Mapping Selective Logging in the Amazon with Artificial Intelligence and Sentinel-2

Jailson S. de Souza Filho [1], Camila da S. Damasceno [1], Dalton R. Ruy Secco Cardoso[1], Carlos M. Souza Jr. [1]

[1] Amazon Institute of People and the Environment (Imazon), Belém, Pará, Brazil - (jailson, camila, dalton, souzajr)@imazon.org.br

Keywords: Logging, Amazon, Artificial Intelligence, Forest Disturbance.

Abstract

The Amazon forest, the largest tropical forest in the world and marked by its rapid change in forest cover, has suffered from intense anthropogenic phenomena such as deforestation and forest degradation, this one caused mainly by fires and selective logging. This study explores a U-NET model to accurately identify selective logging infrastructure (roads, skid trails, storage yards) using Sentinel-2 imagery. Our goal is to improve the SIMEX (System for Monitoring Timber Harvesting) in the Brazilian Amazon, reducing the human workload and increasing the system's accuracy. Data from 780 SIMEX registration polygons (2021-2022) were used, with stratified sampling creating a training data set. The U-NET model, optimized with specific hyperparameters and data augmentation, analyzed six spectral bands (two-year RGB). We achieved an F1 score of ~81% with high precision (73.7%) and recall (90.31%) on the test set, indicating strong performance and generalization. Our model excels at accurately predicting logging infrastructure and potential damage to forest canopies. It provides detailed detection of roads and stockyards, offering a comprehensive view compared to models that generalize explored areas. This refined approach increases its usefulness for forest conservation and management efforts.

1. Problem Statement

The Brazilian Amazon Rainforest, the largest tropical forest in the world, is experiencing rapid changes due to deforestation and forest degradation (Souza et al., 2023). Deforestation is the total removal of forest cover, with areas of native forest being converted into pastures or agricultural fields (Lambin, 1999). Forest degradation is the partial removal of trees without changing the land cover, caused by predatory logging, fires, road construction, and forest fragmentation. Drawing a clear distinction between deforestation and forest degradation proves to be a complex endeavor. Most land use and land cover change analyses in tropical forests have focused on the causes and effects of deforestation; however, other less well-studied anthropogenic disturbances also threaten the future of tropical forests, such as logging, which often results in transient effects, with its distinct indicators disappearing within a relatively brief period of one to two years. This ephemeral nature renders degraded forests indistinguishable from undisturbed ones in spectral terms. (Souza Jr. et al. 2024; Lapola et al., 2023).

Selective logging in the Amazon can be carried out in two ways: through sustainable forest management or illegally and predatorily. Sustainable management is aimed at the controlled extraction of resources, guaranteeing ecosystem regeneration and preserving the integrity of the forest canopy. This practice involves the construction of planned infrastructure, such as stockyards and skid trails, designed to minimize damage to the soil and forest structure. In contrast, illegal logging has been expanding at an alarming rate, especially in protected areas, where improvised and destructive infrastructure leads to habitat fragmentation and significant degradation of biodiversity (Lapola *et al.*, 2023; Schulze *et al.*, 2006; Asner *et al.*, 2004). Recent studies show that even selective logging, when poorly managed, can result in long-term forest degradation, with persistent impacts on the forest canopy and biomass, which compromises the recovery of these areas and contributes to the encroachment of indigenous lands and conservation units (Dos Santos *et al.*, 2019; Piponiot *et al.*, 2019). Distinguishing between these practices is fundamental for formulating public policies that protect forests and combat illegality.

In this study, we focused specifically on identifying selective logging infrastructure. These structures are characterized by roads (primary and secondary), skid trails, and timber storage yards. Various methods detect these areas remotely (Souza Jr., 2013), but many still depend on human intervention. As SIMEX does, based on Landsat images (30 meters), the scars in the forest canopy and the roads are identified. Artificial Intelligence (AI) has recently been successfully applied to mapping roads, deforestation, and forest degradation (Dalagnol et al., 2023; Botelho et al., 2022). Using U-Net models, which are simpler architectures with less computational and training effort, it has been applied with remote sensing images to find spatial patterns in high spatial resolution data. Here, we present the results of our U-NET model to detect selective logging infrastructures in the Amazon region with Sentinel-2 images. We aim to improve the Logging Monitoring System (SIMEX) in the Brazilian Amazon to reduce human interpretation efforts and increase mapping accuracy.

2. Methods

2.1 Reference Dataset

We utilized SIMEX, a logging monitoring system developed by Imazon in 2008, which integrates satellite data with official logging documents to map logging activities in the Amazon region (SIMEX, 2024). Using Spectral Mixture Mix

techniques, SIMEX generates Normalized Difference Fraction Index (NDFI) images to enhance logging detection complemented with visual interpretation and validation of the results (Souza et al., 2005). Our dataset for 2021-2022 covered the Legal Amazon region, comprising 4,416 logging polygons totaling 718,000 hectares. We selected 20% of the dataset using stratified sampling (figure 1), amounting to 780 polygons covering 140,000 hectares.

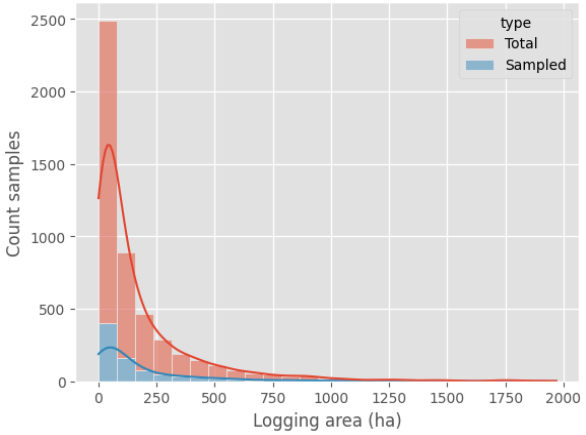

Figure 1 - Histogram of total and sampled logging dataset.

Most logging polygons (601 samples) of the partitioned dataset are concentrated up to 190 ha while large logging areas (> 1k ha) totals only 11 units. We also checked the spatial distribution of the data. Below is the map showing the density of logging along the Amazon biome.

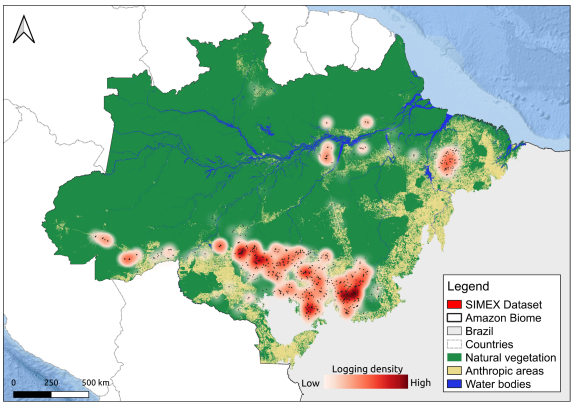

Figure 2 - Amazon biome subset dataset.

2.2 Image CHIPs Generation

The logging polygons were converted into 512x512 image chips for AI modeling. A Sentinel-2 image representing the year of detection (t1) was selected for each logging polygon. To capture the previous year's conditions (t0), images with cloud cover ≤ 30% were filtered, and a median mosaic was generated. The polygons were further refined through visual inspection within the Google Earth Engine (GEE) environment. Additionally, data augmentation techniques, including rotation and horizontal flipping, were applied to the training dataset. As a result, the final training dataset comprises 1,404 samples.

2.3 Selective Logging Model

We employed the conventional U-NET architecture, customizing its hyperparameters to align with the specific requirements of our task. The architecture consists of encoder and decoder blocks, where the encoder gradually reduces the spatial dimensions while capturing features, and the decoder progressively reconstructs the spatial dimensions to produce the final output.

We integrated 25% dropout layers immediately following the 3 x 3 convolutional layers to mitigate overfitting. This strategy helps ensure that the model generalizes well to unseen data. Additionally, we chose to replace the standard Rectified Linear Unit (ReLU) with the Leaky Rectified Linear Unit (LeakyReLU) activation function. This decision was driven by the enhanced ability of LeakyReLU to capture subtle features, as demonstrated by Maas, Hannun, and Ng (2013).

Our model processes inputs consisting of the RED, GREEN, and BLUE spectral bands captured at two distinct time points, t0 and t1, resulting in a total of 6 input data channels. These channels are critical for capturing the temporal changes in vegetation and other surface features, which are essential for identifying selective logging. The architecture culminates in a dense layer equipped with a sigmoid activation function, which plays a pivotal role in generating the probabilistic output required for this classification task.

Once the model generates the output image, it is subsequently converted into a binary image by applying a threshold of 0.5. Pixels with values below 0.5 are assigned a value of 0, while those above the threshold are set to 1. This binarization step is crucial for delineating the areas of interest, enabling a clear distinction between logged and unlogged regions.

2.4 Accuracy Assessment

To conduct the accuracy assessment, we partitioned the refined dataset into three subsets: training (60%), calibration (20%), and test (20%). Subsequently, we utilized the training and validation subsets for model calibration, reserving the test subset for evaluating the model's performance. Below we present the metrics employed in this study.

Precision: precision indicates how many of the predicted positive outcomes were actually correct, making it particularly useful in scenarios where the cost of false positives is high. It is defined as:

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

this metric measures the accuracy of positive predictions. It is the ratio of true positive (TP) predictions to the total number of positive samples - true positives plus false positives (FP).

Recall: this metrics indicates if the model is effective at capturing most of the relevant cases, which is particularly important in scenarios where missing positive cases has significant consequences. It is defined as:

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

It is calculated as the ratio of TP to the sum of TP and false negatives (FN) - actual positive cases that were incorrectly classified as negative.

F1 Score: it is the harmonic mean of precision and recall, providing a single score that balances the trade-off between the two. Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positives among all actual positive cases. The F1-score ranges from 0 to 1, with 1 being the best possible score, indicating perfect precision and recall. It is defined as:

$$F1\,Score = 2 . \frac{Precision . Recall}{Precision + Recall} \quad (3)$$

Loss Function: finally, we chose to utilize the Dice loss function for its ability to handle class imbalance and its sensitivity to small object segmentation, enhancing the model's performance across diverse scenarios. This loss function facilitates robust training by effectively penalizing false positives and false-negative detection, leading to more accurate segmentation results (Sudre et al., 2017).

$$Soft\ Dice\ Loss = 1 - \frac{2\sum i\ pigi}{\sum i\ pi + \sum i\ gi} \quad (4)$$

where pi is the predicted selective logging region for pixel i, and gi is the ground truth region for pixel i.

2.5 Experiment

After training the U-NET model, we conducted a comprehensive experiment to classify logging areas in Pará, Brazil. Leveraging the automation capabilities of the classification process, we analyzed all satellite scenes with less than 80% cloud cover during the SIMEX detection period (August 2022 to July 2023). For each scene, we computed a binary map of selective logging, saved the time information, and stacked the scenes to sum the values, creating a logging frequency image. This approach provided a detailed temporal analysis of logging activity, resulting in a high-resolution map that aligns with the SIMEX data.

3. Results

We trained the U-NET models over 50 epochs, selecting the epoch with the lowest dice loss, which happened to be epoch 43. Figure 3 shows the accuracy metrics results.

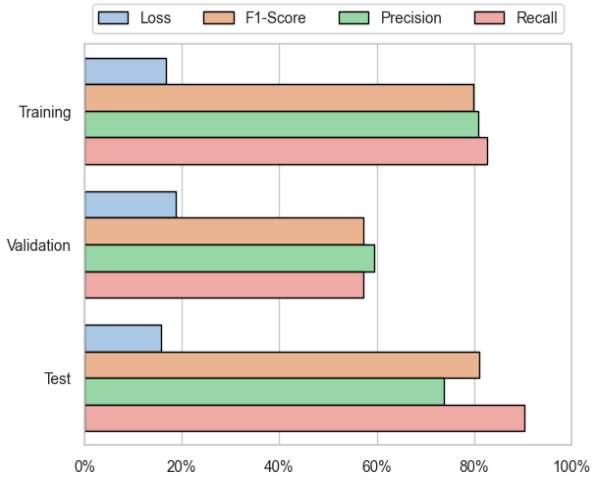

Figure 3. Accuracy of the U-NET model for the training, validation, and test datasets.

The model achieved satisfactory results in its accuracy metrics during training and validation. The model demonstrated solid performance with an average accuracy of around 80% for the F1-Score and consistent precision and recall of around 80% and 82%, respectively. In the test set, the results were promising, with an F1-Score of approximately 81%, a precision of 73.7%, and a great recall of 90.31%. These results indicate that the model can generalize well and maintain consistent performance across different logging intensities (Figure 6).

The analysis of the dice loss metrics (Figure 4) for the selective logging target dataset reveals a promising trend of model improvement over the training epochs. Initially, the loss value for the training dataset exhibits a significant decrease, dropping from 0.93 to 0.14, which suggests that the model effectively learns to predict selective logging patterns in the early stages. This decline in loss is particularly notable given the complexity and variation within the dataset, which includes 1,404 samples with diverse logging road segments. The stabilization of loss values in later epochs further indicates that the model is converging and can generalize well across the varied sample set.

In parallel, the validation loss follows a similar declining trend, demonstrating the model's capacity to maintain its predictive accuracy on unseen data. However, the existing dataset, while yielding encouraging results, might need to be revised to capture the variability inherent in selective logging scenarios fully. The performance could benefit from including more samples, which would provide a broader representation of logging conditions and enhance the model's robustness. Expanding the dataset would likely improve the

model’s generalization capabilities, making it more reliable for practical applications in monitoring and managing selective logging activities.

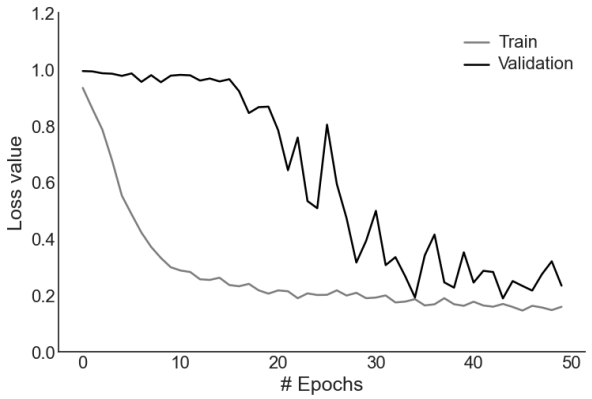

Figure 4. Dice loss values over epochs.

3.1 Frequency of Logging Detection

In total 1,811 scenes were processed. The data indicates a correlation between the number of Sentinel-2 images processed per month and the detected area of selective logging in the state of Pará. It is observed that in months with a higher number of processed images, such as July 2023 (1,151 images), the largest areas of selective logging were also identified (75,873 hectares). Conversely, in months with fewer processed images, such as March 2023 (527 images), the detected area of selective logging was significantly smaller (8,187 hectares).

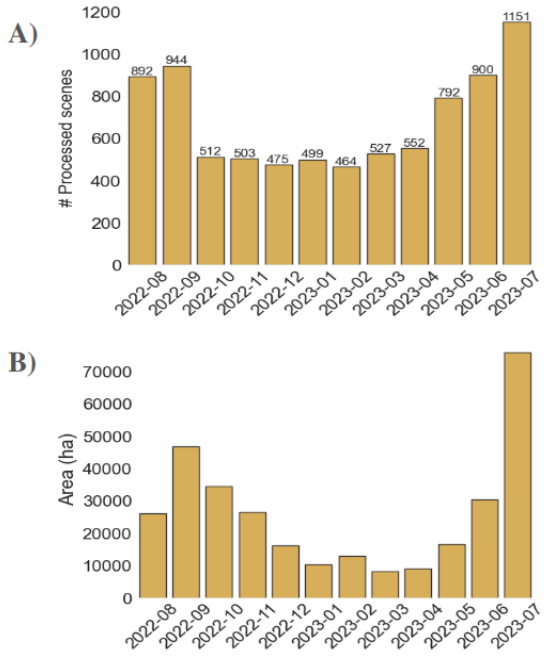

Figure 5. A) Number of Sentinel-2 scenes processed, B) logging detection by month.

The results, illustrated in Figure 7, highlight areas of high logging activity (marked in red), corresponding to regions with significant forest loss. In contrast, areas shaded in white and yellow indicate regions with minimal impact. Notably, there are regions that SIMEX did not detect, which may be attributed to the limited number of images analyzed by visual interpretation.

4. Conclusion

In conclusion, our analysis underscores the efficacy of the prediction model in accurately identifying logging infrastructure and assessing the potential damage to forest canopies. Its meticulous detection of roads and stockyards offers a comprehensive perspective, distinguishing it from other models that often generalize logging areas in forests. The model's refined and specific approach enhances its utility for forest conservation and management efforts.

The automation allows for the processing of large volumes of data, providing temporally detailed information about the logging dynamic. It may also be used as an alert system for regulating and controlling illegal logging by government agencies.

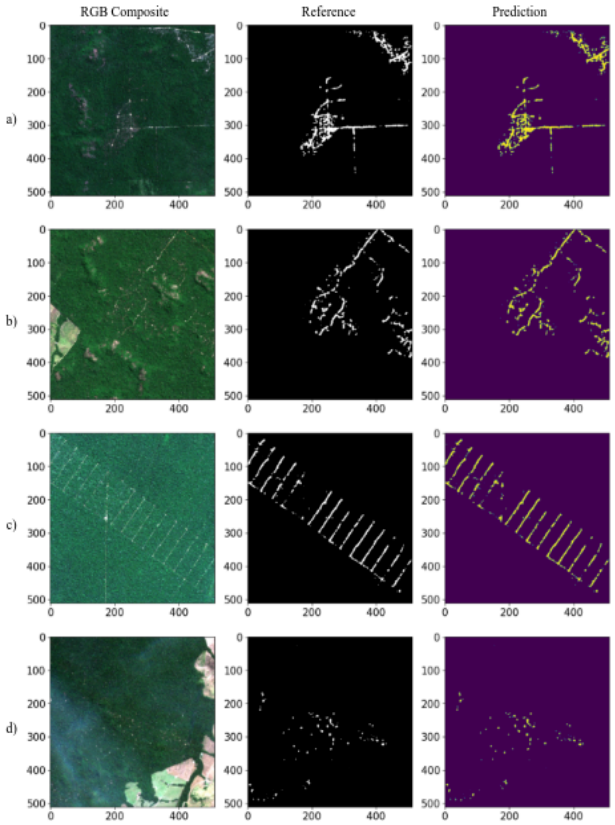

Figure 6. Predictions of different patterns of logging. A-b) unplanned logging, c) planned extraction, and d) low-impact logging.

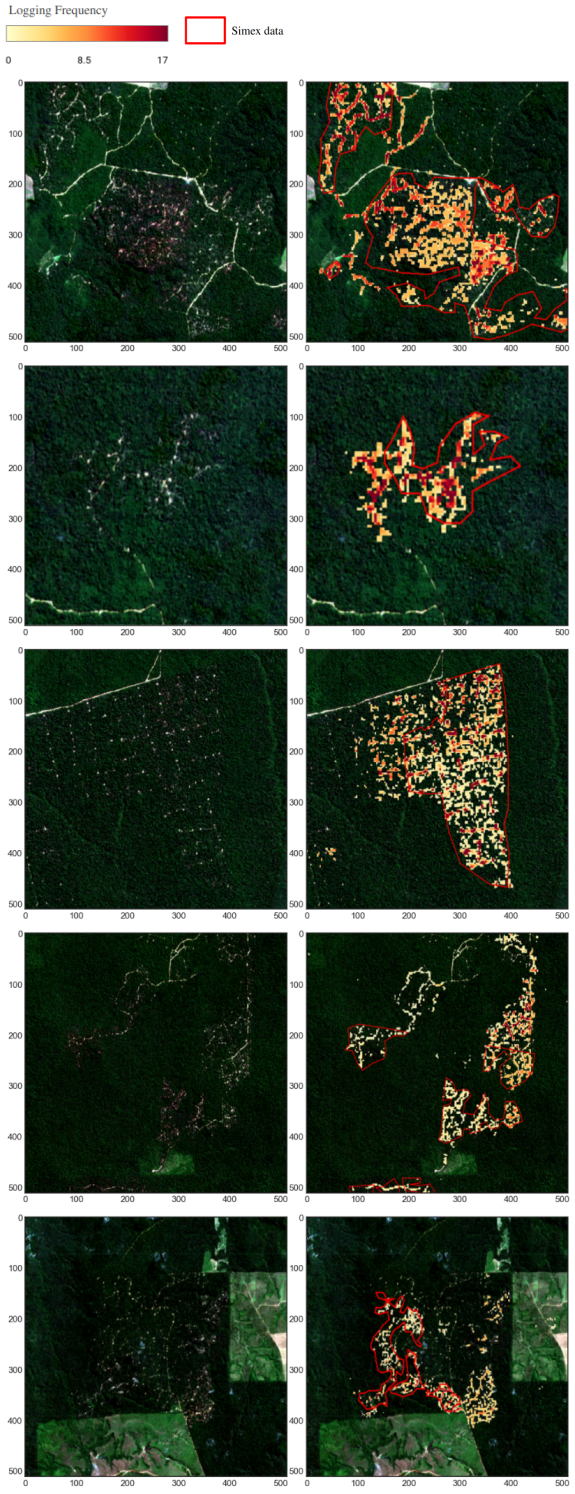

Figure 7. Logging frequency detected by AI model and SIMEX data.

References

Asner, G.P., Keller, M., Silva, J.N. 2004. Spatial and temporal dynamics of forest canopy gaps following selective logging in the eastern Amazon. Glob. Chang. Biol. 10, 765–783.

Botelho, J., Costa, S. C. P., Ribeiro, J. G., Souza, C. M. 2022. Mapping Roads in the Brazilian Amazon with Artificial Intelligence and Sentinel-2. Remote Sensing 14(15), 3625. https://doi.org/10.3390/RS14153625.

Dalagnol, R., Wagner, F. H., Galvão, L. S., Braga, D., Osborn, F., da Conceição Bispo, P., Saatchi, S. 2023. Mapping tropical forest degradation with deep learning and Planet NICFI data. Remote Sensing of Environment, 298, 113798.

Lambin, E. F. 1999. Monitoring forest degradation in tropical regions by remote sensing: some methodological issues. Global Ecology and Biogeography, 8(3–4), 191–198. https://doi.org/10.1046/j.1365-2699.1999.00123.x.

Lapola, D. M., Pinho, P., Barlow, J., Aragão, L. E., Berenguer, E., Carmenta, R., Walker, W. S. 2023. The drivers and impacts of Amazon forest degradation. Science, 379(6630), eabp8622.

Maas, A.L., Hannun, A.Y., Ng, A.Y. 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models. In Proceedings of the ICML Workshop on Deep Learning for Audio, Speech and Language Processing, Atlanta, GA, USA, 16 June 2013; Volume 28.

Pinagé, E., Keller, M., Duffy, P., Longo, M., dos-Santos, M., Morton, D. 2019. Long-Term Impacts of Selective Logging on Amazon Forest Dynamic from Multi-Temporal Airborne LiDAR. Remote Sensing 11, no. 6: 709. https://doi.org/10.3390/rs11060709.

Piponiot, C., Rutishauser, E., Derroire, G., Putz, F., Sist, P., West, T., Descroix, L., Guedes, M., Coronado, E., Kanashiro, M., Mazzei L., d'Oliveira, M., Peña-Claros, M., Rodney, K., Ruschel, A., de Souza, C. 2019. Optimal strategies for ecosystem services provision in Amazonian production forests. Environmental Research Letters 14, no. 12. Doi: 10.1088/1748-9326/ab5eb1.

Ronneberger, O., Fischer, P., Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer International Publishing.

Sistema de Monitoramento da Exploração Madeireira. Simex. Série de Boletins e Mapas de Exploração Madeireira. Imazon. Belém: Imazon. Available at: http://imazon.org.br/publicacoes.

Schulze, M., Zweede, J. 2006. Canopy dynamics in unlogged and logged forest stands in the eastern Amazon. For. Ecol. Manag. 236, 56–64.

Souza Jr, C. 2013. Monitoring of forest degradation: a review of methods in the Amazon Basin. Global Forest Monitoring from Earth Observation. Boca Raton, FL: CRC Press/Taylor & Francis Group, 171–194.

Souza, C., Tenneson, K., Dilger, J., Wespestad, C., Bullock, E. 2024. Forest Degradation and Deforestation. In: Cardille, J.A., Crowley, M.A., Saah, D., Clinton, N.E. (eds) Cloud-Based Remote Sensing with Google Earth Engine. Springer, Cham. https://doi.org/10.1007/978-3-031-26588-4_49.

Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M. 2017. Generalized Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. Deep. Learn. Med. Image Anal. Multimodal. Learn. Clin. Decis. Support 2017, 2017, 240–248.