Journal of Coastal Research	SI	113	710-714	Charlotte, North Carolina	2024

Automated Mapping of Surface Sedimentary Features in Mixed Sand-Gravel Tidal Inlets Using UAV, XGBoost and U-Net

Jie Gong* **and Helene Burningham**

Department of Geography
University College London
London, U.K.

www.cerf-jcr.org

www.JCRonline.org

ABSTRACT

Gong, J. and Burningham, H., 2024. Automated mapping of surface sedimentary features in mixed sand-gravel tidal inlet using UAV, XGBoost and U-Net. *In:* Phillips, M.R.; Al-Naemi, S., and Duarte, C.M. (eds.), *Coastlines under Global Change: Proceedings from the International Coastal Symposium (ICS) 2024 (Doha, Qatar)*. *Journal of Coastal Research*, Special Issue No. 113, pp. 710–714. Charlotte (North Carolina), ISSN 0749-0208.

Understanding the distribution of surface sedimentary features within tidal inlets is crucial for assessing their morphodynamic response to waves and tides. Sediment size significantly influences sediment transport and deposition dynamics. However, in mixed sand-gravel tidal inlets, classifying sediment distribution poses unique challenges. Previous research has relied on labour-intensive field sampling, while the rapid spatial changes in the surface features challenge traditional surveying methods and reduce the potential for frequent monitoring. Although satellite imagery offers regular observations, low resolutions hinder the accurate classification of detailed surface characteristics. This study integrates consumer-grade UAV technology with XGboost and U-Net (ResNet34) model to develop automated high-resolution mapping models for surface features in a mixed sand-gravel tidal inlet at the mouth of the Deben estuary, based on the RGB images. The results show that both XGBoost and U-Net have good performance and high potential to classify surface sediments and map these at the pixel level in mixed sand-gravel systems, with relatively high accuracy in the prediction of gravel, sand and vegetation cover. These combined methods demonstrate the potential for regular UAV monitoring of tidal inlets over short- and long-term scales, which can enhance our morphodynamic understanding and contribute to the coastal monitoring and management.

ADDITIONAL INDEX WORDS: *Tidal inlet, sedimentary features classification, UAV, XGBoost, U-Net.*

INTRODUCTION

The morphodynamics of tidal inlets and their physical processes (waves, tides, river flow) and sedimentary characteristics have been largely studied through in-situ surveys (Burningham and French, 2007), physical experiments (Sutherland and Soulsby, 2010), numerical modelling (Georgiou *et al.*, 2023) and statistical analysis (Gong and Burningham, 2024). The interaction between rivers, tides, and waves is the primary driver of morphological change in inlets (Boothroyd, 1985), whilst the nature, availability and supply of sediment also influence morphodynamic behaviour and in particular the configuration of sediment bodies (FitzGerald *et al.*, 2002).

Coastal behaviour can be significantly controlled by sedimentology, and this is best evidenced in beach systems where fine sediment supports dissipative morphodynamics and coarse sediment encourages reflective conditions. In mixed sand-gravel inlets, the sedimentological role is thought to be important in controlling aspects of morphology and sediment transport fluxes that could also influence sediment bypassing and morphodynamic behaviour (Burningham and French, 2006). Therefore, regular monitoring of surface sediment dynamics and mapping sediment types and their distribution patterns are crucial for better morphodynamic understanding.

The significant challenges of mapping these surface types are their spatial heterogeneity and temporal variability. The traditional mapping techniques for surface cover usually depend on ground-based sampling (Son *et al.*, 2011), which is time-consuming and labour-intensive and difficult to apply with high frequency and accuracy to map large areas. Remote sensing has been used in land cover classification and object detection (Dong *et al.*, 2018). Its applications at the coast include detecting shorelines and the generation of satellite-derived digital elevation models (Burvingt *et al.*, 2022; Caballero and Stumpf, 2021). However, the relative resolution of satellite images is limited in detecting small-scale features such as cusp and sediment patches, whilst cloud cover and timing relative to suitable tides are also issues.

In the recent decade, near sensing using unpiloted aerial vehicles (UAV) or drones has been increasingly employed in geomorphological studies (Casella *et al.*, 2016). UAVs reduce 80% of ground data collection time and acquire high-resolution orthophotos at low altitudes (Casella *et al.*, 2016; James and Robson, 2012). Additionally, the relatively low cost makes UAVs suitable for repeated surveys to obtain multi-temporal observations at the small-scale over short time intervals.

DOI: 10.2112/JCR-SI113-140.1 received 23 June 2024; accepted in revision 7 August 2024.
*Corresponding author: jie.gong.19@ucl.ac.uk

Another new wave in geomorphological technologies is the application of deep learning to environmental data. For classification tasks like land cover and land use, the tree-based extreme gradient boosting (XGBoost) and convolutional neural networks (CNNs) model U-Net are two algorithms widely used to achieve pixel-level prediction. The UAV-XGBoost combined method has been explored in urban land cover classification (Dong *et al.*, 2018; Sefercik *et al.*, 2023), while it has not been largely tested in coastal sediment analysis. The UAV-U-Net method has been applied to surface sediment classification in a mixed mud-sand tidal flat with the highest averaged F1-score of 0.66 (Kim et al., 2024). Performance can be improved by using pre-trained U-Net models such as the geospatial semantic segmentation U-Net architecture with the pre-trained encoder residual network (ResNet34) used in land cover classification with an averaged F1-score of 0.73 (Sewada and Goyal, 2023). Therefore, this study aims to evaluate the potential of XGBoost and a pre-trained U-Net model with ResNet34 architecture in classifying sediment surface features with high-resolution UAV RGB images as the crucial basis for the automated mapping of mixed sand-gravel coastal systems.

PHYSICAL SETTING

The Deben inlet is a mixed sand-gravel tide-dominated meso-tidal inlet located on the south Suffolk coast, east England (Figure 1). The tidal length of the Deben estuary is around 18 km, the mean spring tidal range is approximately 3.2 m, and the average spring tidal prism is about 12×10^6 m^3 (Burningham and French, 2006). The average offshore wave height is about 0.97 m, but the wave direction is bimodal, predominantly from the NE and SW, which is considered to play a vital role in the ebb delta morphodynamics (Gong and Burningham, 2024). The narrow inlet throat (~180 m) limits wave propagation into the inner estuary, and the low river discharge provides negligible fluvial sediment input. The inlet comprises of a small flood-tidal delta in the immediate inner estuary and a relatively large ebb-tidal delta including an updrift complex (Bawdsey foreland) and shoals, a big middle shoal, and a downdrift complex (Felixstowe Ferry foreland). The primary littoral drift direction is from north (updrift side) to south (downdrift side) via the ebb-tidal delta sediment bypassing process (ebb shoals). Based on the laboratory analysis of inlet surface sediment samples, 82% of sediment is gravel, mostly in the range of 8 - 32 mm diameter, and 17% is sand with a mean grain size of around 1 mm.

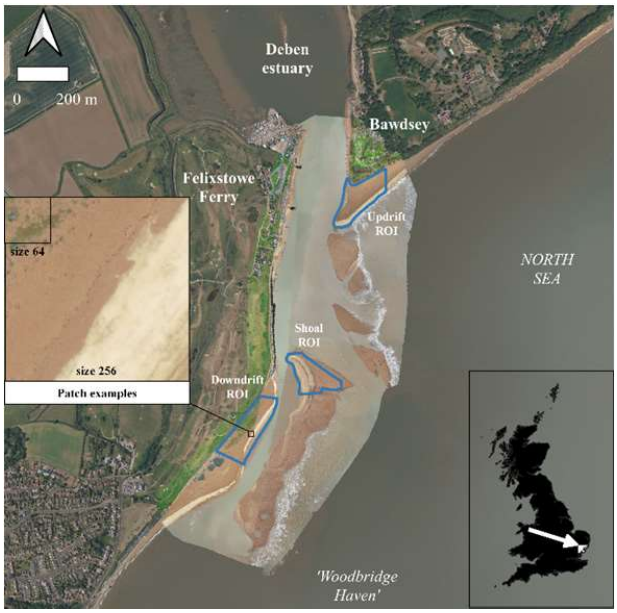

Figure 1. Location map of Deben inlet. Regions of interest (ROI) are marked by the blue polygons. The zoom-in training patch examples are selected from the downdrift coast with size 64 × 64 pixels for XGBoost model and 256 × 256 pixels for U-Net model.

METHODS

UAV Datasets

Low tide UAV images were acquired by using a DJI Phantom 4 RTK on 20th April 2023, which covered most of the ebb-tidal delta features in Deben inlet. To focus on the spatial variability both vertically and horizontally, three regions of interest (ROI) separately located on the updrift, mid-shoal, and downdrift sedimentary complexes were used to extract subset from the large orthophotos for further data preprocessing of deep learning models training. The orthorectified images with a resolution of 0.1 m were generated from Agisoft Photoscan software based on the structure from motion algorithm (Casella *et al.*, 2016). In addition to the RGB colour bands, the visual atmospheric resistance index (VARI) and green leaf index (GLI) were calculated, which are normally used to estimate the plant fraction above the soil. Due to vegetation growth on beach ridges here, especially on the downdrift coast, both or one of them was used in the K-means labelling process and model training. The equations of VARI (Costa *et al.*, 2020) and GLI (Barbosa *et al.*, 2019) are:

$$VARI = (G - R)/(G + R - B) \quad (1)$$

$$GLI = (2G - R - B)/(2G + R + B) \quad (2)$$

Pixel Labelling

Based on field experience and assessment of the UAV images, there are six surface types distributed across the study area – gravel, sand, mixed, vegetation 1 (high density), vegetation 2 (low density or other colour vegetations), and clay. In the updrift ROI, the major classes are gravel, sand and mixed, which is similar to the shoal ROI but that has one extra class clay. Except for clay, all classes can be found in the downdrift ROI. Labelling high-resolution images at the pixel level is a challenging manual operation, due to the mixed classes on the surface. Therefore, K-means was employed to establish the initial pixel-level labelling. Due to the different surface contexts of each ROI, the different combinations of feature input were tested. The satisfied labels in updrift and shoal ROIs were clustered via the features of VARI and GLI, while RGB and VARI were used for downdrift ROI clustering. Some pixel labels in the final ground truth maps were corrected manually in QGIS.

Deep Learning Models

XGBoost is a tree-based supervised ensemble learning algorithm relying on a generalized gradient-boosting method that employs a regularization term to generate accurate models with distributed settings and multicore for regression, ranking and classification tasks (Chen and Guestrin, 2016). An XGBoost model was trained in this study at the pixel level. Another algorithm, U-Net was also trained here and compared to the XGBoost model. Due to the spatial correlation between pixels, the extension of CNN was used to extract spatial features, and the U-Net architecture of the neural network, a fully convolutional network extension that is encoder-decoder-based, was employed in this study (Sewada and Goyal, 2023). The pre-trained encoder, ResNet34, was used as the encoder-decoder in the background for the classification and mapping of surface types, which utilizes residuals from the layers in the connected layers and skips the connections between the encoder and its paired decoder blocks (He *et al.*, 2016).

Pre- and Post-training

The workflow of the whole process is shown in Figure 2. The input features RGB and VARI and the target labels were prepared for feeding the XGBoost and U-Net model. To reduce the influence of spatial correlation in model validation on model performance evaluation, patch processing of large datasets is applied in the data preprocessing of XGBoost (Table 1). Specifically, the patch size of 64 was used to crop the large datasets to reduce no-data pixel samples in each patch and increase the spatial randomness in the subsequent step of splitting training and validation datasets. First, patches with >40% no-data pixels were removed; each remaining patch's dominated class (the pixel number of one class >50%) was calculated to group the patches. Otherwise, the no-class-dominated patches were put into one group. Gravel occupies most patches (>1000) while clay only dominates three patches. Two clay patches were selected for training, and the remaining one used as a validation patch. For other groups, 50 patches were randomly chosen as training patches, and others were used in validation.

Table 1. *Summary of the data preprocessing of XGBoost model.*

Class	Patch number (> 50%)	Train patch number	Train pixel number	Validation Patch number
No data	-	-	5000	-
Gravel	1084	50	50000	1034
Sand	208	50	50000	158
Mixed	85	50	50000	35
Vegetation 1	55	50	50000	5
Vegetation 2	68	50	50000	18
Clay	3	2	6424	1
No-dominated	101	50	-	51
Total	1604	302	261424	1302

To reduce dataset redundancy and training time, 50,000 pixel samples were further selected from each class (gravel, sand, mixed, vegetation 1, vegetation 2) in all training patches. All 6424 training clay pixels were kept, and only 5000 no data samples were selected. In total, over 260,000 pixel samples entered XGBoost model training and hyperparameters tuning.

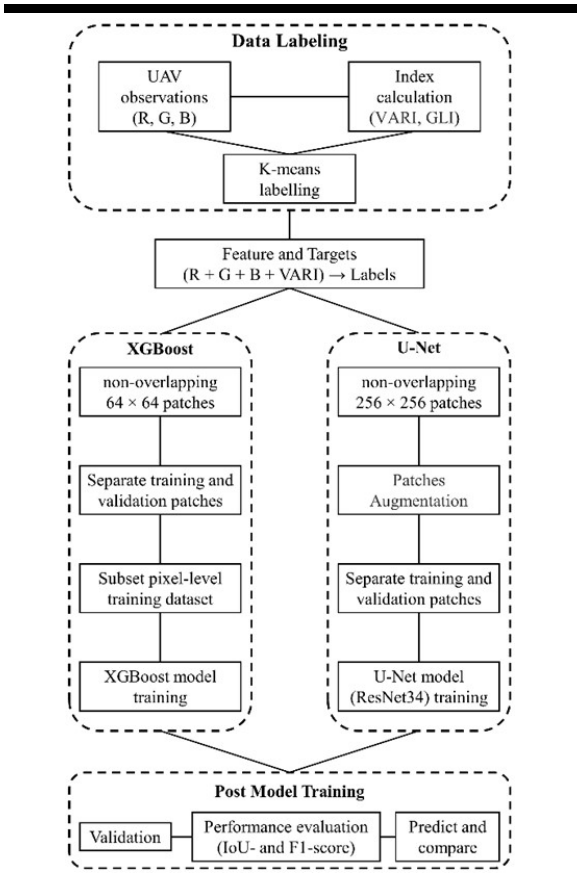

Figure 2. Workflow of pre- and post-training XGBoost and U-Net model.

Regarding the data preprocessing of U-Net model training, a larger patch size (256) was applied to keep the spatial diversity of classes, and patches were removed if the number of no data pixels exceeded 15%. The RGB examples of patch sizes 64 and 256 can be seen in Figure 1. There were 77 non-overlapping patches cropped from the original large datasets, and the total patch number was raised to 616 via data augmentation of rotation (90°, 180°, and 270°) and flipping. The processed patches were divided into 492 training (80%) and 124 validation (20%) patches. In addition, the model performances were evaluated with Intersection over Union (IoU) and F1-score. F1-score is the harmonic mean of precision and recall, whilst the IoU score is the intersection of the predictions and true label divided by the union of selected features (Kim *et al.*, 2024; Sewada and Goyal, 2023). The equations are:

$$\mathrm{F1} = \frac{2\times\mathrm{precision}\times\mathrm{recall}}{\mathrm{precision}+\mathrm{recall}} \quad (3)$$

$$\mathrm{IoU} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \quad (4)$$

Where TP is the true positives, FP is the false positives and FN is the false negatives. The precision = TP / (TP + FP) and recall = TP / (TP + FN).

RESULTS AND DISCUSSION

The best-performing model is evaluated by IoU- and F1-score during validation shown in Table 2. Both XGBoost and U-Net models have satisfied performance for detecting the gravel and sand with IoU-score > 0.94 and F1-score > 0.97. These two models have similar IoU- (0.73 – 0.81) and F1-score (0.85 – 0.89) when predicting the classes of mixed and low-density vegetation (vegetation 2). However, XGBoost (IoU-0.95) has a better performance in classifying vegetation 1 than U-Net (IoU-0.83). Clay is a difficult class to be recognized by XGBoost with a low IoU-score of 0.77, while U-Net did not detect any clay pixel. In addition, the overall accuracy of XGBoost model is up to 0.97, and 0.96 for the U-Net model. The surface sedimentary features of mixed sand-gravel tidal inlet are classified by optimized XGBoost and U-Net model at the pixel level via the non-overlapping patch process (size 256), and the comparison between RGB, ground truth map and model predictions of the ROIs can be seen in Figure 3.

Table 2. *Performance comparison between XGBoost and U-Net models*

Classes	XGBoost		U-Net	
	IoU-score	F1-score	IoU-score	F1-score
No data	0.9998	0.9999	0.9846	0.9922
Gravel	0.9756	0.9876	0.9644	0.9819
Sand	0.9423	0.9703	0.9433	0.9708
Mixed	0.7615	0.8646	0.7330	0.8459
Vegetation 1	0.9510	0.9749	0.8322	0.9084
Vegetation 2	0.8058	0.8910	0.7774	0.8748
Clay	0.7664	0.8677	0	0
Mean	0.8857	0.9365	0.7478	0.7963

Figure 4 shows the patches that have the lowest average F1-score in the XGBoost and U-Net models, and Table 3 provides the model performance of these patches. The locations of these two patches are marked in Figure 3. The worst patches of both model predictions occur in the updrift ROI, which has classes of gravel, mixed and vegetation 2. U-Net predicted all or most of the pixels in these two patches as gravel. XGBoost can detect most of the gravel pixels, part of the mixed pixels, and a few pixels of vegetation 2, but predicted some classes that do not show in the ground truth patches.

The mixed class includes the pixels of thin gravel covering sand, thin clay covering gravel, wood debris, broken shells, and dead seaweed covering gravel etc. It is challenging to label these in more detail at the high-resolution pixel level, which largely determines the accuracy and precision of trained models. XGBoost model was trained using isolated pixel values without any geospatial associations, which highly depends on the RGB band values and consistent light conditions during UAV surveys. In contrast, the U-Net model is prone to ignore classes that are sparsely distributed in gravel patches or that have minimal training samples, due to the geospatial feature computation. However, compared to the self-trained U-Net model (Kim *et al.*, 2024), the pre-trained encoder ResNet34 requires fewer training samples and has better accuracy.

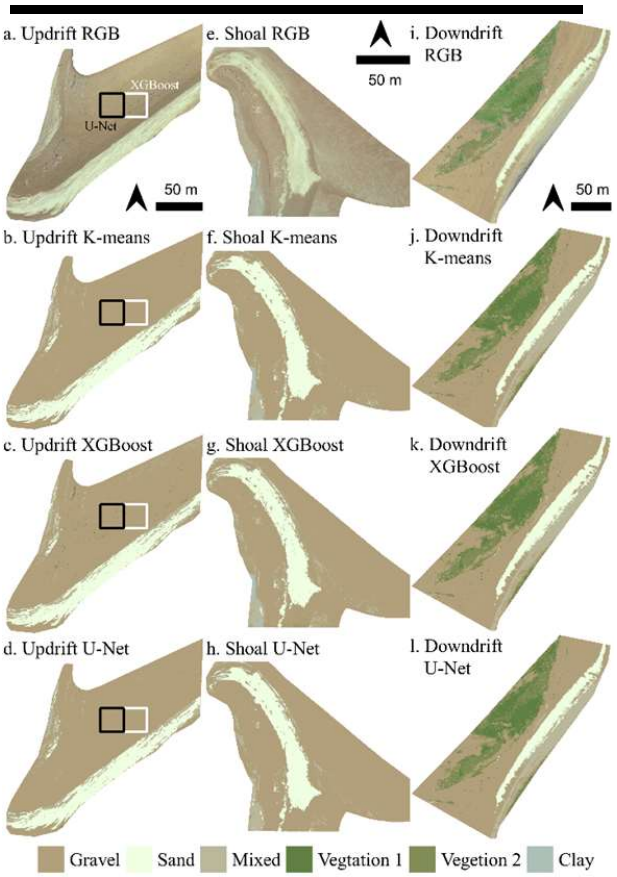

Figure 3. The RGB, K-means labels, XGBoost and U-Net predictions of the ROIs. The white and black squares are the patches having the lowest average F1-score predicted by XGBoost and U-Net, respectively.

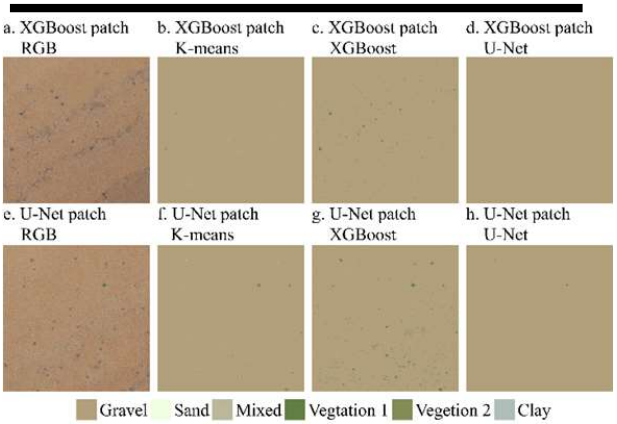

Figure 4. The RGB, K-means labels, XGBoost and U-Net model predictions of the patches (size 256) having the lowest average F1-score in the XGBoost and U-Net models.

Table 3. *F1-score of the worst model predicted patches*

Classes	True class	XGBoost lowest F1-score patch		U-Net lowest F1-score patch	
		XGB	U-Net	XGB	U-Net
Gravel	√	0.9960	0.9995	0.9914	0.9988
Sand	-	-	-	-	-
Mixed	√	0.2424	0	0.0612	0
Vegetation 1	-	0	-	0	0
Vegetation 2	√	0.0304	0	0.0909	0.0787
Clay	-	0	-	-	-
Mean		0.2538	0.3332	0.2859	0.2694

CONCLUSIONS

Based on the RGB UAV images, both XGBoost and U-Net presented a good ability to segment or classify the surface sedimentary features in this mixed sand-gravel inlet at the high-resolution pixel level, specifically detecting gravel, sand and vegetation. These models could be improved by providing more detailed and high-quality labels, and including more geospatial features like digital elevation models and texture information to reduce the light-condition influence on spectral reflectance and broaden the morphosedimentary relevance of predictions. In addition, the pre-trained encoders (ResNet34 used in this study) of U-Net could be a good choice for a small training dataset. Overall, this study represents a significant advancement in coastal monitoring, offering valuable methods to understand sedimentary processes within mixed sediment systems.

ACKNOWLEDGMENTS

The authors wish to thank Xiaolin Hou and Yixuan Chang for data collection, Feng Yin and Thomas Keel for model development, and the anonymous reviewers for their helpful comments.

LITERATURE CITED

Barbosa, B.D.S., Ferraz, G.A.S., Gonalves, L.M., Marin, D.B., Rossi, G., 2019. RGB vegetation indices applied to grass monitoring: a qualitative analysis. *Agronomy Research*, 17, 349–357.

Boothroyd, J.C., 1985. Tidal inlets and tidal deltas. *Coastal Sedimentary Environments*, 445–532.

Burningham, H., French, J., 2006. Morphodynamic behaviour of a mixed sand-gravel ebb-tidal delta: Deben estuary, Suffolk, UK. *Marine Geology*, 225, 23–44.

Burningham, H., French, J., 2007. Morphodynamics and sedimentology of mixed-sediment inlets. *Journal of Coastal Research*, SI 50, 710 – 715.

Burvingt, O., Lerma, A.N., Lubac, B., Mallet, C., Senechal, N., 2022. Geomorphological control of sandy beaches by a mixed-energy tidal inlet. *Marine Geology*, 450, 106863.

Caballero, I., Stumpf, R.P., 2021. On the use of Sentinel-2 satellites and lidar surveys for the change detection of shallow bathymetry: The case study of North Carolina inlets. *Coastal Engineering*, 169, 103936.

Casella, E., Rovere, A., Pedroncini, A., Stark, C.P., Casella, M., Ferrari, M., Firpo, M., 2016. Drones as tools for monitoring beach topography changes in the Ligurian Sea (NW Mediterranean). *Geo-Marine Letters*, 36, 151–163.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, *in: Proceedings of Knowledge Discovery and Data Mining. Presented at the 22nd Acm Sigkdd International Conference*, San Francisco, CA, USA, 785–794.

Costa, L., Nunes, L., Ampatzidis, Y., 2020. A new visible band index (vNDVI) for estimating NDVI values on RGB images utilizing genetic algorithms. *Computers and Electronics in Agriculture*, 172, 105334.

Dong, H., Xu, X., Wang, L., Pu, F., 2018. Gaofen-3 PolSAR Image Classification via XGBoost and Polarimetric Spatial Information. *Sensors*, 18, 611.

FitzGerald, D.M., Buynevich, I.V., Davis, R.A., Fenster, M.S., 2002. New England tidal inlets with special reference to riverine-associated inlet systems. *Geomorphology*, 48, 179–208.

Georgiou, I.Y., Messina, F., Sakib, M.M., Zou, S., Foster-Martinez, M., Bregman, M., Hein, C.J., Fenster, M.S., Shawler, J.L., McPherran, K., Trembanis, A.C., 2023. Hydrodynamics and Sediment-Transport Pathways along a Mixed-Energy Spit-Inlet System: A Modeling Study at Chincoteague Inlet (Virginia, USA). *Journal of Marine Science and Engineering*, 11, 1075.

Gong, J., Burningham, H., 2024. Multivariate analysis of inlet morphodynamics: The mixed sand-gravel ebb-tidal delta of Deben Estuary, Suffolk. *Geomorphology*, 454, 109163.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning for Image Recognition, *in: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Las Vegas, NV, USA, 770–778.

James, M.R., Robson, S., 2012. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, 117, F03017.

Kim, K.-L., Woo, H.-J., Jou, H.-T., Jung, H.C., Lee, S.-K., Ryu, J.-H., 2024. Surface sediment classification using a deep learning model and unmanned aerial vehicle data of tidal flats. *Marine Pollution Bulletin*, 198, 115823.

Sefercik, U.G., Kavzoğlu, T., Çölkesen, İ., Nazar, M., Öztürk, M.Y., Adali, S., DiNç, S., 2023. 3D positioning accuracy and land cover classification performance of multispectral RTK UAVs. *International Journal of Engineering and Geosciences*, 8, 119–128.

Sewada, R., Goyal, H., 2023. Landuse Landcover Change Detection Using Geospatial Semantic Segmentation U-Net-ResNet Architecture, *in: Proceedings of 2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10-HTC)*, IEEE, Rajkot, India, 620–625.

Son, C.S., Flemming, B.W., Bartholomä, A., 2011. Evidence for sediment recirculation on an ebb-tidal delta of the East Frisian barrier-island system, southern North Sea. *Geo-Marine Letters*, 31, 87–100.

Sutherland, J., Soulsby, R., 2010. Guidelines for the physical modelling of sediment dynamics, *in: Proceedings of the Third International Conference on the Application of Physical Modelling and Coastal Protection. Presented at the Coastlab 2010*, Barcelona.