Deep Learning for Mapping Forests with TanDEM-X

José-Luis Bueso-Bello, Daniel Carcereri, Carolina González, Michele Martone, and Paola Rizzoli
Microwaves and Radar Institute, German Aerospace Center (DLR), 82234 Wessling, Germany

Abstract

In a bistatic SAR system such as TanDEM-X, characterized by the absence of temporal decorrelation, the interferometric coherence adds valuable information to the common amplitude images, typically acquired by monostatic SAR systems. The interferometric SAR dataset, acquired to generate the TanDEM-X global Digital Elevation Model (DEM), represents a unique data source to derive land classification maps at global scale, such as the TanDEM-X Forest/Non-Forest Map and the TanDEM-X Water Body Layer. Both maps have as main input the interferometric coherence and are based on a supervised fuzzy clustering algorithm and on the watershed segmentation algorithm, respectively. Single images are classified with the corresponding algorithm and a final weighting mosaicking strategy of overlapping coverages is necessary to improve the final accuracy of the generated classification maps. In this work, we now investigate the capabilities of using a state-of-the-art convolutional neural network (CNN) with TanDEM-X interferometric data for forest and water mapping at large scale. An ad-hoc training strategy has been developed to train a U-Net-like architecture, which aims at balancing the training data set with respect all possible acquisition geometries that can be found in TanDEM-X acquisitions. The Amazon rainforest has been used as region of interest (ROI) to compare the improvement in image classification with respect to the reference fuzzy-clustering approach. On forest classification, a significant performance improvement with respect to the clustering approach, with an f-score increase of 0.13, has been measured. This classification improvement of the forested areas, as well as the capabilities of the U-Net to accurately classify water bodies without the necessity of mosaicking overlapping acquisitions to improve the final classification accuracy, make it possible to generate up to three time-tagged mosaics over the Amazon rainforest by utilizing the nominal TanDEM-X acquisitions between 2011 and 2017.

1 Introduction

The TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement) mission currently maps the Earth's surface providing single images at high resolution from the recorded backscattered signal. The advantage of flying two satellites in close formation, constituting a single-pass interferometric SAR (InSAR) system, adds valuable information to the amplitude data, such as the the interferometric phase and the interferometric coherence.

The TanDEM-X interferometric phase has been exploited for the generation of the global DEM at a spatial resolution of 12 m x 12 m [1]. The interferometric coherence was the main input for the generation of the global TanDEM-X Water Body Layer (WBL), based on a watershed segmentation algorithm [2]. The interferometric coherence, defined as the normalized complex correlation coefficient between the two InSAR acquisitions, gives information about the amount of noise in the interferograms and is sensitive to different decorrelation sources such as signal-to-noise ratio losses and volume scattering mechanisms. This last parameter, the volume correlation factor (γ_{Vol}), was the main input feature for the generation of the global TanDEM-X Forest/Non-Forest (FNF) map using a fuzzy clustering algorithm [3].

For information extraction and forest mapping in the context of TanDEM-X SAR images, the potentials of deep learning (DL) have been demonstrated in [4], where three state-of-the-art CNN architectures were compared, showing very promising performance. The U-Net [5] solution demonstrated to be the most effective one. This previous study investigated different combinations of TanDEM-X SAR input features and was performed on single images at full-resolution on a temperate forest area where very high-resolution reference data from lidar and optical data where available.

Since our focus lies on the generation of global and large scale land classification maps with TanDEM-X data, we have extended the input features of the previous work to consider the different acquisitions geometries from the TanDEM-X InSAR system. We have kept the sigmoid as activation function to generate only binary classification maps. The U-Net has been selected to classify both, forest and water, separately, by using two different training data sets. Both classification outputs have been then combined in a final product with three layers: forest, non-forest, and water. For testing our approach, we have considered the unique environment provided by the Amazon rainforest and the TanDEM-X images acquired over this region of interest between 2011 and 2017.

The paper is organized as follows: in Section 2 a brief introduction on the TanDEM-X mission and the used dataset is presented. In Section 3 the baseline approaches for forest mapping and water bodies detection with TanDEM-X, including previous investigations with DL methods, are summarized. The further developments done in the present analysis to extend the land classification of TanDEM-X images on a large scale are presented in Section 4. The obtained results over the Amazon rainforest, including some comparisons with other contemporary land classification maps, are resumed in Section 5. Finally, Section 6 summarizes the advances done in the actual work and draws some guidelines for further studies.

2 TanDEM-X mission and dataset

Since the end of 2010, the two twin satellites TerraSAR-X and TanDEM-X, flying in a controlled close formation, have been operationally acquiring interferometric SAR images in bistatic configuration with a typical resolution (azimuth and range) of about 3 m. The main objective of the TanDEM-X mission was the generation of a global and consistent DEM at a final independent posting of 12 m x 12 m. In order to achieve such a demanding goal and improve the overall product accuracy, it has been necessary to acquire at least two global coverages of the Earth's landmasses and multiple acquisitions over selected regions, such as mountainous and forested areas. The global TanDEM-X DEM has been finalized and delivered in September 2016 [1].

During the whole mission, global mosaics from TanDEM-X quicklook data have been exploited as a helpful tool for performance monitoring and acquisition planning optimization [6]. The TanDEM-X quicklook images are generated at a ground resolution of 50 m x 50 m by applying a spatial averaging process to the corresponding operational TanDEM-X interferometric data at full resolution (12 m x 12 m). Working with such quicklook data allows for the exploitation of the TanDEM-X dataset at a global scale with a limited computational burden and a huge reduction in terms of data volume, memory usage, and processing time [2, 3]. The TanDEM-X quicklook dataset, including backscattering and interferometric coherence, have been used during the present investigations to generate large-scale maps and test the proposed architectures.

3 Baseline classification approaches

3.1 Forest mapping with TanDEM-X

In a bistatic SAR system such as TanDEM-X, with the absence of temporal decorrelation, the γ_{Vol} adds valuable information for discriminating between forested and non-forested areas, thanks to its sensitivity to vegetated areas. γ_{Vol} was selected as main feature for the generation of the global TanDEM-X FNF map, based on a supervised geometry-dependent fuzzy clustering classification approach as explained in [3]. This approach represents the baseline approach for forest mapping with TanDEM-X.

For the generation of the global TanDEM-X FNF map, due to the variability of the interferometric coherence at X band among different forest types, mainly due to changes in forest structure, density, and tree height, it was necessary to derive three different sets of cluster centers, depending on the specific type of forest. Moerover, only acquisitions with a height of ambiguity (h_{amb}) < 100 m were classified in this case. The h_{amb} is a good descriptor of the acquisition geometry and is related to the phase-to-height sensitivity in the interferogram. On images acquired with higher h_{amb} values, the forest classification was too ambiguous due to the acquisition geometry characterized by smaller perpendicular baselines between the satellites, which reduce the sensitivity of γ_{Vol} on vegetated areas. The classified images were mosaicked in a final single forest map following a weighted mosaicking process, which was necessary to improve the final accuracy of the product. On this map, external references were necessary to filter out water bodies as well as urban areas [3].

For image segmentation and forest mapping in the context of TanDEM-X bistatic SAR images, the potentials of DL were demonstrated in [4], where three different state-of-the-art CNN architectures were utilized, namely ResNet, DenseNet, and U-Net, while training and testing were performed on a limited dataset of TanDEM-X full resolution images acquired over the state of Pennsylvania, USA. Additionally, different combinations of input features were investigated as well, including, backscatter, coherence, volume decorrelation, and local incidence angle. Overall, the U-Net [5] demonstrated to be the most effective one for such a task, achieving the best performance among all investigated architectures. The specific implementation of the U-Net for TanDEM-X works on four scale levels, with two chained convolutional layers located on both encoder and decoder sides at each level. The network head has an additional 1 × 1 convolution, used to map 64 features in a single probability channel. This U-Net implementation constitutes the starting point for the investigations presented in this paper.

Regarding the performance assessment, the f1-score (f1) has been selected as main accuracy metric, since it is specially useful when dealing with imbalanced datasets, where the accuracy may present too optimistic results. The f-score is defined as the harmonic mean of precision and recall of a model and ranges between 0 and 1. It can be expressed as

$$\text{f1} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}}, \tag{1}$$

where the terms on the right-hand side of the equation are the classical and widespread measures used for detection: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The overall accuracy (OA) is used, too. It represents the overall correctly classified pixels, with respect to the total number of classified pixels, considering all land cover classes and is defined as:

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

3.2 Water mapping with TanDEM-X

Two main approaches have been followed for the detection of water bodies in TanDEM-X images. Together with the final TanDEM-X DEM product, some auxiliary maps, such as the Coverage Mask (COV) and the Water Indication Mask (WAM), are provided as well. The purpose of the WAM is to indicate possible noisy water bodies surfaces that remains in the final DEM layer. The WAM approach is based on a thresholding of the amplitude and coherence images together with counting the possible water occurrences [9]. In SAR images, geometric distortions, such as shadow and layover, result in low coherence areas. This effect specially occurs over mountainous terrain and may lead to a wrong classification of water bodies when using approaches based on coherence thresholding such as the WAM.

A possibility to accurately detect water bodies on TanDEM-X images is presented in [2]. For the generation of the global TanDEM-X WBL, a classification method based on the watershed algorithm has been developed. The interferometric coherence is used as the main input. Auxiliary derived maps from the TanDEM-X interferometric dataset, such as the shadow and layover maps [10], are used to improve the classification, specially over mountainous regions. Other auxiliary information is used to deal with low coherence due to forested areas or heavy rain clouds, even visible at X band. After the classification of water bodies on each TanDEM-X single image, they are combined and mosaicked to derive the final global WBL map, including information about the changes in time of the detected water bodies in the temporary and permanent WBL layers.

4 Deep learning methods for forest and water mapping with TanDEM-X

The U-Net presented in [4] has been used as main prototype for the present study. It uses as activation function in the output layer a sigmoid, allowing for a binary classification. We have kept for the present study the same activation layer. Taking advantage of the unique environment presented by the Amazon rainforest, we have used two different U-Nets, to classify forested areas and to classify water bodies, separately.

Both U-Nets have been adapted to work with a larger number of acquisitions than in the previous study, which represent different land cover types, account for a variability of the pixels belonging to each class in each image, and have been acquired with very different geometries. To account for such variability in the TanDEM-X acquisitions, it has been necessary to extend the number of considered input features to the U-Net. Together with the backscattering coefficient, the total interferometric coherence, and the volume correlation factor, used in the previous studies, the local incidence angle and the h_{amb} have been considered as major descriptors of the variability in the TanDEM-X acquisition geometries.

Both U-Nets have been trained from scratch to avoid any type of transfer learning, taking advantage of the increased amount of available data with respect to previous studies. An ad-hoc training strategy for model generalization on all different acquisition geometries has been developed. Different training sets have been used to train the U-Net for fores mapping and for water detection. As external reference, a forest map based on Landsat imagery from 2010 has been used [11]. Accounting for the different acquisition geometries, the acquisition incidence angles θ_i have been divided in three main ranges, as for the TanDEM-X FNF map [3]: a) $\theta_\mathrm{i} < 35°$, b) $\theta_\mathrm{i} \in< [35°, 45°]$, and c) $\theta_\mathrm{i} > 45°$. The h_{amb}, showing typical values values between 20 m and 120 m, has been divided in ranges with an step of 2 m. Mainly acquisitions acquired in 2011 and 2012 have been considered for the training, to assure a good classification agreement with the considered reference map of 2010. Up to 5 images per range (both h_{amb} and θ_i) have been selected, where possible.

For forest mapping, the images have been chosen with a forest content between 30% and 70%, in order to mitigate class imbalance. The number of pixels flagged as forest, with respect to the total number of valid pixels in each image, has been derived from the reference data. For the training of the U-Net for forest classification, a total of 455 TanDEM-X quicklook images has been selected. 20% of these images have been reserved for validation during the training process. To train the U-Net for water detection, a separated set of 376 TanDEM-X images has been used, characterized by a mean water content of 35% according to the reference data and representative of the different types of water bodies, such as open water, lakes and rivers.

The trained CNN has been tested on a set of 976 additional quicklook images with at least 10% forest content and representative for the considered ranges of h_{amb} values and acquisition incidence angles. The test dataset was clearly not used during the training phase. Regarding the reference baseline, the mean f1-score over the test dataset for the CNN forest map is 0.88 and for the clustering approach is 0.75.

Figure 1 depicts the performance obtained by comparing the forest classification with the CNN and the clustering approach described in [3] as a function of the percentage of forest in the image according to the Landsat map used as reference. An improvement of the detection of forest with the CNN is overall observable, but specially noticeable over dense forested areas (percentage of forest samples > 70%). The corresponding performance is summarized in Table 1.

5 Results over the Amazonas

The Amazon rainforest is the largest moist broadleaf tropical forest on the planet, extending by about 5.5 million km^2. It plays a key role for regulating environmental processes on Earth. Indeed, it has a crucial role in the carbon and water cycles and acts as climate regulator, e.g. by producing about 20% of the Earth's oxygen and contrasting global warming [12]. This unique environment offers the ideal scenario to assess the performance of the developed classification methods with TanDEM-X images and DL methods.

Figure 2 shows an example of the reference map (Landsat), TanDEM-X baseline approaches (FNF+WBL), and the U-Net CNN approach. Detected forest is shown in green,

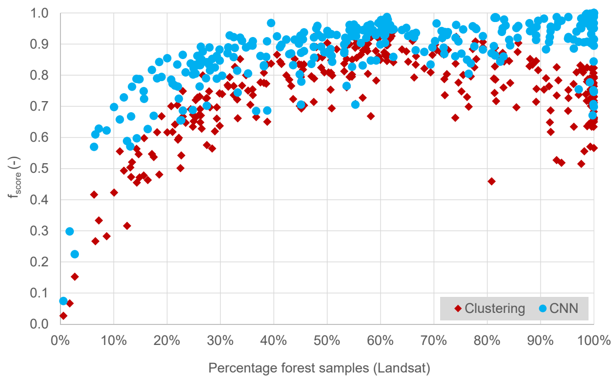


Figure 1 F1-score for the forest class as a function of the forest density. Results obtained with the U-Net are compared with the clustering ones. Forest samples estimated from Landsat reference data.

Table 1 Performance accuracy corresponding to the single image depicted in Figure 2.

		F1-score		
	OA	Forest	Non-Forest	Water
Samples [%] (Landsat)		40.60	54.55	4.85
U-Net CNN	93.14	92.44	93.59	94.11
FNF + WBL	87.75	87.27	88.08	88.57

non-forest areas are depicted in white, and water is represented in light blue. The corresponding performance estimation, including the overall accuracy (OA) and the f1-score for each land cover class, is shown in Table 1. In general, thanks to its two-dimensional capabilities, the U-Net performs better, it is able to generate forest maps with closed forest areas and cleaner non-forest regions. Moreover, the water classification looks more reliable, too.

More than 20,000 TanDEM-X images acquired over the Amazon rainforest between end of 2010 and beginning of 2017 have been classified with the proposed methodology. After combination of the binary outputs of the U-Net with forest/non-forest and water/non-water, a final classification map is available for each TanDEM-X image. This allows for the generation of three time-tagged mosaics over the Amazon rainforest, skipping the weighted mosaicking process of overlapping acquisitions used in the clustering approach, as well as avoiding the use of external reference maps to filter out water bodies. Figure 3 shows the three obtained mosaics. The first mosaic (a) corresponds to the first TanDEM-X global acquisition from end of 2010 up to March 2012. The second mosaic (b) represents the forest map from the TanDEM-X images acquired during the second global coverage (from March 2012 up to the end of 2013), and the last mosaic (c) has mainly input images from 2016 and 2017. Note, that no TanDEM-X global coverage is available for this last mosaic.

For the validation of the results obtained with the U-Net, the generated TanDEM-X mosaics of 2013 and 2016 have been compared with a global land cover map, based on op-

		TDM 2013	TDM 2016
	Nr. Geocells	591	316
	OA	90.58%	87.89%
F1-score	Forest	86.60%	87.07%
F1-score	Water	74.31%	72.89
F1-score	Non-Forest	67.63%	67.34%

Table 2 Performance comparison between TanDEM-X mosaics and FROM-GLC. A geocell is considered as a (1° x 1°) in latitude/longitude coordinates.

tical imagery, the Finer Resolution Observation and Monitoring - Global Land Cover (FROM-GLC) [13]. This map has a resolution of 10 m x 10 m, it is based on Landsat imagery up to to 2015 and it was updated to 2017 with Sentinel-2 images. The deforestation in the Amazon rainforest accounts for a certain variability and differences on the classification maps acquired at different epochs. Nevertheless, as shown in Figure 4 for a geocell over the Amazon rainforest with on-going deforestation activities (Mato Grosso State, Brazil), a good agreement between TanDEM-X and FROM-GLC, as well as with an optical Landsat image is appreciable.

For the generated TanDEM-X large-scale mosaics, the complete performance accuracy is shown in Table 2. The overall accuracy (OA) approaches the 90% on both TanDEM-X mosaics. Taking into account the possible differences in both maps due to the time of acquisition of the images, for the forest class a high agreement, with a f1-score close to 87% is observed in both mosaics, pointing out the potential of DL methods for land cover classification with TanDEM-X. With respect to the water class, the U-Net seems to suffer from the resolution of the used TanDEM-X quicklook images (50 m x 50 m) to detect narrow rivers, which are also typical in the Amazon rainforest ecosystem. For the non-forest class, the f-score is lower than for the other classes. Specially in areas close to the rivers, where the vegetation presents different characteristics as a dense forest, as well as areas with secondary forest, they are seen at X band as non-forest areas, but classified as forest on the reference map. TanDEM-X images acquired over these areas show a low classification accuracy with respect to the reference data but a high agreement between the U-Net classification and the TanDEM-X FNF map obtained with the clustering approach.

The different time-tagged mosaics over the Amazon rainforest opens new possibilities to a reliable and accurate deforestation monitoring with TanDEM-Xdata. An example is shown in Figure 5 for a geocell over the Rondônia State, Brazil, with visible deforested areas (red) between 2011 and 2016. Typical deforestation patterns are well recognized. Between 2011 and 2013, a total of 14500 ha were deforested over this geocell.

6 Conclusions and outlook

In this paper we present the possible improvements achievable when using DL methods with TanDEM-X imagery for

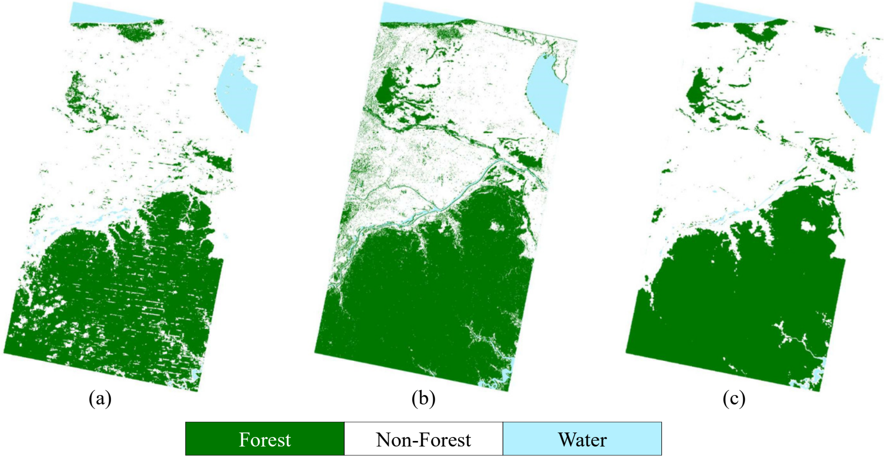


Figure 2 Classification of a TanDEM-X image acquired on Januar 13th, 2013 in the upper part of the Amazon river, which includes all three classes: forest, non-forest and water. (a) Reference map based on Landsat data, (b) classification obtained with both baseline approaches for forest and water mapping with TanDEM-X, (c) classification reached with TanDEM-X and DL methods.

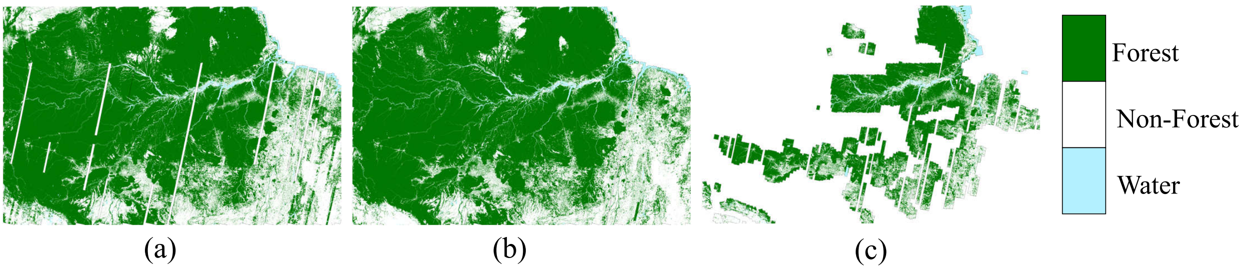


Figure 3 Time-tagged mosaics, including forest, non-forest and water land cover classes, obtained with TanDEM-X data and DL methods. (a) Mosaic obtained with TanDEM-Xdata mainly acquired in 2011 (first global coverage), (b) mosaic obtained with TanDEM-Xdata mainly acquired in 2013 (second global coverage), (c) mosaic obtained with TanDEM-Xdata mainly acquired in 2016 (no global coverage available).

performing forest and water mapping in the Amazon rainforest. Previous works have been extended by considering new input features and utilizing a state-of-the-art CNN to classify forest/non-forest, as well as water bodies on large-scale maps. On single TanDEM-X images acquired over the Amazon rainforest, the f1-score shows an improvement of 0.13 with respect to the clustering approach used for the generation of the global TanDEM-X FNF map. With the presented approach based on DL methods it is possible to accurately classified TanDEM-X images acquired with h_{amb} > 100 m. This improvement makes it possible to generate three time-tagged mosaics of the whole Amazon basin, just by averaging the single image maps classified by the ad-hoc trained CNN. No mosaicking weights were necessary, thanks to the increased accuracy of the single-scene classification obtained with the CNN. Moreover, no external references are necessary either to filter out water bodies, as done for the global TanDEM-X FNF map. The new approach based on DL techniques opens new possibilities to a reliable, very fast, and more accurate deforestation monitoring with TanDEM-Xdata.

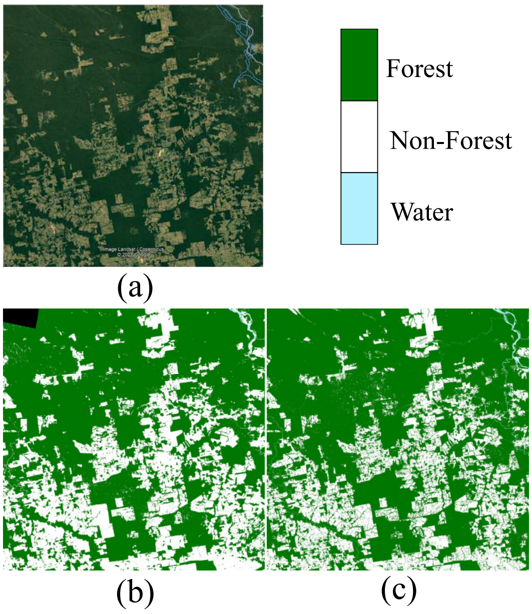
7 Literature

[1] Rizzoli, P.: Martone, M.: Gonzalez, C.: Wecklich, C.: Borla Tridon, D.: Bräutigam, B.: Bachmann, M.: Schulze, D.: Fritz, T.: Huber, M.: Wessel, B.: Krieger, G.: Zink, M.: Moreira, A.: Generation and performance assessment of the global TanDEM-X digital elevation model, ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 132, Oct. 2017, pp. 119-139

[2] Bueso-Bello, J.L.: Sica, F.: Valdo, P.: Pulella, A.: Posovszky, P.: González, C.: Martone, M.: Rizzoli, P.: The TanDEM-X Global Water Body Layer. 13th European Conference on Synthetic Aperture Radar, EUSAR, 2021.

[3] Martone, M.: Rizzoli, P.: Wecklich, W.: González, C.: Bueso-Bello, J.L.: Valdo, P.: Schulze, D.: Zink, M.: Krieger, G.: Moreira, A.: The global forest/non-forest map from TanDEM-X interferometric SAR data. Remote Sensing of Environment; Vol. 205, Feb. 2018, pp. 352-373

[4] Mazza, A.: Sica, F.: Rizzoli, P.: Scarpa, G.: TanDEM-



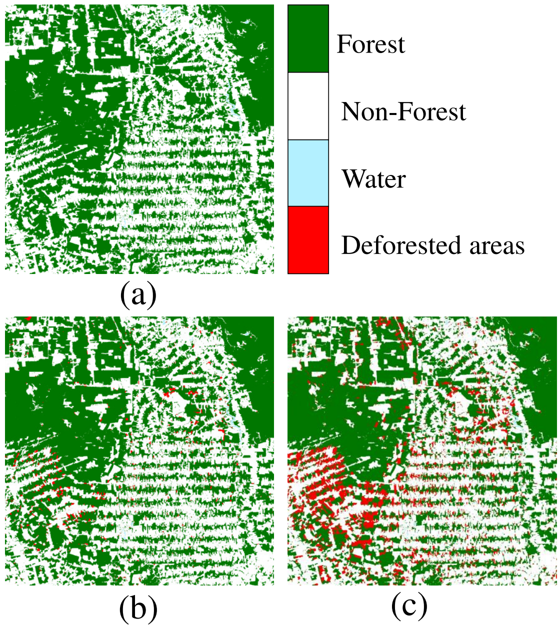


Figure 4 Large-scale mosaics comparison for the Geocell S10W058. (a) Optical image from Google Earth (C), (b) Land cover mosaic obtained with TanDEM-X and DL methods, (c) FROM-GLC reference map.

Figure 5 Deforestation for the Geocell S10W064. (a) TanDEM-Xmosaic 2011, (b) deforestation 2013, (c) deforestation 2016.

X Forest Mapping using Convolutional Neural Networks. Remote Sens. 2019, 11(24), 2980.

[5] Ronneberger, O.: Fischer, P.: Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Munich, Germany, 5-9 October 2015; Navab, N.: Hornegger, J.: Wells, W.M.: Frangi, A.F.: Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

[6] Rizzoli, P.: Martone, M.: Bräutigam, Global Interferometric Coherence Maps From TanDEM-X Quicklook Data. IEEE Geoscience and Remote Sensing Letters, Vol. 11, No. 11, Nov. 2014, pp. 1861-1865

[7] Du, Y.: Feng, G.: Zhiwei, L.: Xing, P.: Ren, Z.: Jian-Jun, Z.: A Method for Surface Water Body Detection and DEM Generation With Multigeometry TanDEM-X Data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 12, No. 1, Jan 2019, pp. 151-161

[8] González, C.: Bachmann, M.: Bueso-Bello, J.L.: Rizzoli, P.: Zink, M.: A Fully Automatic Algorithm for Editing the TanDEM-X Global DEM. Remote Sens. 2020, 12(23), 3961.

[9] Wendleder, A.: Wessel, B.: Roth, A.: Breunig, M.: Martin, K.: Wagenbrenner, S.: TanDEM-X Water Indication Mask: Generation and First Evaluation Results. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 6, No. 1, Feb. 2013, pp. 171-179

[10] Valdo, P.: Sica, F.: Rizzoli, P.: Improvement of TanDEM-X Water Mask by Exploiting the Acquisition Geometry. Proceedings of the European Conference on Synthetic Aperture Radar, EUSAR. Jun. 2018

[11] Hansen, M.C.: Potapov, P.: Moore, R.: Hancher, M.: Turubanova, S.: Tyukavina, A.: Thau, D.: Stehman, S.: Goetz, S.: Loveland, T.: Kommareddy, A.: Egorov, A.: Chini, L.: Justice, C.O.: Townshend, J.: High-Resolution Global Maps of 21st-Century Forest Cover Change. Science, 342. 850-853, 2013.

[12] Sano, E.E.: Rizzoli, P.: Koyama, C.N.: Watanabe, M.: Adami M.: Shimabukuro, Y.E.: Bayma, G.: Freitas, D.M.: Comparative Analysis of the Global Forest/Non-Forest Maps Derived from SAR and Optical Sensors. Case Studies from Brazilian Amazon and Cerrado Biomes. Remote Sens. 2021, 13(3), 367.

[13] Gong, P. et al.: Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017, Science Bulletin, Vol 64, No. 6, 2019, pp. 370-373.