https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

cigChannel: A massive-scale 3D seismic dataset with labeled paleochannels for advancing deep learning in seismic interpretation

Guangyu Wang[1, 2], Xinming Wu[1, 2], and Wen Zhang[1, 2]

[1]School of Earth and Space Sciences, University of Science and Technology of China, Hefei 230026, China
[2]Mengcheng National Geophysical Observatory, University of Science and Technology of China, Hefei 230026, China

Correspondence: Xinming Wu (xinmwu@ustc.edu.cn)

Abstract. Identifying buried channels in 3D seismic volumes is essential for characterizing hydrocarbon reservoirs and offering insights into paleoclimate conditions, yet it remains a labor-intensive and time-consuming task. The data-driven deep learning methods are highly promising to automate the seismic channel interpretation with high efficiency and accuracy, as they have already achieved significant success in similar image segmentation tasks within the field of computer vision (CV). However, unlike the CV domain, the field of seismic exploration lacks a comprehensive benchmark dataset for channels, severely limiting the development, application, and evaluation of deep learning approaches in seismic channel interpretation. Manually labeling 3D channels in field seismic volumes can be a tedious and subjective work and most importantly, many field seismic volumes are proprietary and not accessible to most of the researchers. To overcome these limitations, we propose a comprehensive workflow of geological channel simulation and geophysical forward modeling to create a massive-scale synthetic seismic dataset containing 1,200 256×256×256 seismic volumes with labels of more than 10,000 diverse channels and their associated sedimentary facies. It is by far the most comprehensive dataset for channel identification, providing realistic and geologically reasonable seismic volumes with meandering, distributary, and submarine channels. Trained with this synthetic dataset, a convolutional neural network (simplified from the U-Net) model performs well in identifying various types of channels in field seismic volumes, which indicates the diversity and representativeness of the dataset. We have made the dataset, codes generating the data, and trained model publicly available for facilitating further research and validation of deep learning approaches for seismic channel interpretation.

1 Introduction

Paleochannels are buried river channels that have been preserved in the geological record. They can serve as reservoirs for hydrocarbons (Clark and Pickering, 1996; Bridge et al., 2000; Hein and Cotterill, 2006) and provide insights into paleoclimate conditions (Leigh and Feeney, 1995; Nordfjord et al., 2005; Sylvia and Galloway, 2006). Paleochannels can be identified in seismic volumes by their distinct shapes and sedimentary structures that differ from the surrounding rock formations. Although paleochannels are considered as geobodies, interpreters are limited to view them slice-by-slice in seismic volumes. This limitation significantly increases the complexity and time of interpreting paleochannel bodies in large seismic volumes.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

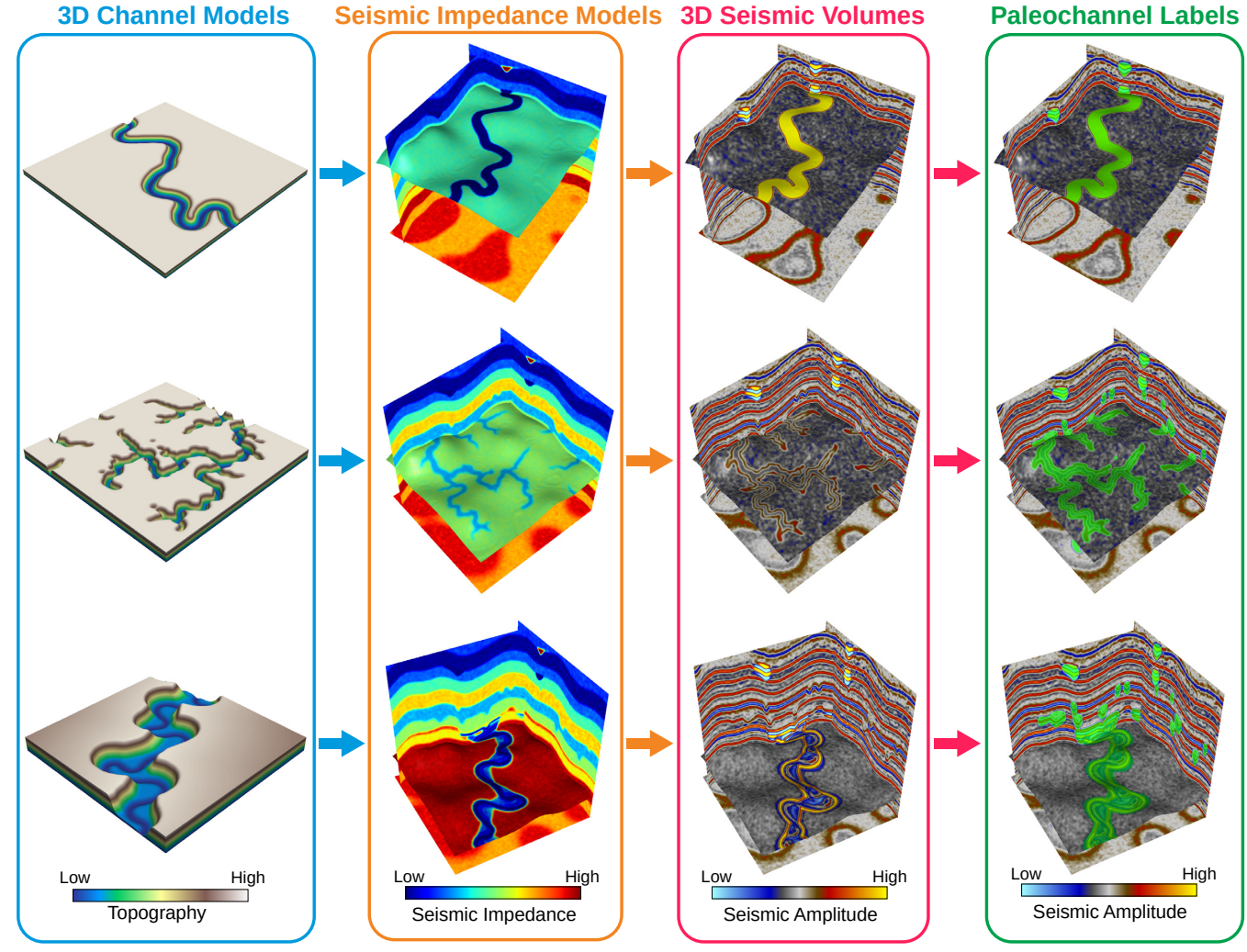


Figure 1. Workflow for generating the *cigChannel* dataset. First, we create topographic models of meandering, distributary and submarine channels. Second, we build 3D seismic impedance models with layered structure and place the channels at layer boundaries as impedance anomalies. Third, the impedance models are used to calculate seismic reflection coefficients, which are subsequently convolved with the Ricker wavelet to create seismic volumes. Finally, seismic reflections of the paleochannels are automatically labeled.

Moreover, the historical tectonic movement may introduce deformations such as foldings to the paleochannels, making them even more difficult to recognize.

To address those issues, automatic paleochannel identification methods based on 3D convolutional neural networks (CNNs) (Pham et al., 2019; Gao et al., 2021) are developed. The 3D CNNs are designed to capture volumetric features by performing 3D convolutions (Ji et al., 2012). They treat paleochannels as bodies rather than slices as human interpreters typically see, which gives them the advantage in identifying paleochannels deformed by historical tectonic movements. Another notable advantage

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

is their efficiency. Once trained, the network can rapidly identify paleochannels in a large seismic volume. However, the main limitation of applying CNNs for paleochannel identification is the lack of labeled paleochannel samples for training. Unlike deep learning for computer vision, which benefits from numerous large datasets with labeled images such as ImageNet (Deng et al., 2009) and COCO (Lin et al., 2014), currently there is no publicly available dataset of field seismic volumes with labeled paleochannels. To create such a dataset, one needs to have access to a large amount of field seismic volumes and correctly label the paleochannels. However, many field seismic volumes are proprietary and not available to most of the researchers (Vizeu et al., 2022). Besides, the complexity of field seismic volumes adds difficulty to correctly labeling the paleochannels. The label noise produced by mislabeling will deteriorate the performance of supervised learning (Pechenizkiy et al., 2006; Nettleton et al., 2010).

While training the networks with a large amount of labeled field seismic volumes is currently not an option, an alternative solution is to use the synthetic seismic volumes, which are generated by a series of simulation processes in order to mimic the field seismic volumes. Although lacking in sophisticated features, the synthetic seismic volumes are controllable, allowing us to tailor the objectives that we want the network to learn. Moreover, mislabeling can be avoided in synthetic seismic volumes since the locations of objectives are known during the simulation process. Synthetic seismic volumes have been proven effective as training data for networks to identify various objectives in field seismic volumes, such as faults (Wu et al., 2019; Zheng et al., 2019), seismic horizons (Bi et al., 2021; Vizeu et al., 2022), paleokarsts (Wu et al., 2020b; Zhang et al., 2024) and paleochannels (Pham et al., 2019; Gao et al., 2021). As for paleochannel identification, the synthetic seismic datasets created by Pham et al. (2019) and Gao et al. (2021) only simulate stacked and single meandering channels, respectively, while the frequently observed distributary (Payenberg and Lang, 2003; Li et al., 2016) and submarine (Deptuck et al., 2007; Gee et al., 2007) channels are not included. Considering the diversity of paleochannels in field seismic volumes, creating a dataset with various types of paleochannels is necessary for enhancing the networks' generalizability.

In this paper, we propose a comprehensive workflow (Figure 1) for generating a massive-scale dataset of synthetic seismic volumes and labels of diverse paleochannels. In this workflow, we first build numerous 3D models of meandering, distributary and submarine channels following the modeling methods developed by Howard and Knutson (1984), McDonald (2020) and Sylvester et al. (2011), respectively. Parameters that control the modeling process are randomized within a reasonable range in order to increase the diversity of channel models. Second, we build seismic impedance models with layered structure and place the channels at layer boundaries as impedance anomalies. Third, the impedance models are used to calculate seismic reflection coefficients, which are subsequently convolved with the Ricker wavelet to create synthetic seismic volumes. Finally, channels in the seismic volume can be automatically labeled since their positions are already known. Using this workflow, we have created a benchmark dataset named *cigChannel* for deep learning-based seismic paleochannel interpretation. This dataset, to our best knowledge, is by far the largest one that contains 1,200 $256\times256\times256$ seismic volumes and labels of more than 10,000 diverse paleochannels. The effectiveness of this dataset is validated by training a CNN to identify various types of paleochannels in field seismic volumes. It should be noted that although we have significantly improved the diversity of paleochannels compared with previous datasets, there is no guarantee that this dataset covers every form of paleochannel in

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

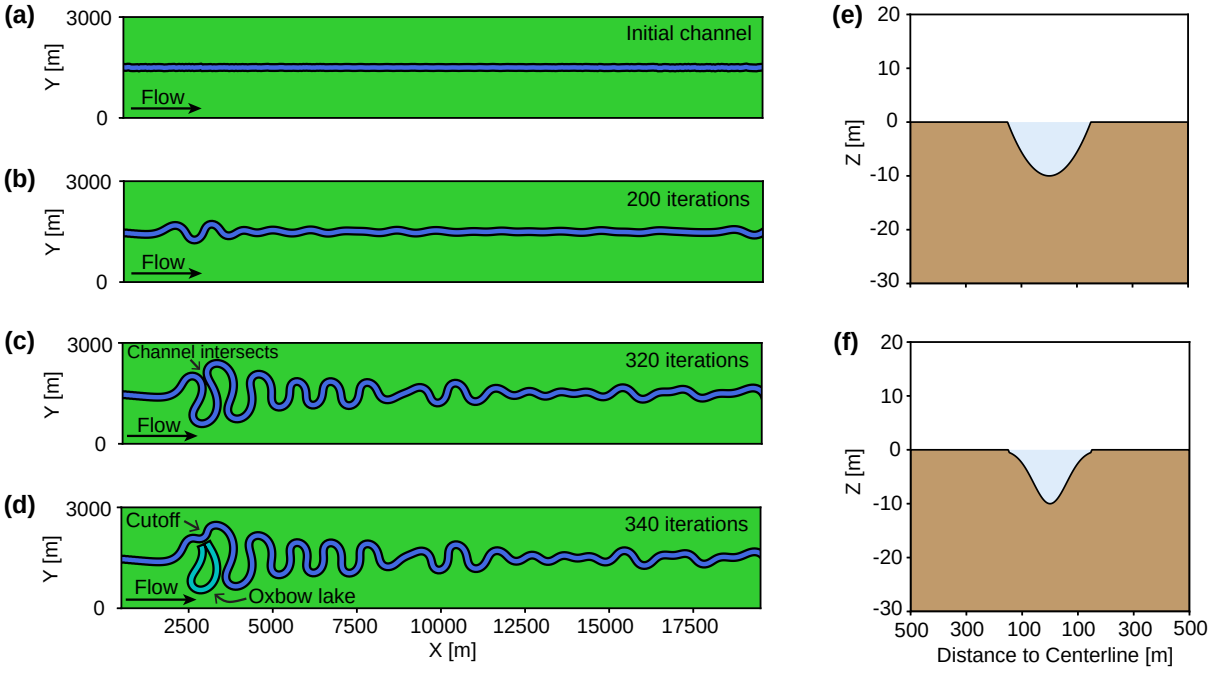


Figure 2. Meandering channel modeling process based on the open-source Python package *meanderpy* (Sylvester, 2021). First, we create (a) a straight channel with some minor perturbations. Then, (b) the channel begins to migrate, leading to the formation of multiple meanders. (c) The channel curvature increases as the migration continues, eventually causing a channel intersection, where (d) the channel cutoff will occur, forming the oxbow lake. Lastly, (e) the U- and (f) V-shaped channel cross-sections are used to define the channel topography.

field seismic volumes. Therefore, a Python package of the dataset generation workflow (see Appendix C for illustrative codes) is also provided for customizing the paleochannels and facilitating further development.

2 Dataset generation workflow

In this section, we will elaborate the dataset generation workflow to explain details of the geological and geophysical modeling in generating the dataset. First, we will describe the modeling process for meandering, distributary and submarine channels. Then, we will explain how to create synthetic seismic volumes based on these channel models, including designing folded impedance models with channels and simulating realistic seismic volumes.

2.1 Meandering channel modeling

Meandering channels are among the most frequently observed river channels distinguished by their sinuous paths. The continuous interaction between water and the riverbed can lead to erosion on the outer bank and deposition on the inner bank, causing the channel to migrate over time and increasing its curvature. The key to create a realistic meandering channel is to simulate its

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

migration. We use the open-source Python package *meanderpy* (Sylvester, 2021) for this purpose, which employs a kinematic simulation method that computes the river migration rate as a weighted sum of upstream curvatures (Howard and Knutson, 1984; Sylvester et al., 2019). The meandering channel simulation process is demonstrated in Figure 2. We start with a straight channel with some minor perturbations, which provide initial curvatures for channel migration (Figure 2a). The channel migrates over time and forms meanders at its upstream (Figure 2b). As the migration continues, curvature of the meander increases and eventually leads to channel intersection (Figure 2c), where the channel cutoff will occur and form the oxbow lake (Figure 2d). The channel migration ends when it reaches the maximum number of iteration. We neglect the oxbow lake and extract the centerline from a random segment of the most recent channel, which has to be long enough to span a 256×256 square grid with a cell size of 25 m after arbitrary rotation.

The centerline is randomly placed on the grid and rotated by a random angle between 0°and 360°. We define the channel topography using the simplified U- and V-shaped channel cross-sections (Figures 2e and 2f). The U-shaped channel is typically found in gentle terrain, formed mainly by lateral erosion. On the contrary, the V-shaped channel usually appears in areas with steep gradients, shaped primarily by vertical erosion. The simplified U-shaped channel is defined as a parabolic function:

$$Z(x) = \begin{cases} 4D_c(x/W_c)^2 - D_c, & x \le W_c \\ 0, & x > W_c \end{cases}, \tag{1}$$

where x is the Euclidean distance from the centerline to any point on the grid, D_c is the maximum depth of the channel (which will be denoted as channel depth hereafter for simplicity) and W_c is the channel width. The simplified V-shaped channel is defined as a combination of Gaussian and parabolic functions:

$$Z(x) = \begin{cases} \min[p(x), g(x)], & x \le W_c \\ 0, & x > W_c \end{cases}, \tag{2}$$

where $p(x)$ is the parabolic function in Equation (1) and

$$g(x) = -D_c e^{-\frac{x^2}{2(W_c/4)^2}}. \tag{3}$$

Although these simplified channel cross-sections may not precisely represent the real ones, they can capture the main features at a low computational cost. We create diverse topographic models of the meandering channel by randomizing the modeling parameters within a reasonable range (see Table A1). Some examples are demonstrated in Figure 5a, showing various meandering channels with different widths, depths and meander wavelengths.

2.2 Distributary channel modeling

Distributary channels are commonly observed in river deltas, where the river channel splits into multiple smaller channels as it approaches the river mouth and spreads out into the sea or lake. Numerous numerical modeling methods based on hydrodynamics and morphodynamics have been proposed to simulate river deltas and the associated distributary channels

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

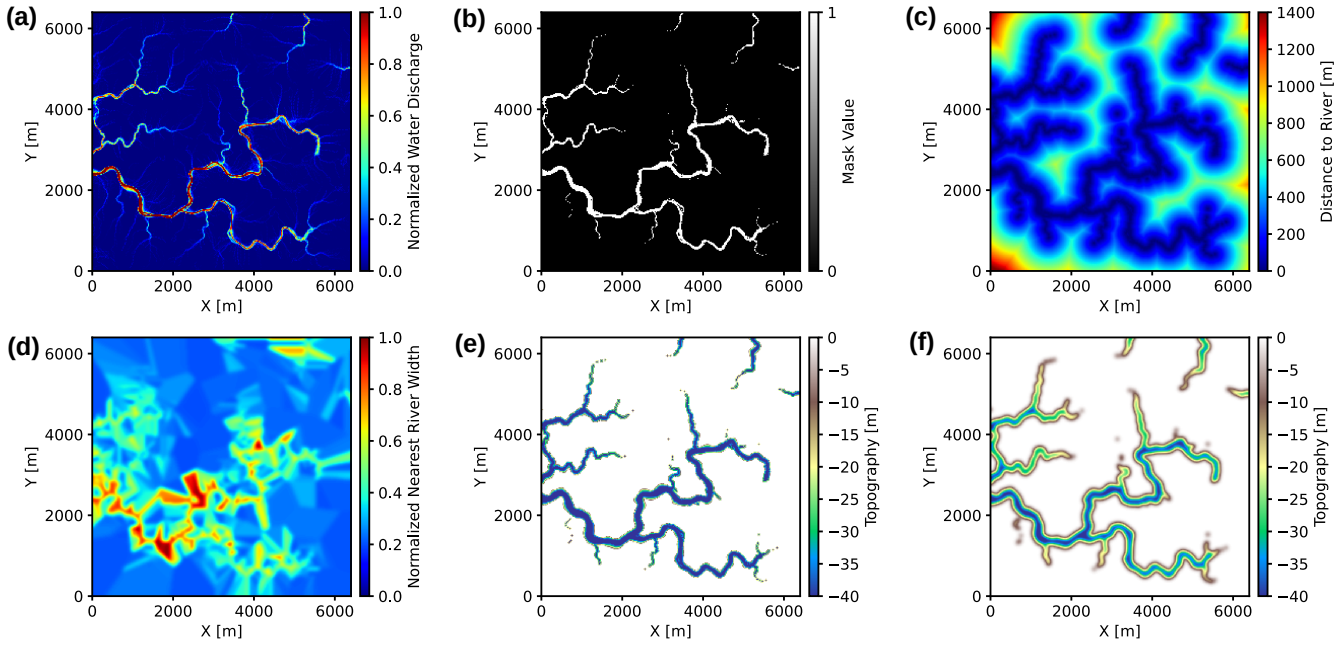


Figure 3. Distributary channel modeling process based on the open-source C++ package *soillib* (McDonald, 2020). First, we generate (a) a map of normalized water discharge using the *soillib* package. Second, we create (b) the river mask by binarizing the normalized water discharge with a threshold value of 0.4, where values greater than this threshold are considered as rivers. Third, we compute (c) the Euclidean distance to rivers and (d) the normalized width of the nearest river, which are subsequently used as parameters in a parabolic function to define (e) the channel topography. Finally, to avoid abrupt topographic shifts, a Gaussian filter is applied to create (f) a smoothed channel topography.

(Seybold et al., 2007; Edmonds and Slingerland, 2007; Geleynse et al., 2011; Liang et al., 2015). However, these methods are time-consuming since they are designed to simulate detailed fluid dynamics. To efficiently generate a large number of distributary channel models, we adopt the open-source C++ package *soillib* (McDonald, 2020), which is a fast implementation of particle-based hydraulic erosion.

The *soillib* is programmed to spawn hundreds of thousands of water particles at random positions on a terrain generated by layered random Perlin noise. The particles move across the terrain following classical mechanics and engage in mass transfer with the surface, eventually forming the distributary rivers. Figure 3a shows the normalized water discharge map of a distributary river network generated by the *soillib* package on a 256×256 square grid with a cell size of 25m. To define the river channel topography, we first binarize the water discharge by a threshold (e.g., 0.4), where values greater than this threshold are considered as rivers (Figure 3b). Next, we compute the Euclidean distance from the river to each point on the grid (Figure 3c) and the normalized width of the nearest river (Figure 3d), which is represented by the normalized water discharge. We then

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

define the channel topography using a parabolic function similar to that in Equation (1):

$$Z_i(x_i) = \min[4D_c(\frac{x_i}{W_c\alpha_i})^2 - D_c, 0], \quad (4)$$

where the subscript i denotes the i-th point on the grid, x is the distance to river, D_c is the channel depth, W_c is the maximum channel width and α is the normalized width of the nearest river. The main modification is replacing the constant channel width W_c with a point-wise channel width $W_c\alpha_i$. By doing so, we are able to create channels with varying widths, as demonstrated in Figure 3e. The variation in channel width is controlled by α, where the mainstream is wider and the distributaries are narrower. However, the channel topography demonstrated in Figure 3e exhibits abrupt shifts at the channel edge due to the inherent width of the river mask. Therefore, we subsequently apply a Gaussian filter to smooth it and the final channel topography is shown in Figure 3f. When implementing the particle-based hydraulic erosion, randomness in the initial terrain and positions of water particles ensure the diversity of distributary channels, which is demonstrated in Figure 5b. Diversity of the channel topographic models can be further increased by using random channel widths and depths within a reasonable range (see Table A1).

2.3 Submarine channel modeling

The submarine channel is a type of underwater channel formed on the ocean floor, particularly on the margin of continental shelf. These channels are primarily carved out by turbidity currents, which carry loads of sediment from shallow coastal areas and move downslope to deeper parts of the ocean under the influence of gravity. Similar to terrestrial river channels which meander across the floodplain, submarine channels also exhibit meandering patterns on the ocean floor, especially in areas of gentle slope. The meanders of submarine channels also migrate laterally and undergo cutoffs. However, a major difference between terrestrial and submarine channels lies in the significant vertical incision and aggradation of submarine channels, driven by the powerful erosive and depositional processes associated with turbidity currents. As a result, submarine channels generally pocess a large-scale erosional surface and a layered structure within the channel, which are discernible in high-resolution seismic profile (Kolla et al., 2007).

To model the large-scale erosional surface and layered structure within the channel, we adopt the modeling method based on submarine channel trajectories (Sylvester et al., 2011), which is also implemented in the *meanderpy*. The modeling process is illustrated in Figure 4. We first simulate the lateral migration of the submarine channel (Figure 4a) using the same algorithm to simulate that of the meandering channel. At each iteration during the migration process, a parabolic function shown in Equation (1) is used to define the surface of channel erosion (Figure 4b), which is followed by the deposition of point bars and natural levees (Figure 4c). Point bars are accumulated sediments on the inner bends of the channel where the flow velocity is lower. Their top surface is defined using a combination of parabolic and Gaussian function as shown in Equation (2) and (3). For modeling convenience, point bars are created on both inner and outer bends, with those on the outer bends will be subsequently eroded. Natural levees are structures that form along the sides of submarine channels when the turbidity currents overflow the channel banks. They typically exhibit a wedge-like shape because the turbidity currents lose energy and sediments as they

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

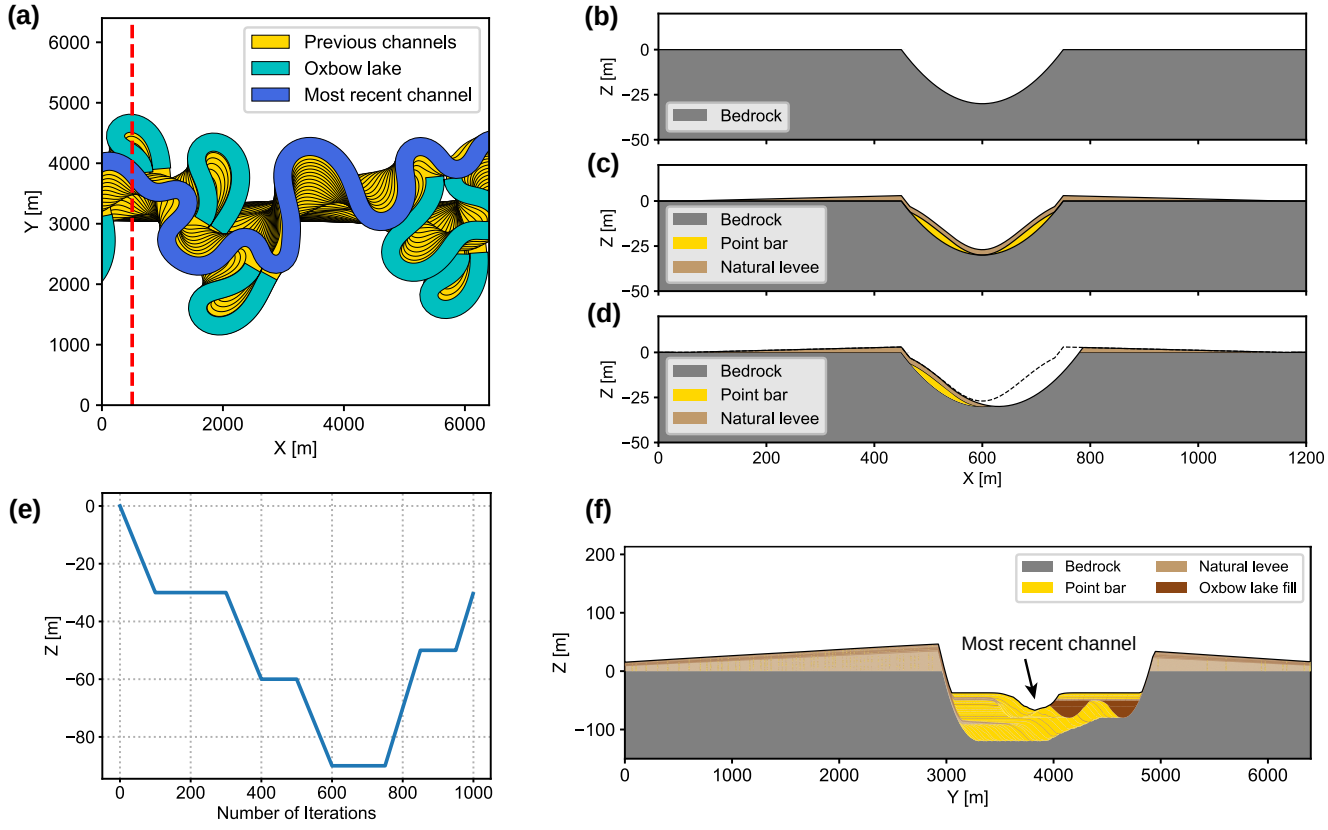


Figure 4. Submarine channel modeling using the open-source Python package *meanderpy* (Sylvester, 2021). (a) Lateral migration of the submarine channel. (b) Channel erosional surface. (c) Deposition of point bars and natural levees. (d) The channel migrates towards the outer bend and erode parts of the sediments. (e) Vertical component of the channel trajectory during the migration process, which is modified from Sylvester et al. (2011), showing an initial channel incision and a later aggradation. (f) The channel cross-section at the red dashed line in (a), showing a large-scale erosional surface, a layered structure within the channel and a wedge-like natural levee after 1,000 times of migration.

145 move away from the channel margins. The natural levee thickness is defined as follows:

$$T(x) = \begin{cases} \frac{T_{\max}}{W_l}(x - \frac{W_c - W_l}{2}), & x \geq W_c \\ T_{\max}, & x < W_c \end{cases}, \tag{5}$$

where x denotes the distance to channel centerline, $T_{\max}$ is the maximum levee thickness, W_l is the levee width on one side of the channel and W_c is the channel width. After the deposition of point bars and natural levees, the channel will migrate towards its outer bends and erode parts of these sediments (Figure 4d). The erosion and deposition processes repeat until the channel
150 migration ends. In the meantime of lateral migration, the channel also experience vertical incision and aggradation (Figure 4e). At the end of migration, the submarine channel will exhibit a large-scale erosional surface, a wedge-like natural levee, and a

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

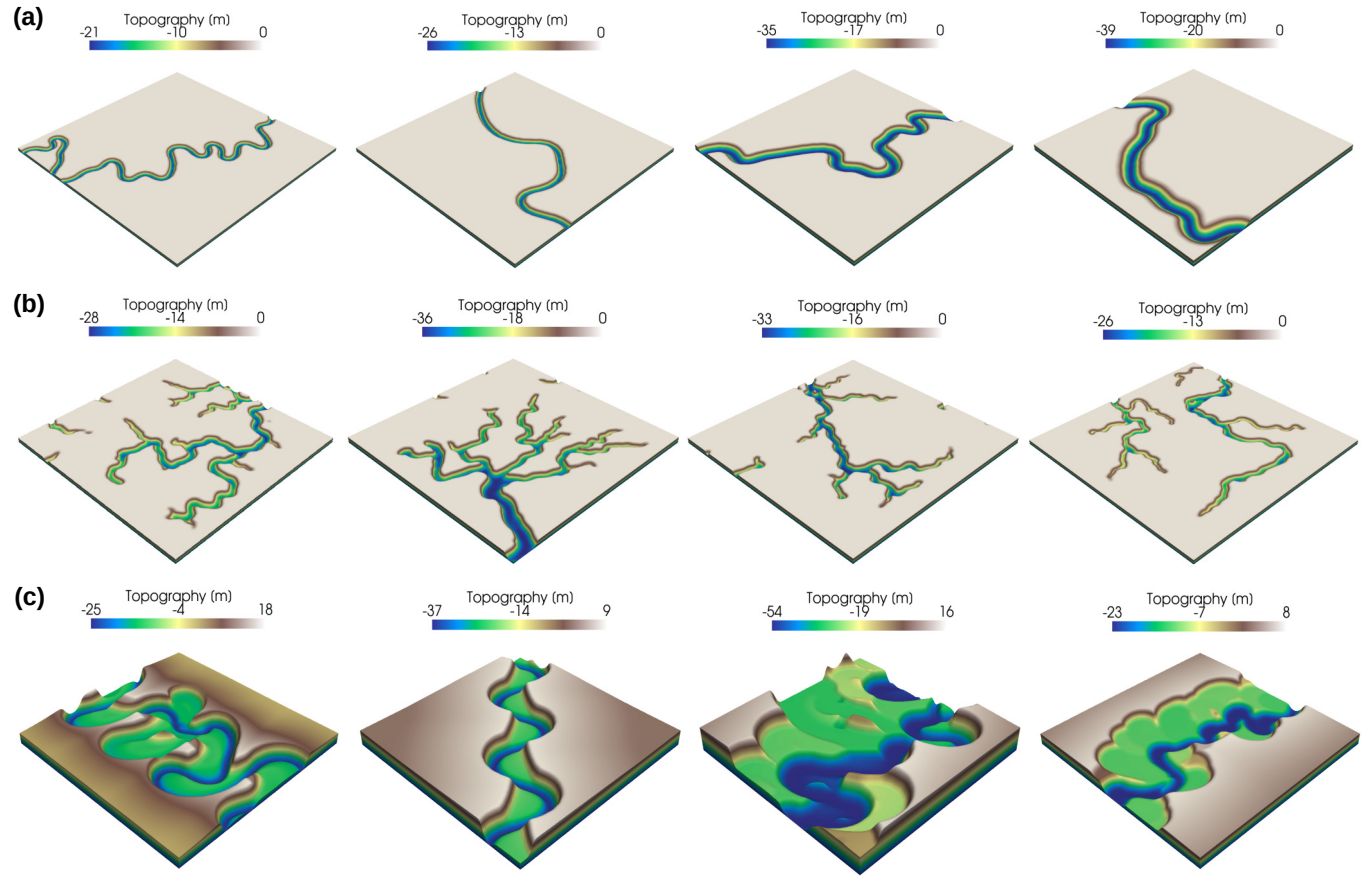


Figure 5. Diverse topographic models of (a) meandering, (b) distributary and (c) submarine channels.

layered structure within the channel, which consists of oxbow lake sediments and interbedded layers of point bars and natural levees (Figure 4f). To create diverse submarine channels, we use a random set of modeling parameters within a reasonable range (see Table A1), and some of the resulting topographic models are demonstrated in Figure 5c.

2.4 Seismic volume simulation

After constructing over 10,000 topographic models covering meandering, distributary and submarine channels, we proceed to create synthetic seismic volumes based on these models. The first step is to define the seismic impedance, which is a crucial parameter for simulating seismic events. In seismic exploration, seismic waves from an artificial source travel through the subsurface rock mass, and part of the waves will be reflected back to the surface at the boundaries of two geological layers with a contrast in seismic impedance. The reflected seismic waves will form the seismic events, which are considered to be representatives of layer boundaries, and their amplitudes are related to the contrast in seismic impedance. We start by generating 3D seismic impedance models with horizontal layers. In each layer, we add some minor random perturbations to the

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

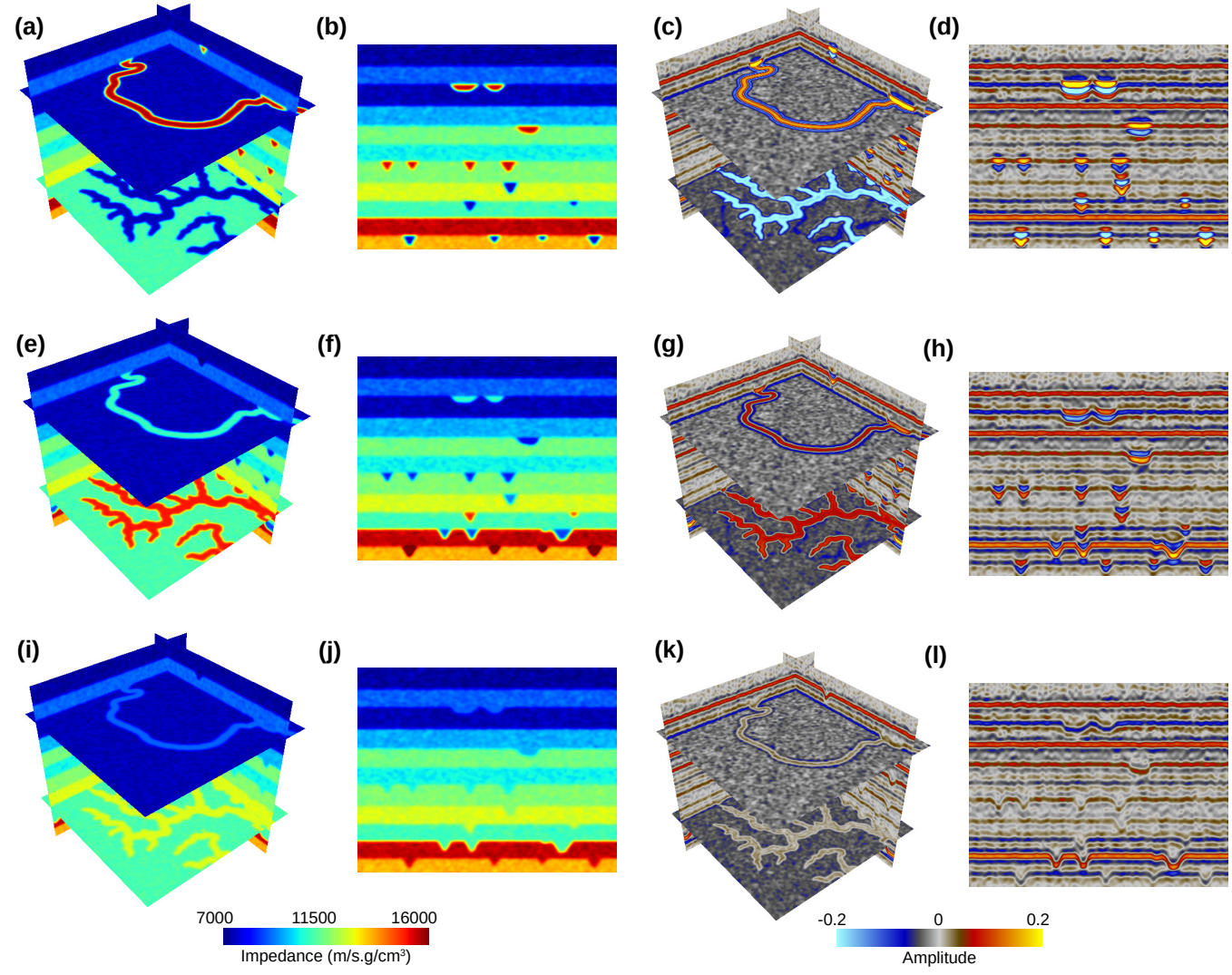


Figure 6. Seismic impedance and amplitude volumes of meandering and distributary channels, showing different levels of impedance contrast between the channel and its covering layer. (a) to (d) correspond to channels with high impedance contrast, (f) to (h) correspond to channels with low impedance contrast, and (i) to (l) correspond to channels with no impedance contrast.

impedance to make it more realistic. Details about the configuration of the impedance model are listed in Table D1. The channel topographic models are then placed at the layer boundaries, and the seismic impedance of the channel is defined according to the channel type.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Within meandering and distributary channels, we fill them with uniform impedance. The impedance value is determined by a parameter ε, which is defined as the impedance contrast between the channel and its covering layer:

$$\varepsilon = \frac{|Z_f - Z_u|}{Z_u}, \tag{6}$$

where Z_f denotes the impedance filling in channels, and Z_u denotes the impedance of the covering layer of the channel. The value of ε varies between zero and one, with the value of one indicating the highest impedance constrast between the channel and its covering layer, and the value of zero indicating the impedance of channel is the same as that of its covering layer. Figures 6a and 6b demonstrate the horizontal and vertical slices of a 3D impedance model, which consists of meandering and distributary channels with high impedance contrast ($\varepsilon = 1$). The impedance model is then used for computing the seismic reflectivity as follows:

$$R_i = \frac{Z_{i+1} - Z_i}{Z_{i+1} + Z_i}, i = 1, 2, ..., N-1, \tag{7}$$

where the subscript i denotes the i-th point in the vertical direction of the model, and N denotes the total number of points in the vertical direction. The reflectivity model is subsequently convolved with the Ricker wavelet (see Figures 8a and 8b for examples), which is commonly used to create synthetic seismic data. The mathematical expression of the Ricker wavelet in depth-domain is:

$$f(s) = (1 - 2\pi^2 k_m^2 s^2) e^{-\pi^2 k_m^2 s^2}, \tag{8}$$

where s denotes the distance and k_m denotes the peak wavenumber of the wavelet. Figure 6c shows the synthetic seismic volume corresponding to a high impedance contrast between the channel and its covering layer. We can observe that the channels have strong seismic amplitudes, appearing as bright spots on the vertical slice of the seismic volume (Figure 6d). As the value of ε decreases to 0.2, the impedance contrast between the channel and its covering layer becomes lower, as shown in Figures 6e and 6f. The corresponding seismic volume (Figure 6g) also indicates a reduction in seismic amplitude of the channels, which exhibit an infilling feature on the vertical slice of the seismic volume (Figure 6h). When the value of ε is set to zero, the impedance of channel will be the same as that of its covering layer (Figures 6i and 6j). As a result, the channels show no seismic response except at their erosion boundaries (Figure 6k), and an incision feature can be observed on the vertical slice of the seismic volume (Figure 6l).

The impedance of submarine channels is determined based on their sedimentary facies, which include point bars, natural levees and oxbow lakes. Figure 7a shows the sedimentary facies of a submarine channel, which is primarily filled with layers of point bars as a result of continuous channel migration. Additionally, the channel is also filled with oxbow lake sediments and inner natural levees. As shown in Figure 7b, the point bars are assigned lower impedance because they generally consist of sand, whereas the natural levees and oxbow lake sediments are assigned higher impedance due to their muddy composition. The reference impedance ranges of the point bars, natural levees and oxbow lakes are listed in Table D1. It should be noticed that an impedance discrepancy exists between the neighboring layers of point bars, such that the channel will exhibit a layered feature on the vertical slice of seismic volume and a meander belt on the horizontal slice, as shown in Figure 7c.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

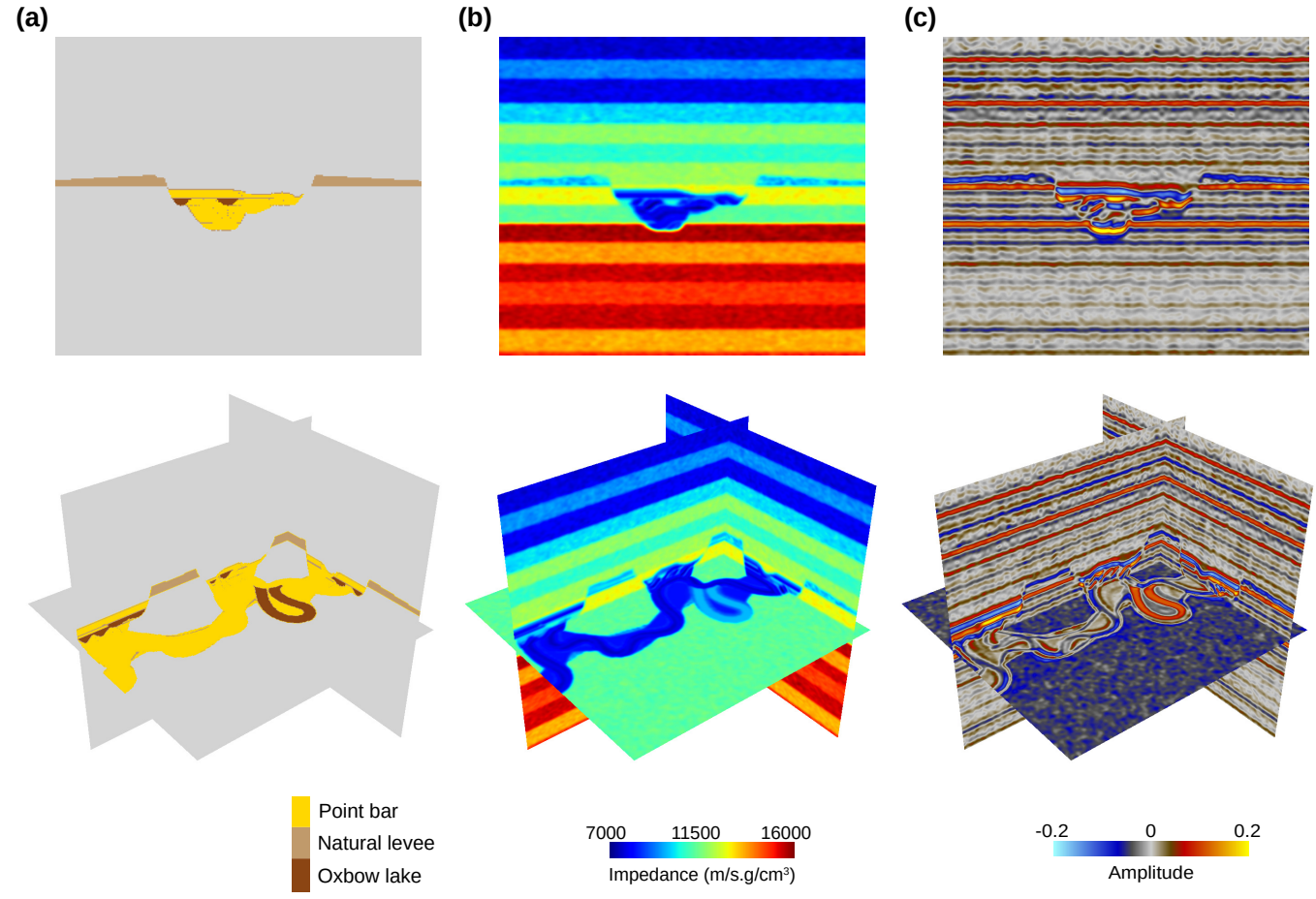


Figure 7. Illustration of creating a seismic impedance model and seismic amplitude volume containing a submarine channel according to its sedimentary facies, showing vertical and orthogonal slices of the (a) sedimentary facies, (b) seismic impedance and (c) amplitude volume.

By far, all the channels and layers in the impedance model are horizontal. However, the channels and layers in practice often undergo structural deformations, such as inclination and folding, which can be observed in many field seismic volume. To increase the diversity and realism of synthetic seismic volumes, we introduce inclination and folding into the impedance model following the workflow proposed by Wu et al. (2020a). An example of the resulting impedance model with inclined and folded layers is shown in Figure 8c. Another way to increase the diversity of synthetic seismic volumes is to use wavelets with various peak wavenumbers. This is also necessary because the peak wavenumber of seismic waves reflected by the channel can be diverse in field seismic volumes. It depends on various factors, such as the absorption effect of subsurface media and the characteristics of the seismic source. Figure 8 shows two synthetic seismic profiles with different wavelets computed from the same impedance model. Using a wavelet with small peak wavenumber (Figure 8a) will generate a low-resolution seismic profile with thick seismic events (Figure 8d), where some thin layers within the submarine channel at the bottom part of the

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

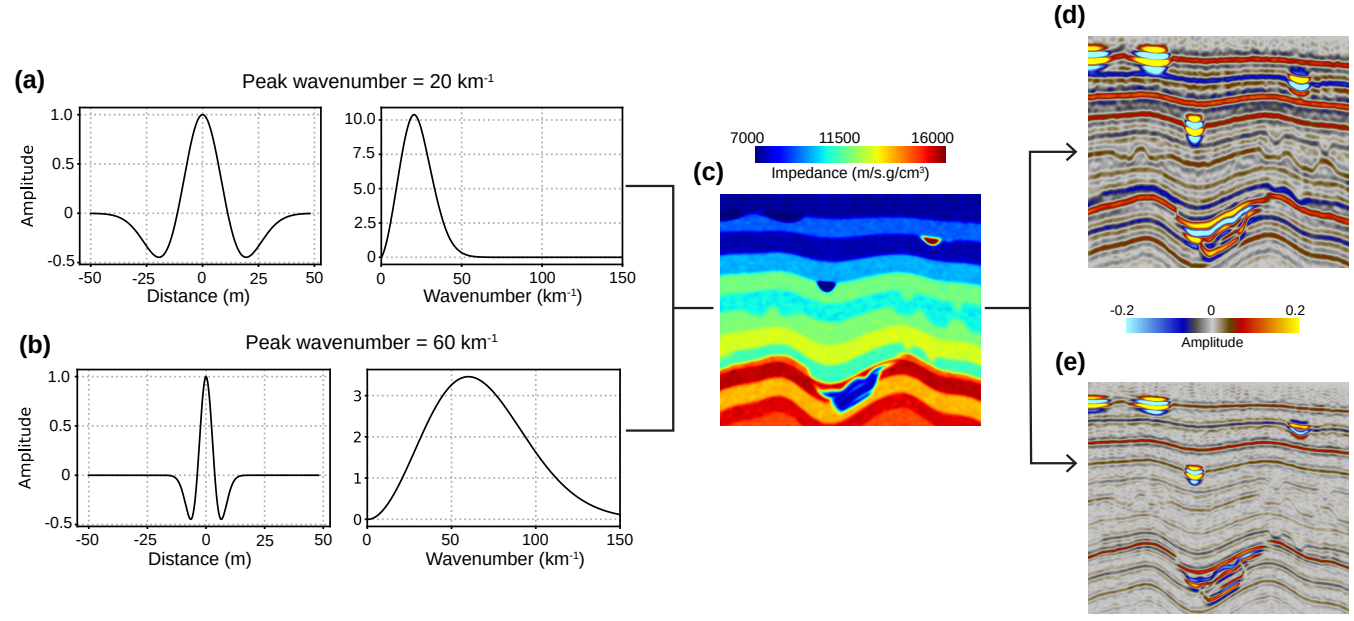


Figure 8. Synthetic seismic profile with different wavelets computed from the same seismic impedance model. (a) A small-wavenumber Ricker wavelet with a peak wavenumber of 20 km^{-1} in depth-domain and wavenumber-domain. (b) A large-wavenumber Ricker wavelet with a peak frequency of 60 km^{-1} in depth-domain and wavenumber-domain. (c) Seismic impedance model with inclined and folded structure. (d) Low-resolution seismic profile generated by using the small-wavenumber wavelet. (e) High-resolution seismic profile generated by using the large-wavenumber wavelet.

profile is hard to distinguish. On the contrary, using a large-wavenumber wavelet (Figure 8b) will create a high-resolution seismic profile (Figure 8e), where those thin layers within the submarine channel become discernible. The peak wavenumber range of the Ricker wavelet that used to generate the synthetic seismic volume is listed in Table D1.

3 Results

Using the aforementioned workflow, we create the *cigChannel* dataset containing 1,200 synthetic seismic volumes with more than 10,000 labeled paleochannels. Each seismic volume has a size of 256×256×256. Four task-specific subsets are included in the *cigChannel* dataset, namely the meandering, distributary, submarine and assorted channel subsets, whose detailed components can be found in Table B1.

Aiming to train deep learning models to identify specific types of channels, each of the meandering, distributary and submarine channel subsets provides 300 seismic volumes containing the corresponding type of channel. Binary class labels are provided in these subsets, which are designed solely to distinguish between channels and the background (i.e. the non-channel areas). As shown in Figure 9, each subset contains seismic volumes featuring horizontal, inclined and folded structures, serving

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

as training data for deep learning models to identify channels with various structures. The inclined and folded structures are randomly generated to introduce variability in the seismic volumes. The number of channels in each individual seismic volume varies according to the size of channel. A single seismic volume may contain multiple meandering or distributary channels yet no more than three submarine channels.

The assorted channel subset contains 300 seismic volumes with multi-class channel labels. It is designed to train deep learning models not only to identify but also to distinguish terrestrial and submarine channels in seismic volumes, which is important because they are indicators for different environments. As shown in Figure 9d, the terrestrial channels, which are represented by meandering and distributary channels in this dataset, have different characteristics from those of submarine channels. The most apparent one is their difference in size. Submarine channels are generally larger than terrestrial channels for many reasons. For instance, the turbidity currents that form the submarine channels are denser than their terrestrial counterparts, and the absence of vegetation on the ocean bottom eliminates a main limitation on channel erosion and sediment transport. Regarding the potential problems of the class imbalance problem and the size discrepancy between terrestrial and submarine channels, we simulate multiple terrestrial channels but only one submarine channel in a single seismic volume in order to make their voxel amounts as balanced as possible. However, there is still a huge gap in voxel amounts between channels and the background. Therefore, it is suggested to adopt strategies for addressing the class imbalance problem when using the dataset to train a deep learning model, such as employing the weighted loss functions.

4 Applications

We use the *cigChannel* dataset to train a simplified U-Net and apply it to identify paleochannels in field seismic volumes. This is a preliminary test mainly to verify the effectiveness of the dataset for training a deep learning model to distinguish between channels and the background in a field seismic volume. Therefore, the multi-class labels in the assorted channel subset are converted into binary labels like those in the other subsets. Architecture of the simplified U-Net is demonstrated in Figure 11, which has fewer convolutional layers and feature maps than its original architecture proposed by Ronneberger et al. (2015) to save memory and computational costs. The input is a $256\times256\times256$ seismic volume. Gaussian random noise is added to the seismic volume to make the training process more robust and reduce the tendency towards overfitting. The noisy seismic volume goes through the contracting path and expansive path of the U-Net for feature extraction. The final output layer is a $1\times1\times1$ convolutional layer followed by a sigmoid activation to map the feature vector into channel probability values. Regarding the huge gap in voxel amounts between channels and the background, we use the balanced cross-entropy as the loss function for network training.

We first use the trained U-Net to identify meandering channels in a seismic volume from the Parihaka seismic survey, which is a publicly available dataset provided by the New Zealand Crown Minerals. As demonstrated in Figure 11a, the seismic volume shows several meandering channels and the sediments of their river mouths where they enter the ocean. Channel identification result of the U-Net is shown in Figure 11b, where the meandering channels are all mapped with moderate to high probability. However, there is a mistaken identification of a fault at the bottom left corner of the image. This is probably

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

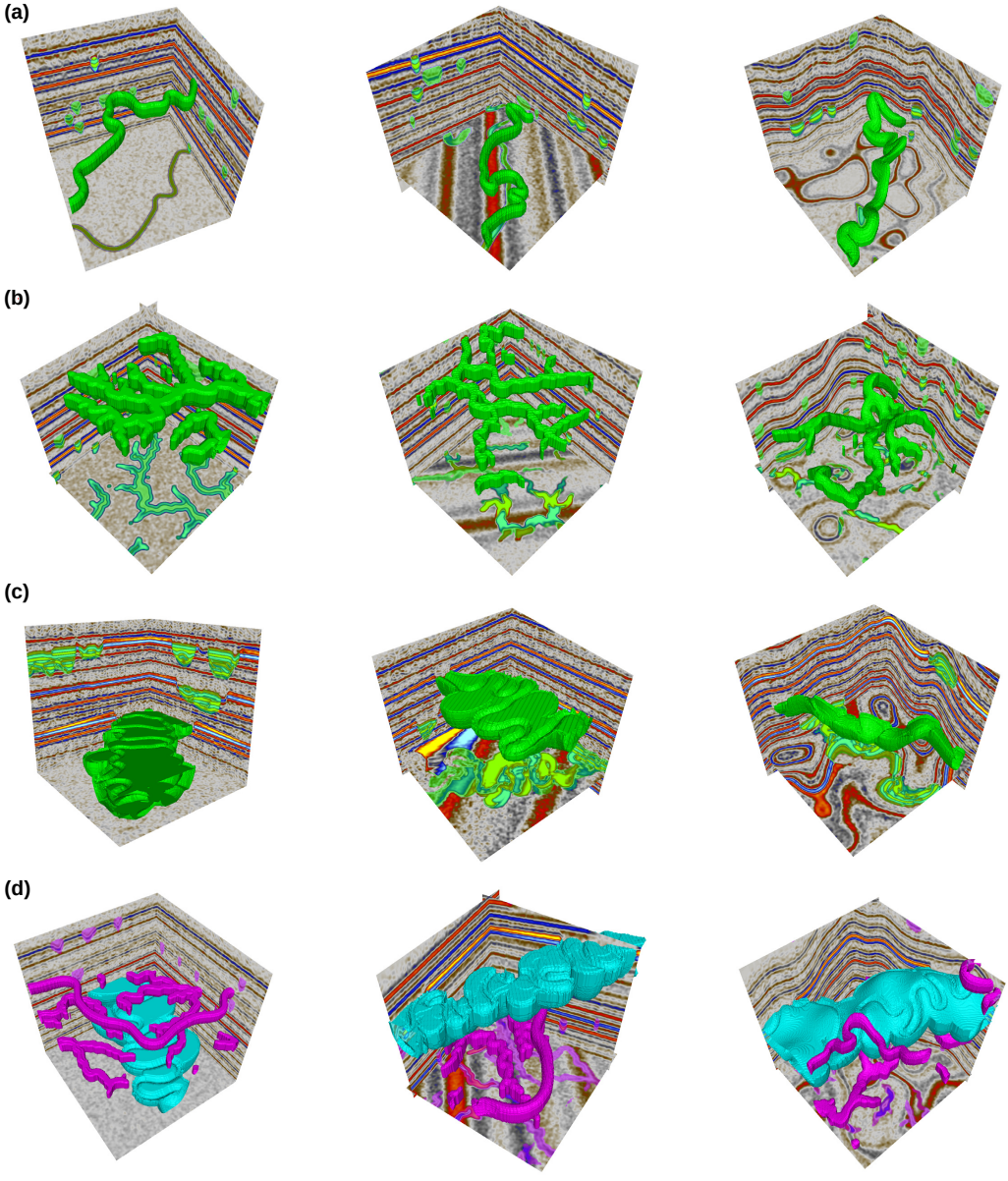


Figure 9. Synthetic seismic volumes and paleochannel labels (visualized as coloured masks and bodies) from (a) the meandering, (b) distributary, (c) submarine and (d) assorted channel subsets of the *cigChannel* dataset, showing horizontal, inclined and folded structures. Each of the meandering, distributary and submarine channel subsets provides binary class labels to distinguish between channels and the background (i.e. the non-channel areas), while the assorted channel subset provides multi-class labels to distinguish between terrestrial, submarine channels and the background.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

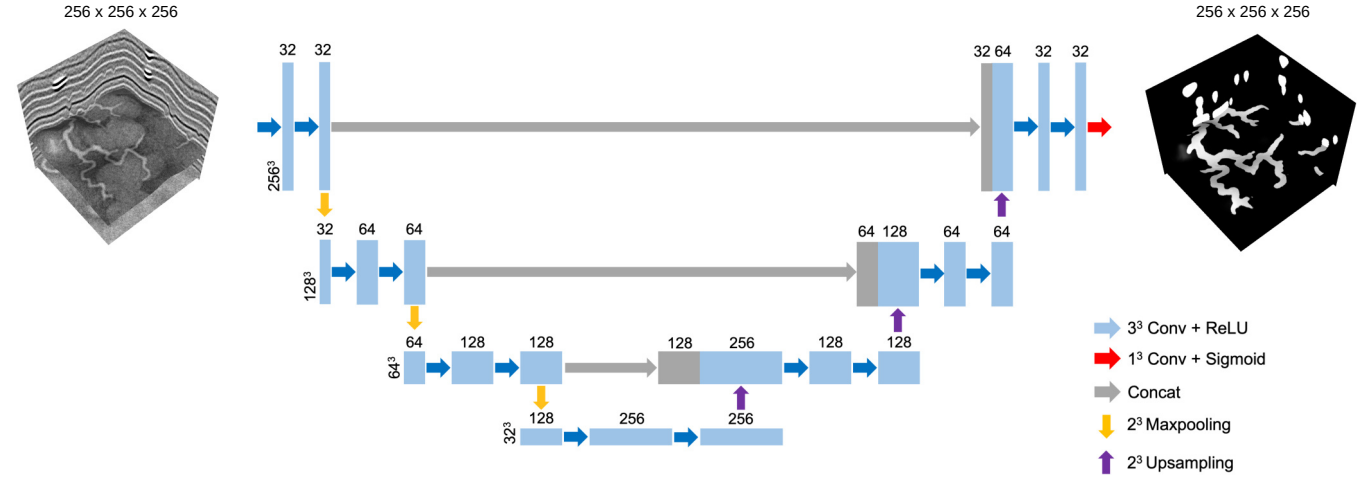


Figure 10. A simplified U-Net for paleochannel identification in seismic volumes.

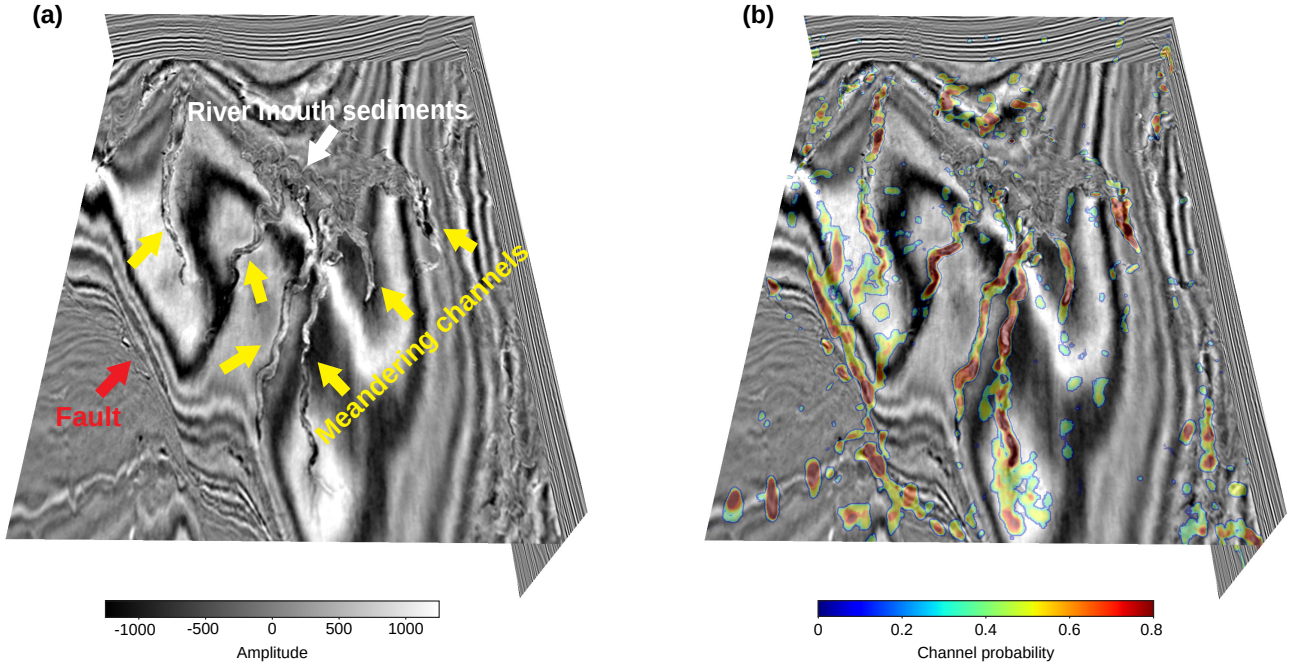


Figure 11. (a) Field seismic volume from the Parihaka seismic survey (courtesy of New Zealand Crown Minerals), showing multiple meandering channels (indicated by the yellow arrows), their river mouth sediments (indicated by the white arrow) and a nearby fault (indicated by the red arrow). (b) The channel identification result of the U-Net trained by the *cigChannel* dataset.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

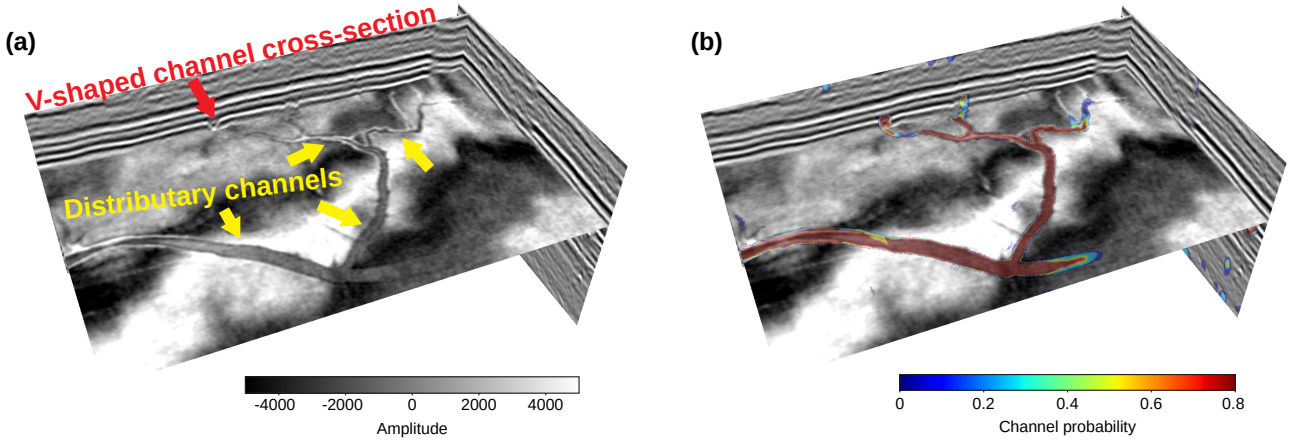


Figure 12. (a) Field seismic volume acquired in the Tarim basin (courtesy of China National Petroleum Corporation), showing several distributary channels (indicated by the yellow arrows) with a V-shaped cross-section (indicated by the red arrow). (b) The channel identification result of the U-Net trained by the *cigChannel* dataset.

because the layers are dragged downward by the normal faulting, making them exhibit an incised feature like the channels. Other noisy clusters with high channel probability may indicate segments of channels which are separated by folds or faults.

In the second example, the network is applied to identify distributary channels in a seismic volume acquired in the Tarim basin, which is provided by China National Petroleum Corporation. As demonstrated in Figure 12a, this seismic volume shows several distributary channels with a V-shaped cross-section. Seismic amplitudes within the channel are homogeneous, indicating a relatively uniform seismic impedance within the channel as we designed in our dataset generation workflow. The channel identification result of the U-Net is demonstrated in Figure 12b, showing that most of the channels are correctly identified except some extremely narrow branches.

In the last example, we identified a submarine channel in the seismic volume from the Parihaka seismic survey, which is pointed out by the yellow arrows in Figure 13a. Its large-scale erosional surface can be seen on the vertical slice of the seismic volume, which is distinct from the small-scale erosional surface of terrestrial channels, such as the one indicated in Figure 13a. This submarine channel has a medium to low seismic amplitude compared with that of its surrounding layer, which indicates a low discrepancy in seismic impedance within the channel. However, a layered structure is still visible within the channel. The horizontal slice that intersects this submarine channel shows its meander belt with a notable boundary. Figure 13b demonstrates the channel identification result of the U-Net. We observe that most areas of the submarine channel are correctly mapped with high probability.

These applications also reveal some limitations of this dataset. As indicated in Figure 11, the network trained by our dataset cannot discriminate faults and channels, which is likely due to that faults are not included in the seismic volumes in this dataset. Therefore, adding faults to the seismic volume and labeling them as the background would help reduce the network's

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

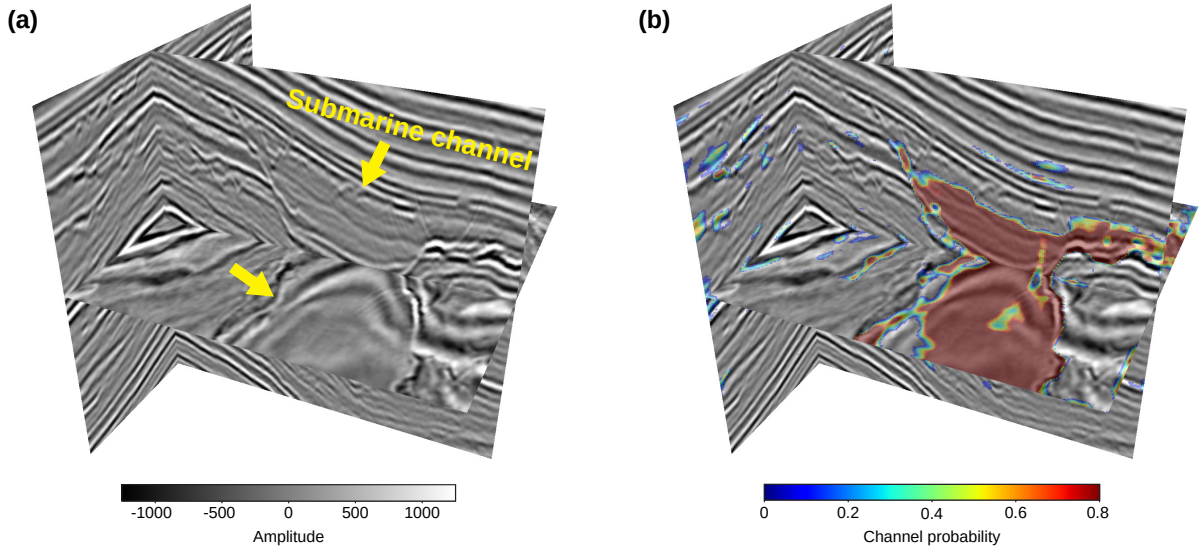


Figure 13. (a) Field seismic volume from the Parihaka seismic survey (courtesy of New Zealand Crown Minerals), showing a large-scale submarine channel (indicated by the yellow arrows). (b) The channel identification result of the U-Net trained by the *cigChannel* dataset.

tendency to mis-identification between faults and channels. It can also be seen that the channel identification result in Figure 11 is not as good as that in Figure 12, where the distributary channels in Figure 12 are mapped with uniformly high probability while some parts of the meandering channels in Figure 11 are mapped with moderate probability. It is probably because that the distributary channels in Figure 12 are filled with uniform seismic amplitude as we designed in our dataset, while the meandering channels in Figure 11 are filled with heterogenous seismic amplitude, which is an exceptional case for our dataset. Therefore, the identification performance of channels with heterogeneous seismic amplitude would be improved if meandering and distributary channels with heterogeneous seismic amplitude can be included in this dataset. As we mentioned, these are preliminary tests mainly to find out whether this dataset can help the network discriminate channels and non-channel areas. Future work could involve using this dataset to train a network to classify terrestrial and submarine channels, or to interpret the sedimentary facies of the submarine channels.

5 Conclusions

The *cigChannel* dataset is dedicated to overcome the shortage of training data for deep learning-based paleochannel identification in seismic volumes. It provides a more comprehensive collection of paleochannels than its predecessors. Workflow for generating this dataset was designed to produce synthetic seismic volumes with realistic characteristics of paleochannels, which exhibit large variability due to the randomization of many parameters that control the workflow. The effectiveness of this dataset is

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

demonstrated by its application on several field seismic volumes, which shows that even a simplified U-Net works well in identifying paleochannels after being trained with our dataset.

Other than providing training data for deep learning models to identify paleochannels in seismic volumes, this dataset can also serve as a publicly available benchmark dataset for validating the performance of various deep learning models and training strategies. This dataset can be further improved by incorporating new elements into the dataset generation workflow, such as adding faults to create a more complex structure and introducing heterogeneous seismic amplitude to the meandering and distributary channels. As the codes corresponding to the dataset generation workflow are also made publicly available, users can customize the controlling parameters and create datasets that used to identify specific forms of paleochannels.

6 Code and data availability

Codes corresponding to the dataset generation workflow are provided on GitHub (https://github.com/wanggy-1/cigChannel) and the *cigChannel* dataset (Wang et al., 2024) can be accessed via Zenodo (meandering channel subset: https://doi.org/10.5281/zenodo.11078794, distributary channel subset: https://doi.org/10.5281/zenodo.11073030, submarine channel subset: https://doi.org/10.5281/zenodo.11079950, and assorted channel subset: https://doi.org/10.5281/zenodo.11044512). Seismic data of the Parihaka seismic survey can be accessed via SEG wiki (https://wiki.seg.org/wiki/Parihaka-3D). Seismic data from the Tarim basin are confidential and cannot be released.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

Appendix A: Channel modeling parameters

Table A1. Channel modeling parameters with their reference values.

Channel type	Parameter	Value
Meandering channel	Width	200 m - 500 m
	Maximum depth	20 m - 50 m
	Strike	N0°E - N360°E
	Migration rate constant*	40 m/year - 50m/year
	Chezy's friction fractor*	0.06 - 0.08
	Iteration time step*	0.1 year
	Number of iterations*	1000 - 2000
Distributary channel	Maximum width	200m - 400 m
	Width/depth ratio	10 - 12
	Maximum number of iterations†	8192
	Number of Particles for early-termination†	0
Submarine channel	Width	300 m - 400 m
	Maximum depth	30 m - 40 m
	Strike	N0°E - N360°E
	Migration rate constant*	50m/year - 60m/year
	Chezy's friction factor*	0.07 - 0.08
	Iteration time step*	0.1 year
	Number of iterations*	500 - 2000
	Natural levee maximum thickness*	0.5m/iteration
	Natural levee width	6000 m - 8000 m
	Incision rate*	8 m/year
	Aggradation rate*	8 m/year

* Inputs of *meanderpy*

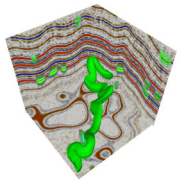
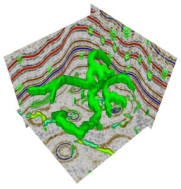
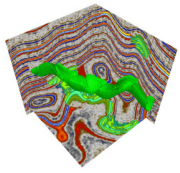
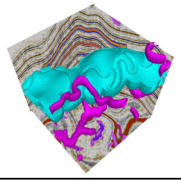
† Inputs of *soillib*

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

Appendix B: Components of the *cigChannel* dataset

Table B1. Components of the *cigChannel* dataset.

Name	Sample amount	Contents	Features	Example
Meandering channel subset	300	1. Seismic volumes 2. Binary label volumes 3. Seismic impedance volumes	1. Meandering channels only 2. Horizontal, inclined and folded structures 3. Noise-free	
Distributary channel subset	300	1. Seismic volumes 2. Binary label volumes 3. Seismic impedance volumes	1. Distributary channels only 2. Horizontal, inclined and folded structures 3. Noise-free	
Submarine channel subset	300	1. Seismic volumes 2. Binary label volumes 3. Seismic impedance volumes 4. Sedimentary facies volumes	1. Submarine channels only 2. Horizontal, inclined and folded structures 3. Noise-free	
Assorted channel subset	300	1. Seismic volumes 2. Multi-class label volumes 3. Seismic impedance volumes	1. Meandering, distributary and submarine channels 2. Horizontal, inclined and folded structures 3. Noise-free	

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Open Access Earth System Science Data Discussions

Appendix C: Illustrative codes of the dataset generation workflow

```
# Import all functions.
from functions import *

# Number of models.
n_model = 300
# Data generation.
for i in range(n_model):
    # Initialize the model.
    model = GeoModel()
    # Assign P-wave velocities.
    model.add_vp()
    # Add meandering channels.
    model.add_meandering_channel()
    # Add distributary channels.
    model.add_distributary_channel()
    # Add submarine channels.
    model.add_submarine_channel()
    # Add inclination.
    model.add_dipping()
    # Add folds.
    model.add_fold()
    # Resampling model's z-coordinates.
    model.resample_z()
    # Compute P-wave impedance.
    model.compute_Ip()
    # Compute reflection coefficients.
    model.compute_rc()
    # Make synthetic seismic data.
    model.make_synseis()
    # Save data.
    model.Ip.tofile()  # Impedance volume.
    model.seismic.tofile()  # Seismic volume.
    model.seis_label.tofile()  # Channel label volume.
    model.facies.tofile()  # Sedimentary facies volume.
```

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Appendix D: Parameters of the seismic impedance model and Ricker wavelet

Table D1. Parameters of the seismic impedance model, Ricker wavelet and their reference values.

	Parameter	Value
Model extension	X	0 m - 6400 m
	Y	0 m - 6400 m
	Z	0 m - 1280 m
	Grid spacing	25 m × 25 m × 5 m (X × Y × Z)
Layer	Seismic impedance	7000 m/s.g/cm^3 - 16000 m/s.g/cm^3
	Impedance perturbation	300 m/s.g/cm^3 - 500 m/s.g/cm^3
	Thickness	60 m - 150 m
Meandering channel	Impedance contrast with covering layer (ε)	0 - 1
Distributary channel	Impedance contrast with covering layer (ε)	0 - 1
Submarine channel	Point bar impedance	6000 m/s.g/cm^3 - 8400 m/s.g/cm^3
	Natural levee impedance	8400 m/s.g/cm^3 - 14400 m/s.g/cm^3
	Oxbow lake impedance	8400 m/s.g/cm^3 - 14400 m/s.g/cm^3
Ricker wavelet	Peak wavenumber	20 km^{-1} - 60 km^{-1}

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Author contributions. Guangyu Wang wrote the manuscript and the Python package of the dataset generation workflow. Xinming Wu were involved in conceptualisation and manuscript preparation. Wen Zhang conducted the experiments on field application of this dataset and co-wrote the Application section.

Competing interests. The authors declare no competing interests.

Acknowledgements. The authors thank the USTC supercomputing center for providing computational resources for this project and Jintao Li for providing the Python package *CIGVis* to visualize the 3D seismic images. The authors also thank Hang Gao and Jiarun Yang for their useful suggestions on the training strategy of the U-Net.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

References

Bi, Z., Wu, X., Geng, Z., and Li, H.: Deep relative geologic time: A deep learning method for simultaneously interpreting 3-D seismic horizons and faults, Journal of Geophysical Research: Solid Earth, 126, e2021JB021 882, https://doi.org/https://doi.org/10.1029/2021JB021882, 2021.

Bridge, J. S., Jalfin, G. A., and Georgieff, S. M.: Geometry, lithofacies, and spatial distribution of Cretaceous fluvial sandstone bodies, San Jorge Basin, Argentina: outcrop analog for the hydrocarbon-bearing Chubut Group, Journal of Sedimentary Research, 70, 341–359, https://doi.org/https://doi.org/10.1306/2DC40915-0E47-11D7-8643000102C1865D, 2000.

Clark, J. D. and Pickering, K. T.: Architectural elements and growth patterns of submarine channels: application to hydrocarbon exploration, AAPG bulletin, 80, 194–220, https://doi.org/https://doi.org/10.1306/64ED878C-1724-11D7-8645000102C1865D, 1996.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L.: Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, pp. 248–255, Ieee, https://doi.org/10.1109/CVPR.2009.5206848, 2009.

Deptuck, M. E., Sylvester, Z., Pirmez, C., and O'Byrne, C.: Migration–aggradation history and 3-D seismic geomorphology of submarine channels in the Pleistocene Benin-major Canyon, western Niger Delta slope, Marine and Petroleum Geology, 24, 406–433, https://doi.org/https://doi.org/10.1016/j.marpetgeo.2007.01.005, 2007.

Edmonds, D. A. and Slingerland, R. L.: Mechanics of river mouth bar formation: Implications for the morphodynamics of delta distributary networks, Journal of Geophysical Research: Earth Surface, 112, https://doi.org/https://doi.org/10.1029/2006JF000574, 2007.

Gao, H., Wu, X., and Liu, G.: ChannelSeg3D: Channel simulation and deep learning for channel interpretation in 3D seismic images, Geophysics, 86, IM73–IM83, https://doi.org/https://doi.org/10.1190/geo2020-0572.1, 2021.

Gee, M., Gawthorpe, R. L., Bakke, K., and Friedmann, S.: Seismic geomorphology and evolution of submarine channels from the Angolan continental margin, Journal of Sedimentary Research, 77, 433–446, https://doi.org/https://doi.org/10.2110/jsr.2007.042, 2007.

Geleynse, N., Storms, J. E., Walstra, D.-J. R., Jagers, H. A., Wang, Z. B., and Stive, M. J.: Controls on river delta formation; insights from numerical modelling, Earth and Planetary Science Letters, 302, 217–226, https://doi.org/https://doi.org/10.1016/j.epsl.2010.12.013, 2011.

Hein, F. J. and Cotterill, D. K.: The Athabasca oil sands—a regional geological perspective, Fort McMurray area, Alberta, Canada, Natural Resources Research, 15, 85–102, https://doi.org/https://doi.org/10.1007/s11053-006-9015-4, 2006.

Howard, A. D. and Knutson, T. R.: Sufficient conditions for river meandering: A simulation approach, Water Resources Research, 20, 1659–1667, https://doi.org/https://doi.org/10.1029/WR020i011p01659, 1984.

Ji, S., Xu, W., Yang, M., and Yu, K.: 3D convolutional neural networks for human action recognition, IEEE transactions on pattern analysis and machine intelligence, 35, 221–231, https://doi.org/10.1109/TPAMI.2012.59, 2012.

Kolla, V., Posamentier, H., and Wood, L.: Deep-water and fluvial sinuous channels—Characteristics, similarities and dissimilarities, and modes of formation, Marine and Petroleum Geology, 24, 388–405, https://doi.org/https://doi.org/10.1016/j.marpetgeo.2007.01.007, 2007.

Leigh, D. S. and Feeney, T. P.: Paleochannels indicating wet climate and lack of response to lower sea level, southeast Georgia, Geology, 23, 687–690, https://doi.org/https://doi.org/10.1130/0091-7613(1995)023<0687:PIWCAL>2.3.CO;2, 1995.

Li, X., Chen, Q., Wu, C., Liu, H., and Fang, Y.: Application of multi-seismic attributes analysis in the study of distributary channels, Marine and Petroleum Geology, 75, 192–202, https://doi.org/https://doi.org/10.1016/j.marpetgeo.2016.04.016, 2016.

Liang, M., Voller, V., and Paola, C.: A reduced-complexity model for river delta formation–Part 1: Modeling deltas with channel dynamics, Earth Surface Dynamics, 3, 67–86, https://doi.org/https://doi.org/10.5194/esurf-3-67-2015, 2015.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L.: Microsoft coco: Common objects in context, in: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13, pp. 740–755, Springer, https://doi.org/https://doi.org/10.1007/978-3-319-10602-1_48, 2014.

McDonald, N.: soillib, https://github.com/erosiv/soillib, 2020.

Nettleton, D. F., Orriols-Puig, A., and Fornells, A.: A study of the effect of different types of noise on the precision of supervised learning techniques, Artificial intelligence review, 33, 275–306, https://doi.org/https://doi.org/10.1007/s10462-010-9156-z, 2010.

Nordfjord, S., Goff, J. A., Austin Jr, J. A., and Sommerfield, C. K.: Seismic geomorphology of buried channel systems on the New Jersey outer shelf: assessing past environmental conditions, Marine Geology, 214, 339–364, https://doi.org/https://doi.org/10.1016/j.margeo.2004.10.035, 2005.

Payenberg, T. H. D. and Lang, S. C.: Reservoir geometry of fluvial distributary channels—Implications for Northwest Shelf, Australia, deltaic successions, The APPEA Journal, 43, 325–338, https://doi.org/https://doi.org/10.1071/AJ02017, 2003.

Pechenizkiy, M., Tsymbal, A., Puuronen, S., and Pechenizkiy, O.: Class noise and supervised learning in medical domains: The effect of feature extraction, in: 19th IEEE symposium on computer-based medical systems (CBMS'06), pp. 708–713, IEEE, https://doi.org/10.1109/CBMS.2006.65, 2006.

Pham, N., Fomel, S., and Dunlap, D.: Automatic channel detection using deep learning, Interpretation, 7, SE43–SE50, https://doi.org/https://doi.org/10.1190/INT-2018-0202.1, 2019.

Ronneberger, O., Fischer, P., and Brox, T.: U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234–241, Springer, https://doi.org/https://doi.org/10.1007/978-3-319-24574-4_28, 2015.

Seybold, H., Andrade Jr, J. S., and Herrmann, H. J.: Modeling river delta formation, Proceedings of the National Academy of Sciences, 104, 16 804–16 809, https://doi.org/https://doi.org/10.1073/pnas.0705265104, 2007.

Sylvester, Z.: meanderpy, https://github.com/zsylvester/meanderpy, 2021.

Sylvester, Z., Pirmez, C., and Cantelli, A.: A model of submarine channel-levee evolution based on channel trajectories: Implications for stratigraphic architecture, Marine and Petroleum Geology, 28, 716–727, https://doi.org/https://doi.org/10.1016/j.marpetgeo.2010.05.012, 2011.

Sylvester, Z., Durkin, P., and Covault, J. A.: High curvatures drive river meandering, Geology, 47, 263–266, https://doi.org/https://doi.org/10.1130/G45608.1, 2019.

Sylvia, D. A. and Galloway, W. E.: Morphology and stratigraphy of the late Quaternary lower Brazos valley: Implications for paleo-climate, discharge and sediment delivery, Sedimentary Geology, 190, 159–175, https://doi.org/https://doi.org/10.1016/j.sedgeo.2006.05.023, 2006.

Vizeu, F., Zambrini, J., Tertois, A.-L., da Graça e Costa, B. d. A., Fernandes, A. Q., and Canning, A.: Synthetic seismic data generation for automated AI-based procedures with an example application to high-resolution interpretation, The Leading Edge, 41, 392–399, https://doi.org/https://doi.org/10.1190/tle41060392.1, 2022.

Wang, G., Wu, X., and Zhang, W.: cigChannel: A massive-scale 3D seismic dataset with labeled paleochannels for advancing deep learning in seismic interpretation, https://doi.org/10.5281/zenodo.10791151, 2024.

Wu, X., Liang, L., Shi, Y., and Fomel, S.: FaultSeg3D: Using synthetic data sets to train an end-to-end convolutional neural network for 3D seismic fault segmentation, Geophysics, 84, IM35–IM45, https://doi.org/https://doi.org/10.1190/geo2018-0646.1, 2019.

Wu, X., Geng, Z., Shi, Y., Pham, N., Fomel, S., and Caumon, G.: Building realistic structure models to train convolutional neural networks for seismic structural interpretation, Geophysics, 85, WA27–WA39, https://doi.org/https://doi.org/10.1190/geo2019-0375.1, 2020a.

https://doi.org/10.5194/essd-2024-131
Preprint. Discussion started: 25 July 2024

Wu, X., Yan, S., Qi, J., and Zeng, H.: Deep learning for characterizing paleokarst collapse features in 3-D seismic images, Journal of Geophysical Research: Solid Earth, 125, e2020JB019 685, https://doi.org/https://doi.org/10.1029/2020JB019685, 2020b.

Zhang, Z., Li, H., Yan, Z., Jing, J., and Gu, H.: Deep carbonate fault–karst reservoir characterization by multi-task learning, Geophysical Prospecting, 72, 1092–1106, https://doi.org/https://doi.org/10.1111/1365-2478.13460, 2024.

Zheng, Y., Zhang, Q., Yusifov, A., and Shi, Y.: Applications of supervised deep learning for seismic interpretation and inversion, The Leading Edge, 38, 526–533, https://doi.org/https://doi.org/10.1190/tle38070526.1, 2019.