Progress and Prospects for Wave Equation MVA

William W. Symes

TRIP Annual Review, January 2007

Agenda

- MVA, Nonlinear Inversion, and Extended Modeling
- MVA with two different extensions: the coherent noise issue
  - DSVA-NMO and Land Data (Verm & S., SEG 06)
  - DSVA-DSR and migration noise (Khoury et al., SEG 06)
- Projects:
  - Kirchhoff-based DS via Eikonal Solvers
  - RTM-based DS
  - Nonlinear Inversion and MVA
  - Other issues: attenuation, sources

MVA, Nonlinear Inversion, and Extended Modeling

Extended Modeling

Modeling and Inversion:

- M = model space, D = data space
- $F : M \rightarrow D$ = modeling operator, aka forward map, aka simulator,...
- Inversion: given $d \in D$, find $m \in M$ so that $F[m] \simeq d$

Extended modeling and inversion:

- $\bar{M}$ = *extended* model space
- $\chi : M \rightarrow \bar{M}$ = extension operator, 1-1, "$\bar{M}$ contains M". $\chi[M] \subset \bar{M}$ = "physical models".
- $\bar{F} : \bar{M} \rightarrow D$ = extended modeling operator: $F[m] = \bar{F}[\chi[m]]$.
- Extended inversion: given $d \in D$, find $\bar{m} \in \bar{M}$ so that $\bar{F}[\bar{m}] \simeq d$ - physically meaningful only if $\bar{m} = \chi[m]$.

Extended Inversion

Since extended model space has more degrees of freedom, ambiguity is more likely.

Same old, same old: look for $\bar{m}$ so that (1) $\bar{m}$ is in range of χ, i.e. $\bar{m} = \chi[m]$ for some m, and (2) $\bar{F}[\bar{m}] \simeq d$. Then m is a solution of original inverse problem - nothing has been gained.

Possibility of genuinely different problem: invent a function $\gamma : \bar{M} \rightarrow \mathbf{R}^+$ so that $\gamma[\bar{m}] = 0 \Leftrightarrow \bar{m} = \chi[m]$ for some $m \in M$. Then inverse problem becomes *optimization problem*:

$$\min_{\bar{m}} \gamma[\bar{m}] \text{ subj } \bar{F}[\bar{m}] \simeq d$$

Many such functions - some of these optimization problems may be qualitatively better (closer to quadratic) than others.

Familiar Extensions

M = positive functions on X = subsurface domain, $\bar{M}$ = functions on $X \times H$, H = additional degrees of freedom.

Common Acquisition Parameter: Indep. models for each offset or planewave or... Expl: common offset extension. H = {range of surface offsets}, $\bar{F}$ = independent modeling for each offset, χ = repeat same model at each offset ("flat gathers").

Space/Time Shift: H = X, $\bar{M}$ = distributions on $X \times X$, interpreted as kernels of SPD operators. $\bar{F}$ = "action at distance" modeling operators obtained by replacing physical positive definite fields (density, Hooke tensor,...) with SPD operators. χ = operators are multiplcation by corresponding physical fields (kernels are concentrated on diagonal).

Typical: limit degrees of freedom in op. to same number as data: eg. $\bar{\rho}(\mathbf{x}_1, \mathbf{x}_2) = \tilde{\rho}(x_1, y_1, x_2, y_2, (z_1 + z_2)/2)\delta(z_1 - z_2)$ - Claerbout 1971. Other possibilities: Sava-Fomel SEG 2005.

MVA as Extended Inversion

MVA based on linearized (tangent) modeling of nonlinear physics $\mathcal{F} : \mathcal{M} \to D$:

- M = tangent space of nonlinear model = pairs $(m_0, \delta m) \in \mathcal{M} \times \mathcal{M}$;
- $F = D\mathcal{F}$ = Born modeling;
- For inversion choose "$\simeq$" to mean "close" (in natural norm on D);
- For migration choose "$\simeq$" to mean "has right phases but maybe wrong amplitudes";
- $\bar{M}$ = tangent space of nonlinear extension = pairs $(\bar{m}_0, \delta\bar{m}) \in \bar{\mathcal{M}} \times \bar{\mathcal{M}}$ - might as well limit to linearization about physical models, i.e. $\bar{m}_0 = \chi[m_0]$;
- $\bar{F} = D\bar{\mathcal{F}}$ = Born extended modeling;
- χ typically linear, so its linearization is $D\chi[(m_0, \delta m)] = (\chi[m_0], \chi[\delta m])$.

Nature of Linearized Extended Modeling

Fundamental results about $\bar{F}$:

Nolan & S. 1996 (SEG), Stolk & S. 2004 (*Geophys*, also Brandsberg-Dahl & de-Hoop 2004 (*Geophys*.): Common Acquisition Parameter extended inverse problem not uniquely solvable in general: in presence of multipathing, can find multiple solutions for same m_0 (eg. flat and non-flat events in gathers for same velocity - "kinematic PSDM artifacts").

deHoop & Stolk 2001, Stolk, deHoop & S. 2005: solution of Space/Time Shift extended inverse problem with full 3D data uniquely determined by m_0 (some provisos). m_0 (incl. v) kinematically correct $\Leftrightarrow$ only solutions in range of χ.

Lesson: the superiority of "wave equation migration" lies in formulation (use of Space/Time Shift extension) and adequate data, not computational method.

Semblance

Upshot: only Space/Time Shift extension generally obeys

> Semblance Principle: correct $m_0 \Leftrightarrow$ solutions $(\chi[m_0], \delta\bar{m})$ of extended linearized inverse problem all in range of χ.

Leads to various optimization formulations: choose annihilator $W : \bar{M} \rightarrow$ ... so that $\bar{m} = \chi[m] \Leftrightarrow W[m] = 0$. Then define $\gamma[m] = \frac{1}{2}\|Wm\|^2$, solve

$$\mathrm{min}_{\bar{m}}\gamma[\bar{m}] \text{ subj } \bar{F}[\bar{m}] \simeq d$$

MVA setting: assuming m_0 determines extended inversion, eliminate $\delta\bar{m}$: $\bar{F}[(\chi[m_0], \delta\bar{m})] = d \Leftrightarrow \delta\bar{m} = \bar{G}[m_0, d]$. Optimization formulation of MVA:

$$\mathrm{min}_{m_0} J[m_0, d] = \frac{1}{2}\|W\bar{G}[m_0, d]\|^2$$

Stolk & S. 2003: J smooth $\Leftrightarrow W$ (pseudo)differential$\Rightarrow$ DSVA

A (Very?) Special Case

Model: M = midpoint-dependent $v(t_0, \mathbf{x}_m) \times$ midpoint-dependent reflectivity ("stacked section"), F = convolutional model (Born approximation!)

Extended Model - special case of CAP: $\bar{M}$ = midpoint dependent v $\times$ ensemble of CMP gathers, χ = repeat zero-offset trace, $\bar{F}$ = convolutional model (INMO), $\bar{G}$ = NMO

Annihilator W = offset divided differences within CMP gathers, leads to DSVA-NMO - most recent implementation released to sponsors 1/06.

S. 1999, 2001: All stationary points of the DSVA-NMO J are global minima.

Nonlinear MVA = Inversion

Big Lesson of MVA context: $\bar{F}$ not smooth as a function of $\bar{m}$, so $\bar{G}$ not smooth as function of m_0 (cycle skipping, loss of derivatve - Stolk 2000) BUT can still find smooth functions of $\bar{G}[m_0, d]$.

Substitute for "velocity control" m_0 in nonlinear context: full bandwidth data. Rationale: appears likely that will determine $\bar{m}$. Evidence incomplete but suggestive:

- numerics - OLS nonlinear inversion, eg. Santosa-Sacks 1987, Bunks et al. 1995, Shin et al. 2001.
- mathematics - Ramm's uniqueness theorem (Ramm, 1989)

Major need: theoretical framework for nonlinear extended inverse problems, both types of extensions. For example: what model classes $\bar{M}$ are (a) computationally tractable and (b) have uniquely solvable extended inverse problems?

MVA with two different extensions: the coherent noise issue

DSVA-NMO and Land Data (Richard Verm & William W. Symes, SEG 06)

Thanks: Geokinetics

(see SEG presentation)

DSVA-DSR and migration noise (Alexandre Khoury, William W. Symes, Paul Williamson and Peng Shen, SEG 06)

Thanks: Total E&P USA

DSVA-DSR

Based on Space/Time Shift extension, Claerbout's special case (diagonal in z).

This time $\bar{G}$ is DSR ("shot-geophone") migration. Similar results for shot profile: P. Shen thesis 2004, SEG 2003 and 2005. Annihilator W = multiply by offset h. Presumption: if v kinematically consistent with data $\Leftrightarrow$ p focused at $h = 0$ $\Leftrightarrow$ $J_{\mathrm{DS}}[v;d]$ minimized over v, where

$$J_{\mathrm{DS}}[v;d] = \frac{1}{2}\sum_{m,h,n} |hp_n(m,h,0)|^2$$

$p_n(m,h,0)$ = depth-extrapolated shot-geophone field in midpoint m and offset h coordinates, depth z_n, time lag = 0.

The Algorithm

Because J_{DS} is smooth and (perhaps) unimodal, can use rapidly convergent *quasi-Newton* algorithm.

- Limited memory variant of Broyden-Fletcher-Goldfarb-Shanno (Nocedal, 1980 - available through Netlib);
- Velocity parametrization - bicubic splines, sigmoid representation to enforce bounds;
- Gradient computation - adjoint state method applied to DSR.

[See abstract, references for details]

Example 1: Marmousi reflectivity, linear velocity

Data (both examples) generated by time domain FD method for Born modeling. Source wavelet = 4-10-25-35 Hz zero phase trapezoidal bandpass filter.

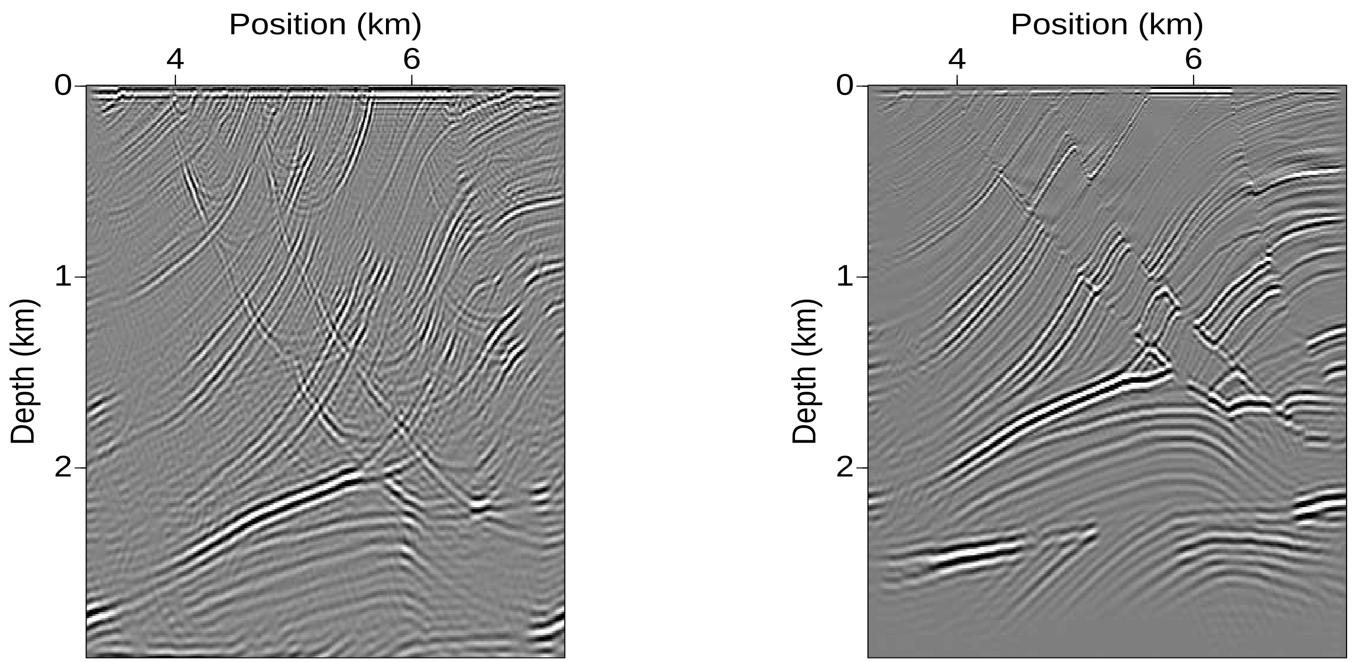
Target velocity, used to generate data: linear, = 1.5 km/s at $z = 0$, = 4.5 km/s at $z = 3km$, represented on bicubic spline grid of 6 nodes in x (Δx = 1.8 km) and 5 nodes in z (Δz = 0.75 km).

Initial velocity also linear, = 2 km/s at $z = 0$, = 4.5 km/s at $z = 3$ km.

Reflectivity = Marmousi velocity model (Versteeg and Grau, 1991) minus 20 m smoothing.

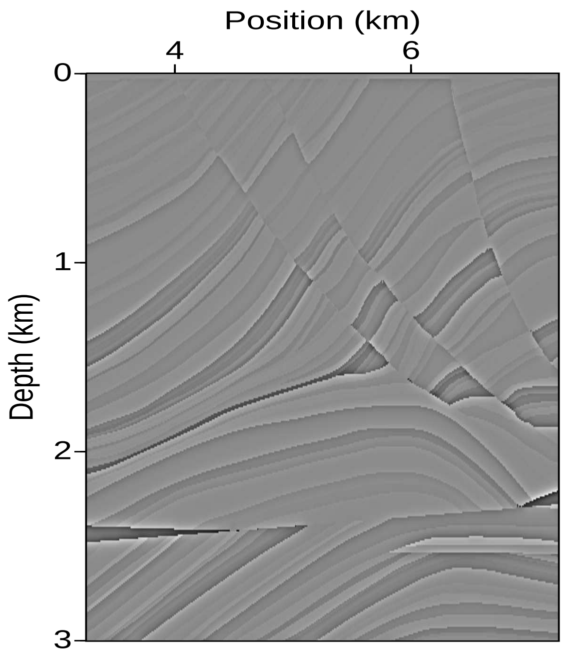
Data geometry same as original.

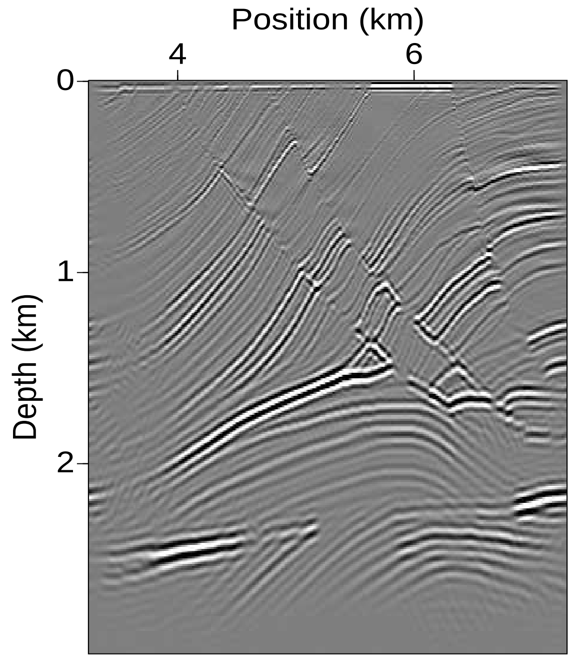
Example 1: Marmousi reflectivity, linear velocity



Left: Migrated image at initial v. Right: after 20 LBFGS v updates.

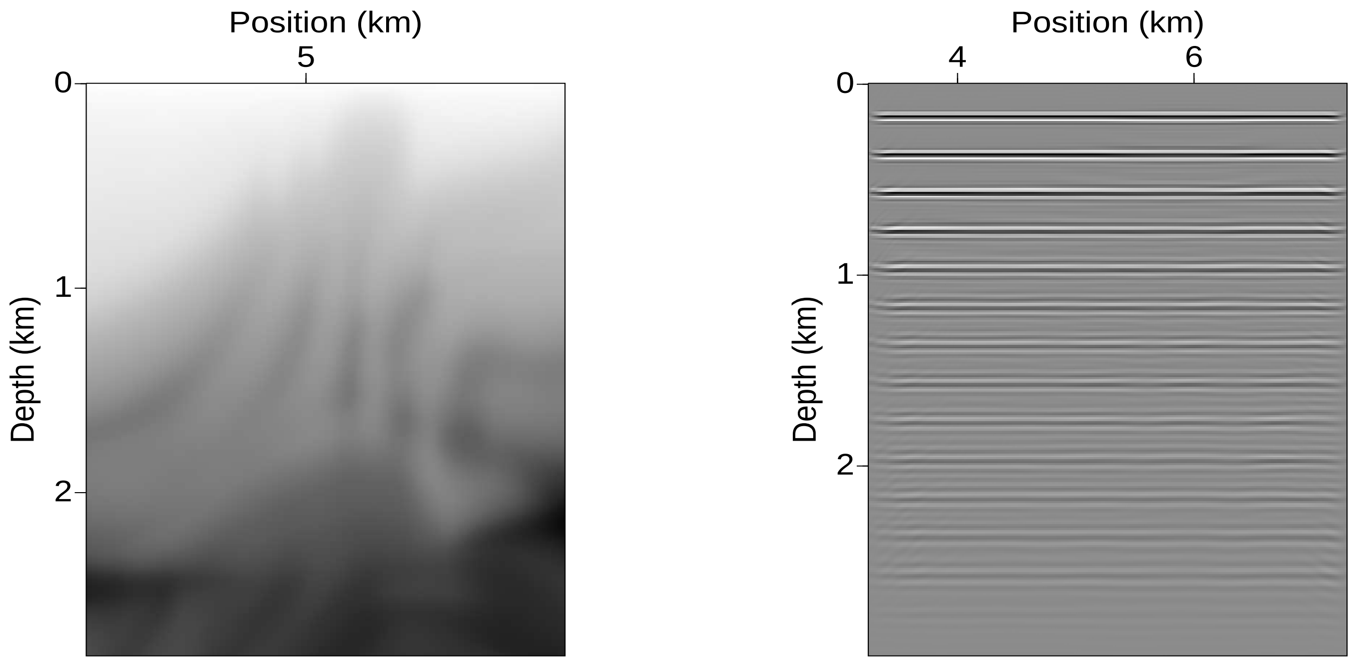
Example 1: Marmousi reflectivity, linear velocity





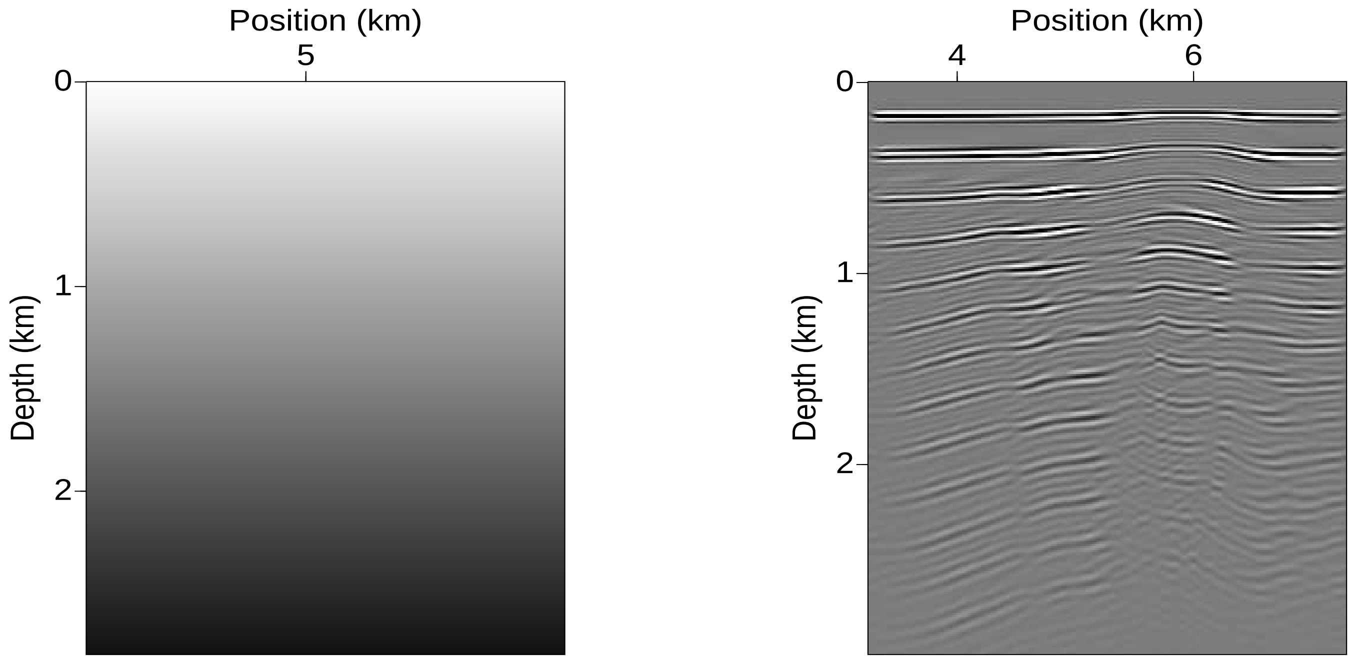
Left: Input reflectivity ("true" image). Right: image from DS velocity analysis. Good focusing and geometry in center. Some residual "sag" from initial velocity error remains on sides. Suggests that larger aperture $\Rightarrow$ more accurate v_{DS}.

Example 2: Layered reflectivity, smoothed Marmousi velocity



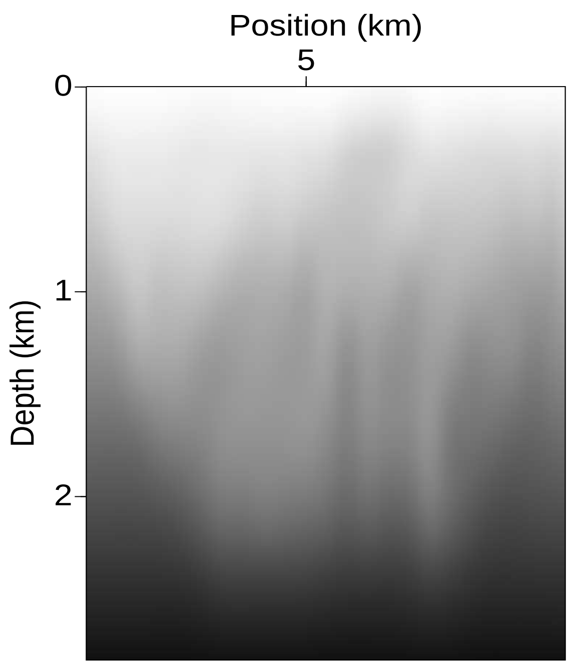
Modeling Inputs. Left: Smoothed Marmousi velocity (160 m smoothing width). Right: layered reflectivity.

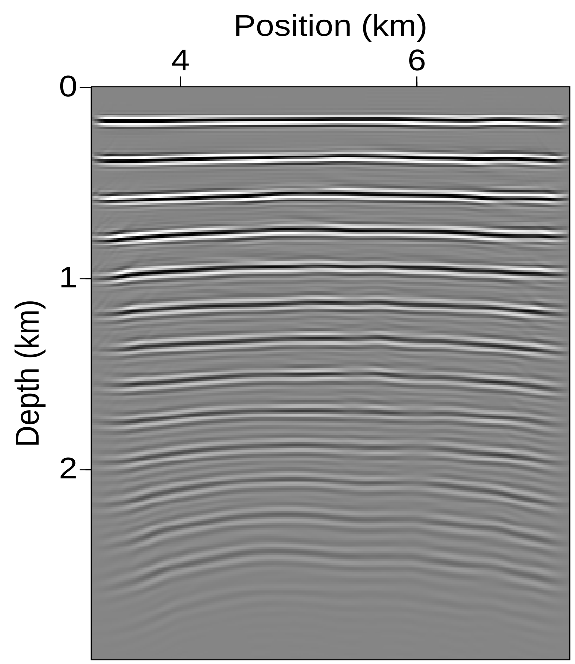
Example 2: Layered reflectivity, smoothed Marmousi velocity



Left: Linear velocity, initial guess for optimization. Right: image.

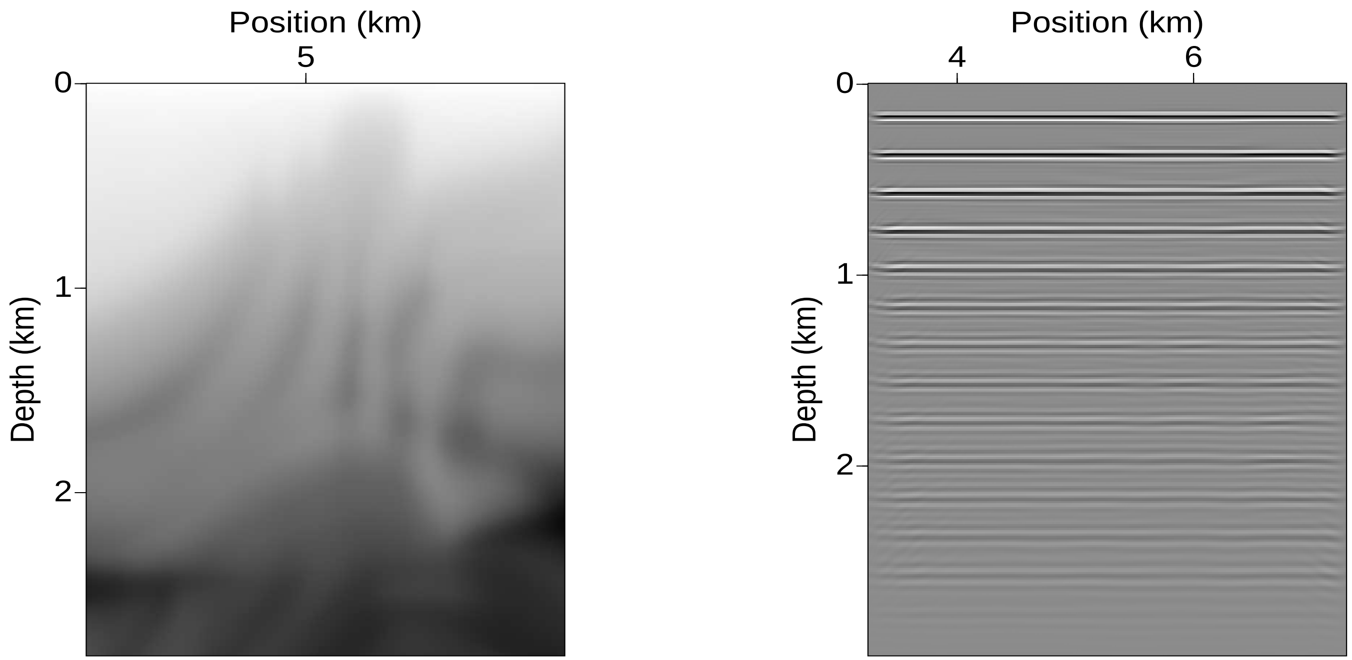
Example 2: Layered reflectivity, smoothed Marmousi velocity





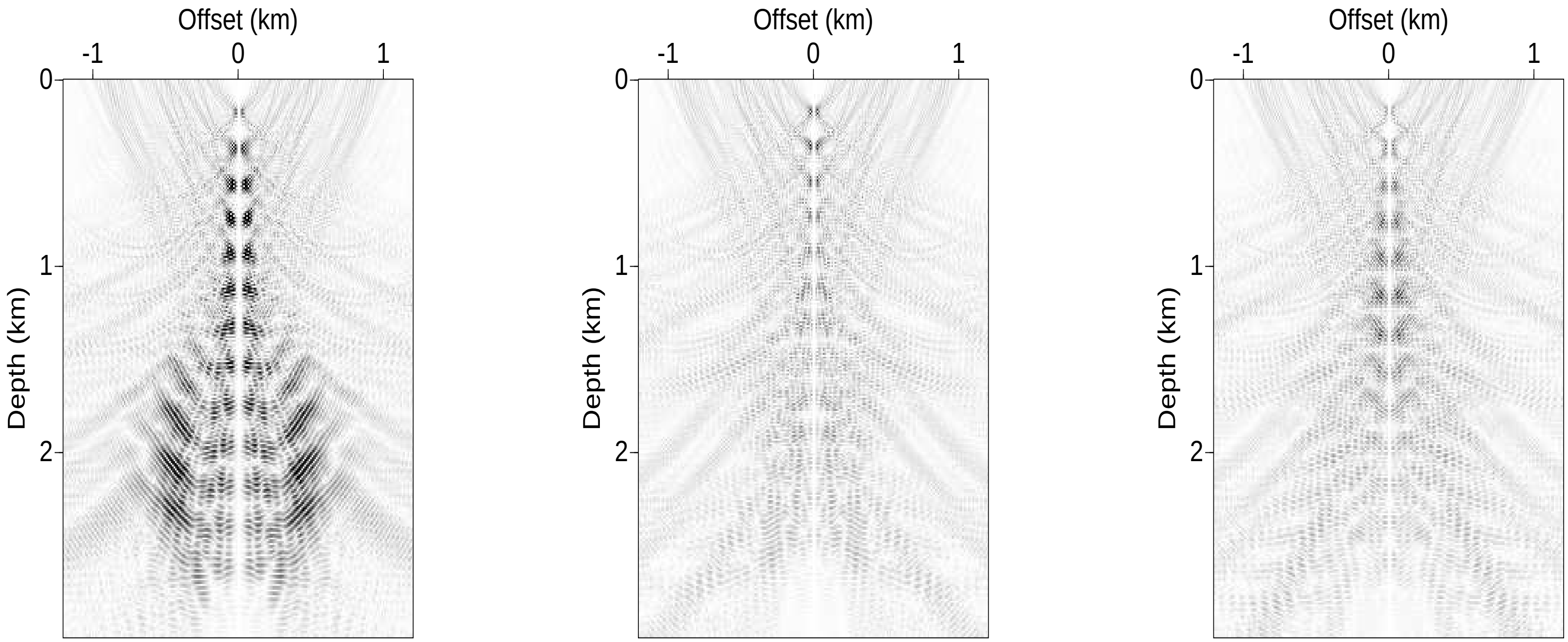
After 20 LBFGS iterations. Left: velocity. Right: image.

Example 2: Layered reflectivity, smoothed Marmousi velocity



Modeling Inputs. Left: Smoothed Marmousi velocity (160 m smoothing width). Right: layered reflectivity.

Example 2: Layered reflectivity, smoothed Marmousi velocity



Plot of image gather scaled by h and squared (sum of all such scaled, squared gathers = $J_{\rm DS}$). Left: initial model. Center: 20 LBFGS updates. Right: input ("true") model. $J_{\rm DS}$ for updated model actually *smaller* than for input model!

Example 2: Layered reflectivity, smoothed Marmousi velocity

How can a "wrong" model generate better focus than a "right" model?

Cause: coherent noise in image gathers.

Possible causes of larger $h \neq 0$ signal in image gathers for "true" velocity:

- coherent noise in data [various remedies, but unlikely to be a factor here];
- edge diffractions [remedy: taper on all axes];
- *mismigration of high angle events* - consistent with appearance of gathers. [remedy: better propagator? Two-way RTM?].

Conclusions - DSVA-DSR

- Have demonstrated a DSR-based automated VA prototype, extensible to 3D via common azimuth approximation.
- Cost of velocity analysis $\simeq$ a few 10's of migrations.
- With present components, algorithm makes large model updates and greatly improves focusing of images.
- Available aperture affects accuracy (cf. tomography).
- Coherent noise in image volume, from data or imaging operator, can degrade accuracy.

Action items for further research: assess effect of better (more expensive!) extrapolators in reducing operator-induced coherent noise; investigate effect and mitigation / modeling of coherent data noise; quantify aperture influence on velocity resolution.

Current Projects

Kirchhoff-based DS via Eikonal Solvers

Attraction: best chance of extension of DSVA-NMO theory, plus interesting numerics (eikonal package), plus applicable to sediment imaging problems.

Theory: main step in "all stationary points" proof uses hyperbolic moveout to *localize* the influence of moveout error in velocity. Possible substitute in PSDM case: time migration (esp. theory incl. error estimates developed by Cameron, Fomel, Sethian SEG 06).

Numerics: probably have to go beyond 1st order scheme implemented in current package to gain enough traveltime accuracy.

Status: Jintan Li MA project to demonstrate basic imaging code, DS computation.

RTM-based DS

Concept introduced by S., also Biondi & Shan, in 02. Introduce shift in crosscorrelation

$$I(\mathbf{x},\mathbf{h}) = \sum_n \Delta t^2 v(\mathbf{x})(Lu^n)(\mathbf{x}-\mathbf{h})w^{n+1}(\mathbf{x}+\mathbf{h})$$

so usual RTM output obtained with $\mathbf{h} = 0$ (recall: u^n = source wavefield, w^n = receiver wavefield, L = Laplacian approximation).

Prospects for (i) layered, (ii) 2D, with high efficiency parallization:

- accuracy regulated by FD scheme; high angle waves propagated equally well;
- judge focussing of finite frequency wavefields;
- generic source-receiver migration, admits arbitrary propagation angles, vertical and horizontal offset gathers $\Rightarrow$ VA for arbitrarily complex *nonreflecting* backgrounds.

Nonlinear Inversion and MVA

Major source of coherent noise: multiple reflection.

SO USE IT!!!!!

Proposed 2005: nonlinear extension and semblance concept.

What stands in the way:

- Theory: when is extended IP invertible?
- Even for layered problems: strong reflection $\Leftrightarrow$ major reflectors *in background model* - invertibility? (No known mathematical results!)
- For CAP extensions: how to avoid small heterogeneities $\Rightarrow$ multipathing;
- For S/TS extensions: how to deal mathematically, computationally with SPD operator coefficients (for layered media, these are convolutional so no sweat).

Other issues: attenuation, sources

Viscoelasticity: formerly a major TRIP emphasis (Blanch, Robertsson, Minkoff).

Minkoff thesis: account for (1) velocity (DSO), (2) elastic P-P reflection, (3) viscoelastic propagation, (4) source anisotropy, and you can fit seismic data (90%). Don't and you can't, and moreover you get the wrong answer (AVO-wise).

Joint inversion source-reflectivity (Minkoff 1997, Winslow MA 1999) should be extended to nonlinear inversion.

Nonlinear acoustic/source inversion: Lailly (2004, 2005, 2006) suggests this is not possible, but we have our doubts...