Bootstrapping reliable noise measure in time-gated nuclear magnetic resonance data

Trevor P. Irons*
University of Utah
Dept. Civil & Environmental Engineering
Energy & Geoscience Institute
423 Wakara Way, Suite 300
Salt Lake City, UT 84108 USA
tirons@egi.utah.edu

Brian J.O.L. McPherson
University of Utah
Dept. Civil & Environmental Engineering
Energy & Geoscience Institute
423 Wakara Way, Suite 300
Salt Lake City, UT 84108 USA
b.j.mcphserson@utah.edu

M. Andy Kass
Aarhus University
Dept. of Geoscience
C.F. Møllers Alle 4
DK-8000 Aarhus C Denmark
andy.kass@geo.au.dk

**presenting author asterisked*

SUMMARY

Time gating is a commonly used approach in the pre-processing of nuclear magnetic resonance (NMR) data before Laplace inversion. Gating suppresses spurious signals that can degrade recovered decay time distributions and therefore often stabilizes inversion. However, care must be taken in applying this technique to real world data where both non-Gaussian and correlated noise decrease the efficacy of noise reduction through stacking. If not properly accounted for, unreliable noise estimates introduce inversion artefacts. Fortunately, noise realization proxies obtained through data phasing can be used to bootstrap reliable confidence intervals for the windowed data. Benefits of the approach are demonstrated through inversion of synthetics as well as borehole data from a deep carbon capture and sequestration application. We also introduce an open source cross platform data processing utility with these capabilities which interested persons can use to explore the impacts of various processing workflows.

Key words: hydrogeophysics, data-processing, inverse theory, statistics.

INTRODUCTION

Nuclear magnetic resonance (NMR) methods provide a unique means by which to directly quantify important hydrological properties including porosity, permeability, and phase saturation. Since the 1960's borehole NMR has been relied upon in the oil and gas industry where it has proven invaluable at fluid-typing and porosity and permeability quantification (e.g. Dunn et al. 2002). More recently, NMR methods have established themselves as an important tool in environmental studies (Behroozmand et al. 2015). One reason is that active nuclear sources (as required by neutron and γ-γ tools, but not NMR) are prohibited from municipal and agricultural aquifers by some jurisdictions. Additionally, miniaturized borehole probes can be used in harsh conditions using hand augured slim holes inaccessible to drill rigs. Finally, earth's field surface NMR measurements do not require any drilling and can probe depths of up to 100m non-invasively.

In geophysical applications NMR data are primarily collected, processed, and inverted in the time domain; although notable exceptions do exist (Irons & Li 2014; Hein et al. 2017). For borehole data, acquisition typically involves CPMG echo trains where each echo peak is retained to form the exponentially decaying signal of interest. The CPMG pulses are utilized to refocus the spins which quickly dephase due to magnetic field inhomogeneity imposed by the high field permanent magnet used for polarisation (B_0). For surface NMR, the earth's field is utilized for B0 which is much more uniform. Although CPMG pulses can be collected with surface NMR instruments (Grunewald & Walsh 2013), in most cases the physical limitations of such measurements results in free induction decay measurements having more utility. These data oscillate at the Larmor frequency which is generally removed using a form of quadrature detection scheme. In both cases, after processing, the ensuing induced NMR data take the form of a complex-valued time series (V_N) decay envelope

$$\boldsymbol{V_N = A_0 e^{j\zeta - \frac{t}{T_2}} + \epsilon}. \tag{1}$$

In Equation 1 A_0 represents the initial amplitude of the data, ζ is a phase term, and T_2 is the transverse plane decay time constant (in sNMR $T_2^* \rightarrow T_2$ for the purposes of this discussion). As a result of the similarity, these two dataforms are often processed using the same techniques and approaches; although we will show the noise characteristics of the two (ϵ) are quite different. The phase term can be due any number of factors, non-exclusively including: instrument loading effects, coil geometries, transmitter phase, and retardation of the electromagnetic fields due to electrical conductivity. In the case of sNMR the phase dependence on electrical conductivity can be used to enhance inversion resolution (Braun et al. 2005; Irons & Li 2014). Even in the case of complex inversion, the presented approach can be followed to provide noise characterisation. Most commonly, the signal is 'phased' in order to rotate all signal to either the imaginary or real channel to avoid introducing bias (e.g. Legchenko & Valla 1998).

$$\boldsymbol{\tilde{\epsilon} = \mathfrak{I}\left(V_n\, e^{-j\zeta}\right)} \tag{2}$$
$$\boldsymbol{V_N = \mathfrak{R}(\, V_N e^{-j\zeta})}. \tag{3}$$

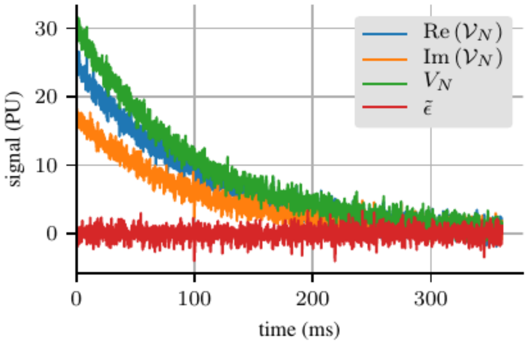


Figure 1: Example of quadrature envelope representation of NMR signal contaminated with noise $N(0,\sigma)$. The real and imaginary components both contain signal due to a non-zero phase term (ζ). The signal can be rotated (phased) so that all signal is contained in the real channel, leaving a noise residual $\tilde{\epsilon}$ which is a realization of the noise contaminating the signal.

Using this convention, Equation 3 contains the real-valued phased signal channel that will be inverted while Equation 2 provides a realization ($\tilde{\epsilon}$) of the noise contaminating the signal. Often the noise residual is dismissed following phasing; however, we will exploit this channel to provide reliable noise estimates. If $\epsilon \in N(0,\hat{\sigma})$ then $\mathrm{std}(\tilde{\epsilon}) \rightarrow \tilde{\sigma} \approx \hat{\sigma}$. Figure 1 presents an illustration of the quadrature data and phasing process with additive Gaussian noise of one porosity unit (PU = 1%).

In order to speed inversion performance, the entire signal shown in Figure 1 is rarely used. Rather, the signal is resampled to a lower sampling frequency. To retain the full nature of the early time signal, gate integrating the signal has become common (Beehroozmand et al., 2012). However, the variable window width results in non-uniform noise conditions across the signal that must be accounted for if accurate inversions are to be realized.

In the remaining sections, we demonstrate the use and utility of using the phasing residual to estimate noise levels through an inversion algorithm which relies on the discrepancy principle to determine the level of regularisation to apply. We consider Schlumberger CMR borehole data from a CO_2-EOR project in northern Texas shown as well as surface NMR data from Laramie Wyoming collected using a Vista Clara GMR. The utility of the noise residual approach across seemingly disparate datasets is interesting, however care must be taken to understand the nature of the underlying noise and what assumptions can be safely made.

INVERSION

Of primary interest in most NMR studies is determining both the total water content which is related to A_0 as well as quantifying the decay time which can often be related to the surface area to volume ratio of pores within the media. Generally speaking, NMR data in porous media will be multiexponential and a common approach is to invert for a distribution of partial water concentrations (f_p) across T_2 times. A simple (Tikhonov L^2 norm) formulation for the inverse problem which is illustrative for our purposes follows

$$(\mathbf{K}\mathbf{K}^{T} + \lambda \mathbf{W}_m^T \mathbf{W}_m)\mathbf{f}_p = \mathbf{K}^T \mathbf{W}_d^T \mathbf{W}_d \mathbf{v}^{obs}. \tag{4}$$

$\mathbf{K}$ is a linearised forward operator, $\mathbf{W}_m$ provides smoothness constraints, and $\mathbf{v}_{obs}$ represents the observed field data. The matrix $\mathbf{W}_d$ is a data weighting matrix whose diagonal entries are the inverse of the standard deviations of $\mathbf{v}_{obs}$. Because water content $\mathbf{f}_p$ cannot be negative, it is necessary to further constrain the solution to non-negative solutions (Li & Oldenburg 2000; Calvetti et al. 2004). For simplicity, these constraint terms are omitted from our discussion, inversion results implement a logarithmic-barrier constrained solution following Irons & Li 2014.

Solving Equation 4 can be accomplished with any number of linear system solvers in a relatively straightforward manner. However, determining the appropriate degree of regularisation (λ) to apply is a non-trivial problem and lies at the heart why inversion can be challenging. If λ is too small, the data will be overfit whereas too large of values of λ will overly reflect the prior knowledge that is imposed. Common approaches involve use of the discrepancy principle, L-curve (Hansen 1992), or generalised cross validation. For any of these criteria to be applied, however, it is vital to either have uniform noise across the data, or to have a reliable metric for determining noise levels. In this paper, we use the discrepancy principle which states that the measure of data misfit $\mathbf{W}_d(\mathbf{K}\mathbf{f}_p - \mathbf{v}_{obs}) \approx 1$.

TIME GATING

Time gating is an effective means by which to accelerate and stabilize the inversion of NMR data (Behroozmand et al. 2012). In terms of application, time-gating represents the application of a low pass filter and subsequent decimation over the time signal. A simple moving average filter, or other regular resampling scheme, should be avoided as picking an appropriate length (L) introduces undesirable trade-offs. Instead an adaptive "time gate" filter is often applied where the averaging window varies as a function of time

$$V_G[i] = \sum_{j=k}^{k+L_i} V_N[j] / L_i. \tag{5}$$

In Equation (5) V_G represents the gated time series, with L_i representing the number of data in the summation window. Commonly time gating is specified in gates per decade, a convention we adopt as well. There are several reasons why time gating is effective for NMR data. First, due to the exponential nature of the imaging kernel, the NMR sensitivity to discrimination of slowly decaying signals is very low--commonly $\mathbf{f}$ is solved for in log-space. Second, late time data suffer from low S/N. Finally, since late time signals can be very low and small water contents with long decay times give minimal measure in an L^2 sense, late time noise can easily bias solutions towards long T_2 solutions. For these reasons, applying an increasingly higher order low pass filter as a function of time is an effective pre-inversion data processing step.

Inverting V_G follows the same approach as Equation (4) with the substitution of the time gated data with the dense variant as $\mathbf{V}_{obs}$. Aside from this, the only other consideration comes in constructing the data weighting matrix $\mathbf{W}_d$. An obvious first approach would be to assume that $\epsilon \in N(0, \tilde{\sigma})$. If this assumption holds, variance of noise should decrease like $\sqrt{N}$, which for a gate i starting at k of width L_i follows

$$\widetilde{\sigma_\iota} = \sqrt{\left(\sum_{j=k}^{k+L_i} \tilde{\sigma}^2[j]\right) + \left(\frac{1}{L_i}\right)^2}. \qquad (6)$$

When $\epsilon \in N(0, \hat{\sigma})$ the theoretical reduction in noise as a function of gate width holds and confidence intervals using $\widetilde{\sigma_\iota}$ from Equation (6) are reliable (Figure 2). In this case Equation (6) provides a reliable estimate of $\hat{\sigma}$ with expected levels of misfit. However, we have found that this relationship does not always provide realistic results in practice (

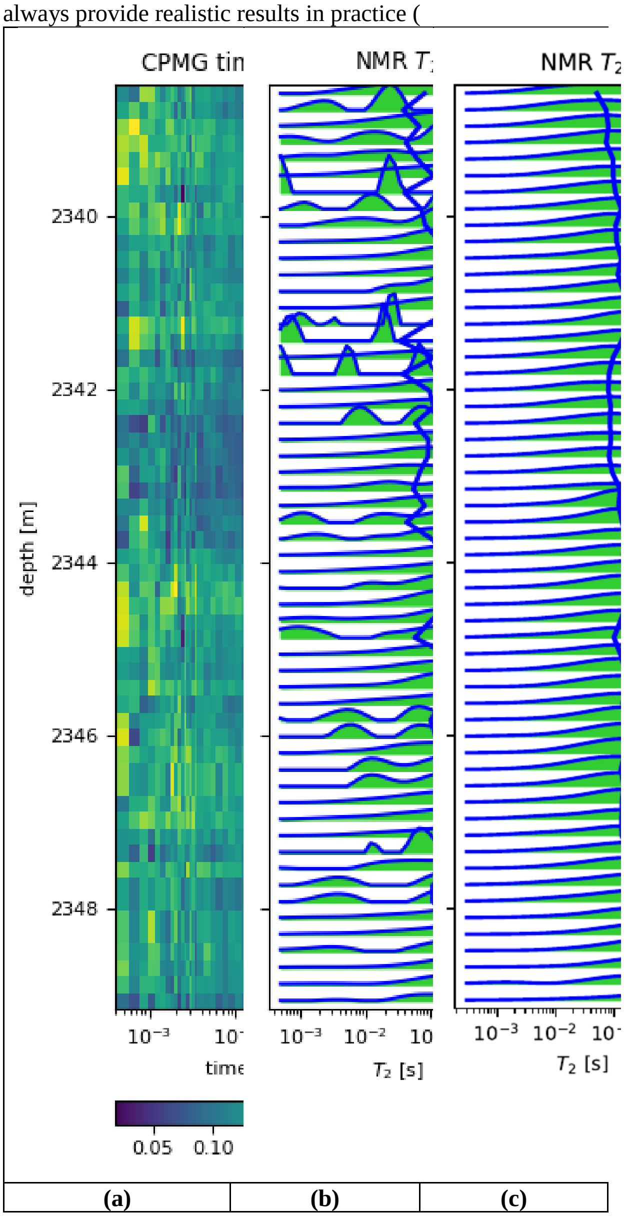


Figure 4(b)). Borehole field data shown in this study were logged using the Schlumberger CMR borehole tool and were collected at well 1310-A in the Farnsworth Unit (FWU) an active CO_2-EOR site (Balch and McPherson 2016). In these results, Equation (6) was applied independently to each CPMG echo timeseries. The resulting inversions are prone to both under and overfitting the data and T_2 distributions are unrealistic at many depths.

Statistical properties of the noise residual offer insight into why inversions using the theoretical gate noise values can be sub-optimal. In Figure 3 (a) probability distributions of $\tilde{\epsilon}$ are shown for the entire CMR dataset are shown alongside the distribution of a single record. From this, it can be seen that the noise distribution of a single record may deviate significantly from a Gaussian distribution. It is postulated this this is the reason for unstable inversion results using Equation (6) as a noise estimate.

BOOTSTRAPPING GATE NOISE

As an alternative to assuming a Gaussian distribution, we propose using a flexible bootstrapping approach towards estimating gate noise. The residual from the phasing step ($\tilde{\epsilon}$) provides a valuable means by which to characterise noise levels. This residual is not itself a direct measure of noise, but rather a realization of the noise distribution. The algorithm proceeds as follows: the noise residual ($\tilde{\epsilon}$) is bootstrap resampled using variable window lengths and starting points, note that the order is not changed. That is to say that data concurrency is maintained such that correlation effects are preserved (Algorithm 1).

Algorithm 1 bootstrap estimate of NMR gate noise

Require: $\tilde{\epsilon}$, the observed noise realization
Require: `nboot`, the number of bootstrap window calculations to perform
Require: L_i, array of window lengths
for all L in L **do**

for `i` in 0 **to** `nboot` **do**
j = random int $\in$ (0, len($\tilde{\epsilon}$))
if j+L < len($\tilde{\epsilon}$) **then**
`eps[i]` = avg($\tilde{\epsilon}$[j:j+L])
else
`eps[i]` = (avg($\tilde{\epsilon}$[j:len($\tilde{\epsilon}$)-1]) + avg($\tilde{\epsilon}$[0:j+L-len(ϵ)])) /2
end if
end for
$\tilde{\sigma}$[L] = std(`eps`)
end for

Application of the bootstrapping algorithm is relatively straightforward. As it is a stochastic process the number of boot iterations has an impact on the results. Increasing `nboot` results in smoother noise graphs. In order to avoid the need to run very large values of `nboot`, a smoothing spline is applied to the resulting bootstrapped gate noise estimates (Figure 3(b)). Application of the bootstrapped noise estimate within inversion results is shown in Figure 4(c). The bootstrapped noise estimates provide more stable inversions throughout the log.

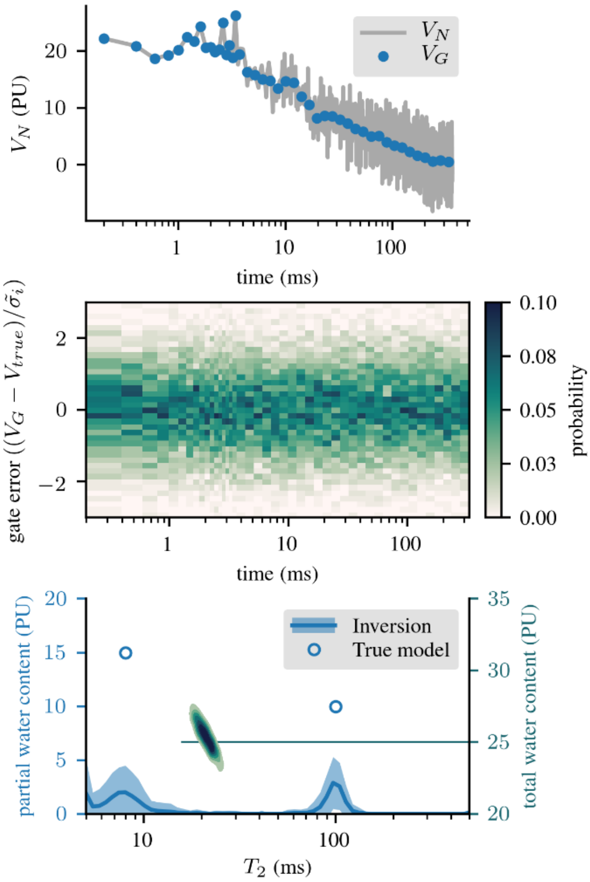
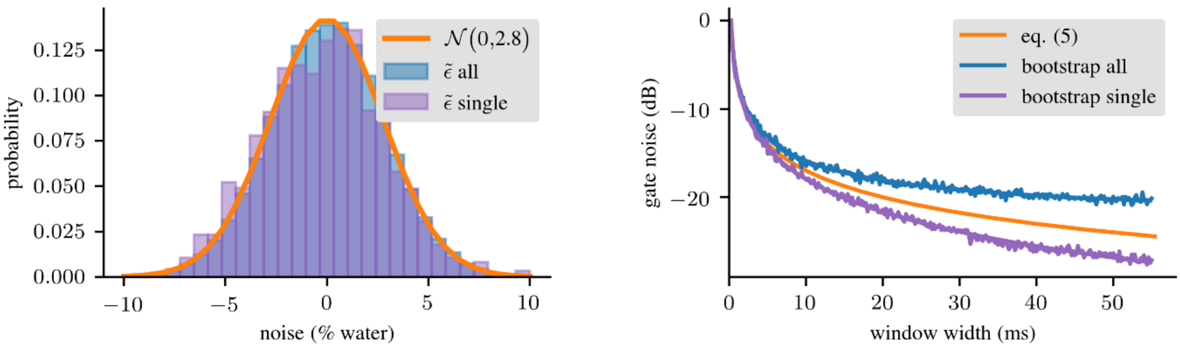


Figure 2 – In the top panel, a synthetic decay is plotted along with 5 PU Gaussian additive noise with an echo spacing of .2 ms, gated data at 14 gates per decade are shown in blue. Probability distribution functions are shown for 200 realizations of Gaussian noise added to exponential signal are shown in the middle panel. The gated values minus the true values map well to the $\tilde{\sigma}$ error metric, from a Frequentist perspective 95\% of the values should fall within two standard deviations of the true value. Inversions (bottom panel) represent smoothed but unbiased estimates of the true values.



(a)	**(b)**

Figure 3- Probability distribution functions for the aggregated CMR log (a) and a single echo train (b). While the noise floor is approximately uniform for the single record compared to the total population, the assumption of Gaussianity is less appropriate.

CONCLUSIONS

Noise level estimation has a large impact on the inversion of NMR data. The residual of data phasing can be used as a noise realization proxy that can be used to estimate noise levels. When gate integration is used, noise levels may not reduce at the theoretical rate due to data correlation. Effective gate noise estimates can be obtained through bootstrapping of the noise residual. While this abstract focused on borehole data, the approach is equally applicable to surface NMR.

ACKNOWLEDGMENTS

Funding for this project is provided by the U.S. Department of Energy's (DOE) National Energy Technology Laboratory (NETL through the Southwest Regional Partnership on Carbon Sequestration (SWP) under Award No. DE-FC26- 05NT42591. Additional support has been provided by site operator Chaparral Energy, L.L.C. and Schlumberger Carbon Services.

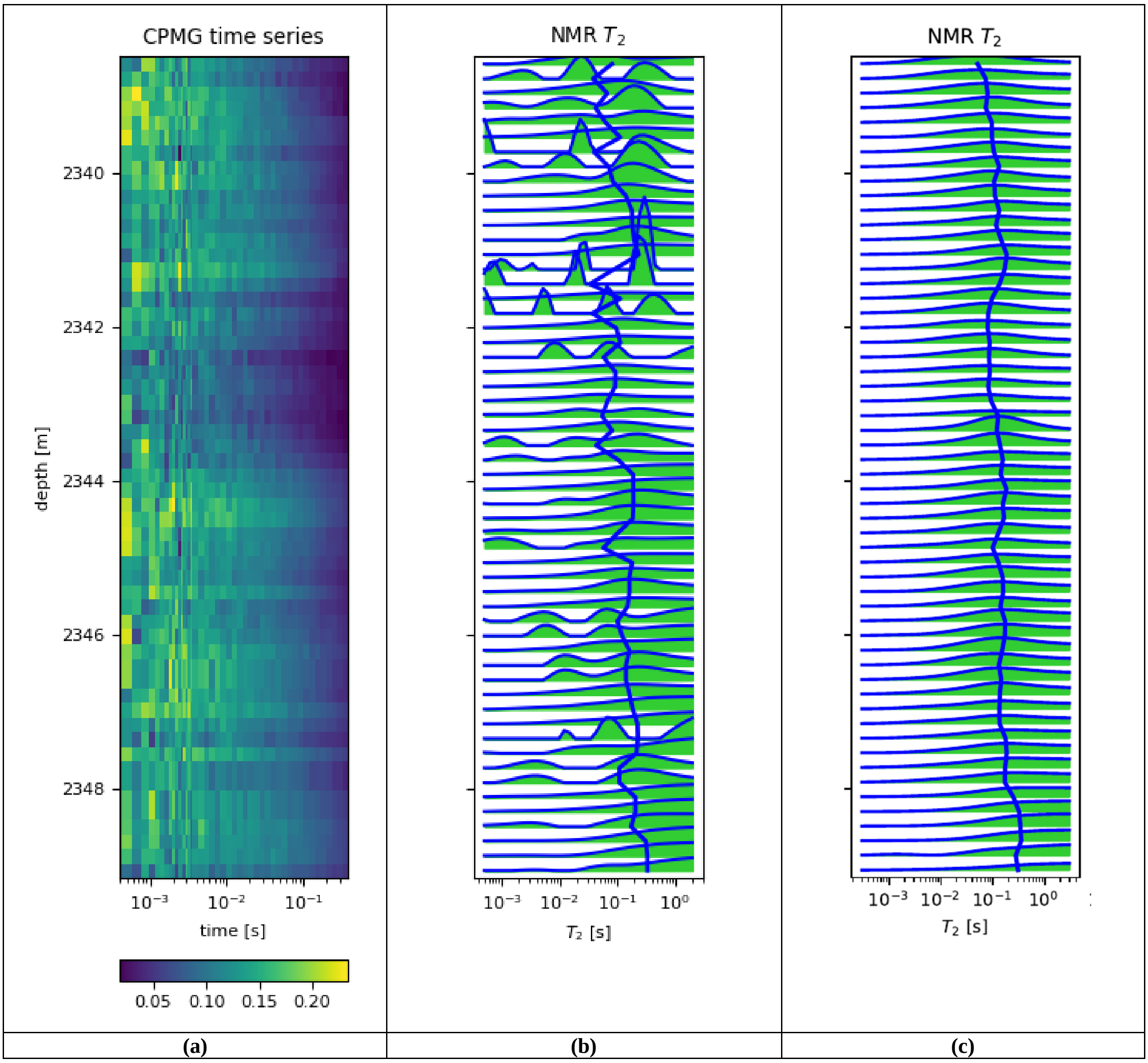


Figure 4 – A Schlumberger CMR log recorded in well 1310-A at the Farnsworth Unit, Texas, an active CO_2-EOR site. The data are time gated at 14 gates per decade (a). These data were inverted for T_2 distributions. Inversions using the theoretical gate noise derived from the phase residual (eq. 6) and are prone to both under-regularised regions (b). Bootstapped error estimates provide more stable inversions throughout the log (c).

REFERENCES

Balch, R. & McPherson, B., 2016. Integrating enhanced oil recovery and carbon capture and storage projects: A case study at Farnsworth field, Texas, Society of Petroleum Engineers annual Meeting.

Behroozmand, A. A., Auken, E., Fiandaca, G., Christiansen, A. V., & Christensen, N. B., 2012. Efficient full decay inversion of mrs data with a stretched-exponential approximation of the distribution, Geophysical Journal International, 190(2), 900–912.

Behroozmand, A. A., Keating, K., & Auken, E., 2015. A review of the principles and applications of the nmr technique for near-surface characterization, Surveys in Geophysics, 36(1), 27–85.

Braun, M., Hertrich, M., & Yaramanci, U., 2005. Study on complex inversion of magnetic resonance sounding signals, Near surface geophysics, 3(3), 155–163.

Calvetti, D., Lewis, B., L.Reichel, & Sgallari, F., 2004. Tikhonov regularization with nonnegatvity contraint, electronic transactions on numerical analysis, 18, 153–173.

Dunn, K., Bergman, D., & LaTorraca, G., 2002. Nuclear Magnetic Resonance: Petrophysical and Logging Applications, Handbook of Geophysical Exploration: Seismic Exploration, Elsevier Science.

Grunewald, E. & Walsh, D., 2013. Multiecho scheme advances surface nmr for aquifer characterization, Geophysical Research Letters, 40(24), 6346–6350.

Hansen, P. C., 1992. Analysis of discrete ill-posed problems by means of the L-curve, SIAM Review, 34, 561–580.

Hein, A., Larsen, J. J., & Parsekian, A. D., 2017. Symmetry based frequency domain processing to remove harmonic noise from surface nuclear magnetic resonance measurements, Geophysical Journal International, 208(2), 724.

Irons, T. P. & Li, Y., 2014. Pulse and Fourier transform surface nuclear magnetic resonance: comprehensive modelling and inversion incorporating complex data and static dephasing dynamics, Geophysical Journal International, 199(3), 1372–1394.

Kenyon, W., Day, P., Straley, C., & Willemsen, J., 1988. A three-part study of nmr longitudinal relaxation properties of water-saturated sandstones, SPE Formation Evaluation, 3(3), 622–636.

Legchenko, A. & Valla, P., 1998. Processing of surface proton magnetic resonance signals using non-linear fitting, Journal of Applied Geophysics, 39(2), 77–83.

Li, Y. & Oldenburg, D., 2000. 3-d inversion of induced polarisation data, Geophysics, 65(6), 1931–1945.