# **Supplementary information 1**

## Steps for 3D reconstruction process

Image preparation steps followed in this study:

1. Extracted 1 frame per second from the video footage.
2. Grouped the images into consecutive seafloor-visible images interrupted only by lifted sediment or smoke obscuring the camera or high altitude.
3. Repeated steps 1) and 2) for both cameras (Rayfin HDE-GigE-6000-DBH13-LO and Kongsberg OE14-522FS-0009) so that each group consisted of 2 subfolders, one for each camera.
4. Color correction was applied to homogenize the colors from the 2 cameras. This consisted in applying transformations to RGB values of the frames from the Kongsberg camera. More specifically, the red band was increased, the blue band decreased, while the green band was kept almost constant. This step was performed in Irfanview.  
   To enhance contrast without increasing the noise we applied the CLAHE method to all images. This was done using library *cv2* in Python, version 3.9.19.

Steps in Agisoft Metashape:

1. In Agisoft Metashape import separately each group folder containing the camera subfolder so that one chunk in Metashape corresponds to one group of images.
2. The images are manually masked to remove noise in the background (e.g., lifted sediment) and visible parts of the vehicle.
3. Start the alignment step. If Metashape doesn’t align all images, the ones left out can either be excluded (e.g., if these are only a couple of images at the beginning or end of the sequence) or some markers can be manually added to distinct object shapes (e.g., tip of the chimney, corner of a rock, etc.).  
   For all chunks the alignment of the cameras is performed on a high accuracy with unlimited number of tie and key point (parameters set to 0). Masks are applied to key points in alignment step.
4. The sparse point cloud is cleaned using the automated filtering procedure developed by (Mohren & Schulze, 2024), based on the guidelines proposed by the USGS (Over et al., 2021). More conservative thresholds were used, as the guidelines by USGS are based on aerial imagery which does not face the limitations of underwater optical data. More specifically, the reconstruction uncertainty threshold was set to 50, the projection accuracy to 8, the reprojection error to 1 and the RMSE minimization to 0.5.
5. Models are created using Depth Maps as the source data and setting a high quality face count. These are then textured based on the source images.
6. The models are cleaned by manually deleting background areas or artifacts which might be introduced by moving objects (e.g., sediment) or during the interpolation processing phase.
7. To scale the models, the distances between the two cameras is used as reference. This was done by extracting the estimated distances from Metashape only for sequences of frames running parallel to each other in the 3D reconstruction. A schematic representation is shown in the Supplementary Fig 1. The distance between each pair of images was set to the physical distance between cameras, measured at 29.3cm. A total of 541 scale bars were placed on the model. This step was carried out in Agisoft Metashape using its integrated Python environment.

Supplementary Figure 1: Model of NUI with a physical measured distance of 29.3 cm (left) and that same distance applied between the estimated camera locations in the 3D model (right). This distance is marked in the 3D model by the orange line between the two cameras (light and dark blue). Each orange line between 2 cameras represents a placed scale bar.

29.3 cm

