SNR based automatic rank determination and its application for denoising using MSSA

Alejandro Quiaro Sandoval

Outline

- Introduction
- Theory
- Synthetic data examples
- Real data examples
- Discussion

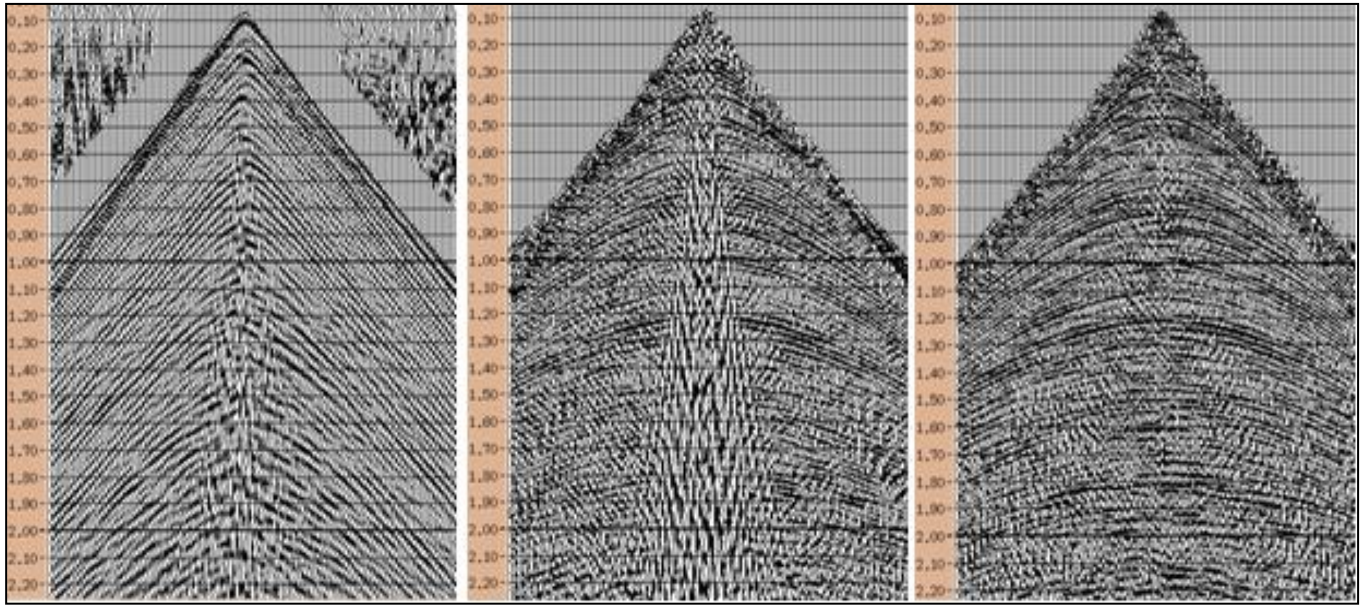
Introduction

Seismic data processing: apply changes to the data in order to facilitate its interpretation. It can also be seen as a set of techniques used to increase the **signal to noise** ratio of a dataset.

Based on the space-time behavior of the noise, it can be classified as:

Coherent:
- Ground-roll
- Air waves
- Near Surface layer reverberations, etc.

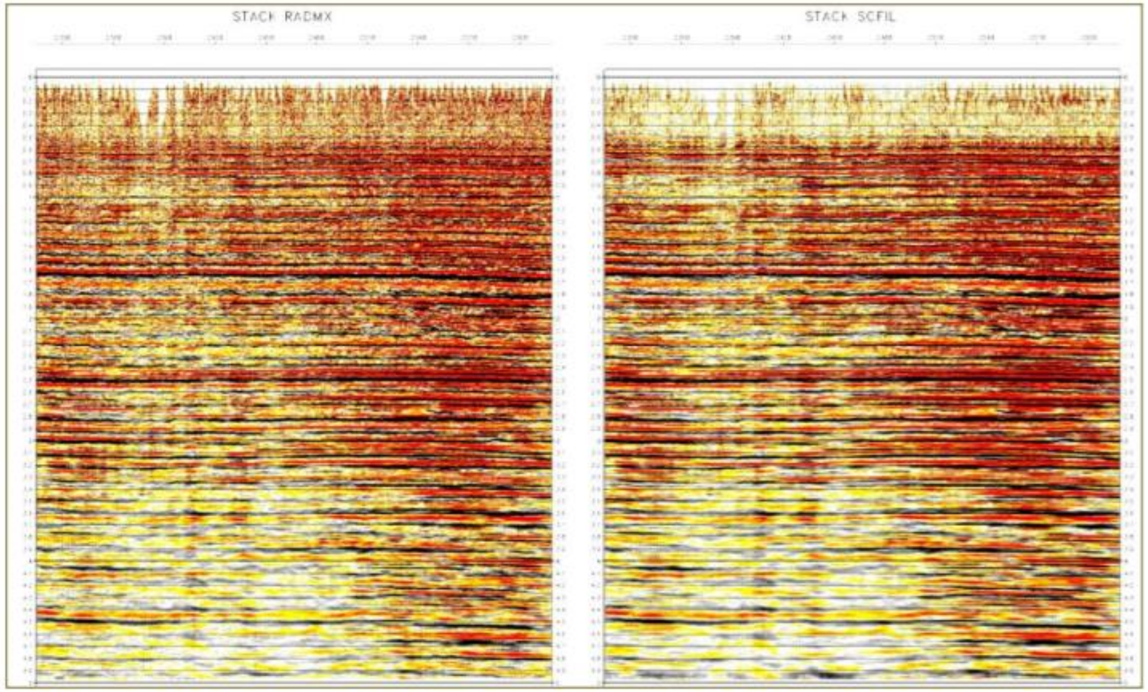
Random:
- Anthropogenic.
- Bad coupling of geophones
- Wind/tree's roots, etc.



Attenuation of coherent noise

Introduction

In the case of random noise, since the space-time behavior of the noise is unknown, usually the attenuation algorithm aim to detect the signal that is laterally uncorrelated or “not-representative”. An example of this is the FX-decon algorithm.



FX-Decon for RNA

Considering random noise as data that in “not-representative” of the underlying geology, allows us to think of the random noise attenuation process as a Rank Reduction problem.

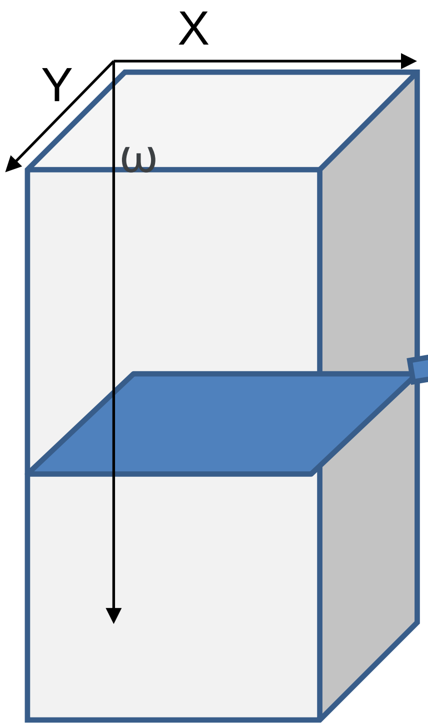
The MSSA (Multi-channel singular spectrum analysis) is a well known process that uses rank reduction to attenuate random noise.

As with other rank reduction problems, the question arises. Is there an objective criteria to choose the appropriate rank for denoising?

Theory

Theory

Consider a 3D seismic cube, which has been Fourier transformed into the F-XY domain:



3D seismic cube

We can consider a single slice of constant frequency, which can be written as the following matrix:

$$S = \begin{pmatrix} S(1,1) & S(1,2) & \dots & S(1,N_x) \\ S(2,1) & S(2,2) & \dots & S(2,N_x) \\ \vdots & \vdots & \ddots & \vdots \\ S(N_y,1) & S(N_y,2) & \dots & S(N_y N_x) \end{pmatrix}$$

Each row of this matrix, represents the variation of energy for a given frequency along the X dimension.

However, each individual row (which has a combination of noise and signal) cannot be rank reduced. So, it is convenient to reshape the information contained in the vector into a matrix that could be rank reduced.

Theory

Consider each row of S, and build a Hankel Matrix (a Matrix whose values along the anti-diagonal are constant):

$$S = \begin{pmatrix} S(1,1) & S(1,2) & \dots & S(1,N_x) \\ S(2,1) & S(2,2) & \dots & S(2,N_x) \\ \vdots & \vdots & \ddots & \vdots \\ S(N_y,1) & S(N_y,2) & \dots & S(N_y,N_x) \end{pmatrix}$$

$$R_j = \begin{pmatrix} S(j,1) & S(j,2) & \dots & S(j,K_x) \\ S(j,2) & S(j,3) & \dots & S(j,K_x+1) \\ \vdots & \vdots & \ddots & \vdots \\ S(j,L_x) & S(j,L_x+1) & \dots & S(j,N_x) \end{pmatrix}$$

F-XY slice for a fixed frequency ω

Hankel Matrix of a row *j* from matrix S

Up to now, we have considered the variations along the x dimension. To incorporate the information from the y direction, a Block Hankel Matrix can be built, using all the Hankel Matrices from the rows of S, in the following way:

$$M = \begin{pmatrix} R_1 & R_2 & \dots & R_{k_y} \\ R_2 & R_3 & \dots & R_{k_y+1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{L_y} & R_{L_y+1} & \dots & R_{N_y} \end{pmatrix}$$

Theory

Adding noise to the data, increases the rank of the matrix **M**. For that reason, the noise attenuation can be considered as a rank reduction problem. We need to find a rank reduced matrix $\mathbf{M_k}$ that approximates the matrix **M** in a least squares sense. The Eckart-Young theorem (Eckart and Young, 1936) states that the rank-k matrix that minimizes the Frobenius norm can be obtained considering the **k-largest singular values of the SVD of matrix M**.

$$M_k = U_k \Sigma_k V_k^H$$

Given the structure of the Hankel Matrix (constant values across the anti-diagonal), the signal que be recovered by averaging along the anti-diagonals of the Hankel Block Matrix.

But we still have one issue to solve. **How many k singular values to consider for the rank reduction?** To solve this, we propose the following signal-to-noise ratio estimator:

Theory

Let's consider the auto-correlation of an observed signal with one seismic event, which is composed of one part of pure signal and additive noise.

$$r(f,x) = s(f,x) + n(f,x) \rightarrow r^T r = s^T s + s^T n + n^T s + n^T n$$

Assuming that signal and noise are uncorrelated (random noise), and taking the expectations at both sides, we can simplify the expresión in the following way:

$$R = E[s^T s] + \sigma_N^2 I$$

In which **R** is the covariance matrix of the signal, **E** is the expectation operator, **I** is the identity matrix and $\boldsymbol{\sigma}^2_N$ is the variance of the noise.

The matrix **R** has the following properties:

- The rank of the expectation of the pure signal is finite and equal to the number of events (1).
- The minimum eigenvalue of R ($\boldsymbol{\lambda_i}$) is given by the variance of the noise.
- The largest eigenvalue ($\boldsymbol{\lambda_1}$) of R is given by the variance of the signal plus the variance of the noise.

Theory

Then we can estimate the covariance of the noise, as the average of the lowest eigenvalues of matrix **R**:

$$\lambda_i = \sigma_N^2 \rightarrow \sigma_N^2 = \sum_{i=2}^{M} \frac{\overline{\lambda_i}}{(M-1)}$$

Where **M** is the maximum number of eigenvalues. Now, the estimation of the variance of the signal may be obtained from the variance of the noise and the largest eigenvalue in the folowing way:

$$\lambda_1 = \sigma_S^2 + \sigma_N^2 \rightarrow \sigma_S^2 = \lambda_1 - \sum_{i=2}^{M} \frac{\overline{\lambda_i}}{(M-1)}$$

Key and Smithson, 1990, suggest that a signal to noise ratio can be estimated based on the ratio of signal and noise variances, and can be expressed as:

$$\boxed{SNR = \frac{\sigma_S^2}{\sigma_N^2} = \frac{\lambda_1 - \sum_{i=2}^{M} \frac{\overline{\lambda_i}}{(M-1)}}{\sum_{i=2}^{M} \frac{\overline{\lambda_i}}{(M-1)}}}$$

Theory

Our proposal is to compute the expression for the SNR, but make it general for the signal represented by the k_{th} largest eigenvalues, transforming the SNR estimation into:

$$SNR_k = \frac{\sigma_S^2}{\sigma_N^2} = \frac{\sum_{i=1}^{k} \frac{\lambda_i}{k} - \sum_{i=k+1}^{M} \frac{\lambda_{k+1}}{(M-k-1)}}{\sum_{i=k+1}^{M} \frac{\lambda_{k+1}}{(M-k-1)}}$$

Where SNR_k represents the signal to noise ratio of the reconstructed signal, using only the k_{th} largest eigenvalues.

The estimated Rank of the matrix, would be the k number of eigenvalues which provide the highest estimated SNR. Unfortunately, this approach requires to calculate the entire range of eigenvalues for the matrix before estimating the appropriate rank of the matrix.

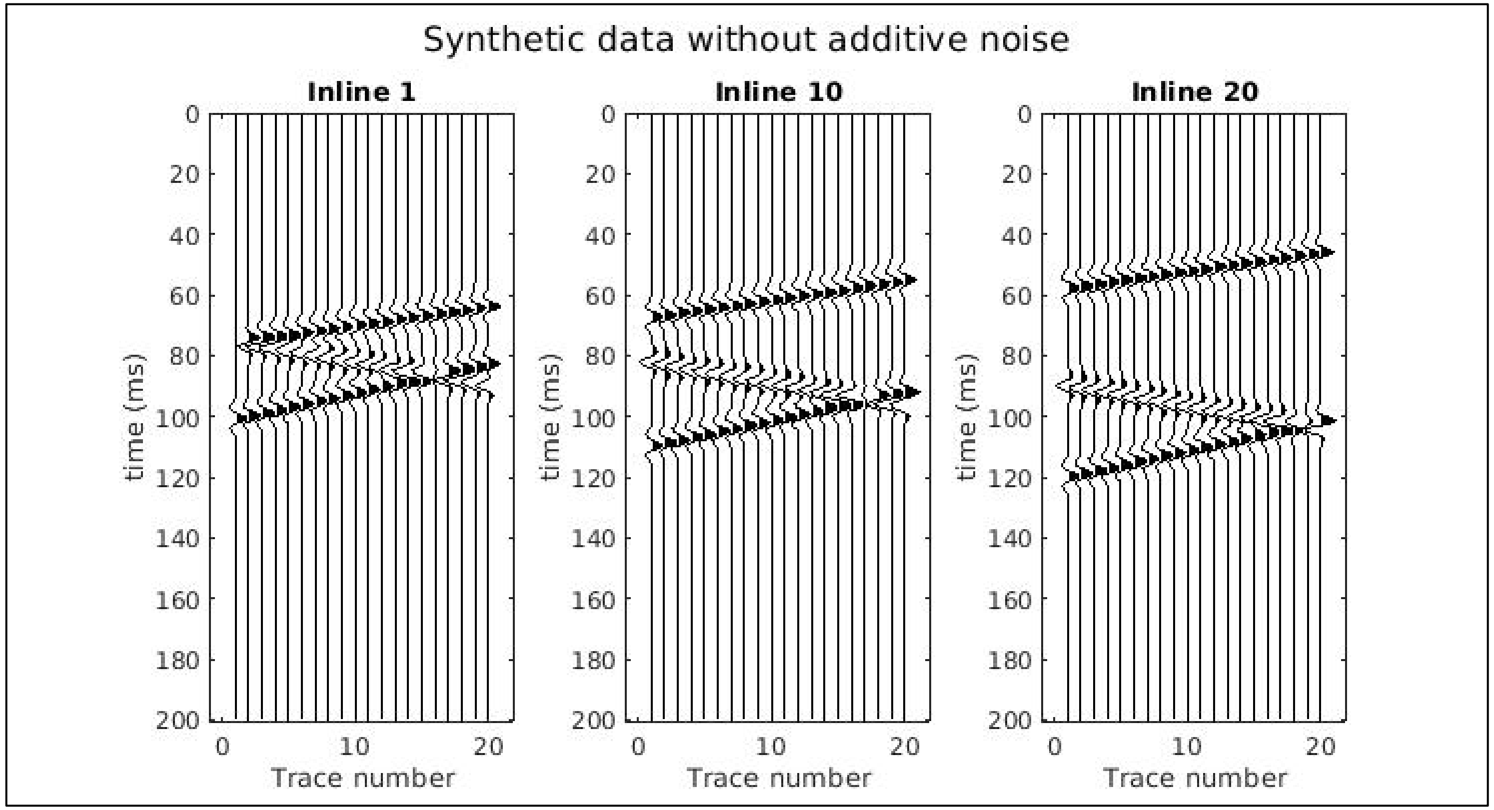
A way to overcome this problem, is to make the estimation only in a narrow frequency band (around the dominant frequency of the input data).

Synthetic Examples

Synthetic Examples

Let’s consider a synthetic seismic cube which has three linear events with different dips. The parameters of the seismic cube are the following:

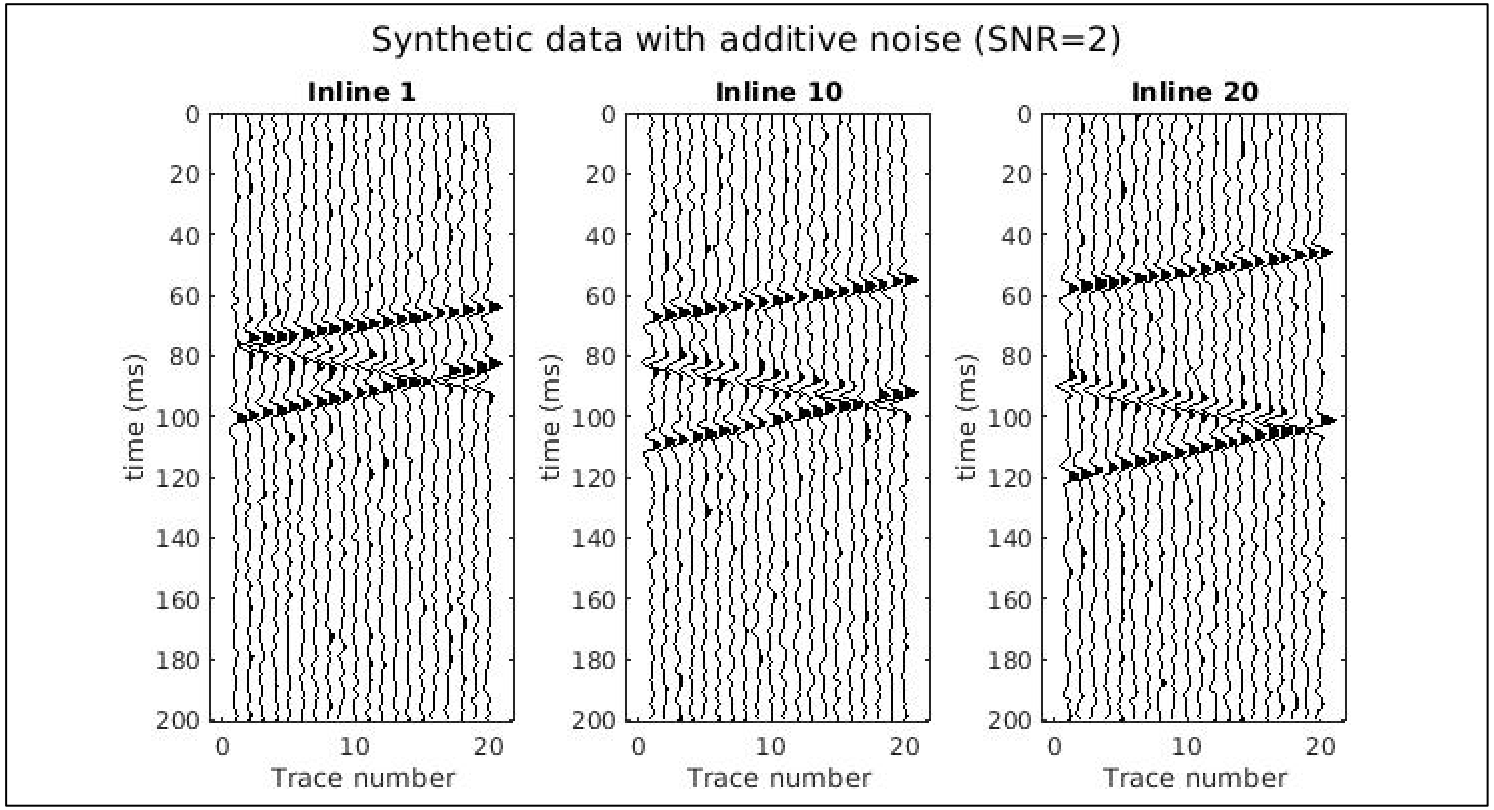
Size: 20x20 traces (distances between traces: 25m)
Sampling rate: 4ms
Trace length: 800ms
Dominant frequency of the wavelet: 30Hz.



Synthetic Examples

Now, let’s add random noise to the seismic data and force an **SNR =2** (twice more data than noise). Later we will analyze the behavior of the workflow for different SNR levels.

Size: 20x20 traces (distances between traces: 25m)
Sampling rate: 4ms
Trace length: 800ms
Dominant frequency of the wavelet: 30Hz.
SNR=2

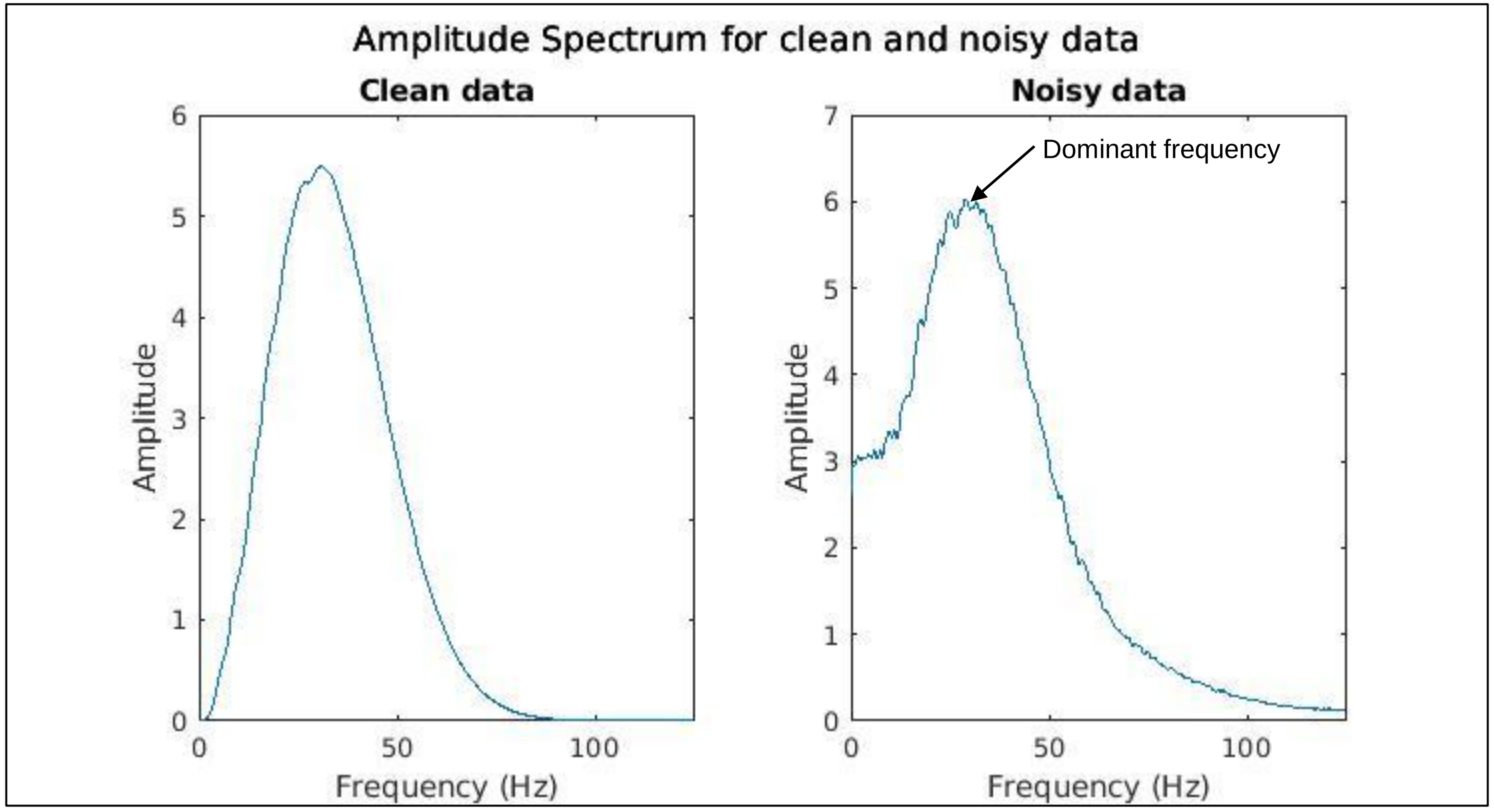


Synthetic Examples

After transforming the data into the F-XY domain, we can calculate the average amplitude spectrum of the data and estimate the dominant frequency of the seismic. In the following picture we can see how the amplitude spectrum of the signal changes when considering random noise:

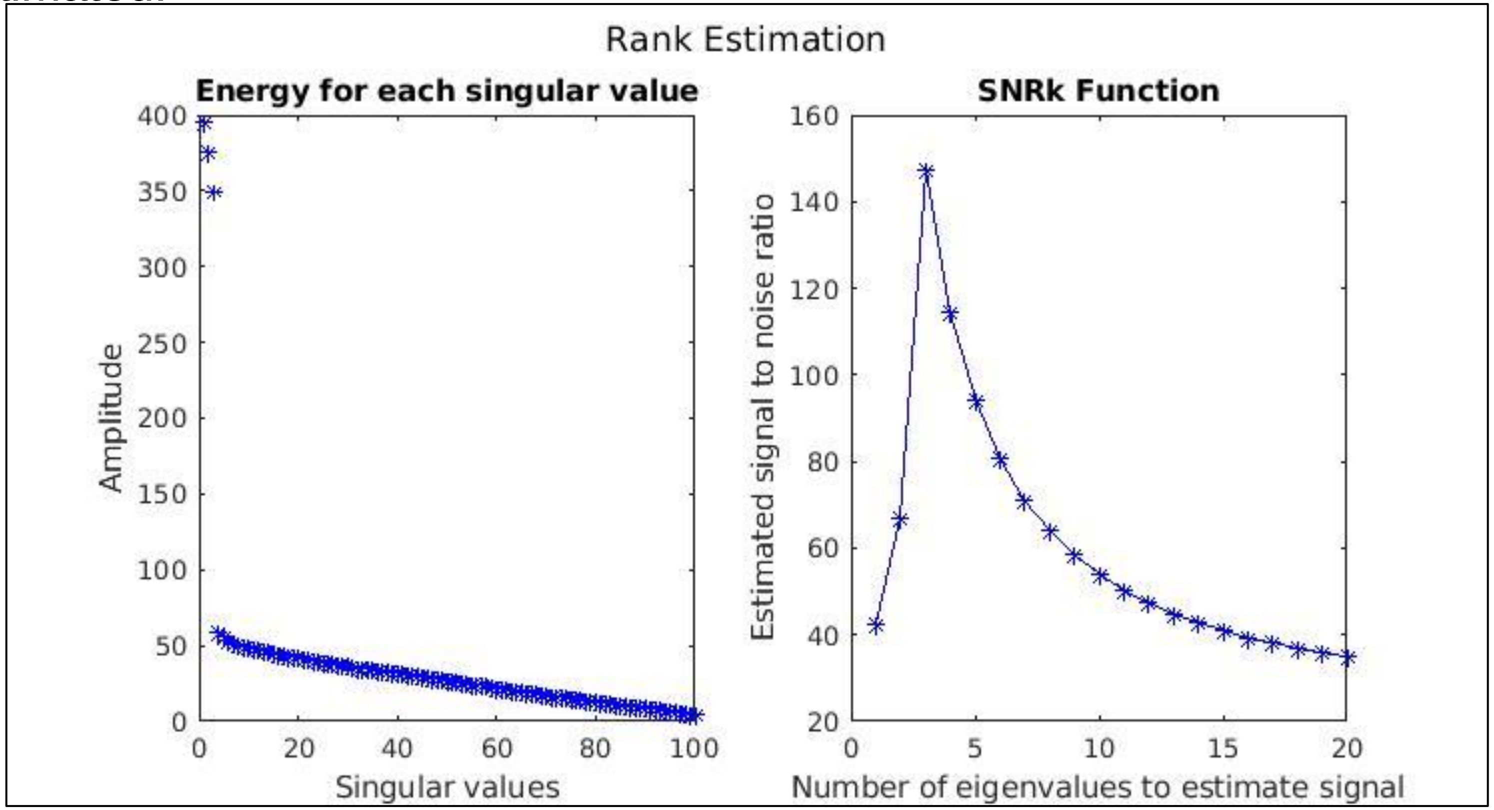
We see that the average amplitude spectrum of the seismic still has the overall behavior of the noise-free data. The dominant frequency can be found around 30Hz, which is the frequency of the Ricker wavelet used.

We still need to keep in mind that changing the SNR of the data may affect the accuracy of the dominant frequency estimation.



Synthetic Examples

The function SNR_k is calculated for the entire range of eigenvalues, and information about rank can be estimated:



In this case, the best signal to noise ratio can be obtained with a signal reconstructed with the first three singular values. This is consistent with the number of events existent in the seismic, as discussed by **Oropeza and Sacchi, 2011**, and other authors.

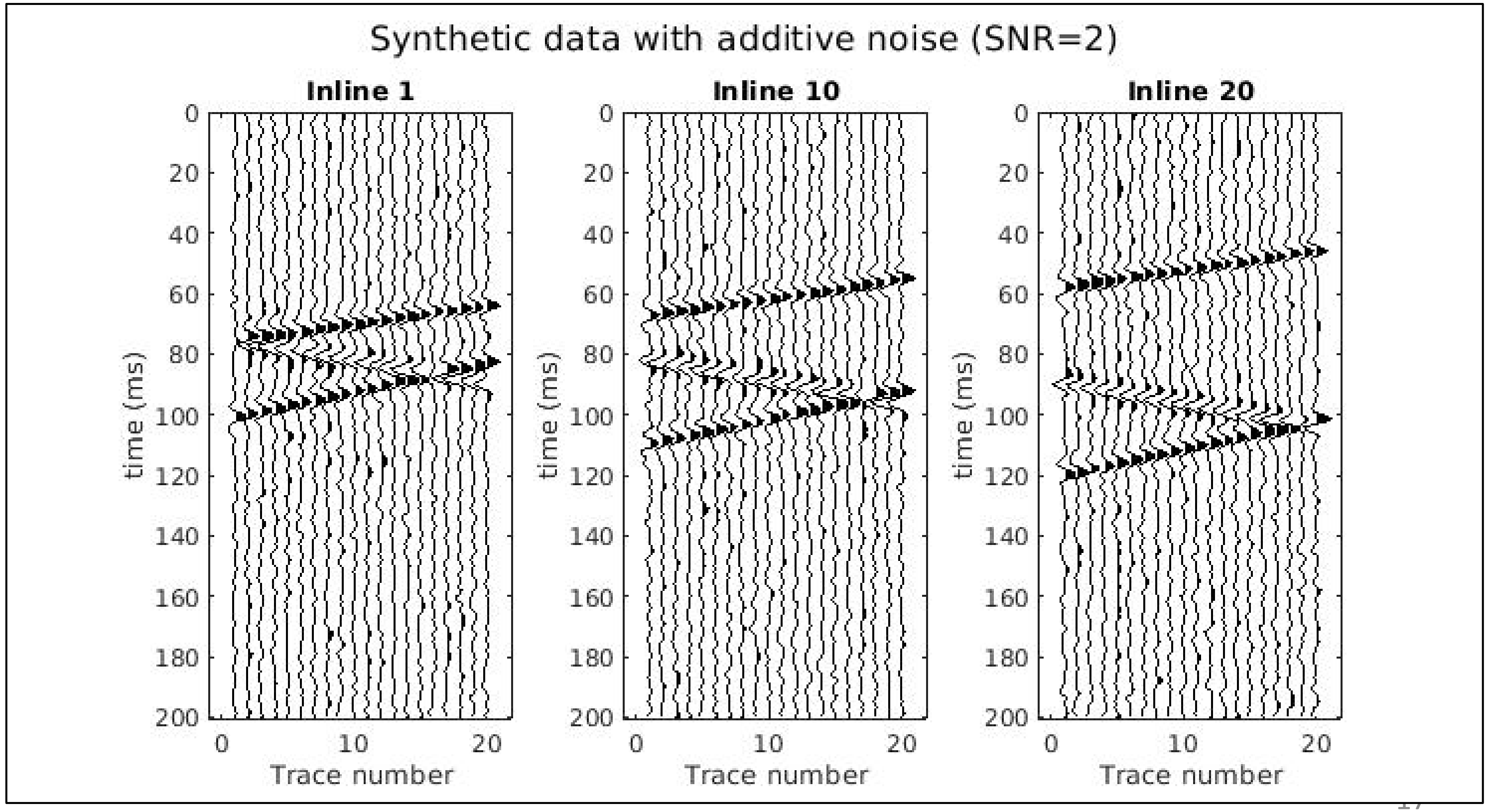
Later, we will discuss how the SNR affects the rank estimation.

The estimated optimal rank of the matrix is 3

Synthetic Examples

From the previous slide, we already know the k_{th} largest eigenvalues needed to reconstruct the data with the highest signal to noise ratio. Remembering the Eckart-Young theorem, the low rank-approximation that we get from the SVD, has the minimum error in terms of the Frobenius norm. After applying the low rank approximation, we can recover the filtered data:

Original data

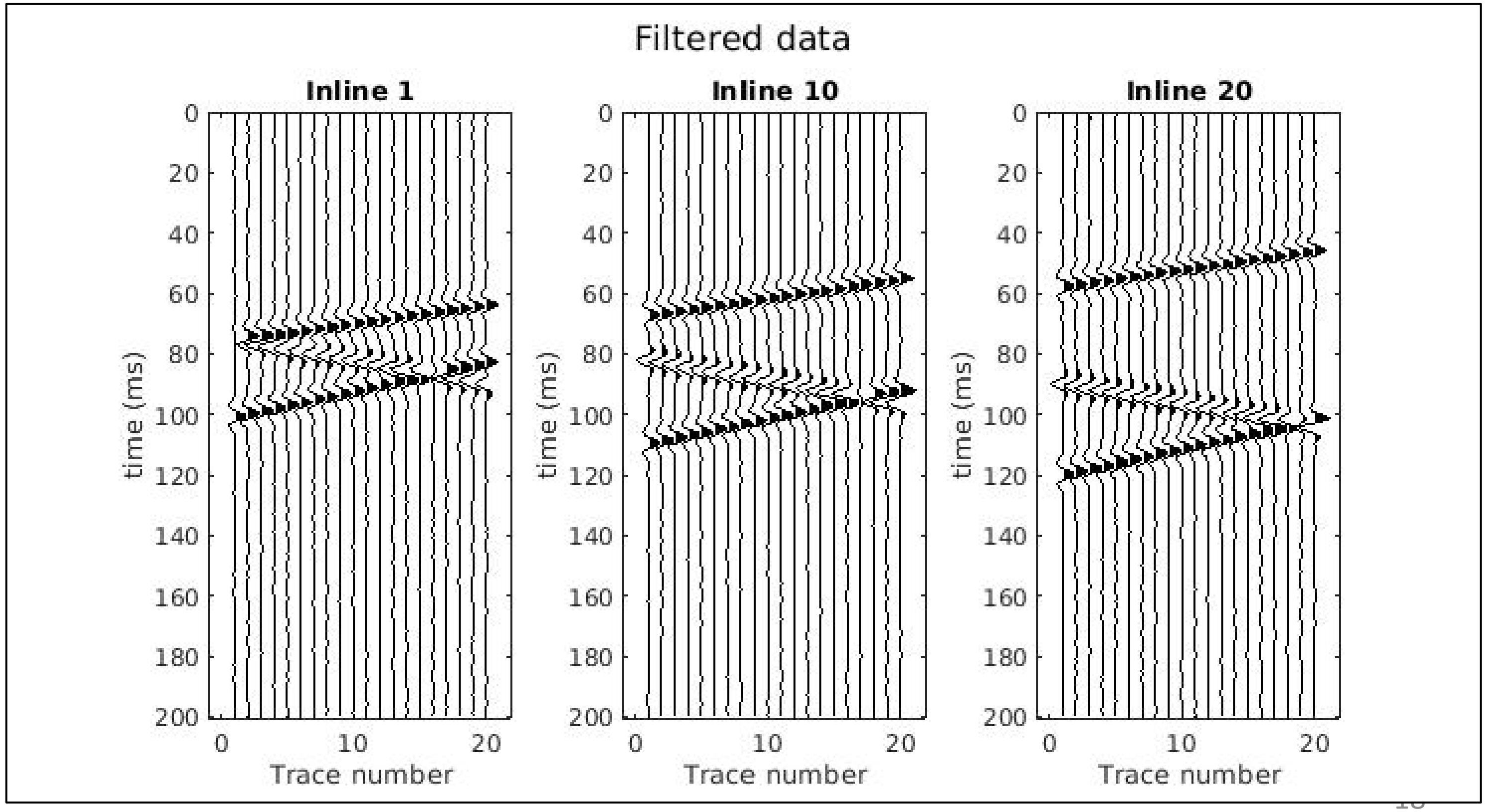


Synthetic Examples

From the previous slide, we already know the k_{th} largest eigenvalues needed to reconstruct the data with the highest signal to noise ratio. Remembering the Eckart-Young theorem, the low rank-approximation that we get from the SVD, has the minimum error in terms of the Frobenius norm. After applying the low rank approximation, we can recover the filtered data:

Filtered data

Notice the improvement in the SNR of the section. The energy of the noise has reduced to almost imperceptible levels.

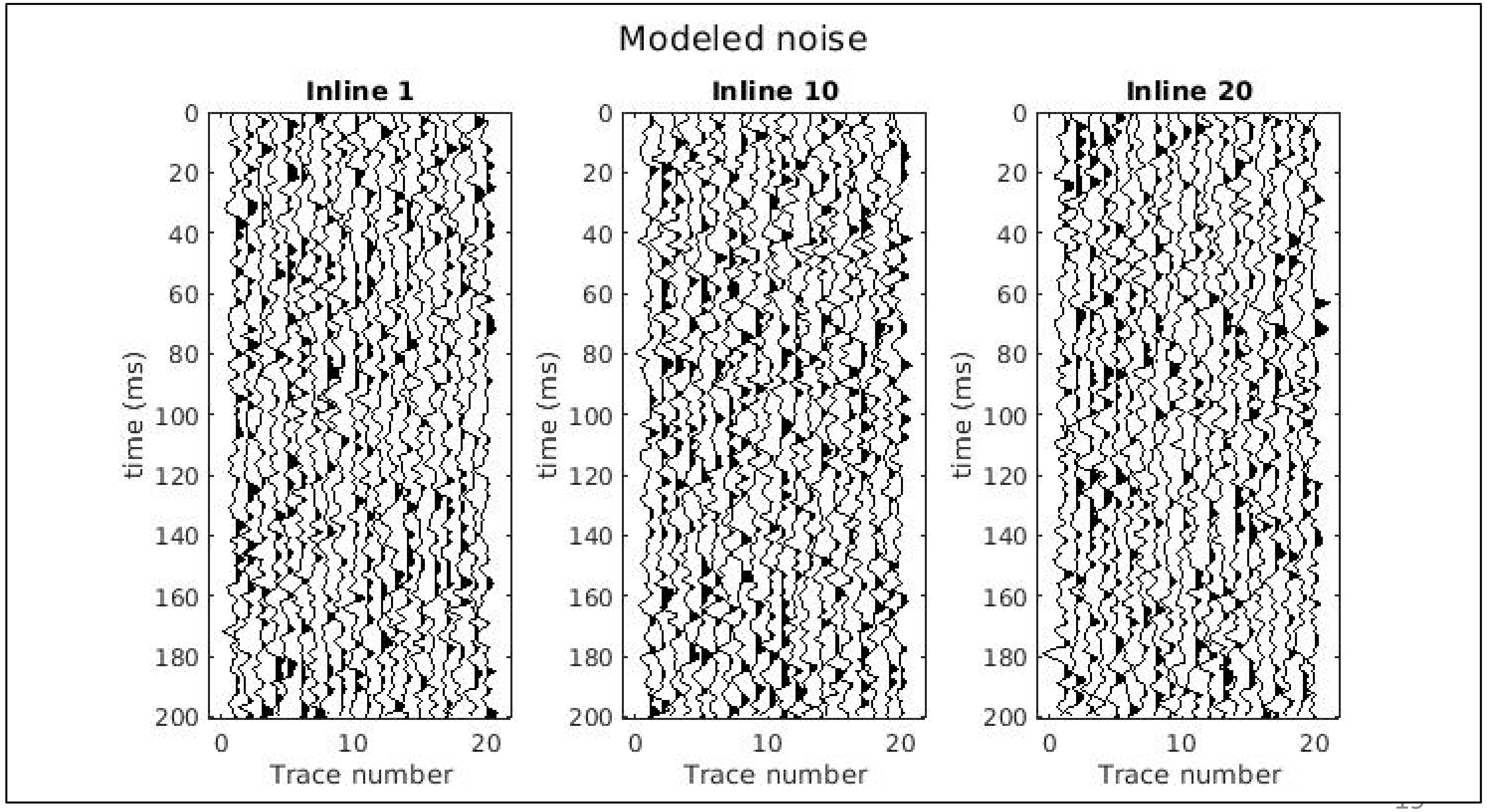


Synthetic Examples

From the previous slide, we already know the k_{th} largest eigenvalues needed to reconstruct the data with the highest signal to noise ratio. Remembering the Eckart-Young theorem, the low rank-approximation that we get from the SVD, has the minimum error in terms of the Frobenius norm. After applying the low rank approximation, we can recover the filtered data:

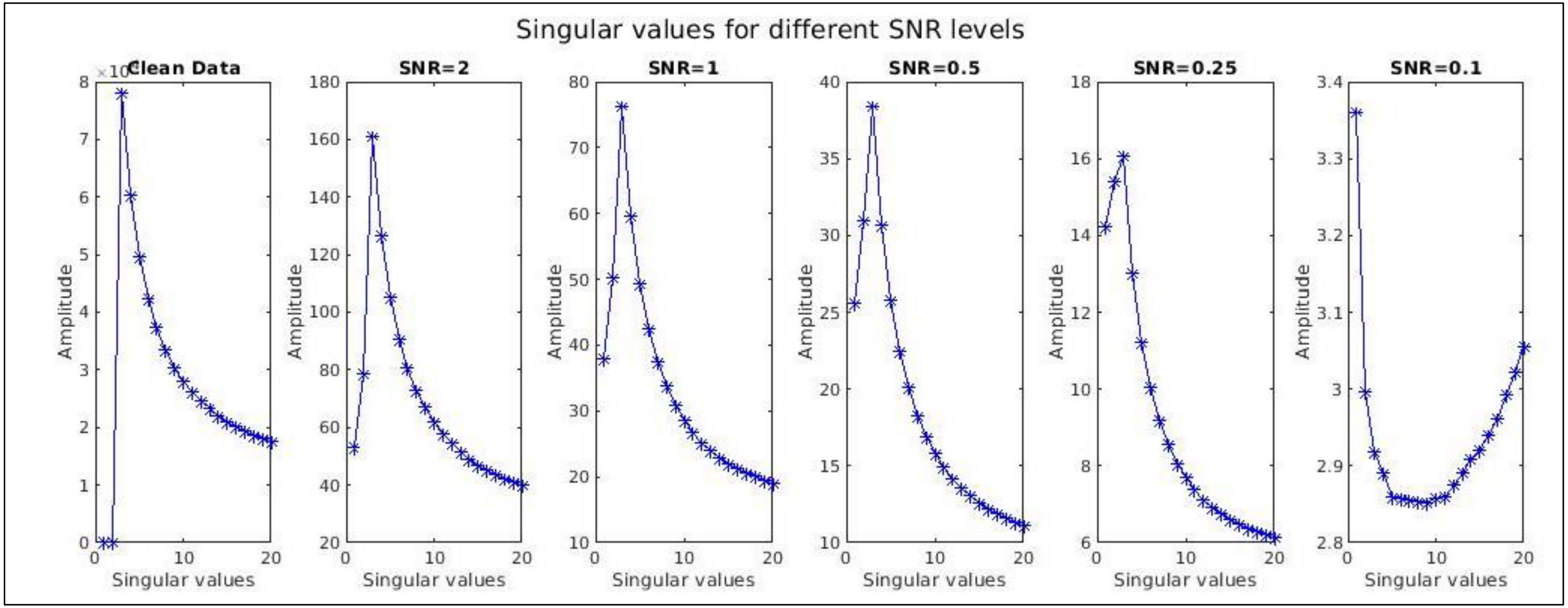
Modeled noise

In the noise model, there is no noticeable leakage of signal (Key QC step in Seismic Data Processing).



Synthetic Examples

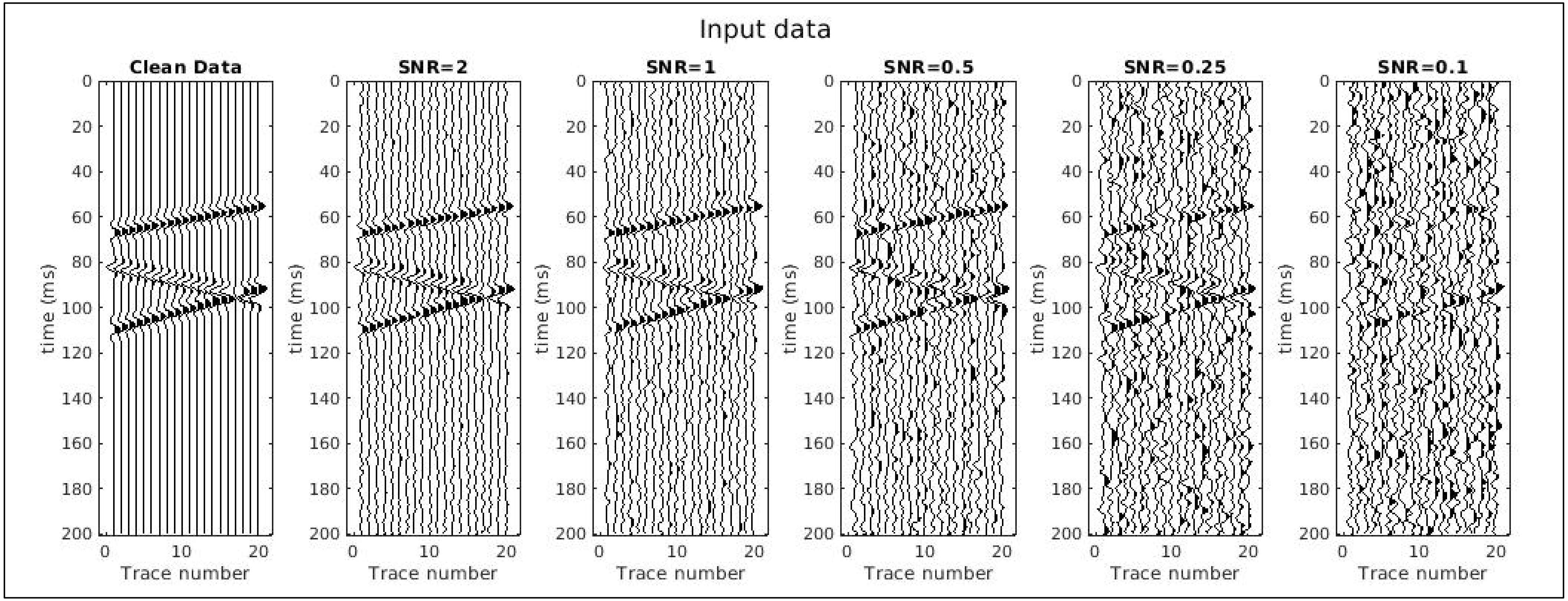
Since the SNR is decreasing, for which SNR we cannot longer estimate accurately the rank of the matrix?



For synthetic data, the method shows a high tolerance to low SNR ratios, being able to properly estimate ranks for SNR as low as 0.25 (4 times more noise than signal).

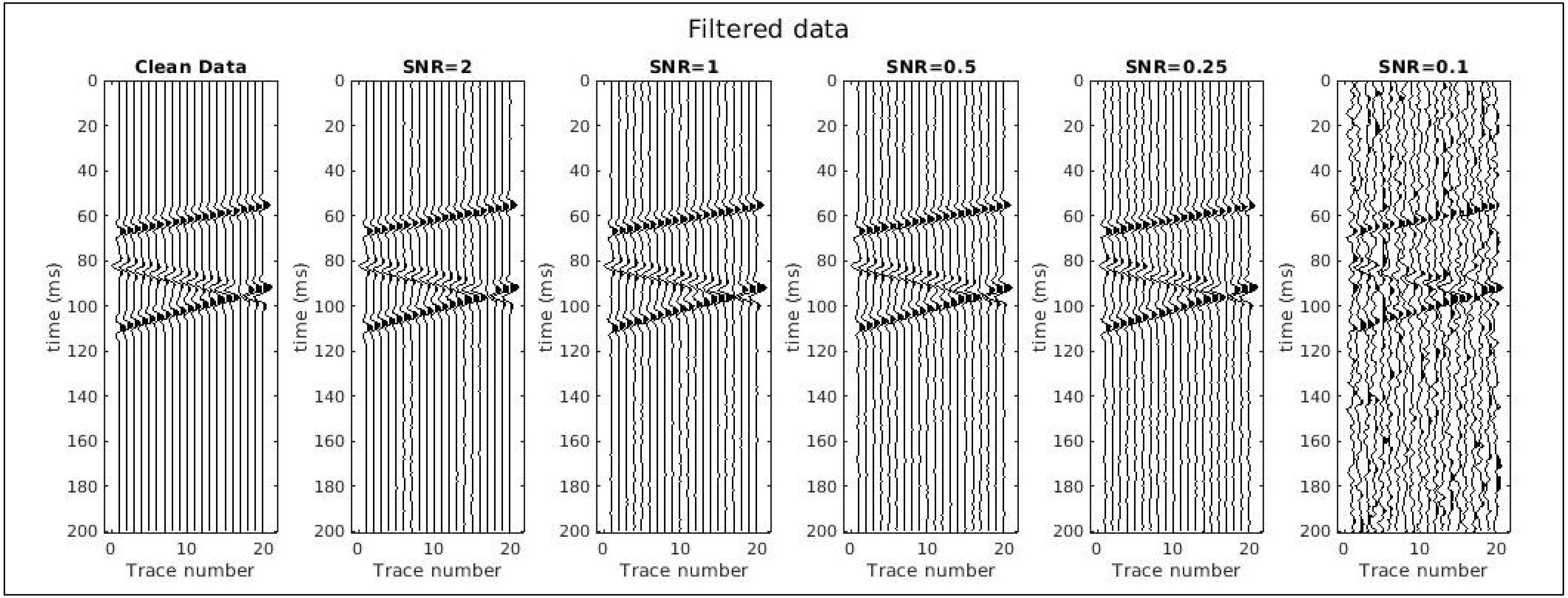
Synthetic Examples

Now let's look the result of applying the entire workflow to the same seismic data set, but with different SNR.



Synthetic Examples

Now let's look the result of applying the entire workflow to the same seismic data set, but with different SNR.

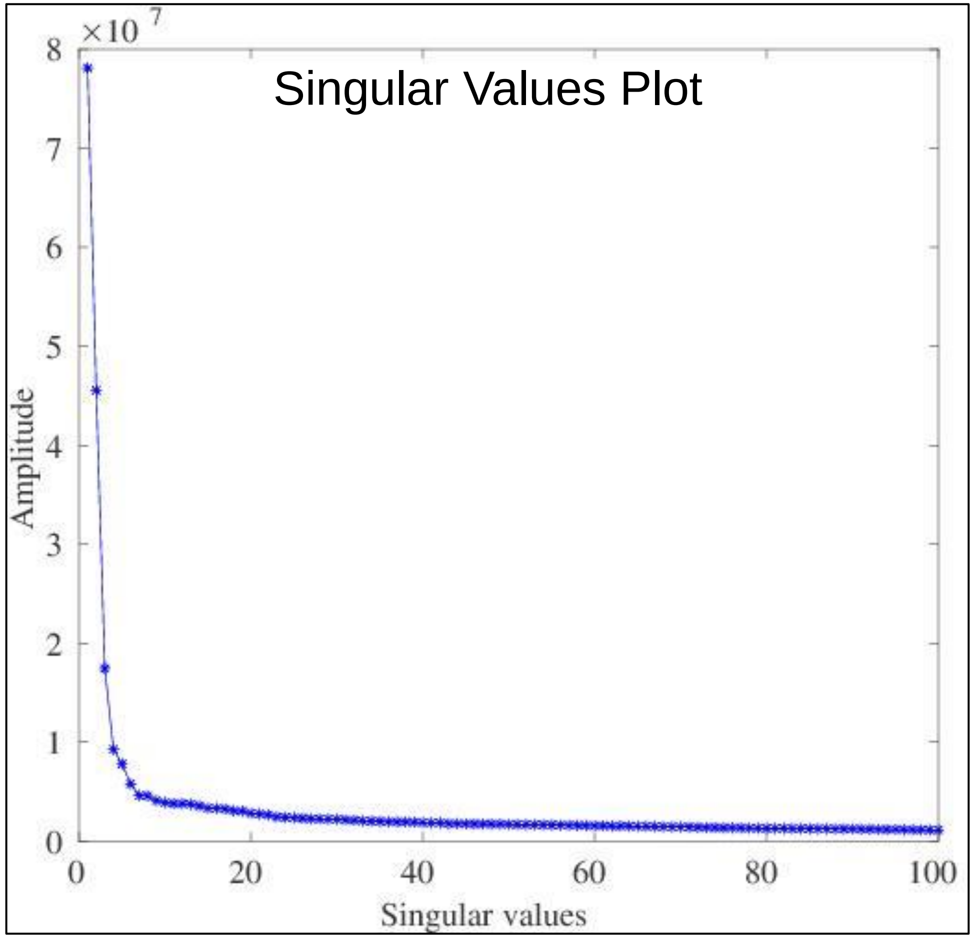


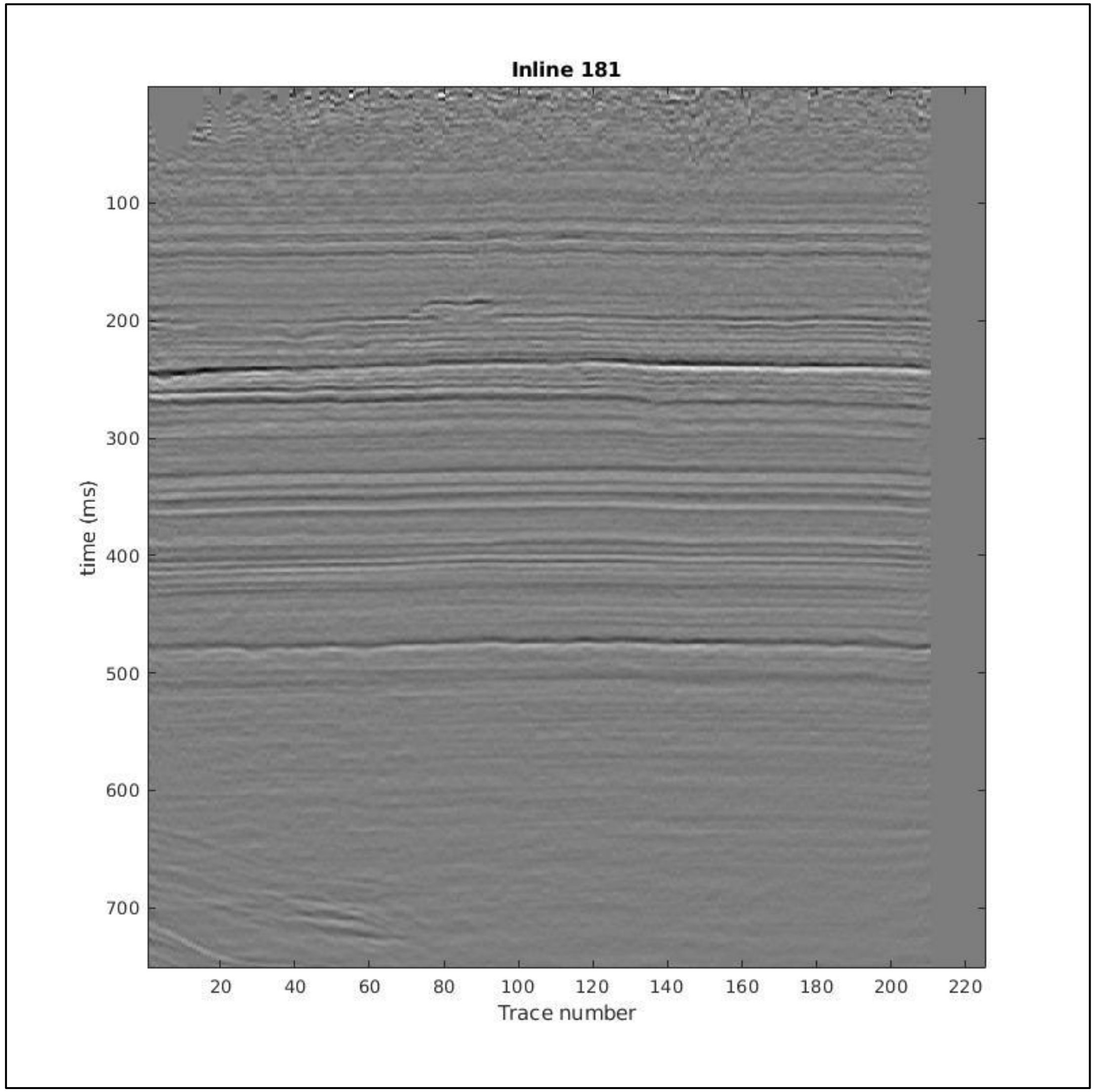
For synthetic data, we can recover good estimation of the input data, for SNR as low as 0.25 (four times more noise than signal).

Real data examples

Real data examples

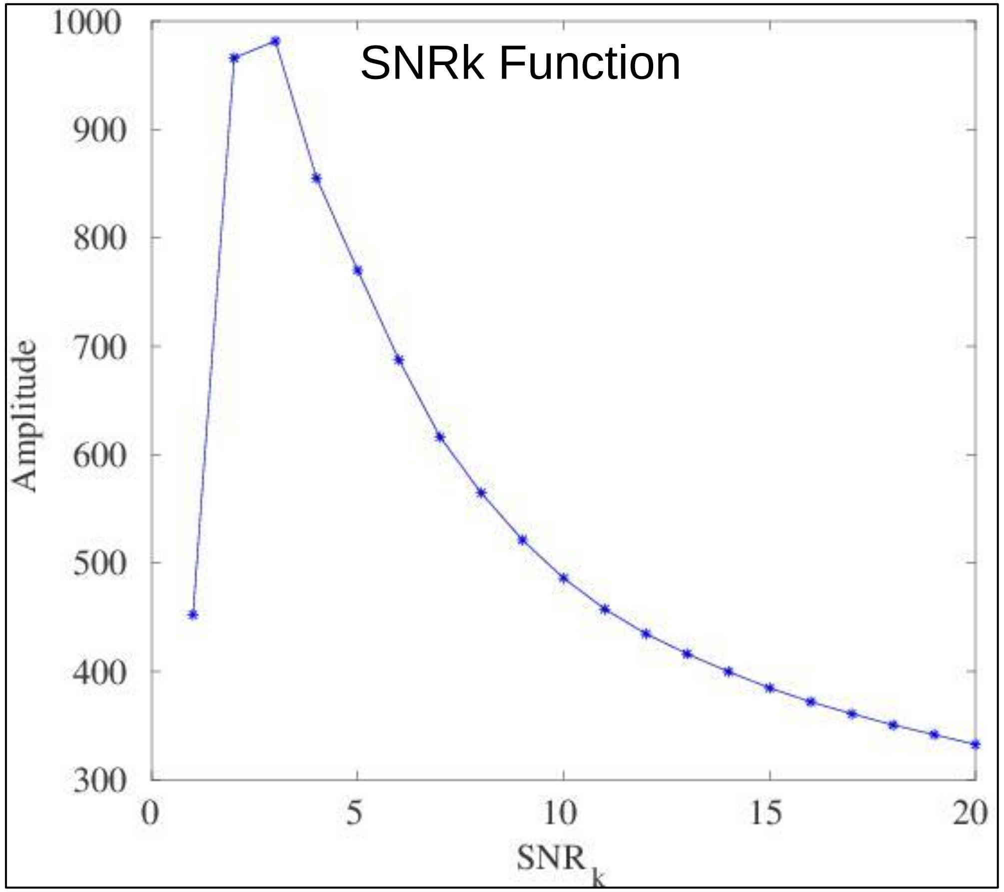
Let's consider the following post- stack seismic cube:

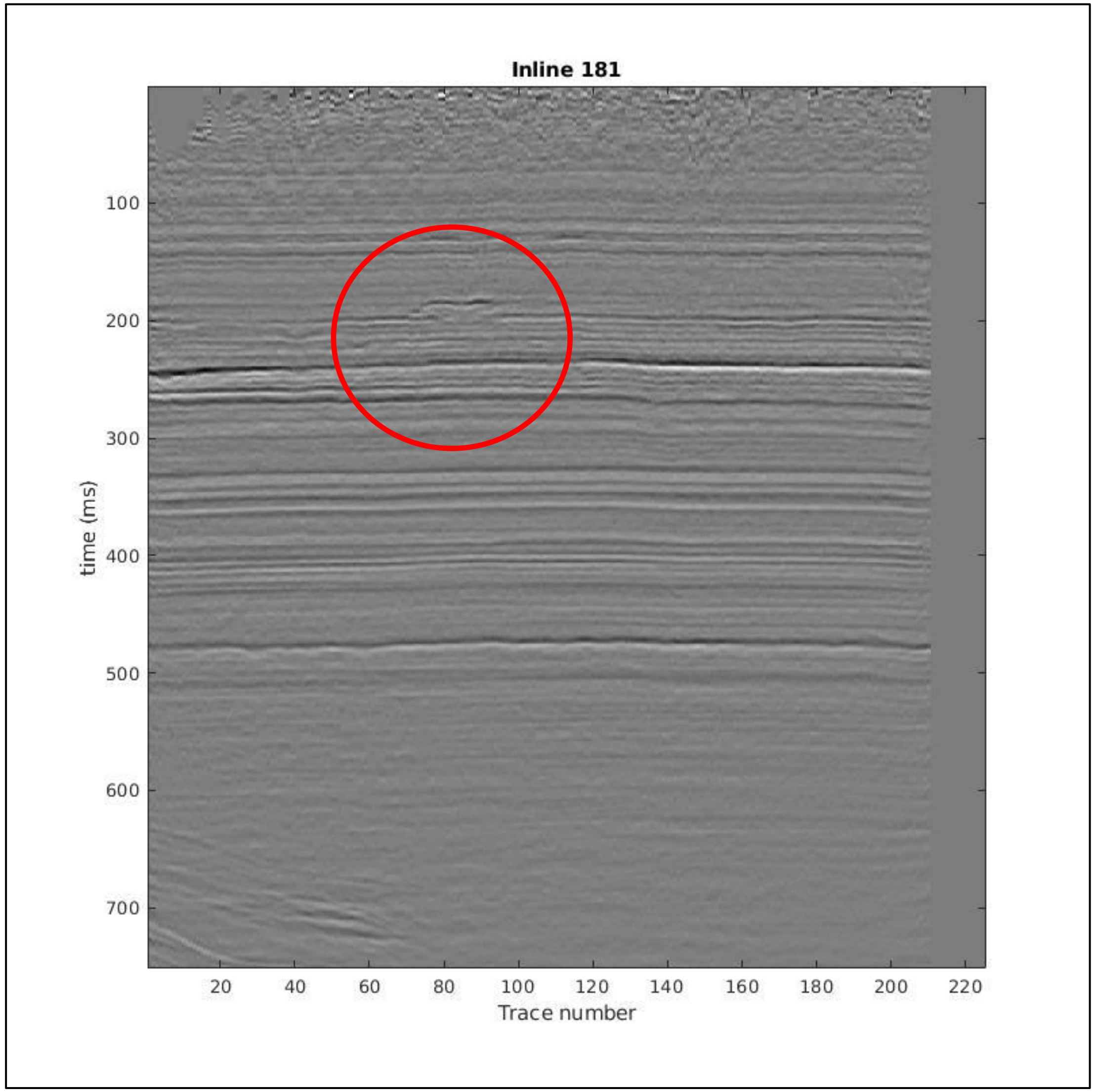




Real data examples

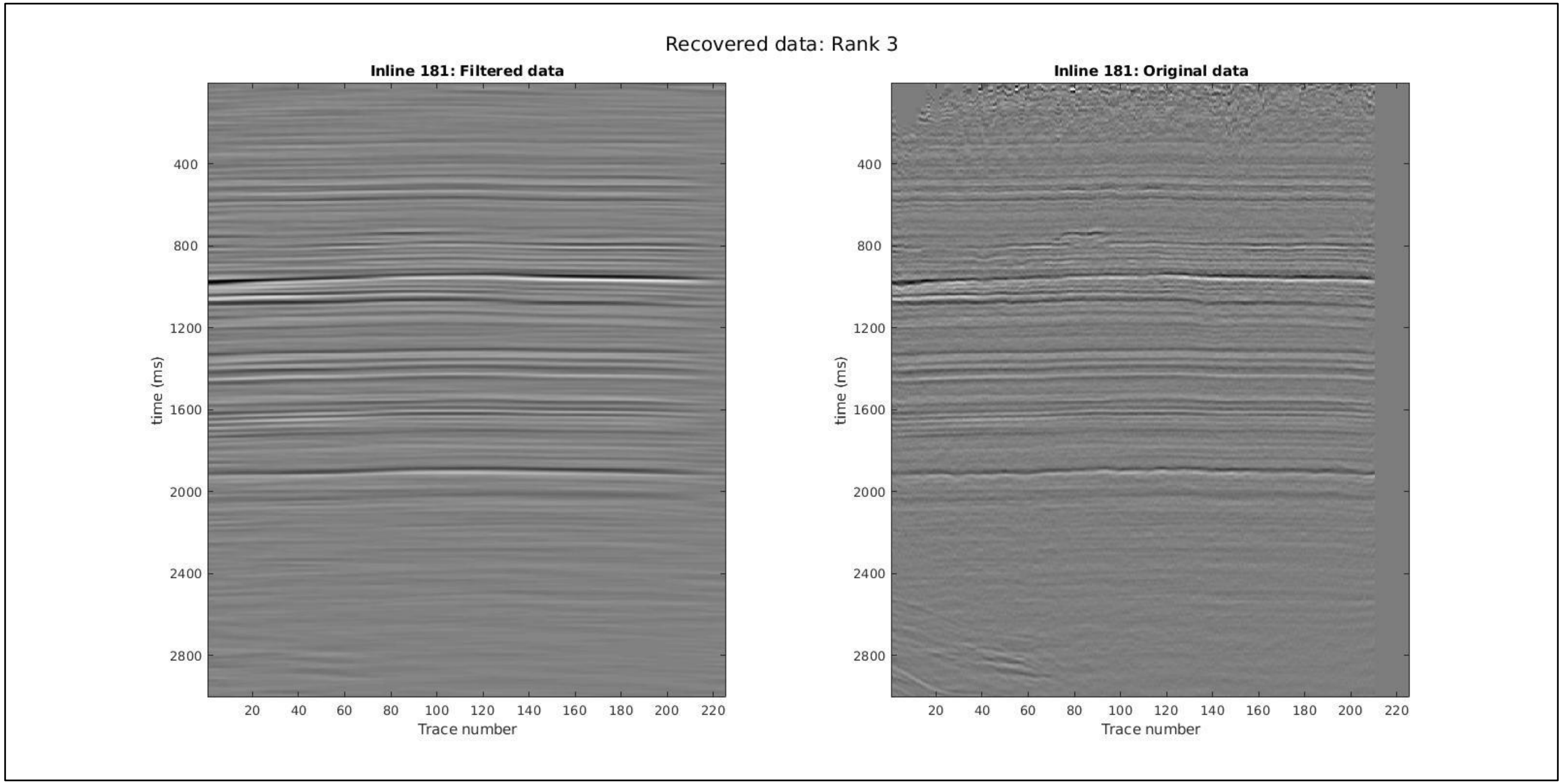
Let's consider the following post- stack seismic cube:





Real data examples

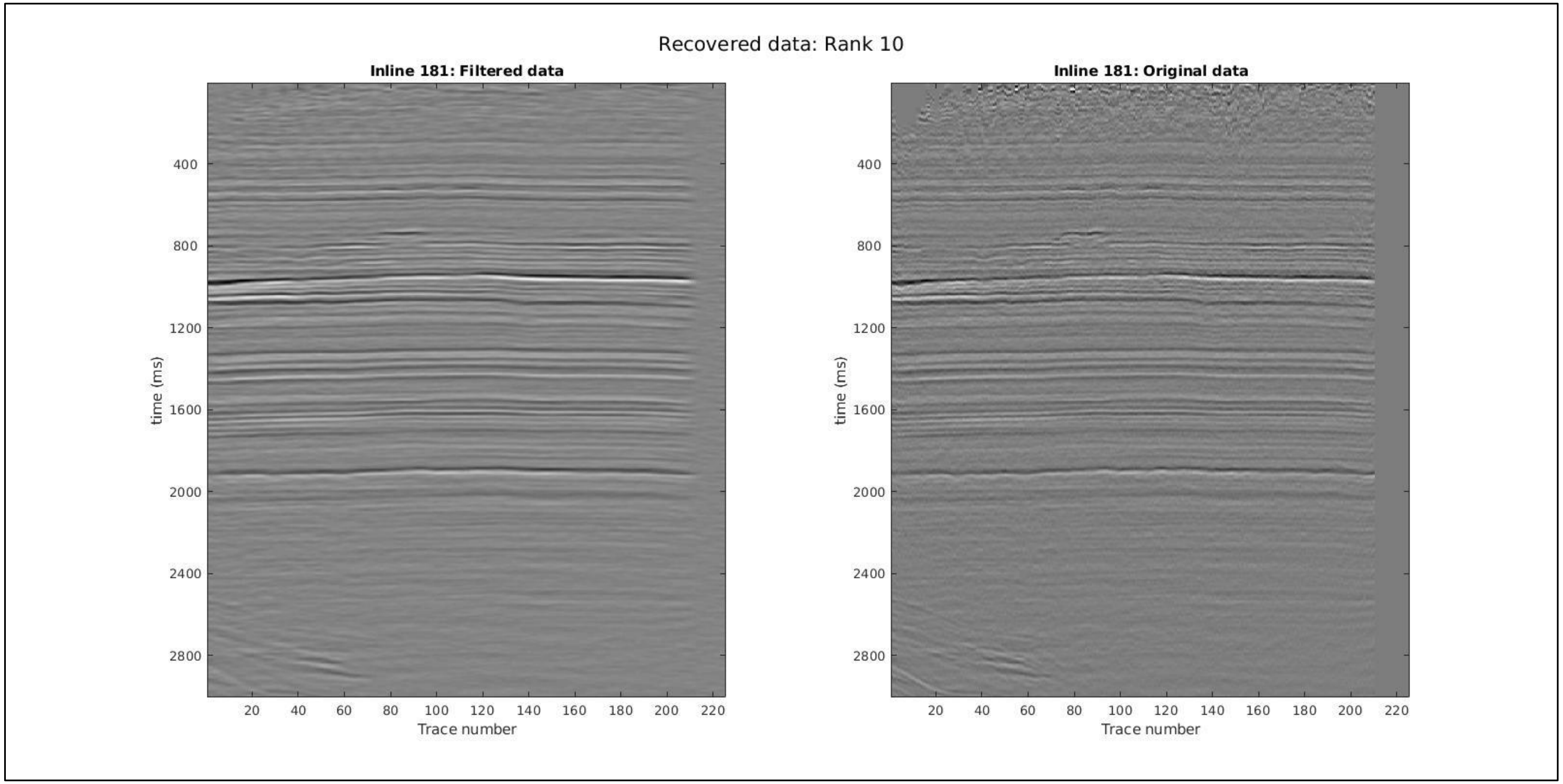
Let’s consider the following post- stack seismic cube:



With the **rank three** reconstruction, the low-frequency statics of the original data is preserved (there is structural consistency). But the energy of the channels is still spread over a larger area than in the original image; thus we can conclude that the result, even with better SNR, has a damaged lateral resolution.

Real data examples

Let’s consider the following post- stack seismic cube:



The **rank ten** reconstruction exhibits the real dips along the CDP axis. There is no noticeable spread of the energy on the channel (improved lateral resolution compared with previous results). The increase of SNR compared with the original data is noticeable.

Conclusions

We have proposed a method to estimate automatically the rank of a matrix based on an estimated signal to noise ratio. This method can be integrated with noise attenuation algorithms based on rank reduction. In this case, we applied the method within a Multichannel Singular Spectrum Analysis noise attenuation workflow.

We proved that for synthetic data, the rank of the matrix can be estimated accurately for SNR levels as low as 0.25 . At lower SNR values, the estimation of the dominant frequency starts to fail, and the energy contained in the noise level eigenvalues becomes too large to allow an accurate determination of the rank.

On the other hand, for real data, there is too much energy in the “noise level” of the singular values, so the rank estimator loses accuracy. The SNR_k function has to be adjusted for the real data case.

Further suggested work may include:

- Include into the SNR_k function a term depending on the slope of the “noise level”. As we saw in the discussion, the slope of the “noise level” increases as we increase de complexity of our input.