Geophysical Prospecting EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

*Geophysical Prospecting*, 2013, **61** (**Suppl. 1**), 346–351 doi: 10.1111/j.1365-2478.2012.01121.x

# Statics-preserving projection filtering

Yann Traonmilin[1]* and Necati Gulunay[2]

[1]*Formerly CGGVeritas,* [2]*CGG Veritas*

Received April 2012, revision accepted July 2012

**ABSTRACT**

Projection filtering has been used for many years in seismic processing as a tool to extract a signal out of noisy data. The effectiveness of projection filtering reaches a limit when seismic events are affected by static shifts. Such shifts degrade the lateral coherency of the data, which is the strongest assumption made by projection filtering. We propose an algorithm to estimate projection filters and static shifts simultaneously in order to perform noise attenuation in the presence of static shifts in the data. We then show results on synthetic and real data to demonstrate the denoising capabilities of our algorithm.

**Key words:** Noise, Signal processing.

## INTRODUCTION

Random noise attenuation has been a subject of interest in seismic processing for many years. We now have a wide variety of tools to perform such a task. From prediction and projection filtering (Canales 1984; Soubaras 1995) to coherency enhancement (Gulunay 2007; Traonmilin and Herrmann 2008), we can efficiently address many situations where random noise contaminates data. These methods rely on the hypothesis that the signal is spatially coherent in order to estimate the signal. A signal is coherent if it is the sum of elementary coherent events. Depending on the denoising method, these elementary events can be defined differently. For prediction and projection filtering, we suppose that the signal is a sum of sparse events in the wavenumber ($k$) domain (seismic events are predictable in the space ($x$) direction). In the case of denoising with Radon transform (Hampson 1987; Gulunay 1990; Herrmann *et al.* 2000), we suppose that the signal is a sum of elementary linear or parabolic events.

However, this coherency assumption is not met when static shifts are present in the data. This problem arises in several situations. In land data, static shifts caused by near-surface propagation effects harm the lateral coherency of seismic events. The structural dip generates static shifts between traces of a midpoint gather for wide-azimuth data unless traces are sorted in increasing azimuth as NMO velocities needed for dipping data are azimuth-dependent. This factor as well as physical azimuthal anisotropy, which manifests itself as statics, may lead to a perturbed velocity analysis and more generally to difficult prestack processing.

If the seismic data have static shifts and are contaminated by noise, we generally face the following problem: on the one hand, if we cannot correct for these shifts, we cannot use conventional noise attenuation techniques that rely on lateral coherency without harming the data. On the other hand, if we cannot attenuate noise, it will be difficult to estimate the shifts. Usual workarounds lead to tedious workflows of denoising and picking iterations (surface-consistent refraction or reflection statics picking and calculation, velocity picking, etc). The objective of this paper is to show that it is possible to break this vicious circle and perform a denoising technique that will not harm the static shifts present in the data. We thus remove the need for denoising algorithms requiring a data set with corrected static shifts.

In this paper, we propose an algorithm to attenuate random noise when seismic events are affected by static shifts. We begin by describing a variant of 2D projection filtering that simultaneously estimates static shifts. We then show on synthetic and real data that it attenuates random noise when seismic events are affected by static shifts.

*E-mail: yann.traonmilin@gmail.com

 

## THEORY

### Definition of the problem

We consider in this paper a generalization of the conventional data model used to apply 2D projection filtering. Our 2D seismic data model is a laterally predictable signal where each trace is shifted and contaminated by noise. We assume that the noise is additive, white and random. We can write it as:

$$d(t,x) = s(t - t_x, x) + n(t,x), \quad (1)$$

where $t$ is the time, $x$ is a spatial dimension, $d$ is the recorded data, $s$ is the spatially predictable signal and $n$ is the noise. The $t_x$ are the static shifts that depend on the spatial dimension $x$. We suppose in this paper that the shifts are random with zero mean. Our objective is to recover the signal $s(t - t_x, x)$.

In the temporal Fourier domain, equation (1) becomes:

$$d(f,x) = T(f,x)s(f,x) + n(f,x), \quad T(f,x) = e^{j2\pi f t_x}. \quad (2)$$

If we multiply this equation by $T(f,x)^{-1}$ (which is equivalent to removing the shifts in the data), we have the usual data model used for projection filtering: a predictable signal with additive random noise. We could then remove the noise by calculating a prediction error filter $A$ and then deriving the corresponding projection error filter $P$ (Soubaras 1995). Thus the only information missing to perform a statics-preserving projection filtering (estimation and application) is knowledge of the static shifts present in the data. We would then be able to denoise the data by removing the shifts, calculating and applying the projection filter and finally re-applying the shifts. We consequently propose to solve for the static shifts and projection filters at the same time.

In what follows, we use notations $A$, $P$ and $T$ for the linear operations of prediction filtering, projection filtering and shifting a data set (which we will write as $d$) in the f-x domain respectively.

We can then write the problem of the least-squares joint estimation of $A$ and $T$ as:

$$\text{find} \quad \arg\min_{A,T} \quad ||AT^{-1}d||^2. \quad (3)$$

If we can calculate $A$ and $T$, we will be able to derive a projection error filter $P$ from $A$ (Soubaras 1995) and use it to obtain an estimate of the noise contaminating the data by using this equation:

$$n = TPT^{-1}d. \quad (4)$$

This operation corresponds to a three-step process: removal of the static shifts, projection filtering and application of the static shifts. This noise estimation equation shows that we can add a constant to a solution for the static shifts and obtain the same results. We take the first trace as the reference (shift is zero) to make shift values.

The relation between the unknown and known parameters in equation (3) is not linear, so we cannot hope to solve this problem with a simple linear least-squares technique. We therefore need to use a non-linear method to solve this problem.

### An iterative algorithm

The most straightforward way to perform non-linear minimization is to use an iterative scheme where we linearize the problem locally. From estimates of $T$ and $P$ at an iteration $i$ (we call these estimates $T_i$ and $P_i$), we need to calculate $T_{i+1}$ and $P_{i+1}$ that decrease the quantity given in equation (3). Let us assume that we have an estimate of the shifts $T_i$. Fixing $T_i$ in the minimization will allow us to obtain an estimate of the prediction error filter $A_{i+1}$:

$$\text{find} \quad \arg\min_{A_{i+1}} \quad \left|\left|A_{i+1}T_i^{-1}d\right|\right|^2. \quad (5)$$

This can be done directly using normal equations. We then derive the projection error filter $P_{i+1}$ corresponding to $A_{i+1}$ and use it to solve the following problem for $T_{i+1}$ :

$$\begin{aligned} &\text{find} \quad \arg\min_{T_{i+1}} \varepsilon_i \\ &\text{with} \quad \varepsilon_i = \left|\left|P_{i+1}T_{i+1}^{-1}d\right|\right|^2, \end{aligned} \quad (6)$$

where $\varepsilon_i$ is the objective function at step $i$. With real data, we have little guarantee that solving this system directly will give correct shifts because coherent noise can bias the solution. There is also a possibility that we may converge to a local minimum. Instead, we use a gradient descent update for the vector of time-shifts $t_{i+1}$ defining $T_{i+1}$ as:

$$t_{i+1} = t_i - \mu_i \overline{\nabla}_{t_i}\varepsilon_i, \quad (7)$$

where $\overline{\nabla}_{t_i}\varepsilon_i$ is the gradient (or conjugate gradient) direction and $\mu_i$ is the step size. The gradient takes a simple form:

$$\overline{\nabla}_{t_i}\varepsilon_i = 2\,Re\left[(P_{i+1}T_i^{-1}d)^H(P_{i+1}(\nabla_{t_i}T_i^{-1})d))\right]. \quad (8)$$

The differentiation of $T_i^{-1}$ is done by differentiating the multipliers $T(f,x)$ from equation (2). With this update method, we can control the values taken by $t_{i+1}$ and we are guaranteed that our objective function will decrease.

Equations (5) and (7) define our iterative minimization algorithm. This algorithm is similar to the primary estimation by sparse inversion (Groenestijn and Verschuur 2009) where successive linear least-squares inversion and gradient step update are performed. We do not need to enforce sparsity but we need to make sure that we do not converge to a local minimum. We provide a method to avoid this in the next chapter. We also define the initialization value of $T_0$ and the step size $\mu_i$ to complete the description of this algorithm.

### Initialization, step size and local minima

We need to set the algorithm parameters to ensure convergence towards the global minimum of the objective function $\varepsilon_i$. We need to define how we initialize the data and how we define the step size along iterations. We must also make sure that the algorithm is not trapped in any local minimum as the problem is not linear.

We first make an observation on the behaviour of the problem with respect to frequency. We will then explain how it relates to the choice of the parameters for our algorithm. We write the recorded data (defined in equation (2)) as a perturbation of the data without shifts:

$$\begin{aligned} d(f,x) &= T(f,x)s(f,x) + n(f,x) \\ &= s(f,x) + s(f,x)(\exp[j2\pi f t_x] - 1) + n(f,x). \end{aligned} \tag{9}$$

The second term in equation (9) is additional noise induced by the static shifts when compared to predictable data contaminated with random noise. We want to see how the energy of this term behaves with respect to the amplitude of the static shifts. If we assume that the shifts have a Gaussian distribution with zero mean and standard deviation $t_m/2$, we can show that (see Appendix for details):

$$E_x|s(f,x)(\exp[j2\pi f t_x] - 1)|^2 < \pi^2 t_m^2 f^2 E_x|s(f,x)|^2, \tag{10}$$

where $E_x|\cdot|^2$ is the energy along the $x$ direction. Thus static shifts cause less deviation from the ideal model at lower frequencies, giving better estimates of projection filters.

Specifically, if statics are low enough or frequencies are low enough:

$$t_m < \frac{1}{\pi f}, \tag{11}$$

we obtain:

$$E_x|s(f,x)(\exp[j2\pi f t_x] - 1)|^2 < E_x|s(f,x)|^2. \tag{12}$$

This means that if condition (11) holds then the term added by the statics is less energetic than the signal without shifts, a well-known phenomenon.

These observations compel us to start solving the low frequencies first and then adding frequencies as the number of iterations increases. This brings two advantages. At low frequencies, we can initialize the statics estimate with $T_0 = \mathit{Identity}$ (initial shifts set to zero) because the estimation of the projection filter will be closer to the solution compared to an estimation taking all frequencies. Starting with low frequencies also avoids local minima, as aliasing caused by the statics is the reason for these local minima (it is similar to cycle skipping). With this method, we can estimate shifts having values as large as half of the inverse of the minimum frequency present in the data. The required number of iterations is typically under 50.

We can use a line search or a heuristic method to determine the step size. The problem that arises for the line search is that we have to apply the projection filtering for every value of $\mu_i$ that we want to test. This process can be very time-consuming. On the other hand, we have some *a priori* information on the static shift values like their distribution or maximum values, we can then use this to set the step size: we start with a large step size and lower it as the statics solution becomes closer to the *a priori* distribution. This saves us from the computational complexity of the line search. By setting the step size this way, we limit the solution space of the statics term in the objective function and guarantee that we converge towards the desired result by avoiding local minima.

## EXAMPLES

### Synthetic data

We generate two synthetic data sets fitting the data model used in our algorithm. They consist in three band-limited linear events, with random static shifts applied and random noise added. Two different magnitudes of shifts are tested. In Fig. 1, shift values are randomly distributed between −4 ms and +4 ms, which is the sampling interval. In Fig. 2, the shifts are randomly distributed between −32 ms and +32 ms (with the same sampling interval). In (Fig. 1), we are able to remove random noise and to preserve the shifts at the same time, even when the shifts are large (Fig. 2). We also show in these figures the limit of conventional projection filtering: irregularities are smeared and primary events are harmed by projection filtering as there is residual signal in the difference sections.

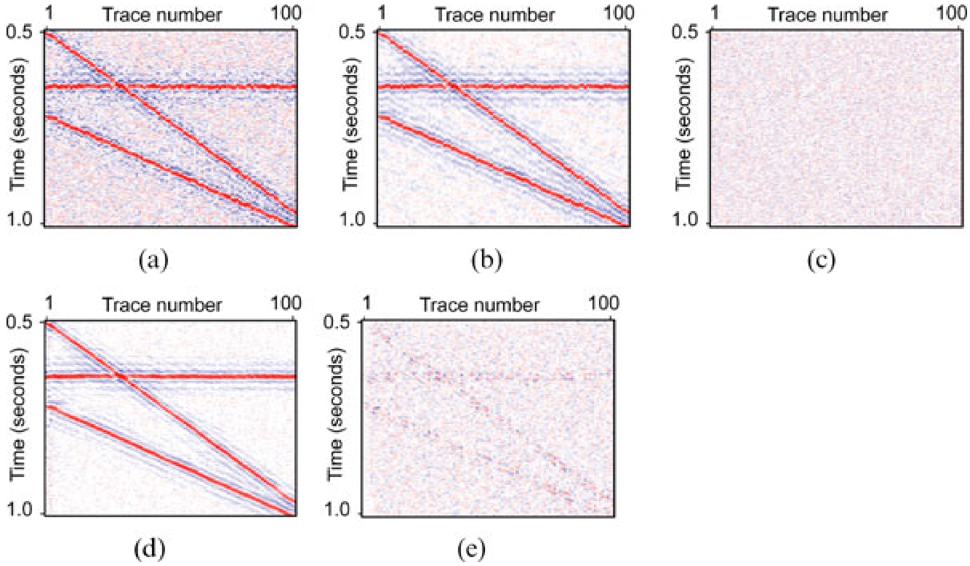


**Figure 1** Denoising of three synthetic linear events with small random static shifts: (a) input data, (b) denoised data using statics-preserving projection filtering, (c) noise removed with statics-preserving projection filtering, (d) denoised data using conventional projection filtering, (e) noise removed by conventional projection filtering.

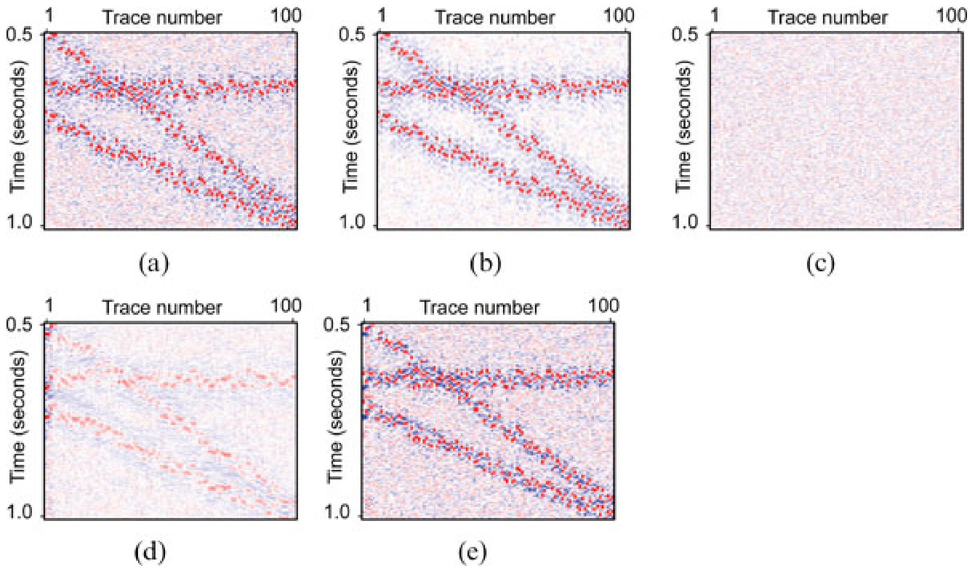


**Figure 2** Denoising of three synthetic linear events with large random static shifts: (a) input data, (b) denoised data using statics-preserving projection filtering, (c) noise removed with statics-preserving projection filtering, (d) denoised data using conventional projection filtering, (e) noise removed by conventional projection filtering.

Our second synthetic data set (Fig. 3) simulates an event in a wide-azimuth CMP gather corrected with normal moveout. The geometry of this WAZ gather is shown in Fig. 3(d). Azimuthal anisotropy caused by a structural dip generates jitter on the event as traces are not sorted in azimuth order and NMO velocities needed for dipping data are azimuth-dependent. (Fig. 3a). They are handled well by our algorithm (Fig. 3b), even as the magnitude of shifts increases at large offsets. The accuracy of the filtering is particularly noticeable on the side lobes of the wavelet that are perfectly preserved. This example highlights the fact that our algorithm is effective even when the amplitude of the shifts is highly variable.

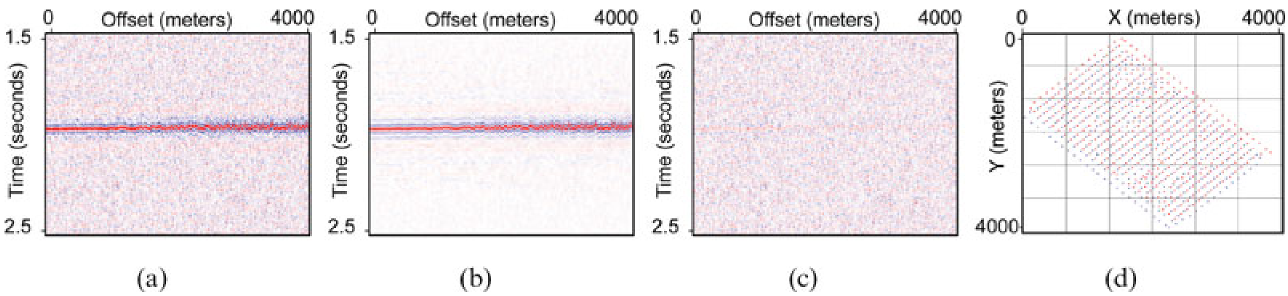


**Figure 3** Denoising of a synthetic WAZ offset gather with statics-preserving projection filtering:(a) input data, (b) denoised gather, (c) removed noise, (d) map of the source (blue) -receiver (red) locations.

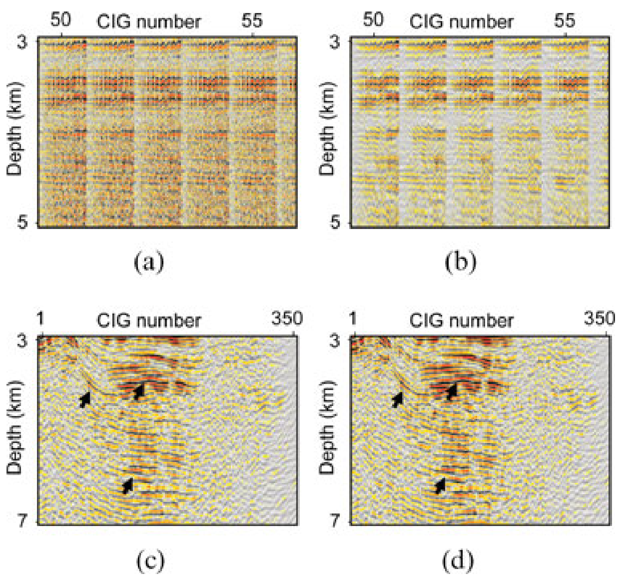


**Figure 4** Preconditioning of a RMO flattening workflow with statics-preserving projection filtering: (a) raw WAZ CIGs, (b) CIGs denoised with statics-preserving projection filtering, (c) stack using conventional RMO flattening sequence, (d) stack using preconditioned input.

### Real data

In Fig. 4, we use our technique to enhance the flattening workflow of land wide-azimuth common image gathers (CIGs) (Gulunay, Magesan and Roende 2007). The objective is to improve the signal-to-noise ratio of CIGs and then process them with a dedicated flattening workflow. As the flattening method is mostly a local process, it will have a problem dealing with areas where the signal is completely overpowered by noise. Therefore, the flattening of the noisy data does not lead to optimal focusing of the stack. After the application of statics-preserving projection filtering on each CIG, the signal-to-noise ratio is improved and static shifts are preserved (Fig. 4a,b). Statics-preserving projection filtering uses bigger spatial windows than conventional trim statics estimation algorithms because we use the same sizes as conventional projection filtering (here 30 traces and 300 ms). Consequently, we are able to recover the signal even in areas where it is very weak with respect to noise. We use CIGs cleaned up in this way for the calculation of time shifts in the flattening process. We apply the resulting shifts to the original (noisy) gathers and stack the result (Fig. 4d). We compare the stack of the conventional gather flattening technique (Fig. 4c) to this result to isolate the effect of the denoising on the estimation of the shifts. The higher signal-to-noise ratio on cleaned gathers leads to a better estimation of time shifts for flattening, which results in a better focusing of the stack especially where events are dipping.

## CONCLUSION

We demonstrated that with statics-preserving projection filtering, we are now able to attenuate random noise when static shifts are present in the data. We also demonstrated that our method brings out a significant uplift in the signal-to-noise ratio when processing wide-azimuth land data.

The natural extension of this work would be to implement a 3D statics-preserving projection filtering. Most of the theoretical developments in this paper still hold in case of 3D projection filtering. In fact, any process that relies on the least-squares minimization of a linear functional of the data (e.g., least-squares Radon transform) could potentially benefit from this type of algorithm, which estimates the best static shifts for the problem.

## ACKNOWLEDGEMENTS

We would like to thank Apache Corporation for permission to show the data given in Fig. 4. We also thank Olivier Hermant and Jean-Paul Gruffeille (CGGVeritas Cairo) for their support. We thank CGGVeritas for giving us permission to publish this paper.

## REFERENCES


Canales L. 1984. Random noise reduction. *SEG Expanded Abstracts* **3**, 525–527.

Groenestijn G.J. and Verschuur D.J. 2009. Estimation of Primaries by Sparse Inversion and its Application to Near Offset Reconstruction. 71st EAGE Conference & Exhibition.

Gulunay N. 1990. *F-X* domain least squares tau-*P*and tau-*Q*. *SEG Expanded Abstracts* **9**, 1607–1610.

Gulunay N. 2007. Coherency enhancement on 3D seismic data by dip detection and dip selection. *SEG Expanded Abstracts* **26**, 2625–2629.

Gulunay N., Magesan M. and Roende H.H. 2007. Gather Flattening. *The Leading Edge* 1538–1543.

Hampson D. 1987. The discrete Radon Transform: A new tool for image enhancement and noise suppression. *SEG Expanded Abstracts* **6**, 141–143.

Herrmann P., Mojesky T., Magesan M. and Hugonnet P. 2000. De aliased, high-resolution radon transforms. *SEG Expanded Abstracts* **19**, 1953.

Soubaras R. 1995. Prestack random and impulsive noise attenuation by f-x projection filtering. *SEG Expanded Abstracts* **14**, 711–714.

Traonmilin Y. and Herrmann P. 2008. Structurally consistent f-x filtering. *SEG Expanded Abstracts* **27**, 2642–2646.


## APPENDIX: DEMONSTRATION OF INEQUALITY IN EQUATION (10)

First, we calculate the energy of the term created by the static shifts. We then suppose that the shifts and the signal are statistically independent, which gives:

$$\begin{aligned} E_x|s(f,x)(\exp[j2\pi f t_x]-1)|^2 &= E_x(|s(f,x)|^2|(\exp[j2\pi f t_x]-1)|^2) \\ &= E_x(|s(f,x)|^2)E_x(|(\exp[j2\pi f t_x]-1)|^2). \end{aligned} \tag{A1}$$

We want to find a bound on $E_x(|(\exp[j2\pi f t_x]-1)|^2)$. We use the fact that $|\sin(x)| \leq |x|$ to obtain :

$$|(\exp[j2\pi f t_x]-1)|^2 = 4\sin^2(\pi f t_x) \leq 4(\pi f t_x)^2, \tag{A2}$$

which we integrate using the Gaussian distribution $t_x \sim \frac{t_m}{2}\sqrt{2\pi}\exp[-\frac{1}{2}(\frac{2x}{t_m})^2]$. We find:

$$\begin{aligned} E_x(|(\exp[j2\pi f t_x]-1)|^2) &\leq \int_R 4(\pi f x)^2 \frac{1}{\frac{t_m}{2}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{2x}{t_m}\right)^2\right) \mathrm{d}x \\ &\leq \pi^2 t_m^2 f^2, \end{aligned} \tag{A3}$$

which shows inequality (10).