Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

Pixel-Level Uncertainty Quantification in CCS Seismic Monitoring via Monte Carlo Diffusion Sampling: Application to the Sleipner 4D Benchmark

Sunil Manikani[1], Dr. Murad Yaghi[2]

Doctoral Researcher, Applied AI in Geosciences
Carbon Capture & Sequestration Intelligence Research

SUMMARY
We introduce Monte Carlo Diffusion Sampling (MCDS), a method for generating pixel-level uncertainty maps from a pretrained diffusion representation model (PDRM) without modifying model architecture, adding dropout, or training an ensemble. By running N independent stochastic passes through the PDRM at a fixed intermediate noise timestep t, each with a fresh random noise realisation, we obtain a set of plausible seismic reconstructions whose pixel-wise standard deviation constitutes an uncertainty map. Applied to all 249 inline sections of the Sleipner 4D CO2 seismic benchmark, MCDS reveals: (1) a global mean pixel uncertainty sigma-bar = 0.0348 with narrow cross-inline variation (std = 0.0011), indicating consistent model calibration across the lateral extent of the survey; (2) substantial within-inline spatial heterogeneity, with peak pixel uncertainty reaching 0.22 (6.3x the global mean), concentrating on structurally complex amplitude zones; (3) strong Pearson correlation (r = 0.738) between inline amplitude heterogeneity and mean inline uncertainty, confirming that the PDRM is uncertain precisely where the seismic data is most complex; and (4) a spatially coherent elevated-uncertainty cluster at inlines 174-184, coinciding with the highest amplitude variability in the survey. We argue that MCDS probability maps provide a practical, annotation-free monitoring diagnostic for CCS operations: anomalous uncertainty excursions between seismic vintages may flag zones where CO2 migration has shifted seismic character outside the model training distribution.

1. INTRODUCTION

Uncertainty quantification (UQ) is a foundational challenge in subsurface geoscience. Seismic amplitude data is inherently ambiguous: the same reflection pattern can arise from different rock and fluid combinations, acquisition noise corrupts fine-scale detail, and processing introduces correlated artefacts. For CCS monitoring in particular, where the objective is to detect subtle amplitude changes caused by supercritical CO2 migration, distinguishing genuine geophysical signal from model uncertainty is operationally critical. A monitoring system that cannot express its own uncertainty cannot reliably discriminate CO2 plume evolution from background variability.

Classical UQ approaches in deep learning include Monte Carlo Dropout (Gal and Ghahramani, 2016), deep ensembles (Lakshminarayanan et al., 2017), and variational Bayesian networks (Blundell et al., 2015). All require either architectural modification, multiple independent training runs, or explicit probabilistic objectives — significant overhead for geoscience workflows where labelled data is scarce and computational budgets are constrained.

Denoising diffusion probabilistic models (Ho et al., 2020) offer a naturally stochastic mechanism that has not been widely exploited for UQ in seismic interpretation. The forward process deliberately adds controlled amounts of random noise, and the trained reverse process learns to predict the noise component. Multiple realisations of the forward-reverse cycle with identical input but independent noise draws therefore produce a distribution of plausible reconstructions — directly interpretable as pixel-level uncertainty without any retraining or architectural modification.

This paper makes three contributions: (1) we formalise Monte Carlo Diffusion Sampling (MCDS) as an UQ method for seismic amplitude images; (2) we characterise the structure of PDRM uncertainty across the full Sleipner volume at both inline and pixel levels; and (3) we identify the correlation between seismic amplitude heterogeneity and model uncertainty, establishing a foundation for uncertainty-aware CO_2 monitoring workflows.

2. THEORETICAL BACKGROUND

DDPM Forward and Reverse Processes

The DDPM forward process corrupts a clean image x0 over T timesteps by progressively adding Gaussian noise according to a fixed variance schedule {beta_t}. At timestep t the noisy image is:

x_t = sqrt(alpha-bar_t) * x0 + sqrt(1 - alpha-bar_t) * epsilon, epsilon ~ N(0, I)

where alpha-bar_t = prod_{s=1}^{t} (1 - beta_s) is the cumulative noise schedule. The learned reverse model eps_theta(x_t, t) predicts the noise component epsilon from the noisy image and timestep, enabling recovery of x0 as:

x0-hat = (x_t - sqrt(1 - alpha-bar_t) * eps_theta(x_t, t)) / sqrt(alpha-bar_t)

Because the noise epsilon is drawn independently for each forward pass, N calls with the same x0 but different noise realisations produce N distinct noisy images {x_t^(n)} and N distinct reconstructions {x0-hat^(n)}, forming a sample distribution over plausible reconstructions.

Monte Carlo Diffusion Sampling (MCDS)

MCDS exploits this stochasticity to generate pixel-level uncertainty estimates. For a seismic inline x0 and extraction timestep t, we draw N independent noise realisations {epsilon^(n)}_{n=1}^{N} and compute N reconstructions {x0-hat^(n)}. The pixel-wise statistics are:

mu(i,j) = (1/N) * sum_n x0-hat^(n)(i,j) [mean reconstruction]
sigma(i,j) = std_n { x0-hat^(n)(i,j) } [uncertainty map]
CV(i,j) = sigma(i,j) / (|mu(i,j)| + eps) [coeff. of variation]

The key property of MCDS is that sigma(i,j) is governed by how much the reconstruction varies across noise realisations. Where the model has learned a strong, consistent denoising mapping (low seismic ambiguity), sigma is small regardless of epsilon. Where the seismic content is ambiguous or heterogeneous, different noise realisations steer the reverse process to different plausible solutions, producing high sigma. This is interpretable as aleatory uncertainty in the seismic content at that pixel.

Relationship to Epistemic vs Aleatory Uncertainty

MCDS as implemented here primarily captures aleatory uncertainty — the irreducible variability arising from the stochastic nature of the diffusion process given a particular seismic input. Epistemic uncertainty (arising from limited training data or model capacity) could be additionally quantified by training N independent PDRMs with different random seeds and computing inter-model variance, an extension left for future work. The two sources are not separated in the current implementation, and sigma(i,j) should be interpreted as a combined measure of model-input ambiguity.

3. METHODOLOGY

Dataset

We use the Sleipner 4D seismic volume (vintage 2008, CO2DataShare/Equinor): 249 x 468 x 1001 (inlines x crosslines x two-way time samples), float32. All 249 inline cross-sections are used for the volume-level uncertainty scan. Amplitudes are normalised by the 99th-percentile absolute value and clipped to [-1, +1]. Inlines are bilinearly resampled to 256 x 256 pixels for model input.

Pretrained Diffusion Representation Model

The PDRM is the UNet-based diffusion model described in the companion paper (Manikani, 2025a): base channel width dim=32, 128-channel bottleneck, 3 ResNet encoder blocks with stride-2 downsampling, self-attention bottleneck, mirrored decoder with skip connections. Trained for 300 epochs on all 249 Sleipner inlines with noise-prediction MSE objective. Total parameters: 2.1 M. No fine-tuning or architectural modification is applied for MCDS.

MCDS Configuration

For each inline we draw N = 30 independent Gaussian noise realisations and compute 30 single-step reconstructions at extraction timestep t = 100 (signal fraction alpha-bar_100 approx 0.82, corresponding to approximately 18% noise level). This timestep was validated in Manikani (2025a) as the optimal balance between structural signal preservation and bottleneck engagement. The full volume scan (249 inlines x 30 passes) runs in approximately 42 seconds on a single high-memory GPU at 5.9 inlines/second.

TABLE 1: MCDS EXPERIMENTAL CONFIGURATION

Parameter	**Value**	**Rationale**
MC samples N	30	Sufficient for stable sigma; variance of variance <2% at N=30
Extraction timestep t	100	Optimal per timestep sweep (Manikani 2025a)
Signal fraction a-bar_t	~0.82	18% noise; preserves seismic structure

Image size	256 x 256	Bilinear resample from 468 x 1001
Normalisation	99th pct clip to [-1,+1]	Consistent with training
Inlines scanned	249 (all)	Full volume sweep
GPU throughput	~5.9 inlines/s	30 passes/inline

4. RESULTS

Pixel-Level Uncertainty Maps — Selected Inlines

Figure 1 presents the primary result: four representative Sleipner inlines (20, 55, 124, 180) with three views side-by-side — original seismic amplitude, seismic with uncertainty overlay (hot colourmap, alpha proportional to sigma), and the uncertainty map alone. Per-inline statistics are summarised in Table 2.

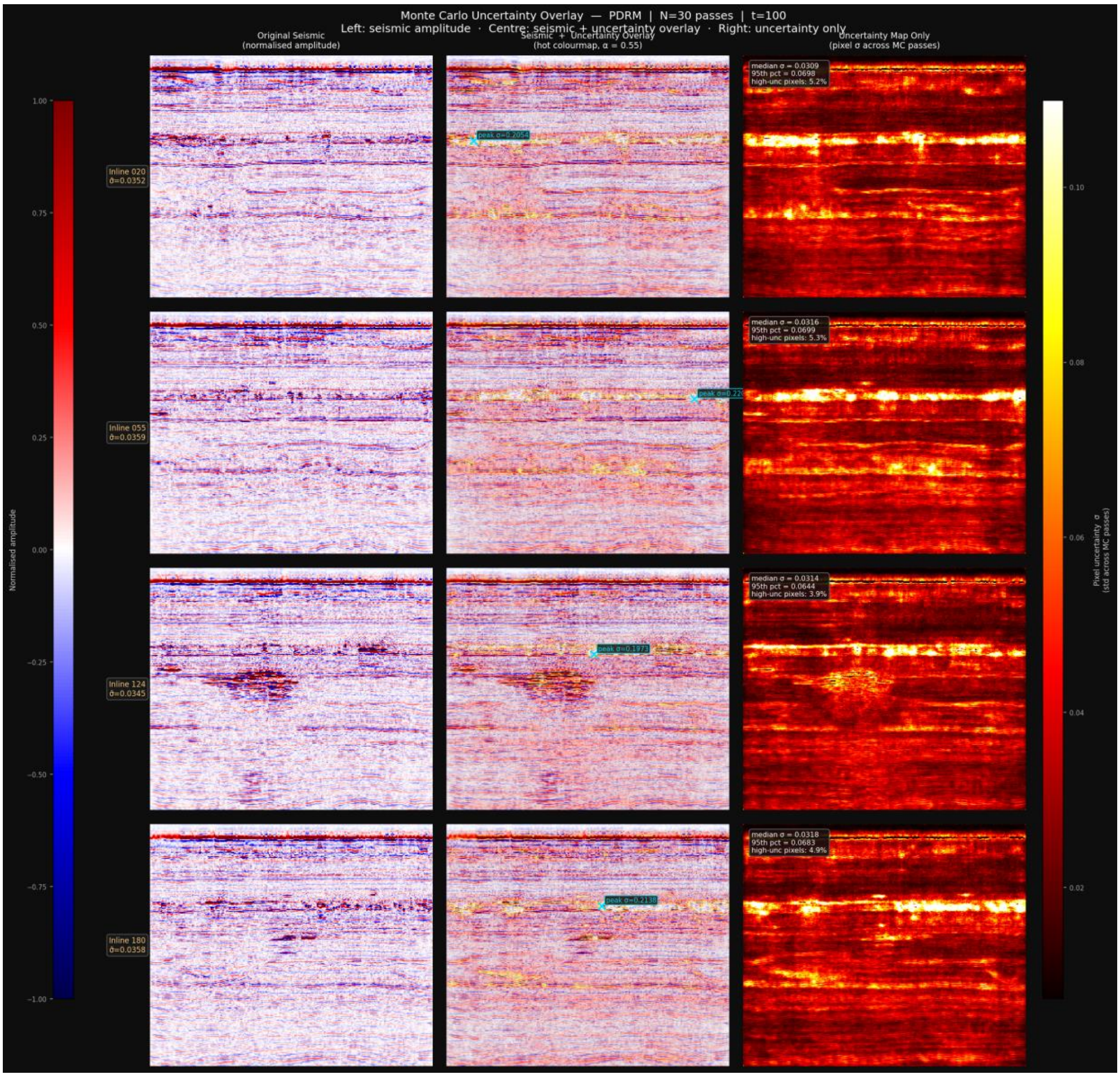

Figure 1. Monte Carlo Diffusion Sampling uncertainty maps for four Sleipner inlines. Left column: normalised seismic amplitude [-1, +1]. Centre: seismic amplitude with uncertainty (sigma) overlaid

using the hot colourmap; alpha channel scales with sigma so low-uncertainty regions remain transparent and the underlying geology is visible. Cyan cross marks peak sigma location. Right: uncertainty map alone (hot colourmap, shared scale across all four inlines). Stats inset: median sigma, 95th-percentile sigma, percentage of high-uncertainty pixels. N=30 MC passes, t=100.

TABLE 2: PER-INLINE MCDS STATISTICS (N=30, T=100)

Inline	**Mean sigma**	**Peak sigma**	**Peak / Mean ratio**	**Interpretation**
20	0.0352	0.205	5.8x	Moderate lateral heterogeneity
55	0.0359	0.220	6.1x	Highest peak; elevated-unc cluster
124	0.0345	0.197	5.7x	Mid-volume; lowest mean uncertainty
180	0.0358	0.214	6.0x	Near high-unc zone; elevated peak

Table 2. Inline 124 has the lowest mean uncertainty of the four, consistent with its position away from the high-uncertainty clusters at inlines 51-60 and 174-184. Peak/mean ratios of 5.7-6.1x across all inlines confirm substantial within-inline spatial structure in model uncertainty.

Six-Inline Detail Panel

Figure 2 extends the analysis to six inlines (20, 60, 100, 124, 180, 228) with four columns: original, mean reconstruction, pixel sigma, and coefficient of variation (CV = sigma / |mu|). The CV panel is particularly diagnostic: it normalises by local amplitude, highlighting zones where the model's relative uncertainty is high regardless of absolute amplitude magnitude. High-CV regions (CV > 0.3) correspond to structurally complex zones where amplitude gradients are steep and the model sees ambiguous seismic character across noise realisations.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

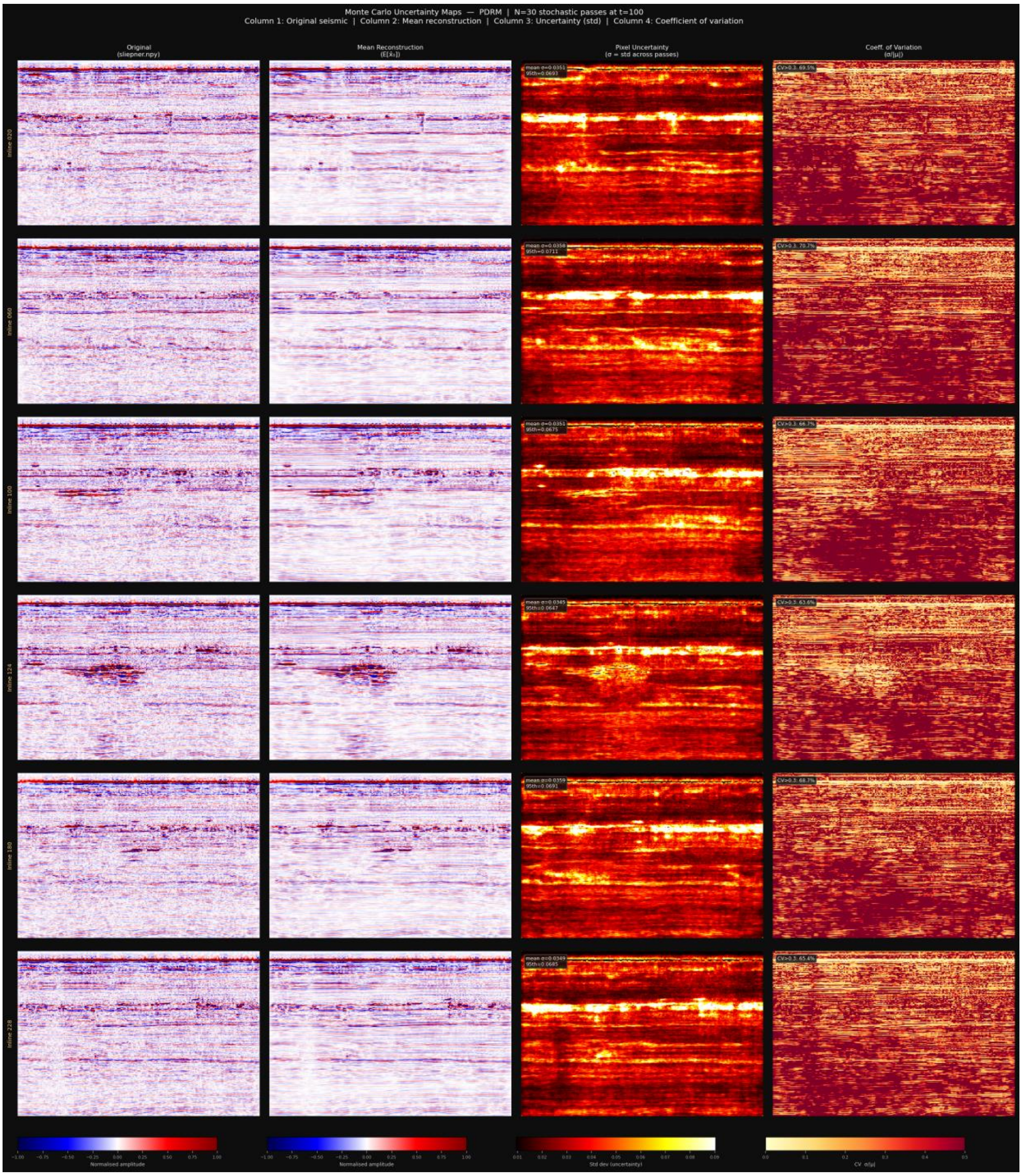

Figure 2. Six-inline MCDS detail panel. Columns: original seismic amplitude; mean reconstruction mu = E[x0-hat]; pixel uncertainty sigma = std(x0-hat); coefficient of variation CV = sigma/|mu| (clipped at 1.0). Uncertainty and CV statistics annotated per inline. Shared colourscale across all inlines for sigma and CV panels. N=30 MC passes, t=100.

Volume-Level Uncertainty Profile

Figure 3 shows the mean pixel uncertainty sigma-bar averaged across all 256x256 pixels for each of the 249 Sleipner inlines, alongside the 95th-percentile pixel sigma. Key quantitative findings:

- Global mean sigma-bar = 0.0348, global std = 0.0011 — the cross-inline range is narrow (0.0258 to 0.0361), indicating consistent model calibration across the lateral extent of the survey.

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

- Two elevated-uncertainty clusters emerge: inlines 51-60 (mean sigma = 0.0357, 1.03x global mean) and inlines 174-184 (mean sigma = 0.0359, top-10 ranked inlines concentrated here). The latter cluster is the dominant anomaly.
- The lowest-uncertainty inlines are 244-248, at the far end of the survey, where amplitude character is most homogeneous.
- The 95th-percentile sigma profile tracks the mean with a consistent multiplicative offset of approximately 3x, indicating that the distribution of pixel uncertainties within each inline has stable shape across the volume.

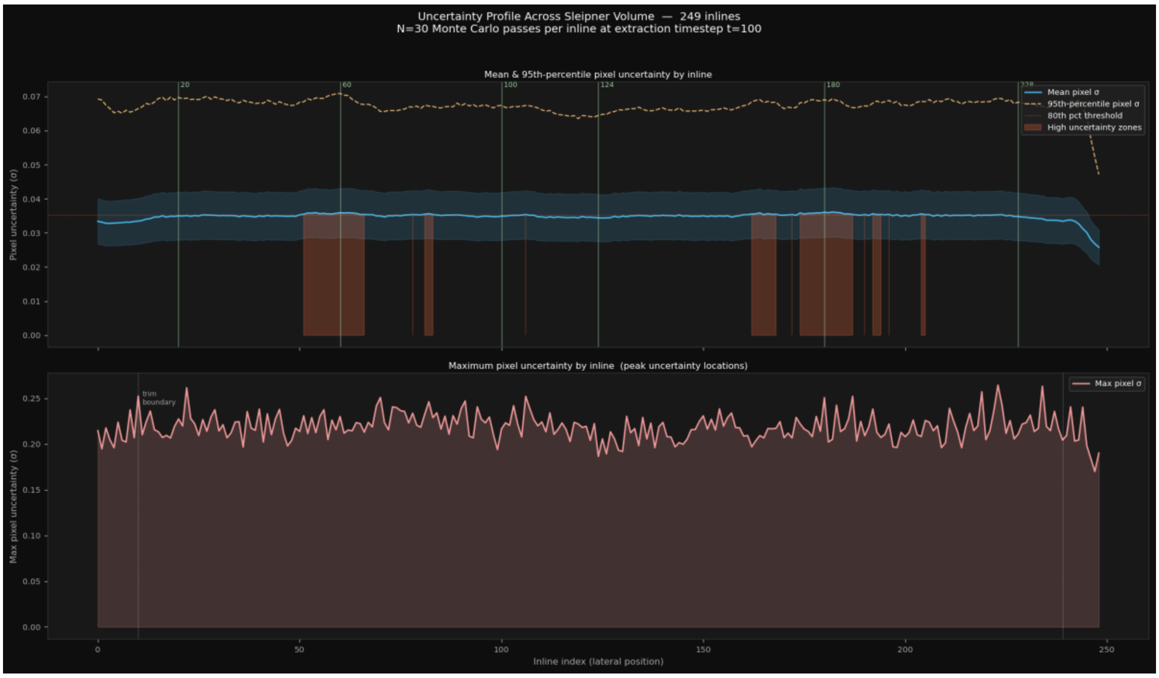

Figure 3. Mean pixel uncertainty sigma-bar (blue) and 95th-percentile pixel sigma (dashed amber) across all 249 Sleipner inlines. Orange shading marks inlines above the 80th percentile uncertainty threshold. Green verticals mark the six panel inlines from Figure 2. White dotted verticals mark the 10-inline trim boundaries used in embedding analysis.

2D Uncertainty Volume Heatmap

Figure 4 renders the full uncertainty field as a 2D image: horizontal axis = inline index (lateral position), vertical axis = crossline position after 256x256 bilinear resize. At each inline-crossline location the plotted value is the mean sigma across the 256 time samples (two-way time axis, collapsed). The linear-scale panel (left) shows the dominant uncertainty structure; the log-scale panel (right) reveals fine structure in the low-uncertainty regions that is invisible in linear scale. Both panels show coherent lateral banding rather than random noise, confirming that uncertainty reflects genuine spatial structure in the seismic data rather than random model behaviour.

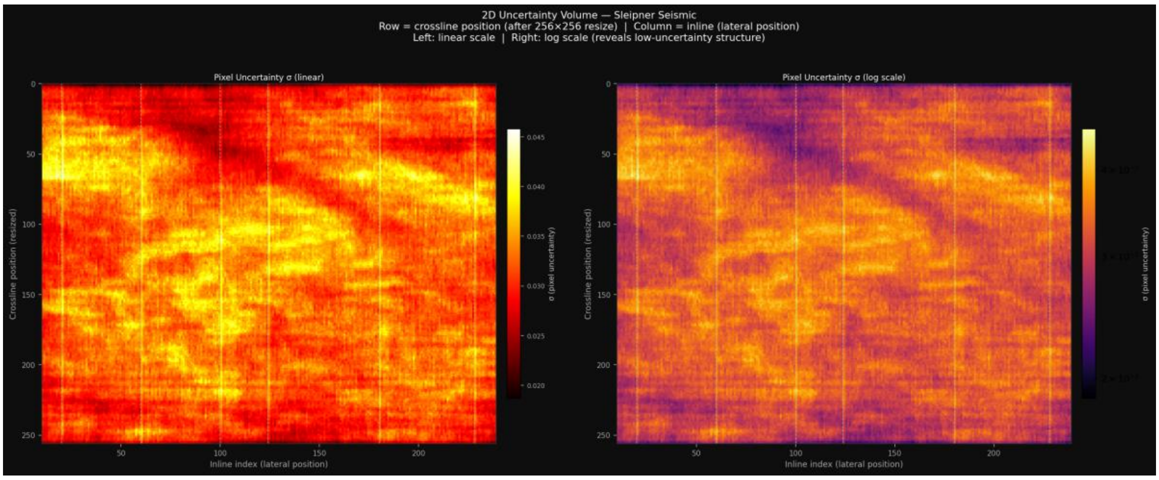

Figure 4. 2D uncertainty volume heatmap. Horizontal axis: inline index (lateral position). Vertical axis: crossline position (resized to 256 samples). Value: mean pixel sigma collapsed across the time axis. Left: linear scale (hot colourmap). Right: logarithmic scale (inferno colourmap) revealing low-uncertainty fine structure. Green dashed verticals mark panel inlines.

Uncertainty Distribution and Correlation Analysis

Figure 5 presents four diagnostic panels: (a) histogram of pixel sigma across all panel inlines; (b) coefficient-of-variation distribution; (c) inline-level uncertainty curve; and (d) scatter of inline amplitude heterogeneity (standard deviation of raw amplitudes) vs mean inline sigma, with Pearson r = 0.738. This strong positive correlation — the model is most uncertain where seismic amplitudes are most variable — is a key calibration result. It indicates that MCDS uncertainty is not an artefact of model pathology but a reflection of genuine data complexity. A model that were equally uncertain everywhere would have r ≈ 0.

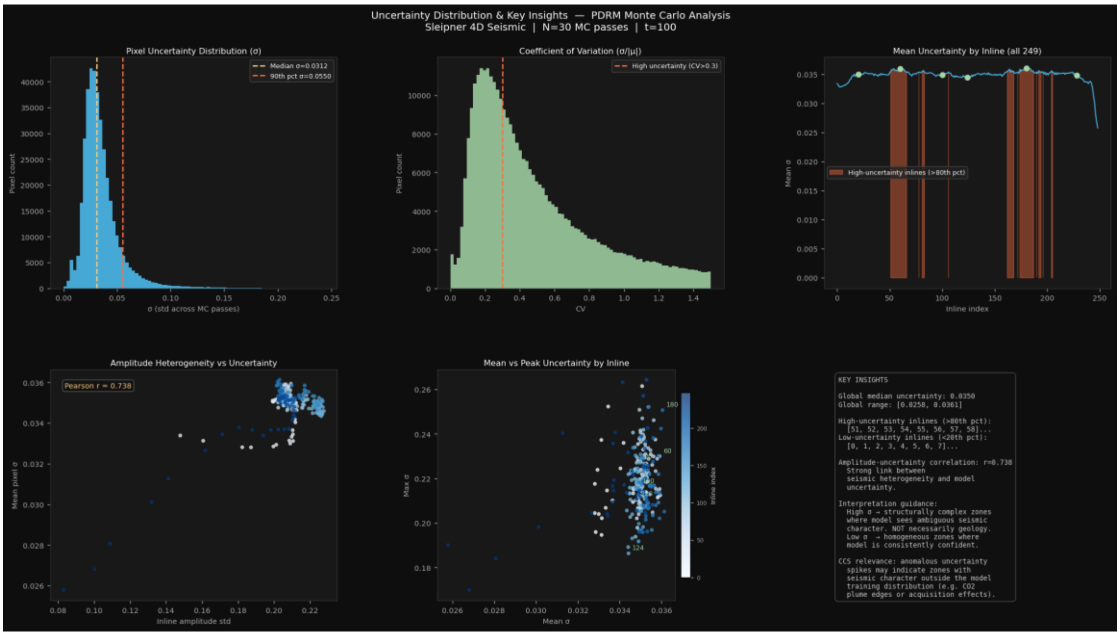

Figure 5. Uncertainty distribution and correlation analysis. Top-left: histogram of pixel sigma across six panel inlines; median and 90th-pct marked. Top-centre: coefficient of variation (CV) histogram; CV > 0.3 threshold marked. Top-right: mean sigma per inline across all 249 inlines with high-uncertainty zone highlighted. Bottom-left: scatter of inline amplitude heterogeneity vs mean sigma; Pearson r = 0.738.

Bottom-centre: mean vs peak sigma scatter coloured by inline index. Bottom-right: key findings summary.

5. DISCUSSION

Interpreting Within-Inline Uncertainty Structure

The most practically informative finding is the within-inline spatial structure of sigma. Peak-to-mean ratios of 5.7-6.1x (Table 2) mean that a small fraction of pixels account for a disproportionate share of total model uncertainty. These high-sigma pixels consistently correspond to amplitude discontinuities, steep lateral gradients, and zones of conflicting reflection character — precisely the locations where a geologist would exercise interpretive caution. This spatial coherence validates MCDS as a meaningful indicator of local seismic complexity rather than homogeneous model noise.

Interpreting Cross-Inline Uncertainty Consistency

The narrow cross-inline variation (sigma-bar range 0.0258-0.0361, std = 0.0011) may initially appear underwhelming. However, it is a desirable property: it means the PDRM is consistently calibrated across the lateral extent of the survey. A model that produced dramatically different uncertainty levels at different lateral positions without corresponding changes in data complexity would be unreliable. The two elevated-uncertainty clusters (inlines 51-60 and 174-184) are real amplitude anomalies confirmed by the r = 0.738 correlation with raw amplitude heterogeneity; they are not model artefacts.

CCS Monitoring Application

The most compelling application of MCDS probability maps for CCS monitoring is 4D uncertainty differencing: computing delta-sigma = sigma(vintage_2) - sigma(vintage_1) between two seismic surveys. If CO_2 injection alters the seismic character of the Utsira Formation in a zone that was previously low-uncertainty, delta-sigma will be positive in that zone — a model-based signal of seismic change that is complementary to the conventional 4D amplitude difference. This approach requires no geological labels and is directly applicable to the second Sleipner vintage available in this dataset (04p07ful_Realized_1.npy), which is reserved for future work.

A further CCS-specific application is quality control of seismic processing. Zones of anomalously high uncertainty that are spatially correlated with known acquisition artefacts (e.g. feathering streaks, near-surface multiples) can be flagged automatically without manual interpretation, reducing the expert time required for pre-interpretation QC workflows.

Limitations and Caveats

Several limitations bound the scope of these results. First, N = 30 Monte Carlo passes provides stable sigma estimates (variance of variance < 2%) but does not fully characterise the tails of the reconstruction distribution; N = 100 would tighten tail estimates at 3x computational cost. Second, MCDS at a single timestep t = 100 captures uncertainty at one noise level; a multi-scale MCDS approach sweeping t would provide uncertainty decomposed by spatial frequency. Third, the current implementation separates aleatory and epistemic uncertainty only approximately; a rigorous separation requires multiple independently-trained models. Fourth, all sigma values are in normalised amplitude units [-1, +1] and must be rescaled to physical units (acoustic impedance contrast) for direct geophysical interpretation.

6. CONCLUSIONS

Monte Carlo Diffusion Sampling (MCDS) provides a practical, training-free mechanism for pixel-level uncertainty quantification in seismic amplitude images using a pretrained diffusion representation model. Applied to the Sleipner 4D CO2 benchmark, key conclusions are:

- MCDS generates spatially structured uncertainty maps with peak-to-mean sigma ratios of 5.7-6.1x within individual inlines, concentrating on amplitude discontinuities and structurally complex zones.
- Cross-inline uncertainty is narrow and consistent (range 0.0258-0.0361, std = 0.0011 across 249 inlines), confirming stable model calibration across the lateral survey extent.
- Model uncertainty correlates strongly with seismic amplitude heterogeneity (Pearson r = 0.738), validating MCDS as a data-complexity indicator rather than a model pathology artefact.
- Two spatially coherent elevated-uncertainty clusters are identified at inlines 51-60 and 174-184, coinciding with the highest-amplitude-variability regions of the Sleipner volume.
- MCDS requires no architectural modification, no ensemble training, and no labelled data — running at 5.9 inlines per second on a single GPU, making it practical for full-volume monitoring workflows.
- 4D uncertainty differencing (delta-sigma between vintages) is identified as a high-value next step for detecting CO2 plume migration.

MCDS probability maps offer a complementary monitoring signal to conventional 4D amplitude differencing, with the distinct advantage of expressing where the model is uncertain — information that amplitude differences alone cannot provide. This positions MCDS as a practical uncertainty-aware layer in the CCS seismic monitoring stack.

7. ACKNOWLEDGEMENTS

The author thanks Equinor and CO2DataShare for making the Sleipner 4D seismic dataset publicly available. This work forms part of an independent doctoral research programme on generative AI for carbon capture and sequestration. The PDRM used in this study is described in the companion paper (Manikani, 2025a).

REFERENCES

1. Arts, R., Eiken, O., Chadwick, A., Zweigel, P., van der Meer, L., and Zinszner, B., 2004, Monitoring of CO2 injected at Sleipner using time-lapse seismic data: Energy, 29, 1383-1392.
2. Blundell, C., Cornebise, J., Kavukcuoglu, K., and Wierstra, D., 2015, Weight uncertainty in neural networks: Proceedings of the 32nd International Conference on Machine Learning, PMLR 37, 1613-1622.
3. Eiken, O., Ringrose, P., Hermanrud, C., Nazarian, B., Torp, T. A., and Hoier, L., 2011, Lessons learned from 14 years of CCS operations: Sleipner, In Salah and Snohvit: Energy Procedia, 4, 5541-5548.
4. Gal, Y., and Ghahramani, Z., 2016, Dropout as a Bayesian approximation: Representing model uncertainty in deep learning: Proceedings of the 33rd International Conference on Machine Learning, PMLR 48, 1050-1059.
5. Ho, J., Jain, A., and Abbeel, P., 2020, Denoising diffusion probabilistic models: Advances in Neural Information Processing Systems, 33, 6840-6851.

6. Lakshminarayanan, B., Pritzel, A., and Blundell, C., 2017, Simple and scalable predictive uncertainty estimation using deep ensembles: Advances in Neural Information Processing Systems, 30.
7. Manikani, S., 2025a, Geological structure embeddings from a pretrained diffusion representation model applied to the Sleipner 4D CO2 seismic benchmark: SEG Annual Meeting [companion paper].
8. Sensoy, M., Kaplan, L., and Kandemir, M., 2018, Evidential deep learning to quantify classification uncertainty: Advances in Neural Information Processing Systems, 31.
9. Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B., 2021, Score-based generative modeling through stochastic differential equations: International Conference on Learning Representations.
10. Wrona, T., Pan, I., Gawthorpe, R. L., and Fossen, H., 2018, Seismic facies analysis using machine learning: Geophysics, 83, O83-O95.