Addressing Spatio-Temporal Non-Stationarity in Land Cover Change.

The case of the deforestation prediction in Brazil.

Global Land Programme / 5th Open Science Meeting
4-8 November

Rodrigo Lopez-Farias

Sergio Ivvan Valdez Peña

Alberto García Robledo

Jesús Arturo Monroy Anieva

CONAHCYT-CENTROGEO. Queretaro, Mexico.

Introduction

There are methodologies for addressing land use and land cover change prediction, however, the research about deforestation is minimal. Therefore, this research proposal explores:

- The problem of predicting temporal non-stationarity deforestation represented by binary data with Convolutional Neural Networks.
- Characterizing the spatial non-stationarity and its modelling with neural networks

This study focuses on the Brazilian state of Bahia to detect and analyze the spatio-temporal non-stationarity of deforestation.

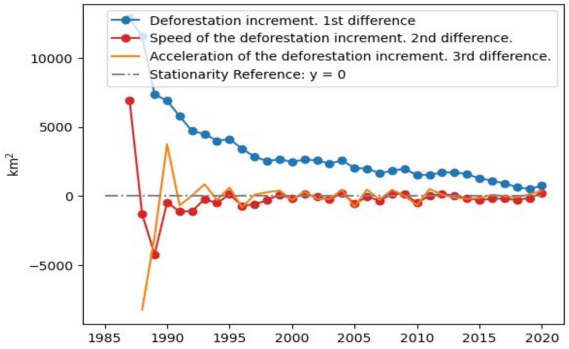
Informed forest management is particularly important in reforestation, wood production, and fire and deforestation control and prevention.

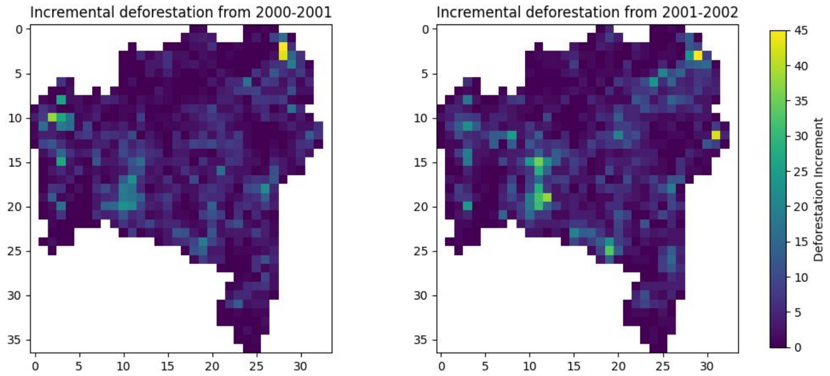
Mathematical models using remote sensing data predict deforestation patterns and guide public policies, but few models predict future forest-degradation risks.

Introduction

Land deforestation, is often considered as a non-stationary process. It is influenced by both the quantity of deforested land and the spatial allocation of this change over time.

To address this, we propose to visualize the data to verify the existence of non-stationarity behavior, and then, use a model to predict deforestation that fits two key components: **temporal deforestation trends** and **shifts in spatial distribution.**





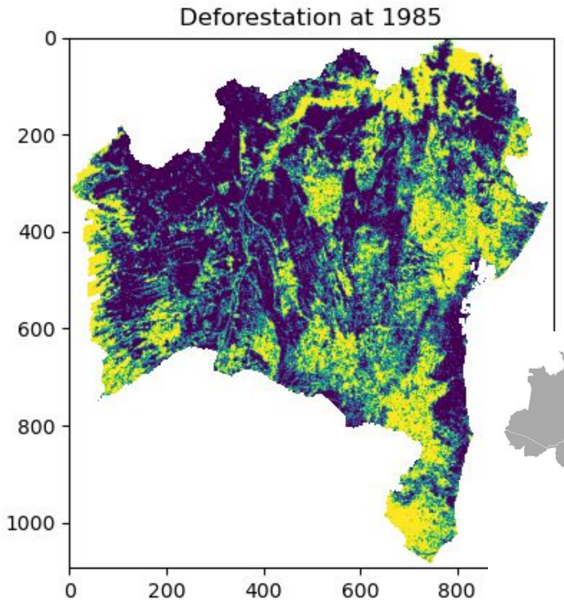
Introduction

For predicting the quantity of land deforestation, we use time-series statistical theory to justify the use of convolution operators to deal with temporal non-stationarity.

For spatial allocation, we analyze distribution changes over time using sequences of spatial density functions that describe the distribution of the deforestation across space at different time steps.

We demonstrate this approach with deforestation data from Bahia, Brazil, predicting from 2003 to 2020. Our results suggest that analyzing multiple observations in non-stationary data enhances predictive accuracy with respect a naïve model.

Database Description



The original database has 36 binary maps representing the presence of deforestation.

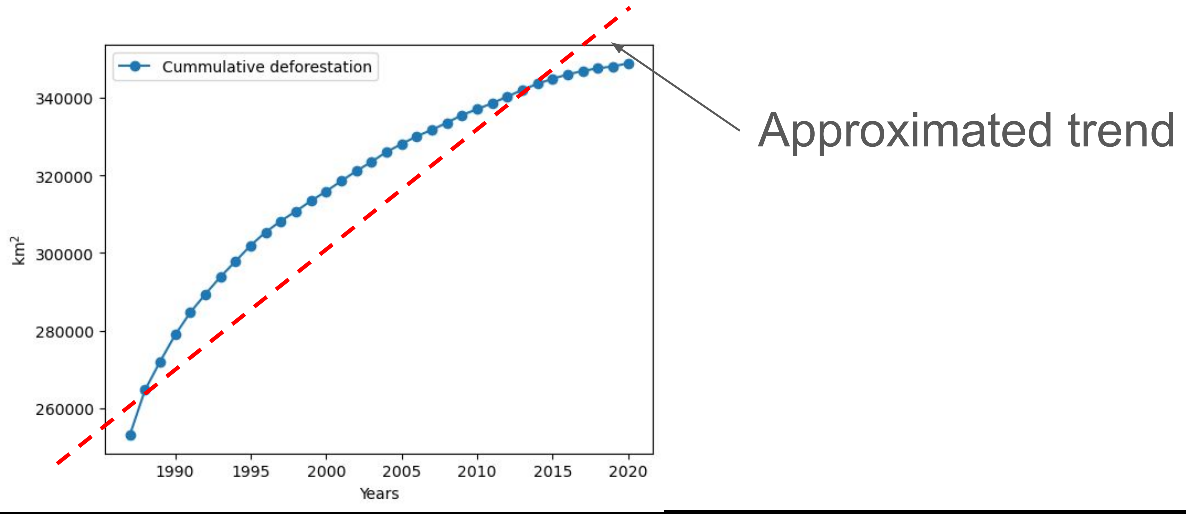
The region of interest has 566 963 cells. Each map pixel represents a region of 1km x 1km.

The deforestation data set is from https://amazonia.mapbiomas.org/

Temporal non-stationarity

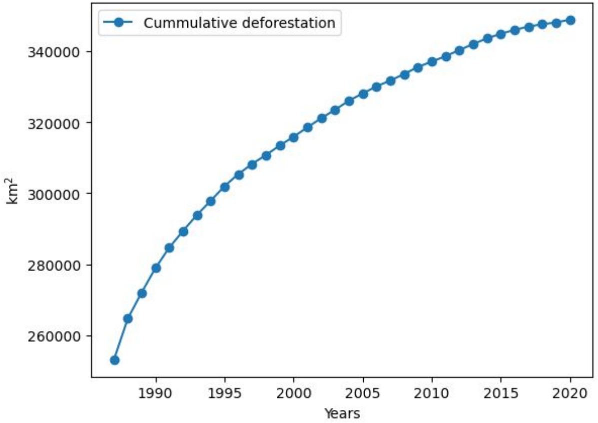
A stationary time series is that one whose properties such as its mean, is approximately the same regardless the time. Thus, time series with trends or seasonality, are not stationary. This is the case of the cumulative deforestation time series of Bahía, Brazil.

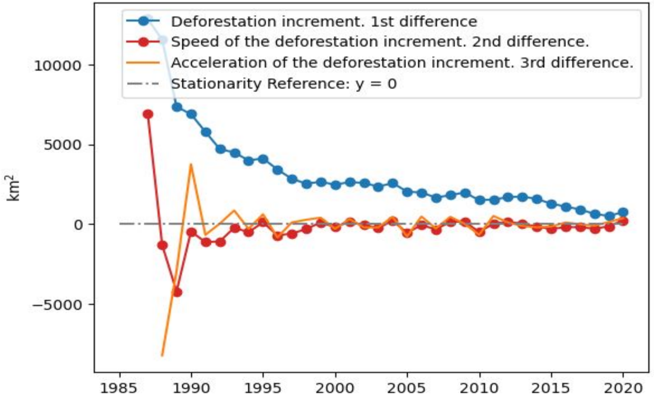
.



Temporal non-stationarity

In order to stationarize the time series, it is necessary differentiate it several times until we obtain a time series with stable mean.





The interpretation of the first difference is the increment of the deforestation in one-year interval. Nevertheless, it still presents a strong negative trend.

Differencing

- First-order differencing

It consists of calculating the difference of every pair of consecutive values of a time series.

$$y'_t = y_t - y_{t-1}$$

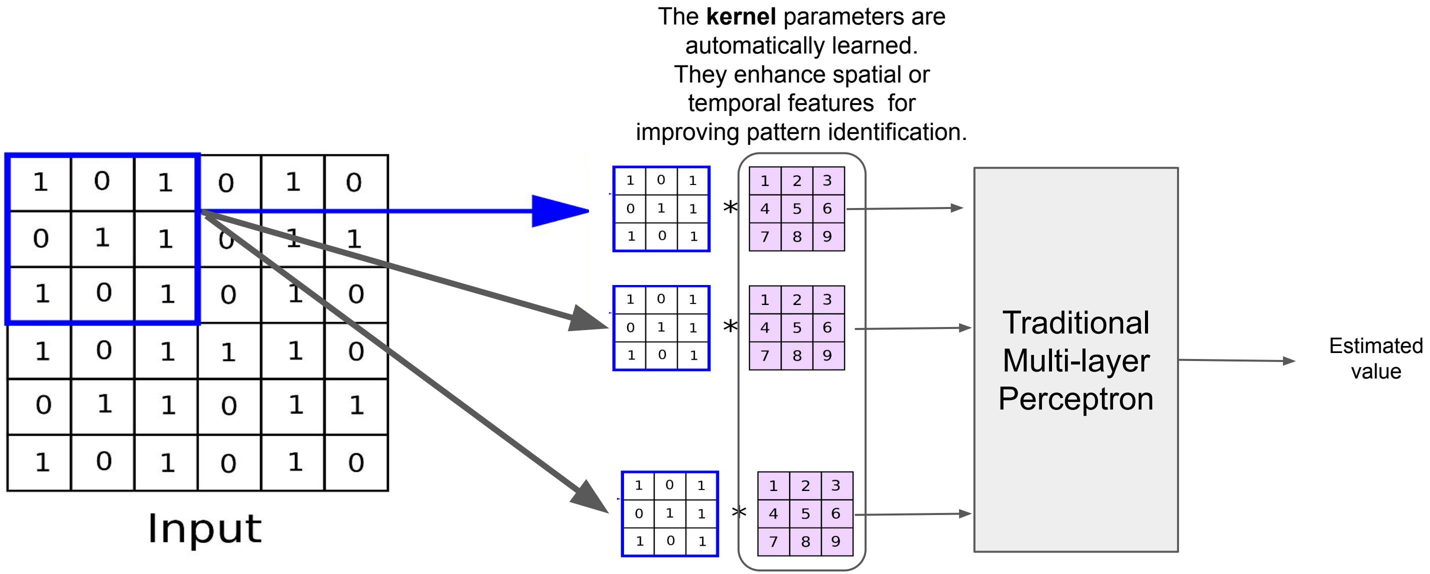
- Second-Order Differencing

It is used when the first differentiated data presents non-stationarity.

$$y''_t = y'_t - y'_{t-1}$$
$$y''_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$
$$y''_t = y_t - 2y_{t-1} + y_{t-2}$$

Convolutional Neural Networks (CNN)

Are a kind of neural networks which considers the elemental structure of the data in the intermediate layers. They can be applied to times series, images, or series of time images



Original Image: https://anhreynolds.com/blogs/cnn.html

The relation of Convolutional Neural Networks with the Differencing

The differencing is a special case of a convolution operation with fixed parameters.

First Order differencing

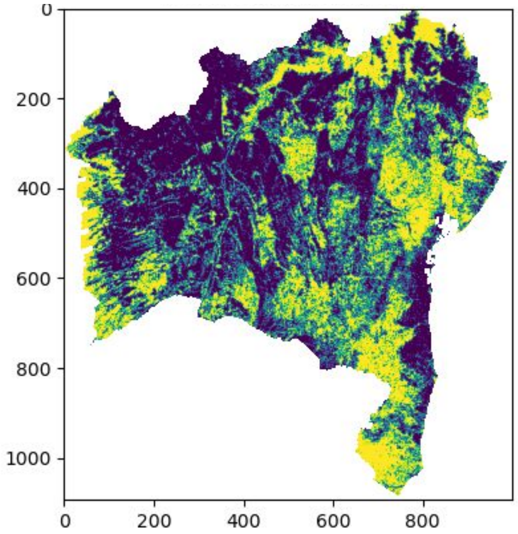
$$y'_t = y_t - y_{t-1}$$
$$y'_t = [y_t, y_{t-1}] * K$$
$$\text{where}$$
$$K = [k_0, k_1] = [1, -1]$$
$$\text{thus}$$
$$y'_t = y_t - y_{t-1} = y_t \cdot k_0 + y_{t-1} \cdot k_1$$

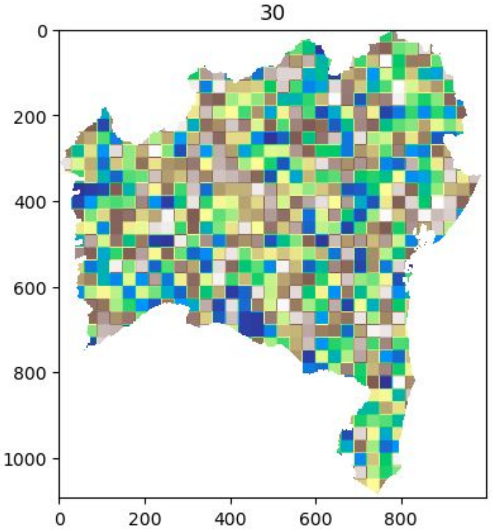
Second Order differencing

$$y''_t = y_t - 2y_{t-1} + y_{t-2}$$
$$y''_t = [y_t, y_{t-1}, y_{t-2}] * K$$
$$\text{where}$$
$$K = [k_0, k_1, k_2] = [1, -2, 1]$$
$$\text{thus}$$
$$y''_t = [y_t \cdot 1 + y_{t-1} \cdot (-2) + y_{t-2} \cdot 1]$$

Transformation of the Data

We group deforested pixels according to a grid cell where each one represents a region of 30x30 km.

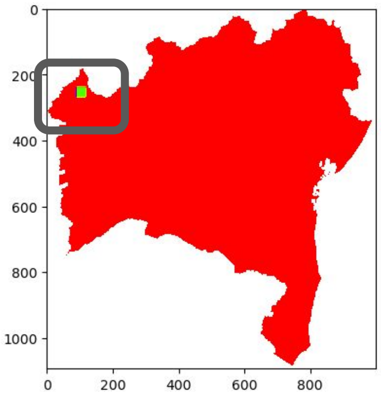


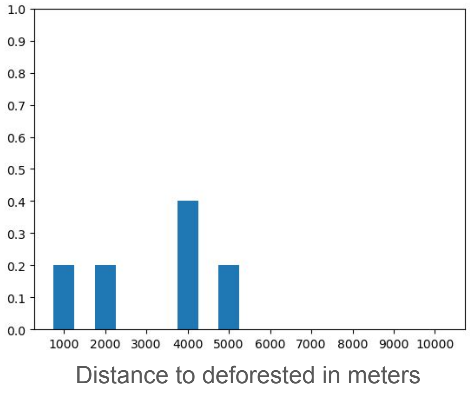


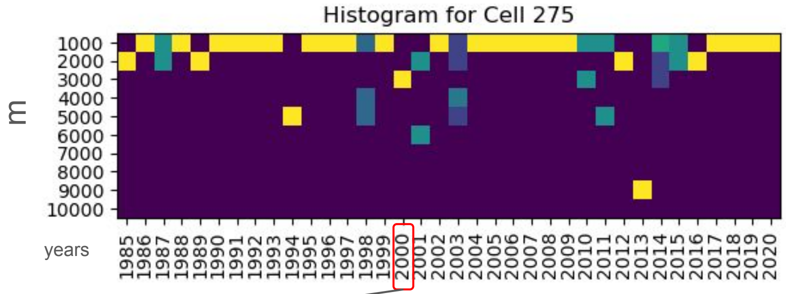
Spatial non-stationarity

Spatial non-stationarity is a condition in which the relations between some sets of spatial variables changes over the course of the time.

Here we show an example describing the spatial allocation shifts with histograms of deforestation in function of the distance to non forest for a region.

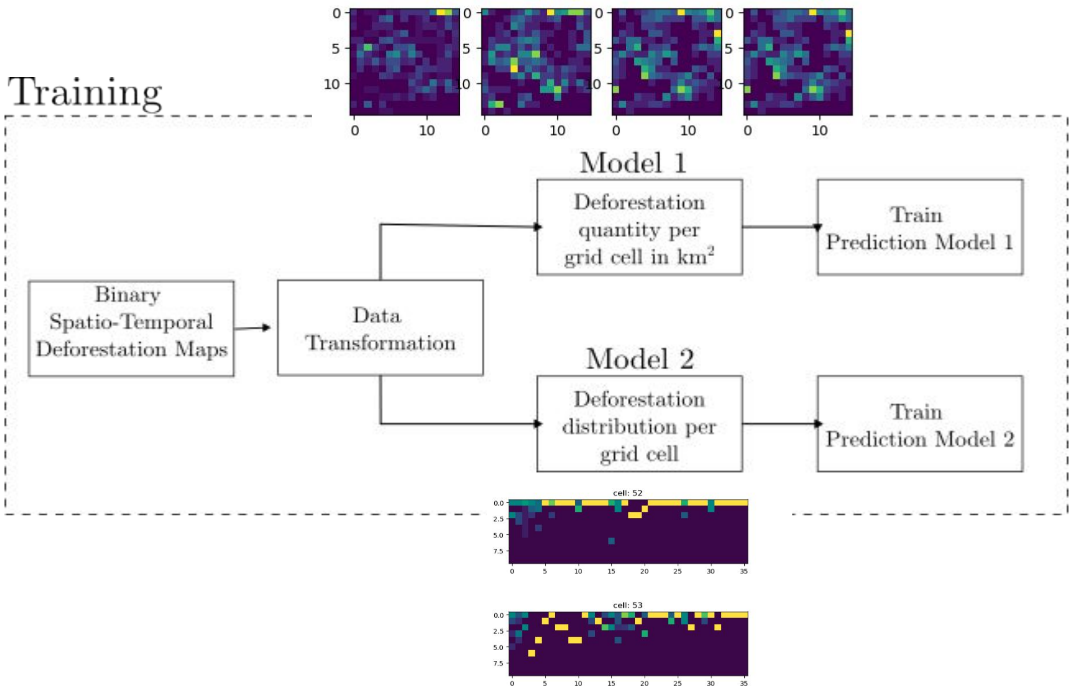




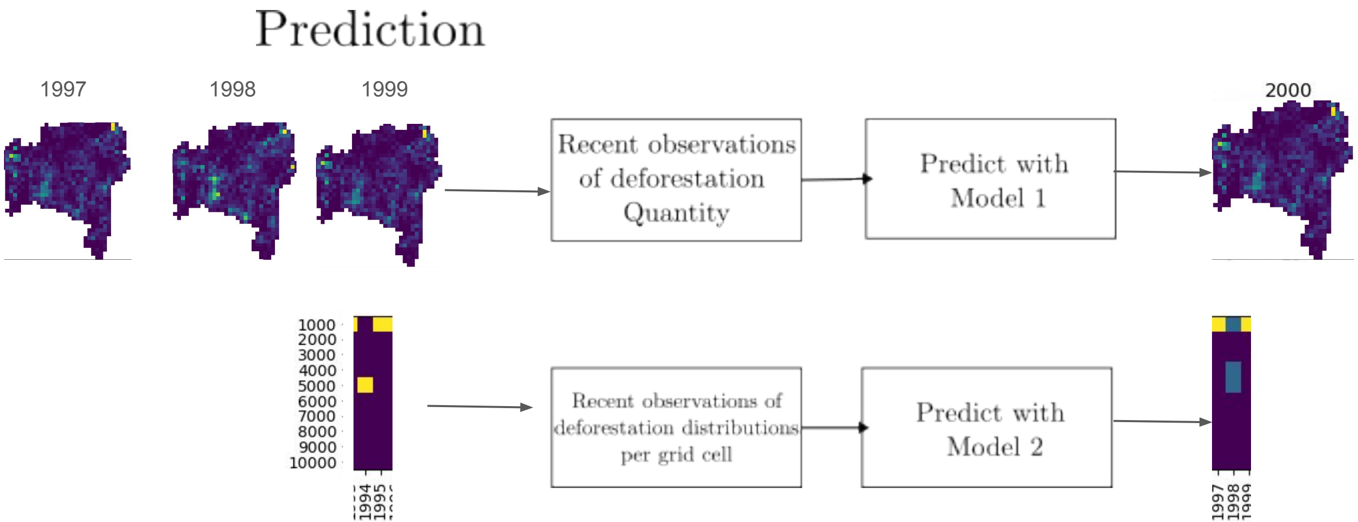


Prediction Model

The model is oriented to predict deforestation increments



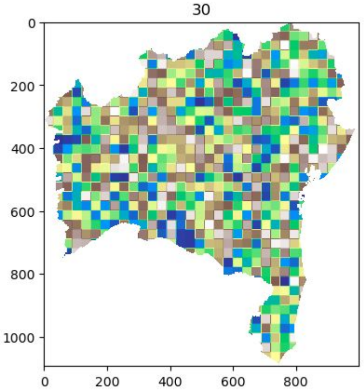
Prediction Model



Experiment 1 Predicting quantity of deforestation increments

We train three CNN models that associate matrices of size 15x15 grid cells of deforestation quantity maps associated with its future values at central location. Each cell of the grid contains the sum of deforested pixels.

Each grid cell represents a squared region of 30kmx30km



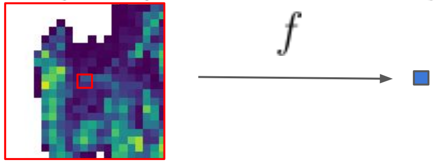
Matrix of deforestation Increments

Future values at central location

$$\min_{\theta} \sum_{y \in Y_{\mathrm{tr}}} \sum_{j \in J} \sum_{i \in I} (f(D^{y}_{(i,j)}, \theta) - d^{y+1}_{i,j})^2$$

$D^{2006}_{(20,10)}$

$d^{2007}_{20,10}$



Experiment 1 Predicting quantity of deforestation increments

- Model 1: A model trained to predict with **one year** of deforestation map.
- Model 2: A Model trained to predict with **two recent** deforestation maps
- Model 3: A Model trained to predict with **three recent** deforestation maps
- A naive model is used as basis for comparison purposes, it produces a prediction completely based in the current observation $d_{i,j}^{y} = \hat{d}_{i,j}^{y+1}$.
- The three models are convolutional neural networks with **2 convolutional layers**, connected to **4 dense feed forward layers**.

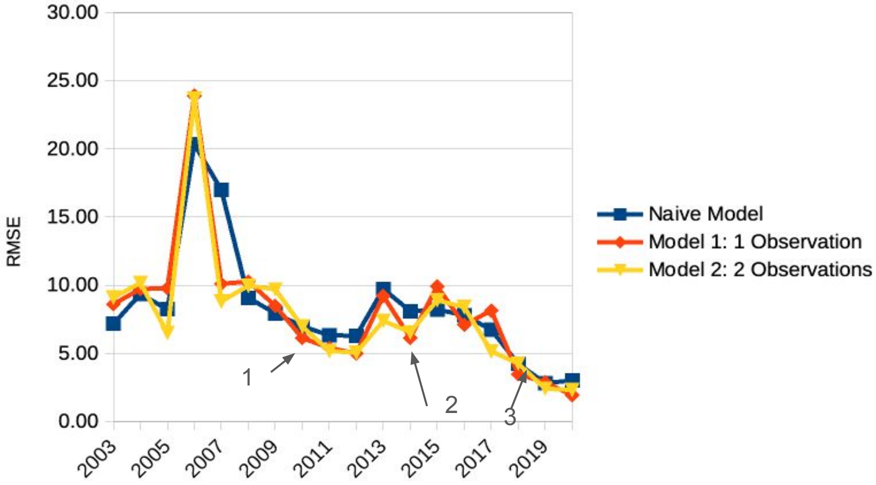
```
x = Conv2D(15, kernel_size=(Psize, Psize), activation='relu')(i)
x = BatchNormalization()(x)
x = MaxPooling2D(pool_size=(3, 3))(x)
x = Conv2D(5, kernel_size=(2, 2), activation='relu')(x)
x = BatchNormalization()(x)
x = Flatten()(x)
x = Dense(40, activation='relu')(x)
x = Dense(30, activation='relu')(x)
x = Dense(20, activation='relu')(x)
x = Dense(1, activation='linear')(x)
```

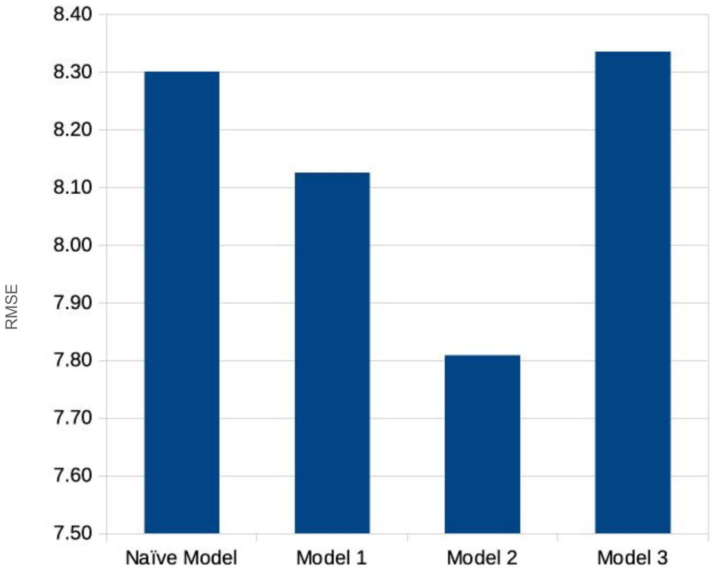
We start predicting 2003 with a training set constructed with information defined by Y_{tr}=[2001-6, 2002), Y_{va}=[2002]

We repeat the training process thought predict 2020.

$$\min_{\theta} \sum_{y \in Y_{\text{tr}}} \sum_{j \in J} \sum_{i \in I} (f(D_{(i,j)}^{y}, \theta) - d_{i,j}^{y+1})^2$$

Results





The proposed models outperform to naive model specifically from 2010 to 2014, and 2018-2020

Experiment for predicting local distribution of the Deforestation.

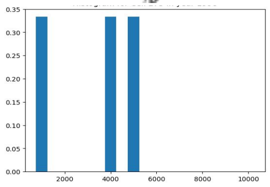
- We train one Feed Forward Neural Network model that relate observed histograms of deforestation to its subsequent histogram in time.
- Each histogram has ten bins. The bin interval represents distance to deforested forest in meters. Each bin contains the proportion of observed deforestation at that distance.
-

```
i = Input(shape=(10,2))
x = Flatten()(i)
x = Dense(40, activation='leaky_relu')(x)
x = Dense(20, activation='leaky_relu')(x)
x = Dense(20, activation='leaky_relu')(x)
x = Dense((10,), activation='sigmoid')(x)
```

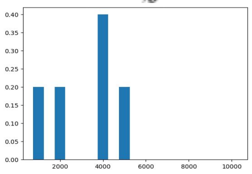
$$\hat{H}^{y+1}_{i,j} = f([H^{y-1}_{(i,j)}, H^{y}_{(i,j)}], \theta)$$

$$\min_{\theta} \sum_{y \in Y_{tr}} \sum_{j \in J} \sum_{i \in I} ((\hat{H}^{y+1}_{i,j})^T H^{y+1}_{i,j})$$

$H^{y-1}_{i,j}$

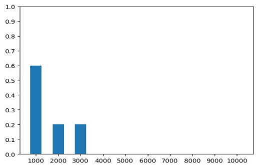


$H^{y}_{i,j}$



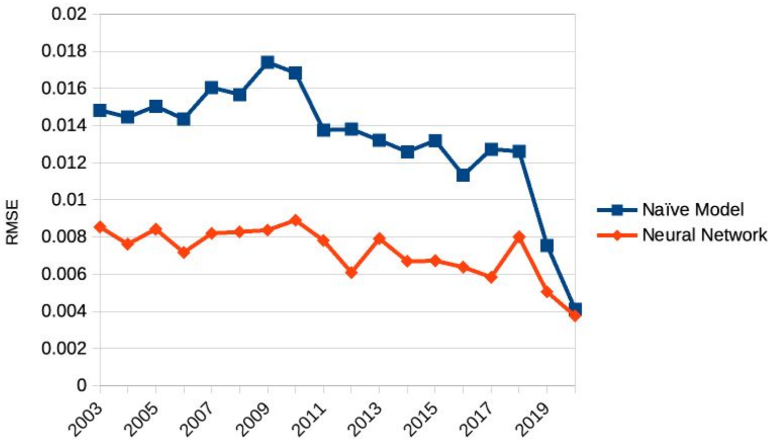
$\xrightarrow{f}$

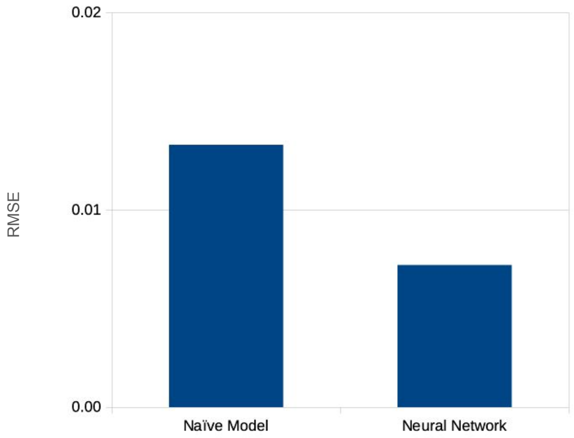
$\hat{H}^{y+1}_{i,j}$



Experiment for predicting local distribution of the Deforestation.

- Each histogram has ten bins. The interval of each bin represents intervals of distance to deforested forest in meters. Each bin contains the proportion of observed deforestation at that distance.





The performance of the neural network is better than naïve model. The quantity of deforestation is predicted with higher accuracy and lower computational cost than its spatial distribution.

Conclusions

Predicting quantity of deforestation

- We observe a slight improvement with respect to the naive model, for model 1 and 2. Incrementing the number of previous observations, leads to performance degradation, possibly due an overfitting effect.
- This fact suggests that the image time series presents a pattern that is difficult to identify due the noise.
- We suggest to measure the performance comparing among models to identify those ones that produce the minimum error prediction at the maximum resolution possible.

Predicting the local distribution of the deforestation

- The neural network found patterns useful to improve the baseline naïve model.
- More spatial variables should be considered as the slope and sea level to have a more detailed description of the deforestation distribution.

Future Work

- To integrate in a single method both: the quantity prediction and the spatial-distribution prediction models.
- To experiment with recurrent and adaptive models such as Long Short Term Memory or Transformers for predicting the deforestation quantity, they are able to retain of forget information automatically according to the recent observations.
- To generate complex histograms that consider more than one variable to increment the allocation accuracy of the deforestation.

References

[1] Fernández Montes de Oca, A.; Gallardo, A.; Ghilardi, A.; Kauffer, E.; Solórzano, J.; Sánchez-Cordero, V. An integrated framework for harmonizing definitions of deforestation. Environmental Science & Policy 2021, 115, 71-78, doi:10.1016/j.envsci.2020.10.007.

[2] Gatti, L.V.; Basso, L.S.; Miller, J.B.; Gloor, M.; Gatti Domingues, L.; Cassol, H.L.G.; Tejada, G.; Aragão, L.E.O.C.; Nobre, C.; Peters, W.; et al. Amazonia as a carbon source linked to deforestation and climate change. Nature 2021, 595, 388-393, doi:10.1038/s41586-021-03629-6.

[3] Garg, P.K. Remote Sensing Theory and Applications; Mercury Learning and Information: Boston, MA, USA, 2024; pp. 326–332. ISBN: 978-1-68392-748-8.

[4] Avila, A. A. O.; Galeana-Pizaña, J.M.; Núñez, J. M. Construction of a Prospective Scenario of Land Use and Cover Change for the Usumacinta River Basin, Indispensable Element for Regional Planning. In Proceedings of the Recent Developments in Geospatial Information Sciences, Cham, 2024//, 2024; pp. 41-52.

[5] Almeida, C.A.d.; Coutinho, A.C.; Esquerdo, J.C.D.M.; Adami, M.; Venturieri, A.; Diniz, C.G.; Dessay, N.; Durieux, L.; Gomes, A.R. High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica 2016, 46.

[6] Vaca, R.A.; Golicher, D.J.; Rodiles-Hernández, R.; Castillo-Santiago, M.Á.; Bejarano, M.; Navarrete-Gutiérrez, D.A. Drivers of deforestation in the basin of the Usumacinta River: Inference on process from pattern analysis using generalised additive models. PLOS ONE 2019, 14, e0222908, doi:10.1371/journal.pone.0222908.

CentroGeo
19°17'30"N 99°13'17"O 2489m

CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES CIENCIAS Y TECNOLOGÍAS

#GLPOSM5
NOVEMBER 2024

Thanks for your attention

Questions

rlopez@centrogeo.edu.mx




Appendix

Meta parameters for Quantity prediction

```
callback = EarlyStopping(monitor='loss', patience = 2, restore_best_weights = True) # se carguen los pesos con la
MDL = model.fit(xtr, ytrb, validation_data = (xval, yvab) , epochs = 20, callbacks = [callback], batch_size = 5)
```

Meta parameters for histogram prediction

```
callback = EarlyStopping(monitor='loss', patience = 1, restore_best_weights = True)
MDL = model.fit(Xtr, Ytr, validation_data = (Xva, Yva) , epochs = 2000, callbacks = [callback], batch_size = 100)
```

1:1 Prediction examples observing the last 2 maps.