Well log foundation model - making promptable AI models for interpretation

[1]*Ben Lasscock,* [1]*Altay Sansal,* [1]*Keyla Gonzalez, and* [1]*Alejandro Valenciano*
[1]*TGS*

Summary

We present a 60M-parameter well log vision transformer foundation model trained using a masked autoencoding framework (ViT-MAE). This model was pretrained on 1.1 million North American well logs for automated well log imputation and subsequently extended to execute prompt-based geologist-guided formation top interpretation as a downstream task. The formation model was fine-tuned on 271,972 human-interpreted formation tops from 44,062 wells across the Permian Basin, covering 37 formations. Unlike classification-based approaches, this prompt-based method for predicting formation depth facilitates training on a vast collection of existing interpreted wells. The foundation model and geologist-guided interpretation method are crucial for accelerating prospect generation by integrating AI-driven geological analysis into exploration workflows.

Introduction

This study explores how emerging AI techniques can transform formation interpretation in the oil and gas sector through large-scale pretraining, and prompt-based inference. Inspired by the scalability of self-supervised Masked Autoencoder Vision Transformer (ViT-MAE) (He et al. 2022) in seismic interpretation (Lasscock et al. 2024b; Sansal et al. 2025a; Sansal et al. 2025b), we introduce a 60M-parameter foundation model pretrained on 1.1 million cleaned North American well logs across multiple basins — one of the largest well-log training datasets. The model learns geological patterns for well log imputation, enabling it to fill in missing curves by leveraging inter-log relationships. We created a prompt-based formation top interpretation model to demonstrate a downstream task built on this foundation model. The model takes as input well log and a formation name as a prompt, then predicts the requested formation depth in feet.

Previous studies have explored AI-driven formation top predictions using support vector machines (Hall et al. 2016), convolutional neural networks (Liu et al. 2018), classification models (Ibrahim et al. 2023), and artificial neural networks (Elkatatny et al. 2019). More recently, functional neural networks (Mahmoud et al. 2024) have been examined for real-time formation prediction. While promising, these methods are often limited to in-domain applications, relying on restricted datasets. In addition to these methods, semi-supervised approaches have been developed to broaden geological interpretations beyond initial formation tops. Gonzalez et al. (2024) applied this technique in the Midland Basin to generate 3D property volumes, while Halder et al. (2024) utilized it to construct an accelerated stratigraphic framework for CO_2 storage assessment. A large-scale training of a time-series transformer model with classification-based formation prediction built on a pretrained model was also conducted (Lasscock et al. 2024a). Unlike classification-based formation interpretation models, the prompt-based approach in this study removes the need for consistent label taxonomies or densely labeled training data, making it adaptable to the largest possible collection of interpreted datasets and extensible to adding new formations.

The formation top interpretation methodology introduced in this study has three key innovations:

- **Transfer Learning Architecture** – Built on a pre-trained ViT-MAE model, the system leverages masked well log reconstruction to develop a deep understanding of subsurface patterns, fine-tuned for formation top prediction.
- **Formation Pattern Recognition** – trained on 271,972 manually interpreted formation tops across 37 geological formations, the model learns characteristic patterns of formation boundaries.
- **Geologist-Guided Interpretation** – Instead of identifying predefined formations, users specify a formation of interest (e.g., "Top Wolfcamp"), allowing targeted predictions.

This approach aligns with industry trends toward promptable foundation models, addressing challenges like incomplete formation picks and reliance on manual interpretation. By enabling accurate depth predictions conditioned on a prompt, the model augments geoscience expertise rather than replacing it.

Dataset

A key feature of this study is building a large-scale AI-driven well log analysis framework trained on 1.1 million wells from the major U.S. basins. The dataset spans 16 basins, including the Permian Basin (Delaware, Central Basin Platform, Midland), Gulf Coast, Anadarko, and Marcellus (shown in Figure 1), providing a diverse geological representation of conventional and unconventional reservoirs.

Before training, we established a well log cleaning pipeline to ensure consistency and quality across the dataset. This

process involved categorizing each log curve, verifying LAS and log header information, splicing and merging logs from different runs, depth shifting for alignment, normalizing logs to account for tool and environmental effects, and removing or editing low-quality data (Gonzalez et al., 2023). A key design principle was to minimize human interpretation, enabling a scalable, basin-wide application. The main curve categories include Gamma-ray, Neutron Porosity, Compressional Sonic, Bulk Density, and Resistivity. The formation-picking model was built using a dataset of 271,972 manually picked formation tops sampled from 44,062 wells across the Permian Basin. This dataset includes 37 interpreted formation tops, covering a broad stratigraphic range from shallow evaporites to deep carbonate and clastic reservoirs. The distribution of tops varies significantly, as illustrated in Figure 2, with higher-density interpretations in key formations such as Wolfcamp, Bell Canyon, Rustler, Bone Spring, Salado, Strawn, Devonian Carbonate, Brushy Canyon, and Glorieta. These formations are thoroughly mapped across the Delaware Basin, Central Basin Platform, and Midland Basin, making them essential reference points for stratigraphic correlation and machine learning-based formation top prediction.

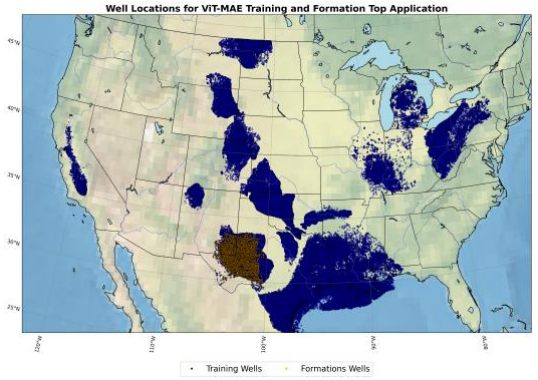

Figure 1. Map showing the spatial distribution of wells used for training (blue) and formation top interpretation (orange) across key U.S. basins.

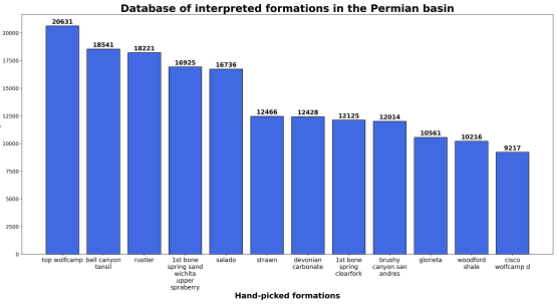

Figure 2. Counts for the 12 (of 37) most frequently interpreted formations in the dataset.

Methodology

Training the well log foundation model

In this paper, we pretrain a foundation model for well log imputation. We train a long-context, anisotropic ViT encoder with multi-head self-attention (12 heads) across 8 layers in a 768-dimensional latent space and a feedforward network (FFN) size of 3072. The input well log can reach depths of up to 24,000 ft (at 0.5 ft per sample) and include one or more of the five primary curve categories. The model tokenizes each well log curve into patches of length 16. It incorporates flash attention for efficiency **and** query-key normalization (QK norm) for stability**,** with a 10% dropout for both attention and feedforward layers. A smaller decoder with the same architecture is used for imputation. The model is trained (self-supervised) with a masking ratio of 60% and is evaluated on the task of well log imputation (Gonzalez et al. 2023).

Downstream application: Formation top picking

An interpreter is intended to guide the inference of a formation depth and input a formation name (one of the 37 formations currently supported by the model). Instead of treating formation names as simple categories, we embed them into the same high-dimensional space as the well data, as shown schematically in Figure 3.

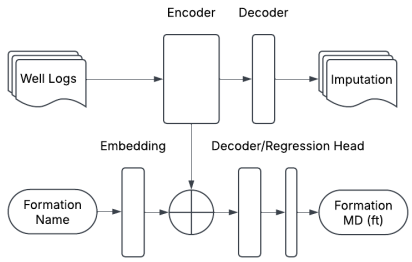

Figure 3. A schematic view of the Well log autoencoder with a formation regression head. The encoder is trained independently and frozen when training the regression model.

This embedding of formation data is integrated into the latent representations produced by the trained, frozen ViT encoder, effectively conditioning the model for the target formation. The conditioned representations are then processed through a decoder that generates token-level features. A learned attention mechanism scores these tokens to emphasize the most relevant ones. By calculating a weighted sum of the token features, we create a context vector that informs the final regression layer, which predicts the formation depth. This blend of transfer learning, formation-specific conditioning, and attention-based

aggregation allows us to train with the existing data without needing a consistent interpretation across all wells. Formation tops interpreted in 8,813 wells, totaling 54,149 formation top picks, were withheld for validation, while the remainder was used for training. This ensures we can assess the model's ability to interpret blind test wells.

Results

Imputation: Filling in Missing Logs

Existing state-of-the-art well log imputation techniques (Gonzalez et al., 2023) ensemble multiple models to fill in various collections of curves on a per-basin basis. The foundation model, on the other hand, is a single model trained on data from 16 basins and can perform imputation with any input collection of curves. Figures 4 and 5 illustrate examples of density curve imputation in the Permian and Marcellus Basins, showcasing the model's ability to generalize across different geological settings. These results indicate that a single generalized model is spatially consistent in predicting missing logs along cross-sections, demonstrating its adaptability for basin-wide applications.

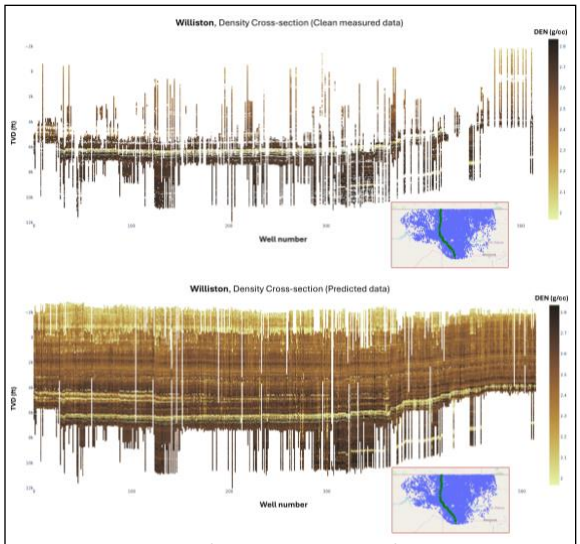

Figure 4. Imputation of density well log data for a Williston Basin cross-section, (upper section) represents the original measured data, and the (lower section) shows the prediction. (inset) Mini map shows the location of the cross-section.

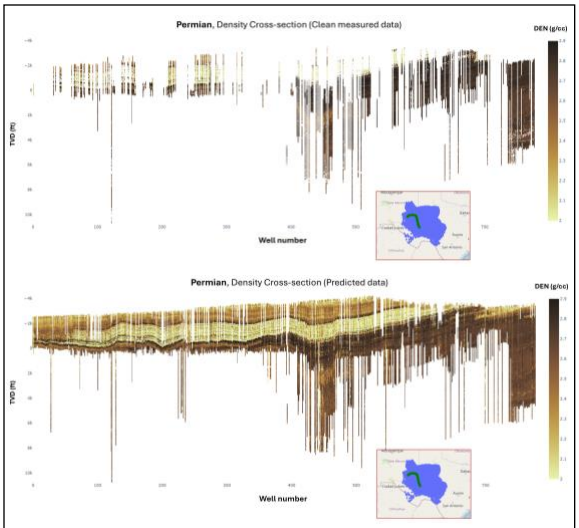

Figure 5. As in Figure 4, but for a cross-section sampled in the Permian basin.

Downstream task: Formation Top Interpretation

We demonstrate a use case where the foundation model can be efficiently fine-tuned to a downstream task of formation top regression. For example, in Figure 6, we show a 2D cross-section of gamma-ray logs sampled from the Midland Basin, featuring conventionally interpreted formations alongside model-predicted formation tops. The gamma-ray cross-section illustrates variations in log response, highlighting the spatial consistency of the formation model predictions and the model's ability to replicate expert-level interpretation of key formations. In Table 1, we validate model performance on the validation dataset for a collection of key formations. We found that the residuals are characterized by a few large outliers; therefore, we examine the trimmed RMSE error (5%) and a median absolute error (MedAE) metric to score each formation. We also compute the mean absolute percentage error and the R^2 coefficient, showing the number of samples for each formation in the validation set. We find that the trained model typically provides excellent performance across the scale of the Permian Basin.

Next, to demonstrate the utility of an automated tool, we replicate the workflow of creating a stratigraphic framework across four counties in the Midland Basin (Glasscock, Howard, Martin, and Midland). The prompt-based model identifies key formation tops in the area of interest by interpreting a total of 18,588 wells. The usefulness of the promptable model is evident here because not every formation in the Permian Basin is also present in the Midland Basin. However, since the interpreter understands what

formations to expect, they can effectively use the prompt mechanism accordingly.

The model can infer formation depth, even in the absence of well data to support it. Unlike the validation data, we cannot be certain that each well has data at the formation location to be identified, which may result in outliers in the predictions. To improve the predictions, large outliers were removed from this collection of tops, and a 2D stratigraphic grid is then created for the basin using biharmonic spline interpolation. Figure 7 (left) displays the Bone Spring – Upper Spraberry and Top Wolfcamp grids derived from the prompt-based model, while the right shows the same grid derived using 1,255 manually selected tops. With denser sampling, the automated tool extracts more detailed structural information and requires a fraction of the manual effort to build (minutes compared to months).

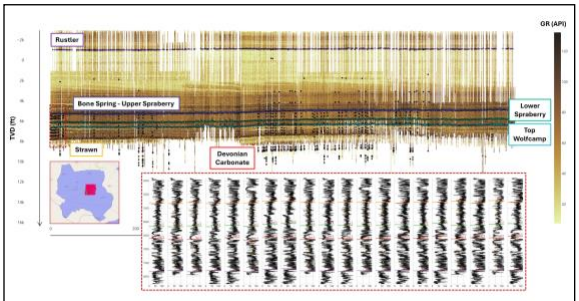

Figure 6. A Gamma-ray cross-section from a 2D well log section in the Midland Basin, showing predicted and manually interpreted formation tops. Colors represent predicted tops: Rustler (purple), 1st Bone Spring - Upper Spraberry (dark blue), Lower Spraberry (green), Top Wolfcamp (light blue), Strawn (orange), and Devonian Carbonate (red), with manually interpreted tops in black. (Inset) A zoomed-in Gamma-ray track view compares automated formation top picks with manual interpretations.

This workflow enabled the rapid construction of stratigraphic grids (Figure 7), showcasing scalability for basin-wide applications and allowing for extra control points in areas with limited interpretations, where uncertainty increases. By systematically transferring subsurface knowledge, the model enhances formation definition, speeds up stratigraphic correlation, and improves structural mapping for reservoir characterization.

Conclusions

This study presents a foundation model for well log interpretation using a ViT-MAE architecture pretrained on 1.1 million well logs. We demonstrate its effectiveness for well log imputation and extend its application to automated formation top interpretation using 271,972 manually interpreted tops from 44,062 wells. The model exhibits expert-level accuracy in identifying structural trends and generating stratigraphic grids. Results indicate that automation significantly reduces the effort required for interpretive tasks compared to traditional methods.

This study outlines a methodology that is scalable to larger models trained on enhanced well log and seismic datasets, improving multi-modal subsurface characterization. By integrating additional geological data and refining the automation process, we strive to further enhance prospect generation, minimizing manual interpretation efforts while preserving geological accuracy.

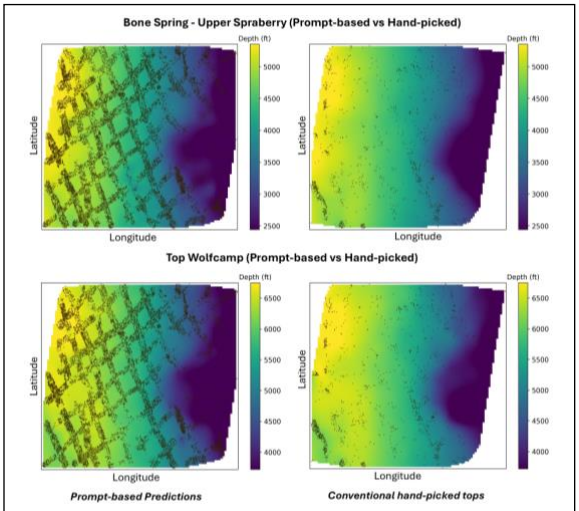

Figure 7. Stratigraphic grids for Bone Spring - Upper Spraberry (Top) and Top Wolfcamp (Bottom). The left panels show predictions from the prompt-based approach, while the right panels display conventional hand-picked formation tops for comparison.

Table 1. Summary statistics measuring the model performance in predicting formation tops across six key formations, in the validation dataset.

Formation	***MedAE (ft)***	***Trim-RMSE (ft)***	***MAPE (%)***	***R^2***	***Support***
Rustler	13.0	18.9	1.36%	0.978	3416
Bone Spring - Upper Spraberry	33.4	54.8	0.83%	0.996	3043
Lower Spraberry Shale	28.6	41.9	0.31%	0.998	1146
Top Wolfcamp	40.0	82.7	1.34%	0.992	3747
Strawn	32.0	65.9	0.68%	0.993	2312
Devonian Carbonate	39.6	65.2	0.77%	0.996	2194