Geo-information Science and Remote Sensing

Internship Report

# WWF FOREST FORESIGHT: PREDICTING DEFORESTATION IN LAOS

Zillah Calle

April 17, 2024

# WWF FOREST FORESIGHT: Predicting deforestation in Laos

Zillah Calle

Registration number: 1015365
Period of Internship: 2023-11-01 – 2024-04-30
Supervisor's Name Internship provider: Jonas van Duijvenbode
Name of Company Institution: WWF Netherlands
Address: Driebergseweg 10, 3708 JB Zeist
MGI Supervisor's Name Internship: Laura Cue la Rosa

April 17, 2024
Amersfoort, The Netherlands

Thesis code number: GRS-80423
Thesis Report: GIRS-2023-51
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

# Management summary

This study addresses the issue of deforestation, which poses a threat to biodiversity and ecosystem services. With a need for early intervention to prevent deforestation events and safeguard the forest, the study focuses on deforestation prediction using the eXtreme Gradient Boosting (XGBoost) model within the Forest Foresight (FF) program. The objectives include optimizing model parameters, assessing feature importance, evaluating predictions in deforestation-free areas, and analyzing seasonal patterns.

The study uses data covering the tropical belt, organized into 10 by 10 degree tiles with a spatial resolution of 400 by 400 meters and monthly temporal resolution from January 2020 to June 2023. The primary focus is on Lao People's Democratic Republic (Laos); however, to address the divergent results observed in Laos and to offer broader insights into the effectiveness of the tested methods on a global scale, the analysis also includes Gabon, Colombia, Peru, and Bolivia. The Global Forest Watch (GFW) Integrated Alerts integrating GLAD and RADD alert systems are used as groundtruth data. Features include static and dynamic variables derived from GFW and GLAD data, as well as accessibility, terrain, anthropogenic activity, and climate variables. The model is evaluated using precision, recall, and F0.5 scores and is compared against a baseline model.

Optimizing model hyper-parameters had only a modest effect on performance, with consistency observed across different landscapes. However, variations in training duration and the use of a dynamic threshold underscored the necessity for different approaches per country. Notably, during our research, correlated and uninformative features were identified for Laos. However, it was also observed that XGBoost exhibited robustness in managing these features. However, predicting deforestation in non-deforested areas is challenging, suggesting the need for alternative methodologies. While seasonal deforestation patterns were captured, detecting trend changes proved difficult, though this was somewhat mitigated by the use of a dynamic threshold. Future investigations may explore alternative algorithms such as deep learning and spatial models, alongside an emphasis on high-confidence alerts and forecasting deforestation extents.

# Contents

# List of abbreviations

**aucpr** Area Under the Curve for Precision-Recall. 8, 17

**FF** Forest Foresight. , 1–3, 5, 9, 17, 21

**GFW** Global Forest Watch. , 5, 6, 18

**GLAD** Global Land Analysis and Discovery. , 6

**Laos** Lao People's Democratic Republic. , 1–3, 5, 7, 11, 13, 15–21

**OSM** OpenStreetMap. 6

**PCA** Principal Component Analysis. 19, 20

**RADD** Radar for Detecting Deforestation. , 6, 20

**WWF** World Wide Fund for Nature. 1

**XGBoost** eXtreme Gradient Boosting. , 3, 7, 18, 20, 21

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Internship organisation background

The World Wide Fund for Nature (WWF), established in 1961 in Switzerland, has evolved into the leading global conservation organization. Committed to preserving nature and mitigating threats to biodiversity, WWF initially focused on endangered species and wildlife conservation. Over the years, its scope has expanded to encompass six key areas: climate, food, forest, freshwater, oceans, and wildlife. As the largest global conservation organisation, WWF strives to reduce humanity's environmental impact and build a future where people coexist harmoniously with nature.

## 1.2 Context and justification of research

Forests play a crucial role in sustaining our planet and its inhabitants. They serve as habitats for more than 80% of terrestrial animal and plant species, making them vital for biodiversity conservation (UNEP 2020). Additionally, forests contribute to climate regulation by capturing and storing carbon. Moreover, forests offer a range of other essential ecosystem services, including soil conservation, water regulation, recreation, and disaster risk mitigation (Jenkins and Schaap 2018).

Regrettably, deforestation persists as a pressing global concern, with an annual loss of 10.2 million hectares recorded from 2015 to 2020 FAO 2020. A country experiencing significant deforestation challenges is Lao People's Democratic Republic (Laos) witnessing a 1.1% decrease in forest cover between 2010 and 2017 (UNEP 2020). The primary driver of forest loss and degradation in Laos is attributed to smallholder farming, particularly shifting cultivation (WWF 2019). However, the growing rural population has been associated with a rise in illegal logging for household consumption (Lestrelin et al. 2013). Addressing the issue of deforestation is crucial for safeguarding the valuable services that forests provide to our environment and its inhabitants.

### 1.2.1 The significance of this topic

Currently, interventions in illegal logging are often too late. Anticipating the locations of illegal logging activities holds the key to early interventions, crucial for preventing deforestation. In collaboration with partners, WWF created the Forest Foresight (FF) program, specifically designed for predicting and preventing deforestation. Notably, this program has proven successful in Gabon and Kalimantan, achieving an 80% user accuracy and a detection rate of up to 46% in Kalimantan (Van Stokkom et al. 2020). Expanding on this success, Introducing the FF model in Laos can be the next logical step. Implementing the Forest Foresight model in Laos will offer insights into potential future deforestation scenarios, empowering local stakeholders with valuable predictions to prevent deforestation effectively.

### 1.2.2 Summary previous research/experiences

In the latest phase of the Forest Foresight (FF) project, a global model has been developed, outperforming the previous model. The main difference between the models is visualized in figure 1.1. The earlier model made deforestation predictions at a high spatial resolution of 15 x 15 meters, later down-sampling them to 480 x 480 meters. In contrast, the newly introduced "aggregated model" enriches high-resolution datasets by aggregating them, creating features at 400 x 400 meters. It then utilizes these features for deforestation predictions, achieving higher accuracy and reducing storage and computational costs by operating at a lower spatial resolution. More details on the model specifications can be found in Section 2.0.3.

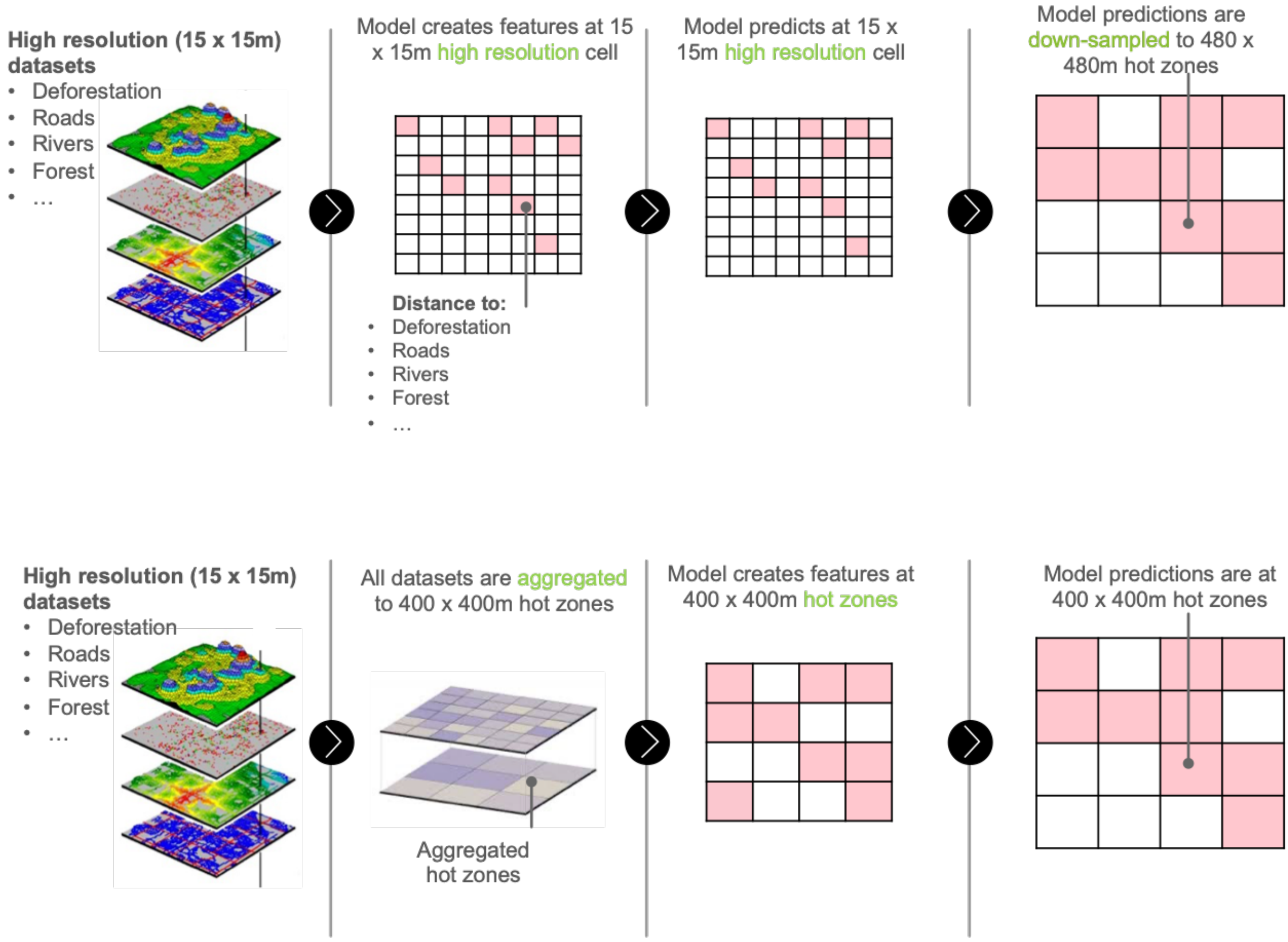



Figure 1.1: A comparison of the previous model (above) and the new aggregated model (below) of the Forest Foresight project.

An additional advantage of the aggregated model is its simplicity, allowing for easy code handover. The R-package "ForestForesight," developed by Jonas Duijvenbode, is specifically designed for implementing the aggregated model. This package facilitates data preparation, model training, and deforestation prediction, with the flexibility to operate at the country, tile (10 by 10 degrees), or country group level.

The datasets, containing 31 features comprising a mix of static and dynamic variables, have undergone preprocessing and are now accessible for 103 tiles via Amazon S3. Further details regarding these features, their sources, and the preprocessing steps applied are discussed in Section 2.0.2 .

### 1.2.3 Contribution/added value of this research

This research contributes to evaluating the applicability of the Forest Foresight project in Laos and identifies areas for enhancement. The insights gained will not only hold the potential to refine the model locally but also offer valuable lessons for implementing the new FF model in diverse global contexts.

Similar to the previous model, the new aggregated model uses eXtreme Gradient Boosting (XGBoost) (Chen and Guestrin 2016) for deforestation prediction (see section 2.0.3). Before running XGBoost, the model parameters need to be set. While the previous model provides initial parameter values, optimizing them for the aggregated model is vital for enhancing accuracy. Additionally, various other factors can influence prediction quality, such as the training time period or the spatial scale at which the model is trained (e.g., country, tile, or group of countries), all of which require testing and evaluation to ensure optimal performance.

An essential aspect of this research involves evaluating the model's performance with the existing set of input features. While the model currently operates with a relatively high number of features, it is important to recognize that more features don not always translate to improved accuracy. The exploration of feature importance and the identification of predictive features are crucial steps. While XGBoost's resilience to uninformative features is acknowledged, optimizing the feature set can be beneficial for reducing computational costs while maintaining or improving accuracy.

Moreover, there is a need to assess the model's quality in areas without previous deforestation. This addresses a possible limitation in the current model, which relies for heavily on past deforestation data. Understanding the model's capability to accurately predict new deforestation hotspots is pivotal and prompts considerations for potential adjustments in approach, features, or model parameters in these regions to ensure robust predictions.

In Laos, a distinct seasonal trend of deforestation is observed. The inclusion of features such as month, sinmonth, and precipitation (see Section 2.0.2) has the potential to capture and incorporate this seasonality into the model. Evaluating whether the model can indeed predict variations in deforestation across seasons is crucial, offering valuable insights for enhancing predictive accuracy.

## 1.3 The aim of internship research/activities

The objective is to forecast and analyze deforestation in Laos, with a primary focus on optimizing the accuracy of the FF model in this region. Additionally, the findings hold the potential to offer valuable insights applicable to the implementation of FF in other countries.

### 1.3.1 Research objectives

This internship study, conducted in Laos, aims to achieve the following research objectives:

1. **Optimization of model parameters:** Optimize the model parameter settings to improve overall prediction accuracy.

2. **Feature importance and selection:** Assess the importance of features and identify the optimal subset to enhance the model's performance.

3. **Prediction in deforestation-free areas:** Evaluate the model's performance in predicting deforestation in regions without a history of previous deforestation, providing insights into its generalization capabilities.

4. **Analysis of seasonality in deforestation:** Investigate the seasonality patterns in deforestation and evaluate the model's ability to capture and predict these temporal variations.

## 1.4 Explanation of technical terms/jargon

**classification threshold** The minimum predicted probability required to classify an observation as deforestation. 8, 9, 11, 15, 17, 18

**datasample** The sample size of the data to be loaded. 7, 10, 11

**eta** The learning rate controlling the step size during training iterations. 7, 10, 11

**evaluation metric** Evaluation metric for validation data, default is aucpr. 7, 8

**gamma** The minimum reduction in loss required to create an additional split. 7, 10, 11

**label threshold** The minimal number of integrated deforestation alerts per 400 by 400 meter area required to label that area as "actual". 6, 8, 12, 17

**min child weight** The minimum sum of instance weights required in a child node. 7, 11

**nrounds** The number of boosting rounds or iterations. 7, 10, 11

**subsample** The fraction of training instances (observations) to be randomly sampled without replacement during each boosting iteration. 7, 11

**tree depth** The maximum depth of a tree. 7, 10, 11

# Chapter 2

# Methodology

This chapter describes the methodological decisions made in the study. Firstly, it discusses the study area (Section 2.0.1) and provides an overview of the data, including groundtruth data and the features used for deforestation prediction and assessment (Section 2.0.2). Next, the Forest Foresight (FF) model is introduced (Section 2.0.3). Then, in Section 2.0.4, the methodology for addressing each research objective is explained.

### 2.0.1 Study area

The country of Lao People's Democratic Republic (Laos) is chosen as the study area (Figure 2.1). Laos is predominantly covered by tropical and subtropical broadleaf forests, with a smaller area in the southwest characterized by tropical and subtropical dry forest. The tropical climate in Laos exhibits two distinct seasons: a wet season (monsoon) from May to October and a dry season from November to April. Typically, the practice of slashing and burning for shifting cultivation occurs during the dry season (Li et al. 2014), resulting in a seasonal pattern of deforestation.

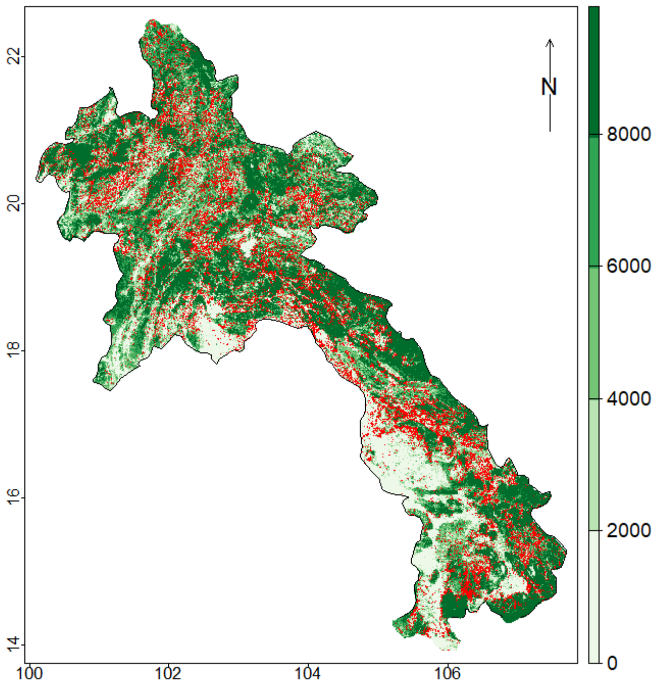



Figure 2.1: Forest cover in Laos (green) and the total deforestation (red, based on the Global Forest Watch (GFW) Integrated Alerts (Berger et al. 2022)) on May 2023.

### 2.0.2 Data

The data is accessible for the tropical belt, between 30 degrees north and 30 degrees south , organized into 10 by 10 degree tiles, with a spatial resolution of 400 by 400 meters, and primarily a monthly temporal resolution. Currently, the data availability extends from January 1, 2020, to June 1, 2023.

#### Groundtruth

The Global Forest Watch (GFW) Integrated Alerts (Berger et al. 2022) (10 m spatial resolution) are used as a proxy for deforestation. These alerts integrate data from both the Global Land Analysis and Discovery (GLAD) (Hansen et al. 2016) and the Radar for Detecting Deforestation (RADD) (Reiche et al. 2021) alert systems (Figure 2.2). The advantage of incorporating GLAD data lies in its extensive spatial coverage, while RADD benefits from the radar's ability to penetrate through clouds, enhancing the dataset's robustness for analysis.


Figure 2.2: Spatial data availability of the GFW integrated alerts. In the tropical belt, between 30 degrees north and 30 degrees south (muted pink), deforestation alerts are based solely on the GLAD system, while in the humid tropics (bright pink), both GLAD and RADD alert systems are used. Source: Berger et al. 2022

To generate groundtruth layers, the integrated alerts, initially at a resolution of 10 by 10 meters, are aggregated to match the processing resolution of 400 by 400 meters. The binary groundtruth for the next six months (groundtruth6mbin, Appendix Table A), serves as the default groundtruth used throughout this internship report. In this groundtruth layer, any 400 by 400 meter pixel where at least one deforestation event occurs (Label threshold=1) is designated as "actual". Consequently, the model aims to predict the incidence of deforestation alerts in the forthcoming six months.

#### Features

A complete list of the current features, their descriptions, and sources can be found in Appendix A. The input dataset comprises static and dynamic features, including both monthly and yearly data. Various features, such as patch density, smoothed total, and time since loss, are derived from the GFW alerts. Additionally, forest height, historic loss, and initial forest cover are obtained from GLAD data. Together, these variables aim to capture deforestation dynamics.

Furthermore, the dataset includes features related to accessibility, such as proximity to roads and waterways, derived from the OpenStreetMap (OSM), as well as terrain variables like slope and elevation. Anthropogenic activity is represented through population increase and nightlights. Climate variables, such as expected precipitation and fire alerts, are also included. These climate variables, along with auto-generated features like month and sine month, have the potential to capture the seasonality of deforestation. Furthermore, coordinates (x and y) are included in the feature dataset to account for spatial variation.

### 2.0.3 The forest foresight model

#### XGBoost

A gradient-boosted tree model, specifically XGBoost (Chen and Guestrin 2016), is trained using the R programming language. This approach involves sequentially training multiple trees, where each subsequent tree is trained to predict the error of the previous tree. In order to controll the learning process, the XGBoost model has multiple hyperparameters, required to be set; e.g. the tree depth, gamma, eta, subsample, min child weight, nrounds, datasample and the evaluation metric. These hyperparameters need to be optimized, as discussed in the next section (2.0.4). Afterwards, model comparisons will always be done with the same set of parameters.

The model will be trained for the full year 2021, ensuring the representation of all months and facilitating the learning of seasonal variations. A 6-month interval is essential for the test dataset due to the unknown groundtruth. Consequently, the model will be evaluated from June 2022 to June 2023. June 2023 serves as the latest feasible month for testing since the groundtruths need to be available 6 months later, in this case, December 2023. This testing duration not only enables an annual examination of seasonal distinctions in the model's performance but also enhances the reliability of the model's performance compared to testing on a shorter period.

#### Model evaluation

To assess the model's performance, its predictions will be compared with the groundtruth, and the precision, recall, and F0.5 scores will be calculated (equations 2.1, 2.2 and 2.3). The choice of the F0.5 score aims to create a unified evaluation metric that balances precision and recall, placing a higher emphasis on precision. This decision aligns with the project's preference for high precision.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2.1}$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{2.2}$$

$$\text{F0.5} = 1.25\frac{\text{Precision} \times \text{Recall}}{0.25 \times \text{Precision} + \text{Recall}} \tag{2.3}$$

To measure the quality of the developed model, these evaluation scores will be compared against a baseline model. The baseline model operates on a straightforward prediction – anticipating deforestation in the next six months if it occurred in the previous six months. This simplistic approach provides a benchmark for assessing the improvement achieved by the developed models.

### 2.0.4 Methods

#### Parameter optimization

The following hyperparameters require optimization: the tree depth, gamma, eta, subsample, min child weight, nrounds, and datasample (detailed descriptions in Section 1.3.1). Additionally, assessing the stability of these hyperparameters across different landscapes is crucial, as it informs the need for hyperparameter optimization per model. As an initial step, a Taguchi analysis is applied to four tiles (00N_010E, 10N_080W, 10N_110E and 10S_070W) representing different landscapes. Subsequent optimization for Laos also involves a Taguchi analysis.

Taguchi analysis is valuable for its ability to optimize multiple hyperparameters efficiently with minimal experiments, facilitating robust and cost-effective model optimization across diverse landscapes. However, it risks oversimplification by focusing on limited factors, potentially overlooking critical interactions or complexities in the data that impact model performance.

Next to tuning the model hyperparameters, various other tests were conducted to optimize model performance:

- Evaluation using F05 instead of Area Under the Curve for Precision-Recall (aucpr) as the evaluation metric.
- Selection of different classification threshold:
  - Based on an equal number of predictions from the last known month.
  - Based on the best classification threshold from the last known month.
  - Based on the best classification threshold from the same month in the previous year.
- Variation of the label threshold: 1, 2, 3, 10.
- Different temporal training methods:
  - Training on a single month; the last month.
  - Training on one year of data.
  - Maximum training (building a new model for each month using the maximum available training data).
  - Iterative training (using the previous month's model as the initial model for the next month).
- Training with and without the forest mask.
- Training with and without geographic coordinates.

### Feature analysis

The Pearson's correlations among the features and between features and the groundtruth was computed (equation 2.4) and visually represented in a correlation plot. This step serves two purposes: (1) identifying highly correlated features that may warrant reconsideration and (2) understanding the correlation between features and the groundtruth, offering initial insights into their predictive capabilities.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (2.4)$$

Another possibility for understanding feature importance involves calculating SHAP (SHapley Additive exPlanations) (Lundberg and Lee 2017) values. The SHAP values determine the contribution of each feature to the model predictions.

Additionally, a forward selection method was employed, iteratively adding features based on their ability to enhance the model. This approach not only offers insights into feature importance but also aids in identifying the optimal subset of features for improved model performance. Initially, features were ordered by importance, and subsequent a model was run with only the two most important features. Following this, additional models were created, with each successive model incorporating the next most important feature.

### Addressing the areas with and without previous deforestation

For each test month, a mask was created to identify areas with previous deforestation (totallossalerts $> 0$, see Appendix A), and the model's performance (precision, recall, and F05-scores) was assessed within and outside this mask. Furthermore, we tested the hypothesis that using a different classification threshold for areas with and without previous deforestation could enhance model performance. Consequently, the model was evaluated across a range of thresholds (from 0.01 to 0.75) for both areas.

### Seasonality

For Laos, the number of deforestation events (groundtruth) and the number of predictions (model trained on the first year) were obtained from January 2021 until May 2023, covering the entire available data time span. Additionally, the average precipitation and the average number of fire alerts in the subsequent half year were obtained for each month. Temporal patterns in these variables were compared to assess seasonality.

## 2.0.5 Problems encountered/limitations and subsequent adaptations

During testing, we discovered that data availability in Laos was limited. Integrated alerts were only accessible from 2021 onward, while features based on integrated alerts required data from up to a year prior (losslastyear). As a result, fair training of the Laos model could only start in 2022. Consequently, the training and testing periods for Laos were adjusted to January 1, 2022, to June 1, 2022, and January 1, 2023, to May 1, 2023, respectively. To assess the impact of the missing data, another model was trained using the data from the first year (2021), and the results were compared.

Another challenge identified during testing is that Laos is not a representative example country. Promising methods observed in Laos do not necessarily translate to success in other countries, and vice versa. Therefore, halfway through the project, it was decided to also consider other countries in the analysis. Specifically, countries already connected to the Forest Foresight project were selected: Gabon, Colombia, Peru, and Bolivia. This broader scope provides a more comprehensive understanding of which methods perform well and how this varies across countries, offering insight into the need for a different approach per country.

The development of the Forest Foresight package coincided with this internship research, leading to some adjustments in the testing process. Initially, models were executed per tile (10 by 10 degrees), but later a country group was adopted as a unit to ensure more consistent predictions within a country. Consequently, model hyperparameters were optimized at the tile level initially, but later models ran at the country group level. It is recommended to re-evaluate hyperparameter selection at the country group level. Notably, training and testing for Laos were consistently conducted at the country level across various tests.

Using coordinates as features in the models can lead to lines in the predictions, as the model may learn that certain latitudes or longitudes correspond to low or high probabilities of deforestation events. While this inclusion of spatial variation is useful, it can result in striped patterns in the predictions, which may appear unrealistic to users. As a result, an additional test is conducted with the exclusion of coordinates to assess the impact on prediction patterns and the models accuracy.

# Chapter 3

# Results

This chapter presents the findings of the internship research, with each section dedicated to describing the results of a specific research objective.

## 3.1 model parameters

### 3.1.1 Hyper-parameters

Overall, performance differences across hyper-parameter sets were modest (Table 3.1). Nrounds, tree depth and datasample size had most influence on the model performance. However, even increasing the nrounds from 10 to 80 only raised the average F05 score by 0.015 across the four tiles. Best parameter settings were generally consistent across the four tiles. Although there were some inconsistencies in parameters like gamma, eta, and relative date, their influence on model performance was relatively low.

Table 3.1: Average F05 scores across various model parameter settings, for different tiles (left) and for Laos (right), derived from Taguchi analysis. The tile models were individually trained at the tile level over a 12-month period (from January 1, 2021, to December 1, 2021) and subsequently tested over a 12-month interval (from June 1, 2022, to May 1, 2023). The Laos model was trained from January 1, 2022, to June 1, 2022, and tested from January 1, 2023, to May 1, 2023, on the county level. To enhance readability, F05 scores are color-coded; the highest values are represented in blue, while lower values are indicated in red.

| | 00N_010E | 10N_080W | 10N_110E | 10S_070W | Average |
|---|---|---|---|---|---|
| **relative date** | | | | | |
| 0 | 0.636 | 0.646 | 0.654 | 0.551 | 0.622 |
| 1 | 0.634 | 0.648 | 0.655 | 0.547 | 0.621 |
| **tree depth** | | | | | |
| 3 | 0.631 | 0.644 | 0.650 | 0.531 | 0.614 |
| 5 | 0.636 | 0.649 | 0.658 | 0.560 | 0.626 |
| 7 | 0.638 | 0.648 | 0.656 | 0.556 | 0.624 |
| **gamma** | | | | | |
| 0 | 0.636 | 0.647 | 0.654 | 0.544 | 0.620 |
| 0.1 | 0.634 | 0.646 | 0.655 | 0.560 | 0.624 |
| 0.3 | 0.635 | 0.648 | 0.655 | 0.543 | 0.620 |
| **eta** | | | | | |
| 0.05 | 0.635 | 0.645 | 0.655 | 0.544 | 0.620 |
| 0.1 | 0.633 | 0.646 | 0.655 | 0.544 | 0.620 |
| 0.3 | 0.637 | 0.649 | 0.654 | 0.559 | 0.625 |
| **subsample** | | | | | |
| 0.3 | 0.637 | 0.648 | 0.657 | 0.558 | 0.625 |
| 0.7 | 0.634 | 0.646 | 0.654 | 0.546 | 0.620 |
| 1 | 0.635 | 0.647 | 0.654 | 0.543 | 0.620 |
| **min child weight** | | | | | |
| 1 | 0.635 | 0.647 | 0.654 | 0.543 | 0.620 |
| 2 | 0.634 | 0.646 | 0.654 | 0.546 | 0.620 |
| 3 | 0.637 | 0.648 | 0.657 | 0.558 | 0.625 |
| **nrounds** | | | | | |
| 10 | 0.630 | 0.641 | 0.652 | 0.526 | 0.612 |
| 80 | 0.639 | 0.650 | 0.658 | 0.562 | 0.627 |
| 200 | 0.637 | 0.650 | 0.655 | 0.559 | 0.625 |
| **datasample** | | | | | |
| 0.1 | 0.633 | 0.644 | 0.651 | 0.541 | 0.617 |
| 0.3 | 0.637 | 0.648 | 0.657 | 0.559 | 0.625 |
| 0.5 | 0.635 | 0.649 | 0.657 | 0.547 | 0.622 |

| Parameter | F05 |
|---|---|
| **tree depth** | |
| 4 | 0.368 |
| 5 | 0.374 |
| **6** | 0.376 |
| **forest mask** | |
| **0** | 0.373 |
| 1 | 0.373 |
| **gamma** | |
| **0.05** | 0.374 |
| 0.1 | 0.372 |
| 0.2 | 0.372 |
| **eta** | |
| 0.2 | 0.365 |
| 0.3 | 0.373 |
| **0.4** | 0.380 |
| **subsample** | |
| 0.2 | 0.371 |
| 0.4 | 0.373 |
| **0.6** | 0.374 |
| **min child weight** | |
| 3 | 0.371 |
| 4 | 0.373 |
| **5** | 0.374 |
| **datasample** | |
| 0.2 | 0.372 |
| **0.4** | 0.373 |
| 0.6 | 0.372 |
| **nrounds** | |
| 50 | 0.361 |
| 100 | 0.374 |
| **150** | 0.383 |

Focusing specifically on the country of Laos, nrounds emerges as the parameter with the largest impact on the F05 score (Table 3.1). Based on the results of the parameter Taguchi analysis, the following model parameters are selected and maintained constant for Laos in subsequent research steps: tree depth = 6, forest mask = no, gamma = 0.05, eta = 0.4, subsample = 0.6, min child weight = 5, datasample = 0.4, nrounds = 150.

### 3.1.2 Additional tests

In Laos, the baseline model achieves an F05 score of 0.520, surpassing the FF model in most conducted tests (Table 3.2, Figure 3.1). The only exception is the test conducted with the optimal classification threshold. Notably, the test using an equal number of predictions yields results similar to those of the baseline. However, it is worth mentioning that the precision of the default model, which is 0.798, outperforms the precision of the baseline, which is 0.700 (Table 3.2).

Table 3.2: F05-score, precision, and recall for additional tests conducted in Laos. All models were trained and tested at the country level. The testing was performed and averaged over the last 5 available months (from January 1, 2023, to May 1, 2023). The model parameters were selected based on the best F05-scores in Table 3.1 and were kept constant across the various tests.

| Test | Description | F05 | Precision | Recall |
|---|---|---|---|---|
| Baseline | Predicted deforestation = deforestation last six months | 0.520 | 0.700 | 0.285 |
| Default | Model trained on 2022-01-01 till 2022-06-01 with a threshold of 0.5 | 0.391 | 0.798 | 0.129 |
| Threshold 0.2 | Selecting the best threshold based on test data | 0.534 | 0.634 | 0.331 |
| Equal number prediction | Selecting threshold based on equal number of predictions last known month | 0.52 | 0.61 | 0.34 |
| Train on tile | Train on tile instead of country | 0.351 | 0.843 | 0.106 |
| Iterative | Initial model (first 18 months, linearly weighted) + iterative training per month | 0.33 | 0.72 | 0.10 |
| Default + 1 month | Default model as initial model and last available month as training data | 0.438 | 0.778 | 0.172 |

Extending the training period did not consistently enhance the model performance across all countries. While maximum training improved the model accuracy for Peru and Bolivia, training over a 12-month period yielded higher accuracy for Laos, Gabon, and Colombia (Figure 3.1). Conversely, training on only one month significantly reduced the accuracy for all countries (Table 3.3). Moreover, switching to F05 as the evaluation metric instead of AUCPR resulted in score improvements only in Gabon, while it led to decreased scores in the other four countries. Similarly, using a forest mask enhanced scores solely in Gabon and resulted in a decrease in F05 score for the other countries. In Peru and Bolivia, the use of the forest mask decreased the F05 score by 0.08 (Figure 3.1). Although selecting the best thresholds proved effective for Laos and Gabon, it reduced the model performance in Colombia, Peru, and Bolivia (Figure 3.1). Excluding geographical coordinates as features resulted in a moderate decrease in F05 values, with an average decrease of 0.07 across the five countries (Table 3.3).

Table 3.3: F05 values for different regions and models. The 1-month model is trained on the data from the last available month (6 months before testing). The 12-month models are trained using data from January 1, 2021, to December 1, 2021, with and without the inclusion of coordinates as features. All models are tested on the period from June 1, 2022, to May 1, 2023 on country group level.

| | 1 month XY | 12 months XY | 12 months no XY |
|---|---|---|---|
| Overall | 0.295 | 0.59 | 0.583 |
| Laos | 0.324 | 0.442 | 0.442 |
| Gabon | 0.195 | 0.673 | 0.614 |
| Colombia | 0.241 | 0.58 | 0.575 |
| Peru | 0.394 | 0.647 | 0.622 |
| Bolivia | 0.234 | 0.501 | 0.543 |

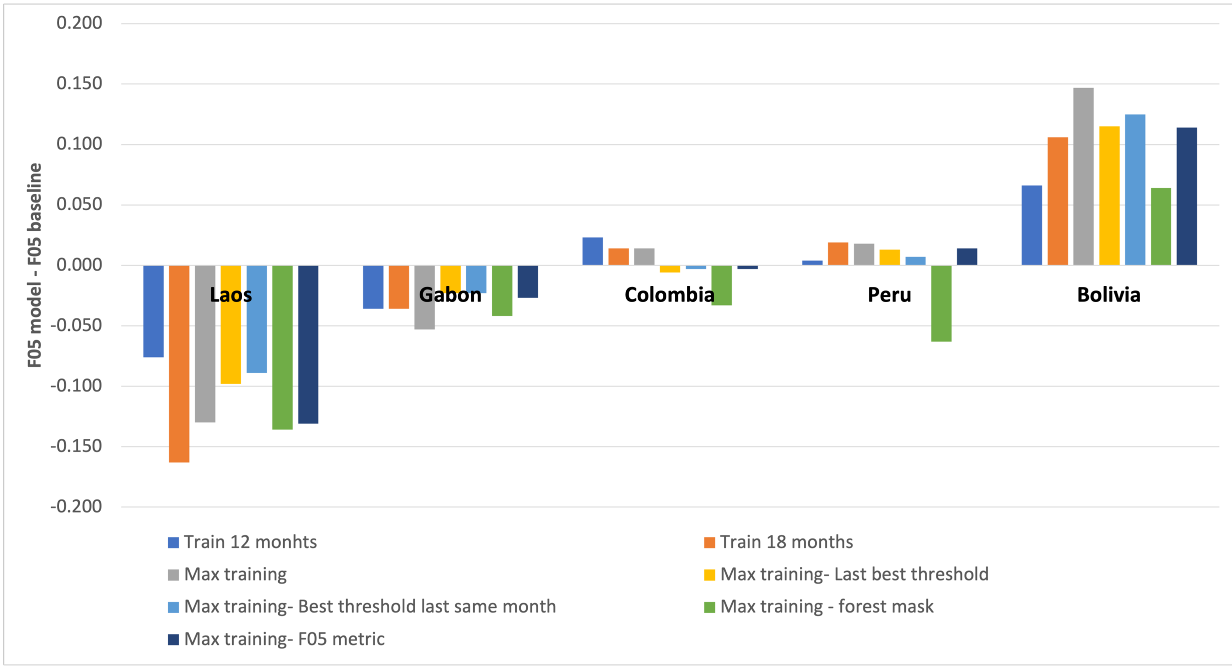



Figure 3.1: F05 scores relative to the baseline for Laos, Gabon, Colombia, Peru, and Bolivia, depicted across different training periods, classification threshold selection methods, and with or without the use of a forest mask. Training durations considered include 12 months, 18 months, and the maximum training duration, with the default classification threshold of 0.5 and AUCPR as the default evaluation metric. Evaluation is based on data from January 1, 2023, to May 1, 2023.

### 3.1.3 Label Threshold

As the label threshold) increased, there was a decrease in precision, recall, and consequently the F05 score for Gabon, Colombia, Peru, and Bolivia, indicating a reduction in overall model performance. However, for Laos, the F05 scores remained equal for label threshold values of 1, 2, and 3. This can be attributed to a slight increase in recall combined with a small decrease in precision. Nonetheless, for a label threshold of 10, the F05 score for Laos also decreased.

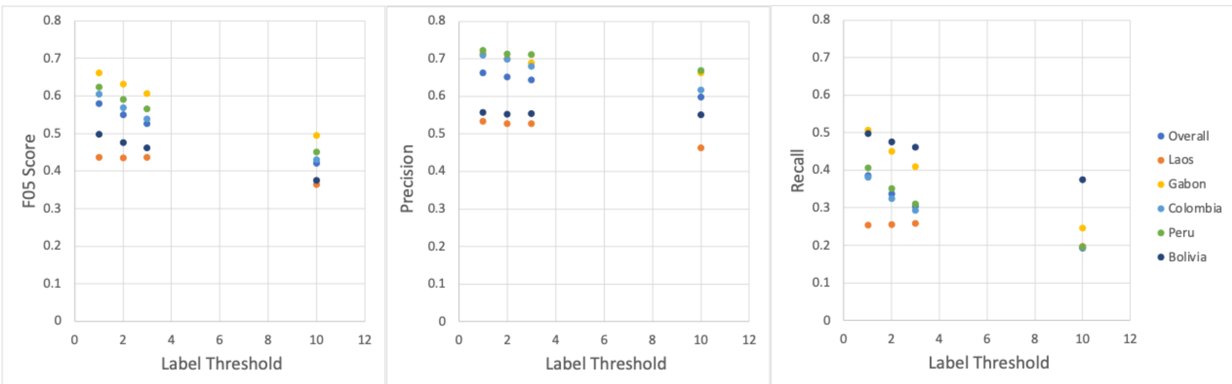



Figure 3.2: Variation of F05 (left), Precision (middle), and Recall (right) with Label Threshold across Laos, Gabon, Colombia, Peru, and Bolivia.

## 3.2 Features analysis

### 3.2.1 Correlations

Within the country of Laos, a strong correlation of 1.00 is observed between the total loss alerts and confidence (Figure 3.3). Additionally, several other feature pairs exhibit high correlations (r > 0.7), including: initial forest cover and forest height (r = 0.89), last three months and last six months (r = 0.75), patch density and last six months (r = 0.71), smoothed total and smoothed six months (r = 0.74), population increase and current population (r = 0.73), x and y coordinates (r = -0.81), and sine month and month (r = -0.75).

Furthermore, the features with the highest correlations with the ground truth are times since loss (r = 0.33), patch density (r = 0.27), smoothed total (r = 0.25), total loss alerts (r = 0.23), confidence (r = 0.23), and last six months (r = 0.20).

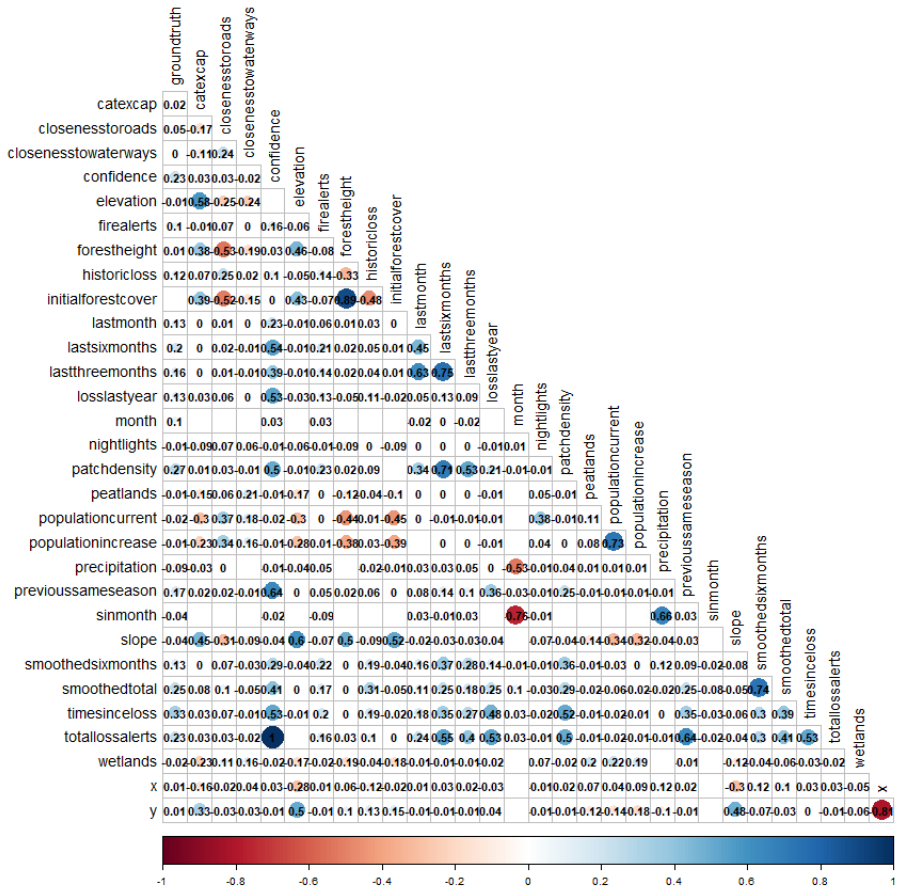



Figure 3.3: Correlation matrix: illustrating the correlations between the features and their correlations with the ground truth in Laos for the year 2022. Positive correlations are depicted in blue, negative correlations in red, with non-significant correlations excluded.

### 3.2.2 Feature importance

The SHAP values provide insights into the most influential features in Laos and help understand the relationship between these features and deforestation alerts. In Laos, the key features identified are smoothed total deforestation (SHAP = 0.788), forest height (0.585), month (0.413), time since loss (SHAP = 0.377), y-coordinate (SHAP = 0.297), historic loss (SHAP = 0.289), slope (SHAP = 0.251), and initial forest cover (SHAP = 0.238) (Figure 3.4).

Using a model with three features (smoothed total, forest height, and historic loss), as opposed to a model with only two features (smoothed total and forest height), led to a significant increase in the F05 score by 0.326 (Figure 3.5). The addition of more features initially showed a slight improvement in model performance, with an average increase in the F05 score of 0.004 per added feature. The highest F05 score of 0.404 was achieved with a model containing 16 features. Using all 31 features resulted in a slightly lower F05 score, with a decrease of 0.006 compared to the maximum F05 score.

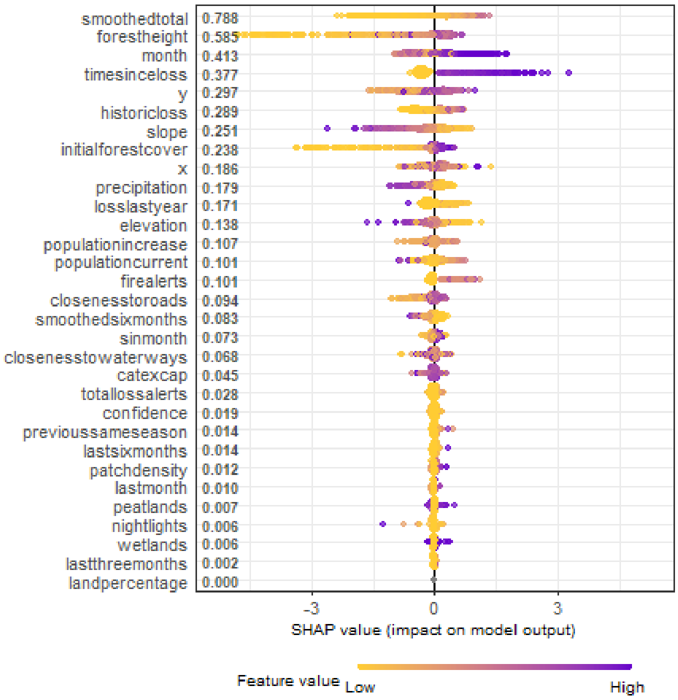



Figure 3.4: Shapley Values for Laos 2022: Positive SHAP values signify a positive impact on deforestation predictions, while negative values indicate a negative influence of the feature. High feature values are visualized in purple, while low feature values are represented in yellow.

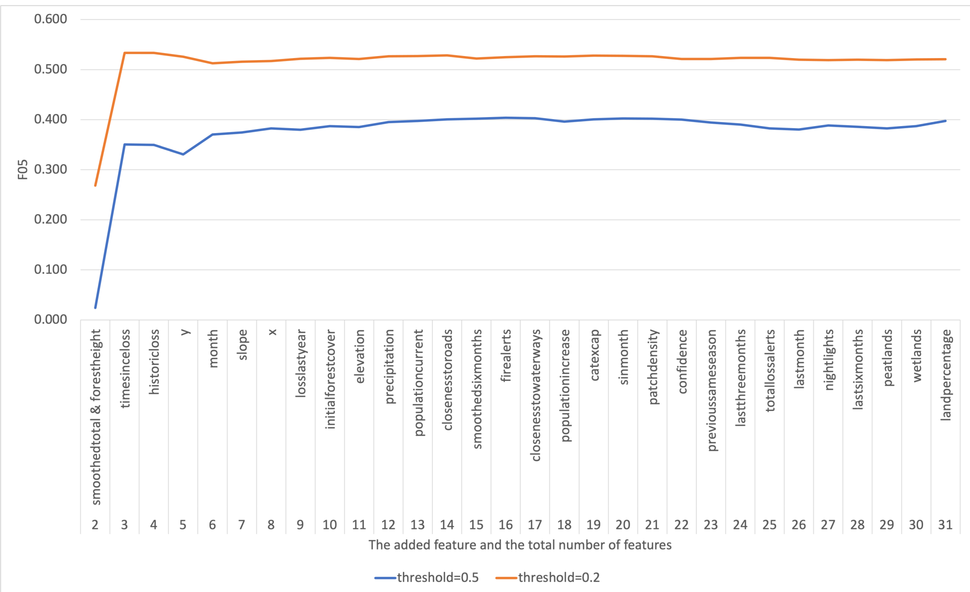


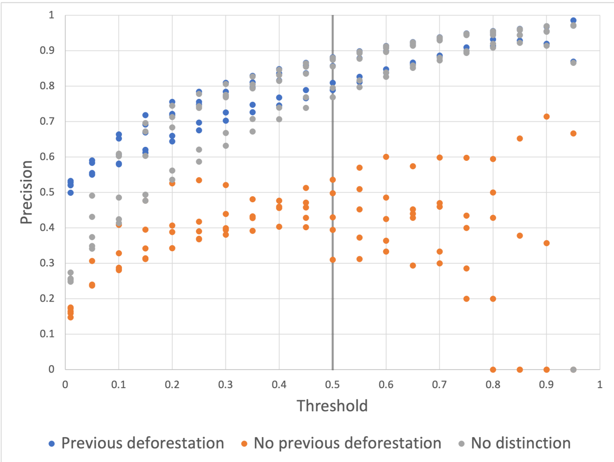

Figure 3.5: F05-scores of Laos for an expanding set of input features, ranked based on their respective feature importance scores. Trained on 2022-01-01 till 2022-06-01 and tested on 2023-01-01 til 2023-05-01.

## 3.3 Distinction areas previous deforestation

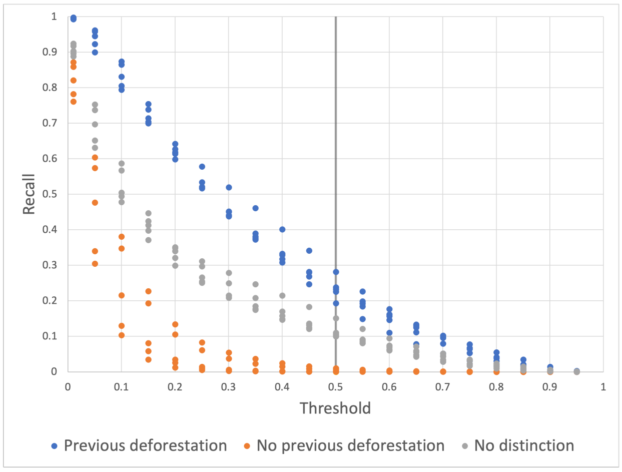
### 3.3.1 Constant classification threshold

At the default classification threshold of 0.5, areas with previous deforestation in Laos exhibited an average precision of 0.843, recall of 0.234, and F05 score of 0.553, while areas without deforestation had an average precision of 0.434, recall of 0.003, and F05 score of 0.015 (Figure 3.6).

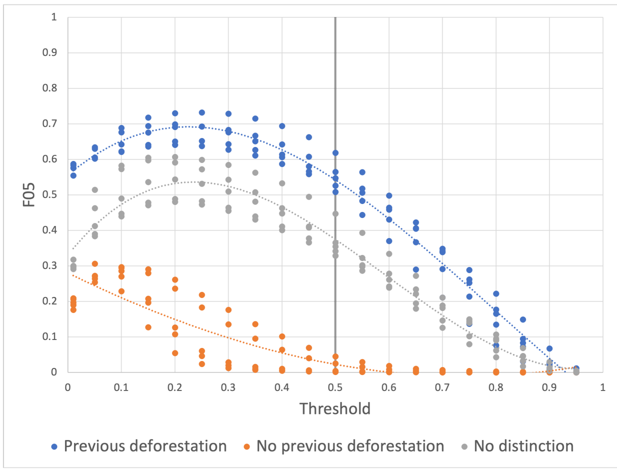
As the classification threshold increases beyond 0.5, precision improves for areas with previous deforestation, while precision for areas without deforestation remains relatively consistent with significant variability across months. Conversely, decreasing the classification threshold increases recall for both areas, resulting in the maximum F05 score for areas with previous deforestation at a threshold of 0.20, and for areas without previous deforestation at a minimum threshold of 0.01 (Figure 3.6 c). When no distinction is made between these areas, the optimal threshold was found to be 0.2.



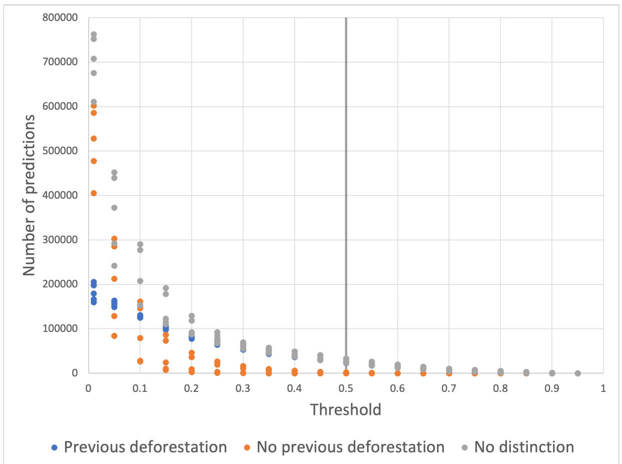

(a) Precision




(b) Recall




(c) F05




(d) Number of predictions

Figure 3.6: The model scores (Precision, Recall and F05) and the total number of predictions in function of the threshold for Laos. Distinctions are made for areas with and without deforestation. The model is trained on 2022-01-01 till 2022-06-01 and tested on 2023-01-01 till 2023-05-01.

### 3.3.2 Different classification thresholds areas with and without previous deforestation

As the classification threshold increases, precision rises while recall declines for both areas with and without deforestation (Table B.1a and B.1b). Particularly, a low threshold in areas with no previous deforestation results in low overall precision, while a high threshold for areas with previous deforestation leads to low overall recall. This trade-off yields low overall F05 scores for high thresholds in areas with previous deforestation and for low thresholds in areas without previous deforestation (Table B.1c). The highest F05 score of 0.559 is achieved with a threshold of 0.15 for areas with previous deforestation and a threshold of 0.45 for areas without previous deforestation. Meanwhile, the highest F05 score for a constant classification threshold is 0.539.

## 3.4 Seasonality

The number of predicted deforestation events in Laos appears to exhibit a seasonal pattern, with the highest number of events occurring from November/December to May/June, and the lowest number occurring from May/June to November/December (Figure 3.7). Additionally, seasonal patterns are also evident in the predictions, with fewer events predicted in the six months following June, and a peak in predictions observed in the six months following February/March/April.

Furthermore, a distinct seasonal pattern emerges in precipitation and fire events. The highest precipitation levels occur during the half-year starting in April, while the peak of fire events is observed in the subsequent half-year following November (Figure 3.7).

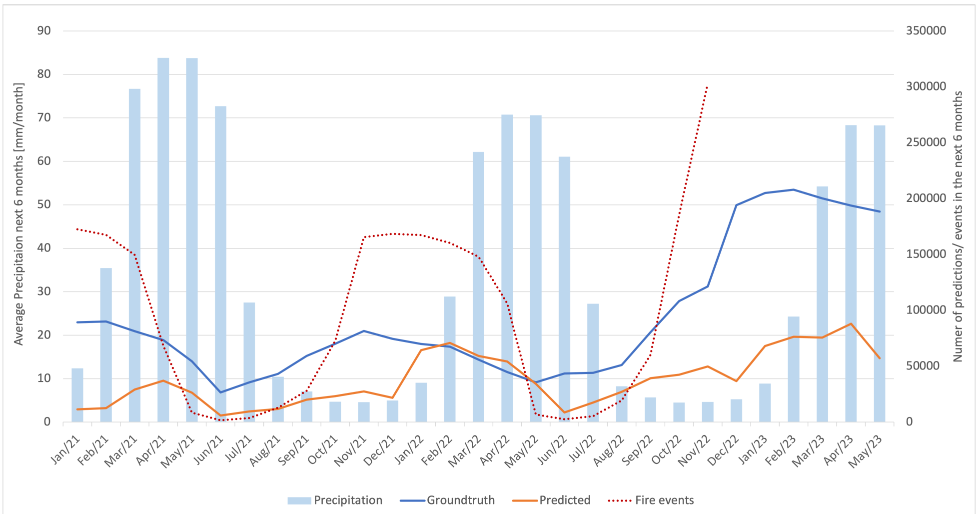



Figure 3.7: Seasonal overview of Laos: total predictions, deforestation events (ground truth), and fire events, along with average monthly precipitation from January 2021 to May 2023. Note: Monthly values represent the sum or average for the subsequent half-year period.

# Chapter 4

# Discussion of results

This chapter discusses the results and provides recommendations for each research objective. The last section (section 4.5) offers additional recommendations for further research.

## 4.1 optimization of model parameters

The impact of hyper-parameter changes on model performance is minimal compared to the variations observed in both spatial and temporal dimensions. Besides, there is little variation in optimal parameters across different tiles. Consequently, setting parameters globally seems acceptable. Moreover, given the limited influence of parameter changes, the absence of identifying the "best" parameter through Taguchi analysis is acceptable. However, these conclusions need reevaluation for the new unit (country group). Moreover, the optimal sample size depends on the sampled area; for instance, a subsample of 0.3 is preferable at the tile level, while a sample size of 0.7 works best for Laos at the country level (Table 3.1). When working with country groups, it is advisable to test whether the sample size can remain constant across groups or should be adjusted based on factors such as the total area or the total forest area of a country group.

The Baseline model outperforms the Forest Foresight model for certain country groups, such as Laos and Gabon. However, for Laos, selecting the optimal classification threshold yields an F05-score surpassing that of the baseline (Table 3.2). This underscores the potential effectiveness of adjusting the threshold in specific scenarios. Yet, determining the best threshold remains challenging without prior knowledge of the ground truth. Although using the best classification threshold from the last known or last same month as a proxy for the current optimal threshold improves model performance, it still falls short of surpassing the baseline accuracy (Figure 3.1). It may be worthwhile to explore methods for predicting the best classification threshold. The method where the threshold is selected based on an equal number of predictions from the last known month performs equally well compared to the baseline and is, for now, the best model for Laos. However, this model's performance across the entire FF implementation area only adds 2% value.

Since none of the countries showed an increase in F05 scores with an increase in the label threshold, it is recommended to keep it equal to 1. The use of F05 as an evaluation metric did not improve the overall model performance and slowed down processing; therefore, it is recommended to continue using aucpr as the evaluation metric.

Counterintuitively, the use of a forest mask for both training and testing decreases model performance. It's possible that there are plantations located outside the forest mask that exhibit predictable alerts. However, it is recommended to double-check the Forest Foresight package for potential errors in the code where the mask is implemented.

Comparing training periods showed that the 12-month option performs better than the 18-month or maximum training methods in countries like Laos and Colombia (Figure 3.1). This suggests the potential influence of seasonal variations on deforestation patterns. Training for a full year ensures an equal representation of all months, whereas extending to 18 months may disproportionately emphasize certain months due to repetition. To address this, maximizing training with a proportional representation of months in the dataset could be a solution, potentially resulting in a more balanced and effective model overall.

Since different methods work well for different countries, it is recommended to select the best methods per country and use this method for the predictions. The recommended methods include training for 12 months or maximal training, with or without a dynamic classification threshold. This results in four distinct methods. However, based on the analysis of the five countries, it is anticipated that this approach will only improve the F05 score per country on average by 0.07. (Figure 3.1).

## 4.2 feature importance and selection

In Laos, several highly correlated features have been identified, particularly the total loss alerts and confidence, which exhibited a correlation of 1.00 (Figure 3.3). For Laos, the initial confidence values provided by the GFW alerts indicate whether there is no alert (0), detected by a single EWS once (2), or detected by a single EWS multiple times (3). For analysis, the average confidence was calculated for 400 by 400 m pixels. The perfect correlation of 1.00 with total loss alerts suggests that the distribution of confidence values remains constant across Laos, regardless of the number of alerts. To reduce correlation between these features, averaging confidence only over the alerts could be considered.

There were several features with correlations above 0.7 (Figure 3.3). The negative correlation observed between the x and y coordinates corresponds to the diagonal shape of Laos (see Figure 2.1). The positive correlation between the deforestation data for the last three months and the last six months is expected, as they partly consists of the same data. It could be beneficial to consider using the deforestation data from three to six months ago instead of the data from the last six months. The correlation between sinmonth and month is understandable since sinmonth is derived from month. Sinmonth is introduced to ensure that January and December are considered adjacent in the analysis. However, XGBoost can learn this relationship, and sinmonth has a relatively low SHAP value (0.073, Figure 3.4). Therefore, it may be worth considering its removal from the feature set.

The correlations of features with the groundtruth indicate the strength of their linear relationships. However, it's important to consider that XGBoost can model nonlinear relationships. As a result, features with the highest correlation with the groundtruth may not always have the highest SHAP values, which quantify the impact of each feature on model predictions. SHAP values account for both linear and nonlinear relationships, providing a more comprehensive understanding of feature importance in XGBoost models.

Based on the SHAP and feature values (Figure 3.4), it's evident that deforestation alerts in Laos are more frequent in areas characterized by several factors. These include high levels of smoothed total deforestation and historic loss, elevated forest heights, and recent deforestation events. Terrain-wise, deforestation events are more prevalent in regions with low slopes and low elevation (aligning with Pujiono et al. 2023). Spatially, deforestation events are concentrated more in the northern parts of Laos, while their distribution in the east-west direction appears more irregular. Furthermore, precipitation has been observed to have a negative influence on the occurrence of deforestation events, which aligns with the documented increase in deforestation during the dry season (Li et al. 2014).

Surprisingly, a model with only 3 features has an F05 score only 0.047 less than the model using all 31 features (Figure 3.5). Adding more features does not necessarily improve the predictions. However, it's worth noting that the model can handle a high number of features, including correlated ones, without a significant decrease in performance. Using all 31 features resulted in a slightly lower F05 score, with a decrease of only 0.006 compared to the maximum F05 score. The advantage of having a high number of features combined with the robust XGBoost model is that different groups of countries might have different important features. While this was only tested in Laos, it could be interesting to conduct similar tests for other country groups.

Since the model with only three features performs almost as well as the model with 31 features, it suggests that many of the features may not be contributing to the predictive performance of the model. Therefore, using Principal Component Analysis (PCA) could make sense. PCA is a dimensionality reduction technique that can help to simplify the model by reducing the number of features while preserving most of the variability in the data. Using PCA can have several benefits: simplification of the model, faster computation, and mitigating overfitting by removing irrelevant features. Although XGBoost was found robust in dealing with correlated and uninformative features, it is recommended to test if the use of PCA improves the overall model performance.

New features could enhance the model's performance. The focus should be on features that don't closely relate to current ones but are strong predictors for deforestation. For example, distance to forest edge has been identified as an important feature (Pujiono et al. 2023), but also expected temperature, or the remaining forest area could be valuable additions. Recently, another intern has started researching which features could be beneficial.

Additionally, refining existing features could make them more effective. As mentioned earlier, adjusting confidence levels and converting deforestation data from the past six months to three to six months ago, could be beneficial. Another example is instead of using all nightlights, focusing only on recent ones to detect changes in forest activity. Similarly, considering only fire events from the previous month, as will be discussed in Section 4.4, might improve predictive capability.

## 4.3 Prediction in deforestation-free areas

The model's performance, as measured by recall, precision, and the F05 score, is low in areas lacking prior deforestation (Figure 3.6). In Laos, the average recall from January to May 2023 for regions without previous deforestation was merely 0.003. This implies that only 0.3% of newly deforested areas were successfully detected, underscoring the challenge of predicting deforestation in such contexts. While this low recall might partly result from the increase in deforestation in Laos (section 4.4), even when employing the optimal classification threshold, the recall in these regions remains modest at 0.06. It could be interesting to examine performance both within and outside the smoothed total deforestation zone. There's a possibility that many of the events detected outside the bounds of previous deforestation areas still occur within this smoothed zone. Further research is useful to develop methods for predicting deforestation in areas lacking prior deforestation. It's possible that employing a different algorithm could enhance performance in these regions. Additionally, communicating the low model performance in these areas to FF users is essential for transparency and managing expectations.

The difference in the optimal classification thresholds between areas with and without previous deforestation hints at the potential advantages of using distinct classification thresholds for these regions (Figure 3.6 (c)). Adopting a low threshold for areas lacking deforestation elevates the F05-score within that area. However, this approach also leads to an abundance of predictions in that region (Figure 3.6 (d)), thus amplifying its influence compared to areas with prior deforestation. Consequently, this results in a reduction of the overall F05-score. The consistent classification threshold resulting in the best F05-score (0.539) was 0.2, which is significantly lower than the default threshold of 0.5. This lower optimal threshold can be attributed to the increase in deforestation events between the model's training and testing phases (Section 4.4). While using separate thresholds can improve the F05 score by 0.02 compared to the best F05 score for a constant classification threshold (Table B.1c), this marginal increase may not justify the need for distinct thresholds. However, this possibility could be explored further across different country groups.

## 4.4 Seasonality

There's a clear seasonality in deforestation events in Laos, with a notable increase during the dry season (Figure 3.7). This trend is also reflected in the SHAP values, indicating a higher number of deforestation alerts when precipitation is low (Figure 3.4). These observations are consistent with the findings of Li et al. 2014, who attribute the seasonal pattern of deforestation to the practice of slashing and burning for shifting cultivation during the dry season. Additionally, the higher number of fire alerts during the dry season (Figure 3.7) further supports this trend. However, it's important to note that the wet season sees more cloud cover, which could potentially impact deforestation detection accuracy. This is particularly relevant given the absence of RADD data in Laos.

The seasonality observed in the deforestation alerts is partly reflected in the predictions as well. The lowest number of deforestation alerts is predicted during the wet season. However, there appears to be a delay of a few months in the peak number of predictions.

For Laos, the number of 400 by 400 m pixels with deforestation alerts over a half-year period remained below 90,000 until the half-year following September 2022. However, thereafter, it exhibited a steep increase, surpassing 200,000 for the half-year period following February 2023 (Figure 3.7). Therefore, the model was trained when deforestation rates were low and tested when they were high. This resulted in a low recall and therefore a low F05 score in Laos. Besides, this explains why a low classification threshold is preferred for this region. The increase in deforestation events aligns with an increase in fire events. It is interesting to understand this gain in fire events; perhaps is it due to an increase in temperature and thereby drought, or to an increase in the practice of slashing and burning for shifting cultivation?

Since there is a half-year gap between the training and predicting phases, it becomes challenging for the model to anticipate an increase or decrease in deforestation during this period. Predicting deforestation over a shorter period, such as one or three months, reduces the gap between training and predicting and provides access to more recent deforestation trends. This adjustment may result in a quicker responding model. Additionally, considering only the fire events from the last few months instead of the entire half-year period may enable the model to demonstrate a faster response to changes in fire events, potentially enhancing its predictive performance.

In addition to the increase in deforestation, several other factors may explain the low scores in Laos. Firstly, there are no RADD alerts available in Southeast Asia, leading to a lower quality of the integrated alerts. Additionally, the integrated alerts from 2020 are missing for Laos, resulting in errors in the features of 2021. Remarkably, the model trained only on this incorrect data of 2021 outperforms the models trained on other time spans. An explanation for this could be that the model learned from this incorrect data to predict more deforestation compared to the previous period, which coincidentally works out well to predict the increased deforestation in 2023.

## 4.5 Remaining gaps/needs for further research

The previous discussion section already provided some recommendations for further research; e.g. testing parameters on country group level, predicting the best classification threshold, applying PCA, refining features and predicting over a different time. This section will introduce a few additional options for further research that may not necessarily correspond to one of the research aims.

First of all, a strategy worth considering is to exclusively train and predict using high-confidence alerts. By prioritizing such alerts, the model can concentrate on more reliable data, potentially enhancing the accuracy of its predictions. Additionally, there's an opportunity to forecast the quantity of pixels with alerts rather than just binary alerts. This approach could provide insights into the extent of deforestation, aiding in the prioritization of interventions. It may be helpful to implement this approach exclusively for the pixels where deforestation is anticipated.

This internship report examined different methods aimed at improving the performance of the XGBoost model, but their impact on the F05 score proved to be limited. This raises doubts about the potential for substantial score improvements solely through the use of XGBoost. Therefore, it's worthwhile to explore alternative algorithms, a task currently undertaken by another intern. Additionally, Wageningen University is developing a deep learning model that exhibits promising results. Moreover, testing the application of a spatial algorithm could be advantageous.

# Chapter 5

# Conclusion

In conclusion, this study on deforestation prediction in Lao People's Democratic Republic using the XGBoost model has provided insights into optimizing model parameters, analyzing feature importance, predicting in deforestation-free areas, and understanding seasonal patterns. From these findings, several key insights and recommendations have emerged.

Adjustments in model parameters had some impact on the model performance, however the improvements were relatively modest across various tests. This highlights the need to explore different methods and algorithms to achieve more significant advancements in prediction accuracy.

In Laos, a simpler model using only three features performed comparably to the model employing all 31 features. Additionally, the feature space included multiple highly correlated features. However, XGBoost demonstrated resilience to correlated and additional features, allowing for the use of the full feature set per country group.

Notably, we found that the model's performance was particularly poor in areas without prior deforestation, underscoring the challenges of predicting deforestation in such contexts. Distinct classification thresholds for deforested and non-deforested areas did not lead to significant improvements. Hence, alternative methods need to be explored to enhance model accuracy outside previous deforestation areas. Additionally, it's crucial to transparently communicate the lower accuracy in these areas to Forest Foresight users.

Moreover, seasonal patterns in deforestation events were observed in Laos, with an increase during the dry period. Although there was a delay in the peak number of events, the model successfully captured this seasonality. However, the model struggled to detect a change in the deforestation trend, leading to a low recall and therefore F05-score for Laos. This was partly solved by the implementation of a dynamic threshold. Additionally, exploring shorter prediction periods and incorporating recent fire events could be tested to enhance the model's responsiveness and accuracy.

Although the low recall resulted in a low overall F05-score in Laos, its precision was satisfactory; when deforestation was predicted, a deforestation event indeed occurred in 80% of the cases. This underscores the usefulness of Forest Foresight for planning interventions.

Moving forward, future research could delve into alternative algorithms, including deep learning models and spatial algorithms, to potentially enhance prediction accuracy. Additionally, investigating the exclusive use of high-confidence alerts and the forecasting of deforestation extent could offer valuable insights for prioritizing interventions.

# Bibliography


Berger, Anika et al. (Mar. 2022). *Looking for the Quickest Signal of Deforestation? Turn to GFW's Integrated Alerts*. Accessed on 2022-03-09. URL: https://blog.globalforestwatch.org/data-and-research/looking-for-the-quickest-signal-of-deforestation-turn-to-gfws-integrated-alerts.

Chen, Tianqi and Carlos Guestrin (2016). "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794. DOI: 10.1145/2939672.2939785.

FAO (2020). *Main report*. Tech. rep. Rome, pp. 1–184. URL: https://doi.org/10.4060/ca9825en.

Hansen, M.C. et al. (2016). "Humid tropical forest disturbance alerts using Landsat data". In: *Environmental Research Letters* 11.3.

Jenkins, Michael and Brian Schaap (2018). "Forest ecosystem services - Background Analytical Study". In: *United Nations Forum on Forests* April, p. 41.

Lestrelin, Guillaume et al. (2013). *The context of REDD+ in the Lao People's Democratic Republic: Drivers, agents and institutions*. CIFOR Occasional Paper 92. Bogor: CIFOR, p. 52. DOI: 10.17528/cifor/004227.

Li, Peng et al. (2014). "A Review of Swidden Agriculture in Southeast Asia". In: *Remote Sensing* 6.2, pp. 1654–1683. ISSN: 2072-4292. DOI: 10.3390/rs6021654. URL: https://www.mdpi.com/2072-4292/6/2/1654.

Lundberg, Scott M. and Su-In Lee (2017). "A Unified Approach to Interpreting Model Predictions". In: *Neural Information Processing Systems*.

Pujiono, Eko et al. (2023). "Factors Contributing to Forest Degradation in the Mountainous Tropical Forest: A Case Study of the Mutis-Timau Forest Complex, Indonesia". In: *Journal of Sustainable Forestry* 42.7, pp. 675–694. DOI: 10.1080/10549811.2022.2123349. URL: https://doi-org.ezproxy.library.wur.nl/10.1080/10549811.2022.2123349.

Reiche, Johannes et al. (2021). "Forest disturbance alerts for the Congo Basin using Sentinel-1". In: *Environmental Research Letters*. DOI: 10.1088/1748-9326/abd0a8.

UNEP, FAO and (2020). *The State of the World's Forests 2020. Forests, biodiversity and people*. Tech. rep. 1. Rome, p. 214. DOI: 10.2307/208372. URL: https://doi.org/10.4060/ca8642en.

Van Stokkom, Anouk et al. (2020). "An Innovative Early Warning System to Tackle Illegal Deforestation An Innovative Early Warning System To Tackle Illegal Deforestation". In: *FIG Working Week 2020* May, pp. 10–14.

WWF (2019). "Deforestation front Laos". In: pp. 32–33.

# Appendices

# Appendix A

# Feature descriptions

Table A.1: The name, periodicity, source, and processing of each feature and the groundtruth layers.

| feature | name | periodicity | source | processing |
|---|---|---|---|---|
| groundtruth-6mbin | binary groundtruth of oncoming six months | monthly | GFW alerts | Every 400x400m pixel where at least 1 deforestation event happens is classified as an actual (value 1) |
| groundtruth1m | groundtruth of oncoming month in amount of pixels | monthly | GFW alerts | Every 400x400m pixel with the total number of deforestation events in the oncoming 1 months as a value (between 0 and 1600) |
| groundtruth3m | groundtruth of oncoming 3 months in amount of pixels | monthly | GFW alerts | Every 400x400m pixel with the total number of deforestation events in the oncoming 3 months as a value (between 0 and 1600) |
| groundtruth6m | groundtruth of oncoming 6 months in amount of pixels | monthly | GFW alerts | Every 400x400m pixel with the total number of deforestation events in the oncoming 6 months as a value (between 0 and 1600) |
| groundtruth12m | groundtruth of oncoming 12 months in amount of pixels | monthly | GFW alerts | Every 400x400m pixel with the total number of deforestation events in the oncoming 12 months as a value (between 0 and 1600) |
| closenessto-roads | closeness to roads | yearly | OSM | This is downloaded using the ohsome package with the filter type:way and highway=* and geometry:line and then further filtered on only lines. The lines are simplified on 10 meters. then the distance in 400x400m pixels is calculated in python. the distance is then converted by the formula 255-20*log(distance+1) |

| feature | name | periodicity | source | processing |
|---|---|---|---|---|
| closenessto-waterways | closeness to waterways | static | OSM | This is downloaded using the ohsome package with the filter type:way and highway=* and geometry:line and then further filtered on only lines. The lines are simplified on 10 meters. then the distance in 400x400m pixels is calculated in python. the distance is then converted by the formula 255-20*log(distance+1) |
| confidence | average confidence of alerts | monthly | GFW alerts | The average confidence of the alerts, which are between 2 (low confidence, detected by one EWS once) and 4 (highest confidence, detected by multiple EWS multiple times) meaning low to highest confidence with NA (not a number) removed |
| elevation | elevation | static | SRTM | The average elevation of the SRTM rasters when aggregated to the 400x400m pixel level in meters above sea level |
| firealerts | forest fire alerts from VIIRS and other satellites | monthly | VIIRS LAADS-DAAP | The data is downloaded for all sensors for historic data and from now on monthly for near real time data. The point shapefiles are rasterized with the sum of all points within the pixel. Thus the number represents the number of forest fires within the pixel of the last 6 months |
| forestheight | height of forest 2020 | yearly | GLAD | This dataset is available till 2020 and represents the average forest height in meters within the pixel, where 0 means no forest |
| historicloss | GLAD deforestation 2001-2018 | static | GLAD FSC alerts | The total number of pixels between 2001 and 2018 being deforested according to the yearly report |
| initialforest-cover | forest mask of 2019 | static | GLAD FSC alerts | The sum of the underlying pixels in the 400x400m pixel where every underlying pixel is a fraction of canopy cover and every pixel that has been deforested between 2001 and 2018 is set to 0 |
| lastsixmonths | deforestation last six months | monthly | GFW alerts | Takes the sum of the underlying 1600 10x10m pixels that have been deforested in the last 6 months |
| lastthree-months | deforestation last three months | monthly | GFW alerts | Takes the sum of the underlying 1600 10x10m pixels that have been deforested in the last 3 months |
| lastmonth | deforestation in the previous month | monthly | GFW alerts | Takes the sum of the underlying 1600 10x10m pixels that have been deforested in the previous month |

| feature | name | periodicity | source | processing |
|---|---|---|---|---|
| losslastyear | GLAD deforestation last year | yearly | GLAD FSC alerts | The total number of underlying pixels as a sum from the previous year |
| month | the number of the month | static | auto-generated | This is autogenerated in the preprocessing |
| nightlights | nighttime activity as measured by the VIIRS satellite | monthly | VIIRS LAADS-DAAP | Monthly cloud-corrected nighttime activity (visible light/SWIR radiation at night) where the values represent the amount of radiation |
| patchdensity | patchiness of last six months of deforestation | monthly | GFW alerts | The deforestation of the last six months is taken as input and processed by giving every group of deforestation pixels a unique value. The total amount of unique values within the 400x400m pixel is the result, representing the amount of deforestation patches in the last 6 months |
| total-deforestation | The total number of deforested pixels according to the integrated alerts | monthly | GFW alerts | Takes the sum of the underlying 1600 10x10m pixels that have been deforested since the beginning of the EWS (normally 1st of January 2020) |
| peatland | areas with a peat soil | static | CIFOR | The data is aggregated to 400x400 and represents the average depth of the peat soil in meters with a maximum of 15 meters |
| population-current | current population | static | EUPOP | This is the POP2025 (GHS-POP) dataset that is first reprojected to the 400x400m pixel resolution with the max function and then a gradual filter of 25 pixels is applied to smooth it over a larger area |
| population-increase | population increase | static | EUPOP | This is the difference between POP2030 (the expected population in 2030, see source GHS-POP) and POP2020 (the known population according to census in 2020, see source GHS-POP) dataset that is first reprojected to the 400x400m pixel resolution with the max function and then a gradual filter of 25 pixels is applied to smooth it over a larger area |

| feature | name | periodicity | source | processing |
|---|---|---|---|---|
| precipitation | predicted precipitation | monthly | CMIP6 precipitation predictions | Averaged (median) over all 16 scenarios and with a cubic resampling resampled to 400x400m pixels. The actual values represent the average (mean) amount of expected precipitation in the 6 months following the date of the dataset in mm/month |
| previous-sameseason | deforestation 6-12 months ago | monthly | GFW alerts | Takes the sum of the underlying 1600 pixels that have been deforested 6-12 months to account for seasonality |
| sinmonth | the sine of the date | static | auto-generated | This is autogenerated in the preprocessing and represents the sine value of the day of the year, between -1 and 1 |
| slope | slope | static | SRTM | Slope is calculated using the terrain function in terra using standard settings and then taken the average of to aggregate |
| smoothed-sixmonths | deforestation last six months smoothed | monthly | GFW alerts | A smoothed version of layer lastsixmonths. This means that the further away from the aggregated 400x400m pixel the pixel in question is, the lower the value gets. |
| smoothedtotal | total deforestation smoothed | monthly | GFW alerts | A smoothed version of layer totaldeforestation (the total number of deforested pixels according to the integrated alerts). This means that the further away from the aggregated 400x400m pixel the pixel in question is, the lower the value gets. |
| timesinceloss | latest moment of deforestation | monthly | GFW alerts | This is classified as highest value (representing latest moment of deforestation) since 1-1-2015 divided by the current date multiplied by 10.000 to make it 16-bit |
| totallossalerts | total deforestation | monthly | GFW alerts | The total amount of the underlying 1600 pixels deforested according to the GFW alerts since the alerts started |
| wetlands | areas with wetlands | static | FAO/WWF | The data is aggregated to 400x400 and reclassified from categorical values to 1/0 for wetland/no wetland by taking any value above 0 |
| x | latitude | static | auto-generated | This is autogenerated in the preprocessing and represents the latitude per pixel between -90 and 90 |

| feature | name | periodicity | source | processing |
|---|---|---|---|---|
| y | longitude | static | auto-generated | This is autogenerated in the preprocessing and represents the longitude per pixel between -180 and 180 |
| landpercent-age | percentage of land cover (as opposed to sea) | static | GADM | This is the percentage landcover in values from 0 (no land) to 255 (100% landcover of the pixel) |
| catexcap | cation exchange capacity at 0-5cm | static | ISRIC | The data was aggregated to 400x400 meter by using the mode of the underlying pixels (most occurring). The cation exchange capacity in mmol(c)/kg is a useful indicator for soil fertility, a proxy for deforestation for agriculture |
| wdpa | WDPA status | yearly | Protected Planet | The data was intersected with land area and the tiles that we have, then simplified to 400 meters and then rasterized. A value of 1 means that the entire area is protected area, 0 means no protected area within the pixel |

# Appendix B

# Separate threshold scores

Table B.1: Precision, Recall, and F05 Score for predictions in Laos using separate thresholds, ranging from 0.01 till 0.75, for areas with and without previous deforestation. The model was trained on data from January 1, 2022, to June 1, 2022, and tested on data from January 1, 2023, to May 1, 2023. Red colors indicate low scores, while green colors indicate high scores.

(a) Precision

| Threshold previous deforestation | Threshold no previous deforestation | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0.01** | **0.05** | **0.1** | **0.15** | **0.2** | **0.25** | **0.3** | **0.35** | **0.4** | **0.45** | **0.5** | **0.55** | **0.6** | **0.65** | **0.7** | **0.75** |
| **0.01** | **0.257** | 0.383 | 0.454 | 0.489 | 0.505 | 0.513 | 0.517 | 0.519 | 0.520 | 0.520 | 0.520 | 0.520 | 0.520 | 0.520 | 0.520 | 0.520 |
| **0.05** | 0.259 | **0.395** | 0.479 | 0.523 | 0.545 | 0.555 | 0.560 | 0.563 | 0.565 | 0.566 | 0.566 | 0.566 | 0.566 | 0.566 | 0.566 | 0.566 |
| **0.1** | 0.255 | 0.401 | **0.501** | 0.556 | 0.585 | 0.600 | 0.608 | 0.612 | 0.614 | 0.616 | 0.616 | 0.617 | 0.617 | 0.617 | 0.617 | 0.617 |
| **0.15** | 0.248 | 0.401 | 0.514 | **0.582** | 0.619 | 0.638 | 0.649 | 0.655 | 0.658 | 0.660 | 0.661 | 0.662 | 0.662 | 0.662 | 0.662 | 0.662 |
| **0.2** | 0.240 | 0.395 | 0.519 | 0.598 | **0.643** | 0.668 | 0.681 | 0.689 | 0.694 | 0.697 | 0.698 | 0.699 | 0.699 | 0.700 | 0.700 | 0.700 |
| **0.25** | 0.232 | 0.387 | 0.519 | 0.608 | 0.661 | **0.691** | 0.708 | 0.718 | 0.724 | 0.728 | 0.730 | 0.731 | 0.731 | 0.732 | 0.732 | 0.732 |
| **0.3** | 0.225 | 0.377 | 0.515 | 0.613 | 0.673 | 0.708 | **0.728** | 0.741 | 0.749 | 0.753 | 0.756 | 0.757 | 0.758 | 0.758 | 0.758 | 0.758 |
| **0.35** | 0.217 | 0.367 | 0.508 | 0.613 | 0.680 | 0.720 | 0.744 | **0.759** | 0.769 | 0.775 | 0.778 | 0.780 | 0.781 | 0.781 | 0.782 | 0.782 |
| **0.4** | 0.211 | 0.356 | 0.498 | 0.610 | 0.683 | 0.729 | 0.757 | 0.775 | **0.787** | 0.794 | 0.798 | 0.800 | 0.801 | 0.802 | 0.802 | 0.802 |
| **0.45** | 0.204 | 0.344 | 0.485 | 0.603 | 0.683 | 0.734 | 0.767 | 0.788 | 0.802 | **0.811** | 0.816 | 0.819 | 0.821 | 0.822 | 0.822 | 0.823 |
| **0.5** | 0.197 | 0.331 | 0.471 | 0.593 | 0.680 | 0.737 | 0.774 | 0.799 | 0.817 | 0.828 | **0.834** | 0.838 | 0.840 | 0.841 | 0.842 | 0.842 |
| **0.55** | 0.192 | 0.319 | 0.454 | 0.580 | 0.673 | 0.736 | 0.778 | 0.808 | 0.828 | 0.842 | 0.850 | **0.855** | 0.858 | 0.860 | 0.861 | 0.861 |
| **0.6** | 0.186 | 0.307 | 0.436 | 0.562 | 0.662 | 0.731 | 0.778 | 0.812 | 0.838 | 0.855 | 0.865 | 0.872 | **0.876** | 0.878 | 0.879 | 0.880 |
| **0.65** | 0.181 | 0.296 | 0.416 | 0.541 | 0.644 | 0.720 | 0.773 | 0.811 | 0.841 | 0.862 | 0.876 | 0.885 | 0.890 | **0.893** | 0.895 | 0.896 |
| **0.7** | 0.177 | 0.285 | 0.396 | 0.515 | 0.621 | 0.703 | 0.761 | 0.805 | 0.840 | 0.866 | 0.884 | 0.896 | 0.903 | 0.907 | **0.910** | 0.912 |
| **0.75** | 0.173 | 0.275 | 0.376 | 0.485 | 0.589 | 0.676 | 0.741 | 0.789 | 0.830 | 0.862 | 0.885 | 0.902 | 0.913 | 0.919 | 0.923 | **0.926** |

(b) Recall

| Threshold previous deforestation | Threshold No previous deforestation | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0.01** | **0.05** | **0.1** | **0.15** | **0.2** | **0.25** | **0.3** | **0.35** | **0.4** | **0.45** | **0.5** | **0.55** | **0.6** | **0.65** | **0.7** | **0.75** |
| **0.01** | **0.905** | 0.721 | 0.603 | 0.541 | 0.511 | 0.496 | 0.488 | 0.484 | 0.482 | 0.480 | 0.479 | 0.478 | 0.477 | 0.477 | 0.477 | 0.477 |
| **0.05** | 0.877 | **0.692** | 0.575 | 0.513 | 0.482 | 0.467 | 0.460 | 0.456 | 0.453 | 0.451 | 0.450 | 0.450 | 0.449 | 0.449 | 0.449 | 0.449 |
| **0.1** | 0.827 | 0.642 | **0.525** | 0.463 | 0.432 | 0.417 | 0.410 | 0.406 | 0.403 | 0.402 | 0.400 | 0.400 | 0.399 | 0.399 | 0.399 | 0.399 |
| **0.15** | 0.774 | 0.589 | 0.472 | **0.410** | 0.379 | 0.364 | 0.357 | 0.353 | 0.350 | 0.348 | 0.347 | 0.346 | 0.346 | 0.346 | 0.346 | 0.345 |
| **0.2** | 0.725 | 0.541 | 0.423 | 0.361 | **0.331** | 0.316 | 0.308 | 0.304 | 0.302 | 0.300 | 0.299 | 0.298 | 0.297 | 0.297 | 0.297 | 0.297 |
| **0.25** | 0.684 | 0.500 | 0.382 | 0.320 | 0.290 | **0.275** | 0.267 | 0.263 | 0.260 | 0.259 | 0.257 | 0.257 | 0.256 | 0.256 | 0.256 | 0.256 |
| **0.3** | 0.649 | 0.464 | 0.347 | 0.285 | 0.254 | 0.239 | **0.232** | 0.228 | 0.225 | 0.224 | 0.222 | 0.222 | 0.221 | 0.221 | 0.221 | 0.221 |
| **0.35** | 0.618 | 0.434 | 0.316 | 0.254 | 0.224 | 0.209 | 0.201 | **0.197** | 0.195 | 0.193 | 0.192 | 0.191 | 0.191 | 0.190 | 0.190 | 0.190 |
| **0.4** | 0.590 | 0.406 | 0.288 | 0.226 | 0.196 | 0.181 | 0.173 | 0.169 | **0.167** | 0.165 | 0.164 | 0.163 | 0.163 | 0.162 | 0.162 | 0.162 |
| **0.45** | 0.564 | 0.380 | 0.262 | 0.200 | 0.170 | 0.155 | 0.147 | 0.143 | 0.141 | **0.139** | 0.138 | 0.137 | 0.137 | 0.136 | 0.136 | 0.136 |
| **0.5** | 0.540 | 0.356 | 0.238 | 0.176 | 0.146 | 0.131 | 0.124 | 0.119 | 0.117 | 0.115 | **0.114** | 0.113 | 0.113 | 0.112 | 0.112 | 0.112 |
| **0.55** | 0.519 | 0.335 | 0.217 | 0.155 | 0.125 | 0.110 | 0.102 | 0.098 | 0.096 | 0.094 | 0.093 | **0.092** | 0.092 | 0.091 | 0.091 | 0.091 |
| **0.6** | 0.500 | 0.316 | 0.198 | 0.136 | 0.106 | 0.091 | 0.083 | 0.079 | 0.077 | 0.075 | 0.074 | 0.073 | **0.073** | 0.072 | 0.072 | 0.072 |
| **0.65** | 0.483 | 0.299 | 0.181 | 0.119 | 0.089 | 0.074 | 0.067 | 0.062 | 0.060 | 0.058 | 0.057 | 0.056 | 0.056 | **0.055** | 0.055 | 0.055 |
| **0.7** | 0.469 | 0.284 | 0.167 | 0.105 | 0.074 | 0.059 | 0.052 | 0.048 | 0.045 | 0.043 | 0.042 | 0.042 | 0.041 | 0.041 | **0.041** | 0.041 |
| **0.75** | 0.456 | 0.272 | 0.154 | 0.092 | 0.062 | 0.047 | 0.039 | 0.035 | 0.033 | 0.031 | 0.030 | 0.029 | 0.029 | 0.028 | 0.028 | **0.028** |

(c) F05

| Threshold previous deforestation | Threshold No previous deforestation | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **0.01** | **0.05** | **0.1** | **0.15** | **0.2** | **0.25** | **0.3** | **0.35** | **0.4** | **0.45** | **0.5** | **0.55** | **0.6** | **0.65** | **0.7** | **0.75** |
| **0.01** | **0.300** | 0.421 | 0.476 | 0.498 | 0.506 | 0.509 | 0.510 | 0.511 | 0.511 | 0.511 | 0.511 | 0.511 | 0.510 | 0.510 | 0.510 | 0.510 |
| **0.05** | 0.301 | **0.431** | 0.494 | 0.520 | 0.531 | 0.535 | 0.537 | 0.538 | 0.538 | 0.538 | 0.538 | 0.538 | 0.538 | 0.537 | 0.537 | 0.537 |
| **0.1** | 0.296 | 0.432 | **0.502** | 0.533 | 0.546 | 0.551 | 0.554 | 0.555 | 0.556 | 0.556 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 | 0.555 |
| **0.15** | 0.287 | 0.425 | 0.500 | **0.534** | 0.548 | 0.554 | 0.557 | 0.558 | 0.559 | 0.559 | 0.559 | 0.559 | 0.559 | 0.559 | 0.559 | 0.559 |
| **0.2** | 0.277 | 0.414 | 0.490 | 0.525 | **0.539** | 0.545 | 0.548 | 0.550 | 0.550 | 0.550 | 0.550 | 0.550 | 0.550 | 0.550 | 0.550 | 0.550 |
| **0.25** | 0.267 | 0.401 | 0.476 | 0.510 | 0.524 | **0.529** | 0.532 | 0.533 | 0.534 | 0.533 | 0.533 | 0.533 | 0.533 | 0.533 | 0.533 | 0.532 |
| **0.3** | 0.258 | 0.387 | 0.459 | 0.490 | 0.503 | 0.507 | **0.509** | 0.510 | 0.511 | 0.510 | 0.510 | 0.510 | 0.509 | 0.509 | 0.509 | 0.509 |
| **0.35** | 0.250 | 0.373 | 0.440 | 0.468 | 0.478 | 0.481 | 0.483 | **0.483** | 0.483 | 0.482 | 0.482 | 0.481 | 0.481 | 0.481 | 0.481 | 0.480 |
| **0.4** | 0.242 | 0.359 | 0.420 | 0.443 | 0.450 | 0.451 | 0.451 | 0.450 | **0.450** | 0.449 | 0.448 | 0.448 | 0.447 | 0.447 | 0.447 | 0.447 |
| **0.45** | 0.234 | 0.344 | 0.398 | 0.415 | 0.418 | 0.416 | 0.415 | 0.413 | 0.412 | **0.411** | 0.410 | 0.409 | 0.408 | 0.408 | 0.408 | 0.408 |
| **0.5** | 0.226 | 0.329 | 0.375 | 0.385 | 0.382 | 0.378 | 0.375 | 0.372 | 0.370 | 0.369 | **0.367** | 0.366 | 0.365 | 0.365 | 0.365 | 0.365 |
| **0.55** | 0.219 | 0.315 | 0.352 | 0.353 | 0.346 | 0.338 | 0.332 | 0.329 | 0.326 | 0.323 | 0.322 | **0.320** | 0.320 | 0.319 | 0.319 | 0.318 |
| **0.6** | 0.213 | 0.302 | 0.329 | 0.322 | 0.308 | 0.296 | 0.288 | 0.283 | 0.279 | 0.276 | 0.274 | 0.272 | **0.271** | 0.270 | 0.270 | 0.270 |
| **0.65** | 0.207 | 0.289 | 0.306 | 0.290 | 0.270 | 0.253 | 0.243 | 0.236 | 0.232 | 0.228 | 0.225 | 0.223 | 0.222 | **0.221** | 0.221 | 0.220 |
| **0.7** | 0.202 | 0.278 | 0.285 | 0.260 | 0.232 | 0.211 | 0.199 | 0.190 | 0.184 | 0.180 | 0.176 | 0.174 | 0.173 | 0.172 | **0.172** | 0.171 |
| **0.75** | 0.197 | 0.267 | 0.266 | 0.231 | 0.197 | 0.172 | 0.156 | 0.146 | 0.139 | 0.134 | 0.130 | 0.128 | 0.126 | 0.125 | 0.124 | **0.124** |