DOI:10.1190/image2022-3735804.1

Closing the loop between acquisition and processing: data-driven volumetric SNR estimation vs acquisition design predictions

Andrey Bakulin and Ilya Silvestrov, EXPEC Advanced Research Center, Saudi Aramco; Pierre Leger, Geophysical Imaging Department, Saudi Aramco

Summary

Designing seismic acquisition is a blend of science and art, especially in a desert environment with low signal-to-noise ratios (SNR) of -20 to -60 dB. Basic design driven by the square-root model may be overly simplistic. At the same time, conventional numerical feasibility studies are lengthy and suffer from processor bias. We propose a new workflow that can streamline the design process, remove human bias, and close the loop between acquisition and processing. A central element is a data-driven approach for deriving absolute SNR directly from the data itself. The resulting SNR volumes can be analyzed as volumes, slices, or statistical distribution. Mean SNR over the volume/subvolume can be compared directly with the predicted SNR from acquisition design, thus allowing to quantitatively close the loop, rather than blindly relying on simple theoretical predictions. For feasibility studies, complete automation can be achieved by applying migration in-lieu of processing and data-driven SNR as evaluation steps. We demonstrate the new approach on SEAM Arid data assessing several orthogonal 3D acquisition geometries with 9-geophone arrays and single sensors.

Introduction

The seismic acquisition represents the lion's share of geophysical expenditure. Consequently, a great deal of effort is going into tinkering with the seismic acquisition. However, final imaging results are often evaluated qualitatively, with human bias playing a significant role. In our view, such a process represents a significant impediment to efficient seismic imaging with an unbiased feedback loop between acquisition and processing. We present a novel data-driven approach that allows estimation of the absolute signal-to-noise ratio and enables direct comparison with the predictions made in acquisition design.

Acquisition design and signal-strength estimate

To establish a feedback loop between acquisition and processing, it is critical to relate the input variables of the acquisition directly to the quality of the seismic image. Considering that seismic image results from a combination of migration and stacking processes, Meunier and Gillot (2000) and Meunier (2011) put forward a very popular formula for a signal-strength estimate (SSE) providing an estimator of signal-to-noise ratio (SNR) of the final seismic image. SSE is derived through the simple square-root combination rule and is expressed as

$$SSE(f) = SS(f)\sqrt{SD\ NR\ RA} = SS(f)\sqrt{TD\ RA}\ \ ,$$

where RA is the area of the receiver station, NR the number of receivers per shot point (SP), SD is the source density (number of SP per surface unit), and SS is the source strength. The product of SD*NR is often referred to as trace density (TD), leading to the second simple expression above. It should be noted that SSE can be converted to theoretical SNR in decibels as $SNR^t_{dB} = 20\log_{10} SSE$.While based on a simple assumption of random noise, such a formula is powerful since it allows direct balancing of different trade-offs occurring in seismic acquisition between various parameters at play. It also incorporates direct effects of the receiver array allowing to compare legacy acquisition with source groups and newer acquisition with point sensors.

While acquisition design heavily relies on the SSE equation, it rarely gets to validate the ultimate accuracy of this formula. Processors, seldom familiar with the SSE equation, prefer to use their more complex metrics. They further tend to overestimate the role of the processing, believing that applying their preferred processing technique could overcome many restrictive mundane considerations embedded in SSE. In summary, every processor is biased towards their own qualitative assessment of the final image. Moreover, these assessments might significantly differ between various experts. As a result, the efficacy of different acquisition elements is often subject to vigorous debates making it hard to reach broader technical consensus and complicating management decisions.

What if we find a way to directly measure SNR of the final data volume using a method acceptable to processors and compatible in assumptions with the SSE equation? We demonstrate a workable method of doing that and for the first time directly evaluate processing outcomes and acquisition design predictions using SEAM Arid elastic data representative of challenging land seismic in the Middle East (Oristaglio, 2012; Regone, 2017).

Data-driven stack-based SNR method

We propose to use data-driven stack-based SNR method (Bakulin et al., 2021, 2022b). Each estimate of SNR requires a 3D data window consisting of superimposed signal and noise

$$d_{ij} = s_i + n_{ij},\ \ i = 1, \dots, N,\ \ j = 1,\ \dots, M,$$

where $s_i = s(t_i)$ is a signal, $n_{ij} = n(t_i, x_j)$ is noise, i is a time sample, j is a trace index. Then experimental SNR is computed by a simple formula using semblance S as an intermediate variable

$SNR^e = \frac{S}{1-S}$; $S = \frac{\sum_{i=1}^{N}(\sum_{j=1}^{M} d_{ij})^2}{M\sum_{i=1}^{N}\sum_{j=1}^{M} d_{ij}^2}$; $SNR^e_{dB} = 10\log_{10} SNR^e$.

A moving window runs throughout the 3D data volume, and the computed SNR is output at the center of the window. As a result, the 3D volume of SNR is generated for each data volume.

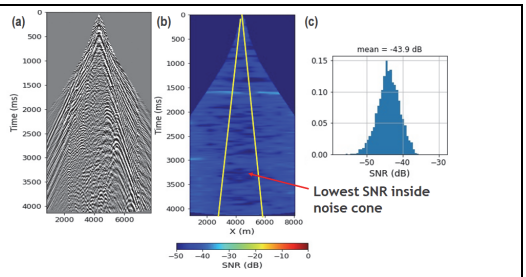

Figure 1: Raw field gather (a) with associated SNR panel (b) and histogram (c).

Application to SEAM Arid data

Let us demonstrate the proposed methodology using a realistic elastic SEAM Arid 3D model replicating challenges from the desert environment (Bakulin and Silvestrov, 2021). Following the Regone et al. (2015) approach, the processing sequence is restricted only to prestack depth migration of raw data volumes with the true velocity model. Figure 1a shows that pre-stack reflections are overwhelmed by near-surface arrivals. Using the stack-based SNR method, we estimate the average SNR of prestack data between -35 and -55 dB (Figure 1b,c). The strongest deep reflectors exhibit higher SNR around -20 dB, for instance, at 1.6 s or 2.5 km depth (Figure 1b). However, the majority of weaker reflectors possess much lower SNR. Note that for reliable estimation of SNR of -40 dB, the ensemble has to contain at least 10,000 traces (Bakulin et al., 2022b).

Let us consider six 3D acquisition geometries listed in Table 1. Geometries G1-G3 resemble high-channel-count (HCC) data conventionally acquired with 9-geophone arrays in a desert environment (Dmitriev et al., 2017). Geometries G5-G6 represent emerging point-receiver surveys acquired with single sensors (SS). Geometry G4 represents Narrow-Azimuth HCC that reduces cross-line offset to ¼ of the inline offset (1.75 km vs. 7 km). It has the same trace density as G3 but has twice as dense a shot spacing in both directions. So it provides a finer sampling of near and mid offsets but compromises the wide-azimuth coverage at far offsets. Figure 2 displays migrated vertical sections from full PSDM volumes for all geometries. The previous study (Bakulin and Silvestrov, 2021) has analyzed one geometry G1 and found it was unable to properly image shallow subsurface (700-2000 m) with or without conventional processing. Figure 2 suggests that increasing density can eventually resolve imaging in the shallow part, however exact progression is non-uniform, and arrays play an important role. Instead of subjective qualitative discussion, let us focus on a systematic comparison of predicted and measured SNR metrics to evaluate the results and, in the process, also validate the usefulness of the square-root equations used for acquisition design.

SNR volumes

The second row of Figure 2 shows cross-sections from SNR volumes. Bakulin et al. (2022b) demonstrated that the stack-based method estimates absolute SNR instead of the SSE equation that only conveys relative SNR. Volumetric representation of SNR is useful for all kinds of analysis. A simple glance at Figure 2 suggests two different "regimes" of lower SNR in the shallow part (700-2000 m) and higher SNR in the deeper part (2000-3750 m). While shallow and deep reflections may have a similar level of noise (Figure 1a), the deeper part is characterized by significantly higher reflectivity driving higher SNR (Bakulin and Silvestrov, 2021). The SNR volume is essentially a calibrated version of the semblance, so coherent events become clearly visible as stripes of higher SNR. As a result, the numerical value of SNR also carries interpretative significance. To demonstrate that, we further present histograms of SNR distribution for shallow and deeper parts (third and fourth rows in Figure 2). While it is possible to focus on a specific event and analyze SNR at the target level, such distributions enable the assessment of volumetric imaging quality over an extensive vertical depth window. For example, the mean value of each histogram for 700-2000 m windows is now posted in Figure 2. The mean value varies from -1.2 dB for G1 to 8.7 dB for G6. A value around 0 dB means that signal and noise are comparable. Negative SNR values (Figure 2a-d) relate well to visual perception of the actual images with missing (-1.2 dB, Figure 2a), ambiguous (1.8 dB, Figure 2b), or intermittent (3.9 dB, Figure 2c) reflections. It appears that a level close to ~5 dB may be a prerequisite for robust structural interpretation. Out of all the volumes, only G4 and G6 cross that threshold and exhibit crisp boundaries and robustly mapped channels shown by the yellow arrow.

Comparison of measured and predicted SNR

Let us now contrast the predicted and measured SNR values. Since the acquisition formula predicts a single "average" value of SNR, let us compare them with the mean values from measured SNR values. Besides, let us do a separate analysis for the shallow and deep parts. Table 1 shows such a comparison. SSE formula only predicts relative SNR between different acquisitions. In addition, the average SNR is clearly different between shallow and deep portions. Therefore, to "calibrate" the predicted and measured SNR scales, let us align them for the last geometry G6 so $SNR^t_{dB} = SNR^e_{dB}$. As such, we cannot make any useful comparison for G6 itself but now can quantitatively relate G6 to the remaining four geometries. For example, going from G6 to G5, shallow SNR drops by 5.8, while prediction

DOI:10.1190/image2022-3735804.1

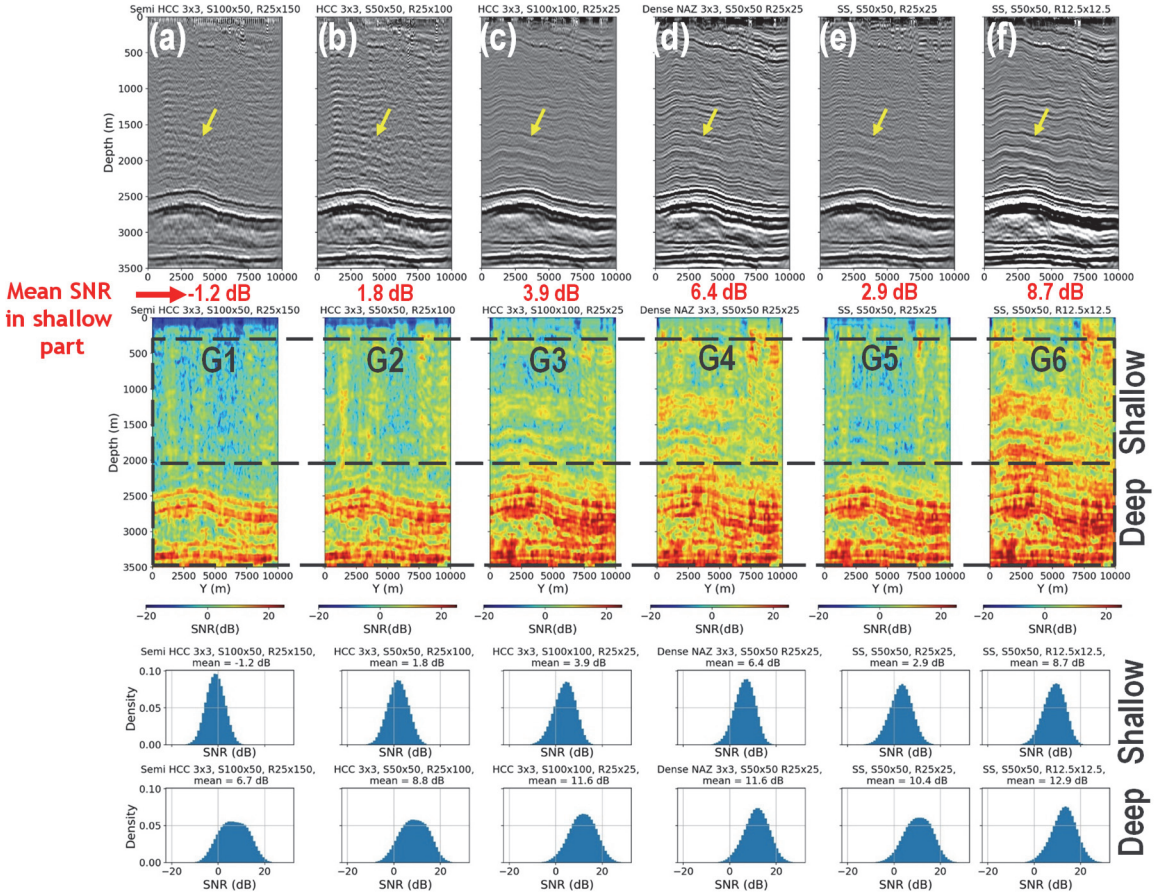

Figure 2: Cross-sections from a 3D PSDM volume (top row) and SNR volumes (second row) obtained by migrating raw data with different acquisition geometries: (a) Semi-HCC (3x3 geophone array) S100x50m, R25x150m; (b) HCC (3x3 geophone array) S50x50m, R25x100m; (c) Nodal HCC (3x3 geophone array) S100x100m, R25x25m; (d) Narrow-azimuth HCC (maximum crossline offset is ¼ of inline offset) S50x50m, R25x25m; (e) SS (intermediate) S50x50m, R25x25m; (f) SS (dense) S50x50m, R12.5x12.5m. Two lowest rows show histograms of measured SNR computed over the shallow 700-2000 m (third row) and deeper 2000-3750 m (forth row) 3D subvolumes.

Name	**Short name**	**Shot spacing**	**Receiver Spacing**	**Trace density**	**SSE**	SNR^t_{dB} (shallow)	SNR^e_{dB} (shallow)	SNR^t_{dB} (deep)	SNR^e_{dB} (deep)
Semi-HCC (3x3 geophone-arrays) (SD=1, NR=1, RA=9)	G1	100x50 m	25x150m	1	3	1.5dB	-1.2 dB	5.7dB	6.7 dB
HCC (denser, 3x3 geophone-arrays) (SD=2, NR=1.5, RA=9)	G2	50x50 m	25x100m	3	5.2	6.3 dB	1.8 dB	10.5 dB	8.8 dB
Nodal HCC (nodal, 3x3 geophone array) (SD=0.5, NR=6, RA=9)	G3	100x100m	25x25m	3	5.2	6.3 dB	3.9 dB	10.5 dB	11.6 dB
NAZ nodal HCC (3x3 geophone array) (SD=2, NR=1.5, RA=9)	G4	50x50m	25x25m	3	5.2	6.3 dB	6.4	10.5 dB	11.6 dB
Nodal SS (intermediate) (SD=2, NR=6, RA=1)	G5	50x50m	25x25m	12	3.5	2.8 dB	2.9 dB	7 dB	10.4 dB
Nodal SS (dense) (SD=2, NR=24,RA=1)	G6	50x50m	12.5x12.5m	48	6.9	8.7 dB reference	8.7 dB reference	12.9 dB reference	12.9 dB reference

Table 1: Summary table with evaluated geometries and their theoretical (SNR^t_{dB}) and experimentally measured (SNR^e_{dB}) signal-to-noise ratios. To calibrate predicted and measured dB scales, we assumed $SNR^t_{dB} = SNR^e_{dB}$ for the last geometry G6.

also suggests a close value of 5.9 dB. Cross-plotting predicted and measured SNR demonstrates good agreement between the two sets, both clustering around $SNR^t_{dB} = SNR^e_{dB}$ line with reasonable deviations.

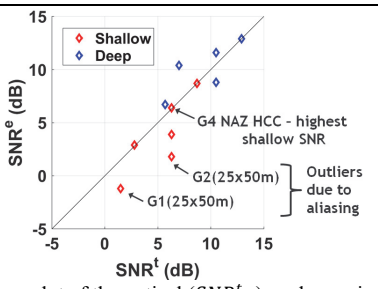

Figure 4: Cross-plot of theoretical (SNR^t_{dB}) and experimentally measured (SNR^e_{dB}) signal-to-noise ratios from Table 1.

More detailed conclusions could be summarized as follows:
1) acquisition sampling and array size both play an equally important role as predicted;
2) two orthogonal directions out of the four acquisition directions must be well sampled (address aliasing) for successful imaging. A similar observation was done by Regone et al. (2015). This study suggests that such a threshold sampling for aliasing is 25 m for data with arrays and 12.5 m for single-sensor data in this particular case.
3) The SSE equation provides a reasonable prediction but fails to describe the effects of aliasing that impedes noise reduction. For example, Figure 4 shows that measured SNR for G1 and G2 represent the furthest outliers because only one direction is well sampled in the domain of best sampling in those geometries (25x50m in the cross-spread domain). Also, G2 and G3 have the same SSE. However, when two directions become well-sampled in G3 (25x25m) – SNR consistently increases by 2.1 dB and 2.8 dB accordingly for shallow and deep (Table 1). Regone et al. (2015) also noted reduced noise reduction during migration of aliased data.
4) G4 (narrow-azimuth HCC) shows the highest shallow SNR of all three geometries (G2-G4) having the same SSE. We interpret that it provides an extra sampling inside the inner noise cone with the lowest SNR (Figure 1b), thus delivering uplift where it is needed most. While the lack of larger offsets may compromise the deeper part, the average deep SNR of G4 and G3 appear identical.

Effect of arrays

The SSE equation suggests that SNR is controlled by the product of trace density and receiver array size. The measured SNR from Table 1 and Figure 4 supports these conclusions. Suppose we desire single-sensor data of equivalent SNR. In that case, we should increase trace density by a factor of nine (for 9-geophone arrays considered here). When instead, we increase trace density by a factor of four by going from G3 to G5, we observe a drop in measured SNR by 4 dB for shallow and 2.5 dB for deep. This is compared with an expected decline of 3.5 dB. A reduced shallow SNR is visually evident when comparing Figure 2c and 2e. Single-sensor geometry G6 with 16 times the trace density of G3 can finally outperform arrayed HCC data of G3 as manifested by an increase in SNR of 4.8 and 1.3 dB for shallow and deep, respectively. These values are close to an expected improvement of 2.4 dB from the SSE equation. While the agreements are not perfect, they show good performance of the simple SSE equation on a rather complex realistic synthetic data at hand. The effect of arrays could be larger for real data due to multiplicative scattering noise (Bakulin et al., 2022a), not included in the simple SSE model above. This perhaps can explain the more significant drop in SNR in real single-sensor data than expectations.

Conclusions

We presented a practical approach to measuring volumetric SNR on imaged seismic data using a stack-based method. We applied it to the SEAM Arid dataset and showed that the volumetric SNR metric captures observed data quality and eliminates any human bias in the assessment process. We further applied this to close the loop between acquisition and processing and contrasted measured SNR against predictions done by the square-root model used during acquisition design. We found overall reasonable agreement between predicted and measured SNR, confirming the usefulness of the SSE model. Please note that the synthetic data at hand possess only "geological near-surface noise" and does not contain any actual random additive noise as hypothesized by the SSE model. Nevertheless, reduction of such organized near-surface noise during stacking or migration appears to reasonably follow a theoretical square-root model. Noticeable deviation seems to occur when three acquisition directions remain aliased. However, when two orthogonal acquisition directions out of four become non-aliased, noise reduction seems to follow the SSE equation quite well. This is in agreement with current practices of orthogonal acquisition designs used in the industry. Since the lowest SNR occurs at near-mid offsets, dense narrow-azimuth geometries may provide immediate uplift addressing this issue. Suggested workflow opens the way to a more automated feasibility study of different acquisition geometries based on data-driven SNR estimation on depth-migrated volumes. In addition, the data-driven SNR approach allows tackling more complex phenomena such as multiplicative noise caused by small-scale near-surface scattering that plays a small role in the SEAM Arid model at hand. We believe such multiplicative noise causes tremendous additional complexity, which explains why single-sensor data from the desert environment is significantly more challenging than expected.

Acknowledgments

We thank Maxim Protasov (Institute of Petroleum Geology and Geophysics) for the assistance in the development and implementation of the stack-based SNR method.

REFERENCES

Bakulin, A., and I. Silvestrov, 2021, Understanding acquisition and processing challenges in the desert environment through SEAM Arid and Barrett models: First International Meeting for Applied Geoscience & Energy, SEG/AAPG, Expanded Abstracts, 2824–2829, doi: https://doi.org/10.1190/segam2021-3583002.1.

Bakulin, A., I. Silvestrov, and M. Protasov, 2021, Quality control of 3D prestack land seismic data with a focus on data enhancement: First International Meeting for Applied Geoscience & Energy, SEG/AAPG, Extended Abstracts, 2929–2934, doi: https://doi.org/10.1190/segam2021-3576612.1.

Bakulin, A., I. Silvestrov, and M. Protasov, 2022, Signal-to-noise ratio computation for challenging land single-sensor seismic data: Geophysical Prospecting, **70**, 629–638, doi: https://doi.org/10.1111/1365-2478.13178.

Dmitriev, M., A. Bakulin, and P. Golikov, 2017, Efficient four-dimensional supergrouping algorithm for enhancement of high-channel count seismic data: 87th Annual International Meeting, SEG, Expanded Abstracts, 4986–4990, doi: https://doi.org/10.1190/segam2017-17668141.1.

Meunier, J., 2011, Seismic acquisition from yesterday to tomorrow: SEG and EAGE, 2011 Distinguished Instructor Course, Distinguished Instructor Series, No. 14.

Meunier, J., and E. Gillot, 2000, 3D seismic survey design: A solution: First Break, **36**, 176–179.

Oristaglio, M., 2012, SEAM Phase II—Land seismic challenges: The Leading Edge, **31**, 264–266, doi: https://doi.org/10.1190/1.3694893.

Regone, C., M. Fry, and J. Etgen, 2015, Dense sources vs. dense receivers in the presence of coherent noise: A land modeling study: 85th Annual International Meeting, SEG, Expanded Abstracts, 12–16, doi: https://doi.org/10.1190/segam2015-5833924.1.

Regone, C., J. Stefani, P. Wang, C. Gerea, G. Gonzalez, and M. Oristaglio, 2017, Geologic model building in SEAM Phase II—Land seismic challenges: The Leading Edge, **36**, 738–749, doi: https://doi.org/10.1190/tle36090738.1.