Introduction, Basic applied reservoir simulation. SPE Textbook Series-Society of Petroleum Engineers,1-10.

Franseen, E. K., A. P. Byrnes, J.R. Cansler, D. M. Steinhauff, T. R. Carr, and M. K. Dubois, 2003, Geologic controls on variable character of Arbuckle reservoirs in Kansas—An emerging picture: Kansas Geological Survey, Open-file Report 2003-59.

Ferill, D. A., and A. P. Morris, 2008, Fault zone deformation controlled by carbonate mechanical stratigraphy, Balcones fault system, Te3xas: AAPG Bulletin, **92**, 359-380.

Gersztenkorn, A., and K. J. Marfurt, 1999, Eigenstructure-based coherence computations as an aid to 3-D structural and stratigraphic mapping: Geophysics, **64**, 1468-1479.

Guo, Y., K. Zhang, and K. J. Marfurt, 2012, Quantitative correlation of fluid flow to curvature lineaments, SEG Technical Program Expanded Abstracts, 1-5.

Harbaugh, A. W., 2005. Chapter 2 – Derivation of the finite-difference equation, MODFLOW-2005. US Geological Survey, 2.1-2.17.

Heath, R.C., 1983. Basic ground-water hydrology, U.S. Geological Survey Water-Supply Paper 2220, 86p.

Höcker, C., and G. Fehmers, 2002, Fast structural interpretation with structure-oriented filtering: The Leading Edge, **21**, 238-243.

Holubnyak, Y., J. Rush, and W. Watney, 2014, Evaluation with dynamic simulations of the volumetric curvature tool for modelling reservoir compartments and leakage pathways in the Arbuckle saline aquifer: Energy Procedia, **63**, 2014, 3544-3552.

Infante-Paez, L., and K. J. Marfurt, 2017, Seismic expression and geomorphology of igneous bodies. A Taranaki Basin, New Zealand case study: Interpretation, **5**, Ahead of Print:1-72.

Istok, J., 1989. Groundwater Modeling by the Finite Element Method. American Geophysical Union.

Jarvie, D. M., R. J. Hill, T. E. Ruble, and R. M. Pollastro, 2007, Unconventional shale-gas systems: the Mississippian Barnett Shale of North-Central Texas as one model for thermogenic shale-gas assessment: AAPG Bulletin, **91**, 475 - 499.

Kirkpatrick, S., C.D. Gerlatt Jr., and M.P. Vecchi, 1983, Optimization by simulated annealing: Science, **220**, 671–680.

Li, L., H. Zhou, and J. Jaime Gómez-Hernández, 2010, Steady-state saturated groundwater flow modeling with full tensor conductivities using finite differences: Computers & Geosciences, **36**, 1211-1223.

Machado, G., A. Alali, B. Hutchinson, O. Olorunsola, and K. J. Marfurt, 2016, Display and enhancement of volumetric fault images: Interpretation, **4**, SB51-SB61.

Machado, G., G. Hickman, M. Gogri, M. J. Pranter, K. J. Marfurt, and Z. A. Reza, 2017, Multidisciplinary characterization of geomechanical properties and flow behavior of the coupled Arbuckle-Basement System, Payne County, Northern Oklahoma: AAPG 2017 Annual Convention and Exhibition (ACE) Expanded Abstracts.

Marfurt, K. J., 2006, Robust estimates of 3D reflector dip and azimuth: Geophysics, **71**, P29-P40.

Morley, C. K., A. Maczak, T. Rungprom, J. Ghosh, J.A. Cartwright, C. Bertonib, N. Panpichityota, 2017, New style of honeycomb structures revealed on 3D seismic data indicate widespread diagenesis offshore Great South Basin, New Zealand: Marine and Petroleum Geology, **86**, 140-154.

Nissen, S. E., T. R. Carr, K. J. Marfurt, and E. C. Sullivan, 2009, Using 3D seismic volumetric curvature attributes to identify fracture trends in a depleted Mississippian carbonate reservoir: Implications for assessing candidates for CO2 sequestration: AAPG Studies in Geology, **59**, 297–319.

Pedersen, S. I., T. Randen, L. Sonneland, and O. Steen, 2002, Automatic 3D Fault interpretation by artificial ants: 64th Meeting, EAGE Expanded Abstracts, G037.

Perez-Altamar, R., "Brittleness estimation from seismic measurements in unconventional reservoirs: Application to the Barnett Shale" PhD dissertation, The University of Oklahoma, 2013.

Senoglu, D., "Reservoir characterization and modeling of the Desmoinesian Series Granite Wash, Buffalo Wallow Field, Anadarko Basin, Texas" Master's thesis, The University of Oklahoma, 2017.

Sverdrup, E., J. Helgesen, J. Vold, 2003, Sealing properties of faults and their influence on water-alternating-gas injection efficiency in the Snorre field, northern North Sea: AAPG Bulletin, **87**, 1437 - 1458.

Sinha, S., D. Devegowda, and B. Deka, 2017, Quantification of recovery factors in downspaced wells: application to the Eagle Ford Shale: Society of Petroleum Engineers. doi:10.2118/185748-MS.

Stuevold, L. M., R. B. Faerseth, L. Arnesen, J. Cartwright, and N. Möller, 2003, Polygonal faults in the Ormen Lange field, Møre Basin, offshore mid Norway: Geological Society London Special Publications, **216**, no. 1, 263–281.

Trumbo, D.B., "A production calibrated reservoir characterization of the Mississippi Lime in a mature field utilizing reprocessed legacy 3D seismic data, Kay County, Oklahoma" Master's thesis, the University of Oklahoma, 2014.

Uruski, C. I., 2010, New Zealand's deepwater frontier: Marine and Petroleum Geology, **27**, no. 9, 2005-2026.

Weber, K.J., 1987, Hydrocarbon distribution patterns in Nigerian growth fault structures controlled by structural style and stratigraphy: Journal of Petroleum Science and Engineering, **1**, 91-104.

Appendix A: Finite-difference equations

This appendix is greatly inspired by the work done by Harbaugh (2005). The potential fluid head (h) in three-dimensions is given by:

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + W = S\frac{\partial h}{\partial t}, \qquad \text{(A1)}$$

where K_x, K_y, and K_z are the conductivities of the media in the x, y, and z coordinate directions, W is a volumetric flux per unit volume (e.g. a pumping or injecting well), S is the specific storage of the media, and t is time. Because my objective is to compute only the steady state solution of this equation, the time derivative of the head is zero, i.e.:

$$\frac{\partial}{\partial x}\left(K_x \frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y \frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z \frac{\partial h}{\partial z}\right) + W = 0. \qquad \text{(A2)}$$

For the finite-difference formulation, I consider that the sum of all flows into and out of a voxel must be equal to the rate of change in storage within the voxel. Because I assume a single fluid with constant density, the balance of the voxel is given by:

$$\sum_{n=1}^{N} Q_n = S\frac{\Delta h}{\Delta t}\Delta V = 0, \qquad \text{(A3)}$$

where

Q_n is the flow rate into the cell coming from the n source with units L^3T^{-1},

S is specific storage using the finite-difference notation with units L⁻¹,

ΔV is the volume of the voxel with units L³, and

Δh is the change in the potential head over a time interval Δt.

Note that because I assume steady state, the time derivative in equation A3 is zero. Now, I need to compute the sum of the flow coming from the neighboring six voxels (Figure A1) where the size of the cells are Δx = cdp increment, Δy = line increment, and $\Delta z = \frac{1}{2} v\Delta t$ in directions $\boldsymbol{i}, \boldsymbol{j}$, and $\boldsymbol{k}$ respectively, where v is the velocity, and Δt is the time sample increment. Figure A2 shows the stencil used for the computations as a 2D plane time slice for easier visualization. For this 2D visualization, I do not draw the neighbors in the time (or depth) direction.
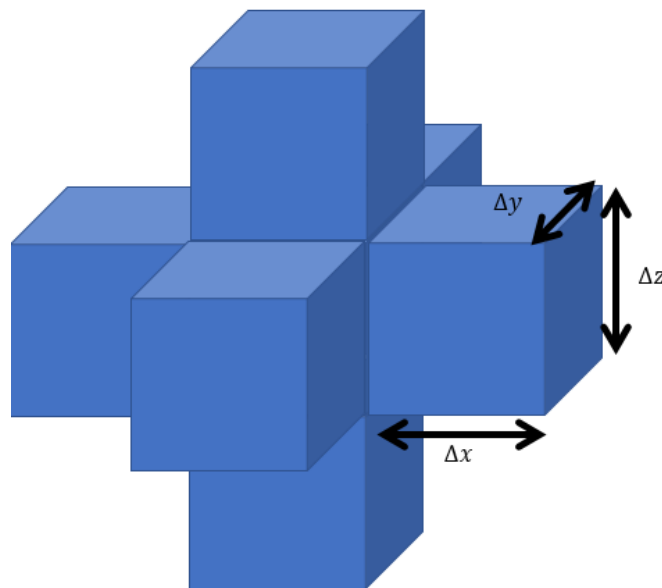

Figure A1: Voxel and its neighbors. I compute the flow for the central voxel (not visible on this image) and the flow depends on the six neighbors.
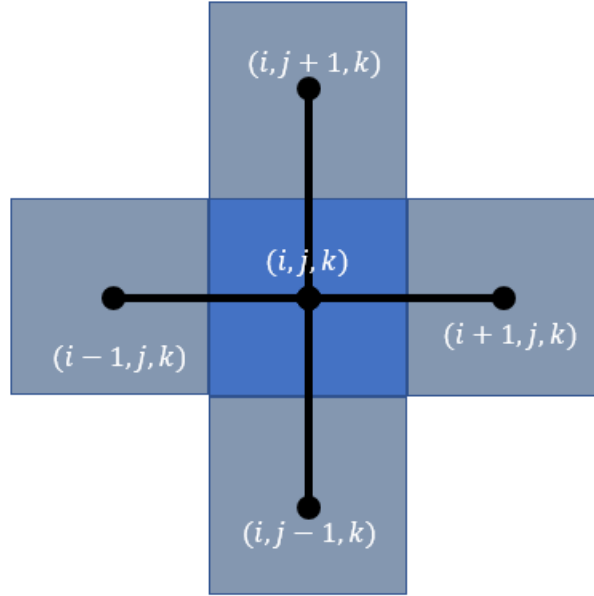
Figure A2: Voxels and neighbors with stencil represented as a 2D time slice plane. Remember that there are to other neighbors, (i,j,k+1) and (i,j,k-1).

Using this naming conventions, the flow coming from the voxel to the right $(i + 1, j, k)$ to the center voxel (i, j, k) (Figure A3) is given by:

$$Q_{i+0.5,j,k} = KX_{i+0.5,j,k}\Delta z * \Delta y \frac{h_{i+1,j,k}-h_{i,j,k}}{\Delta x} \quad \text{(A5)}$$

where

$Q_{i+0.5,j,k}$ is the volumetric flow rate through the face between voxels i,j,k and $i+1,j,k$ (L^3T^{-1});

$KX_{i+0.5,j,k}$ is the hydraulic conductivity along the row between nodes voxels i,j,k and $i+1,j,k$ (LT^{-1});

$\Delta z * \Delta y$ is the area of the voxel normal to the $\mathbf{k}$ direction and Δx is the distance between nodes.

Figure A3: Representation the location of the voxels used for computation of the flow coming from (*i+1,j,k*) to (*i,j,k*). Δx is the distance between the center of two voxels (nodes).

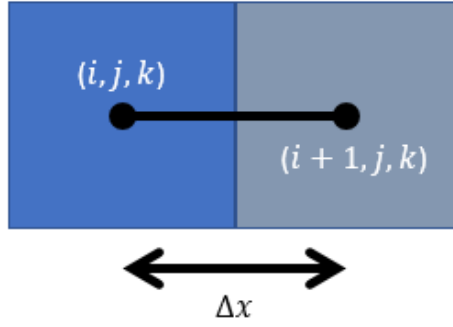For my formulation, the conductivity at ½ grids points is the harmonic mean of the conductivity values of the voxels i,j,k and $i+1,j,k$.

All other voxels will have an analogous flow description. To simplify the notation, the conductivity and the grid dimensions can be grouped together in the conductance:

$$CX_{i+0.5,j,k} = KX_{i+0.5,j,k} \frac{\Delta z * \Delta y}{\Delta x} \quad (A6)$$

where $CX_{i+0.5,j,k}$ is the conductance in the ***i*** direction between nodes i,j,k and $i+1,j,k$ (L^2T^{-1});

Equation A3 states that the sum of all flows coming from neighboring's voxels should be zero. Using the appropriate naming conventions, I have:

$$CX_{i-0.5,j,k}\left(h_{i,-1j,k} - h_{jt,jcdp,jline}\right) + CX_{i+0.5,j,k}\left(h_{i+1,j,k} - h_{i,j,k}\right)$$

$$+$$

54

$$CY_{i,j-0.5,k}\left(h_{i,j-1,k} - h_{i,j,k}\right) + CY_{i,j+0.5,k}\left(h_{i,j+1,k} - h_{i,j,k}\right)$$

$$+$$

$$CZ_{i,j,k-0.5}\left(h_{i,j,k-1} - h_{i,j,k}\right) + CZ_{i,j,k+0.5}\left(h_{i,j,k+1} - h_{i,j,k}\right)$$

$$= 0 \qquad \text{(A7)}$$

where $CX, CY,$ and CZ are the conductance values on the *i,j*, and *k* faces respectively.

Their value is dependent on the index-voxel position.

With this formulation, I can rearrange the terms to solve the system iteratively.

Weighted Jacobi method

To solve equation A7 iteratively, I adapt the procedure used by Li et al. (2010) to implement the weighted Jacobi method. First, I rearrange the equation to isolate the head potential for the central voxel:

$$h_{i,j,k} = [CX_{i-0.5,j,k}\left(h_{i-1,j,k}\right) + CX_{i+0.5,j,k}\left(h_{i+1,j,k}\right)$$

$$+$$

$$CY_{i,j-0.5,k}\left(h_{i,j-1,k}\right) + CY_{i,j+0.5,k}\left(h_{i,j+1,k}\right)$$

$$+$$

$$CZ_{i,j,k-0.5}\left(h_{i,j,k-1}\right) + CZ_{i,j,k+0.5}\left(h_{i,j,k+1}\right)]/$$

$$(-CX_{i-0.5,j,k} - CX_{i+0.5,j,k} - CY_{i,j-0.5,k} - CY_{i,j+0.5,k} - CZ_{i,j,k-0.5} + CZ_{i,j,k+0.5}) \qquad \text{(A8)}$$

Equation A8 shows, as expected, that the current head potential of the voxel i, j, k depends on the values of the head potential of its neighbors.

Equation A8 allows me to use an iterative Jacobi method. Although the Gauss-Seidel and the successive over-relaxation methods converge with fewer iterations, their parallelization is more complicated. Details on the data distribution for parallelization are described in Appendix B.

The Jacobi method relies only on the previously computed heads to compute the head values for the current iteration. Figure A4 shows a simplified flow of the implementation for one iteration. After every voxel has its head value updated, the next iteration starts and the head value is updated again for every single voxel. Iterations continue till convergence or until the maximum number of iterations is reached.
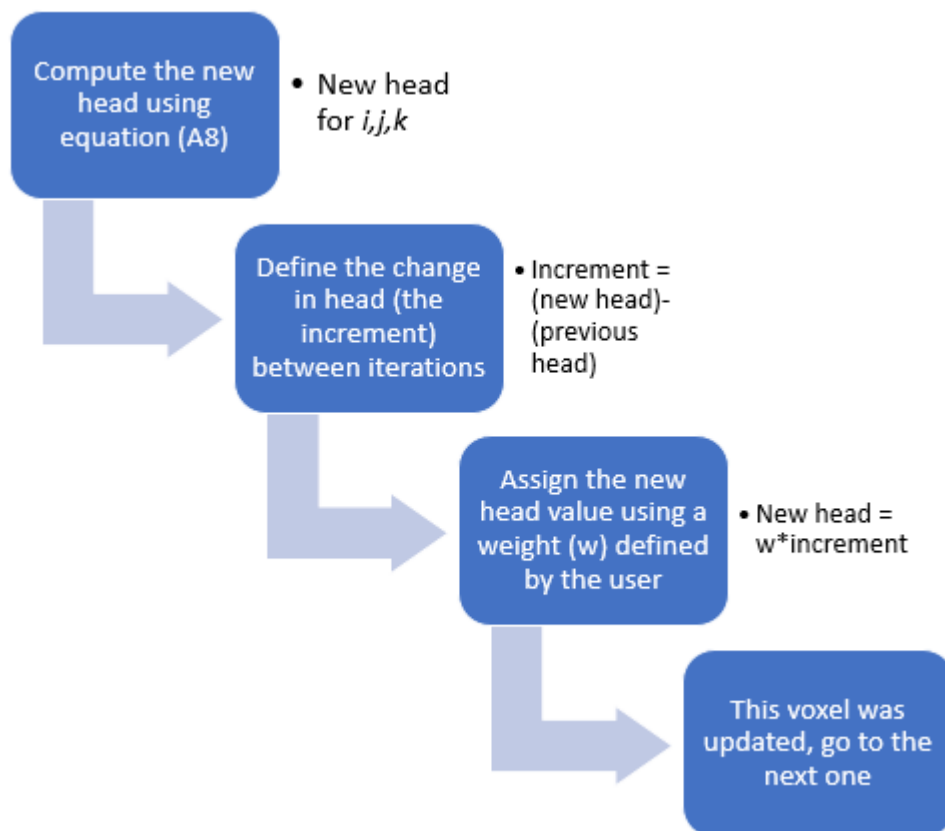

Figure A4: Simplified flow showing how the weighted Jacobi method was implemented.

Appendix B: Parallelization

The head solution is computed using multiple processors simultaneously. As described in Appendix A, the head value for any voxel depends only on the head values (and conductances) of the six adjacent voxels. For this reason, and because of the iterative method implemented, splitting the data across multiple process provides an efficient solution.

The cartoon in Figure B1 shows how to split the data across multiple processors. I use different colors and different numberings to represent the processors and how the technique is implemented. Each processor is responsible for updating a thick vertical slice of the data. Data necessary for adjacent processors are transferred after the weighted Jacobi iteration is completed at each processor. Therefore, for each process, the current iteration uses the computed head value from the previous iteration to update the voxel value. The cartoon in Figure B1 shows a single time (or depth) slice of the data for easier visualization, however the data extend in all three dimensions.
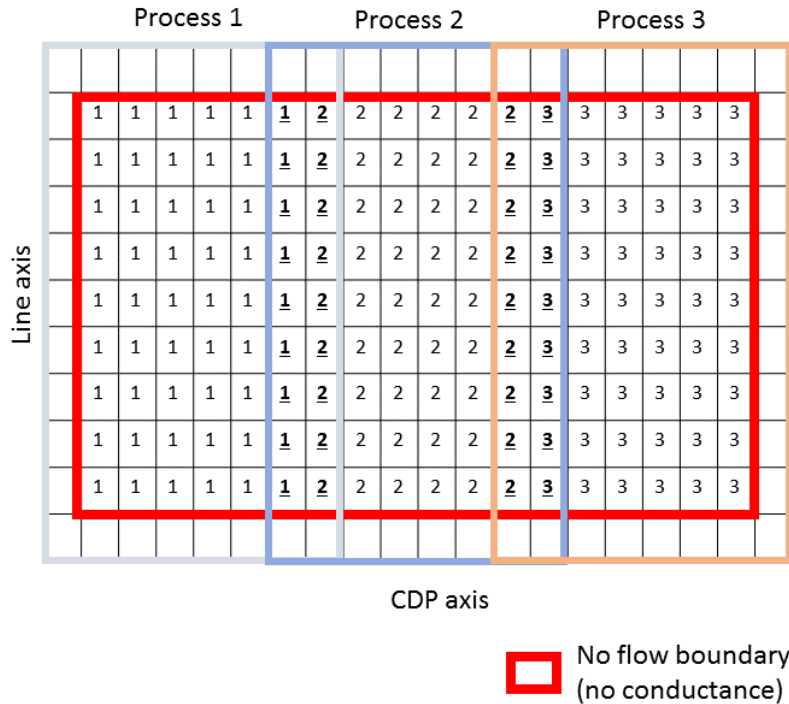
Figure B1: Representation of data splitting for different processors. Note that the no flow boundary extends to the edges of the complete dataset while the walls between processors will not have such characteristics and will be populated with the real computed conductance, therefore requiring communication between processes. Numbers inside voxels (1, 2, and 3) represent the process responsible for updating such voxel. Boldface and underlined numbers show voxels shared between processes. The boldface and underlined "1" voxels are computed by process 1 and sent to process 2, the boldface and underlined "2" voxels are computed by process 2 and sent to process 1 for the next iteration and so forth.

Using multiple processors simultaneously greatly improves algorithm performance. The Great South Basin dataset described in Chapter 3: Results and discussion has dimensions of $Nx = 1001$, $Ny = 651$, and $Nz = 251$ samples, totalizing 163,564,401 voxels. Computing 5000 iterations for this volume using a single processor required an elapsed wall clock time of 30.9 hours. Using the same parameters and 10 processors, the wall clock time dropped to 4.5 hours, a reduction of approximately 85%. Both computations were performed on processors with 3.47 GHz.

Appendix C: Simulated Annealing

Kirkpatrick et al. (1983) noticed the link between statistical mechanics and multivariate or combinatorial optimization and proposed that as a basis for optimization techniques. The simulated annealing (SA) method was named because of the analogy between the way in which a metal cools and freezes into a minimum energy crystalline structure (the annealing process) and the search for a minimum in a more general system (Abbasi and Mahlooji, 2012).

Abbasi and Mahlooji (2012) state that principal advantage of SA's is the ability to avoid getting caught and stuck at local minima. The algorithm randomly searches for a new solution for the problem in question. This new solution may have a smaller or larger error. If the solution has a smaller error than the current one, the algorithm changes to that position. In contrast, if the solution has as larger error, the algorithm decides based on a probability if it changes to the new solution. The probability of change is given by

$$p = e^{-\Delta/T} \qquad (C1)$$

Where Δ is the increase in the objective function and T is the temperature analogue control parameter.

The basic SA structure, as presented by Abbasi and Mahlooji (2012), is:

S = the current solution,

$S*$ = the best solution,

S_n = new solution,

$f(S)$ = the value of objective function at solution S,

n = repetition counter,

T_0 = initial temperature,

I = number of repetitions allowed at each temperature level,

p = probability of accepting Sn when it is not better than S,

The following pseudo code summarizes the simulated annealing process: