geoconvention
Calgary • Canada • May 11-13 2026

Synthetic P-wave Velocity Model Generation using LLMs for Seismic Data Simulation

Marcelo Guarido, Ivan Sanchez, and Kristopher Innanen
CREWES – University of Calgary

Summary

In this work, we explored the use of large language models (LLMs) and multimodal models to generate synthetic P-wave velocity models for seismic data simulation. We tested GPT-5 and GPT-5-mini, combined with extensive prompt engineering, for generating velocity models from textual descriptions, but the outputs were often unrealistic and inconsistent. Improvement attempts using few-shot prompts with examples of complex velocity models did not yield satisfactory results. Maybe using more examples in the few-shot prompt or providing a better-designed system prompt could improve the results, but due to token limitations and resource constraints, we were unable to explore these options further. On the other hand, DALL-E 2 showed promise in generating variations of velocity models based on an initial image and text prompts. It allows the creation of multiple images simultaneously, providing a range of velocity models for seismic data simulation. However, the generated models lacked a direct mapping to velocity values and were not geologically plausible.

Method

Since the launch of GPT-3 by Brown et al. (2020), large language models (LLMs) have revolutionized the field of natural language processing (NLP). These models are trained on massive amounts of text data and can generate human-like text, translate languages, and even answer questions. The architecture of these models is based on the transformer model introduced by Vaswani et al. (2017), which uses self-attention mechanisms to capture long-range dependencies in text data.

LLMs can generate coherent, contextually relevant text by predicting the next word in a sequence based on prior words. LLM can also generate structured data, such as tables, in JSON format, which is equivalent to a Python dictionary and can be easily converted to other formats, such as CSV or Excel, for further analysis. This capability enables the generation of numerical matrices representing velocity models from textual descriptions and is tested in this work using GPT-5 and GPT-5-mini. For that, prompt engineering techniques are employed to guide the model in generating the desired output.

Multimodal models are designed to process and generate data from multiple modalities, such as text, images, and audio. These models leverage the strengths of each modality to improve overall performance on tasks that require understanding and generating information from different sources. For example, a multimodal model could analyze a text description of a scene and generate a corresponding image or video.

Recent advancements in multimodal models have been driven by the development of transformer architectures capable of handling diverse data types. These models often use a combination of

encoders and decoders to process input from different modalities and generate coherent outputs. Examples of multimodal models include CLIP (Contrastive Language-Image Pretraining) (Radford et al., 2021) and DALL-E (Ramesh et al., 2021). DALL-E 2 (Ramesh et al., 2022) is a multimodal model developed by OpenAI that can generate high-quality images from textual descriptions.

In this project, we are using text-to-text LLMs (GPT-5 and GPT-5-mini) and the image-modification capabilities of DALL-E 2 models to generate P-wave velocity models for seismic simulation and machine learning training data.

Results

The models generated using GPT-5 and GPT-5-mini were unsuccessful in creating realistic velocity models, as the outputs often contained inconsistent velocity values and unrealistic geological features. The models struggled to understand the complex relationships between geological formations and their corresponding velocity values, resulting in outputs that did not accurately reflect plausible subsurface structures. We achieved better results using DALL-E 2 to generate variations of velocity models from an initial model and text prompts describing the desired modifications. All the tests were done using the OpenAI API.

Velocity Models Generated from GPT-5-mini and GPT-5

Initial tests were conducted using GPT-4o-mini, but its output size limitations prevented the generation of velocity models with sufficient resolution. Therefore, we switched to GPT-5-mini and GPT-5, enabling us to generate longer outputs. We used a few-shot-chain-of-thought prompt style, providing detailed instructions to the LLM for generating velocity models with complex geologic features, such as faults and dip layers, in JSON format, with the Marmousi model (Versteeg, 1994) as an example output. Results are shown in Figure 1:

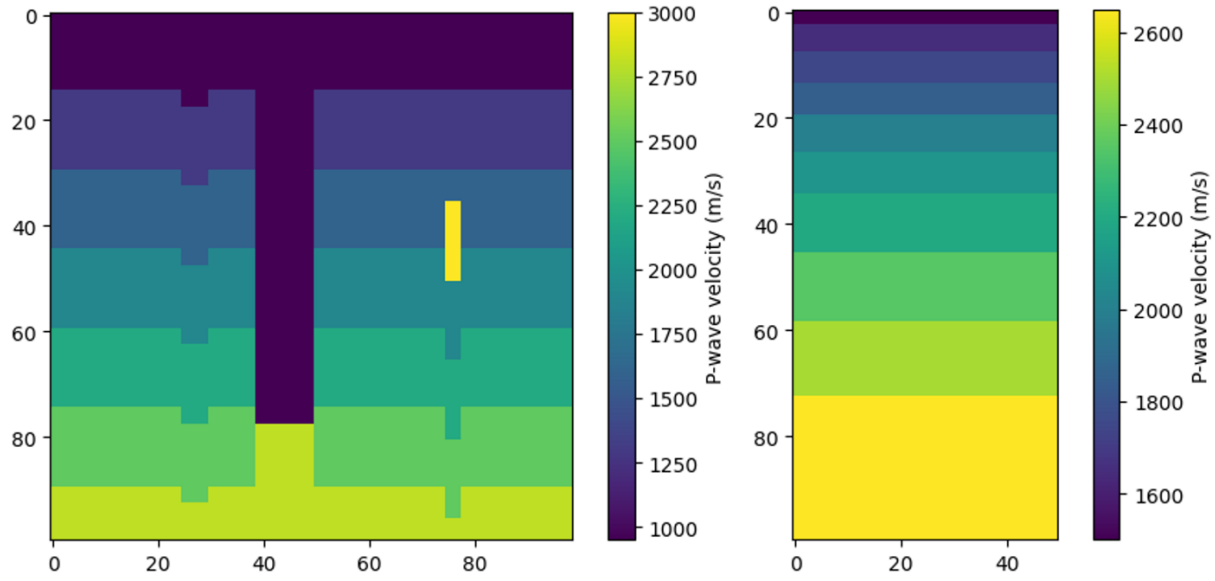


Figure 1: Velocity models generated using GPT-5. The LLM was unable to include geological complexity in the velocity models.

LLMs like GPT-5 are unable to generate velocity models for complex geology because they were not trained for this task. The generated models were too simple and inconsistent, and they included implausible geological features.

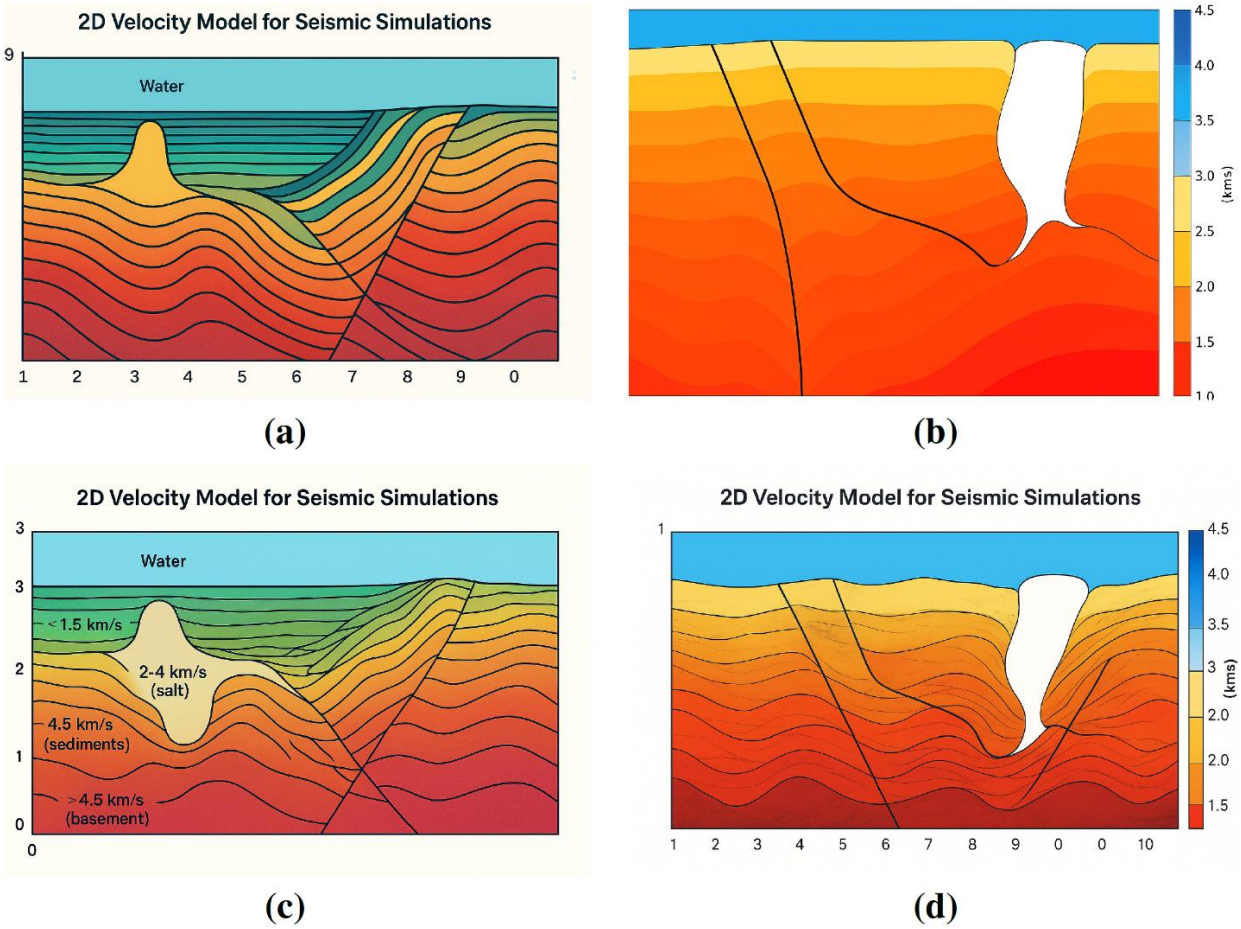


Figure 2: Output examples generated by GPT-5's image capabilities. The images are somewhat structured, but the geology is not accurately represented.

Figure 2 shows examples of velocity models generated by GPT-5 using its image-generation capabilities. The images are somewhat structured, but geology is not accurately represented, with unrealistic layering and velocity variations. Some images have titles, labels, and colorbars, while others do not. This indicates that while GPT-5 can generate images, it may not be well-suited to complex geological modelling.

Velocity Models Generated from DALL-E 2

For the DALL-E 2 model, we used an initial velocity model as a base image and provided text prompts to modify the image and generate variations of the velocity model. The initial velocity model used is shown in Figure 3. It is a complex onshore velocity model with multiple layers and geological features, providing a good starting point for generating variations.

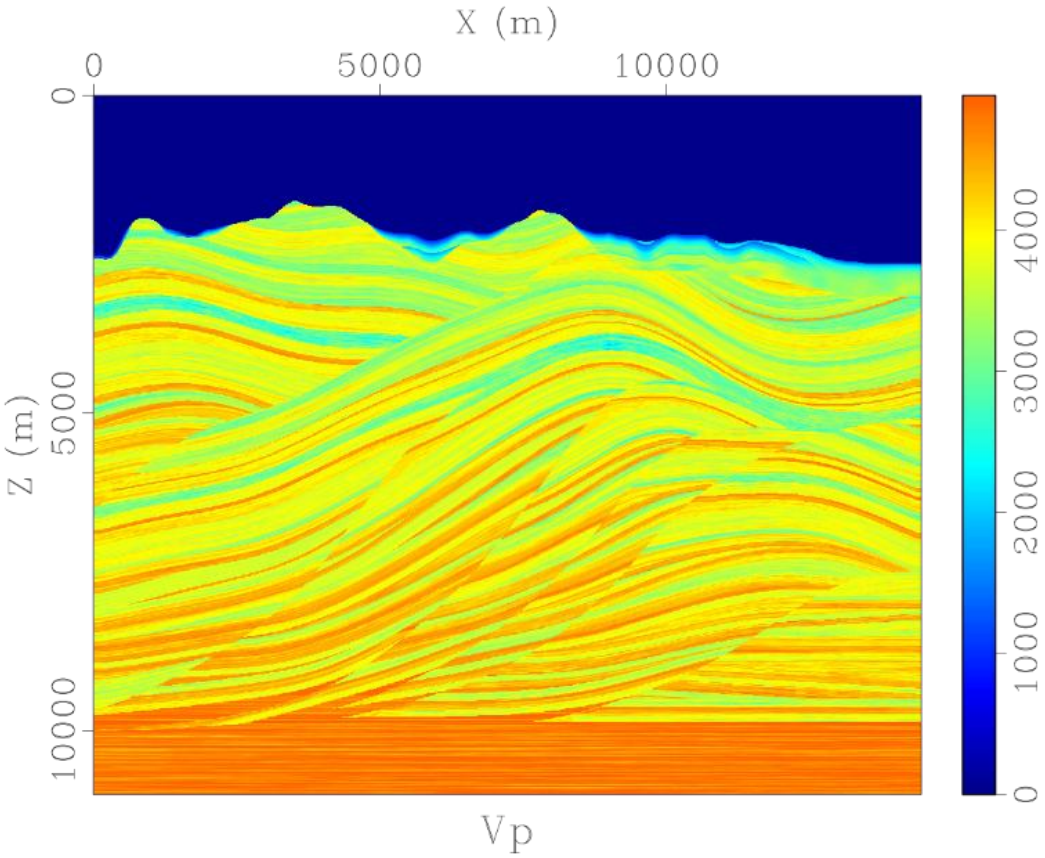


Figure 3: The velocity model provided by Sanchez et al. (2025) was used to generate multiple variations of similarly complex onshore velocity models.

Figure 4 shows a grid of velocity models generated by DALL-E 2 based on the initial model in Figure 3. The first observation is that the velocity models lack a colorbar. DALL-E 2 generates images based on pixel colours, using the RGB scheme, where each colour is represented by a combination of red, green, and blue values ranging from 0 to 255. Therefore, the generated images do not have a specific colour scale that can be directly translated to velocity values. To use these images as velocity models, a mapping between the pixel colours and velocity values would need to be established.

Some of the generated velocity models appear to be geologically plausible, with realistic layering and velocity variations (Figure 4 a, d, and e). However, others contain unrealistic features, such as abrupt velocity changes and inconsistent layering (Figure 4 b, c, and f). This indicates that while DALL-E 2 can generate complex velocity models based on an initial image and text prompts, it may still struggle to consistently produce geologically accurate representations. Although all of these models can be used to generate seismic data for training machine learning models, many numerical tests in the literature on seismic inversion, denoising, and other applications use simple velocity models that do not reflect real geology.

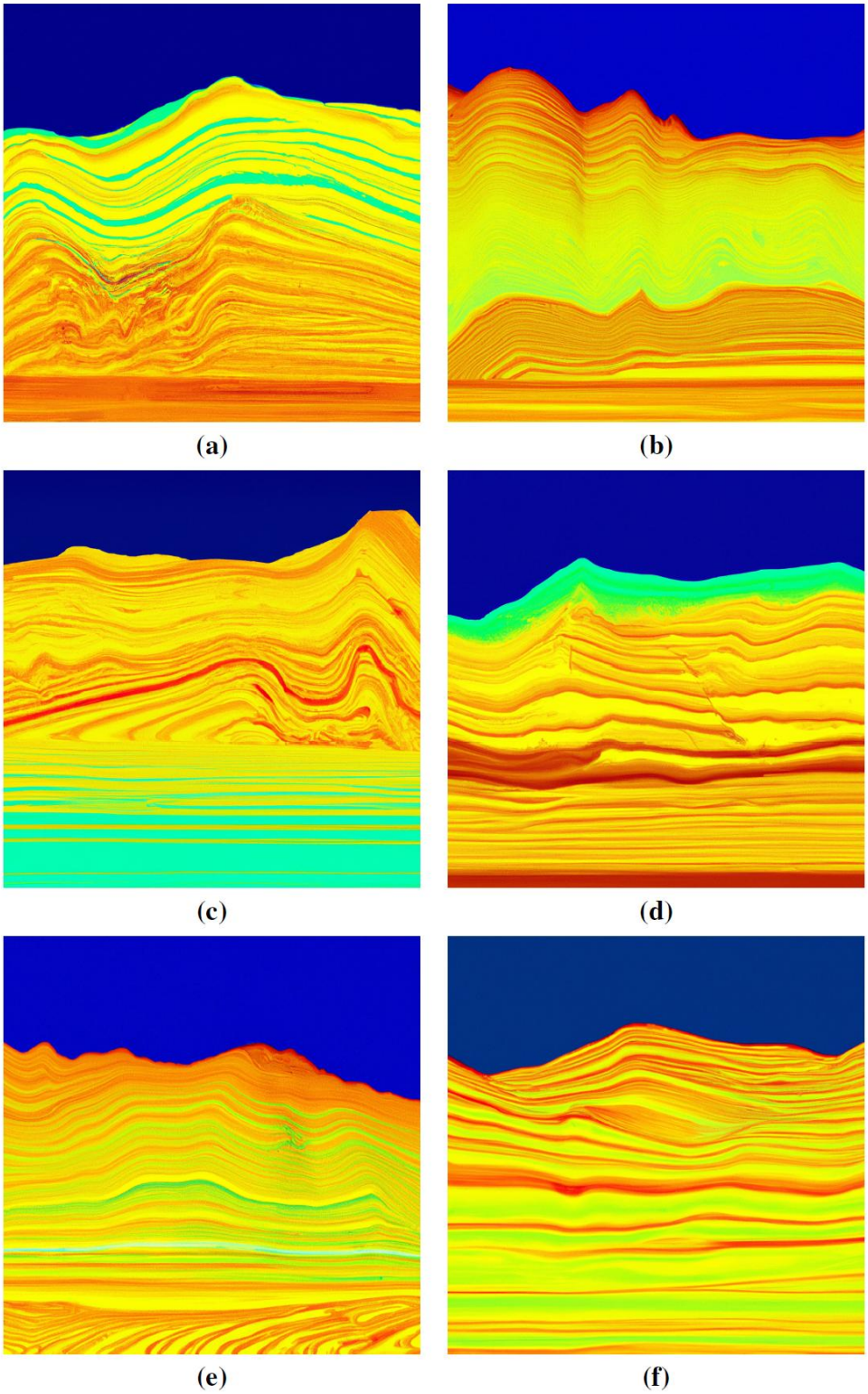


Figure 4: Output examples generated by DALL-E 2. Some velocity models generated are geologically plausible (a, d, e), while others contain unrealistic features (b, c, f).

Conclusions

In this work, we explored the use of large language models (LLMs) and multimodal models to generate synthetic P-wave velocity models for seismic data simulation. Despite extensive prompt engineering with GPT-5 and GPT-5-mini, the outputs were often unrealistic and inconsistent. While attempts with few-shot prompts did not improve results, using more examples or better-designed prompts may have helped, but we faced resource limitations.

On the other hand, DALL-E 2 demonstrated the ability to generate variations of velocity models based on an initial image and accompanying text prompts. This capability allows for the simultaneous creation of multiple images. However, some of these models lacked a clear relationship with velocity values and were not geologically plausible, highlighting the model's limitations in consistently producing velocity models.

Fine-tuning multimodal models like DALL-E 2 for geological applications could improve their realism. This would require a substantial dataset of unique velocity model images and their geological descriptions, which we plan to explore in future work.

Acknowledgements

We extend our heartfelt thanks to the sponsors of CREWES for their invaluable support. This work was made possible by generous funding from CREWES industrial sponsors and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 543578-19. We also acknowledge the additional support provided by Emissions Reduction Alberta through the ACT4-SPARSE project.

References

- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A. et al., 2020, Language models are few-shot learners: Advances in neural information processing systems, 33, 1877–1901.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J. et al., 2021, Learning transferable visual models from natural language supervision, in International conference on machine learning, PmLR, 8748–8763.
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M., 2022, Hierarchical text-conditional image generation with clip latents: arXiv preprint arXiv:2204.06125, 1, No. 2, 3.
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., and Sutskever, I., 2021, Zeroshot text-to-image generation, 2102.12092. URL https://arxiv.org/abs/2102.12092
- Sanchez, I., Guarido, M., and Trad, D., 2025, Simulating synthetic seismic data from AI-generated velocity models: CREWES Research Report, 37.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., 2017, Attention is all you need: Advances in neural information processing systems, 30.
- Versteeg, R., 1994, The Marmousi experience: Velocity model determination on a synthetic complex data set: The Leading Edge, 13, No. 9, 927–936.