# Rock Physics Driven Joint Inversion to Facies and Reservoir Properties

**Michel Kemper**
*Ikon Science Ltd*
*Suite 317 Level 25, Tower 2,*
*Etiqa Twins, 11 Jalan Pinang,*
*50450 Kuala Lumpur,*
*Malaysia*
*mkemper@ikonscience.com*

**James Gunning**
*CSIRO*
*Bayview Avenue*
*Clayton*
*Melbourne VIC 3168*
*Australia*
*james.gunning@csiro.au*

## SUMMARY

In this paper we will first review the industry-standard continuous simultaneous inversion methods and point out some shortcomings. We will then introduce our new joint categorical/continuous simultaneous inversion technology, which reformulates the problem to address these issues. It casts the inversion as a Bayesian problem, which needs to be solved iteratively, as the inversion to categorical and continuous properties cannot be written down in closed form. We present some examples and conclude that the new inversion is able to overcome the shortcomings mentioned.

**Key words:** seismic inversion, Bayes' theorem

## METHOD AND RESULTS

### 1. Simultaneous inversion to date

***A) Deterministic method***

We start in this case with the Fatti (1994) equation (other approximations to the Zoeppritz equation are possible also):

$$R_{pp}(\theta) = aR_{AI} + bR_{SI} + cR\rho \qquad [1]$$

Where:

$$R_{AI} = \Delta Vp/(2Vp) + \Delta\rho/(2\rho)$$
$$R_{SI} = \Delta Vs/(2Vs) + \Delta\rho/(2\rho)$$
$$R_\rho = \Delta\rho/(2\rho)$$
$$a = 1 + \tan^2\theta$$
$$b = -8\ K \sin^2\theta$$
$$c = 4\ K \sin^2\theta - \tan^2\theta$$
$$[\ K = (Vs_{avg}/Vp_{avg})^2\ ]$$

We can turn reflectivity $R_{pp}(\theta)$ in [1] into synthetic seismic $S_{pp}(\theta)$ by convolving with wavelet $W(\theta)$:

$$S_{pp}(\theta) = aW(\theta)R_{AI} + b\ W(\theta)R_{SI} + c\ W(\theta)R_\rho \qquad [2]$$

Subsequently we can use the small contrast approximation $R_{AI} = (R_{AI2} - R_{AI1}) / (R_{AI2} + R_{AI1}) \approx ½\ \Delta AI / AI = ½\ \Delta Ln(AI)$ to rewrite [2] as

$$S_{pp}(\theta) = a/2\ W(\theta)\ \Delta Ln(AI) + b/2\ W(\theta)\ \Delta Ln(SI) + c/2\ W(\theta)\ \Delta Ln(\rho) \qquad [3]$$

Formula [3] is the essence of all *continuous* simultaneous inversion methodologies: plug in a starting model of AI, SI and ρ, take the natural logarithm, difference, for each partial angle stack convolve with the corresponding wavelet $W(\theta)$, scale (using a/2, b/2 and c/2) and for each partial stack incidence angle θ compare the synthetic seismic $S_{pp}(\theta)$ so obtained with the real seismic $S_{real}(\theta)$ which we assume to be true amplitude (with calibrated wavelets available from well ties). Use some optimization apparatus (e.g. the conjugate gradient method, or least squares optimization) to iteratively change AI, SI and ρ until the difference between $S_{pp}(\theta)$ and $S_{real}(\theta)$ is minimized for all partial stack incidence angles θ, typically in some least squares sense (solving for AI, SI and ρ directly is possible also; not discussed here).

Because of the convolution by $W(\theta)$, [3] cannot be used sample-by-sample. Instead it is typically used trace-by-trace (this can be extended to multi-trace; not discussed here). Therefore **Ln(AI)**, **Ln(SI)** and **Ln(ρ)** can be seen as column vectors which we stack on top of one another to obtain **Ln(Z)**. The difference operation Δ, applied to each of Ln(AI), Ln(SI) and Ln(ρ), can be expressed as an (almost diagonal) matrix **D**. And lastly the convolution can be expressed by a (banded) matrix **W** (typically a different **W** matrix per partial angle stack). The product of **W** and **D** we can call system matrix **X**, into which the scaling parameters a, b and c (and the ½ factor) can be subsumed. So [3] in block matrix form reduces to

$$\mathbf{S = X \bullet Ln(Z)} \qquad [4]$$

It is useful to sketch (Figure 1) the sizes of the various block-vectors/block-matrices in the case where we invert for 3 impedances (i.e. AI, SI and ρ) given 4 partial stacks at, say, 5°, 15°, 25° and 35° incidence angle, over a gate of 100 samples:

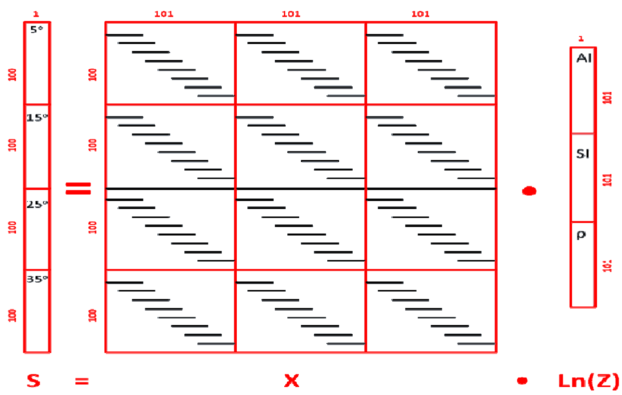

**Figure 1: Matrix sizes (in red) for 4 partial angle stacks, 3 impedances, and a time gate of 100 samples. The black lines show that the blocks within matrix X are banded; in the top right and bottom left are many zeroes (especially if the wavelet is small and/or the gate is long), which speeds up calculation**.

**Note-1:** Each block of the system block matrix **X** contains an almost diagonal difference matrix **D** multiplied by a banded wavelet matrix **W**, making each block banded and thus quite sparse. These blocks are not identical, as vertically the convolution matrix changes (different wavelets for different incidence angles θ), and horizontally the scaling factors change from a/2 to b/2 to c/2 respectively.

**Note-2:** The impedance traces are 1 longer than the synthetic seismic traces, as we need 2 impedances to calculate 1 reflectivity!

So the optimization typically consists of minimizing an objective function $\|\mathbf{Sreal} - \mathbf{X} \bullet \mathbf{Ln(Z)}\|^2$, by changing **Ln(Z)**. This in itself is usually wildly unstable, so normally one adds a regularization term to ensure the impedances do not drift away too much from the initial impedances **Zo**. The total objective function to be minimized is then something like $\|\mathbf{Sreal} - \mathbf{X} \bullet \mathbf{Ln(Z)}\|^2 + \mu\|\mathbf{Ln(Z)} - \mathbf{Ln(Zo)}\|^2$, where μ, the so-called model weight, should be as small as possible to ensure the inversion is driven mostly by the data (i.e. the seismic **Sreal**), and as little as possible by the initial model (**Zo**). Note that the objective function is quadratic, which makes optimization relatively easy.

It is well known that the methodology described in some detail above gives results that are not credible when compared to the impedances from wells. A major reason for this is that the band-limited convolution destroys high and low frequency information, making the inverse problem very underdetermined, and the regularization term introduced above is too simplistic a cure for this. In other words, in the deterministic inversion scheme introduced so far we need to incorporate more sophisticated regularization information for it to give useful results. We'll come back to this in section 2.

***B) Statistical method***
Clearly simultaneous inversion can be seen as a statistical problem, given the noise component of the seismic signal. The most common way to proceed is to cast it as a Bayesian problem, in which the prior information, if well chosen, will provide sufficient regularization. Bayes' theorem in this case can be written as:

$$\pi(\mathbf{Z}|\mathbf{Sreal}) \approx L(\mathbf{S}_{\mathbf{real}}|\mathbf{Z})\ p(\mathbf{Z}) \qquad [5]$$

Where π is the posterior distribution, L the likelihood function and p the prior distribution. Again **Z** represents the impedances (say AI, SI & ρ in case of Fatti)

It is customary (Buland and Omre, 2003) to represent the distributions as being (multi)normal; this is often very reasonable, and makes the mathematics a lot more tractable.

The prior distribution can be obtained from a depth trend of, say, AI and cross-plots between AI vs. SI and AI vs. ρ (all derived from well data, and all with an assessment of uncertainty). These trend fits can be expressed as a multi-normal prior distribution of form

$$p(\mathbf{Z}) \approx \exp\{-½(\mathbf{Z} - \mathbf{Zo})^T C_p^{-1} (\mathbf{Z} - \mathbf{Zo})\} / |Cp|^{½} \qquad [6]$$

Where $C_p$ is the covariance matrix describing the variance of and the correlation between the impedances

The likelihood function can be expressed as

$$L(\mathbf{Sreal}|\mathbf{Z}) \approx \exp\{-½(\mathbf{Sreal} - F(\mathbf{Z}))^T C_d^{-1} (\mathbf{S}_{\mathbf{real}} - F(\mathbf{Z}))\} / |Cd|^{½} \qquad [7]$$

Where F(**Z**) is the function to derive synthetic seismic from the impedances **Z**, as described earlier, and $C_d$ is the covariance matrix representing the 'effective' seismic noise.

The (un-scaled) posterior distribution can be derived using [5], from which we can derive the maximum a-posteriori (MAP) model of Z (kind of like a P50 estimate), or we can use MCMC sampling to obtain marginal distributions of interest (see also Gunning and Glinsky, 2004).

**2. Why not optimal?**
The inversion schemes described above have a number of shortcomings. Firstly, the additional regularization information needed to ensure reasonable comparisons between the deterministic inversion results and well impedance profiles at best contains only global or 'pooled' rock physics trends (i.e. one rock physics trend for all facies) and at worst no rock physics at all. As an example, one 'trick' often applied is to approximate Ln(SI) as linear in Ln(AI): $\mathrm{Ln(SI)} = \alpha_{SI}\,\mathrm{Ln(AI)} + \beta_{SI} + \delta\mathrm{Ln(SI)}$ (same for Ln(ρ)). This changes [3] to

$$S_{pp}(\theta) = a'/2\ W(\theta)\ \Delta\mathrm{Ln(AI)}\ + b/2\ W(\theta)\ \Delta\delta\mathrm{Ln(SI)}\ + c/2\ W(\theta)\ \Delta\delta\mathrm{Ln}(\rho) \qquad [8]$$

Where $a' = a + \alpha_{SI}\, b + \alpha_\rho\, c$

So now we invert to Ln(AI), **δ**Ln(SI) and **δ**Ln(ρ), i.e. for SI and ρ we invert for the *deviation* from the global linearizations. So how accurate are these linearizations? In Figure 2 you see that the relationship between Ln(AI) and Ln(SI) can be reasonably linear, but that the same cannot always be said for Ln(AI) and Ln(ρ)!

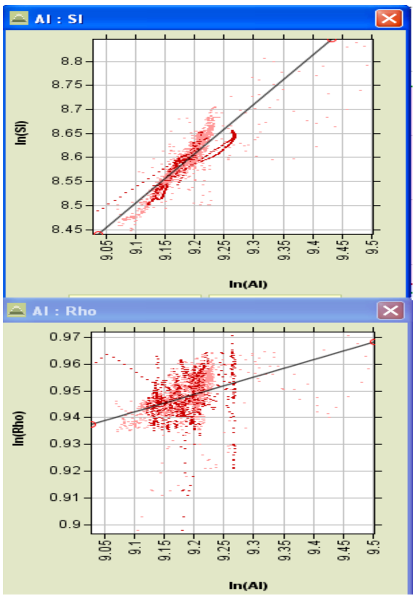

**Figure 2: Typical Ln(AI) vs. Ln(SI) (top) and Ln(AI) vs. Ln(ρ) (bottom) cross-plots over all facies excluding hydrocarbon bearing intervals. Whereas the Ln(AI) vs. Ln(SI) cross-plot is reasonably linear, the Ln(AI) vs. Ln(ρ) relation is more complex**.

Secondly, most energy in the seismic $\mathbf{S}_{\mathbf{real}}$ comes from interfaces between different facies, but the inversion methodologies described so far are entirely continuous (whereas of course facies are categorical). This means that away from facies interfaces, where the energy wanes, the impedance tends to the pooled rock physics trend value (or background value), as shown in the figure below.

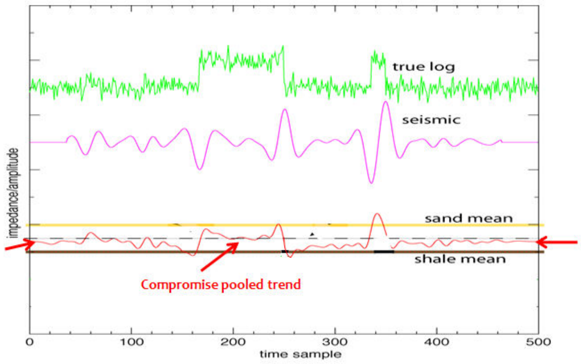


**Figure 3: An impedance model (top), the corresponding synthetic (middle) and the inverted impedance trace (bottom). Note how away from seismic energy (indicated by red arrows) the impedance tends to a value not representative of any rock.**

These two shortcomings lead to the following realizations:

1. We need to ensure the regularization is based on un-pooled, one-per-facies rock physics trends, as a pooled rock physics trend (one rock physics trend to describe all facies) is demonstrably unrealistic/inaccurate
2. Therefore we need to jointly invert for facies **and** for impedances per facies!

**3. Joint categorical/continuous simultaneous inversion**

For the avoidance of doubt, 'joint' here means that we invert both for facies (a categorical variable) and impedances (continuous variables), and 'simultaneous' means that we invert for more than one impedance in one go from multiple partial stacks.

This joint inversion we desire is mathematically demanding, as it cannot be written down in closed form (such as [4] above), nor can it be solved using standard optimization apparatus. We use an iterative method where we first invert for impedances (given a starting facies model), then given these impedances we invert for facies labels, then given these facies labels we re-invert for impedances and so forth, until a suitable solution has been obtained. This method is called the Expectation-Maximization algorithm, or E-M for short. In the M step we estimate the impedances given the expected facies labels, and in the E step we estimate the expected facies labels given the impedances.

Again using Bayes' theorem, the joint-model posterior distribution equivalent of [5] is now more complex

$$\pi(\mathbf{Z},F|\mathbf{S}_{\mathbf{real}}) \approx L(\mathbf{S}_{\mathbf{real}}|\mathbf{Z})\ p(\mathbf{Z}|F)\ p(F) \qquad [9]$$

The likelihood function $L(\mathbf{S}_{\mathbf{real}}|\mathbf{Z})$ is the same as [7], and the prior distribution $p(\mathbf{Z}|F)$ is very similar to [6], the only difference being that the prior mean **Zo** depends on the facies label F:

$$p(\mathbf{Z}|F) = \exp\{-\tfrac{1}{2}(\mathbf{Z} - \mathbf{Zo}(F))^T C_p(F)^{-1} (\mathbf{Z} - \mathbf{Zo}(F))\} / |Cp(F)|^{\frac{1}{2}} \qquad [10]$$

This facies-dependent prior distribution can be obtained from depth trends of, say, AI and cross-plots between AI vs. SI and AI vs. ρ. The difference with section 1 is that now we develop these depth trends and cross-plots *per facies*, each one complete with an assessment of uncertainty.

That leaves p(F), the facies prior distribution. For this we use a discrete Markov Random Field. Expressed simply, any 3D seismic lattice consists of many pixels. We define a set of edges connecting the direct neighbours of each of the pixels in the inversion window. Each pair of connecting neighbours forms a so-called clique, and so each pixel belongs to six cliques (except at edges, corners). The probability of a configuration F of the whole lattice is then defined by the sum of potential energies over all the cliques in what is usually called a Gibbs distribution:

$$p(F) \approx \exp(-\textstyle\sum V_c(F_c)) \qquad [11]$$

where $V_c$ represents the "potential energy" of the set of labels $F_c$ seen by each clique c.

Comparing the central pixel with each of the 6 individual neighbour we can write the potential energy as $V_c = \tfrac{1}{2}\,\beta\, I(F_{centre},F_{neighbour})$ where the discrete indicator function I is 1 if labels $F_{centre}$ and $F_{neighbour}$ are different and is 0 if they are the same, and β is a positive smoothness/continuity parameter. So we can rewrite [11] as ...

$$p(F) \approx \exp(-\textstyle\sum\sum \tfrac{1}{2}\,\beta\, I(F_1,F_2)) \qquad [12]$$

where the first summation is over all pixels and the second summation is over the 6 cliques per centre-pixel

... and thus we penalize (reduce) the probability p(F) if 2 labels are different.

We have implemented different β's for horizontal and vertical smoothness, as geologically horizontal continuity is typically larger than vertical continuity. However, as geology is seldom perfectly horizontal, we use the concept of stratigraphic age to determine neighbours of the same age (figure 4), which may not be the neighbouring pixel at the same time index.

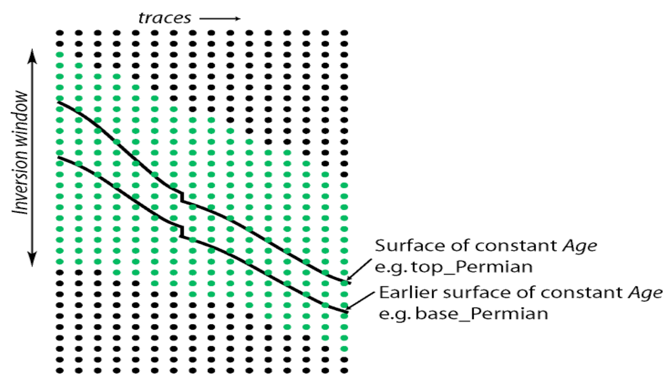

**Figure 4: A stratigraphic age 'field' assists in determining neighbours of the same age.**

In the E-step, the facies distribution needs to be updated based on the current estimate of the impedances (and the per facies rock physics model trends). I.e. we also need to optimize $p(F|\mathbf{Z}) \approx p(\mathbf{Z}|F)\ p(F)$. For this graph-cutting and so-called loopy Bayesian Belief Propagation techniques are used, which are beyond the scope of this paper.

Presently the tool inverts only for the maximum a-posteriori probability (MAP) estimate (i.e. no sampling, such as MCMC). As all 3 probabilities in [9] are of exponential form, determining where the maximum is achieved is equivalent to determining where the minimum of the sum of the exponents is achieved, which simplifies the algebra.

**4) Examples**

The first example shows a trace from a wedge model with a slow ca. 2450 m/s sand encased in fast ca. 2800 m/s shale (black model in Figure 5). It is clear the final iteration (number 7, grey dashed) of the new Joint Categorical/Continuous inversion is superior to the continuous inversion result (number 0, red) as described in section 1, which is the industry standard at present.

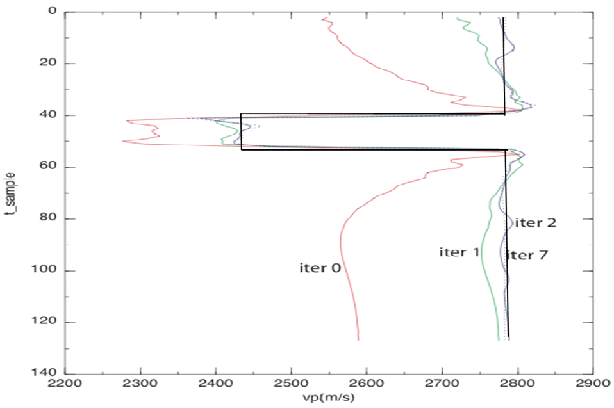


**Figure 5: A sand embedded in a shale. Iterations 0, 1, 2 and 7 (the last one) are labelled. Note that iteration 0 is the normal continuous inversion result, and shows familiar problems: the sand value 'overshoots' (2300 m/s instead of the real ca. 2450 m/s sand value), a dip to higher velocities in the middle of the sand, and away from the shale/sand and sand/shale interfaces the impedance value tends to a neither-sand-nor-shale value.**

In the second example we have applied the new Joint Categorical/Continuous inversion to the 3D seismic of the Stybarrow field, offshore Western Australia, as shown in Figure 6.

The inversion to AI was performed twice, once with horizontal β values equal to 0 (no horizontal continuity), and once equal to 0.5. We see that imposing some horizontal continuity in this case (bottom of Figure 6) improves both the facies inversion (less salt and pepper effect, and therefore more geological) and the density inversion (a less dappled ρ image). The velocity image is less sensitive to the β values, which is understandable, as Vp has most effect on the impedance.

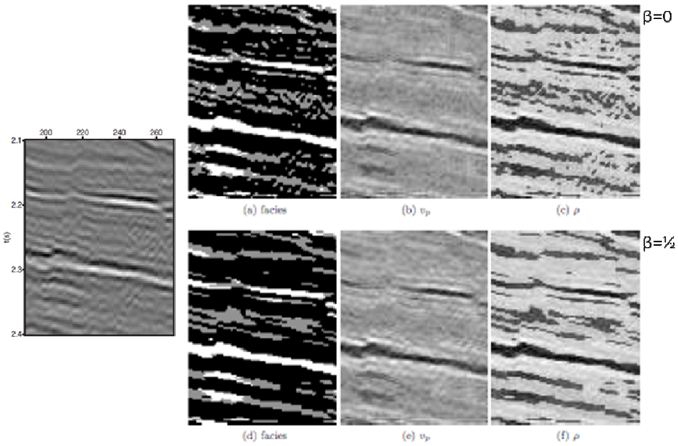


**Figure 6: The normal incidence section (left), inverted results (facies labels, Vp and ρ) with horizontal β values equal to 0 (top), and same inverted results with horizontal β values equal to 0.5 (bottom)**

## CONCLUSIONS

In-depth analysis of the present industry standard continuous simultaneous inversion method highlighted some shortcomings, which require the following remedies:

1. Prior constraints based on un-pooled (i.e. per facies) rock physics trends need to be incorporated in the ideal inversion algorithm, as rock type is crucial.
2. This means the ideal inversion needs to invert not only to continuous impedances, but also jointly to categorical facies!

In this paper we have introduced a new joint categorical/continuous simultaneous inversion method which implements these two remedies, and which does not show the unsightly artefacts of standard methods such as impedances tending to a non-geological value away from seismic energy.

*En passant* the new technique has some other attractive features, such as the ability to hardwire the facies labels at wells. This is quite a unique way to use well data *directly* in seismic inversion (normally it is only used *in*directly, and in inversion QC).

## REFERENCES

Buland, A. and Omre, H., 2003, Bayesian linearized AVO inversion, Geophysics, 68, 185–198.

Fatti, J.L., Smith, G.C., Vail, P.J., Straus, P.J. and Levitt, P.R., 1994, Detection of gas in sandstone reservoir using AVO analysis: a 3D seismic case history using the Geostack technique, Geophysics, 59, 1362-1376.

Gunning, J. and Glinsky, M.E. ,2004, Delivery: an open–source model–based Bayesian seismic inversion program. Computers and Geo-sciences, 30, 619−636.

Gunning, J. and Kemper, M., 2011, Some newer algorithms for joint categorical and continuous inversion problems. Inversion Workshop, 73rd annual conference, EAGE, Vienna.