



HAL
open science

Characterization of normal fault scarp using convolutional neural network: application to Mexico

Lea Pousse-Beltran, Sophie Giffard-Roisin, Laurence Audin, Pierre Lacan, Théo Lallemand, Glenn Cougoulat, Andrés Núñez Meneses

► To cite this version:

Lea Pousse-Beltran, Sophie Giffard-Roisin, Laurence Audin, Pierre Lacan, Théo Lallemand, et al.. Characterization of normal fault scarp using convolutional neural network: application to Mexico. 11th International INQUA Meeting on Paleoseismology, Active Tectonics and Archeoseismology (PATA), INQUA Focus Group Terrestrial Processes Perturbed by Tectonics (TPPT), Sep 2022, Aix-en-Provence, France. hal-04214143

HAL Id: hal-04214143

<https://hal.science/hal-04214143>

Submitted on 21 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Characterization of normal fault scarp using convolutional neural network: application to Mexico

Lea, Pousse-Beltran (1), Sophie, Giffard-Roisin (1), Laurence, Audin (1), Pierre, Lacan (2), Théo, Lallemand (1), Glenn, Cougoulat (3), Andres, Nuñez Meneses (2)

(1) Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, IRD, Univ. Gustave Eiffel, ISTerre, 38000 Grenoble, France. Email address of corresponding author: lea.pousse@univ-grenoble-alpes.fr

(2) Centro de Geociencias, Universidad Nacional Autónoma de México

(3) LJK - Laboratoire Jean Kuntzmann, Université Grenoble Alpes, France

Abstract: Fault markers in the landscape (scarps, offset rivers) are records of fault activity. The geomorphological characterization of these markers is currently a time-consuming step with expert-dependent results, often qualitative and with uncertainties that are difficult to estimate. To overcome those issues, we are developing a bayesian supervised machine learning method using convolutional neural networks (CNN) trained on a database of simulated topographic profiles across normal fault scarps, called ScLearn. From a topographic profile the implemented, ScLearn is able to automatically give the scarp height with an uncertainty, and to show the area of the profile containing the scarp. We apply ScLearn for the characterization of normal active faults in the Trans-Mexican Volcanic Belt. From this specific case study, we will explore the progress (computation time, accuracy, uncertainties) that machine learning methods bring to the field of morphotectonics, as well as the current limits (such as bias).

Key words: machine learning, normal fault scarp

Introduction

Fault marker characterization is necessary to understand past fault activity, the physical processes that govern fault rupture, and hence to improve seismic hazard estimates. Among the examples of characterization, the estimation of marker offset by ruptures that have reached the surface is a parameter directly used to estimate fault rates, or the spatial pattern of past ruptures, or the number of ruptures. Currently the estimation of this offset is done by empirical or semi-automatic approaches (i.e. Hodge et al. 2019; Wolfe et al. 2020). These methods are often time consuming and user dependent.

Today, AI techniques have proven to be efficient in performing automatic tasks in geosciences (i.e. Ren et al. 2020), and in particular Convolutional Neural Networks (CNN), a deep learning architecture specifically designed for processing image or series, have made it possible to perform automatic mapping of fractures (Mattéo et al. 2021). Here we propose to automatize the following fundamental task in morphotectonics analyses by evaluating the ability of a CNN to characterize a normal fault scarp (surface offset, position).

Scope

The CNN ScLearn presented is applied to normal fault scarp on topographic profiles perpendicular to the fault (see Figure 1). The profiles are first extracted by the user from terrain elevation models, each one consisting in a series of 500 spatial points. We also assume that the fault is crossing in the middle of the topographic profile, and that the area is only sparsely anthropized. The ScLearn CNN estimates the total surface offset with a model uncertainty. Furthermore, to explain the results, we produce a visualization of the ScLearn CNN intermediate

steps, providing insights of where the ScLearn CNN focuses its attention (i.e. the scarp area).

ScLearn is trained on realistic synthetic topographic profile catalogs. Those synthetic are created by a simulator developed here. Synthetic profiles characteristics are therefore based on the choices made to simulate (see details in the methodological section). Here the simulator create only one fault branch in the middle of the topographic profile, this fault branch ruptures several times creating fault scarp. At each inter-seismic period, the scarp is subjected to some diffuse erosion (Figure 1), and random perturbations are also added to produce a realistic profile.

After training on the synthetic data, ScLearn is tested on Mexico in the Trans-Mexican Volcanic Belt (Figure 2). This region is affected by more than 600 potentially active faults yet less than 5% have been correctly characterized by paleoseismological studies (Núñez Meneses et al. 2021). In this context, a robust and automatic method to characterize the normal fault active scarp in a global, reproducible, robust (not expert-dependent) quantitative way is very valuable and a great step towards a better characterization of the region seismic hazard. To do so, we sampled real data across Ameca-Ahuisculco fault system (Figure 2). We will compare the ScLearn's results with existing empirical and semi-automatic methods: MCSST (Monte Carlo Slip Statistics Toolkit of Wolfe et al. 2020) and SPARTA (Scarp PARameTer Algorithm of Hodge et al. 2019).

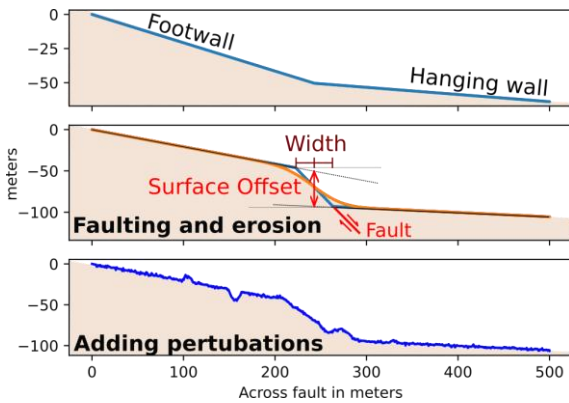


Figure 1: Synthetic normal fault scarp produced by our simulator pipeline to train ScLearn. The total surface offset (in meters) is used as the ground truth label.

Methodology

The scarp “simulator” pipeline:

To train the CNN, we simulate 1000 different topographic profiles with random parameters (Figure 3). They have two slopes, one for the hanging wall and one for the footwall (uniform distribution between 0° and 20°). The simulator breaks a single fault branch, with a dip randomly set between 30° and 50° (unif. distribution). The rupture location is randomly set to ± 5% from the profile center (Gaussian distribution). At each rupture, a fault scarp is created (for a total cumulative throw in a uniform distribution between 0 and 50m). Then, between each rupture a diffusive erosion (Avouac et Peltzer 1993) following Smith et Bretherton (1972)’s equation is simulated with a random constant diffusion (uniform distribution between 0.5 and 10 m²/kyr). The seismic recurrence is also random, but is constrained by a throw velocity between 0.05 and 20 mm/yr. Once the ruptures are produced, we add perturbations to create a realistic morphology using random hyperbolas such as in Hodge et al., (2020) to represent trees, narrow drainage, wide rivers, hills. Finally, we add a Gaussian noise. The resolution of the profile is 500 points. Here we use 1m resolution.

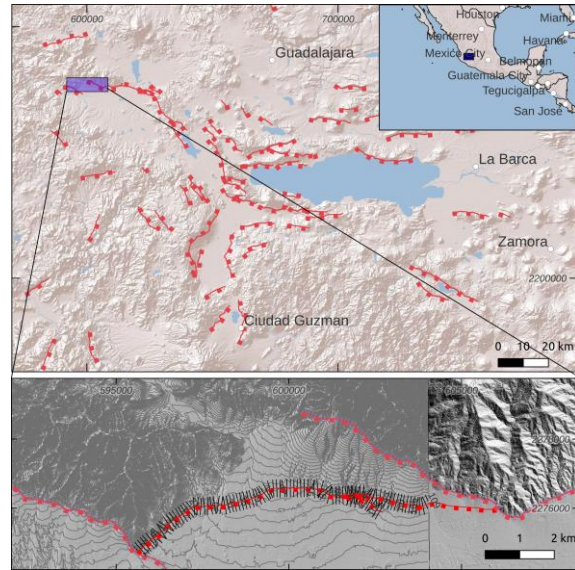


Figure 2: Topographic profile locations across the Ameca-Ahuiscalco fault system in the Trans-Mexican Volcanic Belt. Red lines are active faults from Núñez Meneses et al. (2021), light gray lines are 10m elevations contours, black lines are topographic profiles.

The CNN ScLearn :

In order to weight the final assessments of morphotectonic analyses, uncertainty quantification is crucial, in particular for the scope of probabilistic seismic hazard models. To address it, we estimate the model uncertainty in the CNN using variational Bayesian learning inference (Figure 3). We follow the method of Blundell et al. (2015) which assigns probability distribution on the weights of a neural network. The CNN training then consists in reducing the error between predicted and real offsets while estimating consistent uncertainties. In addition, we also show a visualization of the intermediate steps of the CNN processing (in intermediate feature maps) to understand better how the CNN identifies the fault scarp (Zeiler et al. 2014). The visualization on a processed profile allows making sure that the prediction of the surface offset has been calculated from the fault scarp area. This can be useful in complex areas that are at the limits of the application requirements.

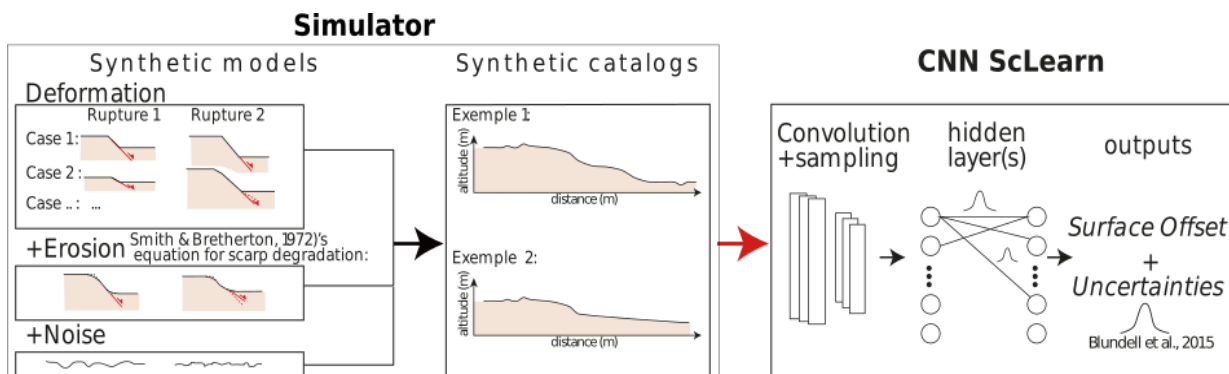


Figure 3: Synthetic normal fault scarp produced by the simulator pipeline to train ScLearn. The total surface offset is fixed as the ground truth.



Results

We first test our CNN model on a second simulated dataset: 1000 profiles created using the same simulator pipeline. Training the CNN on the synthetic data yields a mean accuracy of 4.1 m. Furthermore, we observe that where the predictions are correct the uncertainty bars (3σ) are reduced, while the distant predictions have larger errors allowing encompassing the actual values (Figure 4).

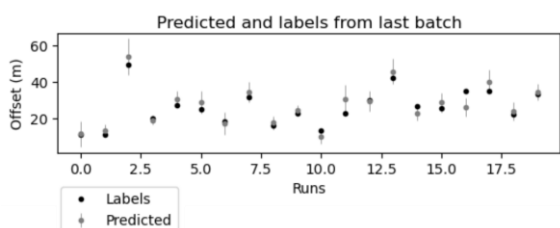


Figure 4: Training of ScLearn. Labels (ground truth) and predictions of the last batch of training. Here uncertainty bars show 3σ .

For validation on real data; we apply ScLearn, SPARTA and MCSST to 110 topographic profiles (Table 1 and Figure 5) extracted along the Ameca-Ahuiscalco fault system. Empirical and semi-automatic methods are time-consuming (several days), do not always produce results and the estimates depend on the user (several choices to make). The ScLearn CNN allows obtaining results within a second, for every profile, independent of the user and therefore reproducible. The ScLearn CNN compared to MCSST it gives a mean residual of 1.9m (median 0.5m) (and a mean 3.0m and a median 1.5m using absolute residual). The residual standard deviation is 4.6m. In addition, there is a constant variance at every level of scarp height. Finally, the CNN predictions overlaps MCSST results at 66.6 % at 3-sigma.

Ameca-Ahuiscalco fault system. * SPARTA does not give a prediction when the profile is too complicated.

Discussions and conclusions

ScLearn is comparable with semi-automatic methods and overlaps the MCSST results on 66.6% of the cases. Although the distribution of residuals is centered around 0.5m (median 0.5 m and mode is 0.5 m), there are complicated profiles where the CNN differs from the MCSST, such as in Figure 6. This is due to the fact that the CNN has been trained by synthetic data that does not take into account this type of configuration: antropized area, with high slope in hanging wall. Although the CNN is automatic, it is always necessary to have an expert overview on the context in which the CNN can be applied or not. Here it depends on the fault scarp model: one fault branch, sparsely antropized, hanging and footwall slope between 0 and 20°. However, once these conditions are fulfilled, the ScLearn allows gaining considerable expert time, to obtain reproducible results, not depending on the user, and to make many measurements with uncertainties. This provides therefore a reliable database on which to perform further fault analysis. This abstract thus shows the potential of the proof of concept of CNN for scarp characterization. Moreover, a future development will consider a larger variety of profiles in order to be able to estimate the scarp in much more contexts. A future study will also estimate other scarp parameters such as two fault branches, antropized area.

	Average surface offset	Mean residual (from the MCSST value)	Proportion of profiles with overlapping uncertainty bars for CNN and MCSST (3σ)
Comparison for all 110 profiles			
MCSST	9.0 m	/	66.6 %
CNN	6.6 m	3.0 m	
Comparison for the subset of profiles (22) estimated by SPARTA*			
MCSST	17.3 m	/	50 %
SPARTA	12.2 m	4.9 m	
CNN	11,4 m	5.6 m	

Table 1: Comparison between ScLearn, SPARTA and MCSST from 110 topographic profiles sampled across the

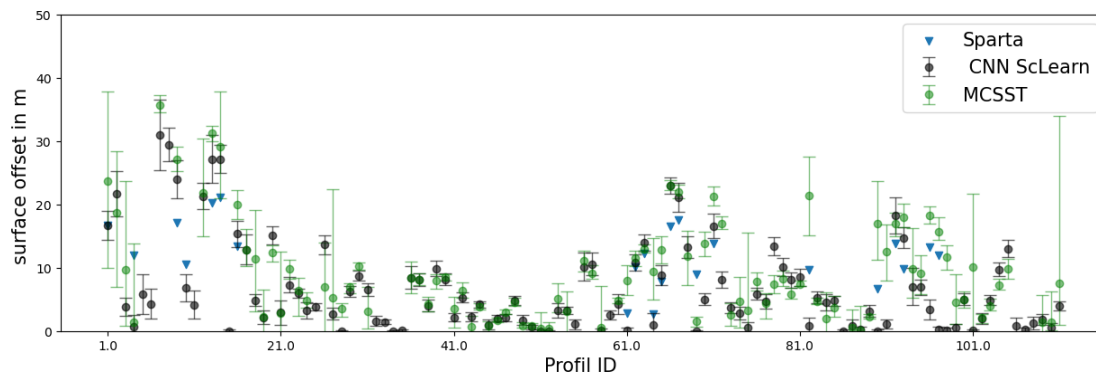


Figure 5: Comparison with ScLearn predictions, SPARTA and MCSST estimations for the 110 topographic profiles across one segment of the Ameca-Ahuiculco fault system. Bars uncertainties show 3σ .

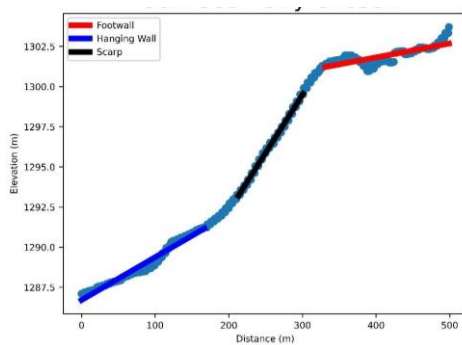


Figure 6: Characterization of the topographic profile 29 through MCSST which allows to estimate a surface offset of 7 ± 0.54 m (2σ), while the CNN predicts a value of 5.6 ± 0.59 m (2σ).

Acknowledgments:

This work has been partially supported by the MIAI@ Grenoble Alpes, (ANR-19-P3IA-0003). Thanks to GRICAD infrastructure (gricad.univ-grenoble-alpes.fr), which is supported by the Grenoble research communities, for the computations. ISTERre is part of Labex OSUG@2020 (ANR10 LABX56). We also thank ISTERre for funding the master internship of Théo Lallemand. We thank the CNES for providing high-resolution optical images. Access to topographic data was granted through the DINAMIS program (<https://dynamis.teledetection.fr/>).

REFERENCES

Avouac, J., et G. Peltzer. 1993. « Active Tectonics in Southern Xinjiang, China: Analysis of Terrace Riser and Normal Fault Scarp Degradation along the Hotan-Qira Fault System ». *Journal of Geophysical Research: Solid Earth* 98 (B12): 21773-807.

Blundell, C., J. Cornebise, K. Kavukcuoglu, et D. Wierstra, 2015. « Weight uncertainty in neural network ». In International conference on machine learning, 1613-22. PMLR.

Hodge, M., J. Biggs, Å. Fagereng, A. Elliott, H. Mdala, et F. Mphemo. 2019. « A Semi-Automated Algorithm to Quantify Scarp Morphology (SPARTA): Application to Normal Faults in Southern Malawi ». *Solid Earth* 10 (1): 27-57.

Hodge, M., J. Biggs, Å. Fagereng, H. Mdala, L. N. J. Wedmore, et J. N. Williams. 2020. « Evidence From High-Resolution Topography for Multiple Earthquakes on High Slip-to-Length Fault Scarps: The Bilila-Mtakataka Fault, Malawi ». *Tectonics* 39 (2)

Mattéo, L., I. Manighetti, Y. Tarabalka, J.-M. Gaucel, M. van den Ende, A. Mercier, O. Tasar, et al. 2021. « Automatic Fault Mapping in Remote Optical Images and Topographic Data With Deep Learning ». *Journal of Geophysical Research: Solid Earth* 126 (4).

Núñez Meneses, A., P. Lacan, F. R. Zúñiga, L. Audin, M. Ortuño, J. Rosas Elguera, R. León-Loya, et V. Márquez. 2021. « First Paleoseismological Results in the Epicentral Area of the Sixteenth Century Ameca Earthquake, Jalisco – México ». *Journal of South American Earth Sciences* 107

Ren, C., C. Hulbert, P. A. Johnson, et B. Rouet-Leduc. 2020. « Chapter Two - Machine Learning and Fault Rupture: A Review ». In *Advances in Geophysics*, édité par Ben Moseley et Lion Krischer, 61:57-107. Machine Learning in Geosciences. Elsevier.

Smith, T. R., et F.. Bretherton. 1972. « Stability and the Conservation of Mass in Drainage Basin Evolution ». *Water Resources Research* 8 (6): 1506-29.

Wolfe, F. D., T. A. Stahl, P. Villamor, et B. Lukovic. 2020. « Short Communication: A Semiautomated Method for Bulk Fault Slip Analysis from Topographic Scarp Profiles ». *Earth Surface Dynamics* 8 (1): 211-19.

Zeiler, M. D., & Fergus, R. 2014. Visualizing and understanding convolutional networks. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer vision – eccv 2014* (pp. 818–833). Cham: Springer International Publishing.