Near Surface Geophysics, 2016, **14**, xxx-xxx doi:10.3997/1873-0604.2016004

Enhancing SNMR model resolution by selecting an optimum combination of pulse moments, stacking, and gating

E. Dalgaard[1,3*], M. Müller-Petke[2] and E. Auken[1]

[1] *HydroGeophysics Group, Department of Geoscience, Aarhus University, Denmark*
[2] *Leibniz Institute for Applied Geophysics, Hannover, Germany*
[3] *Qeye Labs, Copenhagen, Denmark, Formerly at 1*

Received September 2014, revision accepted August 2015

ABSTRACT

For the surface nuclear magnetic resonance sounding method, we investigate the tradeoff between stacking and the number of different pulse moments by analyzing the *a posteriori* model covariance matrix. It is shown that a better determination of the model parameters is obtained by increasing the number of pulse moments compared with increasing the stack size until a certain point. From this point, the determination does not change. A distribution of pulse moments based on the resolution of the surface nuclear magnetic resonance sounding kernel is calculated and compared with a standard logarithmic distribution. Our results show that the resolution-derived distribution performs better overall, and the logarithmic distribution is only slightly better for very shallow layers. Finally, we use gating of the free induction decays in order to suppress noise, and we show that, by using a logarithmic distribution of gates, we obtain maximum resolution of the models by only seven gates per decay.

INTRODUCTION

Surface nuclear magnetic resonance (SNMR) is a tool for exploration and characterization of shallow aquifers. Based on the principles of nuclear magnetic resonance, but operating in the Earth's magnetic field and utilizing large surface-based loops, the technique allows for detecting hydrogen protons in pore water (Semenov 1987).

SNMR has been developed rapidly (Hertrich *et al.* 2007; Weichman, Lavely, and Ritzwoller 2000), through improved forward modelling (Lehmann-Horn *et al.* 2011; Valla and Legchenko 2002), more efficient and comprehensive inversion algorithms (Behroozmand *et al.* 2012; Müller-Petke and Yaramanci 2010), better instrumentation (Walsh 2008), and better design of measurement sequences (Legchenko *et al.* 2011; Walbrecker *et al.* 2011). These improvements has resulted in a more reliable estimation of aquifer properties (Legchenko *et al.* 2002; Lubczynski and Roy 2007; Ryom Nielsen *et al.* 2011). Also important are the recent advances in signal processing (Dalgaard, Auken, and Larsen 2012; Larsen, Dalgaard, and Auken 2013; Müller-Petke and Costabel 2013), which improves the signal-to-noise ratio (SNR) and has enabled a more routinely use of the method in urbanized areas.

SNMR provides depth resolution by changing the pulse moment for the excitation pulses. The pulse moment (q) is the product of pulse amplitude and duration, and the unit is ampere multiplied by time (As). Usually, pulse duration is kept constant, and only the pulse amplitude is changed. A typical dataset for a depth "sounding" may consist of up to some tens of pulse moments in the range between 0.01 As and 20 As. The number of pulse moments in a sounding is referred to as N_Q. In order to improve the SNR for a certain pulse moment, a number of raw recordings are stacked. The number of recordings for a single pulse moment is called the stack size N_S. A typical stack size ranges from less than 10 but may also in some cases exceed 100, resulting in sounding acquisition times between 3 hours and 9 hours. For T_1 soundings, acquisition times will be around three times higher due to the multiple pulse experiments (Müller-Petke, Walbrecker, and Knight 2013). One of the main challenges with SNMR is the long data acquisition time with several factors controlling it: (i) the relaxation processes rebuilding the equilibrium state of the water molecules from one pulse to another, (ii) the number of pulse moments N_Q, (iii) the stack size N_S, and (iv) the data transfer capability of the instrument.

Legchenko and Shushakov (1998) presented a scheme that allows for calculating the minimum number and optimal distribution of pulse moments for a given SNR. They argue that, for a given SNR, it is not efficient to further increase N_Q as this would only increase measurement time but does not further improve the estimated subsurface model. However, the SNR of a sounding also depends on N_S; therefore, the measurement time depends on both N_Q and N_S. The purpose of this paper is to investigate the optimum relation between N_Q and N_S resulting in the best resolution of the underlying SNMR model. This relation is investigated for a sounding where the acquisition time is fixed, thereby a

* ebd@qeye-labs.com

fixed amount of excitation pulses. Within the same time, we could measure a sounding with a high N_Q and with a low N_S (i.e., low SNR for each pulse moment) or we can measure a sounding with a low N_Q but with a corresponding higher N_S (i.e., high SNR). In addition, we will investigate the effect of gating the free induction decay (FID), as discussed by Behroozmand *et al*. (2012). Gating is an efficient way of suppressing noise (Macnae, Lamontagne, and West 1984; Nyboe and Sørensen 2012). We introduce a logarithmic gating and analyse the minimum number of gates per decade (N_G), which must be used to sample the FIDs. Considering the data amount to be handled in the inversion and that the FIDs get visually clearer with a high N_S, one would like to have N_Q and N_G as low as possible, while retaining as much resolving power as possible. Note that logarithmic gating is a common approach in laboratory NMR used to resample typically multi-exponential T_2 relaxation time data (e.g., Whittall, Bronskill, and Henkelman 1991).

Currently, there are two different pulse transmission schemes used by the commercially available instruments. The first scheme transmits all pulse moments, from high to low, during one discharge of the instrument capacitors. In this case, the pulse moments are distributed logarithmically between the lowest and highest pulse moments. We call this logarithmic distribution (LD). The second scheme does all excitation pulses at each pulse moment before going to the next. The benefit of the first scheme is that no recharge of the system is needed between the successive excitation pulses. The recharge time may exceed the time needed to rebuild the equilibrium state, and in this case, it will increase the total acquisition time. However, it is hard to quantify the time spent on recharge as it is specific for the given type of instrument. An investigation of this is beyond the scope of this paper. The benefit of the second scheme is that the pulse moment distribution can be user specified and thereby optimized. We do this by investigating the data importance matrix of the 1D SNMR kernel. In this paper, we call this distribution the resolution-derived distribution (RD).

In the following, we analyse the tradeoff between N_Q and N_S for both the RD and the LD pulse schemes, while also applying gating with different sampling densities to the FIDs. The analysis is done by studying the parameter standard deviations based on the *a posteriori* model covariance matrix. The analyses are conducted both by adding synthetic random distributed noise to the data and by adding noise collected at a field site.

METHODOLOGY

To analyse the optimal distribution of pulse moments and gates, we (i) generate data from a known model, (ii) add simulated or field noise, (iii) apply the same processing scheme to the simulated data as we would do to real data, and (iv) calculate the *a posteriori* model parameter covariance matrix. The latter is based on the data sensitivity (the Jacobian) and the data covariance matrix, including the data uncertainties calculated during data processing.

A review of SNMR data processing

Processing Surface nuclear magnetic resonance (SNMR) data follows five steps as described by Dalgaard *et al*. (2012). These are:

- de-spiking;
- multichannel Wiener filtering;
- stacking;
- envelope detection;
- gating.

Explained in more detail, the spikes from electrical discharges are first removed in the time domain (number 1), followed by a multichannel Wiener filtering (number 2), which suppresses correlated noise such as powerline harmonics. In the case of an ideal de-spiking and an ideal Wiener filtering, the SNMR signal is then only contaminated by random distributed Gaussian noise. The random noise is suppressed by stacking (number 3), i.e., an averaging of the multiple recordings at each pulse moment. When stacking uncorrelated recordings with Gaussian distributed noise with an initial standard deviation STD_{ini}, the standard deviation $STD_{stacked}$ is reduced as (Yin *et al*. 1996):

$$STD_{stacked} = \frac{1}{\sqrt{N_S}} * STD_{ini} \tag{1}$$

The free induction decay (FID) is obtained by an envelope detection (number 4) of the oscillating signal, and finally, the FID are logarithmically gated with a boxcar function (number 5) as gating efficiently suppresses random noise (similar to equation (1)) and acts as a low-pass filter (Macnae *et al.* 1984; Nyboe and Sørensen 2012).

In this paper, we analyse both the influence of synthetic noise and field noise measured with a GMR System from Vista Clara. In the former case, the processing only involves numbers 3–5.

SNMR forward modelling, theoretical data

For the forward modeling of the SNMR data, the entire dataset is simulated at all different pulse moments, and the data points in the free induction decays (FIDs) are sampled at each 0.8 ms corresponding to a sampling frequency of 1250 Hz. The 1D SNMR response is given by (Behroozmand *et al.* 2012)

$$V(q,t) = \int K(q,z) \cdot W(z) \cdot \exp\left(-\left(\frac{t}{T_2^*(z)} \right)^{C(z)} \right) dz \tag{2}$$

where $V(q, t)$ is the quadrature voltage in the receiver as a function of pulse moment q and time t. $K(q, z)$ is the 1D SNMR kernel, z is the depth, and $W(z)$ is the water content distribution. The model is a function of the relaxation time T_2^* and the stretching exponent C for the unimodal T_2^* distribution at each depth interval.

In order to add field noise to the data and to incorporate the influence of the frequency detection in the analysis, all synthetic FIDs $V(q, t)$ are interpolated to 10 kHz. Then, the synthetic FIDs are transformed into their signal $S(q, t)$ counterpart oscillating at the Larmor frequency f_L.

$$S(q,t) = V(q,t) \cdot \cos(2\pi \cdot t \cdot f_L + \varphi) \quad (3)$$

where φ is the phase offset of the signal. In the synthetic noise case, the synthetic signal is perturbed by Gaussian noise with a standard deviation of 1024 nV at the sample frequency of 10 kHz, which in this analysis compares to a SNR of around 0.1, where the signal level is defined as the initial amplitude level. In the field case, the synthetic signal is perturbed by noise recordings obtained in the field.

Pulse moment distributions

Similar to Legchenko and Shushakov (1998), the resolution-derived distribution (RD) of pulse moments are derived by looking at the data importance. The data importance describes the dependence of the measured data, which in our case are the data measured for different pulse moments. For instance, the closer two pulse moments are, the more similar the depth interval screened by this pulse moment will be, and the individual data importance will decrease. We analyse the data importance using the singular value decomposition (SVD) of the 1D SNMR kernel function K (Aster *et al*. 2005):

$$K = U \Lambda V^T \quad (4)$$

where Λ is a diagonal matrix containing the singular values in decreasing order, and U and V are unitary matrices formed by the complete set of data and model eigenvectors, respectively. We then derive the data importance matrix as follows:

$$R_d = U_r U_r^T \quad (5)$$

from a truncated SVD, where r denotes the level of truncation, i.e., only the first r data eigenvectors are taken into consideration. Because data eigenvectors exhibit a natural ordering according to their distinguishability (Friedel 2003), setting r defines the number of the most distinguishable pulse moments, i.e., pulse moments showing the least amount of overlapping depth intervals. While, in practice, r needs to be set according to the data quality, e.g., Legchenko and Shushakov (1998) derived their optimal distributed set of pulse moments for a certain number of eigenvectors (or basis functions) derived from a given SNR, we use r to obtain the best N_Q from a huge set of pulse moments. In a case that all pulse moments are equally important and do not depend on each other, R_d equals the identity matrix. In our case, however, pulse moments are dependent, and the main diagonal value is below one and non-zero entries showing the dependencies as illustrated in Fig. 1. Thus, R_d is used to analyse the data importance and how much correlated information the data contains. Once R_d is calculated for a given r (for instance, if we are targeting to select the N_Q=20 most distinguishable out of a set of 1000 pulse moments, r is set to 20), R_d is analysed starting with the first row. The index of the peak maximum (see Fig. 1b) is derived and becomes the first pulse moment. Second, the index

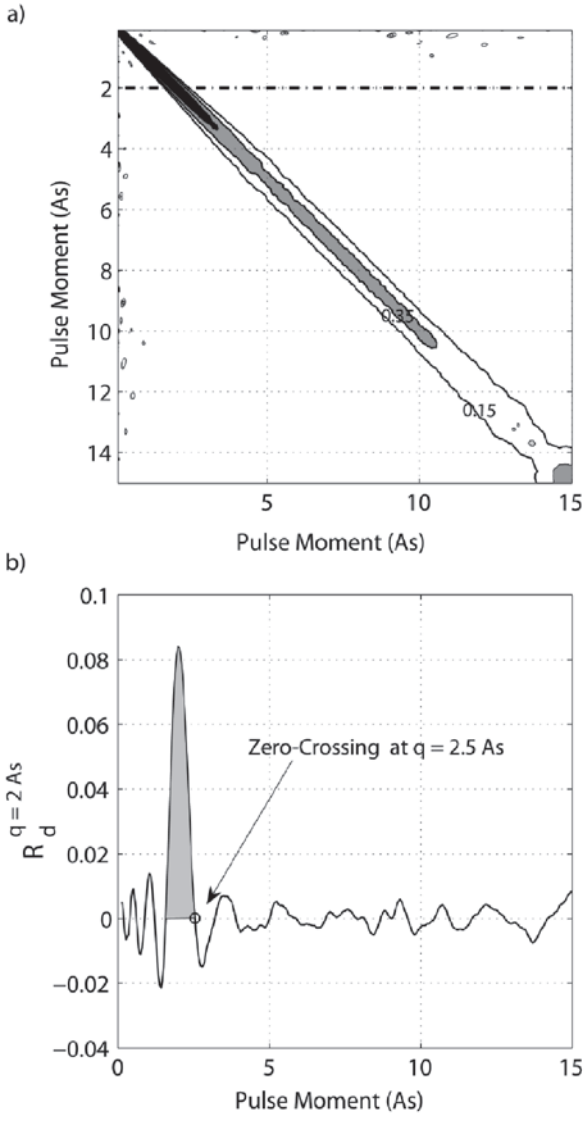

Figure 1 (a) Example of a data resolution matrix (R_d) for a circular loop of 96-m diameter and 100-Ωm half-space resistivity, calculated for a truncation level of 20. (b) Single row of R_d for a pulse moment of 2 As. For a completely independent pulse moment, this would be a delta distribution. The grey-filled area shows dependent pulse moments according to the selected pulse moment of 2 As. The zero-crossing indicate the next pulse moment to be analysed.

of the first zero-crossing (or alternatively a value below a certain threshold) is determined. This zero-crossing index becomes the index of the next row to analyse. This procedure is continued for the complete matrix.

The distribution of pulses obtained is apparently a mixture of linear distributed low pulse moments, whereas for high pulse moments, it follows an LD (Fig. 2a). The distributions in Fig. 2 are calculated for a circular loop of 96-m diameter and 100-Ωm half-space resistivity. Several different values for the resistivity have been tested and have shown similar results. Calculating the

distributions of singular values for 1D SNMR kernel functions consisting of RD, LD, and linear distribution of pulse moments

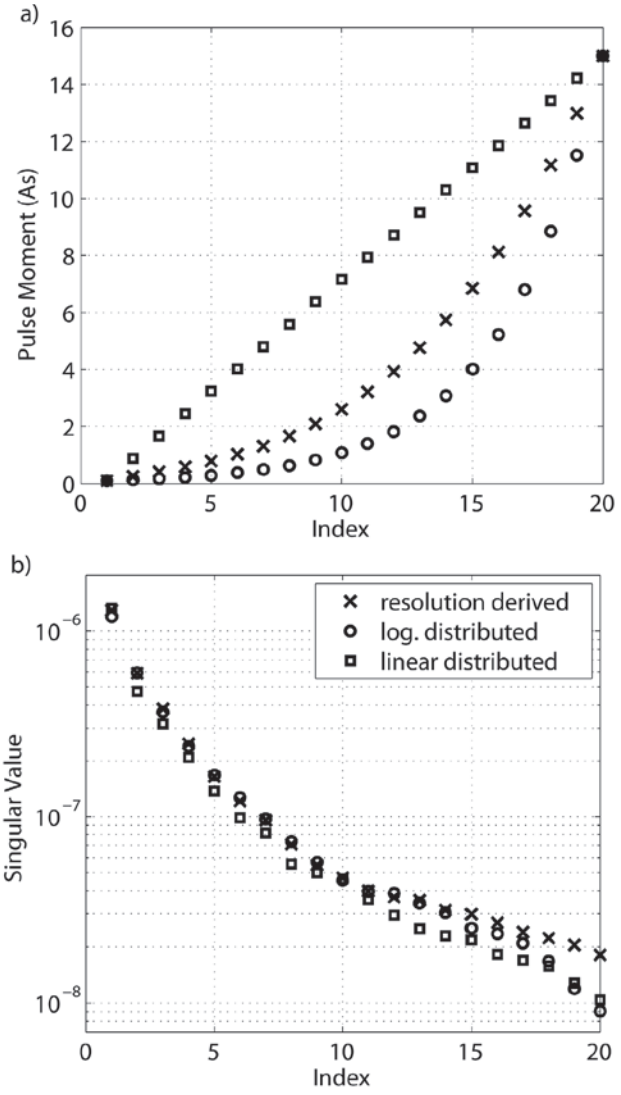

Figure 2 (a) Distribution of 20 pulse moments for different schemes (RD, linear distribution, and LD). (b) Singular value distribution for 1D SNMR kernel function consisting of 20 pulse moments. The distributions are calculated for a circular loop of 96-m diameter and 100-Ωm half-space resistivity.

allows for a first comparison of resolution capabilities. According to an inversion using a truncated SVD, resolution depends on the number of singular value above a certain truncation. The truncation level depends on the data quality. The singular values for the RD fall off slower compared with logarithmic or linear distribution (Fig. 2b); thus, for the same truncation level, more singular values can be included. Because the linear spaced pulse distribution is not implemented in current instrumentation, we focus our evaluation on the RD and the logarithmic-spaced distribution.

We have additionally calculated RDs for several different half-space resistivities and found the sequence to be mostly independent of the assumed resistivity.

Model parameter uncertainty

For the sensitivity analysis of the SNMR parameters, few-layer 1D models are used, and no *a priori* information or regularization was applied to any of the model parameters. Based on a linear approximation to the *a posteriori* model covariance matrix C_{est}, the estimation of the model parameter uncertainty is given by (Auken and Christiansen 2004; Menke 1989; Tarantola and Valette 1982)

$$C_{est} = \left[G^T C_{obs}^{-1} G \right]^{-1} \tag{6}$$

where G is the Jacobian matrix of the forward mapping containing the partial derivatives of the mapping and C_{obs} are the covariance matrices of the observed data. The parameter uncertainty estimates are then obtained by the square root of the diagonal elements of C_{est}. A similar approach for accessing the parameter uncertainty has been reported in (Christensen and Dodds 2007; Christensen and Lawrie 2012).

The analyses were carried out on the logarithm of the model parameters, which provides a standard deviation factor (STDF), on the parameter m_i, given as follows (Behroozmand *et al.* 2013):

$$STDF\left(m_i\right) = \exp\left(\sqrt{C_{est(i,i)}}\right). \tag{7}$$

Therefore, under a lognormal assumption, it is 68% likely that a given model parameter m falls in the interval:

$$\frac{m}{STDF_m} < m < m \cdot STDF_m. \tag{8}$$

Table 1 Thickness, resistivity, water content, relaxation time (T_2^*) and stretching parameter (C) for models A, B, and C. The difference between the models is the thickness of the first layer. A three-layer case reflects an unsaturated top sand layer, a saturated sandy mid layer, and a saturated clay bottom layer.

	Model A Thickness	Model B Thickness	Model C Thickness	Resistivity (Ωm)	Water Content (%)	T_2^* (ms)	C
Layer 1	20 m	5 m	30 m	500	5	300	1
Layer 2	5 m	5 m	5 m	100	35	300	1
Layer 3	Inf	Inf	Inf	10	45	10	1

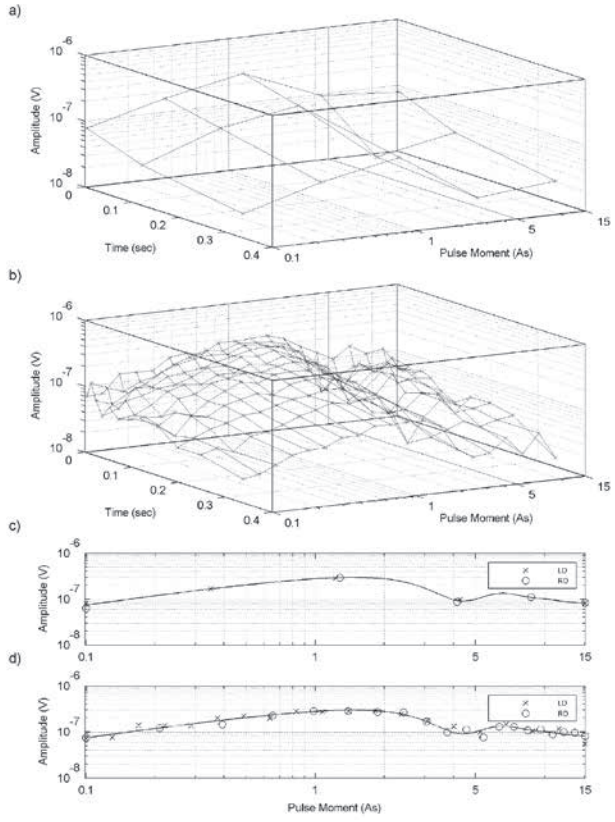

Figure 3(a) Data cube described by five logarithmic distributed pulse moments and three time gates. (b) Data cube described by 20 logarithmic distributed pulse moments and 12 time gates. (c) Sounding curve at the first gate (4 ms); the solid line represents the synthetic data response compared with the data points from both the LD (x) and RD (o) schemes. Top plot shows data points by having five pulse moments. Bottom plot shows data points by having 20 pulse moments. Data calculated from Model A.

An STDF below 1.1 correspond to a very well determined parameter, and a STDF above 2.0 is said to correspond to a completely undetermined parameter.

RESULTS

The base model, model A, is a three-layer model (Table 1), imitating a 20-m unsaturated sand layer, followed by a 5-m sandy aquifer and then a water-saturated clay layer. The pulse moments are chosen between 0.1 As and 15 As, transmitted into a 50-m-side square loop. The pulse duration is set to 20 ms, and the dead time of the system is 10 ms. Both the logarithmic distribution (LD) and resolution-derived distribution (RD) pulse schemes are analysed. The premise for the analyses is that the sounding contains 1500 individual records. These 1500 records are measured for a number of pulse moments N_Q, each consisting of a number of recordings N_S so that

$$N_Q \cdot N_S = 1500\,. \tag{9}$$

Two extreme cases are either a situation with 1500 pulse moments with a stack size of 1 or a situation where the sounding has 1 pulse moment with a stack size of 1500. In the second case, the noise level would be $\sqrt{1500} \cong 39$ times smaller than in the first case (given uncorrelated Gaussian distributed noise). The first analysis will answer the question of whether a high stack size or many pulse moments are desirable. The second analysis will be related to gating. Gating is defined as the number of gates per decade (N_G) each individual free induction decay (FID) consists of. This analysis will investigate the tradeoff between having many gates with a relative low data quality and having fewer observations with a higher data quality.

Figure 3 shows data from model A with different realizations of N_Q, N_S, and N_G. Figure 3(a) shows 15 data points, three gates for each FID, and five pulse moments distributed logarithmically (LD). Figure 3(b) shows the data sampled by 20 pulse moments and 12 gates for each FID. The total number of data points is 240. Figure 3(c, d) shows the sounding curve for gate 1 sampled with the RD and LD schemes for the two realizations above. It is clear from Fig. 3(c) that the LD scheme has gates distributed denser at the lower values compared with the RD scheme. In the case where N_Q=5, although the data have a low noise contribution, due to the intensive stacking of N_S=300, the gates are not sampling the sounding curve at the higher pulse moments. At low moments, the sounding curve varies slowly, and the gates sample the sounding curve well. Considering N_Q=20, we see in Fig. 3(b) that the data points are more scattered as expected, due to the lower stacking number N_S=75, but the data follow the trend of the sounding curve throughout the pulse moments. From this, we can expect shallow layers to be well determined by having both a low and a high number of pulse moments, whereas deeper layers will be less determined in the case where N_Q=5. This is investigated in the analysis in the succeeding section.

Stacking and number of pulse moment

Figure 4 shows the results of the analysis of water content (W) and relaxation in the transverse plane (T_2^*) for each layer in the synthetic noise case. Data are calculated from model A. In general, it is seen that the uncertainty, i.e., the standard deviation factor (STDF), of the model parameters decreases significantly until a certain N_Q. From this N_Q and above, the STDF stays at the same level, implying that the parameter cannot be determined better. For the first layer, W and T_2^* have reached the steady level for both for the RD and the LD schemes at N_Q=3. The STDF of the second layer reaches a steady level for the RD scheme for a N_Q of 6 to 10, although the corresponding number for the LD scheme is about 15. For the water content, it is a general trend that the STDF gets higher with depth. This is clearly seen for the third layer, where again the steady level for the RD is between 6 and 10 pulse moments, but with a STDF well above 2. The STDF for the thickness of the first layer is highly correlated to the water content (Legchenko *et al*. 2004); thus, we see the steady level of the STDF after four pulse moments both for the RD and LD. The STDF for

the thickness of the second layer is also highly correlated to the water content of the same layer; thus, we see diverging patterns for the RD and LD, and again, the RD reaches a steady level at 6–10 pulse moments, whereas the LD are at a steady level from 15 pulse moments. We see for the two uppermost layers, where T_2^* is high, that the STDF is low and, thus, $T_2^*$is well determined. In the third

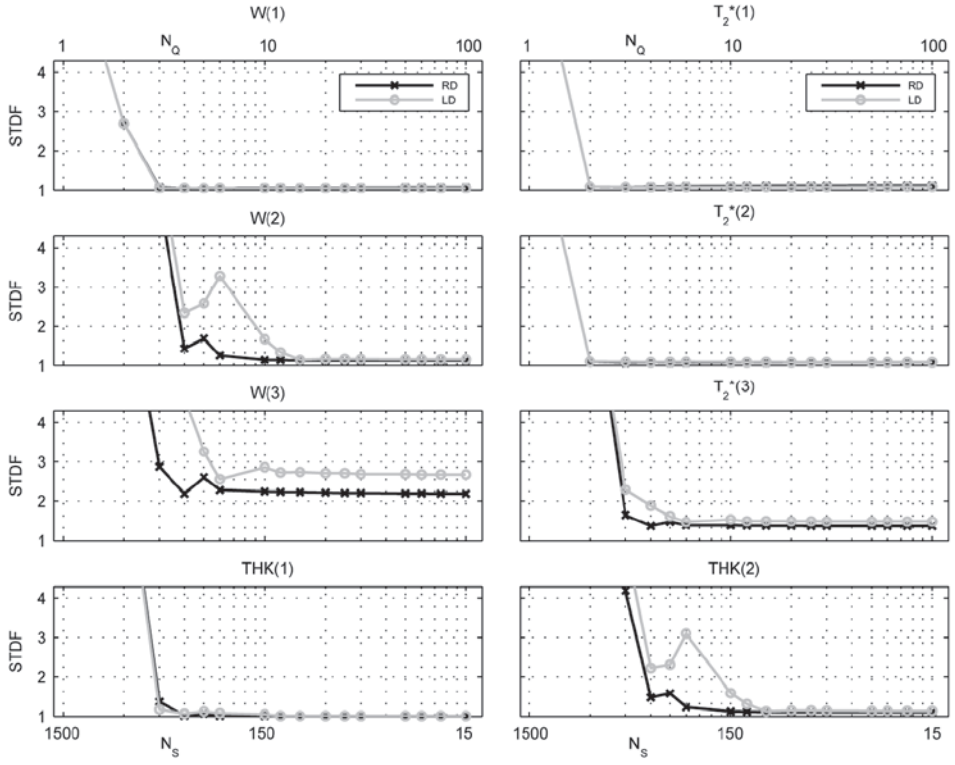

Figure 4 Analysis result for the model parameters: water content (W), relaxation time (T_2^*), and thickness (THK). Data, calculated from Model A, are perturbed by synthetic Gaussian noise. *X*-axis plots the number of pulse moments N_Q, with the corresponding stack size N_S. *Y*-axis plots the standard deviation factor STDF.

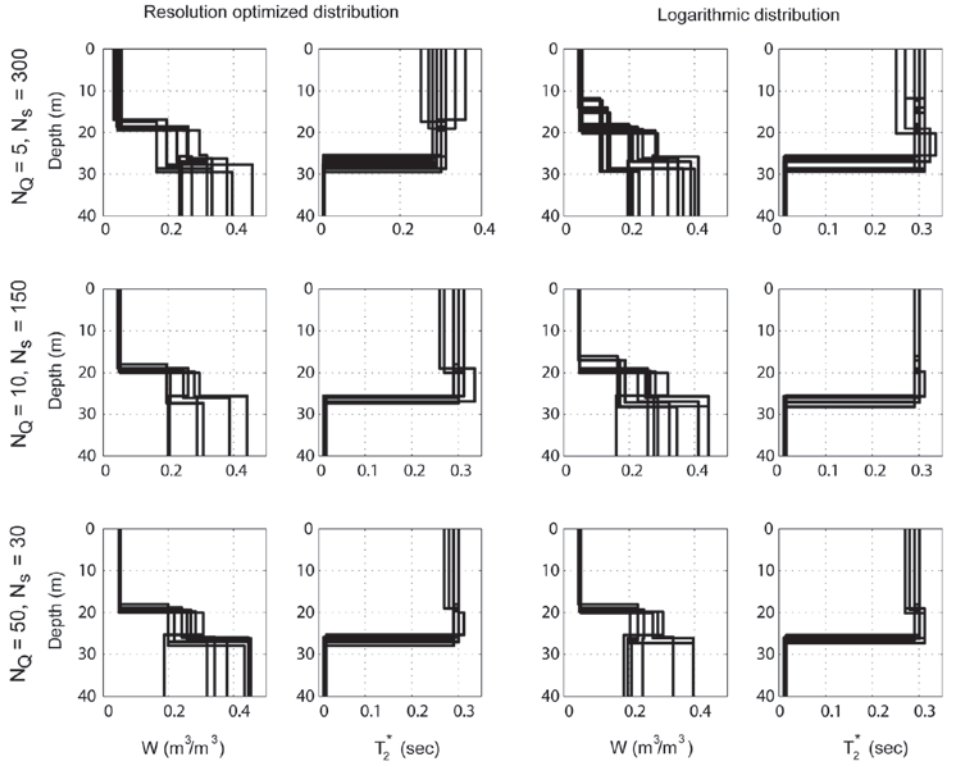

Figure 5 Sets of inversion results using a GA for synthetic data of Model A; 5, 10, and 50 pulse moments, both LD and RD. Results from fixed and free layer thicknesses are presented together.

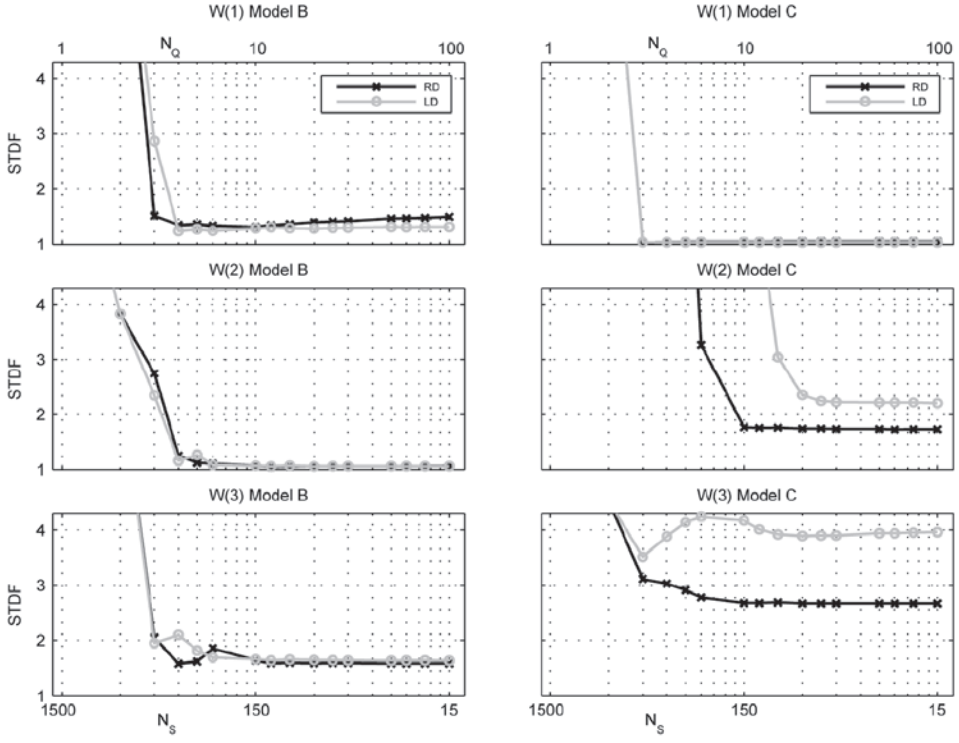

Figure 6 Analysis result for the model parameter water content (W) for models B and C. Model B corresponds to model A with the layer boundary to the second layer shifted up to 5 m below the surface. Model C corresponds to model A with the layer boundary to the second layer shifted down to 30 meters below the surface. *X*-axis plots the number of pulse moments N_Q, with the corresponding stack size N_S. *Y*-axis plots the standard deviation factor STDF.

layer, where T_2^* is low, the steady level is reached at a higher number of pulse moments, and the STDF is also higher.

The surface nuclear magnetic resonance (SNMR) inverse problem is non-linear (equation (2)), and although the methodology of using the covariance matrix is strictly valid only for linear inverse problems, it has been successfully applied for non-linear inverse problems (Behroozmand *et al.* 2013; Christensen and Dodds 2007; Christensen and Lawrie 2012). We decided to investigate the model space using a genetic algorithm (GA) in order to ensure that the linearized covariance calculation gives acceptable results. Some early examples of GAs in geophysical parameter estimation are described in Stoffa and Sen (1991) and a recent applications for SNMR parameter estimation by Akca *et al.* (2014). GAs are global search algorithms that not only allows for providing a single solution reflecting the global minimum of a misfit function but allows for investigating model uncertainty by providing sets of solutions. As GAs are relatively computational intensive, we decided to limit our investigation to three different numbers of pulse moments N_Q (5, 10, 50), both for RD and LD. In addition to running the inversion without limiting the parameter search space, we conducted several inversion runs varying a fixed layer thickness. The latter approach ensures that the model uncertainty is captured as complete as possible because, to our experience, extreme model parameter values are less well represented by an unlimited search space. However, with this procedure, a standard deviation from the set of obtained solutions cannot be calculated. Figure 5 shows a set of solutions for each case instead of calculating an STDF. Qualitatively, the results of these GA-based analyses nicely match the findings of the STDF analyses for Model A: (i) for a low number of pulse moments, the RD outperforms the LD, and (ii) model uncertainty decreases until a certain N_Q and becomes steady for further increasing N_Q. Consequently, as the GA analysis confirms the results of the covariance-based analyses but is significantly slower, we will in the remaining analysis calculate the covariance matrix.

As discussed earlier, the depth sensitivities between the RD and LD pulse schemes are different. This is investigated by changing the depth of the layer boundaries in the model. For this

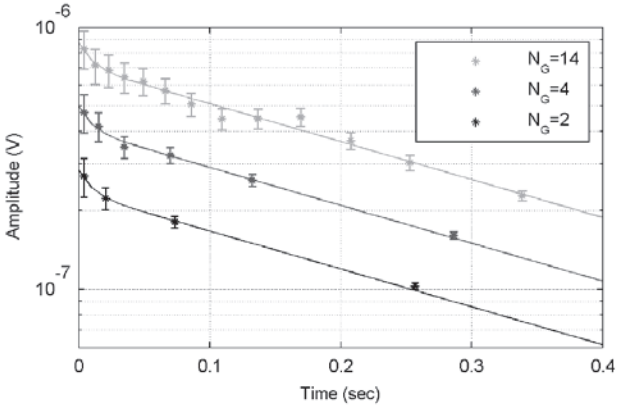

Figure 7 An example of a FID sampled with 2, 4, and 14 gates per decade. All gates are plotted with their standard deviation. The width of the first gate is 8 ms. The response for the three examples are the same and plotted as the solid line. The example with N_G=4 is shifted with a factor of 1.75 above N_G=2. The example with N_G=14 is shifted with a factor of 1.75 above N_G=4.

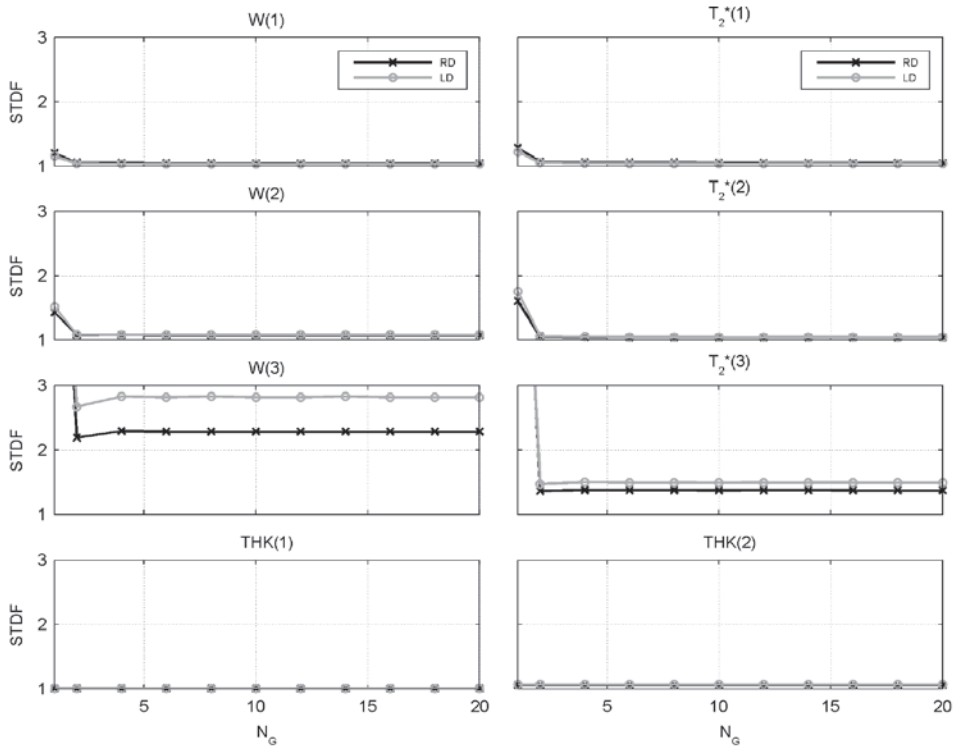

Figure 8 Gating analysis result for the model parameters water content (W), relaxation time (T_2^*), and thickness (THK). *X*-axis plots the number of gates per decade N_G. *Y*-axis plots the standard deviation factor STDF. The first gate is in all scenarios fixed at 4 ms.

analysis, we introduce two new models, model B and C, only differing from model A by the thickness of the first layer, which again affects the depth to the second layer. For model B, the thickness of the first layer is set to 5 m, and in model C, it is set to 30 m. Figure 6 shows the analysis for the water content of all layers. For model B, we observe that the lowest STDF is obtained for the second layer, and in this layer, the RD and the LD pulse schemes perform equally well. In the uppermost layer, W is slightly better determined by the LD scheme compared with the RD. In the third layer, the water content is slightly better determined by the RD scheme compared with the LD. In model C, (Fig. 6), we observe a bigger difference between the two pulse schemes. In the second layer, the RD scheme determines the water content with an STDF below 2, whereas the LD has an STDF above 2. For the water content in the third layer, the difference is bigger: The STDF for the LD is around 4, whereas the STDF for the RD is below 3. For the first layer in Model B, there is no detectable difference between the two pulse schemes. In conclusion, we see that LD is slightly better for resolving thin near-surface layers, whereas RD performs much better for deeper layers.

Gating

In the following, we will investigate the results from varying the number of gates used to sample each FID. The idea of using gates is that they efficiently suppress noise (Macnae *et al*. 1984), and we thus want them as wide as possible but no so wide that we lose information on the model.

The gates are distributed logarithmically with the first gate having a width of 8 ms, with the number of gates per decade being the varying parameter. The model to be analysed is model A (Table 1), and we use 15 pulse moments for both the RD and

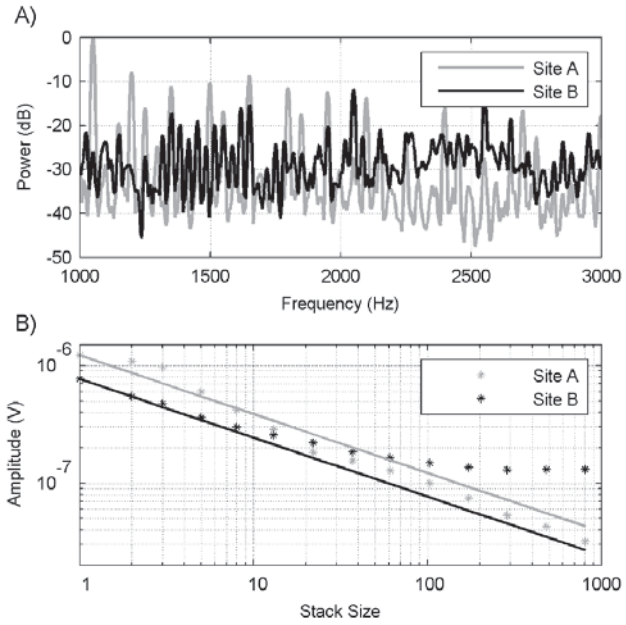

Figure 9 (a) Noise spectra from two different sites in Germany. (b) Stack size versus amplitude for the data collected at the two different German sites. The solid lines represent the theoretical stack relation for random Gaussian noise. The dots represent the actual value after stacking.

LD pulse schemes. An example of an FID sampled with 2, 4, and 14 gates per decade are shown in Fig. 7. The coarsely sampled FID has relatively low standard deviations for each gate, whereas the denser sampled FID has higher gate standard deviations. Results for the analysis are shown in Fig. 8. We see a clear trade-off between many gates with a high standard deviation and fewer gates with a corresponding lower standard deviation. The result shows that we need two gates per decade to sample the curve. From analyses with various other models, it is our experience that the model parameter with the biggest influence on the number of gates is the relaxation time. Therefore, in some cases where the relaxation is short (10 ms), the minimum number of gates has to be as high as five to seven per decade. Thus, balancing between noise suppression by the gate and model resolution, we recommend using seven gates per decade. Note that, in laboratory nuclear magnetic resonance, about 30 gates per decade are common (Whittall *et al*. 1991). This is due to the detectability of short relaxation signals (below 1 ms) and higher SNR. It can be expected that with increasing detectability of short relaxation signals and further improvements to obtain higher SNR in SNMR, the optimal number of gates will increase.

Field noise

In order to investigate the validity of equation (1), we have taken noise recordings (field noise) obtained by the GMR instrument (Vista Clara Inc.) and processed it following the processing steps discussed in the methodology section. The processed noise data are then plotted with N_S against the STD after stacking. Together with the actual STD after stacking, the theoretical STD is plotted. The theoretical STD is calculated from Equation 1, where STD_{ini} is obtained from the very first recording.

Two German sites, A and B, are analysed in order to see if the noise after processing behaves like Gaussian noise. The spectra from the two sites are shown in Fig. 9(a) and demonstrate two quite different behaviours in terms of electromagnetic noise. Site A is strongly dominated by harmonic noise, whereas the spectrum of site B is more flat but with a higher base level and less dominated by harmonic noise. As shown in Fig. 9(b), for site A, the actual STD after processing follows the theoretical relation, indicating that the processed data are nearly Gaussian distributed. At site B, it is evident that the processed data do not follow the theoretical relation. This indicates that field noise is not necessarily Gaussian, and an analysis of examples where field noise is added to the data before taking final conclusions is needed.

The analysis from Section 3.1 is repeated with data calculated from model A but this time perturbed by field noise. The field noise is measured at site B, Schillerslage, German Site, and results are shown in Fig. 10. It is seen that the trends is the same for field noise as for the synthetic noise. The water content W of the first layer, and T_2^* for layers 1 and 2, are well determined after sampling the data space with N_Q equal to 3. For W of the second and third layers, we see different results with the RD and the LD pulse schemes. For the RD, we reach the steady STDF level at 6–10 pulse moments, and for the LD scheme, the steady level is reached at around 15 pulse moments. For the T_2^* of the third layer, the RD is at the steady

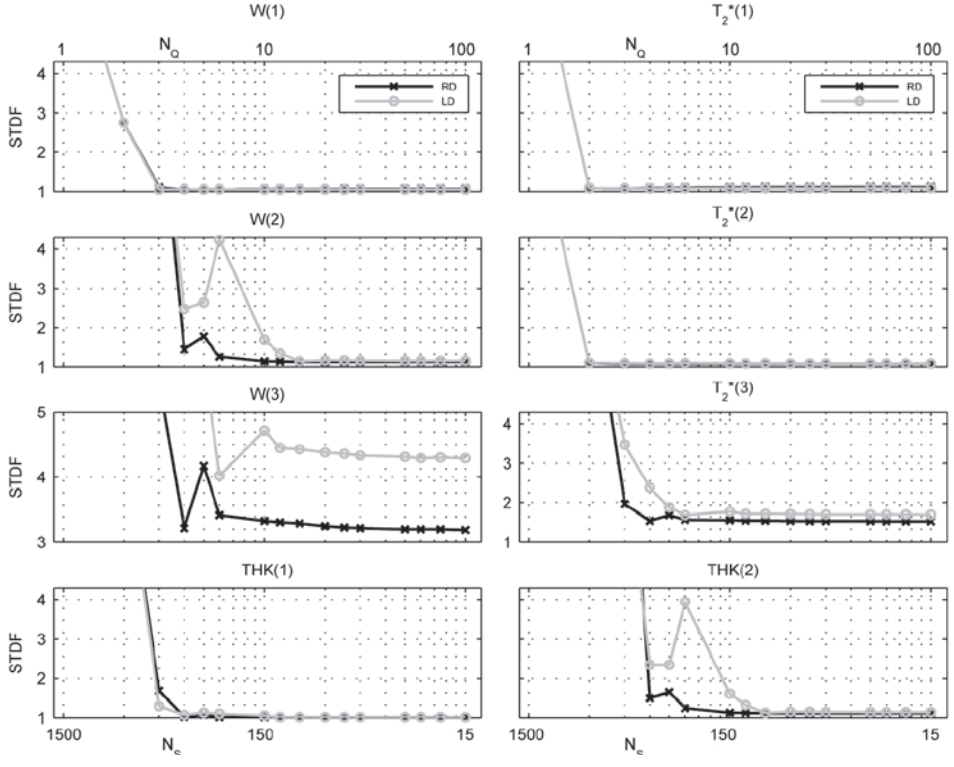

Figure 10 Analysis result for the model parameters water content (W), relaxation time (T_2^*) and thickness (THK). Data are perturbed by noise collected with the GMR instrument from Vista Clara at Schillerslage in Germany (Fig. 9, site B). *X*-axis plots the number of pulse moments N_Q, with the corresponding stacks size N_S. *Y*-axis plots the standard deviation factor STDF.

level at 4–6 pulse moments, and the LD is at steady level around 6–10. As in the case with the synthetic noise, we see in this analysis a high correlation between the water content and the thickness of the first layer, and a high correlation between the water content and the thickness of the second layer. In the third layer, the RD scheme performs better than the LD when the steady level is reached, and in the other layers, the two schemes perform equally well. Thus, it is concluded that the results from using synthetic and field noise are comparable and show the same overall trends.

CONCLUSIONS

In this study, we have analysed how well the model parameters are determined by varying the number of pulse moments and stack size by keeping the total number of recordings fixed to 1500. In this framework, we have introduced a resolution-derived pulse moment distribution and analysed how the distribution of pulse moments influences the parameter determination. Furthermore, we have introduced a logarithmic gating process for the free induction decay (FID) and analysed the number of gates necessary to sample the signal while also decreasing the noise as much as possible. The number of necessary gates has been tested for various models.

It was shown that the resolution-derived distribution (RD) pulse scheme performed overall better than the logarithmic distribution (LD) scheme and only for very shallow layers is the LD scheme better than the RD scheme. Both for the synthetic and for the field noise case that a better determination of the model parameters is obtained by increasing the number of pulse moments $N_{Q,}$ compared with increasing the stack size N_S, until a certain point. From this point, a steady level is reached, and one can increase either N_Q or N_S and obtain the same determination of model parameters.

The specific number of pulse moments and gates to apply on surface nuclear magnetic resonance soundings is model dependent, but the analyses show that one should apply 12–14 pulse moments and seven gates per decade using the RD sampling scheme. For models where very shallow water-bearing layers are present, we recommend to increase the sampling density for the low pulse moments.

ACKNOWLEDGEMENTS

This work was supported by the Danish Council for Strategic Research funded project “Nitrate Reduction in Geologically Heterogeneous Catchments” (NiCA).

We would like to thank Jakob Juul Larsen and Ahmad Behroozmand from Aarhus University, and Raphael Dlugosch and Thomas Günther from Leibniz Institute for Applied Geophysics (LIAG) for the very fruitful discussions and helpful comments during the work. We further thank the associate Editor Anatoly Legchenko, Jean-Francois Girard, and an anonymous reviewer for their helpful comments improving the clarity of the manuscript.

REFERENCES

Akca İ., Günther T., Müller-Petke M., Başokur A.T. and Yaramanci U. 2014. Joint parameter estimation from magnetic resonance and vertical electric soundings using a multi-objective genetic algorithm. *Geophysical Prospecting* **62**(2), 364–376.

Aster R.C., Borchers B. and Thurber C.H. 2005. *Parameter Estimation and Inverse Problems.* Elsevier Academic Press.

Auken E. and Christiansen A.V. 2004. Layered and laterally constrained 2D inversion of resistivity data. *Geophysics* **69**, 752–761.

Behroozmand A.A., Auken E., Fiandaca G., Christiansen A.V. and Christensen N.B. 2012. Efficient full decay inversion of MRS data with a stretched-exponential approximation of the T2* distribution. *Geophysical Journal International* **190**, 900–912.

Behroozmand A.A., Dalgaard E., Christiansen A.V. and Auken E. 2013. A comprehensive study of the parameter determination in a joint MRS and TEM data analysis scheme. *Near Surface Geophysics* **11**, 557–567.

Christensen N.B. and Dodds K. 2007. 1D inversion and resolution analysis of marine. CSEM data **72**, WA27–WA38.

Christensen N.B. and Lawrie K.C. 2012. Resolution analyses for selecting an appropriate airborne electromagnetic (AEM) system. *Exploration Geophysics* **43**, 213–227.

Dalgaard E., Auken E. and Larsen J.J. 2012. Adaptive noise cancelling of multichannel magnetic resonance sounding signals. *Geophysical Journal International* **191**, 88–100.

Friedel S. 2003. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach. *Geophysical Journal International* **153**, 305–316.

Hertrich M., Braun M., Günther T., Green A.G. and Yaramanci U. 2007. Surface nuclear magnetic resonance tomography. *IEEE Transactions on Geoscience and Remote Sensing* **45**(11), 3752–3759.

Larsen J.J., Dalgaard E. and Auken E. 2013. Noise cancelling of MRS signals combining model-based removal of powerline harmonics and multichannel Wiener filtering. *Geophysical Journal International* **196**(2), 828–836.

Legchenko A. and Shushakov O.A. 1998. Inversion of surface NMR data. *Geophysics* **63**(1), 75–84.

Legchenko A., Baltassat J.M., Beauce A. and Bernard J. 2002. Nuclear magnetic resonance as a geophysical tool for hydrogeologists. *Journal of Applied Geophysics* **50**, 21–46.

Legchenko A., Baltassat J.-M., Bobachev A., Martin C., Robain H. and Vouillamoz J.-M. 2004. Magnetic resonance sounding applied to aquifer characterization. *Groundwater* **42**, 363–373.

Lehmann-Horn J.A., Hertrich M., Greenhalgh S.A. and Green A.G. 2011. Three-Dimensional Magnetic Field and NMR Sensitivity Computations Incorporating Conductivity Anomalies and Variable-Surface Topography. *IEEE Transactions on Geoscience and Remote Sensing* **49**(10), 3878–3891.

Lubczynski M.W. and Roy J. 2007. Use of MRS for hydrogeological system parameterization and modeling. *Boletín Geológico y Minero* **118**(3), 509–530.

Macnae J.C., Lamontagne Y. and West G.F. 1984. Noise processing techniques for time-domain EM systems. *Geophysics* **49**, 934–948.

Menke W. 1989. *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, San Diego.

Müller-Petke M. and Costabel S. 2013. Comparison and optimal parameter setting of reference-based harmonic noise cancellation in time and frequency domain for surface-NMR. *Near Surface Geophysics* **12**, 199–210.

Müller-Petke M., Walbrecker J.O. and Knight R.J. 2013. The inversion of surface-NMR T1 data for improved aquifer characterization. *Geophysics* **78**(6), EN83–EN94.

Müller-Petke M. and Yaramanci U. 2010. QT inversion – Comprehensive use of the complete surface NMR data set. *Geophysics* **75**, WA199–WA209.

Nyboe N.S. and Sørensen K.I. 2012. Noise reduction in TEM: Presenting a bandwidth- and sensitivity-optimized parallel recording setup and methods for adaptive synchronous detection. *Geophysics* **77**, E203–E212.

Ryom Nielsen M., Hagensen T.F., Chalikakis K. and Legchenko A. 2011. Comparison of transmissivities from MRS and pumping tests in Denmark. *Near Surface Geophysics* **9**, 211–223.

Semenov A.G. 1987. NMR Hydroscope for water prospecting. *Proceedings of the seminar on Geotomography, Indian Geophysical Union*, 66–67.

Stoffa P.L. and Sen M.K. 1991. Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms. *Geophysics* **56**, 1794–1810.

Tarantola A. and Valette B. 1982. Generalized nonlinear inverse problems solved using a least squares criterion. *Reviews of Geophysics and Space Physics* **20**, 219–232.

Valla P. and Legchenko A. 2002. One-dimensional modelling for proton magnetic resonance sounding measurements over an electrically conductive medium. *Journal of Applied Geophysics* **50**, 217–229.

Vouillamoz J.M., Legchenko A. and Nandagiri L. 2011. Characterizing aquifers when using magnetic resonance sounding in a heterogeneous geomagnetic field. *Near Surface Geophysics* **9**, 135–144.

Walbrecker J.O., Hertrich M., Lehmann-Horn J.A. and Green A.G. 2011. Estimating the longitudinal relaxation time T1 in surface NMR. *Geophysics* **76**(2), F111–F122.

Walsh D.O. 2008. Multi-channel surface NMR instrumentation and software for 1D/2D groundwater investigations. *Journal of Applied Geophysics* **66**, 140–150.

Weichman P.B., Lavely E.M. and Ritzwoller M.H. 2000. Theory of surface nuclear magnetic resonance with applications to geophysical imaging problems. P*hysical Review E – Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **62**, 1290–1312.

Whittall K.P., Bronskill M.J. and Henkelman R.M. 1991. Investigation of analysis techniques for complicated NMR relaxation data. *Journal of Magnetic Resonance (1969)* **95**(2), 221–234.

Yin L., Yang R., Gabbouj M. and Neuvo Y. 1996. Weighted median filters: A tutorial. *IEEE Transactions on Circuits and Systems II. Analog and Digital Signal Processing* **43**, 157–192.