

CLIFF CHANGE DETECTION USING SIAMESE KPConv DEEP NETWORK ON 3D POINT CLOUDS

I. de Gélis^{1,2*}, Z. Bessin^{3,4*}, P. Letortu⁴, M. Jaud^{5,3}, C. Delacourt³, S. Costa⁶, O. Maquaire⁶, R. Davidson⁶, T. Corpetti⁷, S. Lefèvre²

¹ Magellium, F-31000 Toulouse, France

² IRISA UMR 6074, Université Bretagne Sud, F-56000 Vannes, France - iris.de-gelis@irisa.fr

³ Geo-Ocean - UMR 6538, Univ Brest, CNRS, F-29280 Plouzané, France - zoe.bessin@univ-brest.fr

⁴ LETG - UMR 6554, Univ Brest, CNRS, F-29280 Plouzané, France

⁵ European Institute for Marine Studies (IUEM) - UMS 3113, Univ Brest, CNRS, F-29280 Plouzané, France

⁶ IDEES - UMR 6266, Normandie Univ, UNICAEN, CNRS, F-14000 Caen, France

⁷ CNRS, LETG UMR 6554, F-35000 Rennes, France

Commission III, WG III/7

KEY WORDS: Erosion detection, change detection, cliffs, 3D point clouds, deep learning

ABSTRACT:

Mainly depending on their lithology, coastal cliffs are prone to changes due to erosion. This erosion could increase due to climate change leading to potential threats for coastal users, assets, or infrastructure. Thus, it is important to be able to understand and characterize cliff face changes at fine scale. Usually, monitoring is conducted thanks to distance computation and manual analysis of each cliff face over 3D point clouds to be able to study 3D dynamics of cliffs. This is time consuming and inclined to each one judgment in particular when dealing with 3D point clouds data. Indeed, 3D point clouds characteristics (sparsity, impossibility of working on a classical top view representation, volume of data, ...) make their processing harder than 2D images. Last decades, an increase of performance of machine learning methods for earth observation purposes has been performed. To the best of our knowledge, deep learning has never been used for 3D change detection and categorization in coastal cliffs. Lately, Siamese KPConv brings successful results for change detection and categorization into 3D point clouds in urban area. Although the case study is different by its more random characteristics and its complex geometry, we demonstrate here that this method also allows to extract and categorize changes on coastal cliff face. Results over the study area of Petit Ailly cliffs in Varengeville-sur-Mer (France) are very promising qualitatively as well as quantitatively: erosion is retrieved with an intersection over union score of 83.86 %.

1. INTRODUCTION

The dynamics of cliff erosion is a complex phenomenon triggered by various factors whose relative contribution is still difficult to estimate. Since cliff erosion is likely to increase with sea level rise due to climate change (Slott et al., 2006; Ashton et al., 2011; Masson-Delmotte et al., 2021), understanding changes in coastal cliffs in order to better manage them would ensure the safety of communities and infrastructure threatened by erosion. In cliff erosion, we distinguish mass movements and debris falls. According to Varnes (1978), what we call debris falls refer to tiny rocks falling from the cliff face, while mass movements refer to rock falls corresponding to larger scale movements of part or all of the cliff. To quantify cliff retreat rates over decades, the simplest way is to measure the distance from different cliff top locations using aerial photographs or historical maps (Lee and Clark, 2002; Brooks and Spencer, 2010). But since changes at the cliff top may not be representative of changes over the entire cliff face (Young et al., 2009), comparisons between cliff face 3D Point Clouds (PCs) are also used. For this second approach, methods such as terrestrial laser scanning (TLS), airborne laser scanning (ALS), mobile laser scanning, unmanned aerial vehicle photogrammetry (UAVP), or terrestrial photogrammetry (TP) are used to monitor the cliff face (Young et al., 2010; James and Robson, 2012; Michoud et al., 2014; Letortu et al., 2018). Surveys are taken over time and PCs are compared using tools such as Cloud to Cloud (C2C)

(Girardeau-Montaut et al., 2005) in CloudCompare software (Girardeau-Montaut, 2016). Manual analysis of the differences is then performed to determine areas of erosion on the cliff face and accumulation at the cliff foot. PCs differences provide cliff retreat rates and allow for estimation of eroded volume (Letortu et al., 2015). Different methods trying to automate cliff faces changes extraction and cliff top and toe delineation exist, however most of them do not process 3D data directly but 2.5D rasterization of PCs and finally use Digital Elevation Model (DEM) differences (Young and Ashford, 2006; Swirad and Young, 2021). For example, in Young and Ashford (2006), the elevation variability is calculated through DEM difference where negative cells represent erosion and positive cells represent accumulation, enabling to highlight significant erosion areas on cliffs with slight slope. In Swirad and Young (2021), inventories of erosion and deposition objects are retrieved with a method consisting in four steps: 1) PC processing, 2) cliff face identification, 3) change object inventories, and 4) object classification. Erosion rates are calculated using DEM 2.5D analysis combined with vertical and planimetric detection thresholds, Normalized Difference Vegetation Index (NDVI), machine learning processes and manual quality control. These methods, although relevant, are not applicable on vertical or very steep cliffs. A direct 3D approach allows to study vertical cliffs and to limit the loss of information generated by the data rasterization.

Nowadays deep learning is the leading framework in classification and even change identification in remote sensing data. Recently, Siamese KPConv network was proposed to highlight and

* Corresponding authors

categorize changes into urban 3D PCs (de Gélis et al., 2021). In the following study we also aim to detect and categorize changes but in cliffs. Although the field of study is very different, last decades show that as long as a training step is possible, deep learning methods can adapt to very different cases of study. For instance, the U-Net architecture (Ronneberger et al., 2015) has been successfully applied to multiple fields from biomedical to computer vision and remote sensing. Thus in the following study, we aim to experiment the Siamese KPCnv network for cliff change extraction and classification. To do so, we base our study on Petit Ailly cliffs in Varengeville-sur-Mer (France), described in Section 2. In Section 3, materials and the method are detailed. Then, results are presented and discussed in Section 4 and 5 respectively, before conclusions in Section 6.

2. STUDY AREA

For this study, we focus on a section of cliffs at Varengeville-sur-Mer, located in Normandy, Seine-Maritime (NW France), along the English Channel. Cliffs of Seine-Maritime cover 120 km of coast (from Le Havre to Le Tréport), and are around 60–70 m high. Geologically, they are part of the sedimentary Paris Basin and made of Upper Cretaceous chalk interbedded with flint bands (Pomerol et al., 1987; Costa, 1997; Laignel, 2003; Mortimore and Duperret, 2004). Along the cap d’Ailly, where Varengeville-sur-Mer is located, cliffs are specific: the residual flint formation over Santonian chalk strata has been replaced by a bed of clay and sandy sediments from the Paleogene age (Bignot, 1962). The cap d’Ailly is an erosion hotspot where retreat rates can locally exceed 0.80 m/y (Letortu et al., 2014). In Varengeville-sur-Mer, the Petit Ailly cliff face (250–350 m long, 40 m high, slope from 70° to overhang) (Letortu et al., 2018; Jaud et al., 2019) (see Figure 1) is monitored from 2010 every 4–5 months (Letortu et al., 2015, 2019) in the framework of DYNALIT French Observatory Service by terrestrial laser scanning and photogrammetry (Letortu et al., 2018) to improve knowledge on retreat rates and main triggering mechanisms for rock fall activity. This large dataset highlights that rock fall activity is intense over 250–350 m, with a retreat rate of 0.36 m/y and 4 rock falls over 1 000 m³ (Letortu et al., 2019).

3. MATERIAL AND METHOD

3.1 Data

PCs used in this study are taken from several sources: TP and TLS, with various densities (see Table 1). The PC from TP survey contains a significant larger number of points than PCs from TLS surveys, hence TP PC has a larger point density than TLS PCs (9 440 pts/m² for 2016 TP PC and a maximum of 381 pts/m² for 2017 TLS PC). Thereby the following study is based on multi-sensor acquisitions.

Date	Type	Nb of points (pts)	Mean surface density (pts/m ²)
2013-09-25	TLS	2 148 462	202
2016-01-28	TP	123 702 868	9 440
2017-11-02	TLS	4 154 905	381
2018-01-16	TLS	4 038 505	334
2020-04-14	TLS	3 293 548	276

Table 1. Varengeville-sur-Mer PCs dataset.

3.1.1 Terrestrial Laser Scanning survey The TLS used for this study is a Riegl® VZ-400 emitting a laser pulse in the near-infrared (1 550 nm), using the time-of-flight of laser pulse to measure the position of a point. Scan data provided by this instrument have a theoretical accuracy of 0.005 m and a precision of 0.003 m at a range of 100 m. The Riegl® VZ-400 is equipped with a Nikon D800 camera with a fisheye lens providing photographs that can be used to texturize the 3D point cloud. The georeferencing of the PCs is performed with the Trimble M3 total station to measure several reflective targets used as ground control points (GCPs). The advantage of the use of a total station is that it allows to measure targets closer to the base of the cliff without being affected by mask effects as a Differential Global Positioning System (DGPS) would be. PCs are projected in RGF 93 – Lambert 93 and IGN69, i.e. official planimetric coordinate system and vertical datum in France, respectively. Numerous TLS targets (see Figure 1) are used (redundancy of measurements) and disposed at different distances from the scanner to limit the alignment error (Letortu et al., 2018).

3.1.2 Terrestrial Photogrammetry survey The terrestrial photogrammetric device used in this study is a Nikon D800 reflex camera configured with a 35 mm focal length. The camera positions should be a short distance apart and at least 20 m from the cliff face. Photographs are acquired from multiple positions at 10–20° angular intervals over a wide range of angles to cover the area (James and Robson, 2012). To ensure a quality result, an overlap of at least 60 % (80 % is recommended for SfM photogrammetry) is required between each photograph, and each scene must be taken from various points of view (Letortu et al., 2018). In order to georeference the models, GCPs are also needed for TP survey. We use TP targets and measure their absolute coordinates with a total station (see Figure 1). To limit doming effect due to radial distortion and erroneous camera model, GCPs should be numerous (James and Robson, 2014; Jaud et al., 2017). Photographs of the cliff and target coordinates are used to derive a georeferenced 3D PC using the Structure-from-Motion Multiview Stereo Photogrammetry (SfM-MVS) algorithm in Agisoft Metashape. Accuracy is not measured directly for TP survey because it requires geodetic references (surveyor’s nails) that are not visible in the 2016 PC. Comparison between TLS and TP 2016 PCs reveals a mean error value from 0.013 m to 0.03 m. Further details on the TP survey method used are described in Letortu et al. (2018).

3.1.3 Point clouds annotation Manual annotation is performed between the PC pairs (see Table 2) using C2C tool (Cloud-Compare) in order to assign a label to each point of the second PC of the pair. Four classes are defined: “unchanged”, “erosion”, “accumulation” and “no data to compare”. The “unchanged” class designates points whose position is unchanged or almost unchanged (with a tolerance margin of 20 cm) between the two surveys of each PC pair. Points of the most recent PC of the pair are considered as “erosion” when they appear behind the oldest PC and as “accumulation” when they are located in front of the oldest PC and rather at the cliff foot or on slight slopes (below 15°) (see Figure 2). A “no data to compare” class is also defined to address the problem of an occlusion (due to different methods of data acquisition) only visible in the oldest PC of the pair of PCs. Indeed, in this case, the C2C distance computed between the two PCs is inconsistent, thus preventing us to evaluate whether it is accumulation, erosion or no change. We recall that the C2C distance has just been used to help the manual annotation task. To annotate the datasets, with respect to the distance separating the two PCs of each pair, a threshold

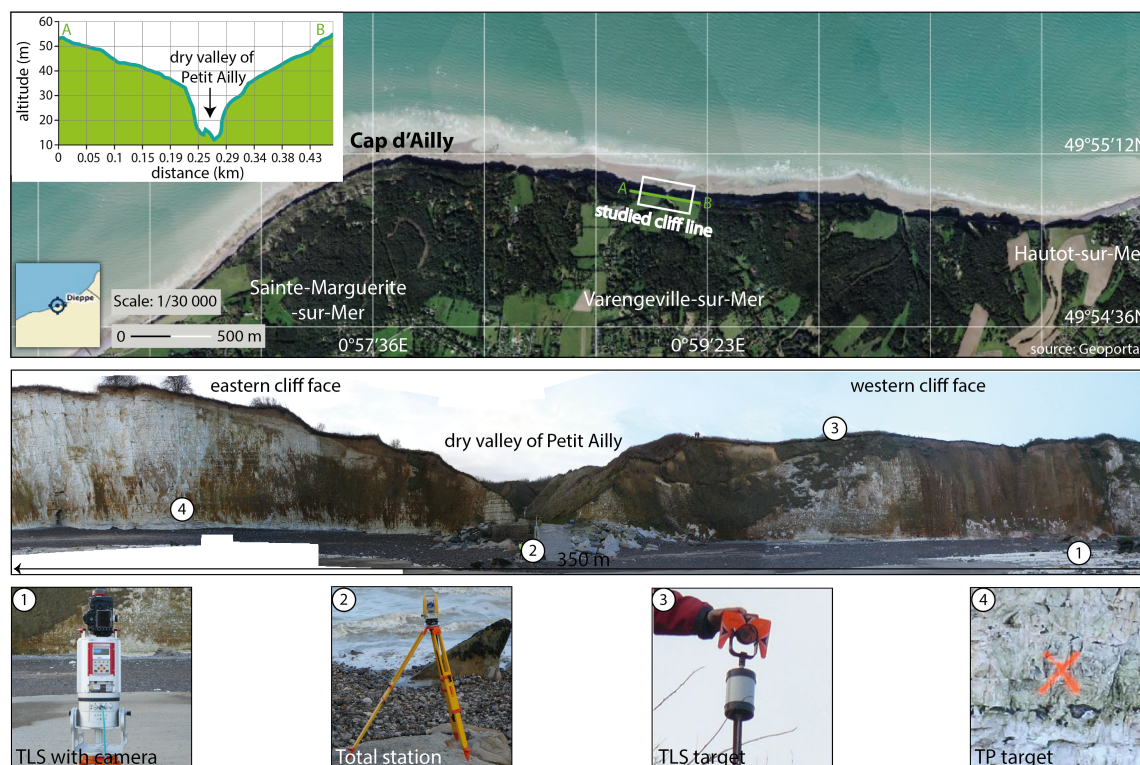


Figure 1. Panorama and aerial photography of Petit Ailly cliff (Varengeville-sur-Mer) and instrumentation used for the TLS survey (2016-01-28) (modified from Letortu et al. (2018)).

PC pairs	Type	Class distribution (%)			
		Unchanged	Erosion	Accumulation	“No data to compare”
2013-09-25 – 2016-01-28	TLS – TP	51.77	40.57	1.32	6.33
2016-01-28 – 2017-11-02	TP – TLS	66.07	22.43	0.47	11.02
2017-11-02 – 2018-01-16	TLS – TLS	79.65	16.20	3.19	0.96
2018-01-16 – 2020-04-14	TLS – TLS	69.67	29.92	0.00	0.41

Table 2. Varengeville-sur-Mer PC pairs.

of 50 cm is used, and refined to 20 cm for some parts. Thereby in this study, we mainly focus on extracting movements greater than 20-50 cm, thus smaller movement like debris falls may be missed.

3.2 Siamese KPConv network

We propose here to use and assess the recent Siamese KPConv network for cliff change extraction. This model relies on a deep Siamese architecture, that has proved its ability to detect and categorize changes even for remote sensing purposes (Daudt et al., 2018; Jiang et al., 2020) with 2D satellite images. In particular Siamese architectures contain an encoder with two branches. Usually composed of a succession of convolution and max-pooling operations, each branch extracts features coming from each input data. To extend the Siamese principle to 3D PCs, a convolution operation suitable with 3D PCs characteristics should be used. Indeed, traditional convolution used for image processing can not be applied here as PCs are not contained into matrices and therefore, the access to neighbors involved in convolutions is less obvious than for structured data on regular grids as images. Thus, de Gélis et al. (2021) rely on Kernel Point Convolution (KPConv). These convolutions are specially designed to extract features from raw 3D PCs by applying weights thanks to kernel points dispatched into a 3D

ball (Thomas et al., 2019). Max-pooling operations are here replaced by strided KPConv acting as a down-sampler of 3D PCs, thus features can be extracted at different scales.

Siamese KPConv architecture is presented in Figure 3. Like a usual encoder-decoder with skip connections, Siamese KPConv also contains skip links between encoder and decoder. Indeed, at each scale of the decoding part, the difference of extracted features associated with the corresponding encoding scale (see Figure 3) are concatenated in the decoder part. Conversely to 2D images where pixels of both images can be easily associated, in 3D Siamese KPConv the computation of the feature difference is not obvious since PCs do not contain the same number of points and are not defined at the same positions, even in non-changed areas. Authors of Siamese KPConv suggest here to compute this difference in each point of the second PC by retrieving features of the corresponding nearest spatial point in the first PC. More details can be found in their original paper (de Gélis et al., 2021). Weights between the two branches of the encoder are shared as it is done when input data are quite similar. Moreover, sharing weights forces the network to be less focused on one type of data. As explained in Section 3.1, our PCs do not come from the same modality between each acquisition. As finally training, validation and test sets do not come from the same pair, we would like to avoid that each branch of the encoder only specializes in extracting information for only

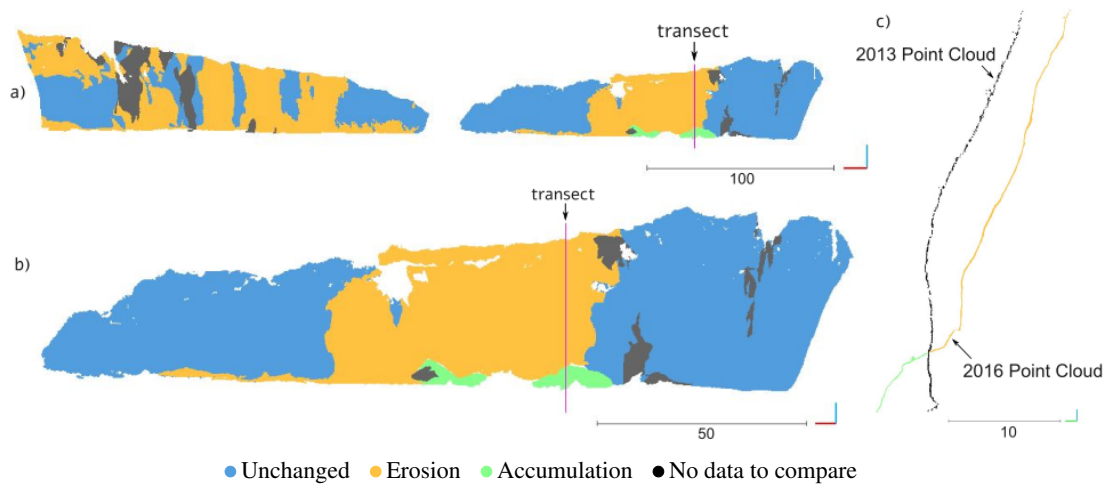


Figure 2. a) Annotation of the PC from 2016-01-28 with the PC from 2013-09-25, b) West part of Ailly cliff and transect position c) Transect showing erosion of the cliff with accumulation at the cliff base.

one type of PC. Thus, sharing weights in the encoder allows us to strengthen the generalization capability of our method, to deal with both TLS and TP data.

4. RESULTS

4.1 Experimental protocol

4.1.1 Dataset configuration In order to conduct our experimentation, we divide our dataset into three parts dedicated to training, validation and testing. We recall that data are annotated according to the previous acquisition forming 4 pairs of PCs. As can be seen in Table 2, each class is not equivalently represented in each PCs. As a matter of fact, accumulation class is very rare compared to eroded and even unchanged areas. Thus the division of the dataset is made in such a way that each split is as representative as possible of each class of changes. Thereby, the split of pairs of PCs in each training, validation and testing set is made as indicated in Table 3. Accumulation class is less represented in point clouds, thus we divided the 2017-2018 pair into an eastern and a western part at the dry valley of Petit Ailly (see Figure 1) to have examples of accumulation in both the training and in the testing set.

Training set	Validation set	Testing set
2013-2016 2017-2018 West	2016-2017	2017-2018 East 2018-2020

Table 3. Split of the dataset into training, validation and testing sets. For each pair of PCs, years relates to the year of acquisition given in Table 1. The annotation is always given for the second PC of the pair with regards to the first PC.

4.1.2 Experimental set-up To set-up experiments, we have to choose the initial sub-sampling rate of PCs, called dl_0 . Thus, no matter the type of input PC (TLS or TP), the input PC's resolution at the first layer is always the same. Notice that final results are interpolated back to initial PC resolution given in Table 1. In our mind, dl_0 should be chosen as small as possible to stick with PC initial density, while in the same time fitting with available memory in the Graphical Processing Unit (GPU). Finally, this first sub-sampling rate is directly linked with the

size of input given to the network. Indeed, as for satellite images divided into patches, PCs are also sub-divided to feed the network. Conversely to urban change detection experiments made with Siamese KPConv, we choose here to divide PCs into spheres and not vertical cylinders. Indeed, in urban areas, the motivation of considering cylinders instead of spheres is to be sure to always include ground in vertical cylinders (de Gélis et al., 2021). The case of study here is quite different, and spheres appear more suited than cylinders because they contain less points, allowing to choose a larger radius. We recall that considering too many points as input for the network leads to memory capacity issues. Spheres are centered on a point of the second date PC, thus the radius should not be chosen too small to ensure to give also points of the first PC in case of large changes as well as providing enough context. In Thomas et al. (2019), the authors chose the radius of input sphere of 50 times dl_0 . However according to our own experiments, best results were obtained with spheres of 10 m in diameter and dl_0 set to 0.15 m. As far as other network hyper-parameters are concerned, we use the same configuration as in de Gélis et al. (2021): a Stochastic Gradient Descent (SGD) with momentum to minimize a point-wise negative log-likelihood loss, with a batch size of 10, a momentum of 0.98 and an initial learning rate of 10^{-2} . The learning rate is scheduled to decrease exponentially. A probability dropout of 0.5 in the last classification layers is set. Also, in order to prevent from over-fitting, we set a L2 loss regularization with a factor of 10^{-6} . As the dataset is largely imbalanced, the loss is weighted according to training set class distribution.

In the following experiment, we decided to normalize input spheres along X and Y axis by retrieving the minimum value into the sphere of X and Y axis respectively. As for the vertical Z axis, we do not perform a normalization with regards to the minimum value in the whole cliff, so to keep information related to elevation. Indeed, this may help for change classification, in particular for accumulation class since it is mainly located at the cliff foot.

We provide a comparison with a popular distance-based method, Multiscale Model to Model Cloud Comparison (M3C2), giving a mean surface change along a normal direction (Lague et al., 2013). Based on this distance, we apply an empirical thresholding at -0.2 m and 0.2 m to extract accumulation, unchanged and erosion areas. Even though M3C2 is not specifically designed

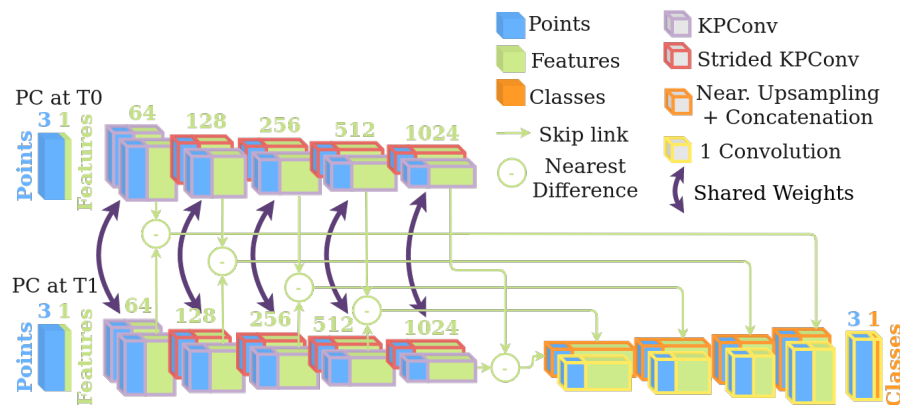


Figure 3. Siamese KPConv network architecture (de Gélis et al., 2021).

to categorize surface changes, this allows to compare Siamese KPConv quantitative results to another method, it is discussed in Section 5.1.

4.1.3 Metrics Finally, in order to provide quantitative assessment, we measure the Intersection over Union (IoU) for each classes as well as the mean IoU (mIoU). The IoU is given by the following equation:

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

where TP, TN, FP and FN respectively stand for True Positive, True Negative, False Positive and False Negative.

4.2 Qualitative and quantitative results

Qualitative results are presented in Figures 4, 5 and 6 corresponding to each part of the testing dataset. As we can see, the predictions provided by Siamese KPConv prediction are close to the ground truth. Erosion great structures are well recognized, and even smaller parts of erosion seem to be highlighted as it can be seen on the right side of Figure 5. Quantitative results are shown in Table 4. We report results for the three classes of interest: unchanged, erosion and accumulation. We remind that the class “no data to compare” is not a class of change and is a bit subjective as it depends on annotator confidence in existing surrounding points in an area. In particular, per class IoU indicates that unchanged area and erosion are mainly well classified as already shown on qualitative results. Main differences with the ground truth appear at boundaries of erosion parts and in some more intricate areas like in the top middle left side of the Figure 5 where erosion, unchanged and “no data to compare” classes are almost mixed up and cliff structure is more complex. Accumulation class obtains a lower score in comparison to erosion, surely explained by the only few accumulation examples available in the whole dataset. Indeed training set contains only 1.51 % of points for accumulation whereas the erosion represents 33.83 % of points. Worth noting that testing set follows the same trend (see Figure 4(a), 5(a) and 6(a)). However, and as shown in Figure 4, it remains quite well retrieved. Main confusions for this class appear also at the boundary of accumulation zone and in “no data to compare” areas where there are some isolated points predicted as accumulation. Concerning the “no data to compare” class, main parts classified as “no data to compare” in the ground truth are also classified like this by Siamese KPConv. However there seem to be more areas identified as “no data to compare” by the network. Indeed, in some parts, only a few points allow the annotator to

be sure that there is or not a change whereas in some other parts, the identification of change is trickier, or even impossible, so it has been marked as “no data to compare”. Thereby, ground truth and prediction should not be strictly compared. Finally, in Table 4 we also report the mean of IoU over classes of interest ($mIoU_{int}$) so the mIoU without the subjective “no data to compare” class.

	Per class IoU(%)			
	Unchanged	Erosion	Acc.	$mIoU_{int}$
Siam. KPConv	91.94	83.86	70.28	82.03
M3C2 + threshold	95.06	87.23	48.20	68.12

Table 4. Quantitative results for Siamese KPConv and M3C2 methods for change detection and categorization. Acc. stands for accumulation.

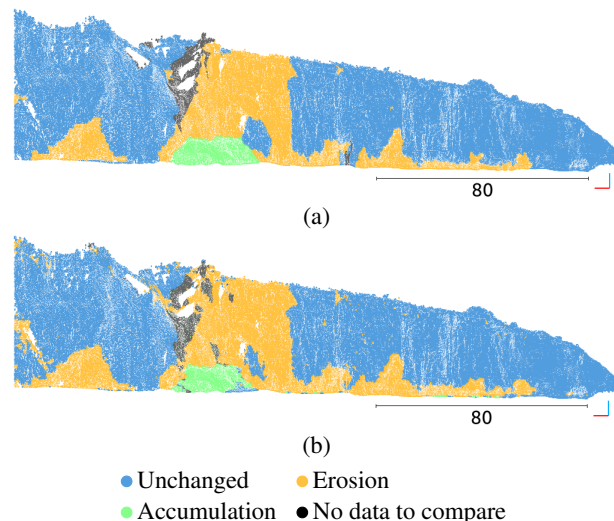


Figure 4. Results on east part of the cliffs between 2017 and 2018 acquisition. Points of the ground truth (a) and in Siamese KPConv results (b) are colored according to the change.

5. DISCUSSION

5.1 Comparison with M3C2 results

In Table 4, a comparison with the distance-based method M3C2 is made. To distinguish between types of changes, a threshold is applied on this distance. According to Table 4, even if quantitative results for unchanged and erosion classes are higher for

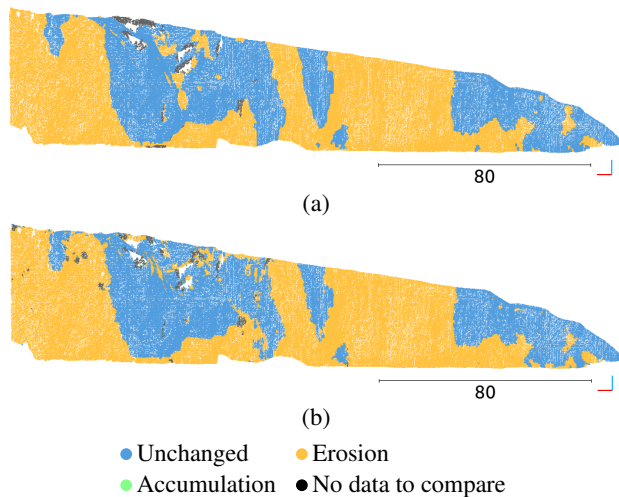


Figure 5. Results on east part of the cliffs between 2018 and 2020 acquisition. Points of the ground truth (a) and in Siamese KPConv results (b) are colorized according to the change.

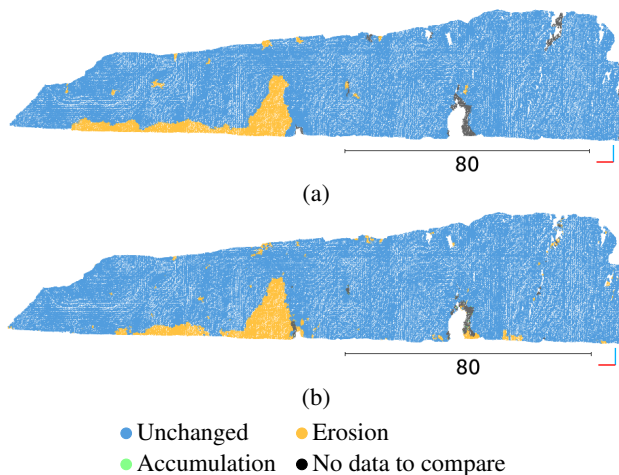


Figure 6. Results on west part of the cliffs between 2018 and 2020 acquisition. Points of the ground truth (a) and in Siamese KPConv results (b) are colorized according to the change.

M3C2 method, general results given by $mIoU_{int}$ and score of accumulation class are better for Siamese KPConv method. Let us note that M3C2 results are only based on a measure of distance between two PCs (quantitative evaluation) whereas Siamese KPConv also gives an interpretation of changes by giving directly a categorization. Siamese KPConv automatically extracts features based on points neighborhood at different scales (with Kernel Point convolution operations). For example, for accumulation class, the morphology and position with regards to the cliffs should be taken into account. Indeed the accumulation can appear “behind” the first PC, and therefore not be highlighted as accumulation by M3C2 method (negative distance), when the stock of debris accumulated at the foot of the cliff decreases from one survey to the next. Finally, Siamese KPConv gives a categorization of changes based on surroundings of points. For a generalization on various types of cliffs, other classes of change (e.g. on vegetation) could be added using Siamese KPConv method which would not be possible with a threshold applied on a distance-based method such as M3C2.

5.2 Class representativity

The method developed here can be applied on cliffs with various slope degrees with satisfactory results but it should be noted that the results are dependent to the annotation quality. Notice that the same protocol for the cleaning of 3D PCs and annotation should be applied for each pair of PCs. Despite encouraging results obtained by this method, supervised deep learning still requires a dataset both large and representative of each class during the training phase. Thus, the accumulation class that is under-represented might obtain better results if our dataset contained more examples. The same remarks stand for the “no data to compare” class, which is also (as mentioned in Section 4.2) more subjective. Conversely to urban applications, cliff variability is important (especially in cavity) so information from missing data at the first PC of the pair cannot be interpolated as it is done for hidden building facade in urban environment (de Gélis et al., 2021). Hence the necessity of creating this class in areas where no conclusion about cliff dynamics can be drawn.

5.3 Artifact robustness

Then, it is worth noting that input data are registered together. However after the experimentation, we noticed a Z off-set error on the eastern part of the 2013-2016 pair, probably due to a target movement during the survey or the depression of the TLS during acquisition on a wet sandy foreshore, leading to mislabeling of erosion parts. Although this pair is in the training set, we saw that erosion identification is satisfactory, showing robustness of the method to acquisition artifacts and mislabeled data during the training.

Because of their verticality, Normandy chalk cliffs and more precisely Varengeville-sur-Mer cliffs have a very scarce vegetation cover limiting the training on vegetation growth and retreat. Hence, these classes are ignored. Nevertheless, these classes should be considered if applying the method on vegetalized cliffs.

5.4 Training configurations

In order to select the best training configuration, several settings of input data have been tested. As explained in Section 4.1, input geometry (sphere or cylindrical) as well as first sub-sampling rate dl_0 have an influence on results. Indeed, considering a too large sphere radius implies a higher sub-sampling rate, thus avoiding change extraction. Similarly, choosing a too thin sub-sampling rate requires to diminish input sphere radius implying that not enough context can be taken into account for feature extraction. Furthermore, we noticed that with spheres of a fixed 10 m in diameter, diminishing the sub-sampling rate finally does not improve results: this may be explained by the fact that setting dl_0 to 0.15 m is enough accurate compared to ground truth precision. Thereby, the sub-sampling rate and the size of input sphere should be chosen according to the level of details of the ground truth.

We have reported results with spherical sub-PCs, conversely to the vertical cylindrical inputs used in original Siamese KPConv (de Gélis et al., 2021) that were motivated by the verticality of changes and the need for including ground in PCs. In our cliff context, changes are more likely to happen on horizontal axis, so we also conducted an experiment by taking some cylindrical inputs oriented according to north-south axis as the cliff is oriented along eastern-western axis (see Figure 1). Results were quite similar to those obtained with spherical inputs. Another

idea would have been to select cylinder oriented according to the normal of the cliff face to be more precise in the orientation of cylinder.

Finally, for presented experiments, the training time was about 20 hours on a single GPU (Titan RTX), while the prediction for the whole testing set (see Table 3) takes about 8 minutes.

5.5 Erosion detection scale

The method developed here presents an interest for the long-term monitoring of cliffs since it allows the detection of mass movements of rocks. In order to detect debris falls, more time should be spent to annotate data very accurately and adjustment of the first sub-sampling rate (dl_0) is required. This rate should be chosen equal or thinner than debris size, worth noting that the minimum threshold for detecting erosion and accumulation would be the Ground Sampling Distance (GSD) of the sparser PC of the initial dataset. Thus, initial clouds resolution should be homogeneous and adequate regarding to the size of the minimal debris falls to detect. Indeed, without homogeneity between the resolutions of the different datasets, detection artifacts could appear.

6. CONCLUSIONS

In this study, we tackle erosion detection on cliff faces thanks to a recent deep learning network for change detection and categorization: Siamese KPConv (de Gélis et al., 2021). To the best of our knowledge, this is the first study where changes in cliff faces are extracted and classified thanks to a deep learning method processing raw 3D point clouds (PCs). Despite the limited training set, we have reported very encouraging results, with a mean of Intersection over Union (mIoU) over classes of interest (unchanged area, erosion and accumulation) of 82.03 %. More particularly, erosion detection results (with an IoU of 83.86 %) seems enough for future applications, as discussed further. Indeed, once eroded points have been identified, one can try to determine the eroded volume. To do so, a standard technique consists in using a 3D Convex Hull method (Preparata and Hong, 1977; Johansen and Gram, 1983) to extract eroded area in both PCs of each pair, before estimating the volume between the two dates. Given the promises of deep learning, one can also try to tackle erosion detection and quantification with an end-to-end trainable model. The results presented in this study are obtained from point clouds produced by Terrestrial Laser Scanning (TLS) or Terrestrial Photogrammetric (TP) methods. Even if densities and resolutions of the PCs resulting from these methods are very different, we have reported promising results, demonstrating the robustness to point density variations and encouraging us to integrate PCs from other sources. However, a “measurement bias” class should be added in order to use PCs from new sensors and methods such as satellite imagery. Indeed, density of PCs retrieved from satellite stereo-restitution (Letortu et al., 2020, 2021) is lower (about 2.7 pts/m² for satellite stereo-restitution using Pléiades images) and more irregular than PCs from TLS or TP surveys. Therefore, annotation is more difficult and less accurate since only sufficiently large changes will be visible to the human eye. Thus, the network should be better trained with PCs coming from a large panel of various resolutions. However, integrating satellite photogrammetric PCs in the network would allow to i) work on a data set with heterogeneous resolution and precision and ii) analyze dynamics at a larger scale. Finally, these results would allow to automate the monitoring of coastal changes, ad-

vocated in particular by governments and coastal observatories in order to improve the management of the associated risks.

ACKNOWLEDGEMENTS

This work is part of the Service National d’Observation DYNALIT, via the research infrastructure ILICO. This work was also funded by ARED/BreTel and the ISblue project, the Interdisciplinary graduate school for the blue planet (ANR-17-EURE-0015), and co-funded by a grant from the French government under the program “Investissements d’Avenir”. This research was also supported by the CNES (EROFALITT and HIRACLES projects). It is based on observations with Pléiades satellites. This work was granted access to the HPC resources of IDRIS under the allocation 2021-AD011011754R1 made by GENCI. I. de Gélis work is funded by Magellium and supported by a CNES grant.

References

- Ashton, A. D., Walkden, M. J. A., Dickson, M. E., 2011. Equilibrium responses of cliffed coasts to changes in the rate of sea level rise. *Marine Geology*, 284(1-4), 217–229.
- Bignot, G., 1962. Étude sédimentologique et micro-paléontologique de l’Éocène du Cap d’Ailly (près de Dieppe-Seine-Maritime). PhD thesis, Université de Paris.
- Brooks, S. M., Spencer, T., 2010. Temporal and spatial variations in recession rates and sediment release from soft rock cliffs, Suffolk coast, UK. *Geomorphology*, 124(1-2), 26–41.
- Costa, S., 1997. Dynamique littorale et risques naturels : L’impact des aménagements, des variations du niveau marin et des modifications climatiques entre la baie de Seine et la baie de Somme (Haute-Normandie, Picardie; France). PhD thesis, Paris 1.
- Daudt, R. C., Le Saux, B., Boulch, A., 2018. Fully convolutional siamese networks for change detection. *2018 25th IEEE International Conference on Image Processing (ICIP)*, IEEE, 4063–4067.
- de Gélis, I., Lefèvre, S., Corpetti, T., 2021. 3D urban change detection with point cloud siamese networks. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 879–886.
- Girardeau-Montaut, D., 2016. CloudCompare. France: EDF R&D Telecom ParisTech.
- Girardeau-Montaut, D., Roux, M., Marc, R., Thibault, G., 2005. Change detection on points cloud data acquired with a ground laser scanner. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(part 3), W19.
- 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.
- James, M. R., Robson, S., 2014. Mitigating systematic error in topographic models derived from UAV and ground-based image networks. *Earth Surface Processes and Landforms*, 39(10), 1413–1420.

- Jaud, M., Letortu, P., Théry, C., Grandjean, P., Costa, S., Maquaire, O., Davidson, R., Le Dantec, N., 2019. UAV survey of a coastal cliff face – Selection of the best imaging angle. *Measurement*, 139, 10–20.
- Jaud, M., Passot, S., Allemand, P., Le Dantec, N., Grandjean, P., Ammann, J., Delacourt, C., 2017. Stratégies d'optimisation d'acquisition par drone pour limiter les distorsions lors de la reconstruction 3D par les logiciels Photoscan et MicMac. *Journées CRITEX*, Grenoble.
- Jiang, H., Hu, X., Li, K., Zhang, J., Gong, J., Zhang, M., 2020. Pga-siamnet: Pyramid feature-based attention-guided siamese network for remote sensing orthoimagery building change detection. *Remote Sensing*, 12(3), 484.
- Johansen, G. H., Gram, C., 1983. A simple algorithm for building the 3-D convex hull. *BIT Numerical Mathematics*, 23(2), 145–160.
- Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, 10–26.
- Laignel, B., 2003. Caractérisation et dynamique érosive de systèmes géomorphologiques continentaux sur substrat crayeux. Exemple de l'Ouest du Bassin de Paris dans le contexte nord-ouest européen.
- Lee, M., Clark, A., 2002. *Investigation and management of soft rock cliffs*. Thomas Telford Publishing.
- Letortu, P., Costa, S., Bensaid, A., Cador, J.-M., Quénot, H., 2014. Vitesses et modalités de recul des falaises crayeuses de Haute-Normandie (France): méthodologie et variabilité du recul. *Geomorphologie: Relief, Processus, Environnement*, 20(2), 133–144.
- Letortu, P., Costa, S., Maquaire, O., Davidson, R., 2019. Marine and subaerial controls of coastal chalk cliff erosion in Normandy (France) based on a 7-year laser scanner monitoring. *Geomorphology*, 335, 76–91.
- Letortu, P., Costa, S., Maquaire, O., Delacourt, C., Augereau, E., Davidson, R., Suanez, S., Nabucet, J., 2015. Retreat rates, modalities and agents responsible for erosion along the coastal chalk cliffs of Upper Normandy: The contribution of terrestrial laser scanning. *Geomorphology*, 245, 3–14.
- Letortu, P., Jaud, M., Grandjean, P., Ammann, J., Costa, S., Maquaire, O., Davidson, R., Le Dantec, N., Delacourt, C., 2018. Examining high-resolution survey methods for monitoring cliff erosion at an operational scale. *Giscience & Remote Sensing*, 55(4), 457–476.
- Letortu, P., Jaud, M., Théry, C., Nabucet, J., Taouki, R., Passot, S., Augereau, E., 2020. The potential of Pléiades images with high angle of incidence for reconstructing the coastal cliff face in Normandy (France). *International Journal of Applied Earth Observation and Geoinformation*, 84, 101976.
- Letortu, P., Taouki, R., Jaud, M., Costa, S., Maquaire, O., Delacourt, C., 2021. 3D reconstructions of the coastal cliff face in Normandy (France) based on oblique Pléiades imagery: assessment of ASP® and MicMac® processing chains. *International Journal of Remote Sensing*, 42(12).
- Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, Ö., Yu, R., Zhou, B. (eds), 2021. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Michoud, C., Carrea, D., Costa, S., Derron, M.-H., Jaboyedoff, M., Delacourt, C., Maquaire, O., Letortu, P., Davidson, R., 2014. Landslide detection and monitoring capability of boat-based mobile laser scanning along Dieppe coastal cliffs, Normandy. *Landslides*, 12(2), 403–418.
- Mortimore, R. N., Duperret, A., 2004. *Coastal Chalk Cliff Instability*. Geological Society of London.
- Pomerol, B., Bailey, H. W., Monciardini, C., Mortimore, R. N., 1987. Lithostratigraphy and biostratigraphy of the lewes and seaford chalks: A link across the Anglo-Paris basin at the Turonian-Senonian boundary. *Cretaceous Research*, 8(4), 289–304.
- Preparata, F. P., Hong, S. J., 1977. Convex hulls of finite sets of points in two and three dimensions. *Communications of the ACM*, 20(2), 87–93.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241.
- Slott, J. M., Murray, A. B., Ashton, A. D., Crowley, T. J., 2006. Coastline responses to changing storm patterns. *Geophysical Research Letters*, 33(18), L18404.
- Swirad, Z. M., Young, A. P., 2021. Automating coastal cliff erosion measurements from large-area LiDAR datasets in California, USA. *Geomorphology*, 107799.
- Thomas, H., Qi, C. R., Deschaud, J.-E., Marcotegui, B., Goulette, F., Guibas, L. J., 2019. Kpconv: Flexible and deformable convolution for point clouds. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6411–6420.
- Varnes, D. J., 1978. Slope movement types and processes. *Special report*, 176, 11–33.
- Young, A. P., Ashford, S. A., 2006. Application of airborne LIDAR for seacliff volumetric change and beach-sediment budget contributions. *Journal of Coastal Research*, 22(2), 307–318.
- Young, A. P., Flick, R. E., Gutierrez, R., Guza, R. T., 2009. Comparison of short-term seacliff retreat measurement methods in Del Mar, California. *Geomorphology*, 112(3-4), 318–323.
- Young, A. P., Olsen, M. J., Driscoll, N., Flick, R. E., Gutierrez, R., Guza, R. T., Johnstone, E., Kuester, F., 2010. Comparison of Airborne and Terrestrial Lidar Estimates of Seacliff Erosion in Southern California. *Photogrammetric Engineering and Remote Sensing*, 76(4), 421–427.