Updating uncertainty in hierarchical subsurface models using geophysical data: synthetic case for cross-borehole GPR

Jorge Lopez-Alvis[1,2], Frédéric Nguyen[1], Thomas Hermans[2]

[1]Applied Geophysics, University of Liège; [2]Department of Geology, Ghent University

ABSTRACT

Considering our incomplete knowledge of subsurface systems we have to deal with model uncertainty. In this study, we work with the parameterization or structure of this model to add uncertainty at different levels and, subsequently, update this uncertainty given some observed data using a Bayesian approach. To deal with the high dimensionality of the data space, we use dimension reduction in combination with regression techniques to approximate the posterior probability distribution. We present the application of parameterization by means of a hierarchical Bayesian model and the use of dimension reduction and non-parametric regression to update uncertainty in a subsurface model using geophysical data.

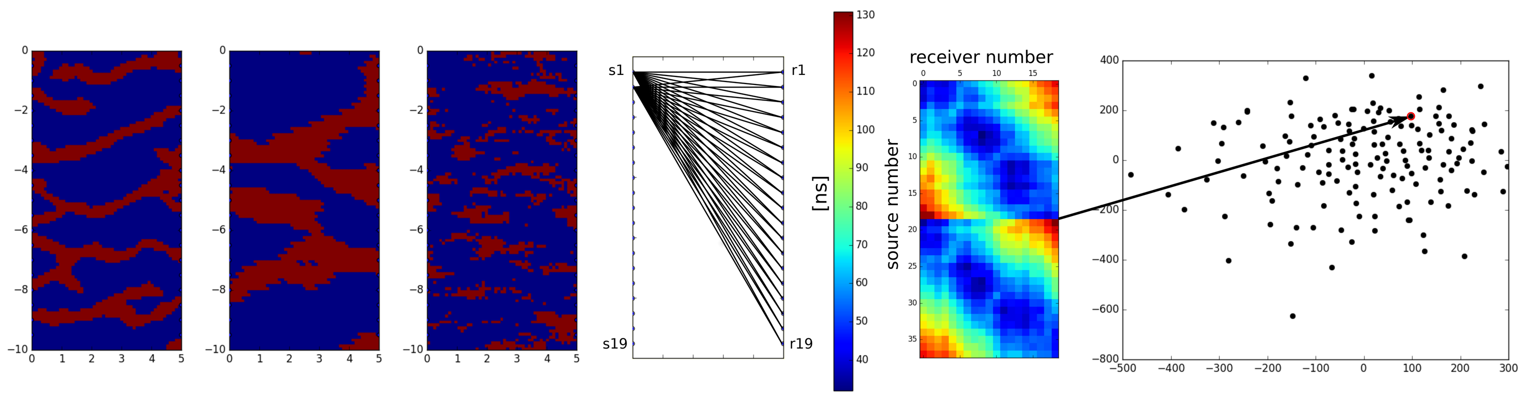
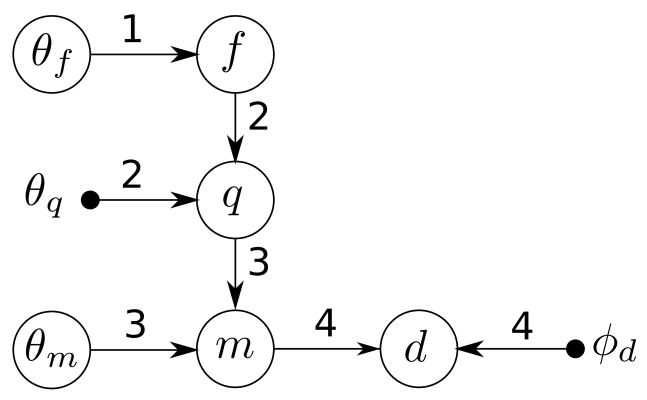


Figure 1. Summary of the workflow, showing as uncertain parameter the geological scenario.

METHODOLOGY

We use a Bayesian approach where we state prior probability distributions of parameters and then obtain their posterior probabilities conditioned on observed data. Uncertain parameters may be considered at different levels using a hierarchical Bayesian model (Wainwright et al., 2014; Feyen & Caers, 2006). The latter can be represented with a graphical model (Fig.2) which explicitly shows conditional dependencies, i.e. the factorization of the joint probability distribution. To obtain samples from this joint probability distribution or its marginals, we sample in topological order (i.e. starting with higher levels).



f - facies

q - porosity

m - geophysical property

$\theta_f = [\alpha, s_k]$

$\theta_q = [\mu_s, \mu_g, \lambda_s, \lambda_g, \sigma_s^2, \sigma_g^2]$

$\theta_m = pe_k$

Figure 2. Graphical representation of hierarchical Bayesian model.

Joint probability distribution

$$p(\theta_f, \theta_q, \theta_m, f, q, m, d) = p(\theta_f)p(f|\theta_f)p(q|f, \theta_q)p(m|q, \theta_m)p(d|m, \phi_d)$$

1. $p(f|\theta_f)$ is sampled through multiple-point geostatistics
2. $p(q|f, \theta_q)$ is given by two Gaussian random fields
3. $p(m|q, \theta_m)$ is given by considering two different petrophysical relations
4. $p(d|m, \phi_d) = MVN(f_g(m), C_d)$

Once a sufficient number of samples is taken, which is dependent on the nature of the problem, we can reduce data dimensionality by means of feature extraction and multi-dimensional scaling. For multi-dimensional scaling we need to define a distance between the samples then an algorithm will project the samples in a lower-dimensional space trying to preserve the original distances (Borg & Groenen, 2005). Finally, in this low-dimensional space we apply kernel smoothing to approximate the posterior probability distribution (Park et al., 2013; Scheidt et al., 2015).

SYNTHETIC CASE

We apply the method to synthetic GPR cross-borehole data. For illustration purposes we only show the effect of one uncertain discrete parameter: the geological scenario. We considered three different geological scenarios, two of them sampled through multiple-point geostatistics and the third through truncated sequential gaussian simulation (Fig.1a).

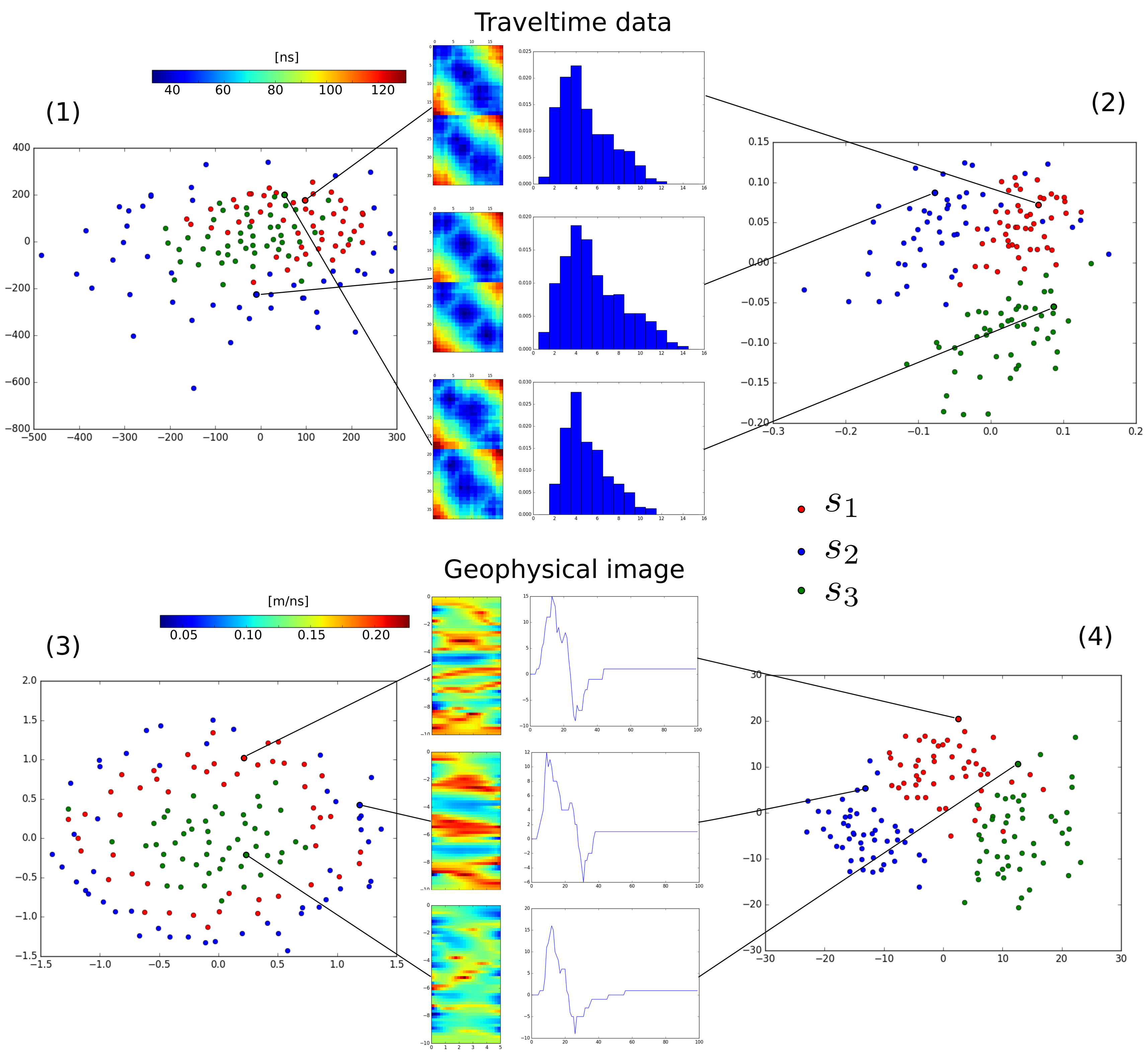


Figure 3. Multi-dimensional scaling applied to different transformations of the data.

The scenarios are built only from two facies, each of which is assigned a constant value for its EM wave velocity. The higher velocity facies presents a degree of continuity in the horizontal direction. In this way samples are transformed into spatial geophysical models and then traveltime data is simulated by means of a forward model and a model error (Fig.1b). To reduce the dimensions of these data, we choose to perform multi-dimensional scaling (Fig.1c) on: (1) traveltime data directly, (2) histograms of traveltime data, (3) "geophysical image", obtained from regularized inversion, and (4) connectivity indicators of the geophysical image (cf. Fig.3). Since we are working with a discrete parameter, we can compute the corresponding posterior distribution by using kernel smoothing separately for each geological scenario (Fig.4).

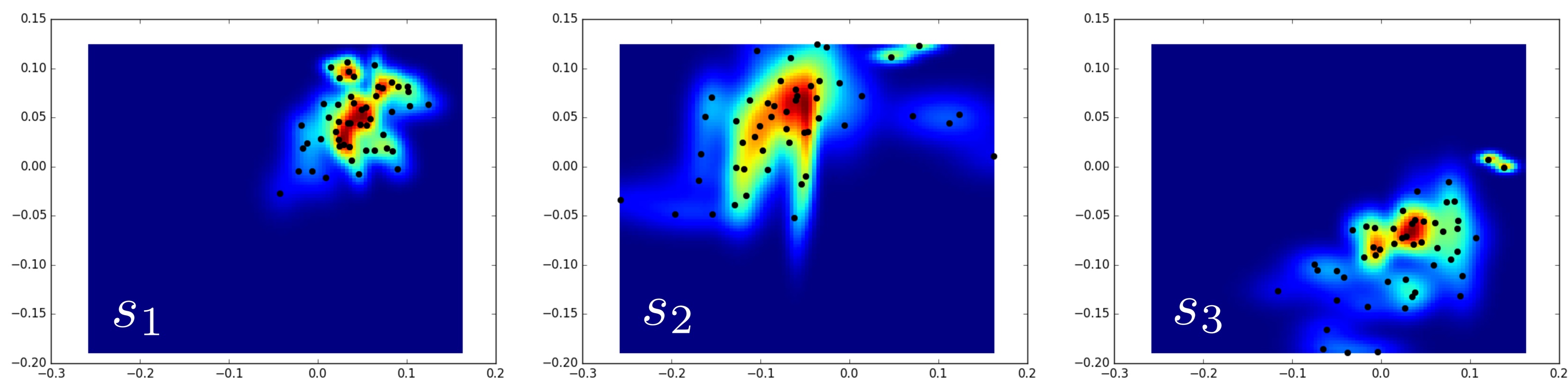
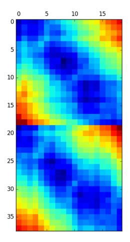
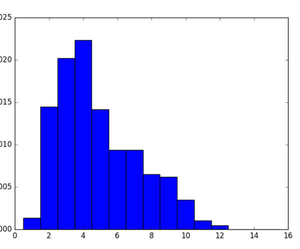


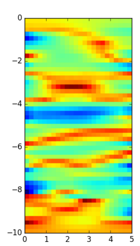
Figure 4. Kernel smoothing applied to low-dimensional projection of the histograms of traveltime data; separate estimations for each geological scenario.



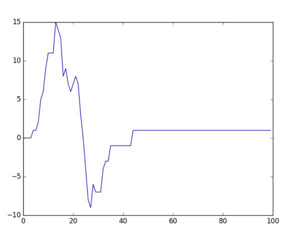
	Classification			Mean updated probability		
	s_1	s_2	s_3	s_1	s_2	s_3
s_1	40	0	10	0.68	0.06	0.26
s_2	8	32	10	0.20	0.60	0.20
s_3	7	0	43	0.19	0.06	0.75



	Classification			Mean updated probability		
	s_1	s_2	s_3	s_1	s_2	s_3
s_1	49	1	0	0.94	0.04	0.02
s_2	5	45	0	0.10	0.90	0.00
s_3	0	0	50	0.01	0.00	0.90



	s_1	s_2	s_3	s_1	s_2	s_3
s_1	21	2	27	0.39	0.16	0.44
s_2	12	25	13	0.30	0.42	0.28
s_3	3	1	46	0.22	0.11	0.66



	s_1	s_2	s_3	s_1	s_2	s_3
s_1	45	0	5	0.84	0.03	0.13
s_2	3	47	0	0.06	0.93	0.01
s_3	2	0	48	0.10	0.00	0.90

Figure 5. Cross-validation applied to each of the previous for cases, showing classification and mean updated probability matrices.

DISCUSSION AND CONCLUSION

We use cross-validation to assess the performance of the method (Fig.5). By taking out one sample (and knowing its true scenario) we compute its posterior probability distribution with the rest of the samples then we classify the geological scenario according to the highest probability value and we also compute the mean posterior probabilities for all the samples corresponding to one true scenario (Hermans et al., 2015). In our synthetic case, the method performs better when we use directly the histograms of traveltime data which is also less computationally demanding than working with the geophysical image. We believe GPR cross-hole traveltime data is very informative in our case due to the presence of horizontally connected facies, as first arrivals correspond to EM waves traveling through higher velocity zones. We conclude that using a hierarchical Bayesian model in combination with an approximate calculation of the posterior probability distribution by application of kernel smoothing and multi-dimensional scaling on features of the traveltime GPR data effectively reduces uncertainty for the geological scenario.

REFERENCES

Borg, I., & Groenen, P. J. F. (2005). Modern multidimensional scaling theory and applications. New York: Springer.

Feyen, L., & Caers, J. (2006). Quantifying geological uncertainty for flow and transport modeling in multi-modal heterogeneous formations. Advances in Water Resources, 29(6), 912–929.

Hermans, T., Nguyen, F., & Caers, J. (2015). Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: Workflow and case study. Water Resources Research, 51(7), 5332–5352.

Park, H., Scheidt, C., Fenwick, D., Boucher, A., & Caers, J. (2013). History matching and uncertainty quantification of facies models with multiple geological interpretations. Computational Geosciences, 17(4), 609–621.

Scheidt, C., Tahmasebi, P., Pontiggia, M., Da Pra, A., & Caers, J. (2015). Updating joint uncertainty in trend and depositional scenario for reservoir exploration and early appraisal. Computational Geosciences, 19(4), 805–820.

Wainwright, H. M., Chen, J., Sassen, D. S., & Hubbard, S. S. (2014). Bayesian hierarchical approach and geophysical data sets for estimation of reactive facies over plume scales. Water Resources Research, 50(6), 4564–4584.

4th Cargèse Summer School: Flow and Transport in Porous and Fractured Media
25 Jun-7 Jul 2018 Cargèse, Corsica (France)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement N° 722028 (ENIGMA ITN).

ITN Enigma