Application of BEL1D for sNMR data interpretation

[1]H. Michel
[1] ULiège, UGent, F.R.S-FNRS
Liège, Belgium
hadrien.michel@uliege.be

[2]T. Hermans
[2] UGent
Gent, Belgium
thomas.hermans@ugent.be

[3]T. Kremer
[3] ULiège, UNantes
Nantes, France
thomas.kremer@unantes.fr

[4]F. Nguyen
[4] ULiège
Liège, Belgium
f.nguyen@uliege.be

SUMMARY

The interpretation of sNMR data is still mainly performed using deterministic or stochastic inversion schemes. sNMR signal to noise ratio is often low regarding electromagnetic noise pollution which coupled to non-uniqueness makes uncertainty quantification challenging. Here, we propose a new Bayesian scheme relying on a learning step and a prediction step to perform the interpretation of sNMR data including uncertainty quantification: BEL1D. With it, it is possible to estimate the uncertainty of models parameters from a given dataset in a rapid manner compared to stochastic inversion and reach an equivalent posterior estimation after iterative prior resampling. The learning step can even be used to multiple datasets to improve performances with only the prediction required. Additionally, BEL1D could be used with any geophysical methods.

Key words: Inversion, Machine Learning, MCMc, Uncertainty, BEL.

INTRODUCTION

Inversion in sNMR is often performed using a deterministic scheme (e.g.: Legchenko & Shushakov, 1998; Legchenko & Valla, 2002; Mueller-Petke & Yaramanci, 2010). Those schemes based on linear (or linearized) inversion offer the advantage to be rapid and provide the user with a single model to interpret. However, the obtained model represents the most-likely model in the least-squared sense. Even if they are likely close to the actual geological model, uncertainty quantification is still lacking. Propagating data uncertainty through the covariance matrix is a cost-effective solution (Tarantola & Valette, 1982). However, using such approach will only represent the propagation of the data uncertainty and not encompass the modelling uncertainty.

Appraisal of uncertainty in geophysical inverse modelling is key to an efficient decision-making process (Scheidt *et al.*, 2018). In sNMR, Andersen *et al.* (2018) proved that the solution to the inverse problem showed large uncertainties, and even sometimes water-content/thicknesses correlations. They showed that only a Markov chain Monte Carlo (McMC) approach was suitable for a reasonable estimation that was able to collect all the features of the uncertainty with correlation between models parameters. However, they stated that using an MCMc approach was approximately 1000 times slower than propagating the covariance matrix in the linearized inversion.

We propose another method to solve the inverse problem and estimate the uncertainty that simplifies the Bayesian problem in a reduced space to speed-up calculations. This method, based on Bayesian Evidential Learning (BEL – Scheidt *et al.*, 2018), is well suited for the 1D imaging of the subsurface from geophysical data and called BEL1D (Michel *et al.*, 2020). In previous works (Michel *et al.*, 2020), we demonstrated that BEL1D was efficient, fast and reliable. However, the uncertainty resulting from this approach was slightly overestimated. Here, we propose the use of Iterative Prior Resampling in order to solve this issue.

METHODS

BEL1D (or Bayesian Evidential Learning 1D imaging) is an adaptation of Bayesian Evidential Learning (BEL – Scheidt *et al.*, 2018) for the direct prediction of 1D models parameters based on geophysical data. This implementation of BEL differs thus in the way the end-result is the models and not a parameter of interest that could be obtained through petrophysics (for example).

The algorithm is simplifying the Bayesian problem in a reduced space. It is described extensively in Michel *et al.* (2020) and can be summarized in 7 steps:

1) Define the prior model space ($f(\boldsymbol{m})$) and sample N models ($\boldsymbol{m}$) out of it.

2) Use the forward model (eq. 1) to obtain the response for each of the N models.

$$\boldsymbol{d} = G(\boldsymbol{m}) \qquad \text{(eq. 1)}$$

3) If required, reduce the dimensionality of the models and the related data using principal component analysis (PCA) (eq. 2 and eq. 3).

$$PCA(\boldsymbol{d}) \xrightarrow{yields} \boldsymbol{d^f} \qquad \text{(eq. 2)}$$

$$PCA(\boldsymbol{m}) \xrightarrow{yields} \boldsymbol{m^f} \qquad \text{(eq. 3)}$$

4) Use canonical correlation analysis (CCA) to obtain a statistical relationship between $\boldsymbol{d^f}$ and $\boldsymbol{m^f}$ (eq. 4).

$$CCA(\boldsymbol{d^f}, \boldsymbol{m^f}) \xrightarrow{yields} (\boldsymbol{d^c}, \boldsymbol{m^c}) \qquad \text{(eq. 4)}$$

5) In the reduced space, approximate the posterior distribution of $\boldsymbol{m^c}$ using kernel density estimation (KDE) for any $\boldsymbol{d^c}$ (eq. 5).

$$KDE(\boldsymbol{d^c}, \boldsymbol{m^c}) \xrightarrow{yields} f(\boldsymbol{m^c}|\boldsymbol{d^c}) \qquad \text{(eq. 5)}$$

6) Apply the PCA and CCA transformations to the field dataset ($\boldsymbol{d_{obs}}$) (eq. 6)

$$CCA\big(PCA(\boldsymbol{d_{obs}})\big) \xrightarrow{yields} \boldsymbol{d^c_{obs}} \qquad \text{(eq.6)}$$

7) Sample models from the posterior distribution $\boldsymbol{m}_{post}^{c} \in f(\boldsymbol{m}^{c}|\boldsymbol{d}_{obs}^{c})$. Each model can be back transformed to the original space (PCA and CCA are both linear transformations).

One of the main advantage of this method is that the field data is only necessary from step 6 onward. Therefore, one can see BEL1D as a machine learning algorithm, where the training consists of describing the Bayesian problem in the reduced space and the interference step is the process of extracting the posterior model space for a new dataset. This makes for extremely fast posterior estimation.

We introduce **Iterative Prior Resampling (IPR)** as an efficient way to converge towards the posterior similar to the one obtained by classical MCMc approaches. IPR is inspired from iterative spatial resampling (Mariethoz *et al.*, 2010). Algorithmically, we are adding the models sampled from the posterior to the prior to re-train BEL1D on a more informed prior. This approach enable to overcome issues with large prior uncertainties that result in difficulties to extract trends between data and models.

RESULTS

In this section, we will discuss some results that we obtained using this approach. We will apply BEL1D to a synthetic dataset created using MRSMatlab (Mueller-Petke *et al.*, 2016). We will use a large prior to force the demonstration that it is rather difficult to converge when very few is known in advance regarding the model.

Table 1: Description of the benchmark model and associated prior

Layer	e_i [m]			W_i [%]			$T_{2,i}^{*}$ [ms]		
#	Min	True	Max	Min	True	Max	Min	True	Max
1	1	25	50	1	5	50	5	100	500
2	1	25	50	1	25	50	5	200	500
H-S	/	Inf	/	1	10	50	5	50	500

The benchmark model along with the used prior is presented in Table 1. The dataset is simulated using a classical transmitter/receiver configuration with a 50 m diameter. The sampling frequency is 500Hz from 0.005 seconds to 0.5 seconds. The prior is barely informative, apart from the input knowledge that three layers can describe accurately the model. By design, the prior has zones where the dataset cannot be sensitive to the data; hence, it makes estimations of uncertainty even more complex.

Results after one iteration

Let us first analyse the results obtained at the first iteration. Since the prior uncertainty is rather large, BEL1D is facing difficulties to retrieve an efficient correlation between the models and the simulated data. However, we are still able to reduce significantly the uncertainty on most of the parameters (*Figure 1*).

From the results at the first iteration, we already see that we are mostly sensitive to the water contents. Then, the relaxation times and, finally, the layers thicknesses are the least sensitive parameters in this configuration. As is also expected, the first layer shows a higher uncertainty reduction at the first iteration for both the water content and relaxation time than the other parameters. This is due to the higher sensitivity of the experiment to this layer.

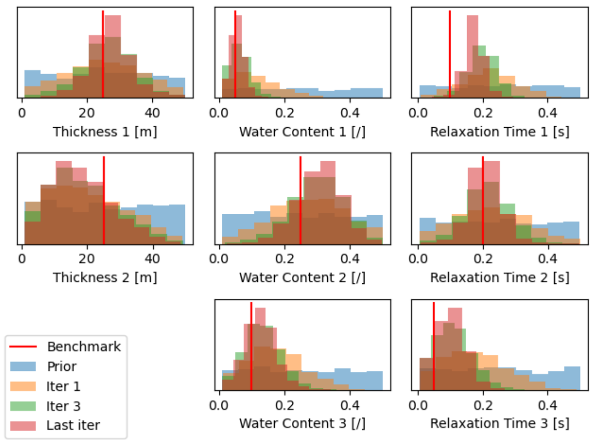

Figure 1: Results obtained from BEL1D with IPR. This graph presents the obtained distributions independently, but the parameter space is explored jointly in BEL1D.

Results after applying IPR

When applying IPR, we are using the information from the previous iterations to better constrain the prior. This leads to a more coherent reduction of uncertainty. Observing the results of the 3rd and 7th (last) iterations (*Figure 1*), we see that the obtained distributions tend towards a more accurate posterior. Nonetheless, some parameters are still lacking sensitivity. This is the case for the relaxation time of the first layer, where the low water content hides this parameter.

If we analyse the correlation between the model's parameters (*Figure 3*), we observe that there is a correlation between the water content of the second layer and its thickness. This result is corroborated by the distributions of the total water content (*Figure 2*). There, we observe that, even though the uncertainty on the water content and the thicknesses remains large, we reduce significantly the total water content.

In this figure, we also propose a comparison between (1) the results from BEL1D with correlation between parameters taken into account and (2) random sampling of the distributions obtained through BEL1D in order to lose the parameters correlations. We see that the correlation that exists inherently between the parameters is crucial to the model estimation.

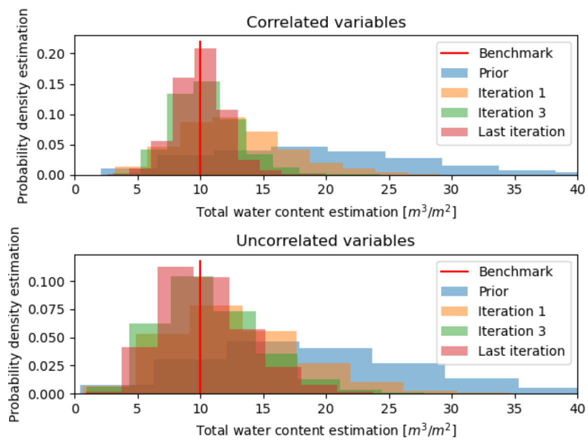

Figure 2: Total water content estimation. The estimation is performed for the correlated case (top – results from BEL1D) and simulated for the uncorrelated case (bottom).

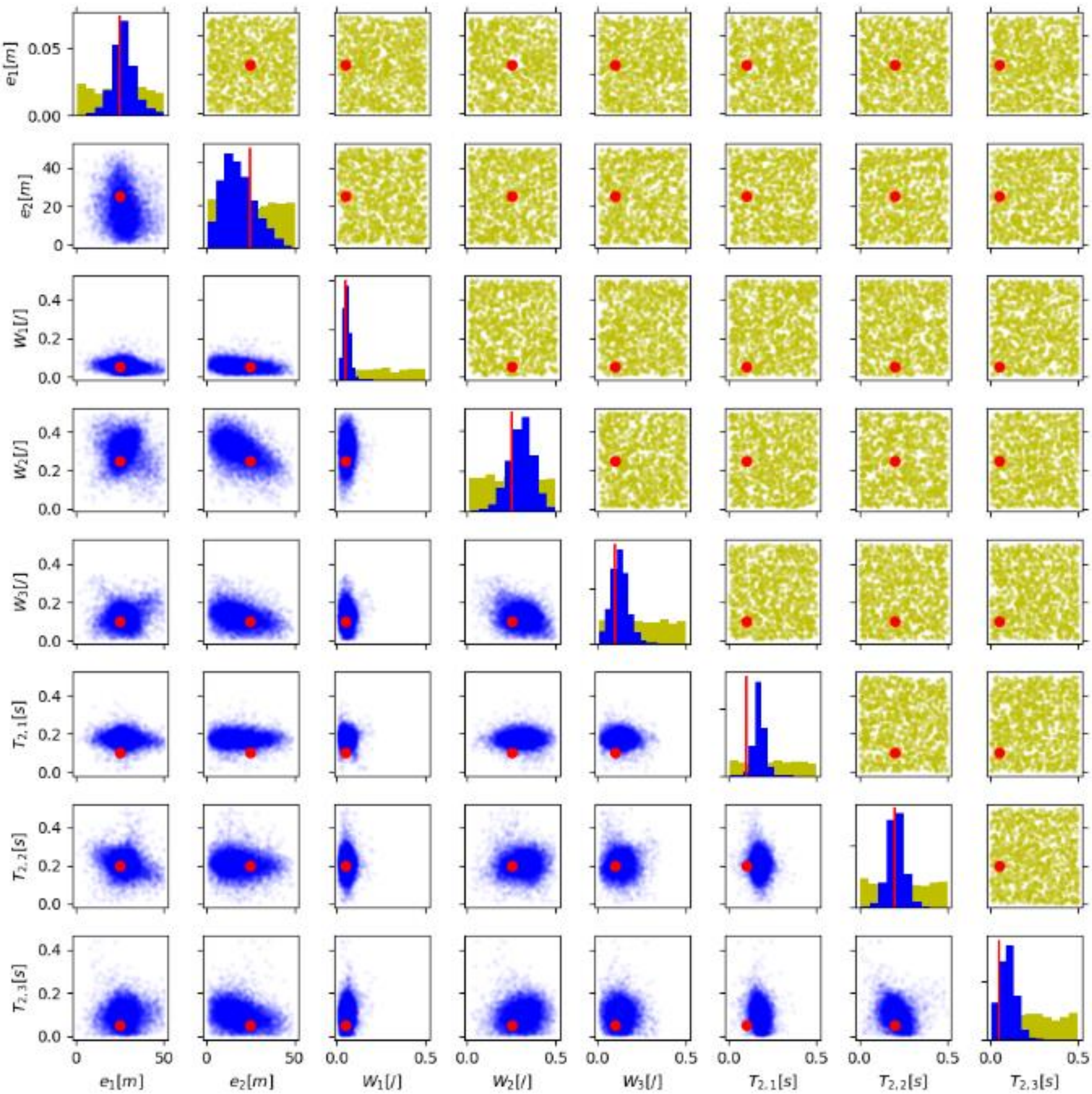

Figure 3: Illustration of the posterior model space after applying BEL1D with IPR. The yellow distributions are the prior and the blue distributions are the posterior after seven iterations of BEL1D.

Finally, we observe that the gain of the latest iterations is marginal. Knowing that the latest iterations are also the ones that are the longest to compute (more models in the informed prior), a user that is interested in a rapid but not especially precise estimation of the uncertainty could use three iterations to gain rapid insight on the uncertainty.

CONCLUSIONS

BEL1D is a new algorithm that can be used to interpret sNMR data in a Bayesian framework. We showed that BEL1D was able to recover reasonable uncertainties, even from large prior model spaces. Moreover, compared to a deterministic scheme with propagation of the uncertainty through linear(ized) inversion (Tarantola & Valette, 1982), BEL1D is able to provide insight on the full behaviour of the posterior, with correlations between the different parameters. This latter is a key aspect as it reduces the uncertainty of joint parameters as well.

Using BEL1D on its own provides a coarse estimation of the uncertainty (especially when dealing with large priors) but can already extract tendencies in the posterior. This iteration is basically free, as the training can be performed prior to any knowledge of the data characteristics. Then, using IPR enables a more precise and accurate estimation of the uncertainty, but this approaches requires CPU time. Using few iterations can already provide a reasonable uncertainty at a reasonable CPU cost. Depending on the degree of precision required, the user could use only those few iterations.

ACKNOWLEDGMENTS

Hadrien MICHEL is a Research Fellow of the *Fonds de la Recherche Scientifique – FNRS*.

REFERENCES

Legchenko, A., & Shushakov, O. 1998. Inversion of NMR data. Geophysics, 63(1), p.75-84.

Legchenko, A., & Valla, P. (2002). A review of the basic principles for proton magnetic resonance sounding measurements. Journal of Applied Geophysics, 50(1–2), p.3–19.

Mueller-Petke, M., & Yaramanci, U. (2010). QT inversion — Comprehensive use of the complete surface NMR data setQT inversion. Geophysics, 75(4), p.WA199–WA209.

Tarantola, A., & Valette, B. (1982). Generalized Nonlinear Inverse Problems Solved Using the Least Squares Criterion. Reviews of Geophysics and Space Physics, 20(2), 219–232.

Scheidt, C., Li, L., & Caers, J. (2018). Quantifying Uncertainty in Subsurface Systems.

Andersen, K. R., Wan, L., Grombacher, D., Lin, T., & Auken, E. (2018). Studies of parameter correlations in surface NMR using the Markov chain Monte Carlo method. Near Surface Geophysics, 16(2), p.206–217.

Michel, H., Nguyen, F., Kremer, T., Elen, A., & Hermans, T. (2020). 1D geological imaging of the subsurface from geophysical data with Bayesian Evidential Learning. Computers and Geosciences, 138, p.104456.

Mariethoz, G., Renard, P., & Caers, J. (2010). Bayesian inverse problem and optimization with iterative spatial resampling. *Water Resources Research*, *46*(11), p.11530.

Müller-Petke, M., Braun, M., Hertrich, M., Costabel, S., & Walbrecker, J. (2016). MRSmatlab — A software tool for processing, modeling, and inversion of magnetic resonance sounding data. Geophysics, 81(4), p.WB9–WB21.