ISPRS Journal of Photogrammetry and Remote Sensing 205 (2023) 1–16

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Mapping intertidal topographic changes in a highly turbid estuary using dense Sentinel-2 time series with deep learning

Chunpeng Chen [a,b], Ce Zhang [c,d,*], Bo Tian [a,*], Wenting Wu [e], Yunxuan Zhou [a]

[a] State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China
[b] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[c] School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK
[d] UK Centre for Ecology & Hydrology, Library Avenue, Lancaster LA1 4AP, UK
[e] Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou University, China

ARTICLE INFO

Keywords:
Intertidal topography
Deep learning
Self-attention
Hybrid loss
Morphological change

ABSTRACT

Intertidal mudflats are an important component of the coastal geomorphological system at the interface between ocean and land. Accurate and up-to-date mapping of intertidal topography at high spatial resolution, and tracking of its changes over time, are essential for coastal habitat protection, sustainable management and vulnerability analysis. Compared with ground-based or airborne terrain mapping, the satellite-based waterline method is more cost-effective for constructing large-scale intertidal topography. However, the accuracy of the waterline method is affected by the extraction of waterlines and the calibration of waterline height. The blurred boundary between turbid water and mudflats in the tide-dominated estuary brings enormous challenges in accurate waterline extraction, and the errors in estuarine water level simulations prevent the direct calibration of waterline heights. To address these issues, this paper developed a novel deep learning method using a parallel self-attention mechanism and boundary-focused hybrid loss to extract turbid estuarine waterlines accurately from dense Sentinel-2 time series. UAV photogrammetric surveys were employed to calibrate waterline heights rather than the simulated water levels, such that the error propagation is constrained effectively. Annual intertidal topographic maps of the Yangtze estuary in China were generated from 2020 to 2022 using the optimized waterline method. Experimental results demonstrate that the proposed deep learning method could achieve excellent performance in land and water segmentation in time-varying tidal environments, with better generalization capability compared with benchmark U-Net, U-Net++ and U-Net+++ models. The comparison between the generated topography and UAV photogrammetric observations resulted in an RMSE of 13 cm, indicating the effectiveness of the optimized waterline method in monitoring morphological changes in estuarine mudflats. The generated topographic maps successfully identified hotspots of mudflat erosion and deposition. Specifically, the mudflats connected to the land predominantly experienced deposition of 10–20 cm over the two-year period, whereas the offshore sandbars exhibited instability and significant erosion of 20–60 cm during the same period. These topographic maps serve as valuable datasets for providing scientific baseline information to support coastal management decisions.

1. Introduction

Intertidal mudflats (also known as tidal flats), located in the critical transition between marine and terrestrial ecosystems, are an important component of the coastal geomorphological system (Gao, 2019). These areas support habitats for waterbirds (Iwamura et al., 2013; Wang et al., 2018), provide land resources for urban development (Tian et al., 2016), and serve as buffers against marine related hazards (Arkema et al., 2013; Schoutens et al., 2019). However, anthropogenic interventions such as coastal reclamation and abrupt reductions in sediment input induced by upstream damming, combined with sea level rise resulting from climate change, have led to widespread erosion and loss of these intertidal areas (Murray et al., 2019; Nienhuis et al., 2020). Accurate and timely mapping of tidal flat topography with high spatio-temporal resolution, and

* Corresponding authors at: School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, UK (C. Zhang).
E-mail addresses: c.zhang9@lancaster.ac.uk (C. Zhang), btian@sklec.ecnu.edu.cn (B. Tian).

https://doi.org/10.1016/j.isprsjprs.2023.09.022
Received 4 May 2023; Received in revised form 12 September 2023; Accepted 22 September 2023
Available online 1 October 2023

tracking of its changes are essential for coastal habitat protection, sustainable coastal management and vulnerability analysis (Kulp & Strauss, 2019; Nicholls et al., 2021). However, the topographic mapping and continuous monitoring of mudflats are extremely challenging since these intertidal environments are spatially complex and temporally dynamic under the compound effects of physical, hydrological and biological processes (Fagherazzi et al., 2012; Wang et al., 2019).

Due to poor accessibility and short exposure times, morphological monitoring of intertidal mudflat relies heavily on remotely sensed techniques, including airborne or terrestrial light detection and ranging (LiDAR) (Xie et al., 2017; Andriolo et al., 2018), unmanned aerial vehicle (UAV) structure-from-motion (SfM) photogrammetry (Kalacska et al., 2017; Chen et al., 2022), video imaging (Holman & Stanley, 2007; Ellenson et al., 2020), and satellite-based waterline method (Heygster et al., 2010; Salameh et al., 2020). These approaches vary in terms of data processing complexity, resulting accuracy, economic cost, and spatiotemporal coverage. Ground- and airborne-based techniques such as LiDAR and SfM tend to have limited spatial coverage, although they can provide terrain with centimeter-level accuracy. They often require a high economic investment for high-frequency observations and are therefore not suitable for continuous and long-term observations of tidal flats at large spatial scales. Consequently, the satellite-based waterline method to date is still the most cost-effective approach for monitoring large-scale intertidal topography and has been successfully applied to tidal flats globally, such as the German Wadden Sea (Wiehle & Lehner, 2015), the Dee Estuary in UK (Bell et al., 2016), the Jiangsu coast in China (Wang et al., 2019), the Australian coast (Bishop-Taylor et al., 2019b), and the Arcachon Bays in France (Salameh et al., 2020).

The waterline method utilizes time-series waterlines recorded at different tidal stages and synchronized *in-situ* observations or simulated water levels to reconstruct intertidal topography (Mason et al., 1995). This method consists of three major steps: waterline delineation, waterline height determination and waterline interpolation. Accordingly, errors in such a method arise from these steps. First, the accuracy of instantaneous waterline extraction depends on the extraction algorithm and the spatial resolution of the input images. Currently, edge extraction techniques (Salameh et al., 2020; Yang et al., 2022), water index thresholding methods (Sagar et al., 2017; Bishop-Taylor et al., 2019a; Tong et al., 2020), supervised classification (Banks et al., 2015), and clustering algorithms (Obida et al., 2019) could provide pixel or even subpixel level accuracy for water and land segmentation. However, the high moisture content of the mudflat surface in estuarine areas and the high concentration of suspended sediment in the surrounding seawater, which varies with river input, make it difficult for the above methods to distinguish the boundary accurately between mudflat and seawater (de Vries et al., 2021). Furthermore, previous practices of the waterline method have typically used simulated water levels to assign waterline heights (Liu et al., 2013; Zhang et al., 2022). These simulated water levels contain an error of up to 30 cm due to the effects of estuarine runoff (Gao et al., 2021). Such an error would propagate into terrain reconstruction, which is unacceptable for monitoring morphological change in estuarine mudflats. In addition, the waterline method needs to consider the mudflat morphological stability during the period of image acquisition (Ryu et al., 2008; Salameh et al., 2020). The use of an insufficient number of images or a prolonged period of image acquisition can lead to inaccuracies in the constructed topography. Previous studies have attempted to increase the number of available images by extending the period of image acquisition, for example by a full year or more (Khan et al., 2019; Jain et al., 2022; Chen et al., 2023b). For the purpose of constructing intertidal topography for use in numerical modelling, the trade-off between the number of images and the period of image acquisition is acceptable. However, estuarine areas are subject to significant intra-annual morphological variability due to the influence of changing river discharge and sediment transport, as well as frequent summer typhoons. The use of topographic maps reconstructed from a full year of satellite observations for inter-annual comparisons may obscure actual morphological changes in the mudflats. It is therefore important to select an appropriate period for the construction of mudflat topography for estuarine areas.

The limitations mentioned above make it impractical to directly replicate the existing approaches of waterline extraction and waterline height calibration for detecting intertidal topographic variability in turbid estuaries. It is therefore necessary to adapt these methods to improve mapping accuracy for large-scale applications. Currently, deep learning (DL) models have been widely used for semantic segmentation of remotely sensed imagery and are gradually being introduced into shoreline detection. For example, Aghdami-Nia et al. (2022), Pucino et al. (2022), Dang et al. (2022) and Seale et al. (2022) implemented shoreline extraction using the U-Net architecture and its variants with significantly higher accuracy than traditional methods. Thus, the integration of DL-based waterline extraction into the waterline method for intertidal topography reconstruction could potentially increase the accuracy of the results. Moreover, local-scale topographic observations have been made in many intertidal areas using LiDAR or UAV photogrammetry to further understand the coastal evolution. The use of these local observations to calibrate waterline elevations instead of low credible water levels can potentially be an ideal solution. In addition, the deployment of the European Space Agency's (ESA) twin Sentinel-2 satellites has made it possible to acquire sufficient images in the short term (avoiding morphological changes) to reconstruct intertidal topography. Thus, the iterations of advanced technology and the accumulation of available datasets offer a new opportunity to monitor intertidal morphological change in highly dynamic and complex estuarine regions.

In this paper, we aim to optimize the waterline method in terms of waterline extraction, waterline quality control and waterline height determination to maximize its accuracy in intertidal topography reconstruction. The specific objectives of this study are 1) to overcome the challenge of accurate extraction of highly turbid estuarine waterlines by developing a deep learning model with a parallel self-attention mechanism and boundary loss optimization; 2) to eliminate the error of waterline height determination by using UAV photogrammetric profiles instead of water levels to assign waterline heights; and 3) to investigate the morphological changes of the Yangtze estuarine mudflats from 2020 to 2022.

2. Study area

The study area is located on the Eastern shore of Chongming Island at the mouth of the Yangtze River in China. The abundant sediment loads from the Yangtze River have formed extensive intertidal mudflats, including Dongtan (DT), Guyuan Shoal (GYS) and Tuanjie Shoal (TJS) (Fig. 1). These areas consist of the largest nature reserve for international migratory birds in East Asia. DT mudflats are approximately 40 km long from north to south, with a mudflat width of approximately 1.5 km in the north, 3 km in the center and 500 m in the south. Tidal surface gradients vary over such a large spatial span, resulting in different elevations along the same instantaneous waterline. Therefore, based on the a priori geomorphological characteristics and terrain slope obtained from UAV photogrammetry, DT mudflats were divided into three sections: North of Dongtan (NDT), Middle of Dongtan (MDT), South of Dongtan (SDT). GYS and TJS are offshore sandbars with a maximum mudflat width of about 4.5 km and 2.5 km respectively. The climate of the Yangtze estuary is monsoonal, with an average temperature of 15.5 °C and an annual precipitation of 1022 mm/year (Hu et al., 2019). The tidal regime in this area is semidiurnal and mixed tides, with mean and maximum tidal ranges of approximately 2.7 m and 5 m, respectively (Yang et al., 2005). Due to the influence of the monsoon climate, the water discharge and sediment load of the Yangtze estuary show seasonal variations. Multi-year observations at the Datong hydrological station demonstrate that water discharge and sediment load during the flood season (May to October) account for more than 70 % and 75 % of the

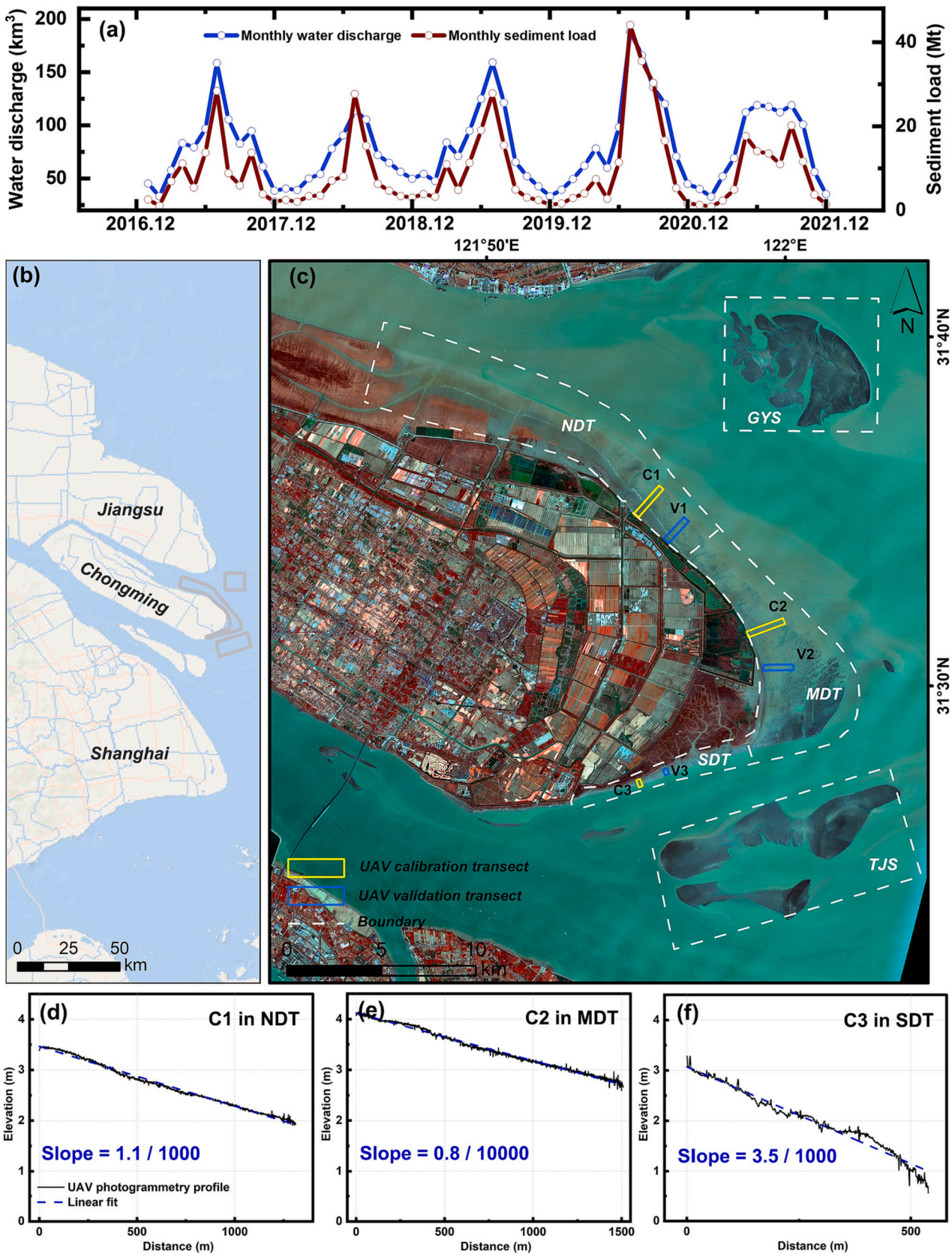

Fig. 1. (a) Monthly water discharge and sediment load from 2017 to 2021 year in Datong station; (b)–(c) Map of the study area for NDT, MDT, SDT, GYS and TJS. The background is a B8, B3 and B2 composite of the Sentinel-2 image taken on 18 January 2021; (d)–(f) The elevation gradient of mudflats acquired by UAV photogrammetry.

year, respectively (Fig. 1(a)). In addition, storm surges and typhoons in the region occur mostly during the flood season. As a result, the mudflat morphology in the Yangtze estuary is relatively unstable during the flood season.

3. Data and methods

3.1. Dataset

The twin Sentinel-2 satellites can acquire high-resolution multispectral imagery with a revisit period of approximately 5 days. In this study, we first filtered the Sentinel-2 Level 2A surface reflectance images in the Google Earth Engine (GEE) platform using a 60 % threshold and masked cloud pixels using the QA60 band. These images were then individually loaded into the GEE interface to check if any masked cloud pixels covered the land–water boundaries, and if so, the images were discarded. Finally, a total of 220 cloud-free or low cloud cover Sentinel-2 images were selected for three periods (Period I: 2019.11–2020.4; Period II: 2020.11–2021.4 and Period III: 2021.11–2022.4) over the study area. The time-series images used for each period were acquired between November and April of the following year. Image acquisition was restricted to the dry season to avoid substantial morphological changes of the estuarine mudflats during the flood season due to high water and sediment discharge and typhoons. Only bands with a spatial resolution of 10 m (i.e., B2, B3, B4 and B8 bands) and 20 m (i.e., B5, B6, B7, B8A, B11 and B12 bands) in the Sentinel-2 imagery were used, and the 20 m bands were resampled to 10 m using nearest neighbour interpolation. In case the number of Sentinel-2 images was insufficient to construct the intertidal topography, Sentinel-1 Synthetic Aperture Radar (SAR) images were used as a backup. Notably, Sentinel-1 SAR images, not affected by cloud cover, were not selected as the primary data source, due to the speckle noise and poor separability of water and mudflats in the estuarine areas (Geng et al., 2016).

Three transects of the DT (i.e., C1, C2, and C3 in Fig. 1(c)) were observed using a DJI RTK-assisted UAV in early May 2020, April 2021, and April 2022. In addition, three other transects (i.e., V1, V2 and V3) were also surveyed by the UAV in April 2022. The width of these transects is approximately 200 m (i.e., the width of 20 Sentinel-2 pixels) and their length is close to the seaward width of the beach at low tide. All UAV aerial surveys were conducted at low tide. After the RTK-assisted UAV acquired images of the above transects according to the planned flight path, digital elevation models (DEMs) of these transects were generated using the structure-from-motion photogrammetric algorithm implemented in the Pix4DMapper software. The accuracy of the UAV-based DEMs was evaluated using ground checkpoints, which yielded a vertical root-mean-square error (RMSE) of 3.1 cm (Chen et al., 2023a). The C1-C3 transects were used to calibrate waterline heights instead of tidal data, and the V1-V3 transects were used to validate the accuracy of the topography constructed from the deep learning-based waterline method.

3.2. Mudflat topography construction with deep learning

Mapping mudflat morphological changes using Sentinel-2 time series and deep learning involves four main steps: (1) image selection and training dataset preparation; (2) DL model building and training; (3) waterline extraction and selection; and (4) topography construction and evaluation. Fig. 2 illustrates the workflow and the platforms used for the corresponding processing.

3.2.1. Image selection and training dataset preparation

DL models trained on large and diverse samples can have better generalization. The diversity of training data in this study is mainly reflected in the variation of water color and tides. Water color is influenced by suspended sediment concentration and tidal waves, which vary from month to month. Therefore, in this study, one Sentinel-2 image was selected from each month of the three periods 2019.11–2020.4, 2020.11–2021.4 and 2021.11–2022.4, for a total of 3 * 6 = 18 images, and one image with highly turbid water bodies, for a total of 19 images used to create the training dataset (Table S1 and S2). These training images covered the entire study area and featured a variety of scenarios, including varying water turbidity, tidal stages over time, and

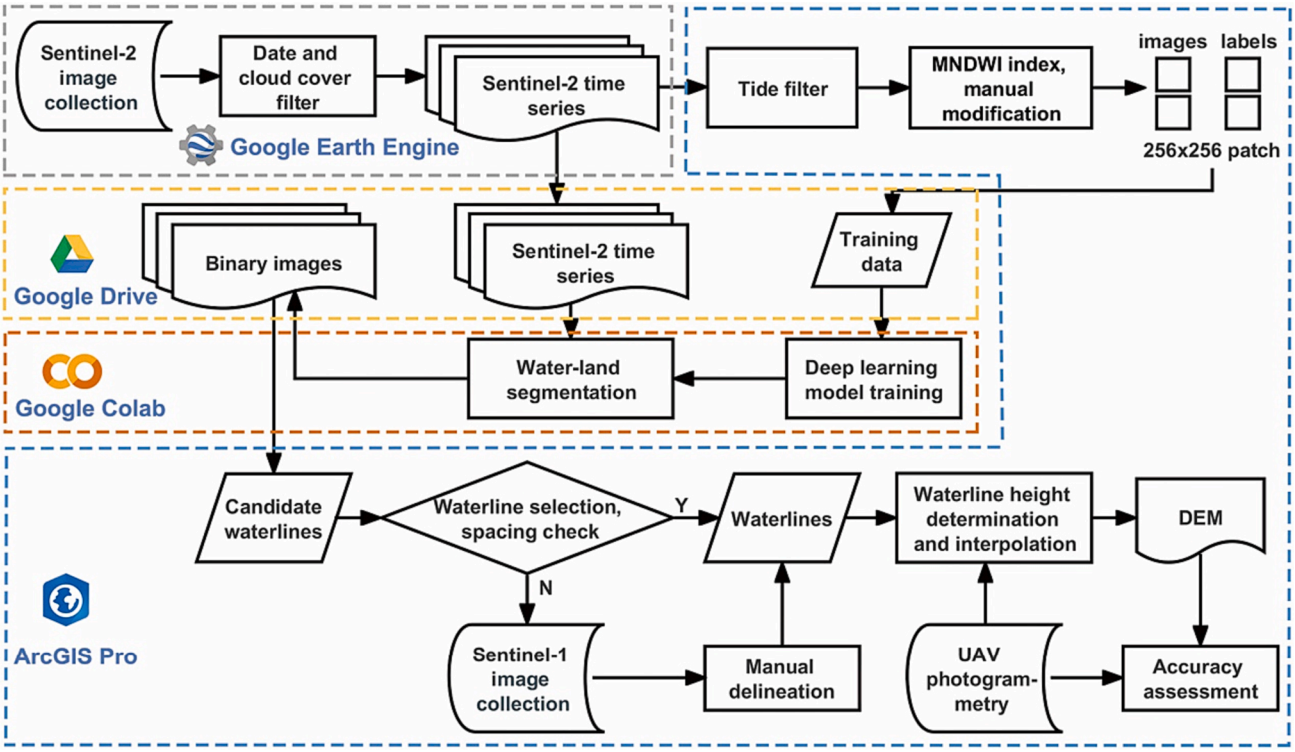

Fig. 2. The workflow for constructing mudflat topography from Sentinel-1/2 time series and UAV photogrammetry with deep learning algorithm.

whitewater caused by wave breaking. To generate the corresponding annotations, we first calculated the Modified Normalized Difference Water Index (MNDWI) (Xu, 2006) from Sentinel-2 imagery, which exploits the optical property that water has strong reflectance in the green band (i.e., B3 band) and high absorption in the shortwave infrared band (i.e., B11 band). Previous tidal flat mapping studies have also shown that MNDWI can provide good initial segmentation of exposed tidal flats after applying a threshold (Tseng et al., 2017; Zhao et al., 2022). We then determined the optimal segmentation threshold through trail-and-error to obtain the initial segmentation results, and manually corrected the misclassified areas, especially the water and mudflat boundaries. Finally, these training images and their corresponding annotations were sliced into patches of 256 × 256 pixels, resulting in a DL dataset with a total of 1112 patches (Fig. 3). This dataset was partitioned 8:2 to generate a training set (889 patches) and a validation set (223 patches). Similarly, a testing dataset of 266 patches was generated from six randomly selected images (Table S3). We did not split the dataset but used different images to generate the training and testing datasets. This is to completely separate the training and testing datasets in order to better assess the generalization ability of the DL model.

3.2.2. *DL model building and training*

(1) **The architecture of DAU-Net model**

U-Net is a fully convolutional neural network that was first introduced in 2015 for biomedical image segmentation tasks (Ronneberger et al., 2015). The architecture of U-Net features an encoder and a decoder. The encoder consists of four blocks, each of which uses 2 convolutional layers and 1 max-pooling layer to extract features from the input image and reduce the spatial dimensions of the feature maps. The decoder includes four blocks, which utilize deconvolutional layers to upsample the feature maps to their original image size. The upsampled features are then concatenated with the corresponding feature maps from the encoder using skip connections, which allows the network to combine high-level semantic features with low-level semantic features. To improve the performance of the classical U-Net architecture and to better match it to our tasks, we made two major adaptations. (1) A parallel dual self-attention mechanism (PDSA) for pixel-wise regression was introduced at the skip connection stage. (2) A hybrid loss function with variable weights combining binary cross-entropy (BCE) and boundary loss was proposed to accurately segment the land–water boundary. The proposed DL model is called dual attention U-Net (DAU-Net) and its architecture is shown in Fig. 4. The DAU-Net used 32, 64, 128, 256 and 512 filters in the encoder and decoder stages. Thus, the difference between the DAU-Net and the standard U-Net is the addition of the PDSA mechanism at the skip connection stage and the construction of the boundary-focused hybrid loss function.

The PDSA is designed for high-quality pixel-wise mapping (Liu et al., 2022). It consists of channel self-attention and spatial self-attention, and its structure is shown in Fig. 4(b), (c). Channel self-attention is used to focus on specific channels of the feature map. Using channel self-attention, the model can selectively weight the most relevant bands and improve the accuracy of the segmentation by exploiting the inter-spectral relationship of feature information (Hang et al., 2020). For the input feature map $\boldsymbol{FM}$, the channel self-attention weight $\boldsymbol{W_C}$ can be expressed as:

$$\boldsymbol{W}_{\mathbf{C}} = Sigmoid[F_C(F_R(F_C(\boldsymbol{FM})) \otimes Softmax(F_R(F_C(\boldsymbol{FM}))))] \quad (1)$$

where F_C, and F_R represent 1 × 1 convolution and tensor reshape operators, respectively;⊗ represents the operation of matrix dot product. Spatial self-attention, on the other hand, is used to focus on specific

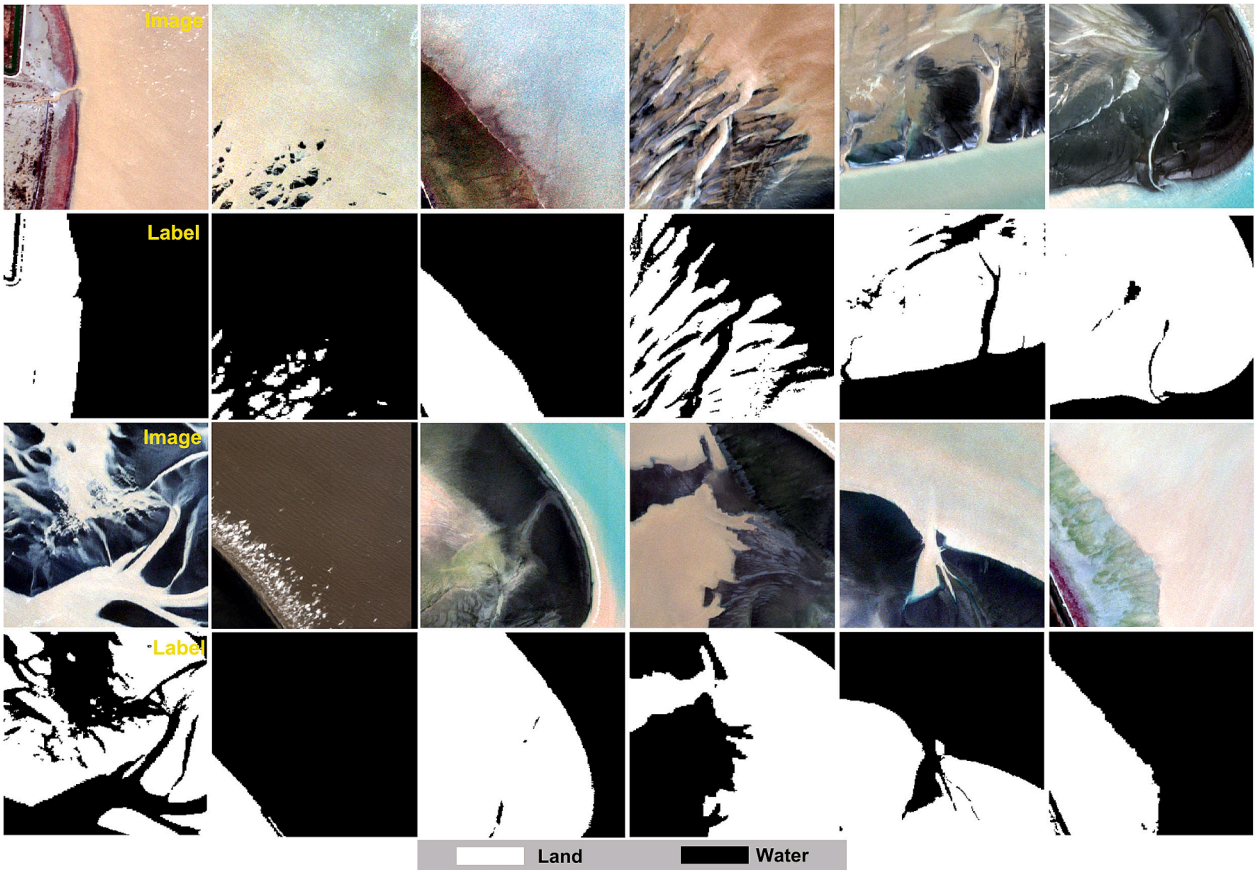

Fig. 3. Examples of Sentinel-2 RGB composite images and corresponding labels. These image patches are 256 × 256 pixels in size.

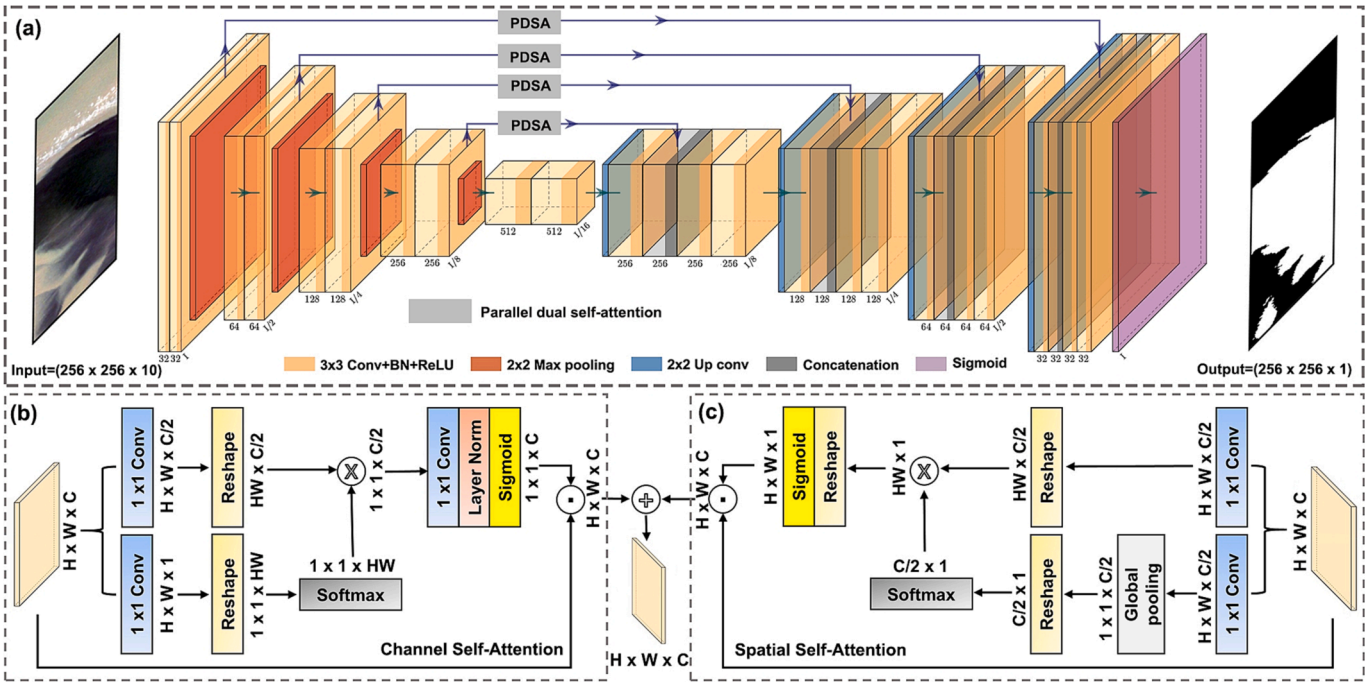

Fig. 4. The structure of the proposed parallel dual self-attention U-Net: (a) the DAU-Net; (b) The channel self-attention module; and (c) The spatial self-attention module.

regions of the input image. Each region is assigned a weight based on its importance for the segmentation task. Given the input feature map ***FM***, the spatial self-attention weight $\boldsymbol{W_S}$ can be expressed as:

$$\boldsymbol{W}_{\mathbf{S}} = Sigmoid[F_R(F_R(F_C(\boldsymbol{FM})) \otimes Softmax(F_R(F_{GP}(F_C(\boldsymbol{FM})))))] \quad (2)$$

where F_{GP} is a global pooling operator. A more discriminative and informative feature representation is produced by combining the channels and spatial self-attention modules in parallel. The output $\boldsymbol{FM}_{\mathbf{PDSA}}$ of the PDSA can be expressed as:

$$\boldsymbol{FM}_{PDSA} = (\boldsymbol{FM} \odot^{C} \boldsymbol{W}_{\mathbf{C}}) \oplus (\boldsymbol{FM} \odot^{S} \boldsymbol{W}_{\mathbf{S}}) \quad (3)$$

where $\odot^{C}$ and $\odot^{S}$ are channel-wise and spatial-wise multiplication operators, respectively; $\oplus$ is the element-wise addition operator.

The BCE loss, which compares the probability of each prediction with the actual category output and then penalizes the probability based on the difference with the expected value, has been widely used in binary classification tasks of remote sensing images (Li et al., 2020; Li et al., 2021; Aghdami-Nia et al., 2022). However, the BCE loss adopts equal weights for every pixel in the same class and does not impose an additional penalty for misclassification of boundary pixels. However, for the task of waterline detection, the focus should be on the accuracy of boundary pixel segmentation. Therefore, we used a boundary loss (Bokhovkin & Burnaev, 2019) to supervise the prediction of boundary pixels. The calculation of the boundary loss is as follows:

$$P = \frac{1}{|B_{pd}|} \sum_{x \in B_{pd}} [\![d(x, B_{gt}) \langle \theta]\!] \quad (4)$$

$$R = \frac{1}{|B_{gt}|} \sum_{x \in B_{gt}} [\![d(x, B_{pd}) < \theta]\!] \quad (5)$$

$$L_{bd} = 1 - \frac{2PR}{P + R} \quad (6)$$

where B_{pd}, B_{gt} represent the boundaries for the predicted binary images and the corresponding ground truth, respectively; d (·) is the Euclidean distance measured in pixels, and θ is a predefined threshold for distance, set to 3 in this study. Thus, to optimize the segmentation performance for DL models, the proposed hybrid loss function used a combination of BCE loss and boundary loss with a weighting parameter α:

$$L_{Hybrid} = \alpha L_{BCE} + (1 - \alpha) L_{bd} \quad (7)$$

where α controls the weight distribution of the two losses. A weight rebalancing strategy was used in model training: the initial value of α was set to 1, and it was self-subtracted by 0.005 at the end of each training epoch until it reached 0. Such a setting could guide the model to gradually optimize the boundary segmentation in the later stages of training to improve the accuracy of waterline detection.

(2) Model implementation and training

The DAU-Net model was implemented using Keras backend with Tensorflow in the Python environment on the Google Colaboratory platform. The training and validation datasets were accessed from Google Drive and the training dataset was randomly shuffled before being fed into the model. To avoid overfitting, the DAU-Net model was preset to train for 1000 epochs with early stopping if the validation loss did not decrease after 15 epochs. The Adam optimizer was adopted to compile the model with an initial learning rate of 1e-4 and the proposed hybrid loss function. The separate BCE loss function was also used in model compilation for comparison. In addition, the learning rate was reduced by a factor of 0.1 when the validation loss plateaued for more than 10 epochs, to avoid an excessive learning rate leading to non-convergence of the loss. At the end of each epoch, the trained model was automatically saved if the validation loss decreased. For comparison, the standard U-Net, and variants of U-Net, U-Net++ and U-Net+++ were trained in the same training configuration. U-Net++ is an extension of the original U-Net proposed by Zhou et al. (2018). It aims to address the limitation of the original U-Net in capturing fine details by introducing nested skip paths in the encoder-decoder architecture. U-Net+++ is another extension proposed by Huang et al. (2020), which further enhances the feature extraction capabilities of the U-Net architecture. It adds more paths in both the encoder and decoder, resulting in a more complex network with multiple levels of nesting. In addition,

Random Forest classification (Murray et al., 2019) and the commonly used Otsu's threshold segmentation based on MNDWI (Jia et al., 2021) were also implemented.

The primary inputs for the U-Net, U-Net++, U-Net+++, and the proposed DAU-Net models were 10-band Sentinel-2 image patches, along with their corresponding labels. Additionally, we investigated the impact of various band combinations, such as utilizing solely the 10 m band from the original Sentinel-2 imagery, and incorporating water indices (i.e., NDWI and MNDWI), on the segmentation performance of the proposed DAU-Net model. A summary of the input data configurations and loss functions employed for training the deep learning models can be found in Table 1. In total, thirteen model training experiments were conducted. All DL models were trained on the Google Colaboratory platform with an NVIDIA Tesla T4 (16 GB) GPU. The source codes and training results developed in this study are openly available at https://github.com/Chunpchen/IntertidalTopoDL.

(3) **Model evaluation metrics**

Precision, Recall, F1-score, Intersection over Union (IoU), and the Matthews correlation coefficient (MCC) are commonly used metrics in binary classification and segmentation tasks because they provide a comprehensive evaluation of model performance. Precision is a metric that measures the accuracy of positive predictions made by a model. Recall, also known as sensitivity, measures the ability of the model to find all positive samples in the dataset. F1-score is the harmonic mean of precision and recall, providing a single metric that balances both precision and recall. IoU is a metric commonly used in segmentation tasks to assess the spatial overlap between the predicted and ground truth regions. MCC is a comprehensive statistical rate that only gives a high score if the prediction scores well in all four confusion matrix categories. They can be calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{F1} - \text{score} = \frac{2 \times Precison \times Recall}{Precison + Recall} \tag{10}$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \tag{11}$$

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{12}$$

where TP, TN, FP, and FN represent the number of true positives (i.e., correctly predicted mudflat pixels), true negatives (i.e., correctly predicted water pixels), false positives (i.e., water pixels predicted to be mudflats), and false negatives at the pixel level (i.e., mudflat pixels predicted to be water), respectively.

3.2.3. Waterline extraction and selection

The filtered Sentinel-2 images from the GEE were exported into Google Drive and then fed into the well-trained DL model for water and land segmentation. Prior to this, each Sentinel-2 image was cropped into a series of 256 × 256 patches and fed into the DL model for patch-by-patch prediction. Once the entire image had been predicted, the outputs of the predicted patches were stitched together and the original georeferencing information was also written using the Python Geospatial Data Abstraction Library (GDAL). A contour tracking algorithm was then used to delineate waterlines from the binary images predicted by the DL model. The contour tracking algorithm, proposed by Suzuki (1985), defines the inner and outer and hierarchical relationships of the contours. As a result, the algorithm allows only the outermost connected waterlines to be extracted, ignoring the edges of the inner hollow water bodies, thus reducing the need for massive post-processing (Fig. 5c). Finally, the extracted contours were converted into vector waterlines.

Because the Sentinel-2 images acquired at different times could correspond to very close tidal levels, there are potentially some redundant waterlines or waterlines that do not cover the entire intertidal range at reasonable intervals. It is therefore necessary to remove redundant waterlines or to additionally supplement the waterlines. Salameh et al. (2020) considered waterlines with a difference in tidal height of less than 15 cm to be redundant. However, the tide stations in the Yangtze estuary are far away from the study area and this metric cannot be used to filter the waterlines. In this study, the transects acquired by UAV photogrammetry accurately recorded the elevation gradient of the mudflats, such that the mudflat width corresponding to the 15 cm elevation value could be calculated according to its slope as a criterion for redundant waterline removal (Fig. 1(d)-(f)). Accordingly, 136 m, 167 m and 39 m were used as discriminating intervals for redundant waterlines in the northern, central and southern Dongtan. In cases where there were insufficient waterlines (i.e., large intervals between waterlines), we manually delineated waterlines from contemporaneous Sentinel-1 imagery to increase the density of the waterlines. To implement this, the waterlines extracted from Sentinel-2 imagery were first uploaded to GEE Assets. The Sentinel-1 images for the corresponding period were filtered through GEE, and then each Sentinel-1 image was loaded into the GEE interface along with the Sentinel-2 derived waterlines. If the waterline in the Sentinel-1 image was found to be positioned between the waterlines extracted from the Sentinel-2 images, the waterline was manually extracted from that image as a supplement.

3.2.4. Topography construction and evaluation

In this study, the waterline heights were calibrated using UAV photogrammetric elevations to generate intertidal topography from multi-temporal waterlines. To achieve this, the spatial resolution of the UAV-based DEMs (i.e., C1, C2, and C3 DEMs) was first resampled to 10

Table 1
An overview of the configurations of input data and loss functions used to train deep learning models.

Model	Input data	Loss function
U-Net	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	BCE
	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	Hybrid
U-Net++	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	BCE
	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	Hybrid
U-Net+++	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	BCE
	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	Hybrid
DAU-Net	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	BCE
	B2, B3, B4, B5, B6, B7, B8, B8A, B11 and B12	Hybrid
DAU-Net	B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12 and NDWI	Hybrid
	B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12 and MNDWI	Hybrid
	B2, B3, B4, and B8	Hybrid
	B2, B3, B4, B8 and NDWI	Hybrid
	B2, B3, B4, B8 and MNDWI	Hybrid

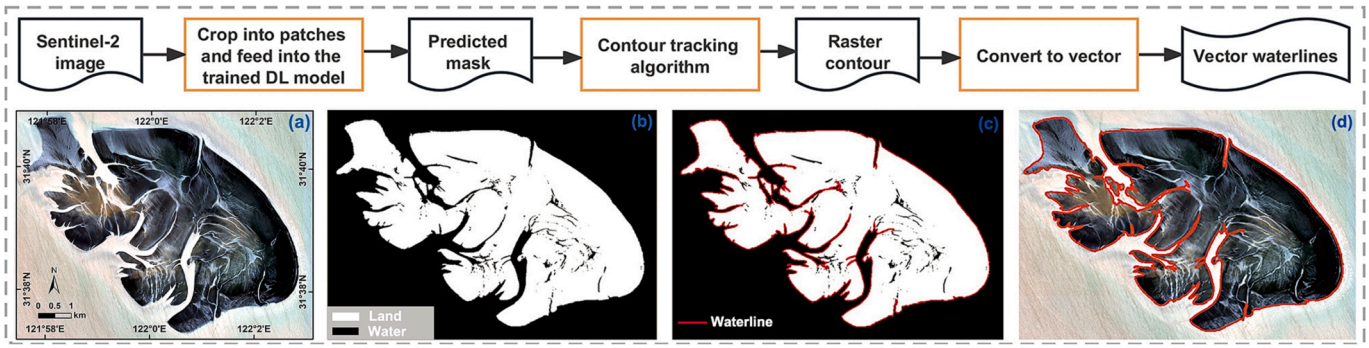

Fig. 5. The procedure for the extraction of the waterlines using the trained DL model and the contour tracking algorithm.

m (in line with Sentinel-2) and then the average elevation of the UAV-based DEM pixels covered by the waterline was calculated as the height of that waterline. The reason for the resampling was that the terrain surface captured by the UAV photogrammetry is very fine and direct overlay can result in poor representativeness of the waterline elevation. Next, the height-assigned waterlines were used to create the triangulated irregular networks (TINs) using the Create TIN tool in ArcGIS Pro. TINs, which are a digital means of representing surface morphology, can vary in resolution depending on the degree of surface variability, resulting in higher resolution in areas with greater variability and lower resolution in areas with less variability, thus accurately reflecting the terrain (Lee, 1991; Gao et al., 2021). Finally, TINs were converted to 5 m resolution raster topographic maps using natural neighbor interpolation. This allowed us to compare the reconstructed terrain with the UAV-based DEMs at a finer spatial scale. Specifically, the topographic accuracy was evaluated by calculating the pixel-by-pixel elevation difference and the RMSE between the three UAV-based transect DEMs (i.e., V1, V2 and V3 DEMs, also resampled to 5 m) and the constructed terrain in 2022.

Topographic changes, also known as DEM of Difference (DoD) maps, are obtained by subtracting the waterline derived DEMs in 2020 from the DEMs in 2022. Due to the presence of topographic reconstruction errors, the DoD map does not necessarily represent the true change in mudflat morphology. To obtain regions of mudflat erosion and deposition under uncertainty, the law of independent error propagation was used to calculate the uncertainty threshold (δ) for the DoD between 2020 and 2022 (Eq. (13) (Wheaton et al., 2010). Absolute elevation differences below the threshold δ are excluded as they are treated as mapping errors rather than actual topographic changes.

$$\delta = \sqrt{(e_{DEM\_2020})^2 + (e_{DEM\_2022})^2} \tag{13}$$

where e_{DEM_2020} and e_{DEM_2022} are the errors of the constructed DEMs in 2020 and 2022, respectively.

4. Results

4.1. Comparison of different DL models for water and land segmentation

Eight trained models using 10-band input were evaluated on the testing dataset of 266 patches. The model trained using the hybrid loss function performed slightly better than the model trained using the binary cross-entropy loss, with each metric 0–2 points higher (Table 2). With the exception of U-Net+++, the other models using the hybrid loss function required fewer epochs and less time to train. The proposed DAU-Net model using hybrid loss achieved the best performance for all metrics except recall, with the highest precision, F1 score, IoU and MCC of 97.4 %, 97.5 %, 95.2 %, and 96.4 % respectively. The comparison results show that the DAU-Net model improves IoU by at least one point compared to the other models. The U-Net and U-Net++ models performed similarly, while the two trained U-Net+++ models performed worse than any other models with equivalent training conditions and spent much more training time. These results show that the addition of

Table 3
Quantitative evaluation of the DAU-Net model trained using different inputs on the testing dataset. The best results for each metric are highlighted in **bold**.

Training			Testing (%)				
Input of DAU-Net	Epoch	Time (h)	Precision	Recall	F1-score	IoU	MCC
10 Bands	390	1.8	**97.4**	**97.6**	**97.5**	**95.2**	**96.4**
10 Bands + NDWI	446	3.5	94.3	97.0	95.6	91.6	92.3
10 Bands + MNDWI	284	2.3	95.8	97.1	96.6	93.3	93.8
B2, B3, B4 and B8	379	1.6	93.1	93.9	93.5	88.5	90.2
B2, B3, B4, B8 and NDWI	282	1.2	90.2	92.9	91.5	85.6	88.9
B2, B3, B4, B8 and MNDWI	424	1.8	94.1	96.8	95.4	91.4	92.9

Table 2
Quantitative assessment of model performance on the testing dataset. The best results for each metric are highlighted in **bold**.

Model	Training			Testing (%)				
	Loss	Epoch	Time(h)	Precision	Recall	F1-score	IoU	MCC
U-Net	BCE	548	1.9	94.2	97.8	96.0	92.3	93.8
U-Net	Hybrid	482	1.9	96.2	97.8	97.0	94.2	94.8
U-Net++	BCE	522	3.8	96.3	96.8	96.6	93.5	94.1
U-Net++	Hybrid	359	2.7	96.9	96.7	96.8	93.9	94.8
U-Net+++	BCE	365	4.6	93.7	**97.9**	95.7	92.0	93.1
U-Net+++	Hybrid	417	5.3	94.4	97.4	95.9	92.2	93.3
DAU-Net	BCE	474	2.1	96.6	97.6	97.1	94.4	95.3
DAU-Net	Hybrid	390	1.8	**97.4**	97.6	**97.5**	**95.2**	**96.4**

the parallel dual attention mechanism to the U-Net model and the use of the hybrid loss function can both improve performance, with the most significant improvements found for the IoU scores. Furthermore, the addition of the attention mechanism and the use of the hybrid loss function did not lead to a dramatic increase in computational time, and instead the DAU-Net model converged faster than the other models.

The evaluation of the DAU-Net models trained with different inputs on the test dataset shows that the model trained with all 10 bands achieved the highest quantitative metrics. Conversely, the model trained solely with four 10 m bands (i.e., B2, B3, B4, and B8) exhibited the poorest performance, showing a significant six-point difference in IoU between the two (Table 3). Furthermore, when incorporating NDWI and MNDWI into the 10-band input, the performance of the trained DAU-Net models deteriorated, with IoU decreasing by 3.6 and 1.9 points, respectively. Moreover, the addition of NDWI to the 4-band training set also resulted in a decline in model performance, while the inclusion of MNDWI improved the model's performance by at least one point for each metric.

We visually compared the results of different DL models and methods for land and water segmentation in three typical water color scenarios. For common water colors (Fig. 6(a)), the coastal segmentation results with different DL models and the Random Forest classifier are very close to the ground truth, while the MNDWI_Otsu method misclassified some mudflat pixels as water bodies. Notably, the proposed DAU-Net is more accurate for coastal segmentation of small isolated islands. In the case of whitewater at the edge of the shore (Fig. 6(b)), our proposed DAU-Net model performed the best and accurately segmented water and land, followed by the standard U-Net, even though the target area in this example was much smaller than the background area. The U-Net, U-Net ++ and U-Net+++ models, the MNDWI_Otsu, and the random forest classifier, however, somehow misclassified the whitewater pixels as land. In extremely turbid estuarine regions (Fig. 6(c)), the coastline boundaries are very blurred, making accurate segmentation a challenging task. A large number of false negatives (i.e., mudflat pixels predicted to be water) were produced by the MNDWI_Otsu and the Random Forest classifier. The U-Net, U-Net++ and U-Net+++ models also produced a small number of false positive and false negative predictions. The DAU-Net model, on balance, was able to detect a greater amount of fine detail consistent with ground truth.

To illustrate the difference in performance between the DAU-Net models trained with the hybrid and cross-entropy loss functions, some randomly drawn examples are shown with the commission and omission errors of the model predictions at the coastal edges. From Fig. 7, the model trained with the hybrid loss function has superior performance on these testing images, where the coastal edges are predicted accurately, with most errors of one pixel and a few errors of more than one pixel. The errors for narrow tidal channels are not informative, as the errors most likely arise from the spatial resolution limitations of Sentinel-2 imagery. Accurate segmentation of water and land minimizes the error in waterline extraction and could therefore reduce the impact of error propagation on the intertidal topography constructed by the waterline method.

4.2. Mudflat DEMs

Waterlines extracted from Sentinel-2 images using the proposed DAU-Net deep learning model, supplemented by a small number of manually delineated waterlines from Sentinel-1 SAR images (Table S4), were used together to create the mudflat DEMs (Fig. S2-S4). Table 4 lists the number of waterlines from Sentinel-1 and Sentinel-2 images acquired for each period for the study sites, the maximum waterline interval, and the elevation difference (Δh) corresponding to the maximum waterline interval estimated from the slope. The maximum waterline interval is used to indicate the worst case for waterline density under the Sentinel 1/2 synergy observations. The Δh corresponding to the maximum waterline interval is less than 0.5 m for the study sites except for MDT. These waterlines were used to construct the annual mudflat topography for the study sites from 2020 to 2022 (Fig. 8). The number of waterlines and Δh used for each period in the same study area are comparable, allowing the constructed topography to be compared at different time periods with the same level of uncertainty to track changes in mudflat morphology. For the NDT, MDT and SDT, the topography maps constructed using the DL-based waterline method accurately record the decreasing elevation from land to sea. For the offshore GYS and TJS, their topography appears high in the middle and low on the sides. The maximum elevation of GYS and TJS is about 3 m, while the maximum tidal level in the Yangtze estuary is over 4 m, which means that both sandbars can be completely submerged by seawater at high tide. We found that both offshore sandbars were not exposed to water in some Sentinel-2 imagery at the time of image selection.

4.3. Accuracy assessment of mudflat DEMs

The elevation differences between the DL waterline-based DEMs and the UAV-based DEMs were calculated (Fig. 9). Positive values indicate that the reconstructed elevations are overestimated, while negative values imply underestimation. The observed elevation differences exhibit a zonal distribution pattern, with similar magnitude differences concentrated in specific areas. A pixel-by-pixel statistical analysis of the errors between DL waterline-based DEMs and UAV-based DEMs is shown in Fig. 9(d). A total of 29492-pixel errors were calculated for comparison. 5.6 % of the elevation difference has an absolute value greater than 0.2 m, and 94.4 % of the elevation difference is distributed in the range of −0.2 to 0.2 m. 77.4 % of the elevation difference is in the range of −0.15 to 0.15 m, and 20.2 % of the elevation difference is in the range of −0.1 to 0.1 m. The RMSE between DL waterline-based DEMs and UAV-based DEMs is 0.13 m.

4.4. Topographic changes in the mudflats

Overall, the mudflats connected to the land were predominantly deposited, with deposition of 10–20 cm over the two years, while the offshore sandbars were unstable and significantly eroded, with erosion of 20–60 cm over the two years, as shown in Fig. 8. By considering the topographic mapping uncertainty, quantified with an RMSE of 13 cm, a threshold of 18 cm was established for the DoD maps between 2020 and 2022. Following this criterion, the topographic changes of the Yangtze estuarine mudflats within the mapping uncertainty between 2020 and 2022 were determined, as presented in Fig. 10. The intertidal mudflats were clearly dominated by deposition at both NDT and MDT, with the greatest deposition occurring in the northernmost part of NDT. The greater deposition at the widest part of the MDT was observed around the tidal channel area. This rapid change may be related to channel migration. Strip erosion was observed in the SDT, which is very close to the salt marsh margins. The morphology of the offshore GYS and TJS was unstable, with alternating erosion and deposition. The DoD maps showed that the erosion hotspots of the GYS and TJS were in the center of the higher topography, while the area around the center was dominated by deposition, showing a sediment transfer pattern.

5. Discussion

5.1. Deep learning for muddy waterline extraction

In this study, a convolutional neural network was used to extract muddy waterlines from Sentinel-2 imagery. The deep learning model with a parallel dual self-attention mechanism and a variable weighted hybrid loss achieved better performance than U-Net, U-Net++ and U-Net+++ models with greater generalization capability. The channel self-attention module in the parallel dual self-attention mechanism can be considered as a spectral self-attention, which can adaptively select bands that are useful for segmenting pixels by enhancing informative

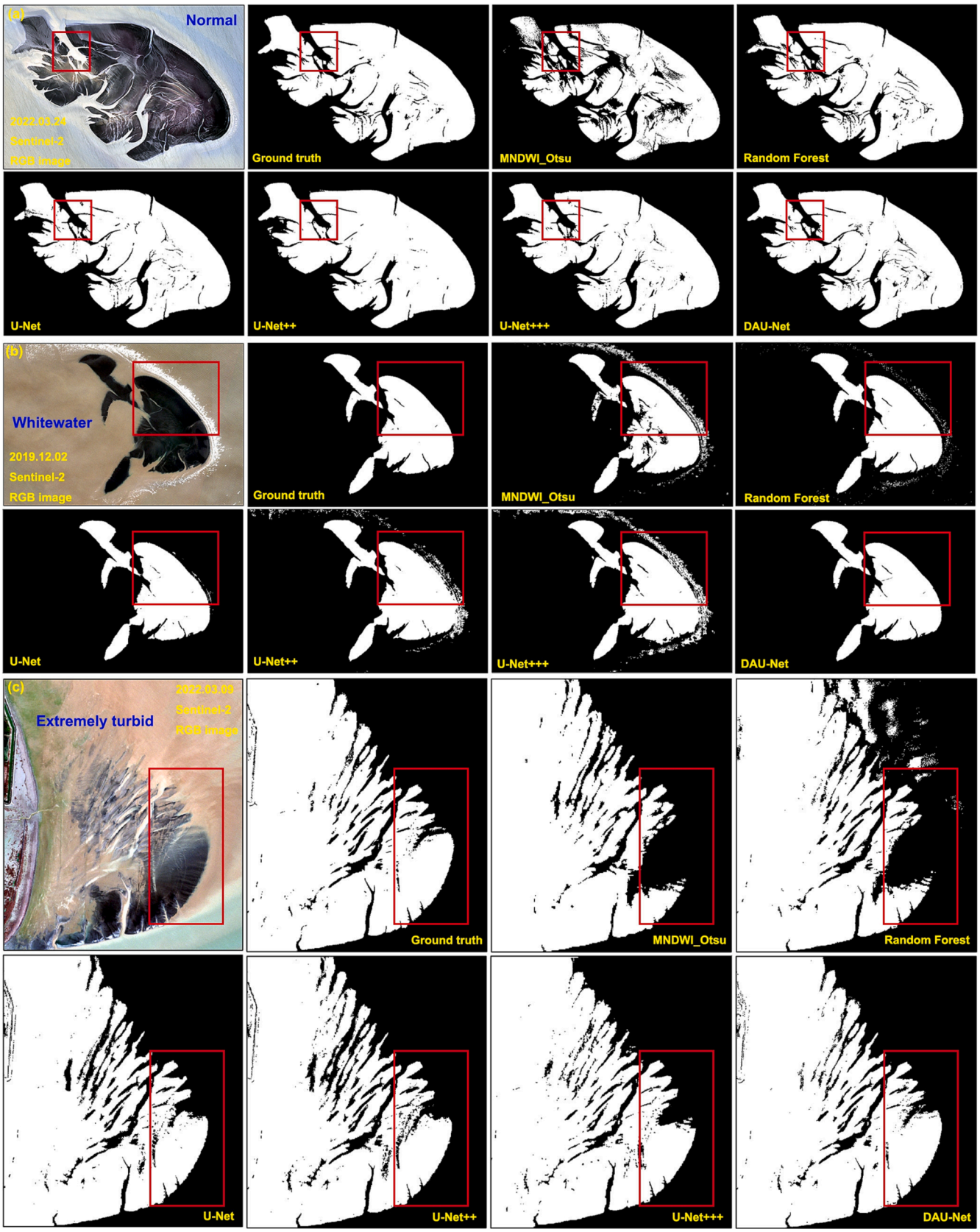

Fig. 6. Comparison of water and land segmentation results from different DL models in different hydrological scenarios.

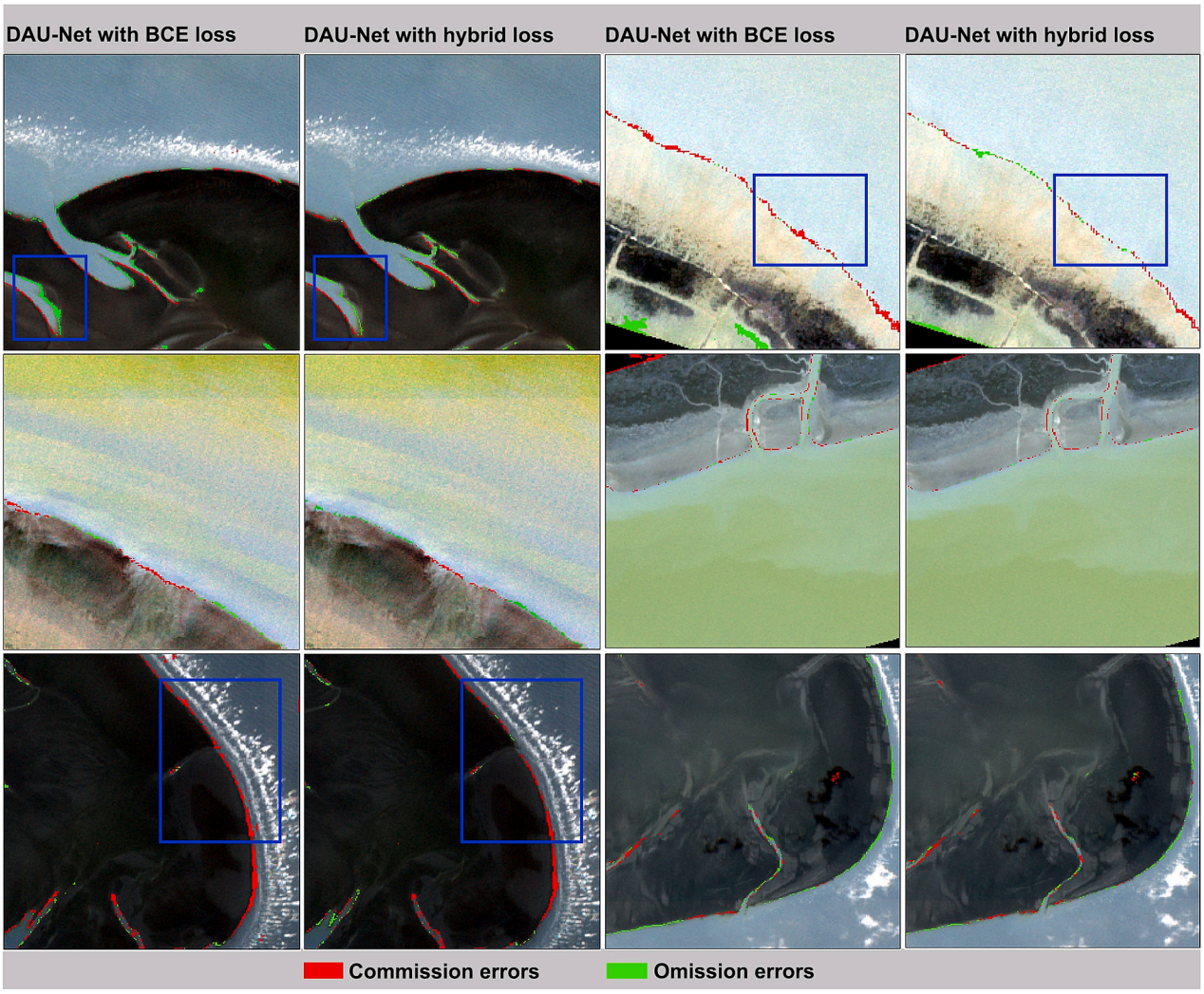

Fig. 7. Comparison of differences in water and land segmentation boundaries for DAU-Net models trained using BCE and the proposed hybrid loss.

Table 4
The number of available waterlines, their maximum spacing and the corresponding maximum height difference (Δh) derived from Sentinel-1 and Sentinel-2.

Region	Period	Number of available waterlines		Maximum waterline interval (m)	Δh(m)
		Sentinel-2	Sentinel-1		
NDT	2019.11–2020.4	11	1	298	0.33
	2020.11–2021.4	10	2	312	0.34
	2021.11–2022.4	10	2	344	0.38
MDT	2019.11–2020.4	10	1	671	0.53
	2020.11–2021.4	9	0	721	0.57
	2021.11–2022.4	10	1	602	0.48
SDT	2019.11–2020.4	7	0	118	0.41
	2020.11–2021.4	7	0	125	0.43
	2021.11–2022.4	7	0	143	0.50
GYS	2019.11–2020.4	13	0	473	0.28
	2020.11–2021.4	13	1	508	0.30
	2021.11–2022.4	12	0	522	0.31
TJS	2019.11–2020.4	11	1	501	0.30
	2020.11–2021.4	11	4	547	0.32
	2021.11–2022.4	8	2	603	0.36

bands and weakening less useful bands (Sun et al., 2019; Zhu et al., 2020). Such a mechanism is important for multispectral remote sensing image segmentation, especially for temporally and spatially varying suspended sediment concentrations and tidally induced scene changes. Our experimental results demonstrate the effectiveness of introducing the attention mechanism, improving the IoU from 92.3 % to 94.4 %. Furthermore, we used a variable-weighted hybrid loss function to guide the model to optimize the accuracy of the boundary segmentation. The boundary loss component of this hybrid function is calculated based on the labels and predicted outputs. This differs from previous approaches that used edge detection operators to calculate edge difference as an auxiliary loss. Such edge detection operators are based on local intensity variation without consideration of semantic context, and waterline detection in changing tidal environments can therefore result in many noisy edges. As a result, DL models with edge operator loss functions have been found to be effective only in structured regions (Cheng et al., 2017; Seale et al., 2022). Our experiments indicate that the introduced hybrid loss function reduces the edge segmentation error and improves the evaluation metrics. This suggests that the DAU-Net model solves the challenge of accurate detection of blurred water-land boundaries in turbid estuarine environments.

Previous studies have shown that training DL models using combinations of bands that make the attributes of objects more obvious, can improve the performance of the model (Konapala et al., 2021; John & Zhang, 2022). Instead, the performance of the trained model decreased

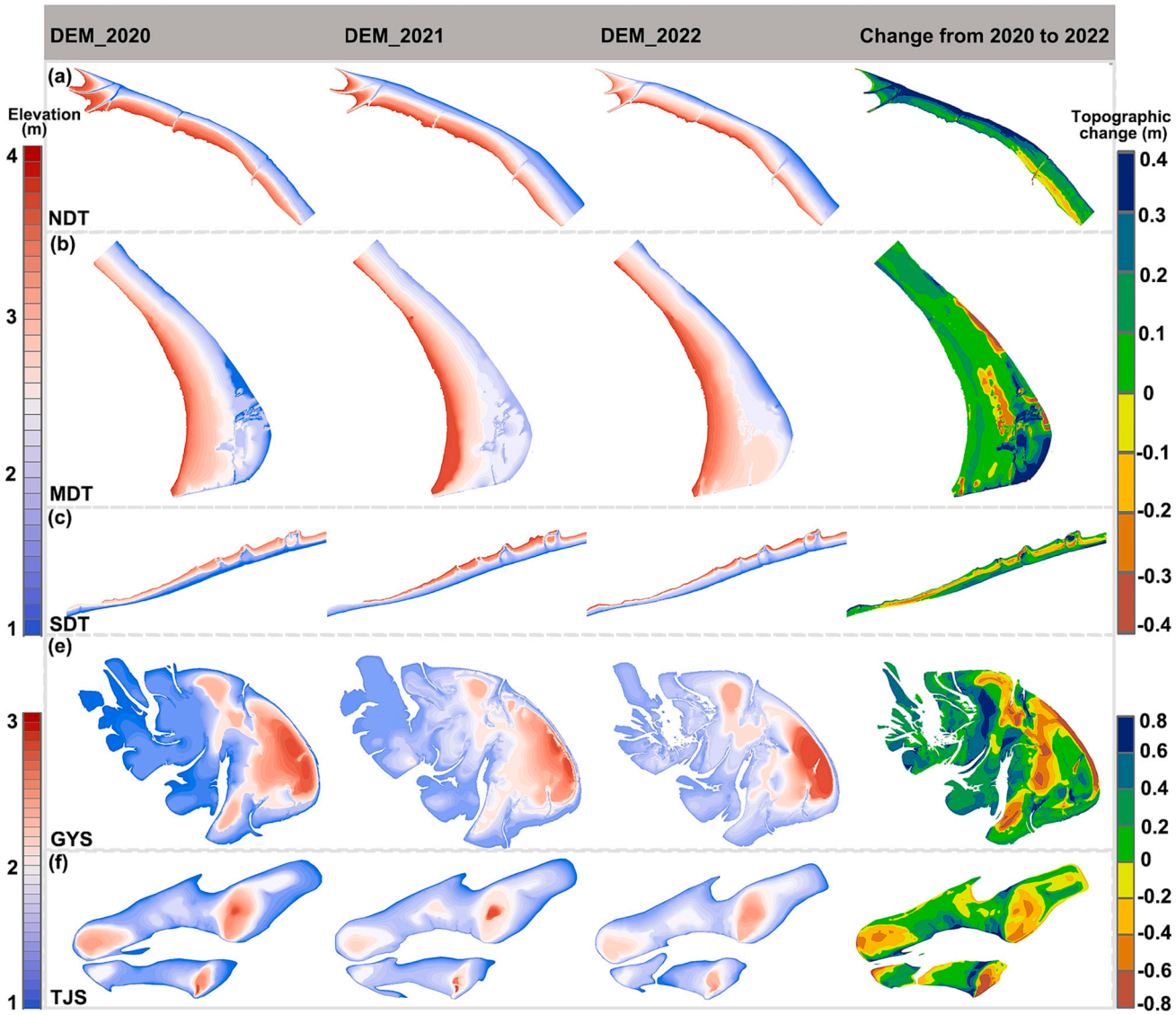

Fig. 8. Topographic maps reconstructed with time-series waterlines and their changes from 2020 to 2022.

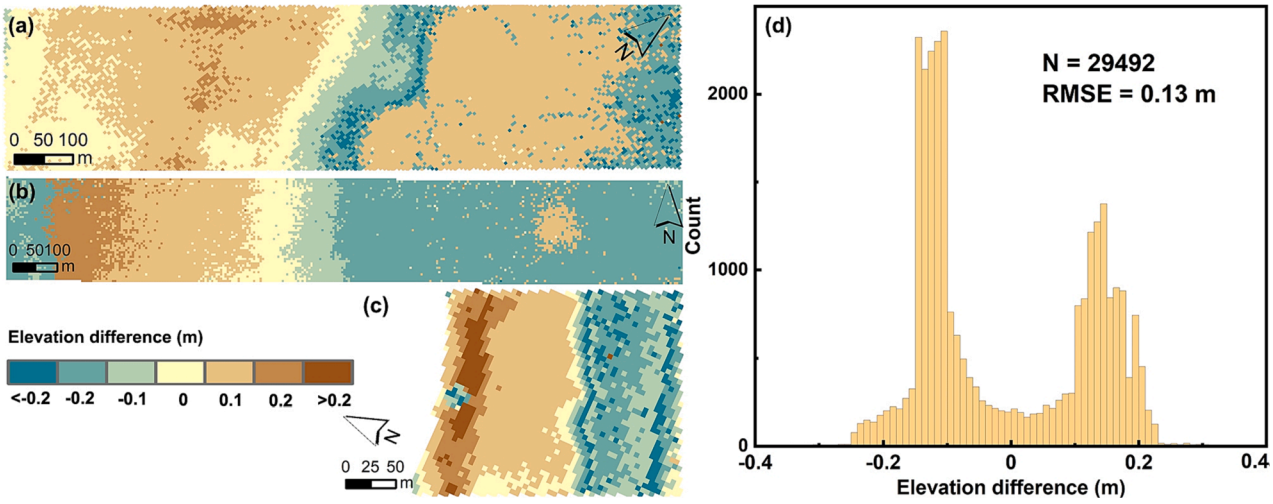

Fig. 9. Elevation difference between satellite-based DEMs and UAV-based DEMs in: (a) V1 transect, (b) V2 transect and (c) V3 transect. (d) The histogram statistics corresponding to the elevation difference.

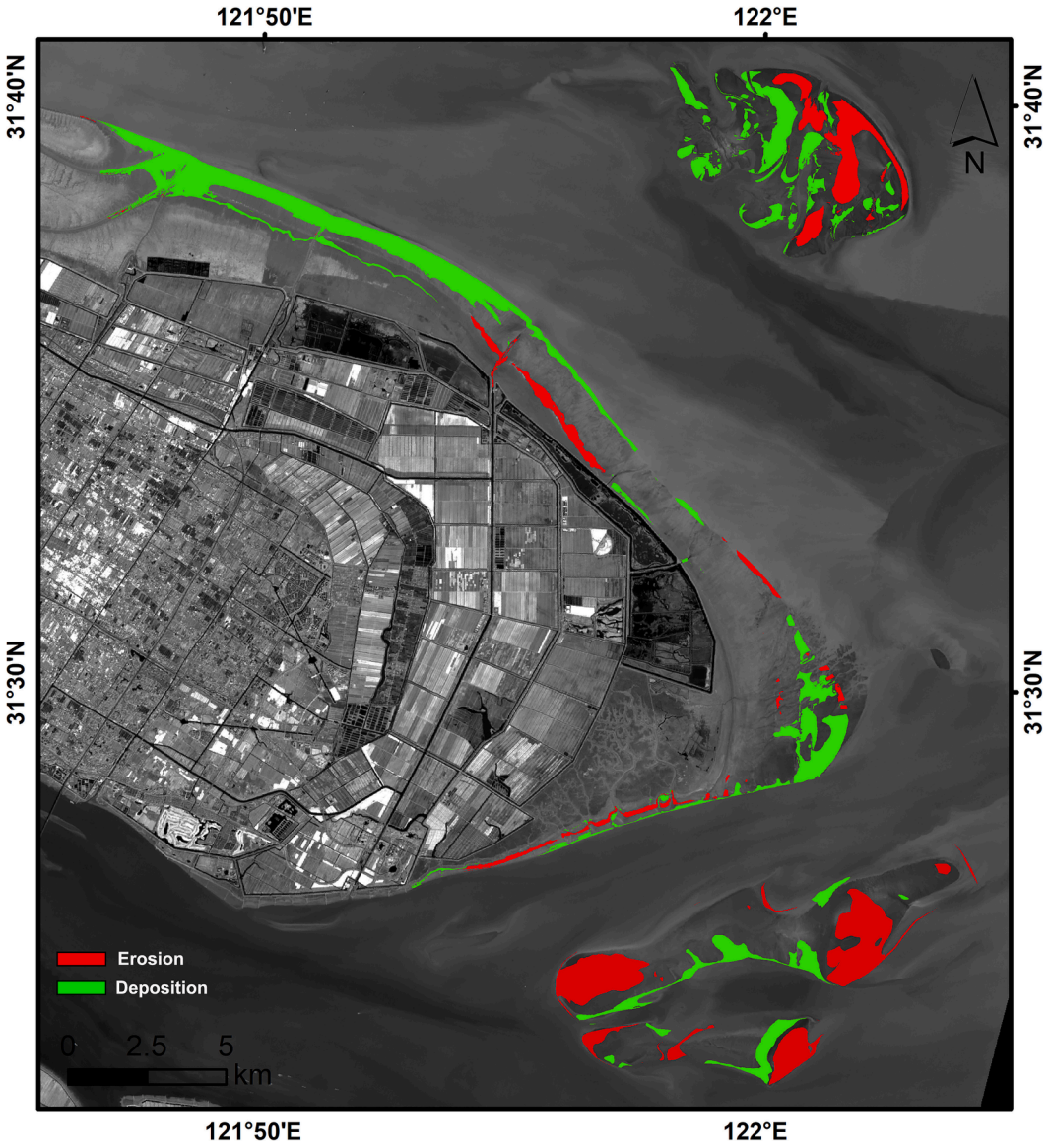

Fig. 10. Mudflat topographic changes in the Yangtze estuary from 2020 to 2022 under mapping uncertainty.

when NDWI and MNDWI, which can highlight water body information, were added to the 10-band data for model training. However, the performance of the model trained on 4-band data was significantly improved with the addition of MNDWI. We selected a Sentinel-2 image to calculate its MNDWI and NDWI, and extracted its waterlines using the trained DAU-Net model, in an attempt to explain why the inclusion of the water indices reduced the accuracy of the model. As shown in Fig. 11, the NDWI values of the turbid water pixels are similar to those of the mudflat pixels, while the MNDWI values of some of the mudflat pixels are even lower than those of the water pixels. This shows the poor separability of turbid water and mudflat pixels in NDWI and MNDWI. Therefore, adding them to the spectrally rich 10-band training dataset introduces more learning uncertainty into the model. However, MNDWI can better detect the segmentation boundary between turbid water and exposed mudflats. Therefore, adding MNDWI to the less spectrally rich 4-band training data could improve model performance.

Deep learning models incorporating attention mechanisms and boundary-focused hybrid loss exhibit superior accuracy in detecting waterlines compared to methods like water index-based thresholding and Random Forest classifiers. However, the process of generating training samples for these models is both labor-intensive and time-consuming. Particularly for large-scale coastal mapping applications, the comprehensive generation of training samples covering diverse coastal types (e.g., muddy, sandy, rocky, etc.), various water color scenarios, tidal conditions, and weather conditions presents considerable challenges. Although some publicly available deep learning training datasets for land-sea segmentation, such as YTU-WaterNet from Landsat-8 imagery (Erdem et al., 2021) and the SWED dataset from Sentinel-2 imagery (Seale et al., 2022), exist, they suffer from limitations in terms of spatial–temporal and semantic scene coverage completeness, as well as the reliability of semantic annotations. Furthermore, the uncertainty in the position of the waterline under time-varying tidal levels complicates the search for high-resolution images with matching imaging moments, essential for enhancing the accuracy and reliability of semantic annotations during training sample generation. Relying solely on visual judgment may introduce uncertainty and adversely impact the deep learning model's performance. To tackle these limitations, some recent research has explored the use of unsupervised deep learning for generating the initial training set or employed generative adversarial networks to expand the training set (Jozdani et al., 2022; Li et al., 2022). However, these techniques are still in their early stages of development in remote sensing sea-land segmentation and require further application

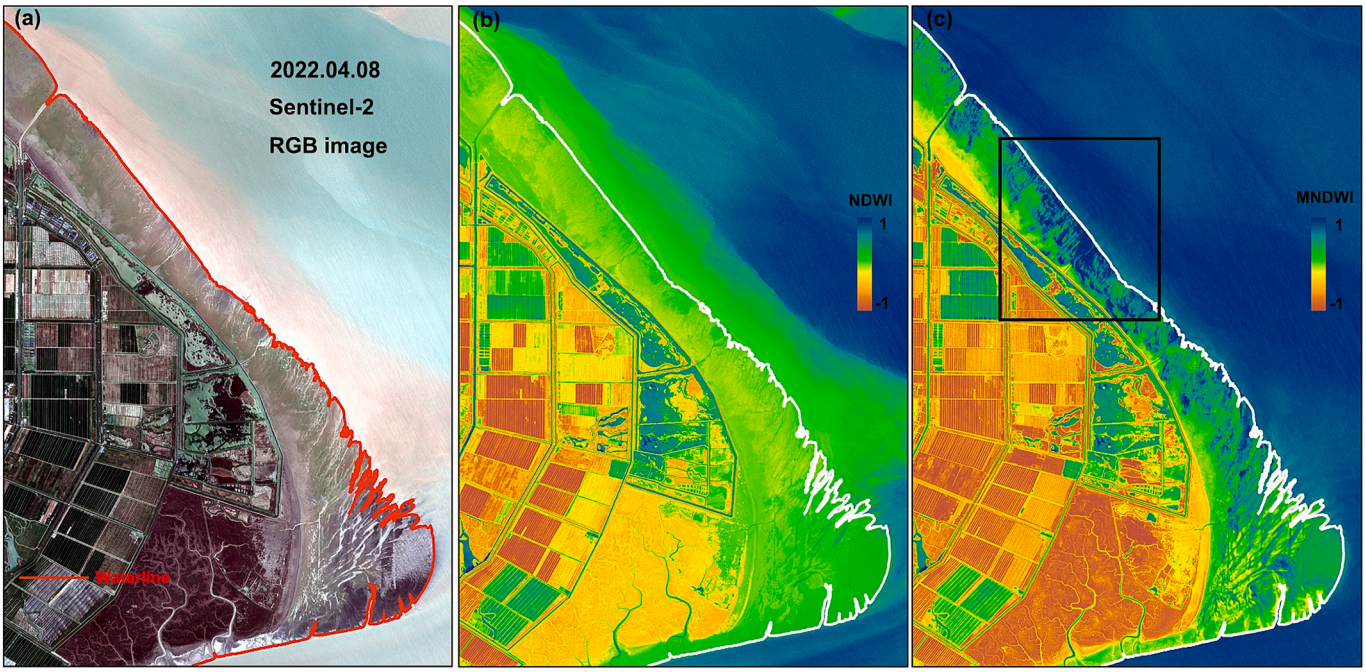

Fig. 11. The poor separability of turbid water and mudflat pixels in NDWI and MNDWI.

and validation in future research endeavors.

5.2. Topographic changes under mapping uncertainty

Uncertainty in the waterline method arises from the parameters associated with the satellite sensor imagery, the waterline extraction technique and the waterline height assignment. A sufficient number of satellite images can provide dense waterlines and the interpolated topography is more detailed. In this study, only Sentinel-2 images acquired during the dry season were used, supplemented by a small amount of Sentinel-1 SAR data. The results show that the waterlines used to construct the topography were dense, with most having a maximum height difference of about 0.3 m. In mudflats with low spatial variability in slope, such height difference introduces less uncertainty after spatial interpolation. Moreover, the short acquisition period, avoidance of extreme water discharge, sediment flux events and frequent summer typhoons, are conditions that are consistent with the assumption that the tidal flat morphology remains stable during the mapping period. Furthermore, the horizontal positional bias due to waterline extraction will lead to errors in the vertical direction. Previous studies have shown that NDWI and MNDWI are the seaward biased and threshold-sensitive indices, with the Otsu thresholding method resulting in an uncertainty of more than three Sentinel-2 pixels (i.e., 30 m) in the shoreline positions (Sagar et al., 2017; Pucino et al., 2022). For mudflats with a slope of 1/1000, such a horizontal error would result in a vertical bias of 3 cm. In comparison, the waterline extraction error is reduced to within one Sentinel-2 pixel using the deep learning method with the proposed hybrid boundary refinement loss. Furthermore, an overestimation of the calibration value of a waterline would lead to an overestimation of the elevation in the corresponding interpolated area, and conversely an underestimation of the calibration value results in an underestimation of the elevation. Consequently, it was observed that errors with the same magnitude exhibited a zonal clustering pattern in the accuracy assessment. However, it is noteworthy that the error associated with the UAV-based DEMs used to assign waterline elevations in this study is merely 3.1 cm, which is notably smaller than the errors in tidal elevations simulated using tidal models as reported in previous studies (Gao et al., 2021; Tsai & Tseng, 2023). As a result of optimizing the waterline method in this study, the constructed topography demonstrates high accuracy, achieving an RMSE of 13 cm.

The areas of mudflat erosion and deposition between 2020 and 2022 were delineated using an uncertainty threshold calculated from error propagation. The results indicate that the Dongtan mudflats were primarily characterized by deposition, with significant deposition observed in the northern region and minor erosion in the southern area. These findings are consistent with previous research by Gao et al. (2021) and Lou et al. (2022), which also reported a dominant deposition trend in the entire Dongtan wetland. Our study further reveals that the mudflats in the northern Dongtan underwent a deposition of approximately 0.4 m during the two-year period from 2020 to 2022, in contrast to the accretion rate of 0.256 m/yr reported by Lou et al. (2022) for the same location. The rapid sedimentation in the northern Dongtan area can be attributed to two main factors. Firstly, it may be linked to the expansion of salt marsh vegetation, particularly the invasive species *Spartina alterniflora*, known for its rapid expansion rate and sediment trapping capability (Zhang et al., 2023). As reported by Liu et al. (2020) *Spartina alterniflora* predominantly occupies the northern Dongtan region. In contrast, other areas of Dongtan had gradually removed *Spartina alterniflora* since the implementation of the ecological restoration project in 2013, and by 2018, it was completely eradicated (Zhang et al., 2020). Secondly, the deposition in the Dongtan mudflats also benefited from sediment recharge originating from offshore sandbars and delta front erosion. This effect becomes more pronounced due to the context of over 70 % reduction in sediment load from the Yangtze River (Yang et al., 2020). These two contributing factors elucidate the variations in topographic changes observed between the northern and southern regions of Dongtan, as well as the offshore sandbars.

6. Conclusions

The accuracy and quantity of the waterline extraction and the accuracy of the associated elevation assignment determine the uncertainty of the reconstruction of the mudflat topography. In this paper, we developed a deep learning model with a parallel self-attention mechanism and boundary-focused hybrid loss to extract turbid estuarine waterlines from dense Sentinel-2 time series. The parallel self-attention mechanism allowed the model to selectively focus on spatial regions and spectral bands that were effective for semantic segmentation, while the

boundary-focused hybrid loss was designed to guide the model to pay more attention to segmentation edges. Compared with U-Net, U-Net ++ and U-Net +++ models, the proposed model exhibits excellent performance in water and land segmentation in time-varying tidal environments. Using the extracted waterlines and UAV photogrammetric surveys, the intertidal mudflat topography of the Yangtze estuary was reconstructed for the years 2020 to 2022. The use of deep learning methods and ground calibration data has maximized the accuracy of the waterline method, allowing more accurate detection of intertidal topographic changes. The comparison of the generated topography with the UAV photogrammetric observations showed an RMSE of 13 cm, an accuracy superior to most existing studies. The reconstructed maps are effective in detecting coastal erosion and deposition hotspots given the uncertainty of the mapping. In the future, with the increasing availability of satellite data, there will be greater potential to detect mudflat topography over even shorter time spans, thereby capturing and analyzing subtle changes more comprehensively.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the project “Coping with Deltas in Transition” within the Programme of Strategic Scientific Alliances between China and the Netherlands (PSA), financed by the Ministry of Science and Technology of the People’s Republic of China (MOST) [grant number 2016YFE0133700], and also sponsored by the China Scholarship Council (CSC).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.isprsjprs.2023.09.022.

References

Aghdami-Nia, M., Shah-Hosseini, R., Rostami, A., Homayouni, S., 2022. Automatic coastline extraction through enhanced sea-land segmentation by modifying Standard U-Net. Int. J. Appl. Earth Obs. 109, 102785.
Andriolo, U., Almeida, L.P., Almar, R., 2018. Coupling terrestrial LiDAR and video imagery to perform 3D intertidal beach topography. Coast. Eng. 140, 232–239.
Arkema, K.K., Guannel, G., Verutes, G., Wood, S.A., Guerry, A., Ruckelshaus, M., Kareiva, P., Lacayo, M., Silver, J.M., 2013. Coastal habitats shield people and property from sea-level rise and storms. Nat. Clim. Chang. 3, 913–918.
Banks, S., Millard, K., Pasher, J., Richardson, M., Wang, H., Duffe, J., 2015. Assessing the potential to operationalize shoreline sensitivity mapping: Classifying multiple wide fine quadrature polarized RADARSAT-2 and Landsat 5 scenes with a single Random Forest model. Remote Sens 7, 13528–13563.
Bell, P.S., Bird, C.O., Plater, A.J., 2016. A temporal waterline approach to mapping intertidal areas using X-band marine radar. Coast. Eng. 107, 84–101.
Bishop-Taylor, R., Sagar, S., Lymburner, L., Alam, I., Sixsmith, J., 2019a. Sub-pixel waterline extraction: characterising accuracy and sensitivity to indices and spectra. Remote Sens 11.
Bishop-Taylor, R., Sagar, S., Lymburner, L., Beaman, R.J., 2019b. Between the tides: Modelling the elevation of Australia’s exposed intertidal zone at continental scale. Estuar Coast Shelf S 223, 115–128.
Bokhovkin, A., Burnaev, E., 2019. Boundary loss for remote sensing imagery semantic segmentation, Advances in Neural Networks–ISNN 2019: 16th International Symposium on Neural Networks, ISNN 2019, Moscow, Russia, July 10–12, 2019, Proceedings, Part II 16. Springer 388–401.
Chen, C., Zhang, C., Schwarz, C., Tian, B., Jiang, W., Wu, W., Garg, R., Garg, P., Aleksandr, C., Mikhail, S., 2022. Mapping three-dimensional morphological characteristics of tidal salt-marsh channels using UAV structure-from-motion photogrammetry. Geomorphology 407, 108235.
Chen, C., Tian, B., Wu, W., Duan, Y., Zhou, Y., Zhang, C., 2023a. UAV photogrammetry in intertidal mudflats: accuracy, efficiency, and potential for integration with satellite imagery. Remote Sens. 15, 1814.
Chen, C., Zhang, C., Tian, B., Wu, W., Zhou, Y., 2023b. Tide2Topo: A new method for mapping intertidal topography accurately in complex estuaries and bays with time-series Sentinel-2 images. ISPRS J. Photogramm. Remote Sens. 200, 55–72.
Cheng, D., Meng, G., Xiang, S., Pan, C., 2017. FusionNet: Edge aware deep convolutional networks for semantic segmentation of remote sensing harbor images. IEEE J. Sel. Topics Appl. Earth. Observ. Remote Sens. 10, 5769–5783.
Dang, K.B., Vu, K.C., Nguyen, H., Nguyen, D.A., Nguyen, T.D.L., Pham, T.P.N., Giang, T. L., Nguyen, H.D., Do, T.H., 2022. Application of deep learning models to detect coastlines and shorelines. J. Environ. Manage. 320, 115732.
de Vries, J., van Maanen, B., Ruessink, G., Verweij, P.A., de Jong, S.M., 2021. Unmixing water and mud: Characterizing diffuse boundaries of subtidal mud banks from individual satellite observations. Int. J. Appl. Earth Obs. 95.
Ellenson, A.N., Simmons, J.A., Wilson, G.W., Hesser, T.J., Splinter, K.D., 2020. Beach state recognition using argus imagery and convolutional neural networks. Remote Sens. 12, 3953.
Erdem, F., Bayram, B., Bakirman, T., Bayrak, O.C., Akpinar, B., 2021. An ensemble deep learning based shoreline segmentation approach (WaterNet) from Landsat 8 OLI images. Adv. Space Res. 67, 964–974.
Fagherazzi, S., Kirwan, M.L., Mudd, S.M., Guntenspergen, G.R., Temmerman, S., D’Alpaos, A., van de Koppel, J., Rybczyk, J.M., Reyes, E., Craft, C., Clough, J., 2012. Numerical models of salt marsh evolution: Ecological, geomorphic, and climatic factors. Rev. Geophys. 50.
Gao, S., 2019. Geomorphology and sedimentology of tidal flats. Coastal Wetlands 359–381.
Gao, W., Shen, F., Tan, K., Zhang, W., Liu, Q., Lam, N.S., Ge, J., 2021. Monitoring terrain elevation of intertidal wetlands by utilising the spatial-temporal fusion of multi-source satellite data: A case study in the Yangtze (Changjiang) Estuary. Geomorphology 383, 107683.
Geng, X., Li, X.-M., Velotto, D., Chen, K.-S., 2016. Study of the polarimetric characteristics of mud flats in an intertidal zone using C-and X-band spaceborne SAR data. Remote Sens. Environ. 176, 56–68.
Hang, R., Li, Z., Liu, Q., Ghamisi, P., Bhattacharyya, S.S., 2020. Hyperspectral image classification with attention-aided CNNs. IEEE Trans. Geosci. Remote Sens. 59, 2281–2293.
Heygster, G., Dannenberg, J., Notholt, J., 2010. Topographic mapping of the German tidal flats analyzing SAR images with the waterline method. IEEE Trans. Geosci. Remote Sens. 48, 1019–1030.
Holman, R.A., Stanley, J., 2007. The history and technical capabilities of Argus. Coast. Eng. 54, 477–491.
Hu, M.-Y., Ge, Z.-M., Li, Y.-L., Li, S.-H., Tan, L.-S., Xie, L.-N., Hu, Z.-J., Zhang, T.-Y., Li, X.-Z., 2019. Do short-term increases in river and sediment discharge determine the dynamics of coastal mudflat and vegetation in the Yangtze Estuary? Estuar Coast Shelf S 220, 176–184.
Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., Wu, J., 2020. Unet 3+: A full-scale connected unet for medical image segmentation. In: ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp. 1055–1059.
Iwamura, T., Possingham, H.P., Chadès, I., Minton, C., Murray, N.J., Rogers, D.I., Treml, E.A., Fuller, R.A., 2013. Migratory connectivity magnifies the consequences of habitat loss from sea-level rise for shorebird populations. Proc. R Soc. B: Biol. Sci. 280, 20130325.
Jain, A., Ramakrishnan, R., Thomaskutty, A., Agrawal, R., Rajawat, A., Solanki, H., 2022. Topography and morphodynamic study of intertidal mudflats along the eastern coast of the Gulf of Khambhat, India using remote sensing techniques. Remote Sens. Appl.: Soc. Environ. 27, 100798.
Jia, M., Wang, Z., Mao, D., Ren, C., Wang, C., Wang, Y., 2021. Rapid, robust, and automated mapping of tidal flats in China using time series Sentinel-2 images and Google Earth Engine. Remote Sens. Environ. 255.
John, D., Zhang, C., 2022. An attention-based U-Net for detecting deforestation within satellite sensor imagery. Int. J. Appl. Earth Obs. 107, 102685.
Jozdani, S., Chen, D., Pouliot, D., Johnson, B.A., 2022. A review and meta-analysis of generative adversarial networks and their applications in remote sensing. Int. J. Appl. Earth Obs. 108, 102734.
Kalacska, M., Chmura, G.L., Lucanus, O., Bérubé, D., Arroyo-Mora, J.P., 2017. Structure from motion will revolutionize analyses of tidal wetland landscapes. Remote Sens. Environ. 199, 14–24.
Khan, M.J.U., Ansary, M.N., Durand, F., Testut, L., Ishaque, M., Calmant, S., Krien, Y., Islam, A.S., Papa, F., 2019. High-resolution intertidal topography from sentinel-2 multi-spectral imagery: synergy between remote sensing and numerical modeling. Remote Sens. 11, 2888.
Konapala, G., Kumar, S.V., Ahmad, S.K., 2021. Exploring Sentinel-1 and Sentinel-2 diversity for flood inundation mapping using deep learning. ISPRS J. Photogramm. Remote Sens. 180, 163–173.
Kulp, S.A., Strauss, B.H., 2019. New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. NatureCommun 10, 1–12.
Lee, J., 1991. Comparison of existing methods for building triangular irregular network, models of terrain from grid digital elevation models. Int. J. Geograph. Informat. Syst. 5, 267–285.
Li, X., Liu, B., Zheng, G., Ren, Y., Zhang, S., Liu, Y., Gao, L., Liu, Y., Zhang, B., Wang, F., 2020. Deep-learning-based information mining from ocean remote-sensing imagery. Natl. Sci. Rev. 7, 1584–1605.
Li, J., Meng, Y., Li, Y., Cui, Q., Yang, X., Tao, C., Wang, Z., Li, L., Zhang, W., 2022. Accurate water extraction using remote sensing imagery based on normalized difference water index and unsupervised deep learning. J. Hydrol. 612, 128202.
Li, M., Wu, P., Wang, B., Park, H., Yang, H., Wu, Y., 2021. A deep learning method of water body extraction from high resolution remote sensing images with multisensors. IEEE J Sel Topics Appl Earth Observ. Remote Sens. 14, 3120–3132.
Liu, Y., Huang, H., Qiu, Z., Fan, J., 2013. Detecting coastline change from satellite images based on beach slope estimation in a tidal flat. Int. J. Appl. Earth Obs. 23, 165–176.

Liu, H., Liu, F., Fan, X., Huang, D., 2022. Polarized self-attention: Towards high-quality pixel-wise mapping. Neurocomputing 506, 158–167.
Liu, Y.-F., Ma, J., Wang, X.-X., Zhong, Q.-Y., Zong, J.-M., Wu, W.-B., Wang, Q., Zhao, B., 2020. Joint effect of Spartina alterniflora invasion and reclamation on the spatial and temporal dynamics of tidal flats in Yangtze River Estuary. Remote Sens. 12, 1725.
Lou, Y., Dai, Z., Long, C., Dong, H., Wei, W., Ge, Z., 2022. Image-based machine learning for monitoring the dynamics of the largest salt marsh in the Yangtze River Delta. J. Hydrol. 608, 127681.
Mason, D., Davenport, I., Robinson, G., Flather, R., McCartney, B., 1995. Construction of an inter-tidal digital elevation model by the 'Water-Line'Method. Geophys. Res. Lett. 22, 3187–3190.
Murray, N.J., Phinn, S.R., DeWitt, M., Ferrari, R., Johnston, R., Lyons, M.B., Clinton, N., Thau, D., Fuller, R.A., 2019. The global distribution and trajectory of tidal flats. Nature 565, 222–225.
Nicholls, R.J., Lincke, D., Hinkel, J., Brown, S., Vafeidis, A.T., Meyssignac, B., Hanson, S. E., Merkens, J.-L., Fang, J., 2021. A global analysis of subsidence, relative sea-level change and coastal flood exposure. Nat. Clim. Chang. 11, 338–342.
Nienhuis, J.H., Ashton, A.D., Edmonds, D.A., Hoitink, A., Kettner, A.J., Rowland, J.C., Törnqvist, T.E., 2020. Global-scale human impact on delta morphology has led to net land area gain. Nature 577, 514–518.
Obida, C.B., Blackburn, G.A., Whyatt, J.D., Semple, K.T., 2019. River network delineation from Sentinel-1 SAR data. Int J Appl Earth Obs 83.
Pucino, N., Kennedy, D.M., Young, M., Ierodiaconou, D., 2022. Assessing the accuracy of Sentinel-2 instantaneous subpixel shorelines using synchronous UAV ground truth surveys. Remote Sens. Environ. 282, 113293.
Ronneberger, O., Fischer, P., Brox, T., 2015. 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*, Springer, pp. 234–241.
Ryu, J.-H., Kim, C.-H., Lee, Y.-K., Won, J.-S., Chun, S.-S., Lee, S., 2008. Detecting the intertidal morphologic change using satellite data. Estuar Coast Shelf S 78, 623–632.
Sagar, S., Roberts, D., Bala, B., Lymburner, L., 2017. Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. Remote Sens. Environ. 195, 153–169.
Salameh, E., Frappart, F., Turki, I., Laignel, B., 2020. Intertidal topography mapping using the waterline method from Sentinel-1 & -2 images: The examples of Arcachon and Veys Bays in France. ISPRS J. Photogramm. Remote Sens. 163, 98–120.
Schoutens, K., Heuner, M., Minden, V., Ostermann, T.S., Silinski, A., Belliard, J.P., Temmerman, S., 2019. How effective are tidal marshes as nature-based shoreline protection throughout seasons? Limnol. Oceanogr. 64, 1750–1762.
Seale, C., Redfern, T., Chatfield, P., Luo, C., Dempsey, K., 2022. Coastline detection in satellite imagery: A deep learning approach on new benchmark data. Remote Sens. Environ. 278, 113044.
Sun, H., Zheng, X., Lu, X., Wu, S., 2019. Spectral–spatial attention network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 58, 3232–3245.
Tian, B., Wu, W.T., Yang, Z.Q., Zhou, Y.X., 2016. Drivers, trends, and potential impacts of long-term coastal reclamation in China from 1985 to 2010. Estuar Coast Shelf S 170, 83–90.
Tong, S.S., Deroin, J.P., Pham, T.L., 2020. An optimal waterline approach for studying tidal flat morphological changes using remote sensing data: A case of the northern coast of Vietnam. Estuar Coast Shelf S, 236.
Tsai, Y.-L.-S., Tseng, K.-H., 2023. Monitoring multidecadal coastline change and reconstructing tidal flat topography. Int. J. Appl. Earth Obs. 118, 103260.
Tseng, K.-H., Kuo, C.-Y., Lin, T.-H., Huang, Z.-C., Lin, Y.-C., Liao, W.-H., Chen, C.-F., 2017. Reconstruction of time-varying tidal flat topography using optical remote sensing imageries. ISPRS J. Photogramm. Remote Sens. 131, 92–103.
Wang, X., Kuang, F., Tan, K., Ma, Z., 2018. Population trends, threats, and conservation recommendations for waterbirds in China. Avian Res. 9, 1–13.
Wang, Y., Liu, Y., Jin, S., Sun, C., Wei, X., 2019. Evolution of the topography of tidal flats and sandbanks along the Jiangsu coast from 1973 to 2016 observed from satellites. ISPRS J. Photogramm. Remote Sens. 150, 27–43.
Wheaton, J.M., Brasington, J., Darby, S.E., Sear, D.A., 2010. Accounting for uncertainty in DEMs from repeat topographic surveys: improved sediment budgets. Earth Surface Process. Landforms: J. British Geomorphol. Res. Group 35, 136–156.
Wiehle, S., Lehner, S., 2015. Automated waterline detection in the Wadden Sea using high-resolution TerraSAR-X images. *J. Sens.* 2015.
Xie, W., He, Q., Zhang, K., Guo, L., Wang, X., Shen, J., Cui, Z., 2017. Application of terrestrial laser scanner on tidal flat morphology at a typhoon event timescale. Geomorphology 292, 47–58.
Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. Int. J. Remote Sens. 27, 3025–3033.
Yang, S.L., Zhang, J., Zhu, J., Smith, J.P., Dai, S.B., Gao, A., Li, P., 2005. Impact of dams on Yangtze River sediment supply to the sea and delta intertidal wetland response. J. Geophys. Res. Earth Surf. 110.
Yang, S.L., Luo, X., Temmerman, S., Kirwan, M., Bouma, T., Xu, K., Zhang, S., Fan, J., Shi, B., Yang, H., 2020. Role of delta-front erosion in sustaining salt marshes under sea-level rise and fluvial sediment decline. Limnol. Oceanogr. 65, 1990–2009.
Yang, Z., Wang, L., Sun, W., Xu, W., Tian, B., Zhou, Y., Yang, G., Chen, C., 2022. A new adaptive remote sensing extraction algorithm for complex muddy coast waterline. Remote Sens 14, 861.
Zhang, S., Xu, Q., Wang, H., Kang, Y., Li, X., 2022. Automatic Waterline Extraction and Topographic Mapping of Tidal Flats From SAR Images Based on Deep Learning. *Geophys. Res. Lett.* 49, e2021GL096007.
Zhang, M., Schwarz, C., Lin, W., Naing, H., Cai, H., Zhu, Z., 2023. A new perspective on the impacts of Spartina alterniflora invasion on Chinese wetlands in the context of climate change: A case study of the Jiuduansha Shoals. Yangtze Estuary. Sci Total Environ 868, 161477.
Zhang, X., Xiao, X., Wang, X., Xu, X., Chen, B., Wang, J., Ma, J., Zhao, B., Li, B., 2020. Quantifying expansion and removal of Spartina alterniflora on Chongming island, China, using time series Landsat images during 1995–2018. Remote Sens. Environ. 247, 111916.
Zhao, B., Liu, Y., Wang, L., Liu, Y., Sun, C., Fagherazzi, S., 2022. Stability evaluation of tidal flats based on time-series satellite images: A case study of the Jiangsu central coast. China. Estuar Coast Shelf S 264, 107697.
Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. Unet++: A nested u-net architecture for medical image segmentation. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings* 4. Springer, pp. 3-11.
Zhu, M., Jiao, L., Liu, F., Yang, S., Wang, J., 2020. Residual spectral–spatial attention network for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 59, 449–462.