

# LANDSLIDE DETECTION IN CENTRAL AMERICA USING THE DIFFERENTIAL BARE SOIL INDEX

Alexander Ariza \*<sup>a</sup>, Norma Angelica Davila<sup>b</sup>, Hannah Kemper<sup>a</sup>, Gerhard Kemper<sup>c</sup>

<sup>a</sup> UN-SPIDER, Bonn/Germany – aariza@ucm.edu.co

<sup>b</sup> UNAM, Mexico City/Mexico – nadavilah@uaemex.mx

<sup>c</sup> GGS GmbH Speyer/Germany – hannah.kemper@ggs-speyer.de

## Commission III

**KEY WORDS:** Landslides, Differential Bare Soil Index, Object-based detection, Disaster Management

### ABSTRACT:

The increasing availability of EO data from the Copernicus program through its Sentinel satellites of the medium spatial and spectral resolution has generated new applications for risk management and disaster management. The recent growth in the intensity and number of hurricanes and earthquakes has demanded an increase in the monitoring of landslides. It is necessary to monitor large areas at a detailed level, which has previously been unsatisfactory due to its reliance on the interpretation of aerial photographs and the cost of high-resolution images.

Using the differential Bare Soil Index for optical imagery interpretation in combination with cloud-computing in Google Earth Engine is a novel approach. Applying this method on a recent landslide event in Oaxaca in Mexico around 62% of the landslides were detected automatically, however, there is a big potential for improvement. Including NDVI values and considering images with a higher spatial resolution could contribute to the enhancement of landslide detection, as the majority of missed events have a size smaller than half a pixel. Landslide detection in Google Earth Engine has become a promising approach for big data processing and landslide inventory creation.

## 1. INTRODUCTION

A landslide is a movements of rocks, earth or debris downhill categorized on the basis of material and type of movement (European Soil Data Centre (ESDAC)). Statistics reveal that landslides caused almost 30.000 fatalities and 40 billion dollars of economic losses in the period of 2000 to 2014 (Valerio Lo Bello, 2017). Several factors have an influence on the occurrence of landslides and are classified whether the trigger is natural like ground vibrations, groundwater pressure or wildfires or human activities being mining, pipe leakages, constructions and soil excavation (Mohan et al., 2020). Landslides are needed to be considered as an important threat, which is growing due to population growth, urbanization in unstable areas of high slopes, deforestation and rising occurrence of weather extremes that could trigger such events.

The Detection of landslides and landslide inventories have been in discussions among scholars for a while. The methods range from geomorphologic field survey and visual analysis of aerial images to remote sensing based approaches like satellite imagery (Mohan et al., 2020). Remote sensing techniques further include Deep Learning and Machine Learning being promising in the subsequent years (Mohan et al., 2020).

The increasing availability of EO data from programs like Copernicus enhance the possibilities to develop new applications of change detection, landslide inventories or the use of Artificial Intelligence. Image classification and the automation of workflows is a great way to detect changes in the land cover.

Considering such changes the Differential Bare Soil Index is a well-known indicator used in studies about mapping drought affected areas or urbanizing settlements (Li & Chen, 2014; Ma et al., 2016; Nguyen et al., 2021).

High soil brightness and low vegetation coverage help to identify uncovered soil areas that could be the result of landslides (Ma et al., 2016). It is crucial to detect such changes in a reliable way, considering the temporal resolution of such images. Abrupt changes in the land surface and land cover are indeed a great way to detect landslide events.

Converting the DBSI into a workflow, data processing in Google Earth Engine (GEE) is realized for enhancing landslide detection by acceleration through higher computational power.

## 2. MATERIALS AND METHOD

### 2.1 Study area

The study zone is located in one of the world's hotspots affected by intense earthquakes and hurricanes that occurred in 2020 in the districts of San Pedro Totolapa and San Francisco de Ozolotepec in the State of Oaxaca in the state of Mexico (Fig. 1).

### 2.2 Data Sources

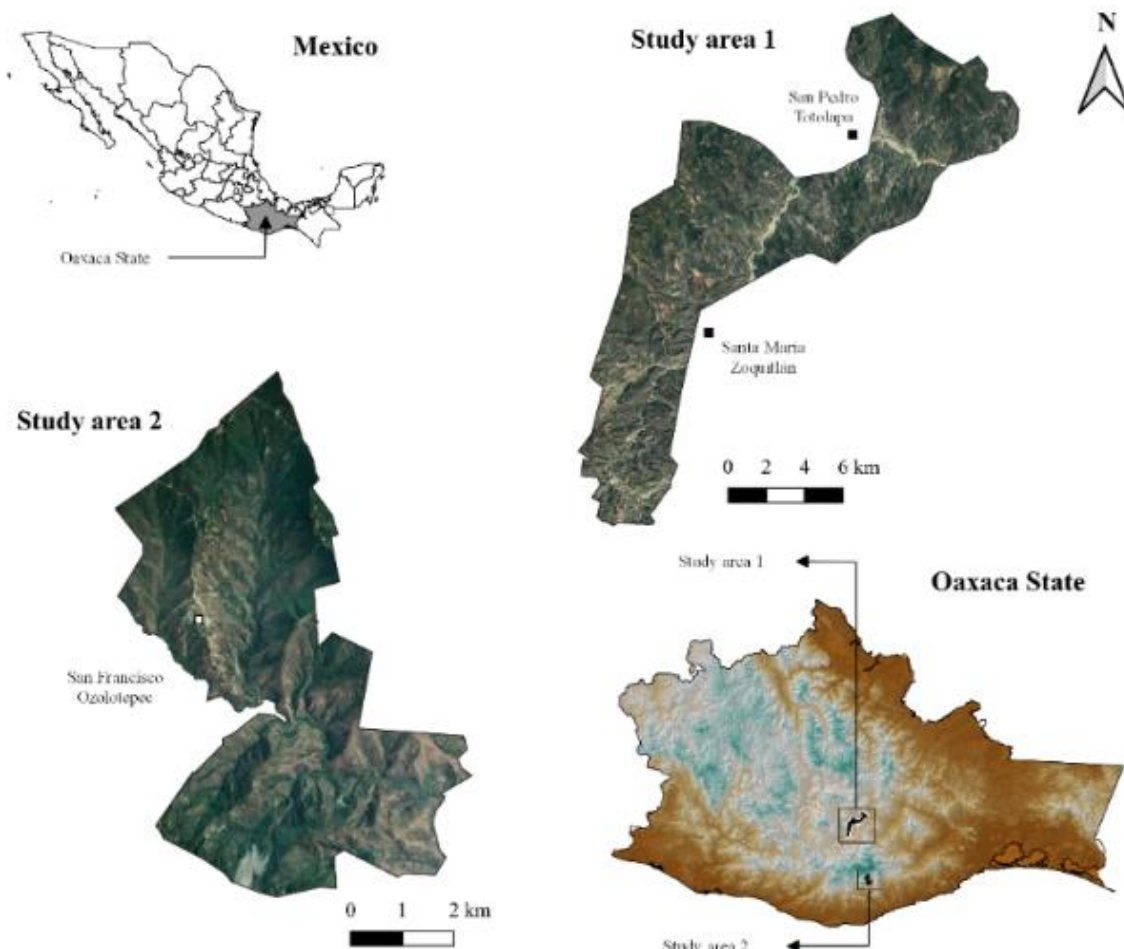
We used the online geospatial and remote sensing cloud computing platform called Google Earth Engine (GEE) (Gorelick et al., 2020), to create Sentinel-2 mosaics. We selected only images between 2020-05-01 and 2020-05-30 for the pre-landslides mosaic, and between 2020-06-20 and 2020-06-28 for post-landslides mosaic in order to

---

\* Contributing author

avoid clouds and cirrus, and yet attempt to include a large enough search window for full spatial coverage of the area. In order to validate the methodology applied to Sentinel-1 data, a landslide inventory was constructed base on SPOT-

7 images (1.5 m resolution) acquired from May before the earthquake. This product implies a fusion process between a high panchromatic resolution (1.5m) and the multispectral imagery (Table 1).



**Figure 1.** Oaxaca State in Mexico

**Table 1.** Details of the satellite data used in this study.

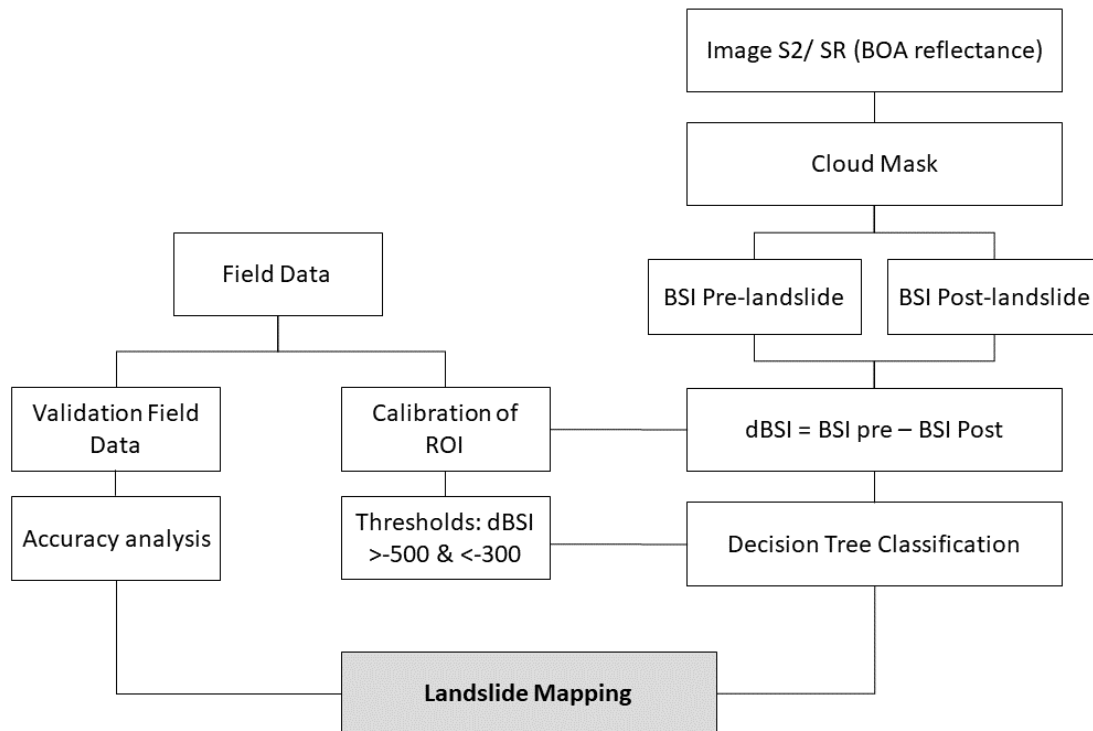
Satellite data	Spatial resolution (m)	Date of acquisition
<b>Sentinel-2 A&amp;B</b>	10 m	03/05/2020, 08/05/2020, 13/05/2020, 18/05/2020, 23/05/2020, 28/05/2020, 22/06/2020, and 27/06/2020
<b>SPOT-7</b>	1.5 m	06/27/2020 and 11/09/2020

### 2.3 Pre-processing of satellite data

The landslide detection that uses space-based data in such cloud platforms has changed the way to see the risk and to manage the disaster dramatically. It aims to facilitate working with big data in the cloud and is therefore an alternative to the utilization of desktop software. The detailed methodology involved in assessing of landslide detection is presented in Fig. 2

The pre-processing of the images involved the filtering of the original image collection by specific dates (pre and post), a cloud masking (mask2clouds) and, finally, the median values. Next, we will select the needed bands to build the spectral indices and to clip the scene images to the study area.

To avoid clouds and cirrus clouds, we used more specifically cloud masks that are calculated from the three QA bands that are present. We used the QA60 band, which is a bitmask band with cloud mask information, where the bits 10 and 11 are clouds and cirrus, respectively.



**Figure 2.** Landslide mapping workflow

### 2.4 Differential Bare Soil Index (dBSI)

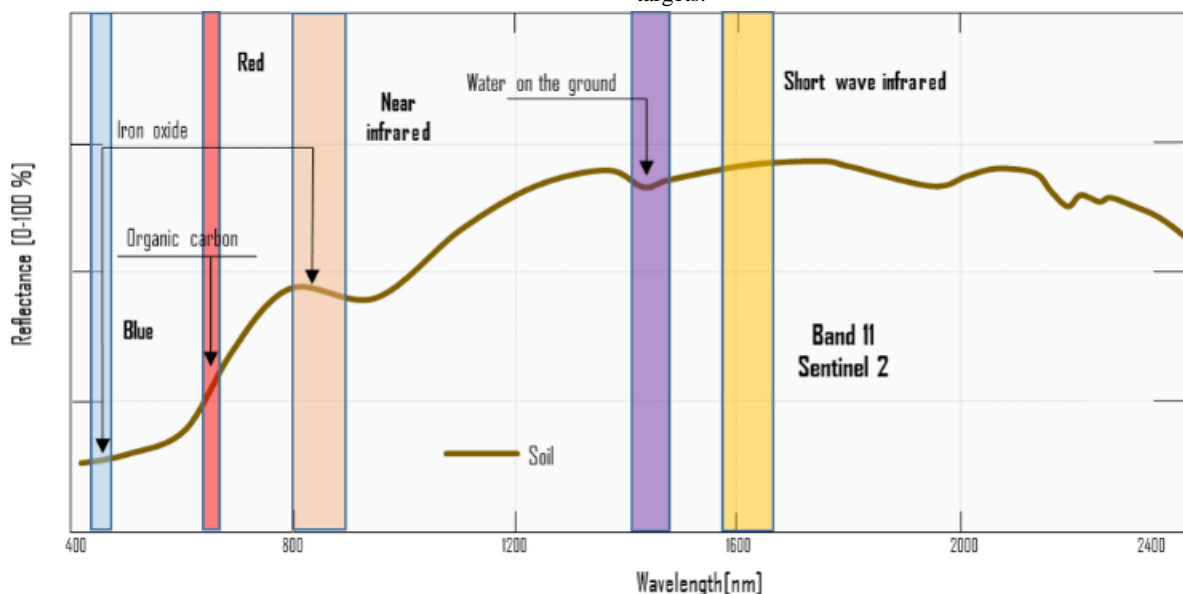
In this study for landslide rapid mapping, we propose to base the response on the Bare Soil Index (BSI) (Roy et al., 1996; Rikimaru & Miyatake, 1997) for the detection of the traces of the soil movements. The BSI is a numerical indicator that combines blue, red, near-infrared and short wave infrared spectral bands to capture soil variations (Figure 3).

These spectral bands are used in a normalized manner. The short wave infrared and the red spectral bands are used to quantify the soil mineral composition, while the blue and the near-infrared spectral bands are used to enhance the

presence of vegetation. BSI can be used in numerous remote sensing applications such as soil mapping, crop identification (in combination with NDVI), etc. To calculate the BSI, the following equation is used (Equation 1):

$$BSI_{S2} = \frac{(SWIR+Red)-(NIR+Blue)}{(SWIR+Red)+(NIR+Blue)} \quad (1)$$

However, the use of the BSI index itself cannot characterize the landslide configurations on terrain, due to after landslide event, the spectral response changes shortly especially over zones with fast reactivation in the vegetation cover. Therefore, to improve the interpretation of the results we use a temporary version or dBSI. This index takes the spectral response before and after the events, to obtain better results of the values of a BSI date linked to the slip targets.



**Figure 3.** The spectral signature of bare soils associated

This index is calculated by subtracting the subsequent values of the BSI from the initial values (Equation 2). In the end, the areas with the presence of landslides will appear with negative values. These results are scaled by a factor of 1000 to facilitate their interpretation and range detection.

$$dBSI = (BSI \text{ pre Landslide}) - (BSI \text{ Post Landslide}) \times 1000 \quad (2)$$

## 2.5 Detection approach

The first step is the calibration of the differential Bare soil index (dBSI), whose values were taken on areas of interest on the field (AOI) for calibration to extract the value of the interval of the related zones with landslides. Mapping is created by using a tree classifier that identifies areas of bare soil related to landslides (Fig. 2).

In this way, we apply statistical analysis of the value of dBSI on the AOI zones, to extract the value of the interval. The value of the slip landslide thresholds will be determined by the mean of the dBSI over the AOI + and - one standard deviation (Equation 3):

$$\left[ Thr_{max} = M_{dBSI} + SD_{dBSI}; Thr_{min} = M_{dBSI} - SD_{dBSI} \right] \quad (3)$$

where  $Thr_{max}$  and  $Thr_{min}$  is the interval of detection, ( $M_{dBSI}$ ) is the mean of dBSI and ( $SD_{dBSI}$ ) standard deviation value on AIO zones respectively.

## 2.6 Accuracy assessment

The accuracy of the classification results derived from the proposed technique by the dBSI, can be expressed using an error matrix (Congalton, 1991), or in terms of branching factor, lack factor, detection percentage and quality percentage (Lee et al., 2003). While the first method is effective when the purpose is to evaluate the accuracy of the classification result of the whole image/area, the second method is effective when the goal is to detect only targets such as landslides from the image. Therefore, in this study, the latter method was selected to estimate the precision figures using the following equations (Lee et al., 2003; Tapas et al., 2012).

$$Branching \ factor \ (BF) = \frac{False \ positive}{True \ positive} \quad (4)$$

$$Miss \ factor \ (MF) = \frac{False \ negative}{True \ positive} \quad (5)$$

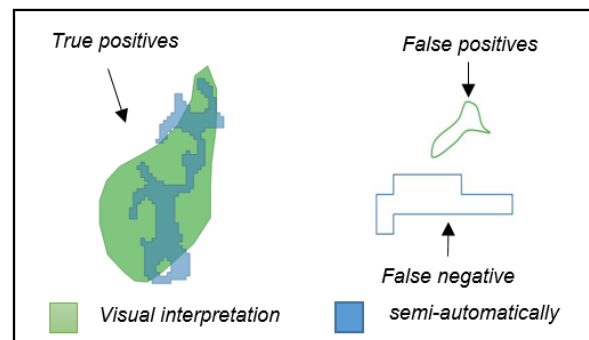
$$Detection \ percentage = \frac{True \ positive}{(True \ positive + False \ negative)} * 100 \quad (6)$$

$$Quality \ percentage = \frac{True \ positive}{(True \ positive + False \ negative + False \ positive)} * 100 \quad (7)$$

The branching and missing factors indicate two types of potential errors, that is, false positives and false negatives.

Those factors can be generated during the automatic detection process. The detection percentage indicates the landslides correctly identified by the proposed methodology. Percent quality, which is the strictest measure of precision among the four previous precision estimates, indicates the probability that a landslide identified by the methodology is true (Tapas et al., 2012).

True positives, false positives and negatives were calculated by comparing the landslide inventories created semi-automatically by the proposed methodology with the reference inventories, which were created manually using the visual interpretation technique on high resolutions images SPOT-7 (Fig. 4).



**Figure 4.** Estimation of the precision: true positives (a), false positives (b) and false negatives (c), for assessing the accuracy of dBSI in landslides.

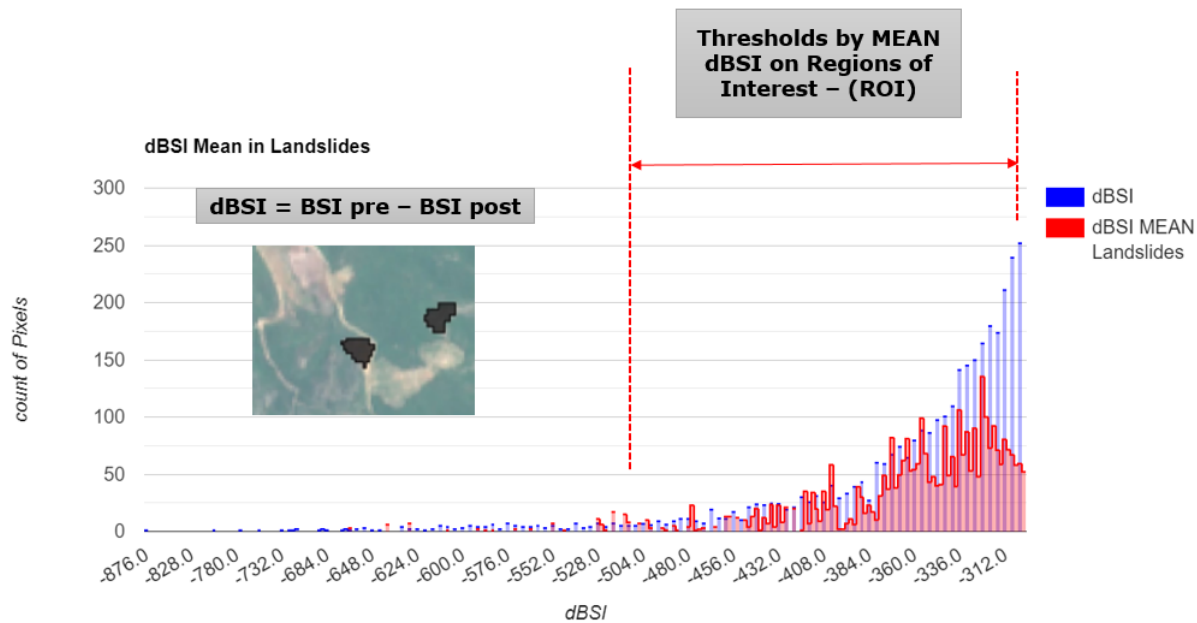
## 3. RESULTS AND VALIDATION

From descriptive statistics analysis of the AOI values observed in the field, it was possible to determine the mean (-336.95) and SD (-336.95) values respectively. This indicates that the majority of pixels related to landslides presents BSI values ( $> 0.7$ ), and in consequence which determines that most of the areas where landslides occur show values between dBSI of -158.17 to 495.12 in the dBSI approximately (Fig. 5).

The results of the accuracy assessment of the dBSI for landslide detection, were obtained for the total number of landslides detected by the decision tree based semi-automatic classification. Wise accuracy figures for all landslides in the Oaxaca state area and their classification into the different precision types measures are provided in Table 2.

**Table 2.** Accuracy of dBSI landside detection classifier in the Oaxaca state for the four different precision types

Analysis of area	3528,66	
Visual detection SPOT 7 (ha)	12,99	
dBSI detection Sentinel 2 (ha)	10,94	
Overlap Analysis	(ha)	%
true positives	6,80	52,35
false positives	6,19	47,65
false negatives	4,14	37,86
Differential bared Soil Index (dBSI)	factor	Value
Branching factor	BF	0.91
Miss factor	MF	0.61
Detection percentage (%)	DP	62.14
Quality percentage (%)	QP	39.69



**Figure 5.** Landslide thresholds of the dBSI

The results obtained were validated through high-resolution images from SPOT 7. The total area analysed summed up to 3,528.66 ha, where 10.9 ha were detected by the dBSI semi-automatic method. Respectively a visual interpretation of the AOI using the SPOT 7 images a total amount of 12.9 ha of landslide affected areas were found, evidencing that the proposed method is capable of detecting approximately 62% of landslides.

The main omissions were those being less than approximately half a pixel wide. This demonstrates the potential implementation of the dBSI for rapid mapping of landslides in earthquakes or intense hurricanes. Using Sentinel-2 images that are free of shadows, clouds and cirrus seem to provide enough information to reliably detect landslide events.

The performance of the methodology for landslide detection is moderately good. Particularly Sentinel 2 images, both for the detection capacity, if not also for the Miss Factor (MF) are reliable data sources for applying the methodology. The miss factor is further lower compared to other methodologies (see Tapas et al., 2011). This is mainly due to the identification of changes in vegetation as false positives.

## CONCLUSION

The differential Bare Soil Index proposed (dBSI) in this study, was developed to detect the traces of the ground movements after earthquake. It can be seen as a complementary method to support traditional interferometric-SAR analysis.

The preliminary results showed that the use of dBSI is sufficient to identify recent traces of soil movements which

are useful for a rapid mapping response during a disaster event. In this case, the use of the dBSI index itself has not been able to characterize the landslide configurations on terrain, due to the short reactivation time of the vegetation cover in the study area and the rapid spectral changes related to this phenomenon. Research areas of quick vegetation recovery, like Mexico would therefore need to be considered in an additional study, whereas research areas in different vegetation zones of slower vegetation recovery could be a suitable AOI to validate the method.

Therefore, to improve the interpretation of the results, we recommended the use of another remote sensing based index like the NDVI or SAVI in the analysis (each with images from pre and post events) to eliminate the vegetation change effects and get better outcomes of BSI values linked to landslide targets.

Unfortunately, it also highlights certain clouds, cirrus clouds, making it difficult to separate areas of bare ground from other covers. These results are highly dependent on vegetation change, seasonal agriculture, and cloud-free images. Considering the study area certain measurements would need to be taken care of to assure detection reliability.

## ACKNOWLEDGMENTS

The completion of field observation on this paper could not have been possible without the participation and assistance of other members of this project. We would like to express our gratitude towards all participants. A special thanks is dedicated to Bonn Office of UN SPI-DER.

## REFERENCES

- Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37 (1), 35–46.
- European Soil Data Centre (ESDAC): Landslides. <https://esdac.jrc.ec.europa.eu/>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Lee, D.S., Shan, J., Bethel, J.S., 2003. Class-guided building extraction from Ikonos imagery. *Photogrammetric Engineering and Remote Sensing* 69 (2), 143–150.
- Li, S.; Chen, X., 2014. A new bare-soil index for rapid mapping developing areas using LANDSAT 8 data. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XL-4,139–144.
- Ma, H.; Xinwen, C.; Chen, L.; Zhang, H.; Xiong, H., 2016. Automatic identification of shallow landslides based on Wordview2 remote sensing images. *Journal of Applied Remote Sensing* 10.1.
- Mohan, Amrita; Singh, Amit Kumar; Kumar, Basant; Dwivedi, Ramji. 2020. Review on remote sensing methods for landslide detection using machine and deep learning. *Trans Emerging Tel Tech*. DOI: 10.1002/ett.3998.
- Nguyen, Can Trong; Chidthaisong, Amnat; Kieu Diem, Phan; Huo, Lian-Zhi, 2021. A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. *Land* 10 (3).
- Valerio Lo Bello (2017): *Landslide Hazard and Risk Assessment*. Ed. UNISDR.
- Rikimaru, A., & Miyatake, S., 1997. Development of forest canopy density mapping and monitoring model using indices of vegetation, bare soil and shadow, presented paper for the 18th ACRS. Kuala Lumpur, Malaysia.
- Rodríguez-Pérez, Q., Márquez-Ramírez, V. H., and Zúñiga, F. R. 2020: Seismicity characterization of oceanic earthquakes in the Mexican territory, *Solid Earth*, 11, 791–806, <https://doi.org/10.5194/se-11-791-2020>, 2020.
- Roy, P. S., Sharma, K. P., & Jain, A., 1996. Stratification of density in dry deciduous forest using satellite remote sensing digital data—An approach based on spectral indices. *Journal of biosciences*, 21(5), 723–734.
- Tapas R. Martha, Norman Kerle, Cees J. van Westen, Victor Jetten, K. Vinod Kumar, 2012. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 67.105-119.
- van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: an overview. *Engineering Geology* 102 (3–4), 112–131.