EFFECT OF DIFFERENT STATISTICAL MODELS IN PROBABILISTIC JOINT
ESTIMATION OF POROSITY AND LITHO-FLUID FACIES
FROM ACOUSTIC IMPEDANCE VALUES

M. Aleardi
University of Pisa, Earth Sciences Department, Pisa, Italy

Introduction. The estimation of petrophysical reservoir properties (i.e. porosity, shale content, fluid saturation) and litho-fluid facies around the target area is a common, highly ill-conditioned problem that is often casted into a Bayesian framework. Independently from the inversion approach adopted (analytical or numerical), the correct choice of the underlying statistical model always plays a crucial role in any geophysical Bayesian inversion. For what concerns the reservoir characterization problem, many authors have demonstrated that such statistical model should be able to correctly capture the facies-dependency of petrophysical and/or elastic properties related to the different lithologic and fluid-saturation conditions. In particular, it has been demonstrated that the accounting for such facies-dependency often provides more accurate descriptions of the uncertainties affecting the sought parameters. However, as the author is aware an in-depth discussion and a comparison of the results provided by different statistical models is still lacking for reservoir characterization studies.

Focusing on this peculiar aspect, I use an inversion approach for the joint estimation of porosity and litho-fluid facies from logged and post-stack inverted acoustic impedance (Ip) values. The inversion approach I employ is a modification of the method proposed by Grana (2018) that is adapted to consider Gaussian-mixture and Gaussian distributions, and to jointly invert porosity and logged or inverted Ip values. This work is mainly aimed at analysing and comparing the results provided by three different statistical assumptions about the underlying joint distribution of the petrophysical model relating porosity and Ip values. To this end, I consider a simple Gaussian assumption that neglects the facies dependency of porosity and acoustic impedance values, whereas an analytical Gaussian-mixture distribution and a non-parametric mixture distribution relate each component of the mixture to a specific litho-fluid facies. In particular, the Gaussian or Gaussian-mixture models are often employed in seismic inversions because they allow for an analytical computation of the posterior model and make it possible to easily include additional constraints (i.e. geostatistical constraints) into the inversion kernel. Differently, a non-parametric distribution is not restricted by any statistical assumption about the underlying statistical model, but it impedes an analytical derivation of the posterior model and also complicates the inclusion of additional regularization operators or geostatistical constraints into the inversion framework.

This work focuses the attention on well log data pertaining to a clastic gas-saturated reservoir. All the three considered statistical models are directly estimated from 5 out of 7 available wells drilled trough the reservoir zone. The kernel density technique is used to derive the non-parametric distribution. One of the two remaining wells is here used as blind test to validate the inversion results, whereas the analysis of the maximum-a-posteriori (MAP) solutions, and the coverage ratio are used to more quantitatively assess the final predictions.

The method. The method used in this work is a modification of that proposed by Grana (2018). If we consider a Bayesian setting, the goal of the inversion is to estimate the probability of the petrophysical properties (m) and the litho-fluid-facies (f) given the logged or inverted elastic properties (d). In this context, the sought probability distributions can be computed as:

\[
p(m, f|d) = \frac{p(m, d|f)p(f|d)}{\int p(m, d|f)p(f|d)dm}
\]

where, in our case \(p(m, d|f)\) is the joint distribution of the porosity and Ip values within each facies, which can be estimated from available well log data. The probability \(p(f|d)\) represents the conditional distribution of facies given the observed data that can be computed as:
\[
p(f = f | d) = \frac{p(f = f) \int p(m, d | f = f) dm}{\sum_{n=1}^{K} p(f = n) \int p(m, d | f = n) dm}
\]

where \(K\) is the total number of facies considered: in the following shale, brine sand and gas sand. The key aspect of this inversion approach is the proper choice of the joint distribution \(p(m, d | f)\). To this end many assumptions can be made, for example one can simply neglect the facies dependency of \(m\) and \(d\), thus using a simple unimodal Gaussian distribution. However, note that in this case it is no more possible performing a facies classification. More generally, the assumed statistical model should honour the multimodality of the \(p(m, d | f)\) distribution, and among the many multimodal distributions, the Gaussian-mixture is often adopted. Another possible, but less common approach, is to directly approximate the joint distribution using a non-parametric technique, such as the kernel density estimation (KDE). The numerical inversion method previously described can be applied to both logged impedance values or \(Ip\) values inferred from a post-stack seismic inversion. In the following, both these cases are analysed: first, I use

---

**Fig. 1 - Non-parametric, Gaussian-mixture, and Gaussian joint \(p(m, d | f)\) distributions (a, b, and c respectively).**
actual logged $Ip$ values to infer porosity and facies. Second, I exploit the well log information to compute synthetic seismic traces that, in a first inversion step are converted into $Ip$ values and associated uncertainties that become the input for the following inversion step aimed at estimating porosity and litho-fluid facies. In this synthetic application I employ a convolutional forward operator to derive the post-stack seismic trace, whereas a simple analytical least-square Bayesian inversion is adopted to estimate the $Ip$ values and the associated uncertainty from post-stack traces. In the seismic examples the Chapman-Kolmogorov equation is used to correctly propagate the uncertainty affecting the estimated $Ip$ values into the uncertainties associated to the final porosity and facies predictions. In addition, a 1D Markov Chain prior model is employed to vertically constrain the predicted facies profile.

**Well log data application.** I first describe the three joint $p(m,d|f)$ distributions derived from 5 out of 7 available wells that reached the reservoir (Fig. 1). The first is the non-parametric distribution estimated through the kernel density technique; the second is the analytical Gaussian-mixture that assumes Gaussian distributed porosity and $Ip$ values within each facies. The third is the simple Gaussian distribution. The two multimodal distributions (non-parametric and Gaussian-mixture) derived for the shale seem very similar, whereas their differences are more prominent for the brine and gas sand facies. In particular, the Gaussian-mixture assumption completely masks the multimodality of the $p(m,d|f={\text{brine sand}})$ distribution that is instead correctly modelled by the non-parametric model. This multimodality could be related to sands with different mineralogic or textural characteristics. In Fig. 1c we observe that the Gaussian prior is not able to reliably model the underlying relation linking porosity and $Ip$ values. In other words, the Gaussian model constitutes an oversimplification of the actual, underlying petrophysical model.

Fig. 2 represents the final results. We observe five significant decreases of the acoustic impedance value that mark the main sand layers. The target, gas saturated reservoir is located between 1400-1420 m. In Fig. 2a I show the results for the non-parametric distribution and we observe that the MAP solution for the porosity closely matches the actual porosity values and correctly captures the fine-layered structure of the reservoir. The outcomes of the facies classification show a satisfactory match with the true facies profile derived from borehole information. In particular, note the high probability that a gas saturated layer occurs at the target depth (1400-1420 m). Fig. 2b shows the results obtained for the same well but employing the Gaussian-mixture $p(m,d|f)$ distribution. We observe that the MAP solution for the porosity is now characterized by a poorer match with the logged porosity values than that obtained with the non-parametric $p(m,d|f)$ model. The facies prediction still shows a satisfactory match with the actual facies profile and, more importantly, the main gas saturated layer is still correctly identified. The MAP solutions derived from the Gaussian model (Fig. 2c), seem still able to capture the vertical porosity variability but the oversimplified statistical $p(m,d)$ model translates into higher posterior uncertainties (i.e. wider posterior distributions) compared to the Gaussian-mixture and the non-parametric $p(m,d|f)$ distributions. In other terms, the suboptimal underlying statistical model results in more inaccurate prediction intervals compared to the previous tests. The 90% coverage probability ratios are equal to 0.92, 0.8884 and 0.7788 for the non-parametric, Gaussian-mixture and Gaussian model, respectively. These values confirm that the non-parametric approach outperforms the other two models, although the Gaussian-mixture distribution provides quite accurate predictions. A direct comparison of posterior distributions, of linear correlation coefficients, and the contingency analysis tools have also been employed to more quantitatively assess the final predictions. However, these analyses are not shown here for the lack of space.

**Post-stack data application.** I now extend the inversion tests on post-stack seismic data. For confidentiality reasons, I limit the attention to synthetic data computed on the basis of actual well log information and adopting a 1D convolutional forward modelling with a 45-Hz Ricker wavelet as the source signature and 0.002 s as the sampling interval. To better simulate
Fig. 2 - Results for the well log data example. From left to right: Logged acoustic impedance; Posterior porosity distribution compared with the MAP solution (white line) and the logged porosity values (black lines); Posterior distribution for litho-fluid facies; Actual facies profile; MAP solution for the facies classification where blue, green, and red code shale, brine sand and gas sand, respectively.
Fig. 3 - Results for the post-stack example. From left to right: Comparison between the observed stack trace (black line) and the predicted trace by the post-stack inversion (red line); b) Post-stack inversion results, where the blue line illustrates the true $I_p$ values, the red line represents the MAP solution, whereas the green lines delimit the 95% confidence interval; Posterior porosity distribution compared with the MAP solution (white line) and the logged porosity values (black lines); (only for the two mixture distributions) Posterior distribution for the litho-fluid facies; Actual facies profile derived from well log information; MAP solution for the facies classification.
a field dataset, Gaussian random noise is added to the synthetic stack traces by imposing a signal-to-noise ratio equal to 10. Fig. 3a represents the results obtained for the blind well when the non-parametric \( p(m, d|f) \) distribution is employed. We observe that the predicted seismic trace perfectly matches the observed trace and that the predicted 1D \( Ip \) profile reliably reproduces the vertical variability of the actual impedance values but, more importantly, the 95% confidence interval always encloses the logged \( Ip \). Note that the filtering effect introduced by the convolutional forward operator produces \( Ip \) predictions with lower vertical resolution with respect to the logged \( Ip \) values. As expected, the filtering effect now translates into less accurate MAP porosity predictions with respect to the well log examples. In particular, the additional uncertainties arising from the seismic inversion yield wider posterior distributions, that is we are now less confident on the final porosity predictions with respect to the previous tests at the well log scale. However, notwithstanding the resolution issue, the inversion still recovers the significant porosity increase occurring at the sand layers. The estimated facies profile still shows satisfactory predictions, although the filtering effect results in final predictions with lower vertical resolution with respect to the previous examples on well logs. Fig. 3b represents the results achieved by the Gaussian-mixture model. By comparing Figs. 3a and 3b we observe that the non-parametric distribution again provides superior porosity estimations and facies profile than the analytical \( p(m, d|f) \). In particular, only the main gas-saturated layer located at 940 ms is correctly identified by the Gaussian-mixture model, while the other sand layers are erroneously misclassified as shaly intervals. The 90% coverage probability values confirm these qualitative descriptions being equal to 0.7687 and 0.6331 for the non-parametric and Gaussian-mixture models, respectively. As expected the Gaussian model (Fig. 3c) achieves less accurate porosity estimations, higher uncertainties, and less reliable prediction intervals resulting in a coverage probability equal to 0.6026; a value lower than those yielded by the Gaussian-mixture and the non-parametric \( p(m, d|f) \) distributions.

**Conclusions.** This work demonstrated that the correct modelling of the facies dependency of porosity and \( Ip \) values could be crucial to achieve accurate estimations and reliable prediction intervals. Both the Gaussian-mixture and the non-parametric distributions provide satisfactory results, although the non-parametric statistical model usually achieves superior porosity estimations and litho-fluid facies classifications. Differently, the Gaussian assumption demonstrated to be a too oversimplified model that, totally neglecting the facies-dependency of the porosity and \( Ip \) values, provides less accurate prediction intervals, poorer match with actual porosity profiles, and higher uncertainties with respect to the other two statistical models. In the seismic experiments, as expected, the filtering effect introduced by the convolutional operator and the additional uncertainties arising from the post-stack seismic inversion, provided less accurate porosity estimations characterized by wider posterior uncertainties and predicted porosity and facies profiles affected by lower vertical resolution with respect to the examples at the well log scale.

The choice of the underlying statistical model is usually complicated because it is not only case-dependent but should constitute a reasonable compromise between the accuracy of the final predictions, the stability of the inversion procedure, the total computational effort, and the actual fitting between the underlying and the considered petrophysical models.

**References**
Grana, D. (2018). Joint facies and reservoir properties inversion. Geophysics, 83(3), M15-M24.