Can Deep Learning help to forecast deforestation in the Amazonian Rainforest?

Tim Engelmann
Group for Sustainability and Technology
ETH Zurich
Zurich, 8092
`tengelmann@student.ethz.ch`

Malte Toetzke
Group for Sustainability and Technology
ETH Zurich
Zurich, 8092
`mtoetzke@ethz.ch`

Abstract

Deforestation is a major driver of climate change. To mitigate deforestation, carbon offset projects aim to protect forest areas at risk. However, existing literature shows that most projects have substantially overestimated the risk of deforestation, thereby issuing carbon credits without equivalent emissions reductions. In this study, we examine if the spread of deforestation can be predicted ex-ante using Deep Learning (DL) models. Our input data includes past deforestation development, slope information, land use, and other terrain- and soil-specific covariates. Testing predictions 1-year ahead, we find that our models only achieve low levels of predictability. For pixel-wise classification at a 30 m resolution, our models achieve an F1 score of 0.263. Only when substantially simplifying the task to predicting if any level of deforestation occurs within a 1.5 km squared tile, the model results improve to a moderate performance (F1: 0.608). We conclude that, based on our input data, deforestation cannot be predicted accurately enough to justify the ex-ante issuance of carbon credits for forest conservation projects. As main challenges, there is the extreme class imbalance between pixels that are deforested (minority) and not deforested (majority) as well as the omittance of social, political, and economic drivers of deforestation.

1 Introduction

Deforestation has a substantial impact on climate change. In the Brazilian Amazonian Biome alone, 21.3 M ha of rainforest were cut down in the years from 2000 to 2020 (Souza et al., 2020) making the rainforest a net-carbon emitter (Gatti et al., 2021). Forest conservation projects help to mitigate deforestation by protecting endangered sites. Often they are financed via voluntary carbon markets by issuing carbon offset certificates upfront. This requires project developers to calculate future emissions reductions achieved by the project by estimating the deforestation risk as a baseline scenario. However, this financing structure is heavily debated, as the deforestation risk is estimated through simplistic methods. Retrospective analyses (West et al., 2020, 2023; Guizar-Coutiño et al., 2022) show that project developers substantially overstated the actual risk of deforestation, thereby achieving only a fraction of the claimed emissions reductions.

Here, we investigate to which extent it is possible to accurately predict the risk of deforestation in advance in order to make well-founded claims about future emissions reductions. We use DL techniques to forecast the location of deforestation in the Amazonian Rainforest. Our analysis focuses on two tasks: 1) *Pixel-wise classification*: For each forest pixel in a given tile (of size 1.5 km), predict if it will be deforested or not in the next year. 2) *Tile-wise classification*: Given a tile, predict if any deforestation will take place in the next year. The code for our analyses is publicly available via `https://github.com/TimEngelmann/future-deforestation`. Our results have important

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023.

implications for deforestation projects in the voluntary carbon markets and question the current practice of issuing carbon credits ex-ante.

2 Related work

Only a few studies have tried to forecast deforestation with different approaches and mixed results. Takahata et al. (2022) uses a Bayesian model to derive dynamic baseline scenarios of deforestation for entire project sites of approximately 30T ha. They account for four different covariates: (1) distance to recent deforestation, (2) distance to urban centers/capitals, (3) distance to roads/highways, and (4) slope/elevation information. However, their ex-ante predictions are merely indicative and could vary significantly within their proposed 90% confidence interval.

Ball et al. (2022) proposes a different approach by forecasting how deforestation will spread on pixel-level. In particular, they predict for each pixel of 30 m resolution if it will be deforested in the subsequent year using satellite data and a convolutional deep-learning model adapted from Li et al. (2017). As a result, they report F1 scores up to 0.715. As we describe in Section 3.1 and 4, we believe that their data splitting and down-sampling approach has led to overconfident test performance results.

3 Methods

3.1 Data and preprocessing

Our analysis is based on the MAPBIOMAS Brasil dataset (Souza et al., 2020; West et al., 2020; Takahata et al., 2022). Specifically, we make use of the yearly deforestation, land use, and pasture quality layers. Additionally, we add slope information from the FABDEM dataset (Hawker et al., 2022; Takahata et al., 2022). We use a resolution of 30 m which is the most granular available resolution. Furthermore, we work in the coordinate system EPSG:6933 and focus on the region of the Brazilian Amazon biome (IBGE, 2019). Appendix A.1 provides further details regarding the choice of data and pixel resolution. Through further preprocessing, we generate the following eight input layers: (1) *Distance to the closest pixel deforested in the last*, (2) *the last 5*, and (3) *last 10 years*; (4) *Distance to the closest urban pixel*; (5) *Slope data*; (6) *Land use in the last input year*; (7) *Pasture quality in the last input year*; (8) *Deforestation data in the last input year*. The target layer includes the next-year state for each forest pixel, which can be either *deforested* or *remains primary forest*. Figure 1 shows the input and target layers of a sample tile of 50×50 px.

We only consider forest areas close to the deforestation line. For each pixel, we determine the Euclidean distance to the nearest pixel that was deforested in the previous 5 years. We then filter for pixels with a deforestation distance of up to 1.5 km (50 px), which captures the large majority of deforestation within a year (90% of the deforestation in the year 2018; see Appendix A.2). The filter helps to reduce the extreme class imbalance between the minority class (*deforested*) and the majority

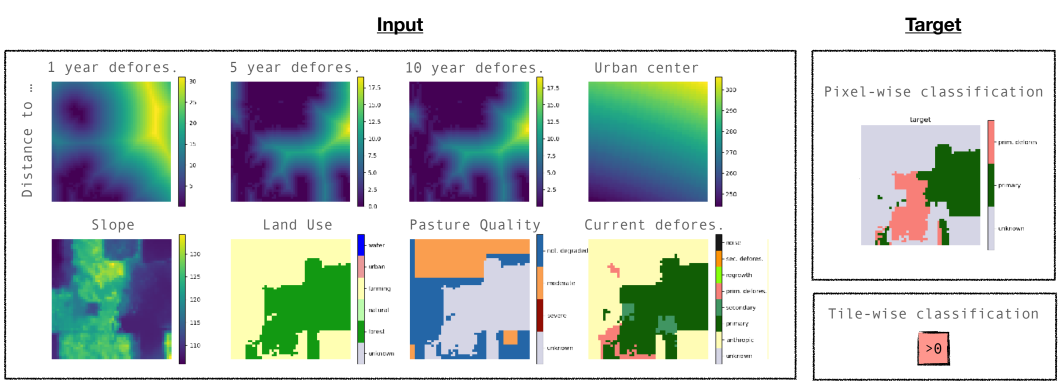

Figure 1: Input and target layers for one tile of 50×50 px (225 ha). The input layer represents the distance to prior deforestation, distance to urban areas, slope, land use, pasture quality, and deforestation state. The target layer represents the occurrence of deforestation in the next year as a binary variable on pixel or tile level.

class (*remains primary forest*). While in the entire biome 0.23% of primary forest was cut down in the year 2018, within our filtered area it was 1.3%.

Lastly, we split our data into spatio-temporally independent training, validation and test sets. Temporal independence is ensured as follows: The training and validation sets consist of input data from 2007 to 2017 and consider 2018 as the target year. In contrast, the time period of the test set is shifted one year ahead with an input range from 2008 to 2018 and the target year in 2019. Spatial independence is ensured by dividing the layers into squared segments of size 55 px (we only keep segments, where the center pixel is covered by primary forest in the year 2017). Each segment is either assigned to the training or the validation/test set. We then sample input tiles of size 50 px from the segments. As the segments are slightly larger compared to the input tiles, it allows for shift and flip data augmentation during training, while maintaining spatial independence. Our final data set consists of 162264 train segments, 40567 val segments, and 37200 test segments. Appendix A.3 and A.4 visualize our data pipeline and final dataset.

Notably, our data splitting approach differs in two main points from related work by Ball et al. (2022). First, Ball et al. (2022) assigned pixels to training or validation/test sets directly instead of using segments. This approach might violate spatial independence by allowing for overlaps in the input layers between training, validation, and test sets. Second, to counter the class imbalance, Ball et al. (2022) uses downsampling where training, validation, and test sets are downsampled to a ratio of 4:1 (non-deforested:deforested pixels). While downsampling is commonly used during training, a downsampled test set follows a non-representative data distribution compared to the real-world task. Both of these differences could have led to overconfident performance estimates by Ball et al. (2022).

3.2 Models

To forecast deforestation 1-year ahead, we define two distinct tasks:*Pixel-wise classification* and *tile-wise classification*.

For the task of *pixel-wise classification*, we predict for each forest pixel in a tile if it will be deforested in the subsequent year. In contrast to the architecture used by Ball et al. (2022), which makes predictions one pixel at a time, we decide on a UNet model. The UNet model, first introduced by Frangi et al. (2015), is commonly used for land usage classification tasks (Wang et al., 2022; Boonpook et al., 2023; Singh and Nongmeikapam, 2023). It consists of a CNN encoder and decoder with intermediate links between the layers. Therefore, it returns predictions for all pixels within the tile in one forward pass. We use the UNet implementation of the Segmentation Models library (Iakubovskii, 2019) and decide on a Dice loss (Cardoso et al., 2017), which works well under class imbalance.

The task of *tile-wise classification* represents a substantially simplified task: We predict if at least one forest pixel in the tile will be deforested. 24.9% of all tiles show no deforestation at all. With this task, we bypass both the exact location within the tile and the overall degree of deforestation in the tile. We reuse the 2D convolutional model (2D CNN) architecture from Ball et al. (2022) and Li et al. (2017). It consists of four convolutional layers with batch normalization, followed by a max-pool, a linear, a dropout, and a sigmoid layer. Finally, the output is compared to the binary target via a weighted binary cross-entropy loss (BCE).

For both tasks, we implement the training and evaluation pipeline in PytorchLightning. We train all models until the validation loss converges and choose the classification threshold with the best F1 score on the validation set. Moreover, we decide on a learning rate of 0.0001 and a dropout of 0.3 where applicable.

4 Results

Figure 2 shows our results for *pixel-wise* and *tile-wise classification*. All scores reported in the following represent performances on the test set. For *pixel-wise classification*, our UNet model achieves an F1 score of 0.263. We find this score well below the scores reported by Ball et al. (2022), which range up to 0.715. However, as described in Section 3.1, we suspect these to be overconfident results, due to the spatial overlaps between the input layers of their training and test sets, as well as their choice to downsample the test set. Therefore, to arrive at a comparable baseline, we implement their described model and train it on our data. When downsampling our test set, it reaches an F1

Task	Information value	Model architecture	Loss function	Validation imbalance	Experiment	Validation Set			Testing Set		
						F1	Prec.	Recall	F1	Prec.	Recall
Pixel-wise Classification	+++	2D-CNN	BCE	1.27%	*Ball et. Al. 2022* but omitting downsampling	0.226	0.213	0.240	0.197	0.171	0.234
		UNet	Dice	1.06%	Segmentation architecture	**0.274**	**0.257**	**0.294**	**0.263**	**0.239**	**0.292**
>0 Tile-wise Classification	+	Naive Baseline	-	24.9%	Naive Baseline, last year state	0.549	0.532	0.567	0.567	0.549	0.586
		Random Forest			Random Forest on center px input	0.598	0.525	**0.695**	0.560	0.477	**0.679**
		2D-CNN	BCE		2D CNN receiving tile wide input	**0.618**	**0.572**	0.671	**0.608**	**0.552**	0.676

Figure 2: Comparison of model performances on the test set for the two described tasks of *pixel-wise* and *tile-wise classification*

score of 0.725. Yet, on our actual non-downsampled test set, it obtains an F1 score of 0.197 after hyperparameter tuning. This is a lower performance compared to our UNet.

On the task of *tile-wise classification*, the 2D-CNN model obtains an F1 score of 0.608. We benchmark this score against two other more naive approaches. First, we train a Random Forest model only on the center pixel of our input layers. The model yields an F1 score of 0.560, which is only slightly lower, than the more complex CNN model. We perform a Mean Decrease in Impurity (MDI) feature importance analysis on the Random Forest model. It reveals that the *distance to 1-year deforestation* feature is by far the most informative. This leads us to our second naive baseline: If deforestation happened in the tile within the last year, we assume that deforestation will occur in the next year as well. This simple rule leads to an F1 score of 0.567.

5 Discussion and conclusion

Overall, our experiments show that predicting deforestation 1-year ahead is a difficult task for DL models. Our models only achieve poor F1 scores on the pixel level. The substantial simplification of tile-wise classification increased the predictive performance. However, based on predictive performance and spatial granularity of predicted outputs, we conclude that both tasks did not yield results that would justify a reliable risk score for deforestation. As a main challenge, we see the extreme class imbalance between deforested and not deforested pixels. Furthermore, our model omits social, political, and economic factors that might impact the spatio-temporal spread of deforestation but are not measured comprehensively.

Our results challenge the current design of voluntary carbon markets offering CO_2 compensations via forest conservation projects. Currently, project developers commonly estimate the emissions reductions of carbon offset projects 5-7 years in advance to issue carbon credits. Our results indicate that, even with more elaborate methods, shorter time periods, and relying on established covariates (e.g., deforestation line development, slope information, soil quality), the risk of deforestation can not be predicted accurately. Therefore, we question whether carbon credits from forest conservation projects should be issued in advance, as currently practiced.

Acknowledgements

We want to thank Sadiq Jaffer and Thomas Swinfield for their helpful advice and constructive feedback throughout the project.

References

Ball, J. G. C., Petrova, K., Coomes, D. A., and Flaxman, S. (2022). Using deep convolutional neural networks to forecast spatial patterns of amazonian deforestation. *Methods in ecology and evolution*, 13:2622–2634.

Boonpook, W., Tan, Y., Nardkulpat, A., Torsri, K., Torteeka, P., Kamsing, P., Sawangwit, U., Pena, J., and Jainaen, M. (2023). Deep learning semantic segmentation for land use and land cover types using landsat 8 imagery. *ISPRS international journal of geo-information*, 12:14.

Cardoso, M. J., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J. M. R. S., Moradi, M., Bradley, A., Greenspan, H., Papa, J. P., and Madabhushi, A. (2017). *Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations*, volume 10553. Springer International Publishing AG.

Frangi, A. F., Hornegger, J., Navab, N., and Wells, W. M. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*, volume 9351. Springer International Publishing AG.

Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Domingues, L. G., Cassol, H. L. G., Tejada, G., Aragão, L. E. O. C., Nobre, C., Peters, W., Marani, L., Arai, E., Sanches, A. H., Corrêa, S. M., Anderson, L., Randow, C. V., Correia, C. S. C., Crispim, S. P., and Neves, R. A. L. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595:388–393.

Guizar-Coutiño, A., Jones, J. P. G., Balmford, A., Carmenta, R., and Coomes, D. A. (2022). A global evaluation of the effectiveness of voluntary redd+ projects at reducing deforestation and degradation in the moist tropics. *Conservation biology*, 36:e13970–n/a.

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C., and Neal, J. (2022). A 30 m global map of elevation with forests and buildings removed. *Environmental research letters*, 17:24016.

Iakubovskii, P. (2019). Segmentation models pytorch.

IBGE (2019). Biomes and coastal-marine system of brazil - 1:250 000.

Li, Y., Zhang, H., and Shen, Q. (2017). Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network. *Remote sensing (Basel, Switzerland)*, 9:67–67.

Singh, N. J. and Nongmeikapam, K. (2023). Semantic segmentation of satellite images using deep-unet. *Arabian journal for science and engineering (2011)*, 48:1193–1205.

Souza, C. M., Shimbo, J. Z., Rosa, M. R., Parente, L. L., Alencar, A. A., Rudorff, B. F. T., Hasenack, H., Matsumoto, M., Ferreira, L. G., Souza-Filho, P. W. M., de Oliveira, S. W., Rocha, W. F., Fonseca, A. V., Marques, C. B., Diniz, C. G., Costa, D., Monteiro, D., Rosa, E. R., Vélez-Martin, E., Weber, E. J., Lenti, F. E. B., Paternost, F. F., Pareyn, F. G. C., Siqueira, J. V., Viera, J. L., Neto, L. C. F., Saraiva, M. M., Sales, M. H., Salgado, M. P. G., Vasconcelos, R., Galano, S., Mesquita, V. V., and Azevedo, T. (2020). Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, 12.

Takahata, K., Suetsugu, H., Fukaya, K., and Shirota, S. (2022). Bayesian state-space scm for deforestation baseline estimation for forest carbon credit. *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*.

Wang, J., Bretz, M., Dewan, M. A. A., and Delavar, M. A. (2022). Machine learning in modelling land-use and land cover-change (lulcc): Current status, challenges and prospects. *The Science of the total environment*, 822:153559–153559.

West, T. A. P., Börner, J., Sills, E. O., and Kontoleon, A. (2020). Overstated carbon emission reductions from voluntary redd+ projects in the brazilian amazon. *Proceedings of the National Academy of Sciences - PNAS*, 117:24188–24194.

West, T. A. P., Wunder, S., Sills, E. O., Börner, J., Rifai, S. W., Neidermeier, A. N., and Kontoleon, A. (2023). Action needed to make carbon offsets from tropical forest conservation work for climate change mitigation.

A Appendix

A.1 Additional Data Exploration

A.1.1 Choice of data type

The MAPBIOMAS dataset provides several different processed layers from which we can infer deforestation rates. Land use data, used and further processed by West et al. (2020); Takahata et al. (2022), transition data, and dedicated deforestation data (only recently added to the collection). The main difference is that transition data has undergone additional spatio-temporal filters and deforestation data differentiates between primary and secondary deforestation. The deforestation rates highly differ according to the choice of data. We decide to use the dedicated deforestation data, as we are particularly interested in primary deforestation.

A.1.2 Choice of px resolution

The data layers can be downloaded in different px resolutions. While West et al. (2020); Takahata et al. (2022) chose a resolution of 250 m/px, the results of Ball et al. (2022) are based on maps with 30 m pixel resolution. We calculate the deforestation rates within the Amazonian Biome for both choices and find that they differ substantially. According to the 30 m layer, from the years 2000 to 2020, 22.1M ha of primary forest were deforested. When using the 250 m layer, we report 4.9M ha less, which corresponds to a difference of 22.3%. It makes sense that the results differ, as in order for a pixel to be considered deforested, more than 50% have to be cleared within a year. For a 250 m pixel, this corresponds to 3.1 ha of deforestation that could potentially be ignored. Refer to Figure A.1 for a comparison of the above-mentioned deforestation rates.

A.1.3 Temporal trends

Calculating the overall deforestation in the Amazonian Biome also shows the variability of yearly deforestation rates. Deforestation rates are much higher in the early 2000s, reaching their peak in 2003 when 0.54% of the entire Amazon Biome area was deforested. Deforestation rates then decline, reaching their minimum in 2010 at 0.11%, and have since slightly increased again. It is important to keep this graph in mind when projecting deforestation rates into the future.

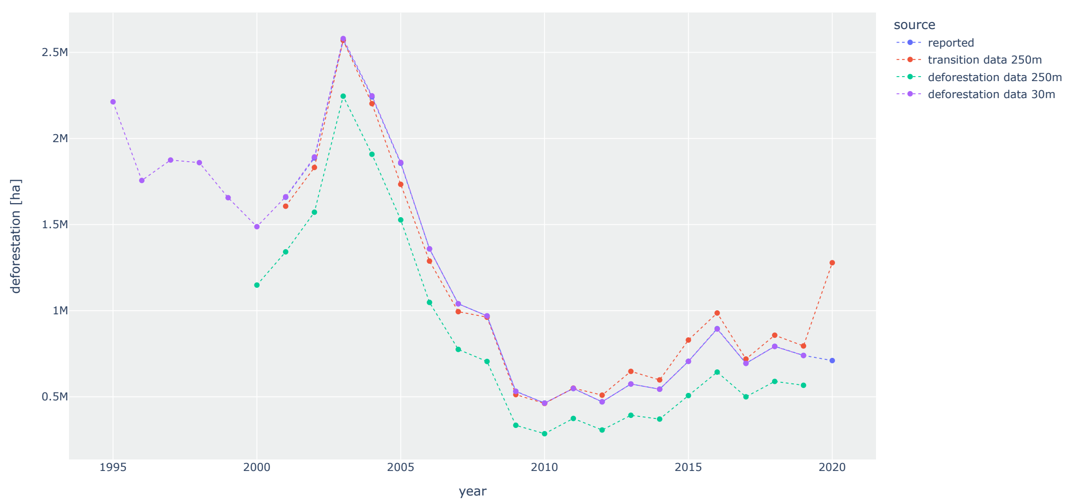

Figure A.1: Deforestation rates in the Brazilian Amazonian Biome, according to the MAPBIOMAS dataset. *reported* corresponds to the primary deforestation reported on the MAPBIOMAS website and coincides with the rates calculated on the *deforestation 30 m* layers. Other data layers and resolutions lead to different deforestation rates.

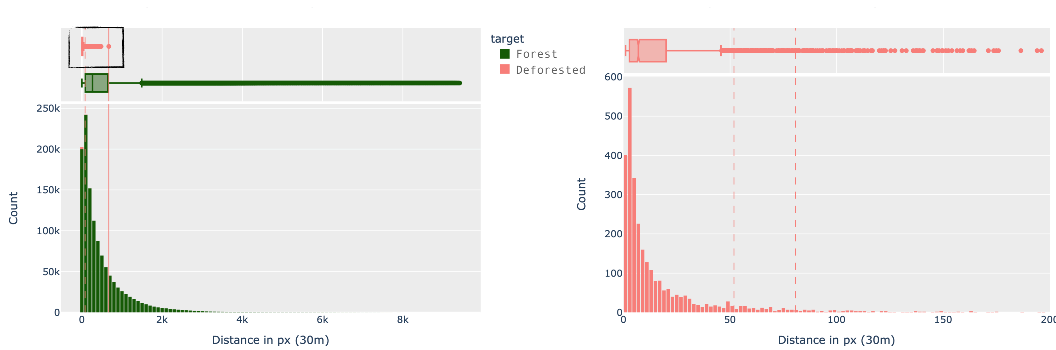

Figure A.2: Distance to recently deforested pixel for the train set. About 90% of pixels that are deforested in the next year are less than 1.5 km (50 px) away from a pixel deforested in the last 5 years. 95% are within a range of 2.4 km (81 px).

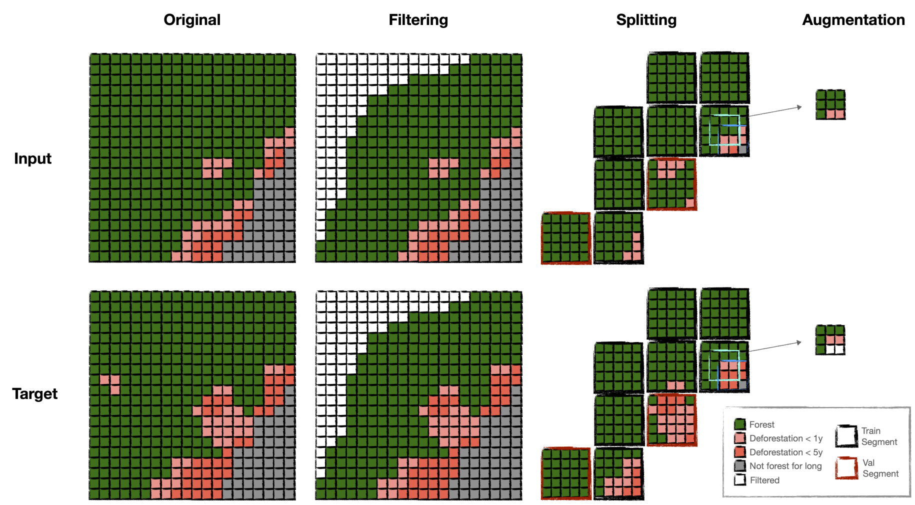

Figure A.3: Main steps of our data pipeline. We filter for pixels in proximity to the deforestation line. We split the data into segments to ensure spatial independence of training and test sets.

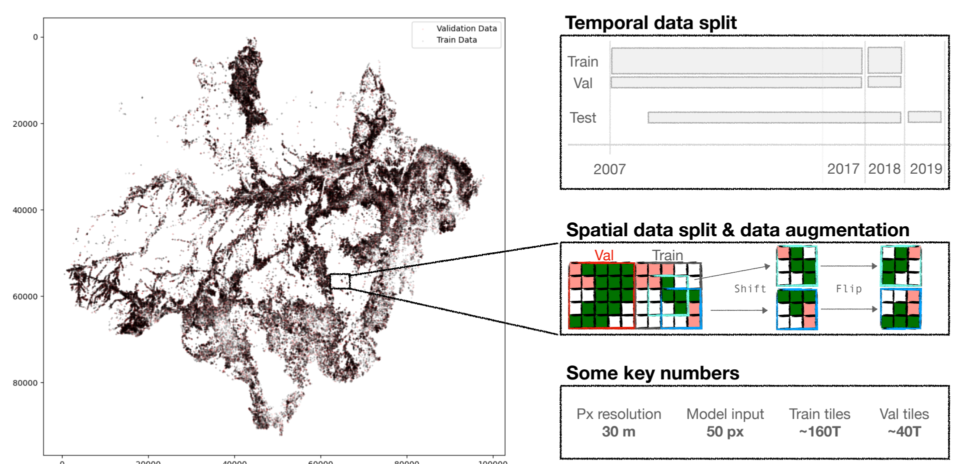

Figure A.4: Overview of our final dataset. The plot on the left shows all training and validation segments. We achieve temporal independence through a time shift.