

Abundance of trace fossil *Phycosiphon incertum* in core sections measured using a convolutional neural network

Kazuki Kikuchi^{a,*}, Hajime Naruse^b

^a*Department of Civil and Environmental Engineering, Faculty of Science and Engineering, Chuo University, 1-13-27, Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan*

^b*Department of Geology and Mineralogy, Graduate School of Science, Kyoto University, Kitashirakawaoiwake-cho, Sakyo-ku, Kyoto, 606-8502, Japan*

*Corresponding author

Email address: kikuchi@kueps.kyoto-u.ac.jp (Kazuki Kikuchi)

Abstract

A convolutional neural network (CNN) was used to construct a semantic segmentation model to examine the abundance of *Phycosiphon incertum* by identifying the trace fossil regions in core section images. The abundance of trace fossils provides information about the past activities of benthic animals affected by paleoenvironmental conditions. To quantify the intensity of bioturbation, it is necessary to extract regions of trace fossils and measure the proportion of bioturbated and the observed area of the outcrop section. In this study, a U-Net-type CNN model was used with residual connections and attention mechanisms to identify the trace fossil *Phycosiphon*. The model was trained to recognize the relationships between core section images from the International Ocean Discovery Program Expedition 362 Site U1480 and manually annotated trace fossil images. After training, the model successfully classified the pixels of the

background, outcrop, and *Phycosiphon* for core section images other than the training data set. The bioturbation intensity estimated from the image predicted by the model was nearly equal to that from the ground truth image. A long-term (approximately past 10 Myr) variation in *Phycosiphon* abundance was estimated by applying the model to the core section images at Site U1480. *Phycosiphon* abundance negatively correlated with the number of sandstone layer intercalations, but it was not affected by the sediment accumulation rates. These findings may reflect resistance of *Phycosiphon* producers to environmental stress. The model developed in this study can be used for other ichnotaxa to reveal the general tendency of variation in bioturbation intensity and ichnodiversity.

Keywords:

ichnology, submarine-fan deposits, IODP Exp.362, U-Net, semantic segmentation, deep learning

1. Introduction

Trace fossils and bioturbation intensity can be used to determine the environmental conditions on ancient seafloors. Bioturbation intensity can provide information on past activities of benthic animals depending on paleoenvironmental conditions such as sedimentation rates, frequency of physical disturbance, or organic matter inputs (Wetzel and Uchman, 1998; Cummings and

Hodgson, 2011; Tonkin, 2012; Uchman and Wetzel, 2012).

Bioturbation intensity has been utilized to quantify bioturbation structures, which are generally defined as the ratio of the bioturbated area to the observed area of the outcrop section (Droser and Bottjer, 1986; Taylor and Goldring, 1993; Miller and Smail, 1997; Rebata-H. et al., 2006). Several studies have investigated changes in bioturbation intensity (Cummings and Hodgson, 2011; Heard et al., 2014). Bioturbation intensity may vary depending on the ichnotaxa because the dominant ichnotaxa in ichnoassemblages can change because of paleoenvironmental changes (Buatois and Mángano, 2013). Thus, the difference in bioturbation intensity among each ichnotaxon may indicate paleoenvironmental conditions.

However, manually quantifying the bioturbation intensity for specific ichnotaxon in outcrop or core section images is tedious. Therefore, bioturbation intensity is generally estimated without distinguishing any particular ichnotaxa based on picture-matching with catalogs showing grades (indices) of bioturbation (Droser and Bottjer, 1986; Taylor and Goldring, 1993). Although the picture-matching method is effective for field workers who need to estimate the bioturbation intensity with the naked eye, it is impossible to quantify it for a specific ichnotaxon. Although there are examples of ichnospecies-specific geological records of changes in occurrence frequency, such as common or rare, the frequency is expressed qualitatively, and few studies have quantitatively recorded the abundance of ichnospecies (Heard and Pickering, 2008; Knaust et al.,

2014).

Image analysis for the automatic extraction of trace fossils may provide a more objective, quantitative, and fast methodology for paleoenvironmental research based on the bioturbation intensity of specific ichnotaxon. The area ratio of each ichnotaxon to the outcrop area can be calculated directly from the image analysis results. Dorador and Rodríguez-Tovar (2014) attempted to extract the region of trace fossils in core section images using a combination of multiple pixel selection methods. However, their methods strongly depend on the appearance of the core section images, and manual operations for ichnotaxa classification are required after extraction.

In this study, we proposed a new automated method for detecting the region of a specific ichnotaxa using a convolutional neural network (CNN). The automatic extraction method for trace fossils was implemented using a CNN model, which is a network architecture that comprises multiple convolutional operations commonly used for filtering in image analysis (Perona and Malik, 1990). Deep learning techniques were employed to optimize the parameters in the convolutional operations, enabling the CNN model to automatically extract features from images. Using this technique, the trained CNN model can predict the bioturbation intensity by learning the relationships between core section images and bioturbation information.

Based on this concept, CNN models have been used to analyze sedimentary geology and trace fossils (Liu and Song, 2020; Timmer et al., 2021; Ayranci et al.,

2021). Liu and Song (2020) constructed a CNN model for fossil recognition from carbonate microfacies images to reconstruct sedimentary environments and paleoecology. Timmer et al. (2021) built a CNN model to classify core section images trimmed into 128×128 -sized tiles to find tiles with bioturbation structures. Their model succeeded in classification with 88% accuracy; however, image segmentation of the bioturbation structures is required to estimate the bioturbation intensity after classification. Ayranci et al. (2021) built a CNN based on VGG-16 (Simonyan and Zisserman, 2014) and attempted to estimate bioturbation index (BI; Taylor and Goldring, 1993) directly from the core section images. They constructed two models, a binary (BI = 0 or not) and three-class classification models (BI = 0, 1–2, or 3–6). Their models achieved over 84.9% in accuracy for the classification. However, in this study, manual extraction of trace fossils was required.

The semantic segmentation model, which is generally implemented as an encoder–decoder type CNN, can solve the problem of extracting and measuring bioturbated areas (Ronneberger et al., 2015). In contrast to previous models, encoder–decoder networks can classify images per pixel. Thus, the model can predict the presence, location, size, and distribution of trace fossils in core section images. Furthermore, the proportion of bioturbated to outcrop areas can be easily calculated from the model output images.

In this study, an encoder–decoder type CNN for automatic extraction of the trace

fossil *Phycosiphon incertum*, which is the most common ichnotaxon in deep sea successions, was constructed (Uchman and Wetzel, 2012). *Phycosiphon* is a small spreite burrow composed of a winding mud tube and surrounding sandy halo (Naruse and Nifuku, 2008; Bednarz and McIlroy, 2009). The semantic segmentation model is suitable for estimating the bioturbation intensity of *Phycosiphon*. *Phycosiphon* abundantly occurs in the vertical section of the core, making the manual extraction of *Phycosiphon* a labor-intensive process. Generally, the concentrated occurrence of *Phycosiphon* could indicate a high frequency of physical disturbances.

Phycosiphon is considered a trace fossil produced by opportunistic benthic animals following physical disturbances (Wetzel and Uchman, 2001). MacEachern and Bann (2020) have reported the ichnofacies characterized by *Phycosiphon*, reflecting the recurring alternations between stable marine conditions and physico-chemically stressed conditions. After the stressed conditions, *Phycosiphon*-forming organisms tend to colonize before other trace fossil formers, resulting in their trace fossils in undisturbed sediments. Other organisms later colonize the area, re-disturbing sediment and overprinting *Phycosiphon* with the trace fossils of subsequent colonizers during inter-event periods. Preserving a high frequency of *Phycosiphon* requires short intervals between physical disturbances preventing the subsequent colonizers from obliterating *Phycosiphon*, and the difference in the response between *Phycosiphon* and other ichnotaxa could reflect the environmental instabilities. Therefore, in this

study, we developed and applied a model to analyze core section images from the International Ocean Discovery Program (IODP) Expedition 362 Site U1480 to estimate the long-term variation in bioturbation intensity associated with the development of a submarine-fan system.

2. Geological setting

Figure 1A shows IODP Expedition 362 drilling offshore at two sites on the Indian Oceanic plate and core samples collected from the core sections of the Nicobar Fan deposits. Site U1480 was approximately 1430 m thick and mainly comprised alternating sandstone and mudstone beds, as well as chalk, calcareous claystone, and basalt (McNeill et al., 2017; Pickering et al., 2020). This site was divided into six lithologic units, labeled Units I to VI from top to bottom (Fig. 1B). The basal part of the succession (Unit VI) comprised Upper Cretaceous basalt, overlain by Paleocene chalk and calcareous claystone (Unit V). Unit IV consists of volcanoclastic coarse-grained deposits that conformably overlay Unit V. Unit III comprised siliciclastic mudstones and tuffaceous silty claystone unconformably, overlying Unit IV. Unit III is characterized by pelagic deposition, indicating prefan deposits (Pickering et al., 2020). Unit II consists of alternating sandstone and mudstone beds, which comprise the main part of the Nicobar fan deposits (Fig. 1B). Unit II is subdivided into three subunits, namely IIA, IIB, and IIC, based on their lithology. Generally, the sandstone beds in Unit II exhibit an

upward-coarsening trend from Units IIC to IIA (Fig. 1B), indicating a progradation of the submarine-fan system (Pickering et al., 2020). Based on the core observation and seismic analysis, Units IIC, IIB, and IIA were interpreted as lobe complex to fan-fringe deposits, lobe complex to channel-levee deposits, and channel-levee deposits, respectively (Pickering et al., 2020). The biostratigraphy of microfossils indicates that Unit II was deposited during the Miocene to Pleistocene (McNeill et al., 2017; Pickering et al., 2020). The uppermost part of Unit I comprises alternating beds of sand, calcareous mud, and intercalated volcanic ash (Fig. 1B).

The depositional age based on biostratigraphy and radiocarbon age indicates that the accumulation rates of Site U1480 fluctuated significantly during the deposition of the Nicobar Fan (the late Miocene to Pleistocene; Pickering et al., 2020). Based on the variation in sediment accumulation rate and depositional facies, Pickering et al. (2020) concluded that the deposition of the Nicobar Fan initiated at approximately 10 Ma at the IODP Exp. 362 sites and ceased at approximately 1.5 Ma due to the depocenter shifting to the western part of Bengal Bay because of the Ninetyeast Ridge closing off the connection to the north.

3. Methodology

In this study, a CNN model was constructed to perform semantic segmentation on regions of trace fossils in core section images. Core section images were taken

from the IODP Exp. 362 Site U1480 Hole F for training and validation of the CNN model. After training, the model was applied to core section images from Site U1480 Holes E, F, and G to classify each pixel into the following three classes: background, outcrop, and *Phycosiphon incertum*. The ratio of *P. incertum* regions to outcrop was calculated to estimate the bioturbation intensity in the core sections throughout the core sites. The procedures are described in detail below.

3.1. Semantic segmentation model

3.1.1. Training and validation datasets

Five core section images from the IODP Exp. 362 Site U1480 Hole F (cores 69X-1, 69X-2, 69X-3, 76X-1, and 79X-1) were selected as training, validation, and test datasets. They were painted in three colors to represent the background (black), outcrop (white), and trace fossil *Phycosiphon* (green) regions. These regions were manually painted with the pencil tool of Adobe Photoshop. The resolution of each core section image was 200 pixels per cm. The appearance of *Phycosiphon* in the core section of Site U1480 consisted of dark-colored, millimetric string-like tubes and surrounding light-colored halos (Fig. 6A). Both the dark-colored and light-colored parts were manually traced. Distinct colors painted on other ichnotaxa were recognized in the core images for each ichnotaxon to obtain area percentage. The core section and painted images were split into tiles of 224×224 pixels (Fig. 2). The tiles of images from the cores 69X-1, 69X-2, 76X-1, and 79X-1 were randomly divided into two sets, 90% and 10%

of the tiles were used for training and validation, respectively. The Core 69X-3 was used for the test of the model. Consequently, the number of training, validation, and test datasets was 1151, 127, and 726 pairs of tiles of core section and labeled images, respectively.

After splitting, all images were converted from BGR (blue, green, and red) to HSV (hue, saturation, and value) color system, and then histogram equalization was applied to the brightness channel (V) of each tile (Acharya and Ray, 2005). After histogram equalization, all images were converted from HSV to BGR color system, and the pixel values in each channel were normalized to adjust the range from -1 to 1 . Each tile of the labeled images were 224×224 in resolution.

3.1.2. *Architecture of CNN*

This study utilized a CNN model that had a U-Net like architecture with residual connections (Zhang et al., 2018; Shang et al., 2021) and attention mechanisms (Oktay et al., 2018) for the semantic segmentation model (Figs. 3 and 4). This model receives images and predicts the class to which each pixel belongs. In other words, the model estimates probabilities in the pixels, whether the pixel shows background, outcrop, or *Phycosiphon*. The prediction is conducted based on image feature values extracted by convolution operations. Consequently, the predicted images painted by up to 3 colors can be obtained by the probability that the model outputs. See the Appendix for the details of the CNN architecture in this study.

Training of the model was conducted to solve an optimization problem to minimize the loss function, which indicated the difference between the predicted and ground truth images of the training dataset. The loss function was differentiated with respect to each weight coefficient in filters of the convolution operations, and then the weight coefficients were changed to lower the loss function. During the training, the training dataset was randomly subdivided into subsets called mini-batches, and the optimization mentioned above was repeated until all mini-batches were used for the optimization. The process to use all data in the training dataset is called epoch. This study set the mini-batch size to 16, and the training was conducted for 150 epochs. The conditions for the training were summarized in Table 1.

3.1.3. Loss function and metrics

This study used the Tanimoto loss with its complement based on the Tanimoto coefficient (Tanimoto loss; Diakogiannis et al., 2020) as a loss function to optimize the CNN. The Tanimoto coefficient showed a similarity between the two sets of vectors. The Tanimoto coefficient T is defined as follows:

$$T(p_i, l_i) = \frac{\sum_{i=1}^N p_i l_i}{\sum_{i=1}^N (p_i^2 + l_i^2 - p_i l_i)} \quad (1)$$

where p_i and l_i are a vector of probabilities and the ground truth for the i -th pixel, respectively, and N denotes the number of pixels. To adjust the class-imbalance problem in the multiple classifications, the Tanimoto coefficient can be modified

as (Diakogiannis et al., 2020):

$$T(p_i^{(c)}, l_i^{(c)}) = \frac{\sum_{c=1}^C w_c \sum_{i=1}^N p_i^{(c)} l_i^{(c)}}{\sum_{c=1}^C w_c \sum_{i=1}^N (p_i^{(c)^2} + l_i^{(c)^2} - p_i^{(c)} l_i^{(c)})}, \quad (2)$$

where C is the number of classes; $p_i^{(c)}$ and $l_i^{(c)}$ represent a vector of probabilities and ground truth of the class c for the i -th pixel, respectively. The parameter w_c denotes the weighting coefficient for the class c , which is defined as

$$w_c = \left(\frac{1}{\sum_{i=1}^N l_i^{(c)}} \right)^2. \quad (3)$$

Diakogiannis et al. (2020) used the Tanimoto coefficient as a region-based loss function for the semantic segmentation model to maximize the overlap between the predicted and ground truth images. This study attempted to maximize the overlap of true positives and negatives between prediction and ground truth. Thus, the Tanimoto coefficient with its complements ($\tilde{T}$) is defined as follows:

$$\tilde{T}(p_i^{(c)}, l_i^{(c)}) = \frac{T(p_i^{(c)}, l_i^{(c)}) + T(\mathbf{1} - p_i^{(c)}, \mathbf{1} - l_i^{(c)})}{2} \quad (4)$$

where

$$\begin{aligned} \mathbf{1} - p_i^{(c)} &= \{1 - p_1^{(c)}, 1 - p_2^{(c)}, \dots, 1 - p_N^{(c)}\}, \\ \mathbf{1} - l_i^{(c)} &= \{1 - l_1^{(c)}, 1 - l_2^{(c)}, \dots, 1 - l_N^{(c)}\}. \end{aligned} \quad (5)$$

To use the Tanimoto coefficients as a loss function for the minimization problem, the Tanimoto loss with its complement $\mathcal{L}_T$ is defined as (Diakogiannis et al., 2020):

$$\mathcal{L}_T = 1 - \tilde{T}(p_i^{(c)}, l_i^{(c)}). \quad (6)$$

The Dice coefficient D was used as a metric to evaluate the model performance in this study, which takes the following form:

$$D = \frac{2 \sum_{i=1}^N p_i l_i}{\sum_{i=1}^N (p_i + l_i)}. \quad (7)$$

3.1.4. Optimization

This study used the rectified Adam (RAdam; Liu et al., 2019) as an optimization method of the weight coefficients for the CNN (Table 1). The hyperparameters of RAdam were set as follows: $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 10^{-7}$. During the training, the learning rate α_t at the t -th epoch was scheduled to change with epochs:

$$\alpha_t = \begin{cases} \frac{\alpha_{\max}}{E_w} t & (t \leq E_w) \\ \frac{\alpha_{\max}}{2} \left(\cos \frac{t - E_w}{E_{\max}} \pi + 1 \right) & (\text{otherwise}) \end{cases} \quad (8)$$

where $\alpha_{\max}$ is the maximum value of the learning rate, E_w represents the length of epochs of the warm-up period, and $E_{\max}$ denotes the maximum length of the epochs. In this study, the parameters were set as follows: $\alpha_{\max} = 0.001$, $E_w = 10$, and $E_{\max} = 150$. As a result, the learning rate α_t increased linearly between the 1st and 10th epoch, and then it decayed according to the cosine curve after the 11th epoch (Fig. 5). The gradual increase in the learning rate in the early epochs of the training (the warm-up period) can prevent the explosive increase in the

adaptive learning rate term in the RAdam algorithm (Liu et al., 2019).

3.2. Estimation of the abundance of *Phycosiphon*

The bioturbation intensity of *Phycosiphon* was evaluated using a ratio of the area of bioturbation structures to that of the core section image. Thus, the bioturbation intensity B_i was defined as follows:

$$B_i = \frac{N_{\text{bio}}}{N_{\text{core}}} \quad (9)$$

Where N_{bio} and N_{core} represent the number of pixels of the bioturbation structures and core section images in the region of interest, respectively.

To evaluate the performance of the image analysis model for estimating bioturbation intensity B_i , the image of Hole F Core 69X Section 3 was subdivided into 14 intervals 10 cm thick from the top to the bottom of the section. B_i was calculated for each interval of the ground truth and predicted images. The root mean squared error (RMSE) of B_i was calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (B_{i,k} - \hat{B}_{i,k})^2} \quad (10)$$

where n denotes the number of the intervals ($n = 14$ in this study). $B_{i,k}$ and $\hat{B}_{i,k}$ represent the bioturbation intensity of the k -th interval in the ground truth and predicted images, respectively. An RMSE close to 0 indicates a high estimation

ability of the model for bioturbation intensity.

3.3. *Analysis of stratigraphic variation in the bioturbation intensity*

The bioturbation intensity indicating the abundance of *Phycosiphon* was estimated on the basis of the predicted images of Site U1480 Holes E, F, and G. The number of the core section images in these Holes E, F, and G were 84, 264, and 338, respectively. This study focused on the bioturbation intensity of *Phycosiphon*, which preferentially occurs in mudstones. Therefore, the regions of the coarse-grained sandstone layers in each predicted image were masked after the prediction to avoid an obvious decrease in the occurrence of trace fossils. The bioturbation intensity B_i was calculated using Equation 9 for each 1.5 m long section. Then, the logarithm of B_i was calculated.

The stratigraphic variation in B_i was compared with the sediment accumulation rates and frequency of turbidite deposition to examine the relationship between bioturbation and paleoenvironmental conditions. The linear sediment accumulation rates (LSAR) of the deposits were calculated using the age-control points and bed thickness, as described in the previous section. The mudstone LSAR (mLSAR), which is the accumulation rate of the mudstones, was estimated by removing the thickness of the coarse deposits from the total deposit thickness. In addition, the number of sandstone layer intercalations (NSLI) was calculated for each section based on the visual core descriptions of the IODP Exp. 362 Site U1480. The moving averages of B_i , LSAR, mLSAR, and NSLI were then

obtained using 150-m-thick sampling windows. The stratigraphic resolution of LSAR ranged in thickness from 100 to 300 m (Pickering et al., 2020); thus, the sampling window of the moving average was set to 150 m thick to obtain adequate smoothness and variations in the above-mentioned parameters.

4. Results

4.1. Trace fossils observed in Unit II of IODP Site U1480

In this study, we used the core section images of Unit II of IODP Site U1480 as the training and validation datasets. Six ichnogenera were observed in the mudstone beds of Unit II (Fig. 6), whereas the sandstone beds rarely yielded trace fossils. *Phycosiphon incertum* was the dominant ichnotaxa in this unit, with area percentages ranging from 73.0% to 96.8% in the training and validation datasets. *Nereites* isp. and *Palaeophycus* isp. were relatively common, whereas *Planolites* isp. and *Scolicia* isp. were rare. The areas of regions for these ichnotaxa ranged from 0.0% to 16.8% (Table 2). Thus, *Phycosiphon* mainly contributed to the bioturbation at the study site, and the area percentage of *Phycosiphon* was considered a representative of the bioturbation intensity.

4.2. Evaluation of the CNN model performance

The training history of the CNN model revealed that the proposed model was

successfully trained without any significant overlearning for segmenting the trace fossil regions in the core images (Fig. 7). The value of the loss function for the training dataset decreased to 0.262 at the last epoch (Fig. 7A), and the validation loss also decreased to 0.610 associated with the epochs. Although the Dice coefficient for the validation dataset significantly fluctuated in the earlier epochs, those for the training and validation datasets increased and followed similar tracks, particularly after the 100th epoch (Fig. 7B). The decline of both training and validation loss associated with epochs suggested that the model did not overfit to the training dataset and lose generality. The Dice coefficient for the test dataset was 0.892, indicating a satisfactory match between the predicted and ground truth images.

The model's ability to predict the distribution of trace fossils was tested by applying it to the test dataset (the core section images of the Hole F Core 69X-3) (Fig. 8). The regions of the background, outcrops, and *Phycosiphon* were well reconstructed compared with the ground truth image. A visual comparison of the predicted image with the ground truth image demonstrated that both images were entirely consistent. The RMSE of B_i calculated from the test dataset was 0.0039 (Fig. 9). The precision of the estimation was nearly constant over the range of B_i tested, and no significant bias in the prediction was detected.

Although the prediction of the cores composed of the mudstones exhibited high accuracy in the prediction of the distribution of *Phycosiphon* (Fig. 10A), as

described above, the segmentation of the trace fossil regions failed in the case of lithology that was not included in the training dataset, such as the chalk, or calcareous or tuffaceous claystones (Fig. 10B). This suggests that the model in this study cannot predict out of range of the training dataset containing only the core images mainly composed of siliciclastic mudstone and sandstone beds. Thus, we analyzed the 546 core section images from Units I and II, mainly composed of mudstones and sandstones. The 140 core sections in Units III, IV, V, and VI were excluded from the analysis because they contained other lithologies.

4.3. Variation of the bioturbation intensity B_i

The bioturbation intensity B_i of *Phycosiphon* ranged from 3.16×10^{-5} to 6.05×10^{-2} (from -10.33 to -2.81 in natural logarithmic scale) in Units I and II (Figs. 11 and 12). The moving average of $\log B_i$ ranged from -6.61 to -4.88 (Fig. 12). B_i slightly increased in the lowermost part of Unit IIC, which was associated with the initiation of the Nicobar Fan deposition.

B_i exhibited dependency on the facies successions. It tended to increase in the mudstone-dominated intervals, so it took higher values throughout Unit IIC. B_i also exhibited higher values in Unit IIA, around 100 mbsf, where the fraction of mudstone layers increased (Fig. 12). However, B_i were relatively low at approximately 600–700 mbsf (Fig. 12), where thin sandstone intercalation was dominated (Pickering et al., 2020).

4.4. Relationships between bioturbation intensity and conditions of sediment accumulation

The LSAR, defined as the deposit thickness between age-control points divided by the depositional period, was high in the lower and upper units of Site U1480. The sediment accumulation rate was relatively low in the middle part of the succession. It was more than 200 m Myr⁻¹ during the deposition of the Unit IIC. In Unit IIB, the LSAR decreased to 65–162 m Myr⁻¹. In the middle part of Unit IIA, the LSAR increased to 362 m Myr⁻¹. LSAR decreased to 40 m Myr⁻¹ in the uppermost part of Unit IIA and declined to 3–17 m Myr⁻¹ around the boundary between Units IIA and I. It increased again to 84 m Myr⁻¹ above 18.17 mbsf in Unit I.

The LSAR of mudstone layers (mLSAR) at Site U1480 ranged from 3 to 194 m Myr⁻¹ (Fig. 12). Generally, the mLSAR tended to decrease upward. The mLSAR in Unit I varied from 3 to 62 m Myr⁻¹. The differences between the mLSAR and the total LSAR were small in Units IIB, IIC and Unit I; however, the total LSAR in Unit IIA was up to 362 m Myr⁻¹ while the mLSAR was only 104 m Myr⁻¹ in maximum due to the deposition of thick-bedded sand layers. The moving average of the number of sandstone layer intercalations (NSLI) ranged from 0.53 to 3.30 layers per meter, with a slight upward-increasing trend (Fig. 12). An peak of NSLI was identified between 600 and 700 mbsf, where core recovery was low because of the high frequency of sandstone deposition (McNeill et al.,

2017).

The relationship between the bioturbation intensity B_I and the previously described parameters LSAR, mLSAR and NSLI was evaluated (Fig. 13). The moving average of $\log B_I$ showed no significant correlations with sediment accumulation rates (LSAR and mLSAR).

In contrast, a negative correlation was observed between the moving average of $\log B_I$ and NSLI (Fig. 13C). The change in the moving average of $\log B_I$ appears to be small because the range of variation of the measured values at each horizon is considerably larger than the range of variation of the moving average; however, Figure 13 in which the moving average values were compared with NSLI clearly indicated that these parameters are negatively correlated. The correlation coefficient Pearson's r was -0.599 with $p \ll 0.01$ ($p = 3.879 \times 10^{-108}$).

5. Discussion

5.1. Performance of the segmentation model proposed in this study

The semantic segmentation model proposed in this study successfully identified the distribution and shapes of *Phycosiphon* in core section images. The proposed model demonstrated a sufficient ability to segment images compared with the segmentation models for other purposes, such as those commonly used in the

medical field. For example, Oktay et al. (2018) demonstrated a U-Net with attention gates to extract the region of the internal organs in 3D abdominal CT scan slice images. Their model segmented the pancreas in the images, achieving 0.840 in the Dice coefficient. Lei et al. (2021) demonstrated a U-Net with the attention gates to classify macerals of coal. Their model achieved an accuracy of 91.56%. Hou et al. (2021) constructed a classic U-Net to classify the CT data of vertebrate microfossils, which achieved 94.39% in IOU (the intersection over union). Yu et al. (2022) built a semantic segmentation model to classify the CT images of dinosaur fossils. They used a Deeplab v3+ (Chen et al., 2016) as the encoder network and a U-Net-like decoder network, and then their model achieved 0.894 in the Dice coefficient. Although the segmentation goals were quite different from those of the previous works, the proposed model exhibited a high Dice coefficient for the test image ($D = 0.893$; Fig. 7) sufficient to the requirement of ichnological analysis.

*5.2. Cause of fluctuation in the abundance of *Phycosiphon**

To the best of our knowledge, this study was the first to compare the continuous and quantitative variation in ichnotaxon abundance at high resolution with other environmental parameters. Among environmental factors, sediment accumulation rate has generally been known to influence bioturbation intensity. For example, high sediment accumulation rates may attenuate bioturbation intensity because the available time for bioturbation per unit volume of sediments becomes short. Indeed,

the condensed sections around the maximum flooding surfaces are characterized by highly bioturbated features due to the long time available for bioturbation (e.g., Wetzel, 2010; Wetzel and Uchman, 2012; Callow et al., 2014). In contrast, a high sedimentation rate may cause the opposite trend because a high rate of organic matter burial can increase bioturbation intensity (Wetzel and Uchman, 1998). In both cases, bioturbation intensity could be affected by sediment accumulation rates. However, the results of this study indicated that *Phycosiphon* abundance does not negatively or positively correlate with sediment accumulation rates. This feature differs from general trends in most types of the trace fossils (e.g., Callow et al., 2014), which may be related to the peculiar behavior of *Phycosiphon* producers.

Another possibility is that the abundance of *Phycosiphon* may correlate with the frequency of geologic events, and therefore an increase in LSAR or mLSAR does not decrease the bioturbation intensity B_i . *Phycosiphon* producers have opportunistic behavior because they often occur at the top of intervals where biological activity is suppressed by processes such as turbidity current inflows or anoxic events (Wetzel and Bromley, 1994; Wetzel and Uchman, 2001). In this study, however, the moving average of $\log B_i$ did not increase with the frequency of turbidites. Instead, it tended to decrease with the number of sand layers NLSI (Fig. 13C). The mudstone layers with low values of the moving average of $\log B_i$ were relatively thin (< 2 cm), intercalated between the turbidite

sandstones, and parallel lamination was common. These characteristics imply that these thin-bedded mudstones were deposited by turbidity currents (Stow and Smillie, 2020). Thus, the abundance of *Phycosiphon* does not necessarily increase with the frequency of physical disturbances.

Although the cause of this negative correlation is unknown, it could be explained by the dominance of muddy intervals of turbidites in the successions. The bioturbation intensity in the mud turbidites is generally low because of their rapid deposition (Uchman and Wetzel, 2012). The *Phycosiphon* producers cannot burrow below the redox boundary because they require a fully oxygenated habitat (Wetzel and Uchman, 2001). Therefore, if pore water did not contain sufficient oxygen, the *Phycosiphon* producer cannot disturb the rapidly deposited mud intervals of turbidites (Stow and Wetzel, 1990). Mudstone intervals with the low B_i may reflect bioturbation attenuation due to the frequent deposition of turbidite muddy intervals (Bouma E(t) division) with low oxygenated pore water. Further investigations are required to discriminate between the hemipelagic mudstones and mud turbidites to determine the cause of *Phycosiphon* abundance fluctuation.

More data are required to reveal the mode of the response of *Phycosiphon* to environmental changes, such as the total organic carbon or other paleoenvironmental proxies. Generally, the abundance and diversity of trace fossils vary with food availability (Wetzel and Uchman, 1998; Dashtgard and

Gingras, 2012). Specifically, the occurrence and geometry of *Phycosiphon* depend on the distribution of organic matter in the deposits because its producers may select and burrow deposits with high organic matter concentrations (Wetzel, 2010). The amount of food available for benthic animals must be estimated from the organic matter contents preserved in deposits; however, it is difficult to reconstruct the condition of organic matter in the environment (Wetzel and Uchman, 2018). Other paleoenvironmental conditions can cause the variation in *Phycosiphon* abundance. For example, it is reported that the abundance of *Phycosiphon* and cooccurred ichnotaxa can vary depending on the cyclicity of oceanographic mechanisms influenced by orbital forcing (Pieńkowski et al., 2021). To explain the variation in *Phycosiphon* abundance, the distribution of organic matter contents in the core sections and long-term time-series analysis of environmental proxies must be investigated at high resolution.

5.3. *Advantages and limitations of the proposed model*

In contrast to previous studies (Ayranci et al., 2021; Timmer et al., 2021), the proposed model has the advantage of directly analyzing the distribution of trace fossils in core section images. The proposed model can classify the core images for each pixel, requiring no additional manual operations to estimate the bioturbation intensity B_i , which can be directly calculated from the predicted image.

The precise estimates of the trace fossil regions are also beneficial for subsequent analyses, such as ichnofaunal or shape analysis for classification and

paleoecology. The proposed model was trained by the data from a single site, so that it is not generalized for every location. However, the model optimized for any location or ichnotaxon can be obtained by preparing a training dataset from the target location or ichnotaxon and using the framework of model construction in this study.

The critical characteristic of the proposed model is its ability to process many core section images to examine a significant number of datasets, thus providing general tendencies of ichnological responses to the surrounding environments. For instance, some deep-marine ichnotaxa interpreted as traces of deposit-feeding animals change their meandering mode depending on the organic matter input conditions (Wetzel, 2010). The morphological parameters of trace fossils, such as diameter, length, and sinuosity, can be easily measured from the predicted images, and the relationships between the morphological variation of these parameters and the paleoenvironmental conditions can be investigated in future studies.

Although this model successfully identified the trace fossil *Phycosiphon*, as described above, the proposed model failed to predict the other types of trace fossils. This difficulty can be attributed to two possible causes: ichnotaxa classification problem and class imbalance. Other ichnotaxa had large sizes relative to the size of the tiles (224×224) compared with *Phycosiphon*, so the model could not learn the characteristics of the entire morphology of the ichnotaxa. In addition, the high

contrast in brightness between the core of *Phycosiphon* and surrounding deposits may have simplified learning for the CNN model. These problems can be solved by optimizing the tile size or preprocessing the images before training to adjust the colors. Furthermore, the amount of data for the ichnotaxa other than that for *Phycosiphon* might have been insufficient to learn from this study. The number of *Phycosiphon* pixels in the training dataset was approximately 8.5 times greater than that of the other ichnotaxa. In future studies, a more robust model that can predict multiple ichnotaxa should be developed to analyze the entire ichnodiversity.

6. Conclusions

The semantic segmentation model was constructed using a U-Net-type CNN to automatically extract the regions of the trace fossil *Phycosiphon* from core section images. The proposed model used residual connections and attention gates to obtain the feature quantities of trace fossils in core section images. The model was trained using the image dataset of IODP Exp. 362 Site U1480 (Nicobar Fan in the Sumatra Subduction Zone), and it achieved 0.889 in the Dice coefficient, the similarity metric between the model prediction results and ground truth, indicating that it can accurately predict the distribution of *Phycosiphon*.

The model developed in this study enabled the evaluation of long-term variation in the abundance of *Phycosiphon* in Nicobar Fan deposits over the past 10 myr. As a result, the pattern of the variation depended on the number of sandstone layer

intercalations. The cause of this variation is unknown, but it could imply that the succession contains many muddy turbidites that were difficult to identify with the naked eye. In the future, a model that can predict the distribution of ichnotaxa other than *Phycosiphon* should be developed. This model cannot recognize other ichnotaxa because of insufficient training datasets. Accumulating training datasets for various ichnofossils to develop the semantic segmentation models will lead to a better understanding of the general trends and diversity in bioturbation intensity and ichnodiversity.

Acknowledgments

Dr. Alfred Uchman and three anonymous reviewers critically read the manuscript and made helpful suggestions. This study was supported by JSPS KAKENHI Grant Number JP20J01599 and JP20H0185, and the Sediment Dynamics Research Consortium.

Data Availability

Dataset and codes for the model training in this study is available in the open access repository here: <https://doi.org/10.5281/zenodo.10400790>.

Appendix

This study utilized a CNN model that had a U-Net-like architecture with residual connections (Zhang et al., 2018; Shang et al., 2021) and attention mechanisms (Oktay et al., 2018) for the semantic segmentation model (Fig. 3). The CNN is a framework of deep learning technique for the multiple dimensional data such as RGB images, which extracts image features by convolutional layers. U-Net (Ronneberger et al., 2015) is an encoder–decoder type CNN that has been used for various semantic segmentation models. Semantic segmentation is a type of classification tasks, which classifies pixels in an image. The encoder–decoder model consists of three parts. The encoder network extracts image features and downsizes them through convolution and max-pooling operations, and the decoder network enlarges the feature maps to match the original image sizes using convolution and upsampling operations. The bridge network connects the encoder and decoder networks. In contrast to other encoder–decoder models, U-Net is characterized by skip connections between feature maps of the same level in the encoder and decoder networks. The skip connections, which copy and concatenate the feature map from the encoder to decoder networks, provide the positional information of the pixels before downsizing.

In the CNN architecture used in this study, the convolutional operations commonly used in normal-type U-Net were replaced by the preactivation-type

residual blocks (He et al., 2016) (Fig. 3). The residual block in this study comprised two subnetworks: the main and residual-connection subnetworks. The main subnetwork consisted of two repetitions of batch normalization, activation, and convolutional layers (Fig. 4A). The convolutional layers are product-sum operations between an image and filters. The convolutional layers in the main subnetwork used 3×3 -sized filters. Batch normalization layers, which are defined as the normalization of feature values based on mean and variance in a subset of data called a mini-batch, were used to maintain the stability of the training of deep-layered networks (Ioffe and Szegedy, 2015). A rectified linear unit (ReLU) was used as the activation function. The residual connection is a pathway by-passing a convolutional layer by adding the feature maps before and after the convolutional layer. The residual-connection subnetwork comprised the convolutional and batch normalization layers (Fig. 4A). The convolutional layer in the residual-connection subnetwork used 1×1 -sized filters to adjust the channel-wise depth to match the resultant tensor of the main subnetwork. At the end of the residual block, the tensor from the residual-connection subnetwork was added to that from the main subnetwork. This operation can prevent gradient disappearance or explosion during backpropagation (He et al., 2016).

The CNN encoder and decoder networks contained four residual blocks. The first batch normalization and activation layers were omitted from the first residual block in the encoder network because the inputs to the first residual block were RGB channels of images without any feature-extracting operations. The max-

pooling operation was performed in the encoder network to downsize the feature map after each residual block. In the decoder network, the upsampling operation and skip connection were inserted immediately before each residual block to stretch the size of the feature map.

The bridge network comprises two subnetworks. The residual connection subnetwork utilized the same structure as that used in the residual blocks of the encoder and decoder networks (Fig. 3). The other subnetwork consisted of three repetitions of the convolutional, batch normalization, and activation layers. The first convolutional layer reduced the channel-wise depth by half with 1×1 -sized filters, and the second convolutional layer performed pixel-wise convolution with 3×3 -sized filters. The third convolutional layer with 1×1 -sized filters increased the channel-wise depth to double that of the input tensor of the bridge network. The activation functions in the bridge network were also ReLU.

In the decoder network, attention gates (Oktay et al., 2018) were used before concatenation (Fig. 3), which indicates the region in images to be considered for pixel classification. The attention gates receive two feature maps: the upsampled feature map from the lower layer of the decoder network and the feature map from the same level in the encoder network (Fig. 4B). The feature maps were processed in the first convolutional layers with 1×1 -sized filters to equalize the number of channels. Subsequently, feature maps were added and activated using ReLU. The convolution operation was performed on the feature map in the second

convolutional layer with a 1×1 -sized filter to reduce the number of channels to 1. The feature map was activated using a sigmoid function to calibrate values in the range 0–1. These values were used as the weighting coefficients w_A multiplied by the feature map provided by the skip connection. Through the above operations, the region in the feature maps where w_A was close to unity was emphasized in subsequent operations.

At the end of the decoder network, a convolutional layer with 1×1 -sized filters and an activation layer using the Softmax function was set. The former compiled the channel-wise depth of the feature maps into three. The latter converts the output values to a one-hot encoded probability distribution.

During training, the training dataset was divided into minibatches containing 16 pairs of tiles. Each time a minibatch was loaded, data augmentation was randomly applied to the tiles. Three types of data augmentation were applied: vertical, horizontal flipping, and Gaussian noise (0.0 in the mean and 0.2 in the standard deviation).

References

Acharya, T., Ray, A.K., 2005. Image enhancement and restoration, In: Acharya, T., Ray, A.K. (Eds.), Image processing: principles and applications. John Wiley & Sons, pp. 105–129.

- Ayranci, K., Yildirim, I.E., Waheed, U.b., MacEachern, J.A., 2021. Deep learning applications in geosciences: Insights into ichnological analysis. *Applied Sciences* 11, 7736. <https://doi.org/10.3390/app11167736>.
- Bednarz, M., McIlroy, D., 2009, Three-dimensional reconstruction of “Phycosiphoniform” burrows: implications for identification of trace fossils in core. *Palaeontologica Electronica* 12, 1–15.
- Buatois, L.A., Mángano, M.G., 2013. Ichnodiversity and ichnodisparity: significance and caveats. *Lethaia* 46, 281–292.
- Callow, R.H., Kneller, B., Dykstra, M., McIlroy, D., 2014. Physical, biological, geochemical and sedimentological controls on the ichnology of submarine canyon and slope channel systems. *Marine and Petroleum Geology* 54, 144–166.
- Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2016. Semantic image segmentation with deep convolutional nets and fully connected CRFs. *arXiv*, 1412.7062 <https://doi.org/10.48550/arXiv.1412.7062>.
- Cummings, J.P., Hodgson, D.M., 2011. Assessing controls on the distribution of ichnotaxa in submarine fan environments, the Basque basin, northern Spain. *Sedimentary Geology* 239, 162–187.
- Dashtgard, S.E., Gingras, M.K., 2012. Marine invertebrate neoichnology, In: Knaust, D., Bromley, R.G. (Eds.), *Trace Fossils as Indicators of Sedimentary*

- Environments. Elsevier. volume 64 of *Developments in Sedimentology*, pp. 273–295.
- Diakogiannis, F.I., Waldner, F., Caccetta, P., Wu, C., 2020. Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing* 162, 94–114.
- Dorador, J., Rodríguez-Tovar, F.J., 2014. Quantitative estimation of bioturbation based on digital image analysis. *Marine Geology* 349, 55–60.
- Droser, M.L., Bottjer, D.J., 1986. A semiquantitative field classification of ichnofabric. *Journal of Sedimentary Research* 56, 558–559.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Identity mappings in deep residual networks, In: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), *Computer Vision – ECCV 2016*, Springer International Publishing, Cham. pp. 630–645.
- Heard, T.G., Pickering, K.T., 2008, Trace fossils as diagnostic indicators of deep-marine environments, Middle Eocene Ainsa-Jaca basin, Spanish Pyrenees. *Sedimentology* 55, 809–844.
- Heard, T.G., Pickering, K.T., Clark, J.D., 2014. Ichnofabric characterization of a deep-marine clastic system: a subsurface study of the middle Eocene Ainsa system, Spanish Pyrenees. *Sedimentology* 61, 1298–1331.
- Hou, Y., Canul-Ku, M., Cui, X., Hasimoto-Beltran, R., Zhu, M., 2021, Semantic

- segmentation of vertebrate microfossils from computed tomography data using a deep learning approach. *Journal of Micropalaeontology* 40, 163–173.
- Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift, In: Bach, F., Blei, D. (Eds.), *Proceedings of the 32nd International Conference on Machine Learning*, PMLR, Lille, France. pp. 448–456.
- Knaust, D., Warchoř, M., Kane, I.A., 2014, Ichnodiversity and ichnoabundance: Revealing depositional trends in a confined turbidite system. *Sedimentology* 61, 2218–2267.
- Lei, M., Rao, Z., Wang, H., Chan, Y., Zou, L., Yu, H., 2021, Maceral groups analysis of coal based on semantic segmentation of photomicrographs via the improved U-net. *Fuel* 294, 120475.
<https://doi.org/10.1016/j.fuel.2021.120475>.
- Liu, L., Jiang, H., He, P., Chen, W., Liu, X., Gao, J., Han, J., 2019. On the variance of the adaptive learning rate and beyond. *arXiv*, 1908.03265.
<https://doi.org/10.48550/arXiv.1908.03265>.
- Liu, X., Song, H., 2020, Automatic identification of fossils and abiotic grains during carbonate microfacies analysis using deep convolutional neural networks. *Sedimentary Geology* 410, 105790.
<https://doi.org/10.1016/j.sedgeo.2020.105790>.

- MacEachern, J. A., and Bann, K. L., 2020, The *Phycosiphon* Ichnofacies and the Rosselia Ichnofacies: Two new ichnofacies for marine deltaic environments. *Journal of Sedimentary Research* 90, 855–886.
- McNeill, L.C., Dugan, B., Petronotis, K.E., Backman, J., Bourlange, S., Chemale Jr., F., Wenhua, C., Colson, T.A., Frederik, M.C.G., Guérin, G., Hamahashi, M., Henstock, T., House, B.M., Hüpers, A., Jeppson, T.N., Kachovich, S., Kenigsberg, A.R., Kuranaga, M., Kutterolf, S., Milliken, K.L., Mitchison, F.L., Mukoyoshi, H., Nair, N., Owari, S., Pickering, K.T., Pouderoux, H.F.A., Yehua, S., Song, I., Torres, M.E., Vannucchi, P., Vrolijk, P.J., Tao, Y., Zhao, X., 2017. Site U1480, In: McNeill, L.C., Dugan, B., Petronotis, K.E., the Expedition 362 Scientists (Eds.), *Proceedings of the International Ocean Discovery Program*, International Ocean Discovery Program. <https://doi.org/10.14379/iodp.proc.362.103.2017>.
- Miller, M.F., Smail, S.E., 1997. A semiquantitative field method for evaluating bioturbation on bedding planes. *Palaaios* 12, 391–396.
- Naruse, H., Nifuku, K., 2008, Three-dimensional morphology of the ichnofossil *Phycosiphon incertum* and its implication for paleoslope inclination. *Palaaios* 23, 270–279.
- Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D., 2018.

- Attention U-Net: Learning where to look for the pancreas. arXiv, 1804.03999. <https://doi.org/10.48550/arXiv.1804.03999>.
- Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 629–639.
- Pickering, K.T., Pouderoux, H., McNeill, L.C., Backman, J., Chemale, F., Kutterolf, S., Milliken, K.L., Mukoyoshi, H., Henstock, T.J., Stevens, D.E., Parnell, C., Dugan, B., 2020. Sedimentology, stratigraphy and architecture of the Nicobar Fan (Bengal–Nicobar Fan System), Indian Ocean: Results from International Ocean Discovery Program Expedition 362. *Sedimentology* 67, 2248–2281.
- Pieńkowski, G., Uchman, A., Ninard, K., Hesselbo, S.P., 2021. Ichnology, sedimentology, and orbital cycles in the hemipelagic Early Jurassic Laurasian Seaway (Pliensbachian, Cardigan Bay Basin, UK). *Global and Planetary Change* 207, 103648. <https://doi.org/10.1016/j.gloplacha.2021.103648>.
- Rebata-H., L.A., Gingras, M.K., Räsänen, M.E., Barberi, M., 2006. Tidal-channel deposits on a delta plain from the Upper Miocene Nauta Formation, Marañón Foreland Sub-basin, Peru. *Sedimentology* 53, 971–1013.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation, In: Navab, N., Hornegger, J., Wells, W.M.,

- Frangi, A.F. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham. pp. 234–241.
- Shang, G., Liu, G., Zhu, P., Han, J., Xia, C., Jiang, K., 2021. A deep residual U-type network for semantic segmentation of orchard environments. *Applied Sciences* 11. <https://doi.org/10.3390/app11010322>.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>.
- Stow, D.A.V., Smillie, Z., 2020. Distinguishing between deep-water sediment facies: Turbidites, contourites and hemipelagites. *Geosciences* 10, 68. <https://doi.org/10.3390/geosciences10020068>.
- Stow, D.A.V., Wetzel, A., 1990, Hemiturbidite: A new type of deep-water sediment. *Proceedings of the Ocean Drilling Program, Scientific Report* 116, 25–34.
- Taylor, A., Goldring, R., 1993. Description and analysis of bioturbation and ichnofabric. *Journal of the Geological Society* 150, 141–148.
- Timmer, E., Knudson, C., Gingras, M., 2021. Applying deep learning for identifying bioturbation from core photographs. *AAPG Bulletin* 105, 631–638.

- Tonkin, N.S., 2012. Deltas, In: Knaust, D., Bromley, R.G. (Eds.), Trace Fossils as Indicators of Sedimentary Environments. Elsevier. volume 64 of *Developments in Sedimentology*, pp. 507–528.
- Uchman, A., Wetzel, A., 2012. Deep-sea fans, In: Knaust, D., Bromley, R.G. (Eds.), Trace Fossils as Indicators of Sedimentary Environments. Elsevier. volume 64 of *Developments in Sedimentology*, pp. 643–671.
- Wetzel, A., 2010. Deep-sea ichnology: observations in modern sediments to interpret fossil counterparts. *Acta Geologica Polonica* 60, 125–138.
- Wetzel, A., Bromley, R.G., 1994. *Phycosiphon incertum* revisited: *Anconichnus horizontalis* is its junior subjective synonym. *Journal of Paleontology* 68, 1396–1402.
- Wetzel, A., Uchman, A., 1998. Deep-sea benthic food content recorded by ichnofabrics; a conceptual model based on observations from Paleogene flysch, Carpathians, Poland. *Palaaios* 13, 533–546.
- Wetzel, A., Uchman, A., 2001. Sequential colonization of muddy turbidites in the Eocene Beloveža Formation, Carpathians, Poland. *Palaeogeography, Palaeoclimatology, Palaeoecology* 168, 171–186.
- Wetzel, A., Uchman, A., 2012. Hemipelagic and pelagic basin plains, In: Knaust, D., Bromley, R.G. (Eds.), Trace Fossils as Indicators of Sedimentary

Environments. Elsevier. volume 64 of *Developments in Sedimentology*, pp. 673–701.

Wetzel, A., Uchman, A., 2018. The former presence of organic matter caused its later absence: Burn-down of organic matter in oceanic red beds enhanced by bioturbation (Eocene Variegated Shale, Carpathians). *Sedimentology* 65, 1504–1519.

Yu, C., Qin, F., Li, Y., Qin, Z., Norell, M., 2022, CT Segmentation of Dinosaur Fossils by Deep Learning. *Frontiers in Earth Science* 9, 805271.
<https://doi.org/10.3389/feart.2021.805271>

Zhang, Z., Liu, Q., Wang, Y., 2018. Road extraction by Deep Residual U-net. *IEEE Geoscience and Remote Sensing Letters* 15, 749–753.

Figure and table captions

Fig. 1. Index map and lithofacies of the study site. (A) Site map showing the drilling location for IODP Exp. 362 Site U1480. (B) Core recovery, depositional age, lithofacies units, and columnar section of Site U1480. Modified after McNeill et al. (2017).

Fig. 2. Examples of dataset. (A) Tiles of the core section image with 224×224

pixels in size. The brightness channel of each tile was histogram-equalized. (B) Tiles of the labeled image corresponding to tiles in A. Black, white, and green parts represent the regions of background, outcrop, and trace fossils, respectively.

Fig. 3. Architecture of the CNN for the proposed model. See Fig. 4 for details of the residual blocks and attention gates.

Fig. 4. Architectures of the residual block (A) and attention gate (B). See Fig. 3 for legend.

Fig. 5. Scheduling function for the base learning rate α_t .

Fig. 6. Core section images containing trace fossils from the study site. (A) *Phycosiphon incertum*. Hole F Core 69X-3. (B) *Palaeophycus* isp. (*Pa*) and *Planolites* isp. (*Pl*). Hole F Core 79X-2. (C) *Scolicia* isp. Hole F 77X-CC. (D) *Nereites* isp. Hole F Core 69X-1. Scale bars: 1 cm.

Fig. 7. Training history of the proposed model. (A) History of loss function. (B) History of the Dice coefficient for metrics.

Fig. 8. Prediction for a part of the test image (lower part of Hole F Core 69X-3). (A) Input image for the proposed model. (B) Ground truth image

painted by manual operation. (C) Model prediction for the image in A. (D) Image showing the difference between the ground truth (B) and predicted (C) images. Blue part represents the pixels correctly predicted and orange part exhibits those failed to predict.

Fig. 9. Scatter plots showing the model ability for estimation of B_i . The test image (Fig. 8) was divided into 14 intervals of 10 cm thickness and B_i was calculated for each interval. The dashed lines represent the equivalent line, and points on the lines indicate perfect reconstructions of B_i .

Fig. 10. Examples of the images predicted by the proposed model. (A) The upper part of Hole F Core 69X-4. (B) Predicted image corresponding to A. (C) Hole G Core 61R-CC. (D) Predicted image corresponding to C.

Fig. 11. Distribution of $\log B_i$ at Site U1480.

Fig. 12. Stratigraphic variation in logarithmic B_i in Site U1480. Data points represent B_i calculated from each core section image. The columnar section (modified after McNeill et al., 2017) and variation of the linear sediment accumulation rate (LSAR; data from Pickering et al., 2020) are also shown. SMA: simple moving average, NSLI: number of sandstone layer intercalations. Abbreviation: mbsf, meter below sea floor.

Fig. 13. Relationships between B_i and the total LSAR (A), the LSAR of mud contents (B), and the number of sand layer intercalations (C). The left columns show the correlation among the parameters. The right columns show the stratigraphic relationships among the parameters.

Table 1 Conditions for model training.

Table 2 Area percentages of each ichnotaxon relative to total bioturbated structures in the core section images for the training and validation datasets.

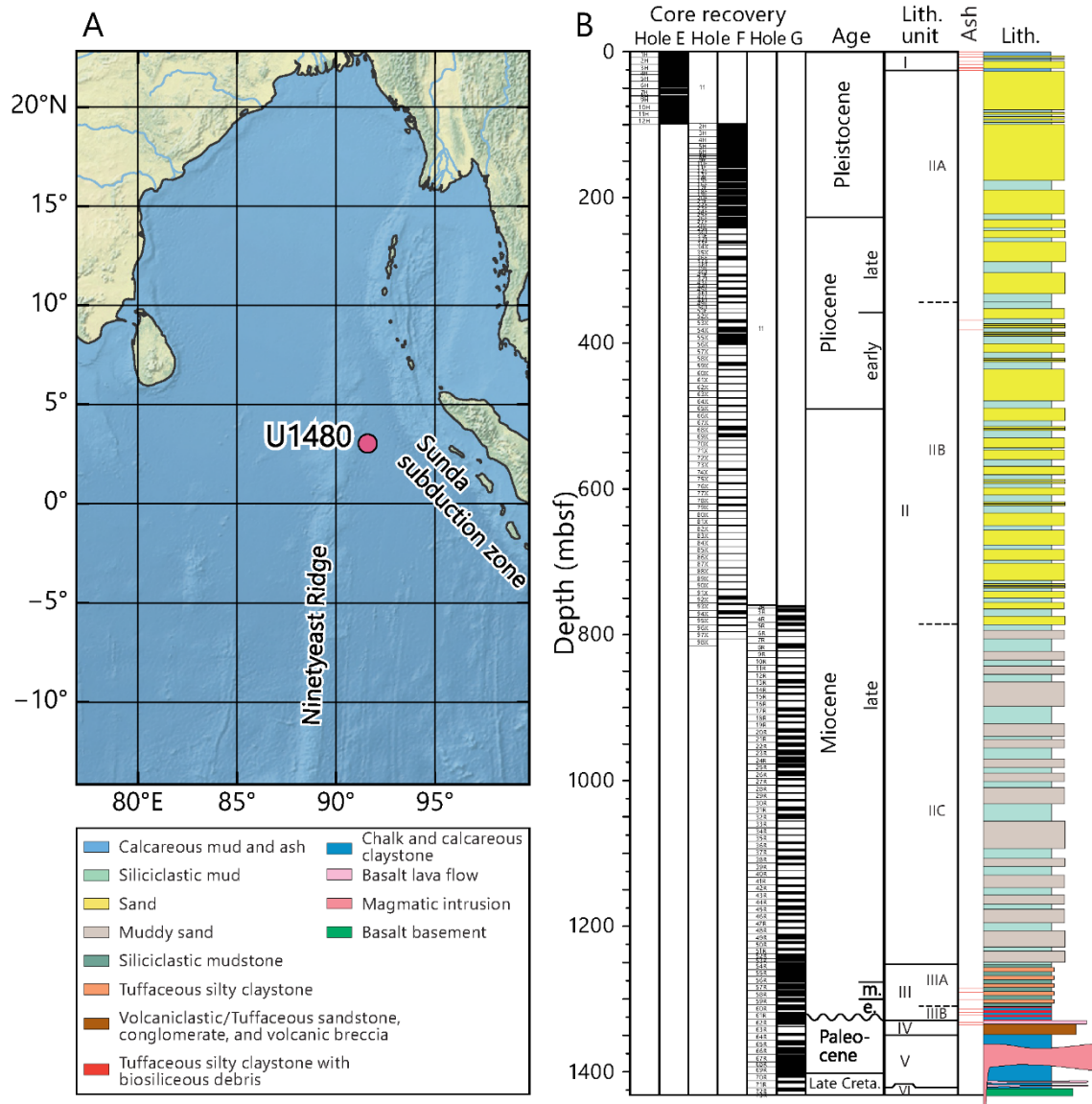


Fig. 1

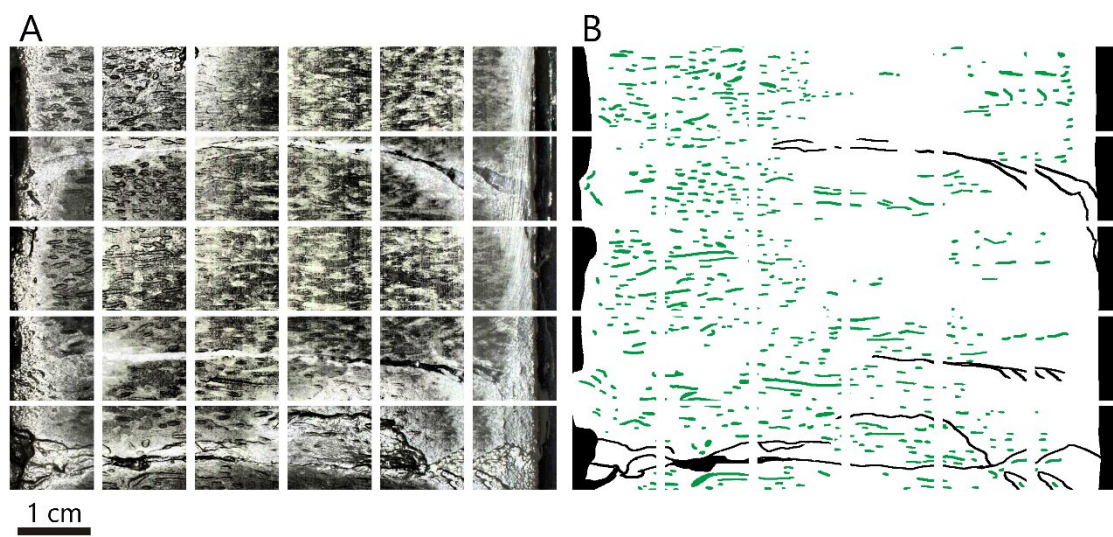


Fig. 2

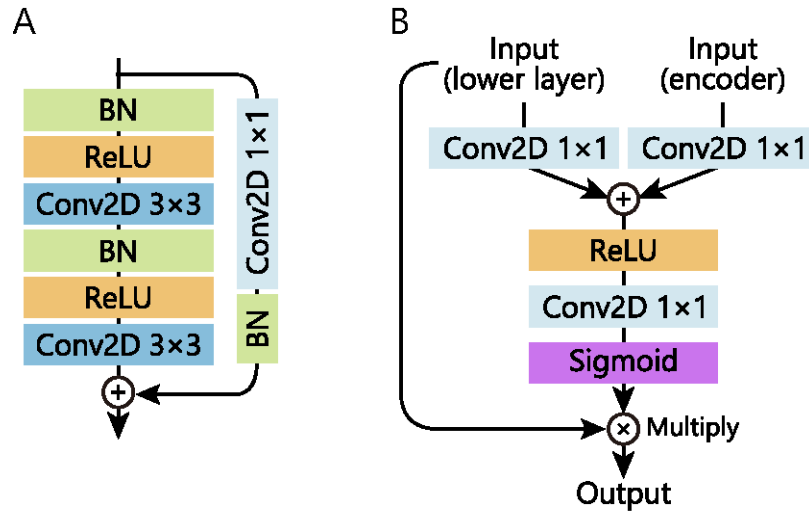


Fig. 4

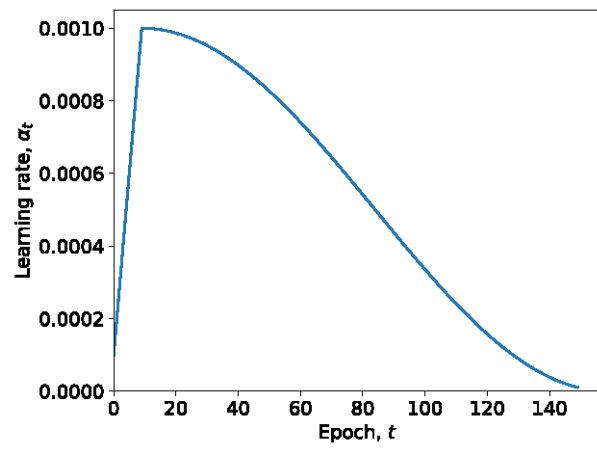


Fig. 5

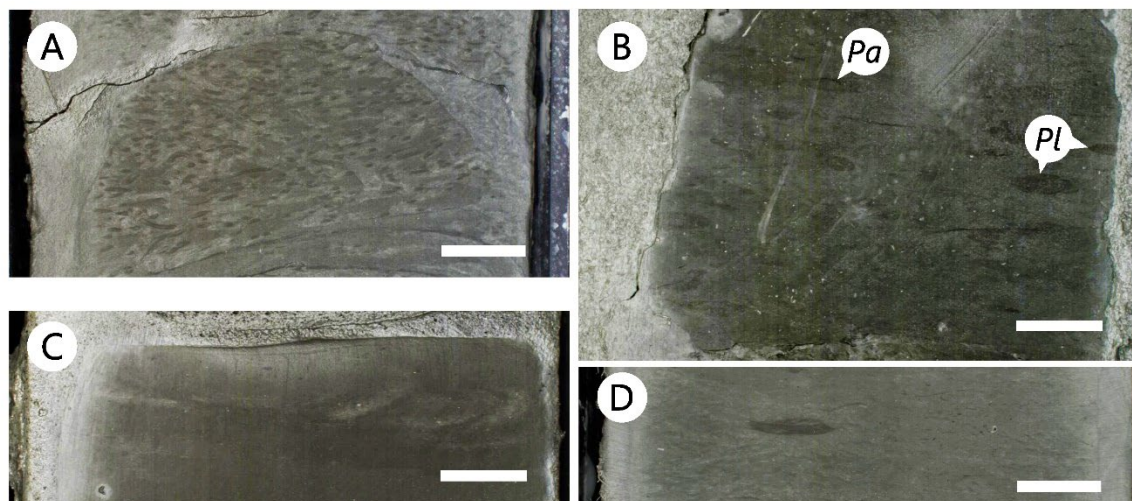


Fig. 6

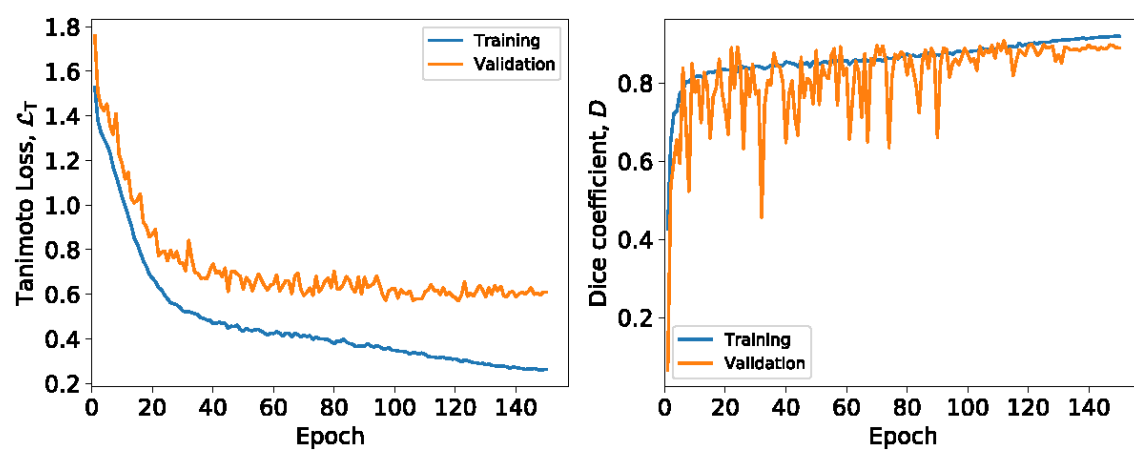


Fig. 7

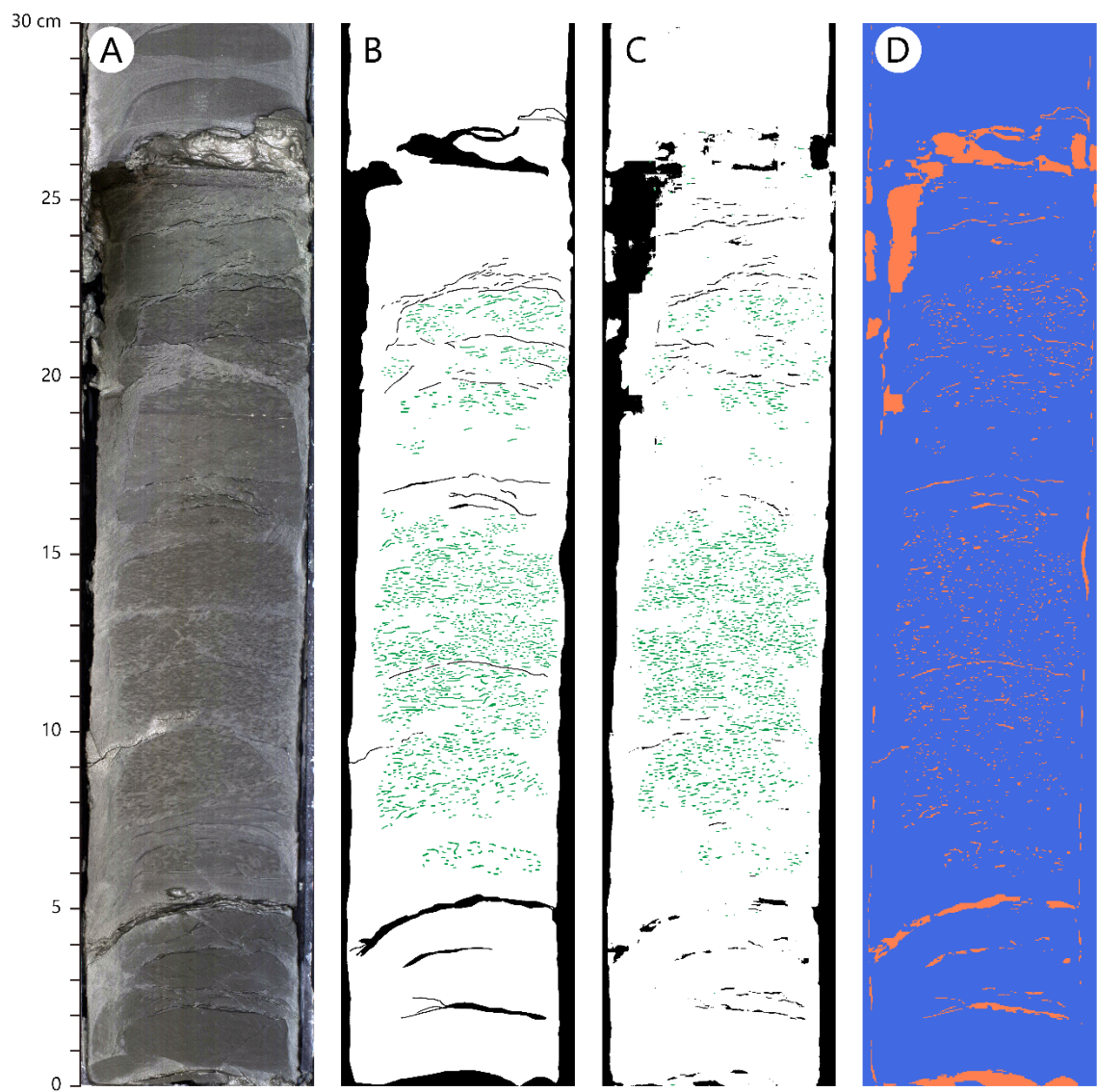


Fig. 8

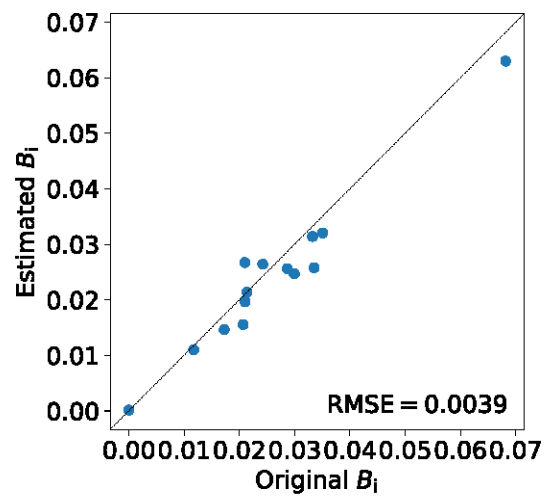


Fig. 9

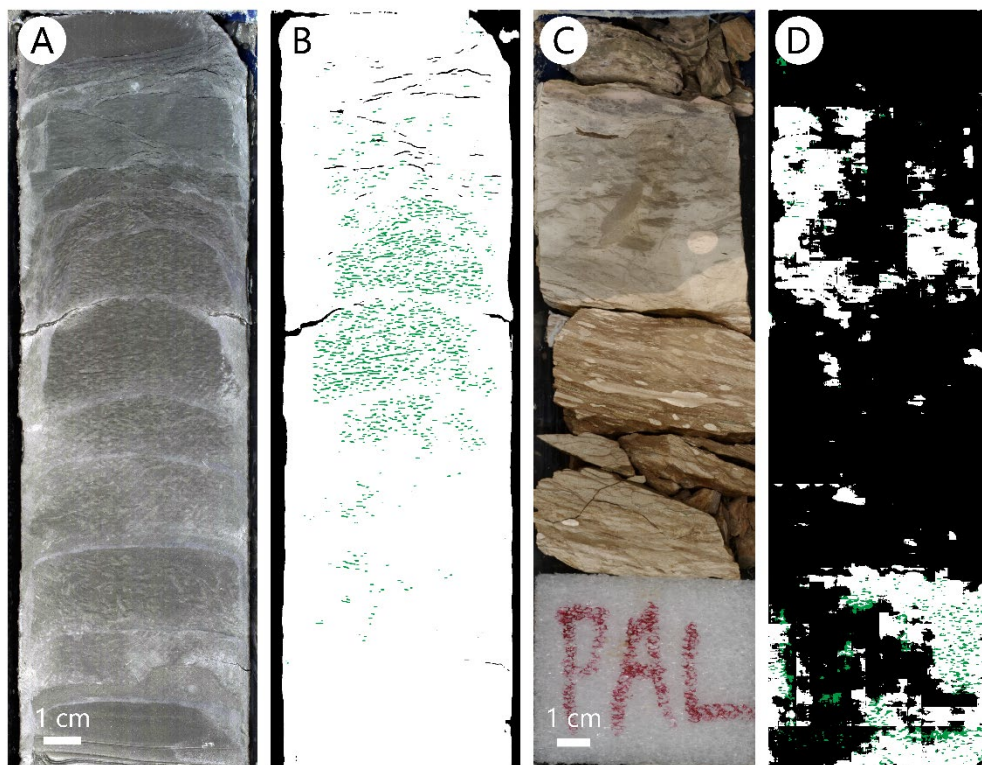


Fig. 10

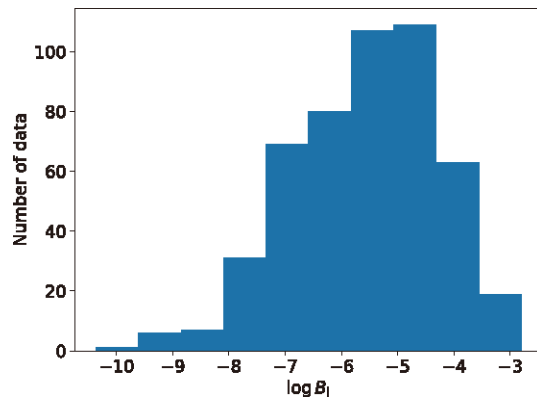


Fig. 11

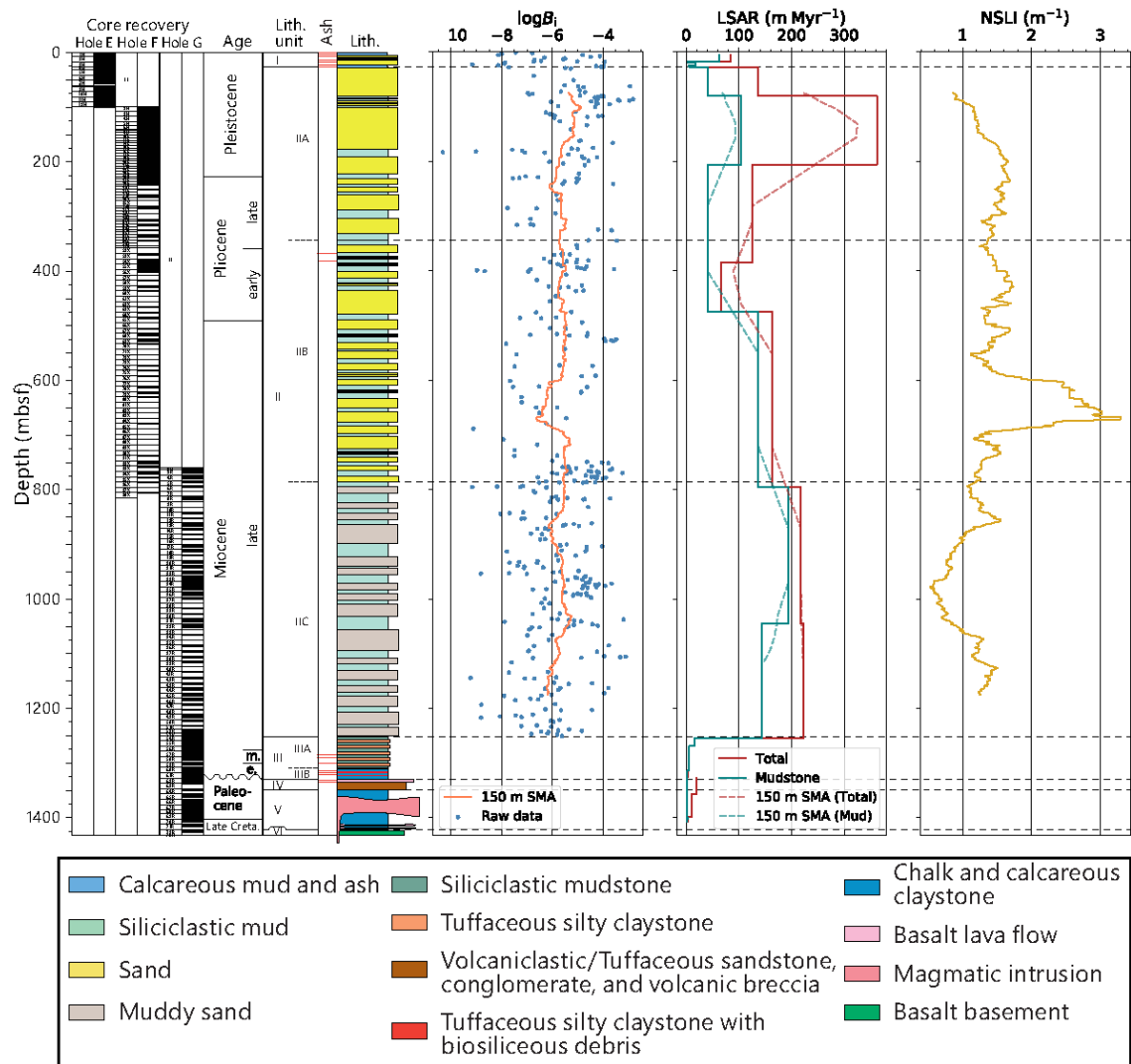


Fig. 12

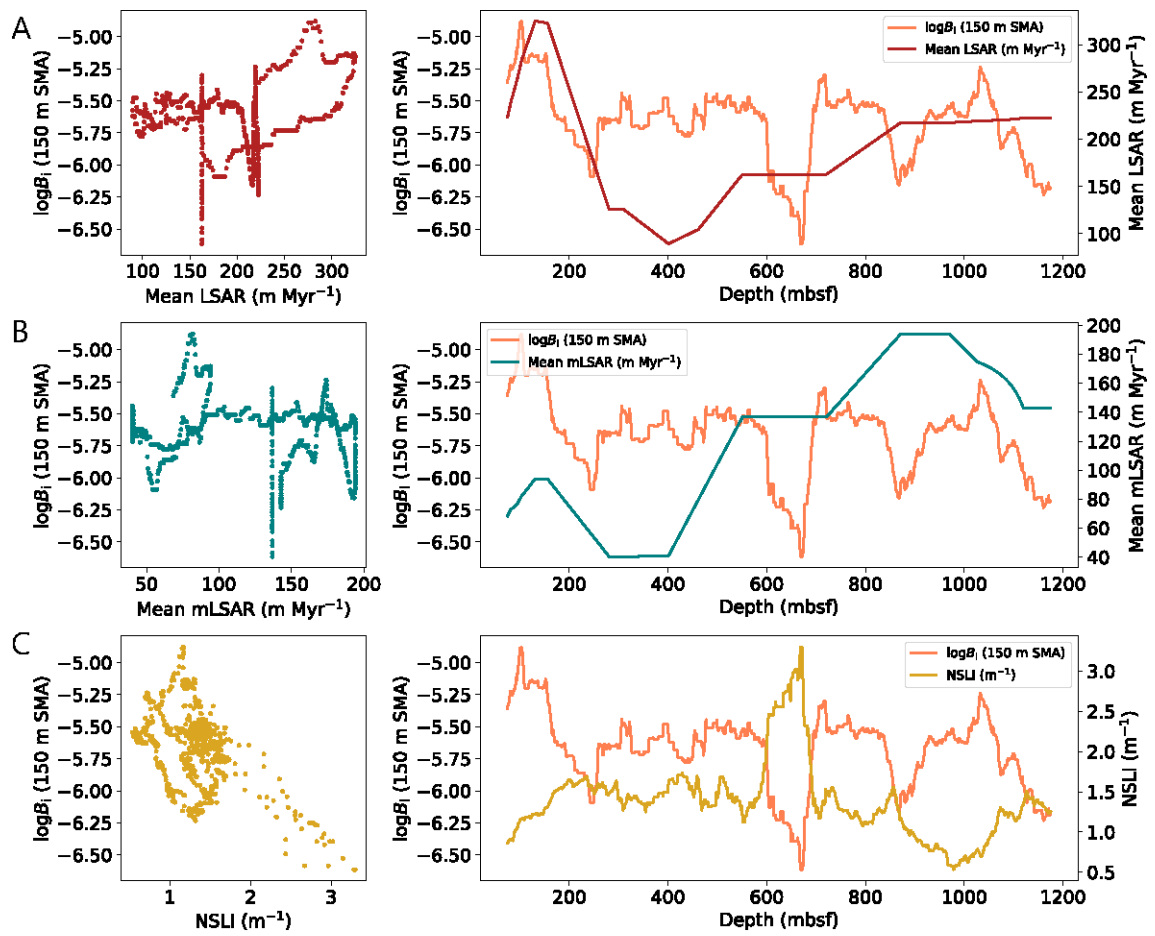


Fig. 13

Table 1

Conditions	
Architecture of CNN	Attention U-Net with residual connections
Input shape	(244, 244, 3)
Number of training data	1151 tiles
Number of validation data	127 tiles
Number of test data	726 tiles
Number of classes	3
Data augmentation	Random vertical or horizontal flipping, Gaussian noise
Epochs	150
Batch size	16
Loss function	Tanimoto loss with complements
Optimizer	RAdam
Learning rate scheduler	Cosine-decay with 10-epoch-warmup
Metrics	Dice coefficient

Table 2

Cores	<i>Nereites</i>	<i>Phycosiphon</i>	<i>Scolicia</i>	<i>Palaeophycus</i>	<i>Planolites</i>
Hole F 69X-1	7.99%	83.61%	1.23%	6.43%	0.74%
Hole F 69X-2	3.29%	92.49%	1.69%	1.33%	1.20%
Hole F 69X-3	0.16%	95.06%	0.00%	1.89%	2.89%
Hole F 76X-1	0.00%	96.78%	0.00%	3.22%	0.00%
Hole F 79X-1	16.75%	73.04%	0.00%	10.21%	0.00%
Total	5.23%	88.70%	0.79%	4.11%	1.17%