

**UNIVERSIDAD AUTÓNOMA DE MADRID**

**ESCUELA POLITÉCNICA SUPERIOR**



**TRABAJO FIN DE GRADO**

**Teledetección aplicada a la deforestación de la Amazonia e impactos del COVID-19 sobre la contaminación en Europa mediante Google Earth Engine.**

**Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación**

**Autor: Martínez Torío, Gonzalo**

**Tutor: Domínguez Carreta, David Renato**

**Junio 2020**



**Teledetección aplicada a la deforestación de la Amazonia e impactos del COVID-19 sobre la contaminación en Europa mediante Google Earth Engine.**



**Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación**

**Autor: Martínez Torío, Gonzalo**

**Tutor: Domínguez Carreta, David Renato**

**Todos los derechos reservados.**

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra sin contar con la autorización de los titulares de la propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (arts. 270 y sgts. del Código Penal).

**DERECHOS RESERVADOS**

© 20 de junio de 2020 por UNIVERSIDAD AUTÓNOMA DE MADRID

Francisco Tomás y Valiente, no 1

Madrid, 28049 Spain

**Gonzalo Martínez Torío**

**Teledetección aplicada a la deforestación de la Amazonia e impactos del COVID-19 sobre la contaminación en Europa mediante Google Earth Engine.**

**Gonzalo Martínez Torío**

IMPRESO EN ESPAÑA – PRINTED IN SPAIN

*A mis padres, Francisco y Elisa*

*“La suerte sonríe a los valientes”*

*Virgilio (70 a. C.-19 a. C.)*

## Agradecimientos

En primer lugar, me gustaría agradecer a mi tutor, David, por haberme propuesto un estudio tan interesante y actual, así como haberme guiado en los momentos que lo requería. Así mismo, gracias a la Escuela Politécnica Superior por haberme ofrecido unos estudios de tan alto nivel, con unos profesores, instalaciones y personal difíciles de igualar.

Gracias a mis padres, Elisa y Francisco, por haberme apoyado en todo momento y haberme dado una educación y valores ejemplares, nunca tendré ni las palabras ni el tiempo para daros las gracias por todo lo que me habéis dado. Gracias a Jaime, mi hermano, por estar a mi lado en todo momento, espero que disfrutes de tu etapa en la universidad tanto como lo he disfrutado yo. Gracias a mis amigos del colegio, con quienes he pasado tantos buenos momentos y con los que siempre podré contar.

Gracias a la carrera que he estudiado, gracias a ella he conocido a gente maravillosa a la que hoy puedo llamar amigos y amigas. Gracias por haberme llevado de Erasmus y haber podido vivir una experiencia así, en la que he tenido la oportunidad de estudiar en otra universidad puntera y en la que he conocido a gente que siempre llevaré en el corazón. Pero, sobre todo, gracias a Ana, por enseñarme a ser mejor persona cada día y por estar junto a mí de forma incondicional, nunca sabrás lo mucho que me aportas.

*Gonzalo Martínez Torío*

*Junio 2020*

## Resumen

Cada vez son más las noticias de diferentes catástrofes naturales que producen un gran impacto en los diferentes bosques del planeta. Muchas de estas catástrofes naturales vienen, en la mayoría de las ocasiones, propiciadas por parte del ser humano. Bien sea por la tala de árboles abusiva, incendios intencionados o, por una de las mayores industrias del planeta, la ganadera. Así mismo, la emisión de gases contaminantes que estos desastres naturales producen supone un grave riesgo para la sociedad.

Este Trabajo Fin de Grado, tiene como objetivo centrar los problemas vistos anteriormente en la mayor región tropical del planeta, la Amazonia brasileña, y en concreto, en las 4 regiones más afectadas por la deforestación, Pará, Mato Grosso, Amazonas y Rondonia. Por otra parte, y en relación con la pandemia provocada por el virus del COVID-19, se ha hecho un estudio sobre la disminución de los gases contaminantes en la atmósfera, centrándose en esta ocasión en los países europeos más afectados por dicha pandemia, España e Italia. Con el gran abanico de herramientas que existen hoy en día para abordar ambos temas, se ha optado finalmente por la plataforma proporcionada por Google, llamada Google Earth Engine.

En los siguientes capítulos se hará un estudio de la herramienta, mostrando las ventajas que proporciona a la hora de abordar un tema de esta dimensión, en el que se requiere una gran cantidad de información en imágenes y capacidad de procesamiento. Así mismo, se hará uso tecnologías complementarias tales como Google Collaboratory como entorno de Jupyter Notebook, Google Cloud para el almacenamiento de la información generada y Tensorflow como biblioteca de código abierto para aprendizaje automático.

## Palabras Clave

Deforestación, Amazonía, Google, Google Earth Engine, Google Collaboratory, Google Cloud, Tensorflow, teledetección, aprendizaje automático, procesamiento de imágenes, COVID-19, contaminación, gases de efecto invernadero.

## Abstract

Nowadays, there is more and more news of different natural disasters that have a huge impact on the different forests of the earth. Many of these natural disasters are, in many cases, caused by human beings. For example, abusive logging, intentional fires, or one of the largest industries on the planet, livestock. Likewise, the emission of polluting gases that these natural disasters produce represents a serious risk for the population.

Many of these natural disasters are, in many cases, caused by human action. Whether it is because of abusive logging, deliberately set fires or because of one of the largest industries on the planet, livestock. On many occasions, it is the last one that produces these deforestations and fires, with the aim of gaining ground in the forests to make space for the production of meat, as a product of the growing demand for these products by society. Likewise, the emission of polluting gases produced by these processes represents a serious risk for society.

This End-of-Degree Project aims to focus the problems seen previously in the world's largest tropical region, the Brazilian Amazon, and specifically in the 4 regions most affected by deforestation: Pará, Mato Grosso, Amazonas and Rondonia. On the other hand, and in relation to the COVID-19 virus pandemic, a study has been carried out regarding the reduction of polluting gases in the atmosphere, focusing on the European countries most affected by this pandemic, Spain and Italy. With the wide range of tools that exist at the present time to tackle these two issues, we have finally decided to use the platform provided by Google, called Google Earth Engine.

In the following chapters, we will study this tool, showing the advantages it provides when approaching a topic of this dimension, which requires a large amount of information in terms of images and processing capacity. Likewise, complementary technologies will be used such as Google Collaboratory as a Jupyter Notebook environment, Google Cloud for storing the information generated and Tensorflow as an open source library for automatic learning.

## Keywords

Deforestation, Amazon, Google, Google Earth Engine, Google Collaboratory, Google Cloud, Tensorflow, teledetection, machine learning, image processing, COVID-19, pollution, greenhouse gases,





# ÍNDICE DE CONTENIDOS

1.	Introducción .....	1
1.1.	Motivación.....	1
1.2.	Objetivos .....	1
1.3.	Organización de la memoria.....	2
2.	Estado del arte.....	4
2.1.	Deforestación en el Amazonas .....	4
2.2.	Contaminación producida por gases de efecto invernadero.....	7
2.3.	Google Earth Engine.....	9
2.4.	Machine Learning .....	11
3.	Diseño.....	13
3.1.	Introducción.....	13
3.2.	Introducción a los conjuntos de datos utilizados para el estudio sobre la deforestación.....	13
3.3.	Introducción a los conjuntos de datos utilizados para el estudio sobre la contaminación.....	15
3.4.	Utilización de la red neuronal.....	16
4.	Desarrollo .....	18
4.1.	Estudio sobre la deforestación.....	18
4.2.	Estudio sobre la contaminación.....	22
5.	Integración, pruebas y resultados.....	24
5.1.	Estudio sobre la deforestación.....	24
5.2.	Estudio sobre la contaminación.....	37
6.	Conclusiones y Trabajo a futuro .....	39
6.1.	Conclusiones.....	39
6.2.	Trabajo a futuro .....	39
	Bibliografía.....	41

## ÍNDICE DE FIGURAS

<b>Figura 1.1:</b> Diagrama de Gantt del desarrollo del Trabajo Fin de Grados.....	2
<b>Figura 2.1:</b> Ejemplo deforestación en espina de pescado en el parque nacional dos campos Amazónicos (Amazonas) .....	4
<b>Figura 2.2:</b> Gráfico ilustrativo del proceso de deforestación .....	5
<b>Figura 2.3:</b> Ratios de deforestación de la Amazonia Legal brasileña por estado [22] .....	6
<b>Figura 2.4:</b> Deforestación acumulada por los estados brasileños [22].....	6
<b>Figura 2.5:</b> Principales gases contaminantes clasificados por grupo, estado físico y sus respectivas fuentes. [6].....	7
<b>Figura 2.6:</b> Pantalla API Code editor GEE. [33] .....	9
<b>Figura 2.7:</b> Estructura simplificada de la plataforma GEE [10].....	10
<b>Figura 2.8:</b> Principio de funcionamiento de una red neuronal. Imagen extraída de [34] ...	12
<b>Figura 2.9:</b> Ejemplo de una red neuronal simple. Imagen extraída de [34] .....	12
<b>Figura 3.1:</b> Bandas High-Resolution Global Maps of 21st-Century Forest Cover Change data. ....	14
<b>Figura 3.2:</b> Bandas USGS Landsat 8 Collection 1 Tier 1 Raw Scenes.....	15
<b>Figura 3.3:</b> Bandas Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide .....	16
<b>Figura 3.4:</b> Bandas Sentinel-5P NRTI NO2: Near Real-Time Nitrogen Dioxide.....	16
<b>Figura 3.5:</b> Diagrama de flujo del trabajo entre las plataformas. ....	17
<b>Figura 4.1:</b> Puntos seleccionados para la clasificación en la región de Pará .....	21
<b>Figura 4.2:</b> Puntos seleccionados para la clasificación en la región de Amazonas.....	21
<b>Figura 4.3:</b> Archivos de training y testing almacenados en plataforma Google Cloud. ....	22
<b>Figura 5.1:</b> Resultado deforestación en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	24
<b>Figura 5.2:</b> Gráfico deforestación por años en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018) .....	25
<b>Figura 5.3:</b> Gráfico deforestación por años en región Pará, datos del INPE [22] .....	25
<b>Figura 5.4:</b> Resultado deforestación en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018) .....	26
<b>Figura 5.5:</b> Gráfico deforestación por años en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	27

<b>Figura 5.6:</b> Gráfico deforestación por años en región Mato Grosso, datos del INPE [22].	27
<b>Figura 5.7:</b> Resultado deforestación en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	28
<b>Figura 5.8:</b> Gráfico deforestación por años en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018) .....	29
<b>Figura 5.9:</b> Gráfico deforestación por años en región Rondonia, datos del INPE [22] .....	29
<b>Figura 5.10:</b> Resultado deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	30
<b>Figura 5.11:</b> Gráfico deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	31
<b>Figura 5.12:</b> Gráfico deforestación por años en región Amazonas, datos del INPE [22]... ..	31
<b>Figura 5.13:</b> Resultado obtenido para la base de datos de FORMA en la región de Pará. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Pará.....	33
<b>Figura 5.14:</b> Resultado obtenido para la base de datos de FORMA en la región de Mato Grosso. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Mato Grosso.....	33
<b>Figura 5.15:</b> Resultado obtenido para la base de datos de FORMA en la región de Rondonia. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Rondonia.....	34
<b>Figura 5.16:</b> Resultado obtenido para la base de datos de FORMA en la región de Amazonas. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Amazonas .....	34
<b>Figura 5.17:</b> Resultados obtenidos para la región de Pará .....	35
<b>Figura 5.18:</b> Resultados obtenidos para la región de Amazonas.....	36
<b>Figura 5.19:</b> Resultados obtenidos para el gas NO <sub>2</sub> en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19.....	37
<b>Figura 5.20:</b> Resultados obtenidos para el gas NO <sub>2</sub> en Italia antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19 .....	37
<b>Figura 5.21:</b> Resultados obtenidos para el gas CO en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19 .....	38
<b>Figura 5.22:</b> Resultados obtenidos para el gas CO en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19 .....	38



## ÍNDICE DE TABLAS

Tabla 5.1: Deforestación por años en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018) .....	26
Tabla 5.2: Deforestación por años en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	28
Tabla 5.3: Deforestación por años en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	30
Tabla 5.4: Deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018).....	32

# 1. Introducción

## 1.1. Motivación

En un mundo en el que el consumismo está a la orden del día y la industria tiene un gran peso en la sociedad, surgen dos problemas importantes relacionados con el deterioro del planeta en el que vivimos. Estos problemas son la deforestación de los grandes bosques, siendo los uno de los mayores productores de oxígeno y “filtros” de aire del planeta, y la contaminación producida por las grandes industrias y vehículos entre otros.

A partir de estos problemas, y con la necesidad de denunciar los mismos, este estudio pretende, a través del tratamiento de imagen, exponer la criticidad de estas dos problemáticas. Así mismo, la situación en la que se ha llevado a cabo este Trabajo de Fin de Grado, en medio de la pandemia provocada por el Coronavirus (COVID-19), ha permitido demostrar los notorios cambios que ha experimentado la naturaleza tras la época de confinamiento que se ha vivido a lo largo de los primeros meses del año, observando notables reducciones en los niveles de contaminación y otros muchos beneficios para la naturaleza.

Por otra parte, cabe destacar la motivación que este Trabajo de Fin de Grado supone a la hora de trabajar con una plataforma puntera, Google Earth Engine, teniendo que aprender desde cero, en el caso que aplica, algunas de las posibilidades que ésta ofrece.

## 1.2. Objetivos

El objetivo principal de este Trabajo de Fin de Grado es el análisis, a través de la plataforma de Google Earth Engine, de cuatro de las regiones más afectadas por la deforestación de la Amazonia brasileña: Pará, Mato Grosso, Rondonia y Amazonas. Así mismo, se tiene como objetivo el estudio del impacto que ha provocado la pandemia de COVID-19 en los países europeos más afectados por dicha pandemia: España e Italia.

Por otra parte, este Trabajo de Fin de Grado tiene el objetivo de servir de aprendizaje en la utilización de la plataforma utilizada en este estudio, Google Earth Engine, analizando las posibilidades que ofrece para llevar a cabo estudios de este tipo, en el que la capacidad de computación y disponibilidad de información son factores determinantes.

## 1.3. Organización de la memoria

### 1.3.1. Fases de realización del TFG

Tareas que realizar:

1. Documentación del estudio: El proceso de documentación ha sido continuo, desde el inicio de éste hasta el final, tratando de detallar lo más posible los procesos a seguir para el correcto seguimiento del Trabajo de Fin de Grado
2. Estudio de las causas y consecuencias de la deforestación y contaminación: Estudio previo a la realización del Trabajo de fin de Grado, para poder tener un contexto lo suficientemente sólido para poder entender ambas problemáticas.
3. Trabajo y entendimiento de la plataforma Google Earth Engine: Este punto incluye el registro en la plataforma y el estudio de las capacidades que ofrece la plataforma, así como el entendimiento del funcionamiento de esta.
4. Selección de las bases de datos a usar: Tiempo desempeñado en la selección de las bases de datos que más se ajustan al estudio, esto incluye las pruebas que se hicieron sobre bases de datos que posteriormente serían descartadas, alcanzando finalmente las bases de datos utilizadas para el presente estudio.
5. Implementación de la red neuronal: Trabajo de entendimiento e implantación del apartado referente a la implementación de la red neuronal, así como la obtención de resultados.
6. Pruebas y resultados: Todas las tareas relacionadas con la obtención de resultados que posteriormente se mostrarían en la memoria.

Diagrama de Gantt											
Num	Tarea	Inicio	Final	noviembre-19	diciembre-19	enero-20	febrero-20	marzo-20	abril-20	mayo-20	junio-20
1	Documentación del estudio	01-11-2019	20-06-2020								
2	Estudio de las causas y consecuencias de la deforestación y contaminación	01-11-2019	20-12-2019								
3	Trabajo y entendimiento de la plataforma Google Earth Engine	20-12-2019	20-05-2020								
4	Selección de las bases de datos a usar	20-01-2020	20-03-2020								
5	Implementación de la red neuronal	01-03-2020	01-05-2020								
6	Pruebas y resultados	01-04-2020	20-06-2020								

Figura 1.1: Diagrama de Gantt del desarrollo del Trabajo Fin de Grados

### 1.3.2. Estructura de la memoria

La memoria se compone de los siguientes capítulos:

1. **Estado del arte**: Este capítulo recoge una descripción de la situación actual que se tiene de los temas que se tratarán posteriormente en el estudio.
2. **Diseño**: Este capítulo expone las características y pasos que se darán para alcanzar los resultados



3. **Desarrollo:** Este capítulo recoge los anteriores pasos descritos detalladamente, con las diferentes consideraciones que se han tenido para poder obtener los resultados esperados.
4. **Integración, Pruebas y Resultados:** Este capítulo recoge los resultados obtenidos a partir de los desarrollos aplicados en el capítulo anterior.
5. **Conclusiones y trabajo a futuro:** El capítulo expone las conclusiones obtenidas a partir de los resultados obtenidos del estudio, así como propuestas de mejora y de puntos de partida a partir de este estudio.

## 2. Estado del arte

En este capítulo se dará una visión de los diferentes puntos sobre los que se trabajará posteriormente, con el objetivo de dar al lector un contexto sobre dichos temas y que la comprensión de estos sea plena.

### 2.1. Deforestación en el Amazonas

Este apartado recoge las principales causas y consecuencias derivadas de la deforestación en la Amazonia. Se parte de la propia definición del concepto de deforestación y que causas tiene esta sobre los entornos en los que se da, se establecerá un marco histórico de la deforestación en la amazonia y se irá descendiendo y concretando en diferentes zonas de la Amazonia.

Se entiende como deforestación el proceso por el que se produce la pérdida de superficie forestal, este proceso normalmente viene provocado por la acción del ser humano. Existen multitud de formas de provocar dicha deforestación, como por ejemplo [1] la tala abusiva y la posterior quema del terreno provocadas por la industria maderera, así mismo, esta tala abusiva muchas veces tiene como objetivo ganar suelo para dejar espacio para la industria ganadera o la agricultura. El problema de la deforestación también viene propiciado por la ineficiencia de su posterior reforestación, ya que no se tiene en cuenta el daño previo que ha sufrido el hábitat de la flora y la fauna, provocando de esta manera la pérdida de biodiversidad y generando aridez en el terreno.

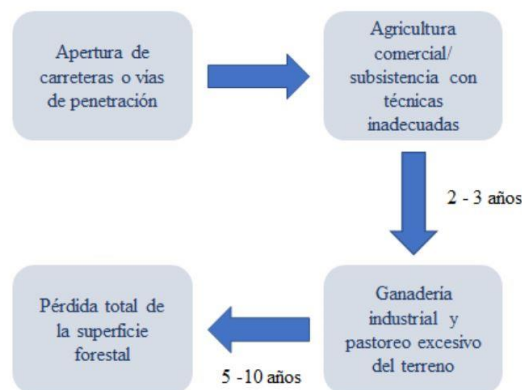
A menudo, la deforestación en estas zonas sigue un claro proceso. Este proceso de deforestación [2] comienza con la construcción de carreteras que atraviesan los bosques y abriéndolos a la tala y la minería, [3] solo la Amazonía brasilera y sectores de la ecuatoriana exhiben el típico patrón de espina de pescado asociado con la construcción de dichas carreteras.



**Figura 2.1:** Ejemplo deforestación en espina de pescado en el parque nacional dos campos Amazónicos (Amazonas)

Una vez los alrededores de la carretera han sido “limpiados” de todo tipo de fauna y flora, los agricultores comerciales o de subsistencia se establecen y comienzan a producir diferentes cultivos. Los suelos de los bosques no están preparados para poder mantener dichos cultivos durante un tiempo prolongado, por lo que al cabo de 2 o 3 años los suelos dejan de ser productivos y los agricultores dejan de plantar en ellos. Es en este momento, en el que los agricultores han agotado el terreno, cuando los ganaderos los toman, ya que adquirirlos supone un coste bajo o en ocasiones nulo (ya que las tierras son abandonadas).

Estos terrenos son idóneos para las industrias ganaderas, ya que poseen hierba con la que alimentar a las cabezas de ganado durante un tiempo. Al cabo de 5 o 10 años, el exceso de pastoreo y la pérdida de nutrientes del terreno, hacen que lo que era un bosque lluvioso y un entorno rico en biodiversidad pase a ser un terreno estéril y sin vida alguna.



**Figura 2.2:** Gráfico ilustrativo del proceso de deforestación

La Amazonia brasileña [4] fue creada como un concepto político destinada a la planificación y desarrollo regional mediante la Ley 1806 del 06/01/1953, modificada por la ley posteriormente 5173 de 27/10/1966 y la ley complementaria 31 11/10/1977. La Amazonia brasileña ocupa un área correspondiente al 59% del territorio brasileño y abarca todos los ocho estados (Acre, Amapá, Amazonas, Mato Grosso, Pará, Rondônia, Roraima y Tocantins) y parte del Estado de Maranhão, ascendiendo de esta manera a 5,0 millones de kilómetros cuadrados (el equivalente a unos 771 campos de fútbol).

A través de la plataforma Terrabrasilis (plataforma desarrollada por el Instituto Nacional de Pesquisas Espaciais (INPE) para centralizar y distribuir los datos generados por el Programa de Vigilancia de la Amazonia brasileña Bosque Satélite (PRODES)) se han obtenido los datos mostrados en las siguientes tablas:

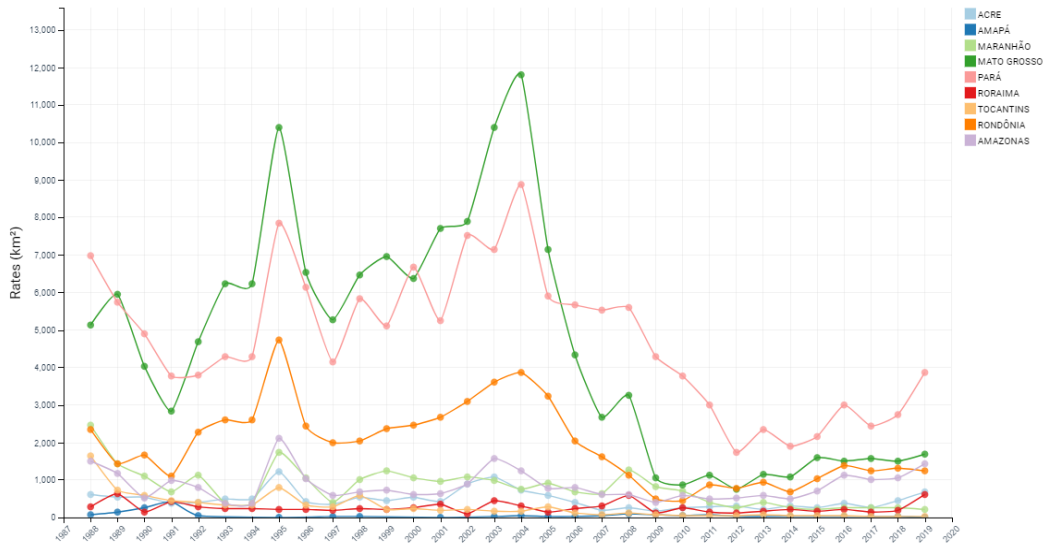


Figura 2.3: Ratios de deforestación de la Amazonia Legal brasileña por estado [22]

En el gráfico anterior se muestran las ratios de deforestación por año en los estados de Brasil incluidos en la Amazonia. El gráfico sitúa a los estados de Pará, Mato Grosso, Amazonas y Rondonia como los más afectados, en ese orden, por la deforestación.

Si se cambia la forma de ver los datos, y se busca una visión de la deforestación acumulada entre los años 1987 y 2019 por los estados brasileños pertenecientes a la Amazonia, se obtiene la siguiente imagen, dando a ver, que los estados más afectados se mantienen respecto a la anterior figura.

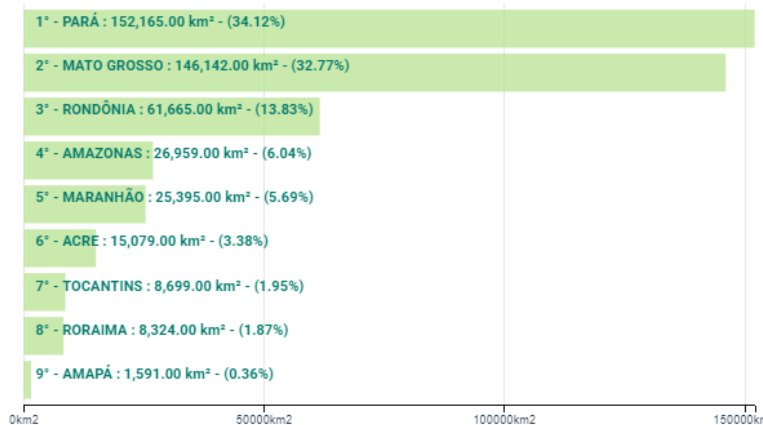


Figura 2.4: Deforestación acumulada por los estados brasileños [22]

## 2.2. Contaminación producida por gases de efecto invernadero

Este capítulo seguirá la misma dinámica que el anterior, se partirá definiendo y estableciendo que se denomina “contaminación” y más tarde se procederá a tratar dos de los gases contaminantes que contribuyen a producir el conocido efecto invernadero efecto invernadero (NO<sub>2</sub> y CO).

Por otra parte, cabe destacar, que este capítulo fue incluido de forma posterior al establecimiento del índice, debido a la pandemia del conocido COVID-19, y la posibilidad de ampliar el estudio a un tema de actualidad como este. Dicha pandemia será brevemente explicada en el punto 1.3 para poder llevar a cabo posteriormente un breve estudio de la relación existente entre la contaminación y la pandemia, en especial en los países de Italia y España, coincidiendo estos con las regiones más afectadas de Europa por dicha pandemia.

Atendiendo a la propia definición de contaminación; [5] la contaminación es el proceso por el cual se altera nocivamente la pureza o las condiciones normales de una cosa o un medio por agentes químicos o físicos. En este estudio se hará hincapié en la contaminación atmosférica, siendo ésta [6] normalmente medida en la atmósfera urbana, proveniente de fuentes móviles (tráfico rodado) y de fuentes fijas de combustión (industrias, usos residenciales -climatización- y procesos de eliminación de residuos). Existen dos grupos de contaminantes; los primarios, procedentes de las fuentes de emisión y los contaminantes secundarios, siendo aquellos que se producen como resultado de la transformación química y física de los contaminantes primarios. A continuación, se muestra una tabla en la que se incluyen los principales contaminantes químicos y sus fuentes más importantes.

Contaminante	Formación	Estado físico	Fuentes
Partículas en suspensión (PM): PM <sub>10</sub> , Humos negros.	Primaria y secundaria	Sólido, líquido	Vehículos Procesos industriales Humo del tabaco
Dióxido de azufre (SO <sub>2</sub> )	Primaria	Gas	Procesos industriales Vehículos
Dióxido de Nitrógeno (NO <sub>2</sub> )	Primaria y secundaria	Gas	Vehículos Estufas y cocinas de gas
Monóxido de carbono (CO)	Primaria	Gas	Vehículos Combustiones en interiores Humo de tabaco
Compuestos orgánicos volátiles (COVs)	Primaria, secundaria	Gas	Vehículos, industria, humo del tabaco Combustiones en interiores
Plomo (Pb)	Primaria	Sólido (partículas finas)	Vehículos, industria
Ozono (O <sub>3</sub> )	Secundaria	Gas	Vehículos (secundario a foto-oxidación de NO <sub>x</sub> y COVs)

PM<sub>10</sub>: partículas con un diámetro inferior a 10 µm  
NO<sub>x</sub>: óxidos de nitrógeno

**Figura 2.5:** Principales gases contaminantes clasificados por grupo, estado físico y sus respectivas fuentes. [6]

Como se apuntó en la introducción de este apartado, se tratarán dos de los gases que aparecen en la anterior tabla, el Dióxido de nitrógeno (NO<sub>2</sub>) y el Monóxido de carbono (CO).

El monóxido de carbono [7] es un gas incoloro e inodoro que se forma por la combustión incompleta de material orgánico, en presencia deficitaria de oxígeno. Se considera uno de los mayores contaminantes atmosféricos y de los mayores problemas en América Latina. La principal fuente de este gas son los vehículos que hacen uso de gasolina o diésel; procesos industriales; incendios forestales y urbanos y las incineraciones de materia orgánica.

En Brasil, el monóxido de carbono es el gas contaminante emitido en mayor cantidad a la atmósfera. En Sao Paulo 1,5 millones de toneladas son lanzadas al aire anualmente; de estos el 78 % son producidos por los automotores a gasolina o diésel, 15 % por automotores con combustible de gas, 3 % por motocicletas, 2 % por taxis y 2 % resultan de procesos industriales.

El dióxido de nitrógeno [8] es un compuesto químico gaseoso amarillento, tóxico e irritante. En la naturaleza se produce por los incendios forestales o erupciones volcánicas. También es producido por la descomposición de nitratos orgánicos. La mayor parte del dióxido de nitrógeno se produce por la combustión de los motores diésel.

### **2.2.1. COVID- 19: impactos en la contaminación**

Este capítulo se enfocará en la pandemia del COVID- 19 y la repercusión que ha tenido en el medio ambiente, y en especial, en la contaminación atmosférica. Cabe destacar que el capítulo se añadió posteriormente al establecimiento del índice, esto fue debido a la posibilidad de poder contribuir a la investigación de uno de los múltiples impactos que ha causado, y está causando en el momento de redacción de este documento, la pandemia del COVID- 19.

Este apartado, como los anteriores, se enfocará en ofrecer una visión general de la pandemia, ofreciendo los datos más significativos para su comprensión. En los capítulos posteriores se tratarán los resultados obtenidos a partir del estudio realizado, así como la forma en la que se han obtenido.

La pandemia del COVID- 19 tiene su primer caso de aparición [9] en el mes de diciembre de 2019, en la ciudad de Wuhan, capital de la provincia de Hubei, en la república popular de China. La pandemia se ocasiona por el virus Coronavirus 2 del síndrome respiratorio agudo grave (SARS-CoV-2). Más tarde, el virus se propagó alrededor del mundo, en concreto, en Europa, los dos países más afectados serían España e Italia, acumulando el mayor número de casos por contagio y fallecidos en el continente. El virus hizo que los Gobiernos de los países tuvieran que imponer el confinamiento de la población. Es por ello por lo que, en los capítulos posteriores, se estudiará los efectos que esto ha provocado en los niveles de contaminación registrados en ambos países.

## 2.3. Google Earth Engine

### 2.3.1. Introducción

Este apartado tiene como objetivo exponer la herramienta con la que se ha llevado a cabo el estudio de los anteriores temas expuestos. En primer lugar, se dará ofrecerá una visión de lo que es la plataforma y su funcionamiento para poder entender plenamente el por qué se ha elegido para la realización de este estudio.

En segundo lugar, se mostrarán en profundidad las diferentes partes de las que se ha hecho uso de esta. Hay que destacar las facilidades que ofrece Google para el trabajo con dicha plataforma, así como la documentación que pone a disposición del usuario para poder entender el funcionamiento de esta y poder llevar a cabo los estudios que se requieran.

Google Earth Engine (en adelante, GEE) [10] es una plataforma geomática en cloud que pone a disposición de sus usuarios el acceso a recursos informáticos de alto rendimiento para procesar conjuntos de datos geoespaciales de gran tamaño, no teniendo que limitarse a la capacidad de computación de las máquinas personales. Así mismo, científicos y diferentes organizaciones sin ánimo de lucro utilizan GEE para llevar a cabo estudios de teledetección remota, predecir brotes de enfermedades, gestionar recursos naturales, etc. Un punto importante a tener en cuenta es la diferencia entre GEE y Google Earth, residiendo dicha diferencia en el análisis de datos y el tratamiento de los mismos.

GEE posee un catálogo de datos satelitales multitemporales del orden de tamaño de petabytes. GEE posee imágenes históricas terrestres de más de cuarenta años de antigüedad. Las imágenes que se recopilan diariamente se ponen a disposición de los usuarios para realizar minería de datos a gran escala.

GEE proporciona una API y otras herramientas para permitir el análisis de grandes conjuntos de datos. A continuación, se muestra la interfaz de la API con la que se ha trabajado.

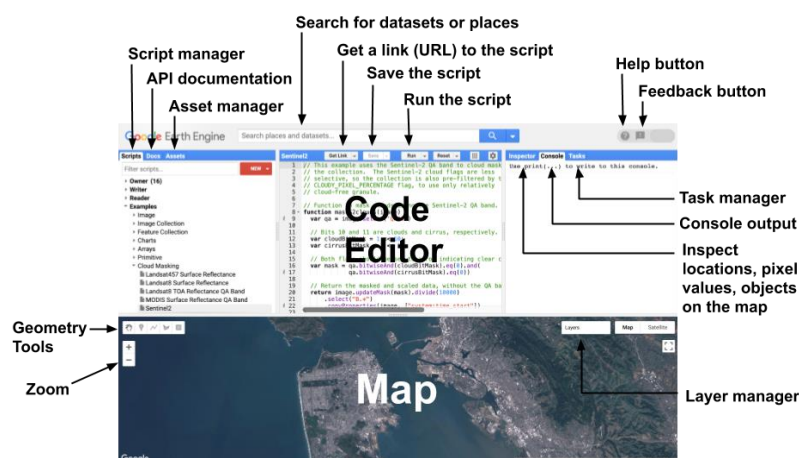


Figura 2.6: Pantalla API Code editor GEE. [33]

### 2.3.2. Arquitectura plataforma Google Earth Engine

GEE está construido [10] [11] sobre una colección de tecnologías habilitadoras que están disponibles dentro del entorno de datos de Google, incluido el sistema de gestión de clúster Borg [12], las bases de datos distribuidas The Bigtable [13] y Spanner [14]; Colossus, el sucesor del sistema de Archivos de Google [15] y el framework Flumejava [16] para la ejecución en paralelo de tareas. Así mismo, GEE también interactúa con Google Fusion Tables [17], siendo esto una base de datos basada en web que admite tablas de datos geométricos (puntos, líneas y polígonos) con atributos.

En la figura 2.7 mostrada a continuación, se ilustra la arquitectura de GEE de manera simplificada. A través del editor de código (mostrado en la foto anterior) o aplicaciones de terceros (ej.: Google Colab [18]) se utilizan las bibliotecas de clientes para enviar consultas a través de una API REST. Dichas consultas, son gestionadas por servidores front end que a su vez envían subconsultas a los compute masters, siendo estos los que manejan la distribución de la computación entre un grupo de servidores de computación. En el caso del sistema de Batch Computation, funciona de forma similar a lo ya explicado, pero utilizando FlumeJava para la gestionar la distribución de la computación.

Ambos servicios de cómputo están respaldados por un servicio de recolección de datos que incluye metadatos de imágenes y proporciona capacidades de filtrado eficiente. Todos los componentes del sistema están gestionados por el software de gestión de clusters de Borg y cada servicio está equilibrado en cuanto a la carga sobre varios agentes. El fallo de cualquier agente individual únicamente provocaría que la persona que ha hecho una consulta tenga que volver a hacerla.

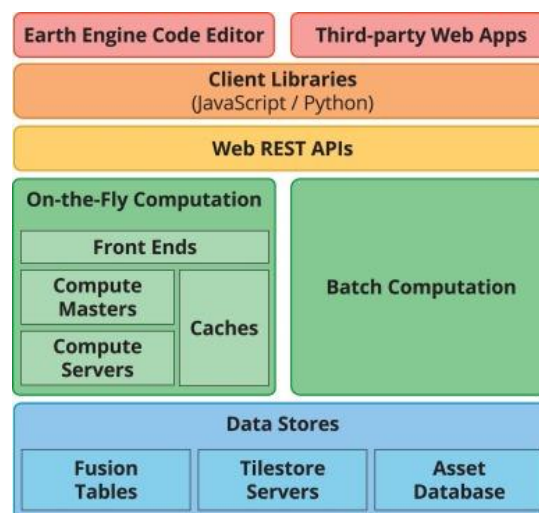


Figura 2.7: Estructura simplificada de la plataforma GEE [10]



### 2.3.3. Catálogo utilizado para el estudio

El catálogo de datos de GEE es una colección compuesta por diferentes conjuntos de datos geoespaciales, teniendo un orden de magnitud de petabytes. En este catálogo se incluye todo el conjunto de imágenes de la tierra proporcionado por los satélites Landsat, así como de los satélites Sentinel 1 y 2. Por otra parte, también incluyen previsiones climáticas, datos sobre la cubierta terrestre y otros conjuntos ambientales.

Un punto importante que destacar, es el ritmo de actualización del catálogo de datos; aproximadamente 6000 escenas son generadas diariamente, con una latencia de unas 24 horas de la adquisición de las imágenes.

### 2.3.4. Datasets utilizados en el estudio

Haciendo uso de los datasets proporcionados por la plataforma bajo el nombre "High-Resolution Global Maps of 21st-Century Forest Cover Change" [19] y "Forest Monitoring for Action" (FORMA) [20] se procederá en los capítulos posteriores a hacer un tratamiento de dichos datasets y poder obtener imágenes sobre la pérdida/ ganancia de cubierta forestal en la Amazonia brasileña. Así mismo, se compararán los resultados arrojados por ambos datasets.

## 2.4. Machine Learning

GEE ofrece la oportunidad de aplicar machine learning a los conjuntos de datos de los que dispone. Para ello hace uso de los paquetes `ee.Classifier`, `ee.Clusterer`, y `ee.Reducer` para poder entrenar e interpretar los resultados arrojados por los modelos con los que se trabajen. GEE propone tanto la opción de aprendizaje supervisado, donde se encuentran los clasificadores CART (Classification and Regression Trees), Random Forest, Naive Bayes y SVM (Support Vector Machine), como la opción de aprendizaje no supervisado.

Así mismo, GEE da la posibilidad de trabajar con la plataforma de Tensorflow para aquellos modelos de machine learning más avanzados.

En este estudio en concreto, se seguirá el modelo propuesto por GEE para la implementación de una red neuronal profunda con una capa oculta, concretando al estudio que aplica.

A continuación, se procede a contextualizar los conceptos de Tensorflow y red neuronal profunda con una capa oculta.

### 2.4.1. Tensorflow

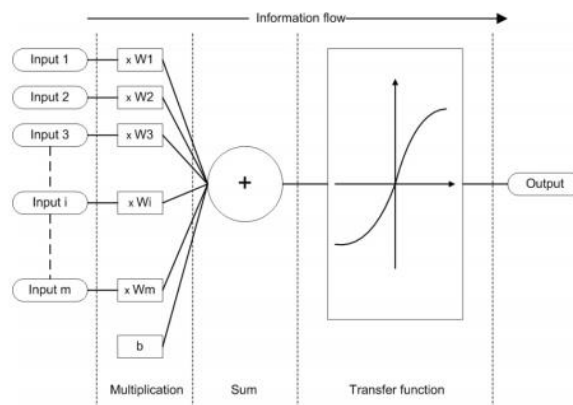
Tensorflow es una [25] biblioteca de software de código abierto dedicada al aprendizaje profundo que tiene como objetivo el definir, entrenar y desplegar modelos de aprendizaje automático. Tensorflow está desarrollada por Google, y fue lanzada en noviembre de 2015.

Su modelo computacional [26] en gráficos de flujo de estado variable. Los nodos de los grafos pueden ser mapeados a través de diferentes máquinas en un cluster, y dentro de cada

máquina a través de las CPU, GPU u otros dispositivos. Como se ha citado anteriormente, Tensorflow sirve está dedicado al aprendizaje profundo, pudiéndose aplicar a reconocimiento de voz, tratamiento de imágenes, robótica, recuperación de información y procesamiento del habla natural.

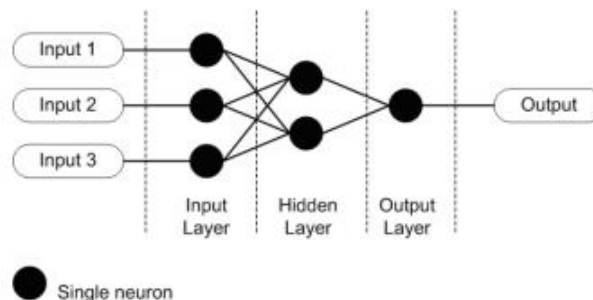
### 2.4.2. Red Neuronal Profunda

Una red neuronal artificial [34] es un modelo matemático que trata de simular la estructura y los comportamientos de las redes neuronales biológicas. El elemento básico de una red neuronal artificial es la neurona artificial, es decir, un simple modelo matemático. Dicho modelo tiene tres conjuntos de reglas simples: multiplicación, suma y activación. En la entrada de la neurona artificial las entradas se ponderan, lo que significa que cada valor de entrada se multiplica con el peso individual. En la sección media de la neurona artificial está la función de suma que suma todas las entradas ponderadas y el sesgo. A la salida de la neurona artificial la suma de las entradas previamente ponderadas y el sesgo está pasando a través de la función de activación que también se llama función de transferencia (Figura 2.8).



**Figura 2.8:** Principio de funcionamiento de una red neuronal. Imagen extraída de [34]

Por lo que, una red neuronal profunda se trata simplemente de una red neuronal artificial con una o más capas ocultas. En la figura 2.9, se muestra una ilustración de una red neuronal simple, siendo éste, el modelo que se ha seleccionado en este estudio.



**Figura 2.9:** Ejemplo de una red neuronal simple. Imagen extraída de [34]

## 3. Diseño

### 3.1. Introducción

Como se ha visto en el anterior capítulo, el objetivo de este estudio está enfocado en dos problemas, por una parte y como objetivo principal del Trabajo de Fin de Grado, la deforestación en los cuatro lugares más afectados de la Amazonia y, por otra parte, la reducción de contaminación en los países de España e Italia. El estudio de ambos temas se afronta con la misma herramienta, por lo que tanto este capítulo, como el siguiente, harán hincapié en la comprensión de esta y el manejo de las herramientas que proporciona.

Como se ha dicho anteriormente, Google proporciona una extensa documentación sobre el uso de esta, proporcionando ejemplos de casos de uso de la herramienta y en especial sobre el tema que se trata en este estudio. Esto ha “facilitado” el aprendizaje y el manejo de la herramienta.

El tema de la deforestación incluirá la parte de machine learning que se trató en el anterior capítulo, explicando los pasos a seguir para la inclusión de este apartado para el análisis de la deforestación.

### 3.2. Introducción a los conjuntos de datos utilizados para el estudio sobre la deforestación.

#### 3.2.1. Hansen et al. (2013) global forest cover and change data.

“High-Resolution Global Maps of 21st-Century Forest Cover Change” [19]; el dataset contiene el conjunto de imágenes del estudio realizado en el que se examinaron imágenes mundiales de Landsat con una resolución espacial de 30 metros para analizar la extensión, la pérdida y la ganancia de los bosques desde 2000 hasta 2018. Cabe destacar, que en el estudio que se ha llevado a cabo, se ha tomado la base de datos que abarca los años entre el 2000 y 2018, teniendo de esta manera datos más actualizados. Remarcar, que, en el dataset, los árboles se entienden como vegetación que supera los 5 metros de altura. El dataset contiene diferentes bandas, las cuales se explican a continuación en la siguiente tabla.

Band Name	Description	Range
treecover2000	Percentage of tree cover in the pixel.	0 - 100
loss	1 if loss ever happen during the study period.	0 or 1
gain	1 if gain ever happen during the study period.	0 or 1
lossyear	The year loss occurred, one-indexed from year 2001, or zero if no loss occurred.	0 - 12
first_b30	The Landsat 7 red band built from the first valid pixels in 2000 (or older if there were no valid pixels in 2000).	0 - 255
first_b40	The Landsat 7 near infrared band built from the first valid pixels in 2000.	0 - 255
first_b50	The first Landsat 7 short wave infrared band built from the first valid pixels in 2000.	0 - 255
first_b70	The second Landsat 7 short wave infrared band built from the first valid pixels in 2000.	0 - 255
last_b30	The Landsat 7 red band built from the latest valid pixels in 2012.	0 - 255
last_b40	The Landsat 7 near infrared band built from the latest valid pixels in 2012.	0 - 255
last_b50	The first Landsat 7 short wave infrared band built from the latest valid pixels 2012.	0 - 255
last_b70	The second Landsat 7 short wave infrared band built from the latest valid pixels 2012.	0 - 255
datamask	No data (0), mapped land surface (1), and permanent water bodies (2).	0, 1, 2

**Figura 3.1:** Bandas High-Resolution Global Maps of 21st-Century Forest Cover Change data.

### 3.2.2. “Forest Monitoring for Action” (FORMA)

“Forest Monitoring for Action” (FORMA) [21]; FORMA es un sistema de alertas de deforestación basado en el instrumento científico MODIS (Moderate-Resolution Imaging Spectroradiometer, incorporado en las sondas puestas en órbita por la NASA (Terra y Aqua), con el objetivo de medir la temperatura del suelo terrestre y océanos, detección de incendios, color de los océanos, detección de cambios en la vegetación, concentración de aerosoles) para los bosques de tipo tropical húmedos. FORMA está diseñado para hacer uso de dos de los productos de MODIS, el NDVI (Normalized Difference Vegetation Index) y FIRMS (Fires Information for Resource Management System). Es importante remarcar que las alertas sólo aparecen en áreas donde la probabilidad de pérdida de la cobertura arbórea es superior o igual al 25%. Los datos poseen una resolución espacial de 250 metros y un intervalo de actualización de 16 días para el NDVI y diario para el FIRM. Las alertas tienen una fecha de inicio en el año 2006.

### 3.2.3. USGS Landsat 8 Collection 1 Tier 1 Raw Scenes

El Landsat 8 [27] es un satélite lanzado el 11 de febrero de 2013, en asociación de la Administración Nacional de Aeronáutica y del Espacio (NASA por sus siglas en inglés) y el Servicio Geológico de los Estados Unidos (USGS). El satélite tenía y tiene como objetivo permitir la detección y caracterización de los cambios globales de la tierra a una escala en la que es posible diferenciar entre las causas producidas por la naturaleza o por la mano del ser humano.

La base de datos que se ha utilizado en el estudio, y en concreto, en el apartado de machine learning, tiene como característica fundamental que las escenas que contiene, captadas por el satélite Landsat 8, son aquellas con la calidad más alta de datos disponibles, es por ello por lo que se colocan en el *Tier 1*. El *Tier 1* incluye datos procesados de Terreno de Precisión de

Nivel 1 (L1TP) que tienen una radiometría bien caracterizada y están intercalados entre los diferentes sensores de Landsat.

El geo-registro de las escenas del Nivel 1 será consistente y estará dentro de las tolerancias prescritas [ $\leq 12$  m de error cuadrático medio (RMSE)]. Todos los datos de Landsat de Nivel 1 pueden considerarse consistentes e inter-calibrados (independientemente del sensor) a través de la colección completa.

En la imagen que se muestra a continuación, se muestran las diferentes bandas que tienen las imágenes contenidas en la base de datos:

Name	Resolution	Wavelength	Description
B1	30 meters	0.43 - 0.45 $\mu\text{m}$	Coastal aerosol
B2	30 meters	0.45 - 0.51 $\mu\text{m}$	Blue
B3	30 meters	0.53 - 0.59 $\mu\text{m}$	Green
B4	30 meters	0.64 - 0.67 $\mu\text{m}$	Red
B5	30 meters	0.85 - 0.88 $\mu\text{m}$	Near infrared
B6	30 meters	1.57 - 1.65 $\mu\text{m}$	Shortwave infrared 1
B7	30 meters	2.11 - 2.29 $\mu\text{m}$	Shortwave infrared 2
B8	15 meters	0.52 - 0.90 $\mu\text{m}$	Band 8 Panchromatic
B9	15 meters	1.36 - 1.38 $\mu\text{m}$	Cirrus
B10	30 meters	10.60 - 11.19 $\mu\text{m}$	Thermal infrared 1, resampled from 100m to 30m
B11	30 meters	11.50 - 12.51 $\mu\text{m}$	Thermal infrared 2, resampled from 100m to 30m
B0A			Landsat Collection 1 QA Bitmask (See Landsat QA page)

Figura 3.2: Bandas USGS Landsat 8 Collection 1 Tier 1 Raw Scenes

### 3.3. Introducción a los conjuntos de datos utilizados para el estudio sobre la contaminación.

#### 3.3.1. Sentinel-5 Precursor

Sentinel-5 Precursor (en adelante, S-5 P) [28] es un satélite lanzado por la Agencia Espacial Europea (ESA) de órbita polar terrestre baja. S-5-P proporciona información y servicios sobre la calidad del aire, el clima y la capa de ozono en el periodo de 2015-2022. La misión del S-5-P está encuadrada en el Programa del Componente Espacial de la Vigilancia Mundial del Medio Ambiente y la Seguridad (GMES). S-5-P equipa el Instrumento de Vigilancia Troposférica (TROPOMI) que tiene como objetivo medir los principales constituyentes atmosféricos, incluidos el ozono, el NO<sub>2</sub>, el SO<sub>2</sub>, el CO, el CH<sub>4</sub>, el CH<sub>2</sub>O y las propiedades de los aerosoles.

El conjunto de los datasets de S-5-P tienen dos versiones (excepto el de CH<sub>4</sub>). Por un lado, la versión Near Real-Time (NRTI) y por otro lado, la versión Offline (OFFL). Las versiones NRTI cubren un área más pequeña que los datasets OFFL, siendo la ventaja de las versiones NRTI la rapidez en la disposición de las imágenes tras su adquisición.

En el estudio se han utilizado dos bases de datos en concreto para el estudio sobre la contaminación:

- *Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide*: El dataset proporciona imágenes de alta resolución casi en tiempo real de las concentraciones de CO. Las imágenes del dataset tienen las siguientes bandas:

Name	Units	Min	Max	Description
CO_column_number_density	mol/m <sup>2</sup>	-279*	4.64*	Vertically integrated CO column density.
H2O_column_number_density	mol/m <sup>2</sup>	-465360*	3.45844e+07*	Water vapor column.
cloud_height	m	-8341*	5000*	Scattering layer height.
sensor_altitude	m	828542*	856078*	Altitude of the satellite with respect to the geodetic sub-satellite point (WGS84).
sensor_azimuth_angle	degrees	-180*	180*	Azimuth angle of the satellite at the ground pixel location (WGS84); angle measured East-of-North.
sensor_zenith_angle	degrees	1*	66*	Zenith angle of the satellite at the ground pixel location (WGS84); angle measured away from the vertical.
solar_azimuth_angle	degrees	-180*	180*	Azimuth angle of the Sun at the ground pixel location (WGS84); angle measured East-of-North.
solar_zenith_angle	degrees	9*	80*	Zenith angle of the satellite at the ground pixel location (WGS84); angle measured away from the vertical.

\* estimated min or max value

**Figura 3.3:** Bandas Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide

- *Sentinel-5P NRTI NO2: Near Real-Time Nitrogen Dioxide*: El dataset proporciona imágenes de alta resolución casi en tiempo real de las concentraciones de NO<sub>2</sub>. Las imágenes del dataset tienen las siguientes bandas:

Name	Units	Min	Max	Description
NO2_column_number_density	mol/m <sup>2</sup>	-0.0006*	0.0096*	Total vertical column of NO <sub>2</sub> (ratio of the slant column density of NO <sub>2</sub> and the total air mass factor).
tropospheric_NO2_column_number_density	mol/m <sup>2</sup>	-0.00064*	0.0096*	tropospheric vertical column of NO <sub>2</sub>
stratospheric_NO2_column_number_density	mol/m <sup>2</sup>	8.7e-06*	0.0001*	stratospheric vertical column of NO <sub>2</sub>
NO2_slant_column_number_density	mol/m <sup>2</sup>	1.4e-05*	0.003908*	NO <sub>2</sub> slant column density
tropopause_pressure	Pa	-0.000644*	0.009614*	tropopause pressure
absorbing_aerosol_index	Dimensionless	-14.43*	10.67*	Aerosol index (at wavelengths 354/388, i.e. the OMI pair) from the AER_AI level 2 product. See Level 2 Algorithms - Aerosol Index.
cloud_fraction	fraction	0*	1*	Effective cloud fraction. See the Sentinel 5P L2 Input/Output Data Definition Spec, p.220.
sensor_altitude	m	828543*	856078*	Altitude of the satellite with respect to the geodetic sub-satellite point (WGS84).
sensor_azimuth_angle	degrees	-180*	180*	Azimuth angle of the satellite at the ground pixel location (WGS84); angle measured East-of-North.
sensor_zenith_angle	degrees	0.098*	67*	Zenith angle of the satellite at the ground pixel location (WGS84); angle measured away from the vertical.
solar_azimuth_angle	degrees	-180*	180*	Azimuth angle of the Sun at the ground pixel location (WGS84); angle measured East-of-North.
solar_zenith_angle	degrees	8*	82*	Zenith angle of the satellite at the ground pixel location (WGS84); angle measured away from the vertical.

\* estimated min or max value

**Figura 3.4:** Bandas Sentinel-5P NRTI NO<sub>2</sub>: Near Real-Time Nitrogen Dioxide

### 3.4. Utilización de la red neuronal

Como se ha explicado en el capítulo anterior sobre el *Estado del Arte*, se ha dedicado una parte de este estudio a trabajar con una red neuronal para poder explorar otra de las múltiples posibilidades que brinda la plataforma de Google Earth Engine. Para la implementación de dicha red neuronal se ha hecho uso de diferentes tecnologías. En primer lugar, se ha utilizado

Google Collaboratory para importar el conjunto de datos de imágenes proporcionado por Google Earth Engine, entrenar nuestro clasificador de imágenes con dicho conjunto de datos y evaluar el modelo. Para esto último, el estudio se ha apoyado en la librería de Tensorflow. Así mismo, se ha utilizado la plataforma de Google Cloud Storage para almacenar todos los archivos que se han ido generando a raíz de la implementación de la red neuronal.

El trabajar con Google Collaboratory ha permitido ejecutar el código en los servidores en la nube de Google, aprovechando de esta manera la potencia del hardware de las máquinas de Google, incluyendo la GPU y TPU, independientemente de la potencia del equipo con el que se ha trabajado para realizar este estudio.

Para poder llevar a cabo la implementación de dicha red neuronal se han tenido que seguir unos pasos previos, siendo estos los siguientes:

- Exportar los datos en formato TFRecord (formato de Tensorflow) proporcionados por la plataforma de Google Earth Engine destinados a ser los datos de entrenamiento y testeo de nuestra red neuronal.
- Preparación de los datos para que puedan ser utilizados por un modelo basado en Tensorflow.
- Entrenamiento y validación del modelo en Tensorflow (utilizando *Keras* para ello).
- Hacer predicciones sobre la imagen exportada en formato TFRecord desde la plataforma Google Earth Engine.
- Hacer la ingesta de datos de la imagen clasificada en formato TFRecord en la plataforma Google Earth Engine para su posterior visualización.

A continuación, se muestra un diagrama de flujo en el que se muestra los diferentes pasos anteriormente enumerados:



**Figura 3.5:** Diagrama de flujo del trabajo entre las plataformas.

En el siguiente capítulo se profundizará en estos puntos para exponer con claridad los pasos que se han seguido para poder implementar la red neuronal.

## 4. Desarrollo

El presente capítulo tiene como objetivo mostrar los pasos que se han seguido para obtener los posteriores resultados relacionados con la deforestación en las áreas de Pará Mato Grosso, Rondonia y Amazonas. Así mismo se mostrarán los pasos seguidos para llegar a las conclusiones relacionadas con la contaminación a raíz de la pandemia provocada por el COVID-19.

### 4.1. Estudio sobre la deforestación

#### 4.1.1. Tratamiento de la base de datos “Hansen et al. (2013) global forest cover and change data.”

En este capítulo se tratarán los pasos que se han seguido para poder modelar la base de datos *Hansen et al. (2013) global forest cover and change data* en vistas a obtener los resultados de deforestación en las zonas de Pará, Mato Grosso, Rondonia y Amazonas.

En primer lugar, surge la problemática de tener que establecer los límites de las diferentes regiones, para ello, a priori se podrían marcar manualmente, teniendo que establecer múltiples puntos para poder delimitar de forma adecuada las fronteras de los estados. En un principio se optó por esta forma de proceder, no obstante, se observó que en múltiples ocasiones la precisión de los puntos en las fronteras no era la adecuada. Para poder dar solución a este problema, y tratar de hacer más eficiente el proceso de delimitación de las fronteras de los estados, se hizo uso de las tablas generadas por el Instituto Brasileiro de Geografía e Estatística, almacenadas y distribuidas por la universidad de Standford, en su apartado *Earthworks* (<https://earthworks.stanford.edu/>), siendo esto un repositorio de datos y mapas de sistemas de información geográfica (SIG, también citado como *GIS* por las siglas de su nombre en inglés). De este repositorio, se hizo uso de las tablas para los estados de Pará [29], Mato Grosso [30], Rondonia [31] y Amazonas [32].

Una vez se han descargado las tablas que contienen las fronteras de los estados sobre los que se trabajará, y se han subido a la plataforma de Google Earth Engine, se procederá a tratar la base de datos de *Hansen Global Forest Change v1.6 (2000-2018)* [33].

Como se mostró en la figura 2.8, las imágenes de la base de datos contienen diferentes bandas; se hará uso de algunas de dichas bandas para poder obtener el resultado esperado.

En primer lugar, se toma la banda “*Loss*”, esta banda contiene píxeles con valor ‘1’ donde se ha producido pérdida de cubierta forestal y ‘0’ en caso contrario. Por otra parte, se toma la banda “*gain*”, siendo esta una banda que es ‘1’ en aquellos píxeles donde se ha producido ganancia de cubierta forestal y un ‘0’ en caso contrario. A partir de estas dos bandas, surge una nueva banda que contiene tanto la ganancia como la pérdida.



Así mismo se ha añadido la banda “*treecover2000*”, conteniendo ésta la cubierta forestal para el año 2000, considerando cubierta forestal aquellos árboles con una altura superior a 5 metros.

La imagen que se obtendrá al final, después de añadir las bandas anteriormente citadas, será la cubierta forestal para el año 2000, y superpuestas, las bandas de “*loss*”, “*gain*” y la unión entre estas dos anteriores.

Una vez obtenida la imagen anterior, se procederá a calcular una estimación del área de pérdida de cubierta forestal en la región de interés deseada. Para ello, se ha de hacer uso de la imagen con la banda “*loss*”, a que “*loss*” está formado por píxeles que tienen un valor de 1 o 0 (por pérdida o no pérdida, respectivamente); la suma de estos valores equivale al número de píxeles de pérdida en la región. Dado que en todos los casos del estudio el número de píxeles de pérdida eran superiores al límite establecido por la plataforma (10 millones), se hizo uso de la función *ee.reducer.sum()* que permite devolver un reductor que calcula la suma (ponderada) de sus entradas. Así mismo, se estableció un máximo de (10 millones) de píxeles que entrasen en el cálculo.

A partir del cálculo anterior, se obtendría el número de píxeles que están considerados como pérdida, pero el dato que se busca en este apartado es hallar los metros cuadrados de pérdida que hay la región de interés. Es por ello, por lo que se hará uso de la función de la plataforma Earth Engine llamada *ee.Image.pixelArea()* que permite generar una imagen en la que el valor de cada píxel es el área del píxel en metros cuadrados. Multiplicando la imagen de pérdida (“*loss*”) con esta imagen de área y luego sumando sobre el resultado nos da una medida del área. De esta manera, ya se tendría el valor de pérdida en la región de interés deseada.

Por último, se ha hecho uso de otra de las bandas que proporciona la base de datos llamada “*lossYear*” que permite tener una estimación del área que se ha perdido y ganada entre los años de 2000 y 2018. Se ha hecho uso de dicha banda para poder mostrar la pérdida entre dichos años y generar gráficos que permitan comparar los datos arrojados por la plataforma con los datos oficiales del INPE.

Los resultados obtenidos en este apartado se procederán a visualizar en el próximo capítulo.

#### **4.1.2. Tratamiento de la base de datos “*Forest Monitoring for Action*” (FORMA)**

En este capítulo se tratarán los pasos que se han seguido para poder modelar la base de datos “*Forest Monitoring for Action*” (FORMA) en vistas a obtener los resultados de deforestación en las provincias de Pará, Mato Grosso, Rondonia y Amazonas.

Con el objetivo de poder enriquecer los potenciales resultados arrojados por la anterior base de datos, se ha tratado la base de datos de “*Forest Monitoring for Action*” (FORMA). Como se ha mencionado en capítulos anteriores, FORMA es un sistema de alertas de deforestación basado en el producto de MODIS para bosques tropicales. La base de datos de FORMA en Earth Engine es una imagen con alertas a partir de enero de 2006, con un periodo de

actualización de esta de un mes. Cada alerta tiene un tiempo asociado en una sola banda llamada “*alert\_date*” en unidades de segundos por época.

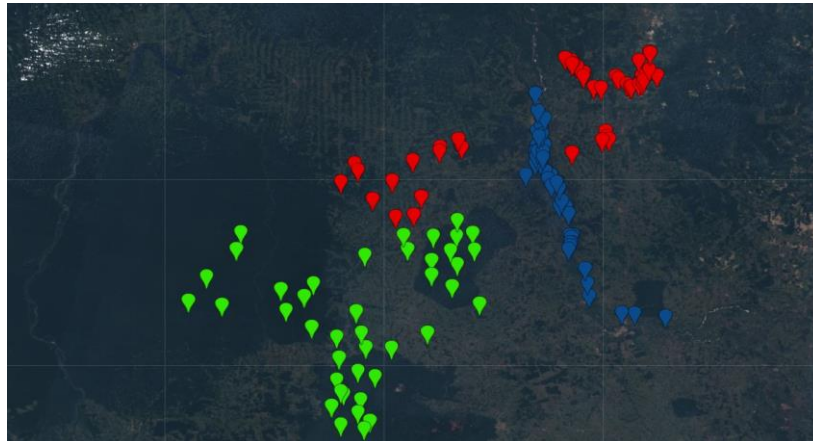
La estrategia que se ha seguido para tratar la base de datos ha sido filtrar por fechas y calcular alertas dentro de las regiones de interés anteriormente citadas. Las fechas que se han utilizado han sido desde 2006 hasta 2018, coincidiendo esta última con la fecha de la que se tienen datos de la anterior base de datos “*Hansen et al. (2013) global forest cover and change data*”. Se ha procedido de esta manera con el objetivo de, una vez se han calculado todos los resultados, poder comparar ambas imágenes y poder tener una visión de la deforestación potencial de las regiones de interés. En el apartado de resultados del estudio se podrá ver la imagen resultado del anterior cálculo.

### 4.1.3. Utilización de una red neuronal

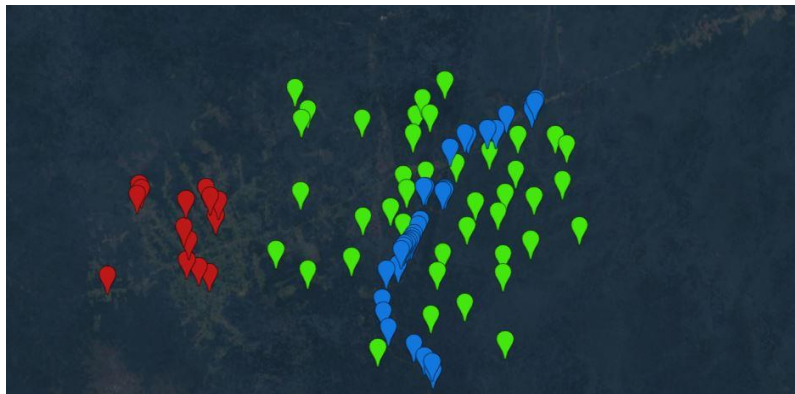
Como se indicaba en el anterior capítulo, el objetivo de utilizar una red neuronal es analizar zonas particulares de cada una de las cuatro zonas de interés planteadas para el estudio de la deforestación. Estas zonas han sido elegidas en base a los resultados obtenidos en los anteriores apartados, mediante el tratamiento de las bases de datos de “*Hansen et al. (2013) global forest cover and change data*” y *Forest Monitoring for Action* (FORMA), escogiendo aquellas zonas en las que aparentemente aparecía haber más deforestación acumulada.

Por otra parte, como ya se comentaba anteriormente, se ha hecho uso de múltiples plataformas para poder articular este apartado. Como tecnología principal en donde se ha planteado todo el código, se ha utilizado Google Collaboratory (en lo sucesivo, Collab). Collab ha permitido unificar las tecnologías de las que se ha hecho uso posteriormente: Google Cloud, Earth Engine y la librería de Tensorflow.

En primer lugar, y como paso necesario, se tienen que definir las clases que existen en la región de interés, ya que el tipo de aprendizaje es supervisado. Para establecer las clases, se ha cargado la base de datos “*USGS Landsat 8 Collection 1 Tier 1 Raw Scenes*” en Google Earth Engine que se utilizará posteriormente para entrenar a la red neuronal. Para las cuatro áreas de interés, se han establecido 3 clases: Vegetación, Deforestación y Agua. En las imágenes que se muestran a continuación, se ven las cuatro zonas de las regiones de interés escogidas para el entrenamiento de la red neuronal con las tres clases anteriormente marcadas.

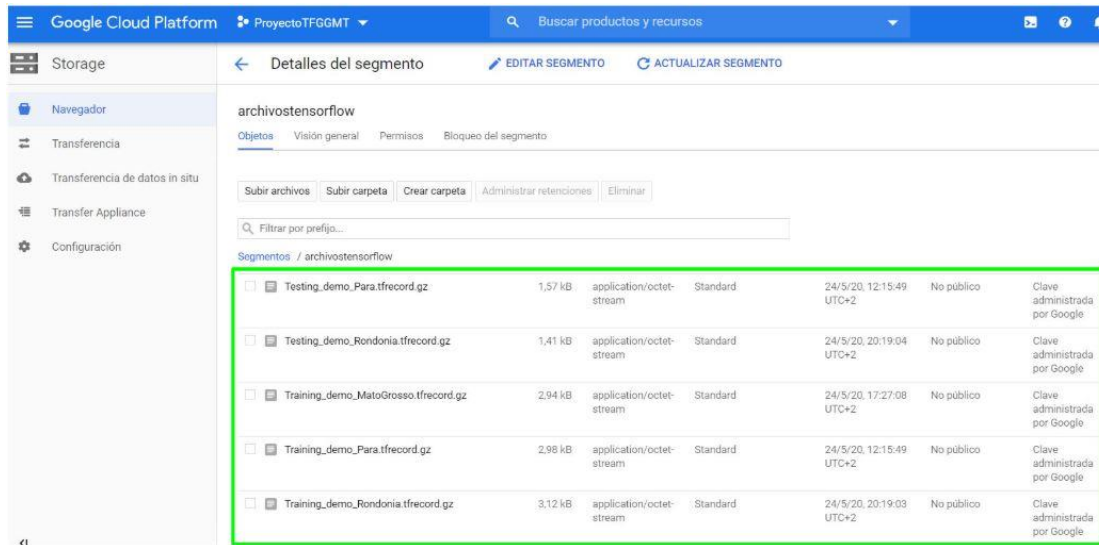


**Figura 4.1:** Puntos seleccionados para la clasificación en la región de Pará



**Figura 4.2:** Puntos seleccionados para la clasificación en la región de Amazonas

Una vez se han generado los puntos para las diferentes áreas de interés, se deben generar los dos archivos dedicados al entrenamiento y al testeo, respectivamente, de la red neuronal. en los cuatro casos se ha optado por destinar un 70% a la parte de entrenamiento y un 30% a la parte de testeo. Los archivos generados se subirán a la plataforma de Google Cloud.



**Figura 4.3:** Archivos de training y testing almacenados en plataforma Google Cloud.

Una vez adaptados los datos al formato de Tensorflow, y añadida y calculada la banda NDVI (*Índice de Vegetación Diferenciada Normalizada*). Se procederá a crear el modelo de red neuronal para más tarde entrenar dicha red neuronal y por último testear su rendimiento.

Para las cuatro regiones de interés se ha optado por una sencilla red neuronal con una capa oculta con 64 nodos, una capa de *dropout* y una capa de salida.

Una vez entrenada la red neuronal se procederá a testear la red con el archivo de testing anteriormente generado.

Por último, los resultados obtenidos se almacenarán en Google Cloud y se subirán a Google Earth Engine para su posterior visualización. Cabe destacar, que este último paso, en el que se deberían subir los archivos de Google Cloud a Google Earth Engine dio bastantes problemas, por lo que finalmente se optó por descargar los archivos en local desde la plataforma de Google Cloud y subirlos manualmente a la plataforma de Google Earth Engine.

Los resultados obtenidos en este apartado se procederán a visualizar en el próximo capítulo.

## 4.2. Estudio sobre la contaminación

### 4.2.1. Tratamiento de la base de datos *Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide*

Como se indicó en capítulos anteriores, el objetivo de tratar este conjunto de datos es analizar el comportamiento de la contaminación y en este subapartado en concreto monóxido de carbono (CO).

Este apartado será desarrollado en su plenitud en la plataforma de Google Earth Engine. En primer lugar, se debe cargar la base de datos y se seleccionará una de las bandas que ésta proporciona, en concreto, la banda “*CO\_column\_number\_density*”.

Como el objetivo del análisis es comparar dos momentos temporales, se han tomado el mes de enero y el mes de marzo para el estudio.

Así mismo se asignaron los máximos y los mínimos a una paleta de colores, en la que el color azul significa contaminación nula el color rojo significa gran densidad de monóxido de carbono.

Los resultados obtenidos a partir de este apartado se visualizarán en el siguiente capítulo.

#### **4.2.2. Tratamiento de la base de datos *Sentinel-5P NRTI NO2: Near Real-Time Nitrogen Dioxide***

Siguiendo la dinámica del apartado anterior, el objetivo de este apartado es tratar el conjunto de datos *Sentinel-5P NRTI NO2: Near Real-Time Nitrogen Dioxide* en donde se analizará el comportamiento del gas NO<sub>2</sub> entre los meses de enero y marzo, coincidiendo estos meses con el inicio de la pandemia y el pico de casos por contagio de COVID-19.

En primer lugar, se ha cargado la base de datos y se ha seleccionado la banda “*tropospheric\_NO2\_column\_number\_density*”. Dicha banda, permite conocer la densidad del gas acumulado NO<sub>2</sub> en la troposfera. Como se decía antes, el objetivo final de este apartado del estudio es comparar dos intervalos temporales, por lo que se procederá a acotar dos momentos. Para ello, como se indica al principio del apartado, se han tomado los meses de enero y marzo.

Así mismo se asignaron los máximos y los mínimos a una paleta de colores, en la que el color azul significa contaminación nula el color rojo significa gran densidad de dióxido de nitrógeno.

Los resultados obtenidos a partir de este apartado se visualizarán en el siguiente capítulo.

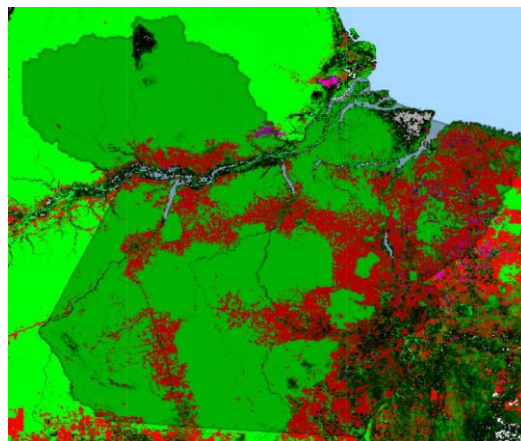
## 5. Integración, pruebas y resultados

En este capítulo se expondrán los resultados obtenidos para los diferentes apartados anteriormente presentados. En primer lugar, se mostrarán los resultados obtenidos en el bloque de del estudio de la deforestación, el tratamiento de la base de datos de *Hansen Global Forest Change v1.6 (2000-2018)* y su comparación con los datos proporcionados por el INPE sobre la deforestación, el tratamiento de la base de datos “*Forest Monitoring for Action*” (*FORMA*), la comparación entre ambas bases de datos, los resultados obtenidos a partir de la aplicación de una red neuronal a las cuatro áreas dentro de los estados del estudio y por último los resultados obtenidos a partir del tratamiento de las bases de datos de *Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide* y *Sentinel-5P NRTI NO2: Near Real-Time Nitrogen Dioxide*.

### 5.1. Estudio sobre la deforestación

#### 5.1.1. Resultados para la base de datos de *Hansen Global Forest Change v1.6 (2000-2018)*

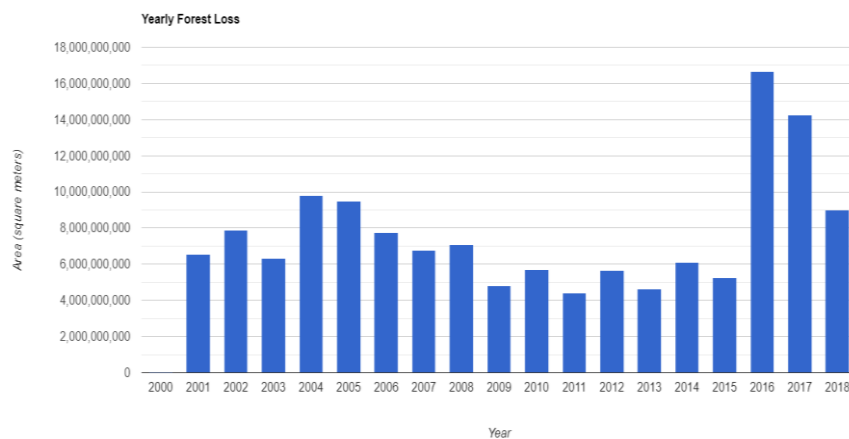
Tras haber tratado la base de datos *Hansen Global Forest Change v1.6 (2000-2018)* y haber aplicado los pasos descritos en el capítulo anterior, se obtuvieron los siguientes resultados para las cuatro regiones de interés:



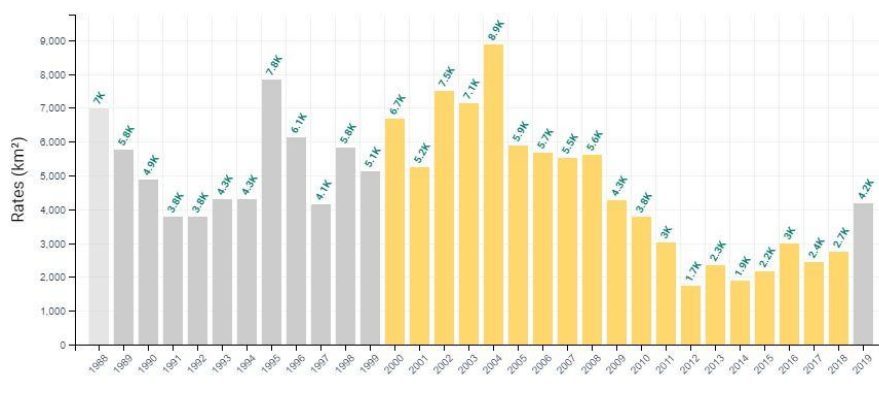
**Figura 5.1:** Resultado deforestación en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

Como se puede observar en la figura 5.1, la región de Pará se ve bastante afectada por el costado Este, acumulando la mayor parte de deforestación por ese frente. Así mismo, se puede apreciar la característica técnica de deforestación comentada en capítulos anteriores llamada “espinas de pescado”. Así mismo, la deforestación avanza hacia el centro de la región a través de dicha técnica mediante carretera ‘230’ que atraviesa el estado de Pará. Se puede observar cómo hacia los lados de dicha carretera la deforestación se va abriendo paso.

A partir del cálculo de la deforestación, se ha obtenido que, para el estado de Pará, los kilómetros cuadrados deforestados ascienden a un total de  $138.079 \text{ Km}^2$ . Contrastando este dato con el proporcionado por el INPE entre estos años  $85.479 \text{ Km}^2$ , no obstante si se compara con la deforestación total entre los años 1988 y 2018 en la región proporcionada por el INPE  $148.303 \text{ Km}^2$  se puede observar que se asemeja más, esto se puede deber a que, como se indica en las notas de usuario [33] de la base de datos, los años anteriores a 2011 todavía no han sido reprocesados con las técnicas más actuales, así mismo, cabe señalar también que no se ha realizado una validación completa de los resultados que incorporan el Landsat 8. Ese análisis puede revelar una capacidad más sensible para detectar y cartografiar las perturbaciones forestales con los datos del Landsat 8. De ser así, habrá una limitación más fundamental a la coherencia de la pérdida interanual cartografiada antes y después de la inclusión de los datos de Landsat 8. A continuación, se muestra una tabla y un gráfico en el que se muestra cómo se reparte la deforestación a lo largo de los años 2001 y 2018.



**Figura 5.2:** Gráfico deforestación por años en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)



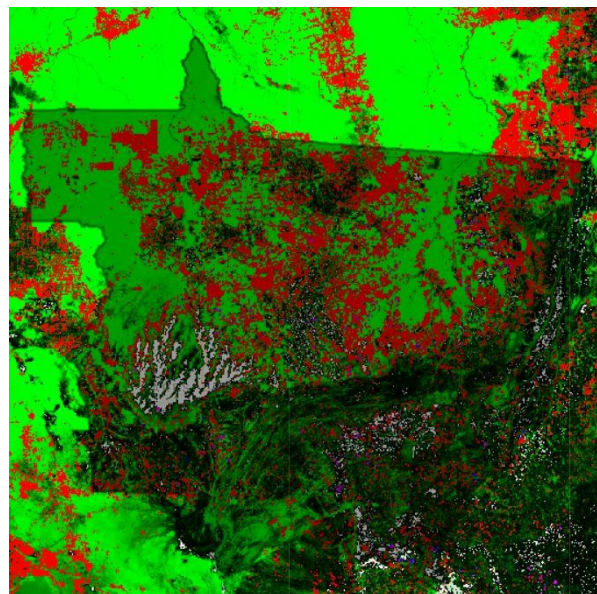
**Figura 5.3:** Gráfico deforestación por años en región Pará, datos del INPE [22]

Año	Deforestación(Km <sup>2</sup> )	Año	Deforestación(Km <sup>2</sup> )
2001	6537,5	2010	5687,1
2002	7867,5	2011	4398,2
2003	6306,7	2012	5641,5
2004	9796,1	2013	4650,3
2005	9474,2	2014	6117,5
2006	7737,0	2015	5255,5
2007	6751,5	2016	16643,5
2008	7105,0	2017	14275,2
2009	4819,6	2018	9015,1

**Tabla 5.1:** Deforestación por años en la región de Pará con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

Como se puede observar en las figuras 5.2 y 5.3, los datos en los primeros años guardan bastante relación, aproximándose bastante entre ellos, mientras que, en los últimos años, para la base de datos utilizada para el estudio, se obtienen valores muy superiores que los arrojados por el INPE. Esto se puede deber a lo que se ha mencionado al principio del capítulo, las técnicas utilizadas en los últimos años no están aplicadas a la totalidad de los años, por lo que es lógico que pueda haber discrepancias.

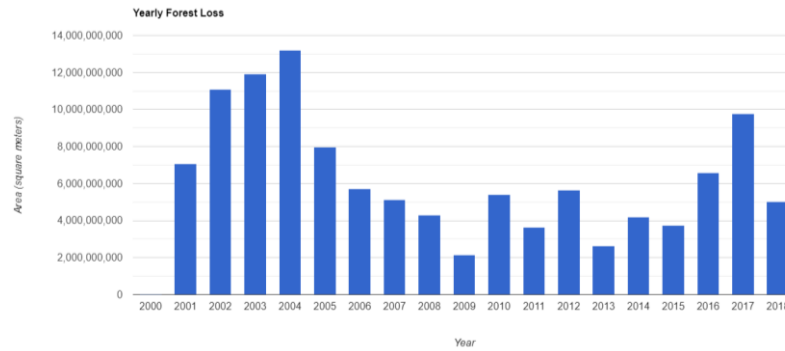
A continuación, se muestra la imagen obtenido referente a la región de Mato Grosso:



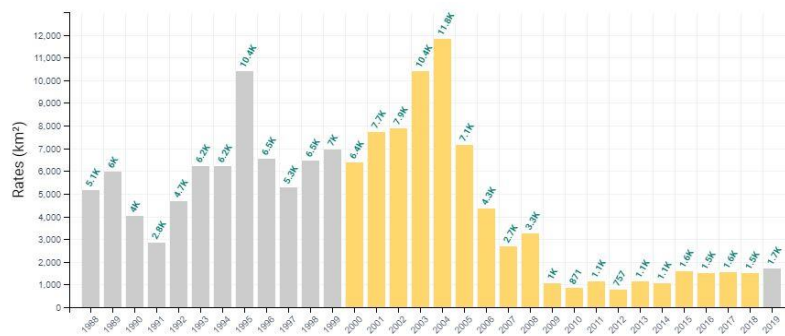
**Figura 5.4:** Resultado deforestación en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)



En la imagen se puede apreciar como en el caso de la región de Mato Grosso, los resultados obtenidos respecto a la deforestación, ésta se acumula en la zona central de la región. A continuación, se muestran los resultados obtenidos respecto a las cifras obtenidas sobre la deforestación repartida entre los años.



**Figura 5.5:** Gráfico deforestación por años en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)



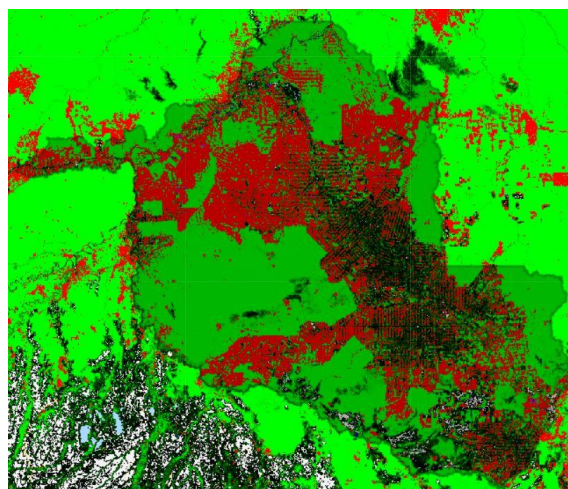
**Figura 5.6:** Gráfico deforestación por años en región Mato Grosso, datos del INPE [22]

Año	Deforestación(Km <sup>2</sup> )	Año	Deforestación(Km <sup>2</sup> )
2001	7068,8775	2010	5389,30644
2002	11100,7207	2011	3622,97622
2003	11926,3481	2012	5665,7097
2004	13220,1444	2013	2632,86412
2005	7976,33753	2014	4210,23632
2006	5720,25481	2015	3746,01213
2007	5141,08747	2016	6595,59853
2008	4281,56271	2017	9777,47035
2009	2148,25335	2018	5025,70481

**Tabla 5.2:** Deforestación por años en la región de Mato Grosso con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

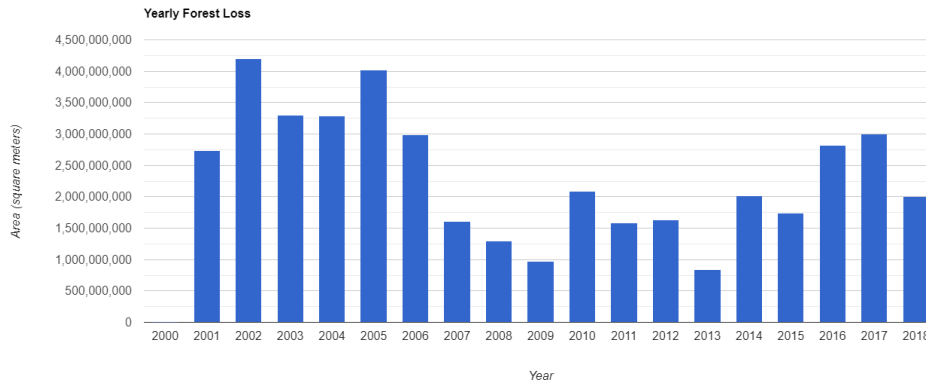
A partir del cálculo de la deforestación, se ha obtenido que, para el estado de Mato Grosso, los kilómetros cuadrados deforestados ascienden a un total de 115.249 Km<sup>2</sup>. Contrastando este dato con el proporcionado por el INPE entre estos años 73.749 Km<sup>2</sup>, no obstante, si se compara con la deforestación total entre los años 1988 y en la región proporcionada por el INPE 144.457 Km<sup>2</sup> se puede observar que se asemeja más, esto se puede deber a la razón anteriormente explicada sobre las notas de usuario. Esto, como se podrá observar, aplicará en las figuras que se muestran a continuación.

A continuación, se muestra la imagen obtenido referente a la región de Rondonia:

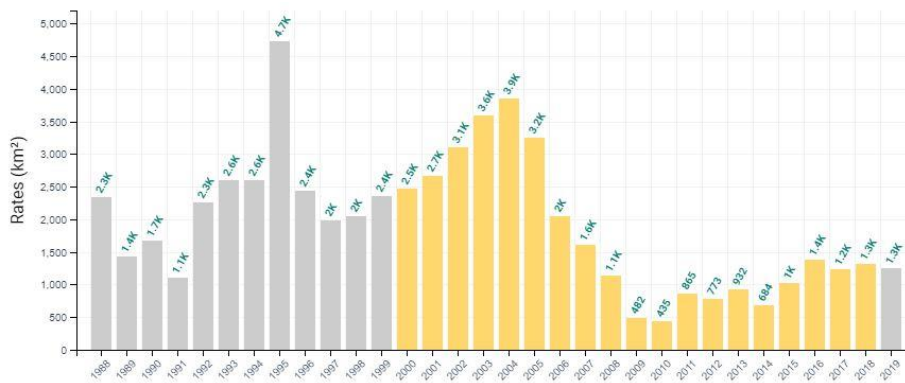


**Figura 5.7:** Resultado deforestación en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

En la imagen se puede apreciar como la deforestación predomina por la zona norte de la región, habiendo una gran concentración de terreno deforestado por dicha zona. Se ha podido observar cómo predomina la técnica de espina de pescado explicada anteriormente. Los datos obtenidos sobre la deforestación repartida entre los años son los siguientes:



**Figura 5.8:** Gráfico deforestación por años en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)



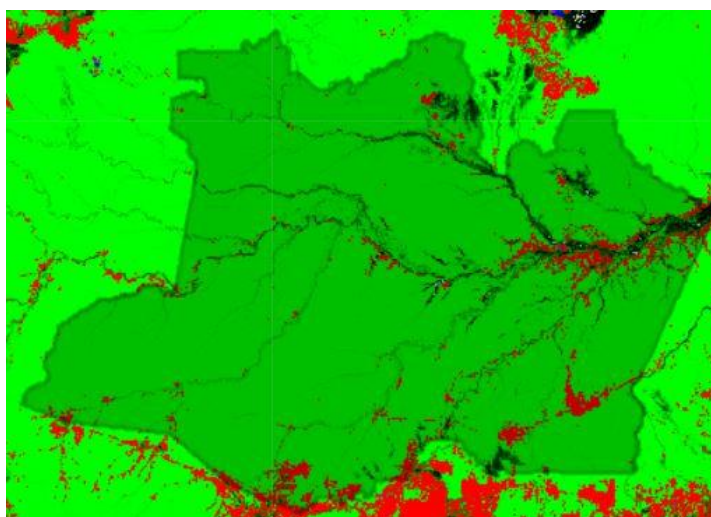
**Figura 5.9:** Gráfico deforestación por años en región Rondonia, datos del INPE [22]

Año	Deforestación(Km <sup>2</sup> )	Año	Deforestación(Km <sup>2</sup> )
2001	2735,34329	2010	2084,75251
2002	4194,46706	2011	1582,13497
2003	3294,556	2012	1632,2387
2004	3289,66627	2013	839,258654
2005	4023,88524	2014	2021,91424
2006	2992,69057	2015	1735,74193
2007	1610,94115	2016	2822,00966
2008	1290,32639	2017	2996,4779
2009	969,115287	2018	2009,20823

**Tabla 5.3:** Deforestación por años en la región de Rondonia con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

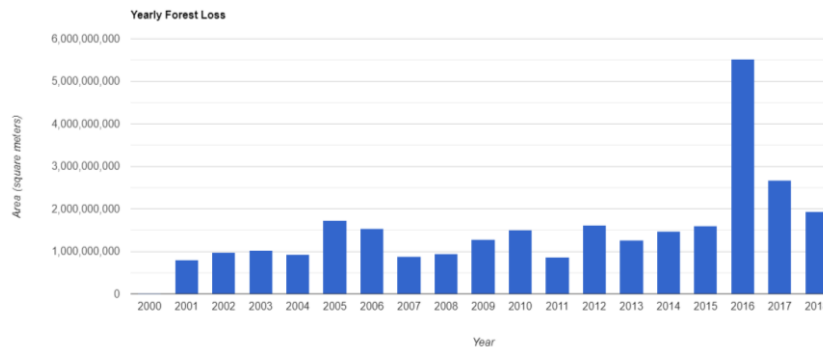
A partir del cálculo de la deforestación, se ha obtenido que para el estado de Rondonia, los kilómetros cuadrados deforestados ascienden a un total de 41.124Km<sup>2</sup>. Contrastando este dato con el proporcionado por el INPE entre estos años 32.868 Km<sup>2</sup>, no obstante, si se compara con la deforestación total entre los años 1988 y en la región proporcionada por el INPE 60.420Km<sup>2</sup> se puede observar que el dato obtenido se trata de un punto intermedio entre ambos datos, como se ha comentado, esto puede ser fruto del diferente procesamiento en las imágenes de los diferentes años.

Por último, se muestran los resultados obtenidos para la región del Amazonas:

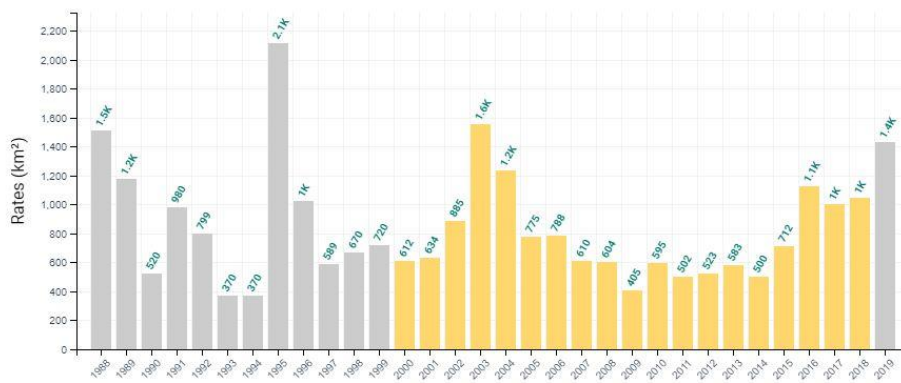


**Figura 5.10:** Resultado deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

En la imagen se puede apreciar como la región del Amazonas, en comparación con las otras tres regiones, es la menos afectada por la deforestación. no obstante, no hay que descuidar el hecho de la deforestación avanza de manera intensa por el sur, por la región de Rondonia, alcanzando los límites de la región del Amazonas. Así mismo, se puede observar que existen indicios de deforestación al este del Amazonas, coincidiendo con la cercanía a la ciudad de Manaos. A continuación, se muestran los resultados por años obtenidos para la región del Amazonas:



**Figura 5.11:** Gráfico deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)



**Figura 5.12:** Gráfico deforestación por años en región Amazonas, datos del INPE [22]

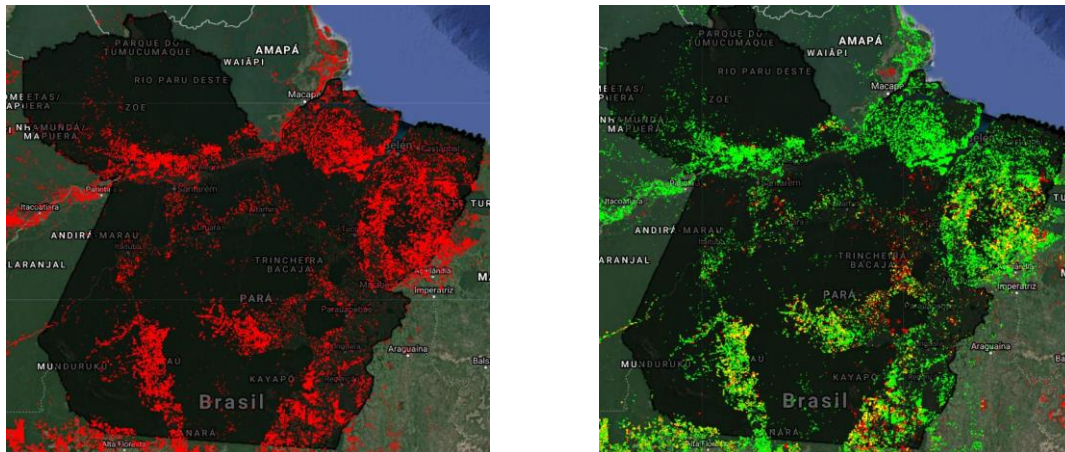
Año	Deforestación(Km <sup>2</sup> )	Año	Deforestación(Km <sup>2</sup> )
2001	800,663383	2010	1497,77647
2002	975,268816	2011	867,192406
2003	1016,40316	2012	1620,89669
2004	934,350011	2013	1268,40754
2005	1732,97855	2014	1479,1349
2006	1539,11909	2015	1600,06822
2007	879,467712	2016	5522,80127
2008	943,055808	2017	2675,54268
2009	1278,95588	2018	1942,29206

**Tabla 5.4:** Deforestación por años en la región de Amazonas con la base de datos de Hansen Global Forest Change v1.6 (2000-2018)

A partir del cálculo de la deforestación, se ha obtenido que, para el estado del Amazonas, los kilómetros cuadrados deforestados ascienden a un total de 28.574Km<sup>2</sup>. Contrastando este dato con el proporcionado por el INPE entre estos años 14.693 Km<sup>2</sup>, no obstante, si se compara con la deforestación total entre los años 1988 y en la región proporcionada por el INPE 25.538Km<sup>2</sup>, como se ha comentado, esto puede ser fruto del diferente procesamiento en las imágenes de los diferentes años.

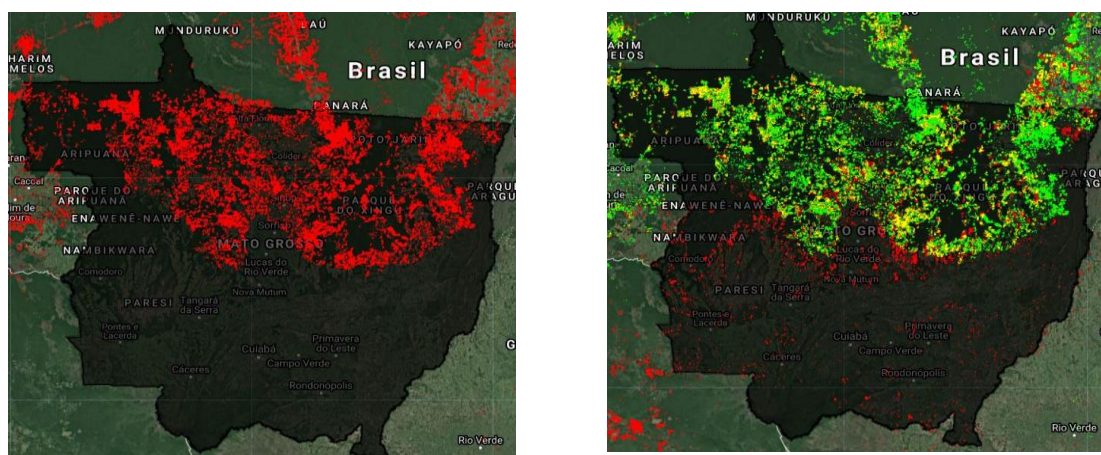
### 5.1.2. Tratamiento de la base de datos “*Forest Monitoring for Action*” (FORMA)

Tras haber tratado la base de datos *Foresto Monitoring for Action (FORMA)* y haber aplicado los pasos descritos en el capítulo anterior, se obtuvieron los siguientes resultados para las cuatro regiones de interés; así mismo, en este apartado se expondrá la comparación realizada entre la base de datos de este apartado y la de *Hansen Global Forest Change v1.6 (2000-2018)*.



**Figura 5.13:** Resultado obtenido para la base de datos de FORMA en la región de Pará. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Pará

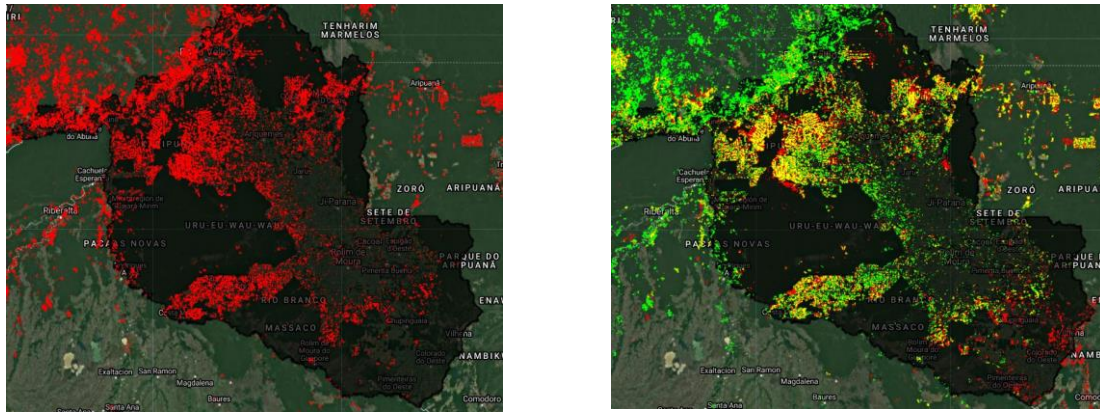
En la imagen se puede apreciar a la izquierda el resultado obtenido para la base de datos de FORMA y a la derecha, la comparación con la base de datos de Hansen para la región de Pará. Para la base de datos de FORMA, se han obtenido 200.176 avisos de deforestación. En la imagen de la derecha, las zonas en verde se corresponden con las zonas con alertas de la base de datos de FORMA, mientras que las zonas rojas se corresponden con las zonas en las que no hay alertas de FORMA, pero la base de datos de Hansen sí que ha detectado pérdida de cobertura forestal. Las zonas en amarillo se corresponden con aquellas partes en las que tanto FORMA, como Hansen han detectado pérdida de cobertura forestal. Como se puede observar, se obtienen más avisos de deforestación de FORMA que zonas identificadas como deforestadas por la base de datos de Hansen.



**Figura 5.14:** Resultado obtenido para la base de datos de FORMA en la región de Mato Grosso. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Mato Grosso.

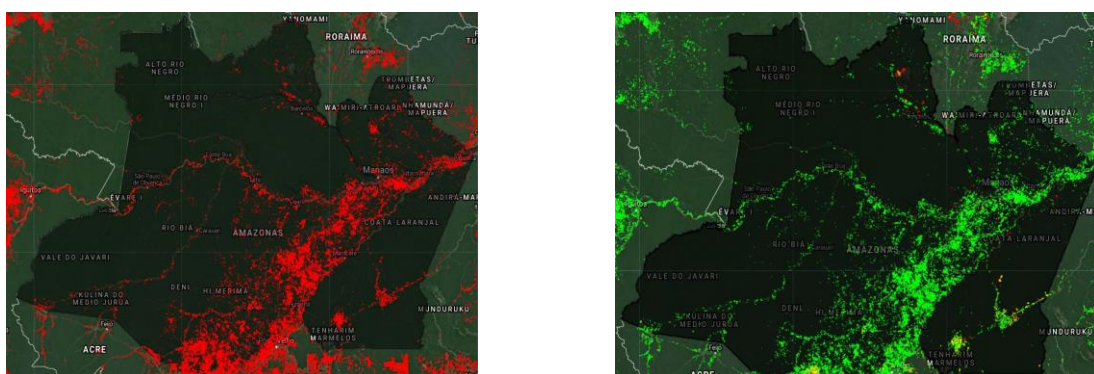
En la imagen se puede apreciar a la izquierda el resultado obtenido para la base de datos de FORMA y a la derecha, la comparación con la base de datos de Hansen para la región de Mato Grosso. Para la base de datos de FORMA, se han obtenido 146.539 avisos de

deforestación. Al igual que en la región de Pará, se puede apreciar que se obtienen más alertas de FORMA que regiones detectadas como deforestadas por la base de datos de Hansen.



**Figura 5.15:** Resultado obtenido para la base de datos de FORMA en la región de Rondonia. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Rondonia.

En la imagen se puede apreciar a la izquierda el resultado obtenido para la base de datos de FORMA y a la derecha, la comparación con la base de datos de Hansen para la región de Rondonia. Para la base de datos de FORMA, se han obtenido 75.811 avisos de deforestación. En la región de Rondonia se puede observar que, a diferencia que, en las anteriores regiones, en esta ocasión se obtiene un resultado parecido entre las alertas de FORMA y las bases de datos de Hansen, obteniendo de esta manera, zonas en las que el color amarillo predomina.



**Figura 5.16:** Resultado obtenido para la base de datos de FORMA en la región de Amazonas. Y comparación entre las bases de datos FORMA y Hansen Global Forest Change v1.6 (2000-2018) en la región de Amazonas

En la imagen se puede apreciar a la izquierda el resultado obtenido para la base de datos de FORMA y a la derecha, la comparación con la base de datos de Hansen para la región del Amazonas. Para la base de datos de FORMA, se han obtenido 146.539 avisos de deforestación. En la región del Amazonas se puede observar que, a diferencia que, en las anteriores regiones, en esta ocasión se obtiene un resultado muy diferente entre las dos bases de datos, obteniendo un número elevado de alertas de la base de datos FORMA, mientras que

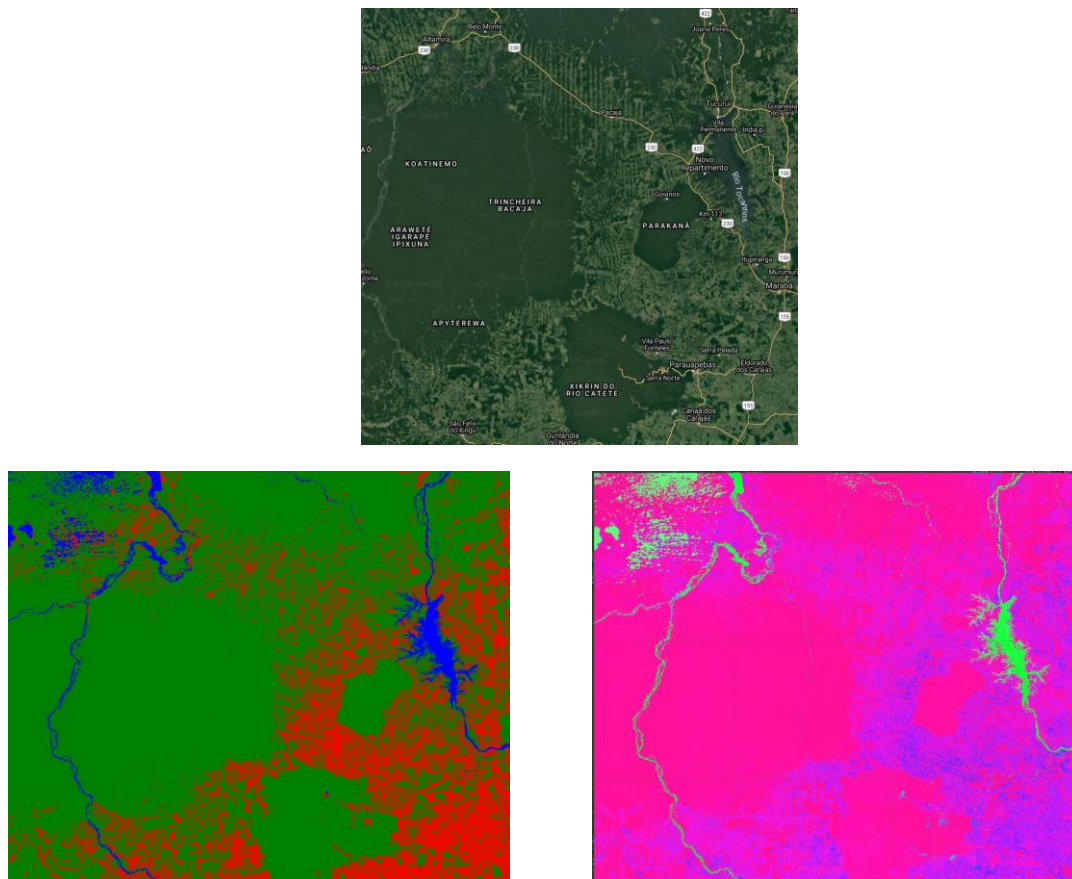


la pérdida de cobertura forestal identificada por la base de datos de Hansen es mucho más baja. Esto se puede deber a la diferencia de resolución entre las bases de datos, siendo de 500 metros en la base de datos de FORMA y de 30 en la de Hansen, pudiendo introducir más error en la primera de estas.

### 5.1.3. Resultados obtenidos a partir de la red neuronal implementada

A continuación, se muestran los datos obtenidos a partir de la implementación de la red neuronal, tras haber seguido los pasos que se expusieron en el anterior capítulo. Para las regiones de interés seleccionadas se expondrá la imagen obtenida a partir de la clasificación, la imagen de predicción obtenida y la zona de interés en concreto.

Para la región seleccionada del estado de Pará, se han obtenido los siguientes resultados:



**Figura 5.17:** Resultados obtenidos para la región de Pará

En primer lugar, se muestra la región de interés seleccionada del estado de Pará, a la izquierda la imagen obtenida a partir de la clasificación y a la derecha la imagen de probabilidad. Es interesante ver cómo en la imagen de la clasificación, la red neuronal parece no tener ningún problema a la hora de detectar zonas de agua y las zonas deforestadas. En la imagen de predicción cabe destacar, al compararla con la primera imagen, que en las zonas en las que

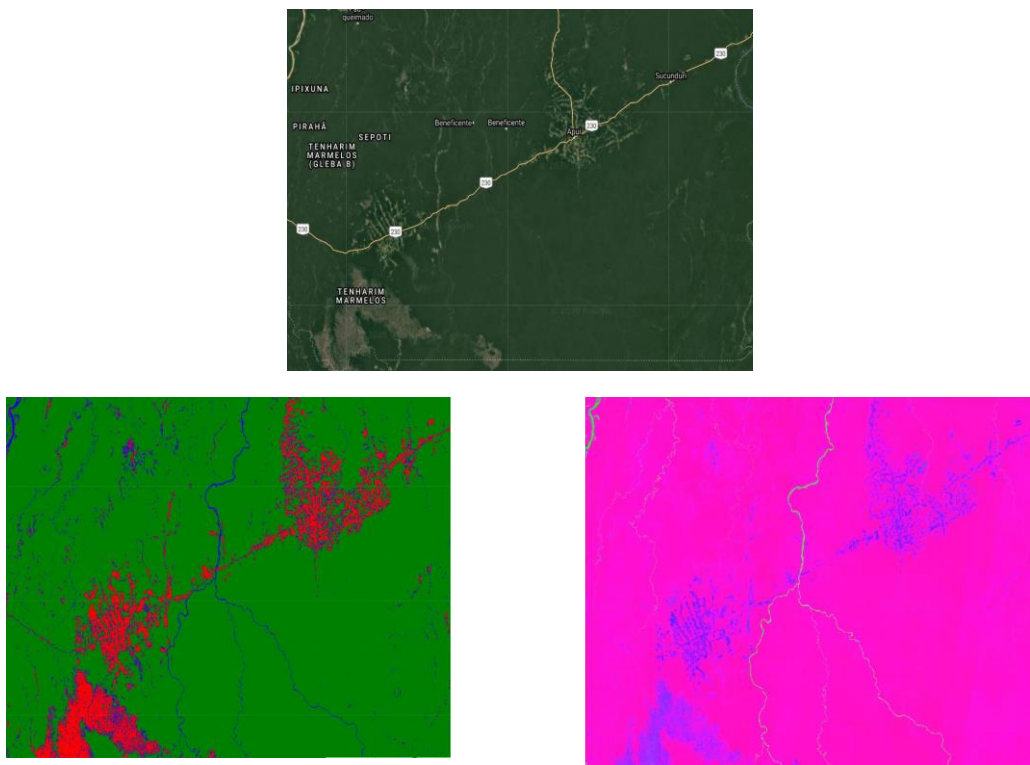
parece haber más deforestación/ menos cobertura forestal, obtenemos una mayor probabilidad, mostrándose esto con colores más oscuros en la imagen de predicción. Los resultados obtenidos sobre la precisión de la red neuronal en el dataset de testeo obtenida son los siguientes:

Loss: 0.3405

Accuracy: 0.9545

Como se puede observar, se ha obtenido un resultado bastante bueno para la red neuronal implementada, alcanzando una precisión del 95%, con una pérdida del 34%.

A continuación, se muestran los resultados obtenidos para la región de interés en el estado del Amazonas.



**Figura 5.18:** Resultados obtenidos para la región de Amazonas

En primer lugar, se muestra la región de interés seleccionada del estado del Amazonas, a la izquierda la imagen obtenida a partir de la clasificación y a la derecha la imagen de probabilidad. En este caso, se puede apreciar a simple vista la característica técnica de deforestación llamada “Espina de pescado”, se puede observar cómo a lo largo de la carretera que atraviesa la imagen por la mitad, la deforestación se va abriendo paso a cada uno de sus lados. Así mismo, se puede observar que el color en la imagen de la derecha es más intenso debido a la probabilidad mayor que se tiene de una cobertura forestal nula. Los resultados obtenidos sobre la precisión de la red neuronal en el *dataset* de testeo obtenida son los siguientes:

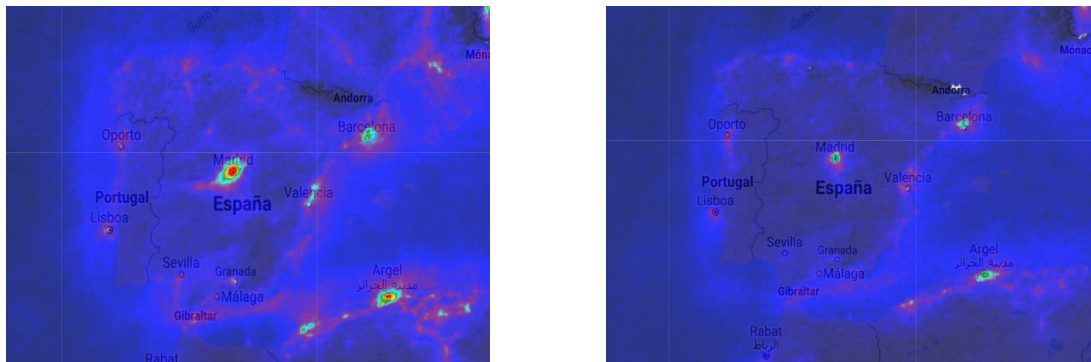
Loss: 0.4129

Accuracy: 0.9474

Como se puede observar, se ha obtenido un resultado bastante bueno para la red neuronal implementada, alcanzando una precisión del 94%, con una pérdida del 41%.

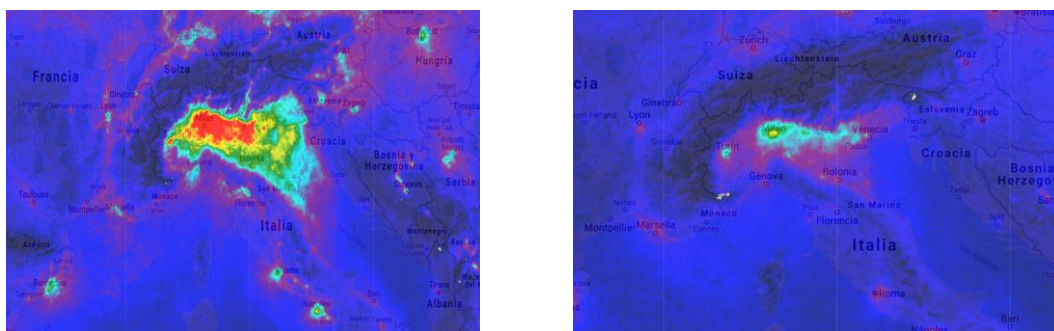
## 5.2. Estudio sobre la contaminación

Este capítulo expondrá los resultados obtenidos a partir del tratamiento de las bases de datos de Sentinel-5P NRTI CO: Near Real-Time Carbon Monoxide y Sentinel-5P NRTI NO<sub>2</sub>: Near Real-Time Nitrogen Dioxide. A continuación, se muestran las imágenes obtenidas para los países de España e Italia para ambos gases contaminantes, en las fechas de enero y marzo.



**Figura 5.19:** Resultados obtenidos para el gas NO<sub>2</sub> en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19

En las imágenes anteriores se puede observar el notable descenso de los niveles de NO<sub>2</sub> en España, a la izquierda la imagen de enero y a la derecha los niveles de marzo. Se puede apreciar como en la capital de España, Madrid, aparecen grandes niveles de concentración de NO<sub>2</sub> en el centro de la capital, viendo una gran mancha roja en el centro de la capital.

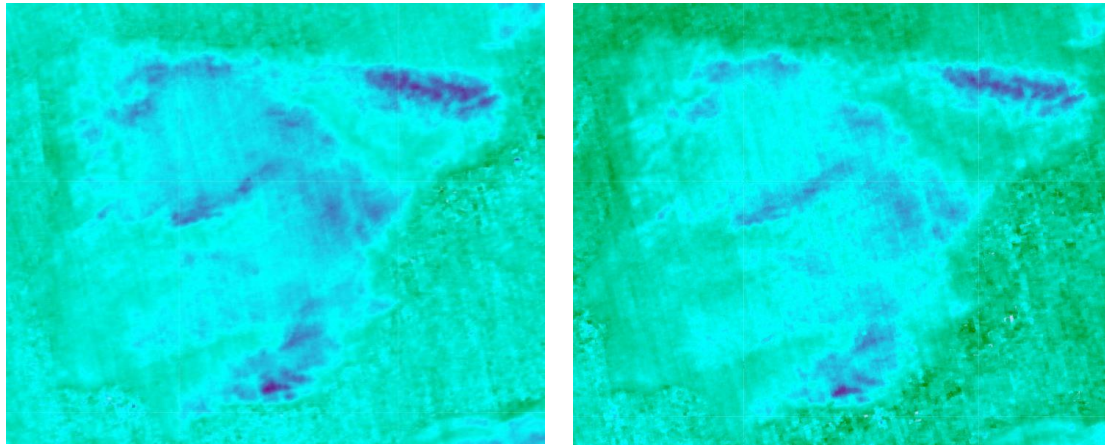


**Figura 5.20:** Resultados obtenidos para el gas NO<sub>2</sub> en Italia antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19

En las imágenes anteriores se muestran los resultados obtenidos para Italia. El cambio es fácilmente apreciable en los niveles de NO<sub>2</sub>. En la zona norte de Italia se puede observar una

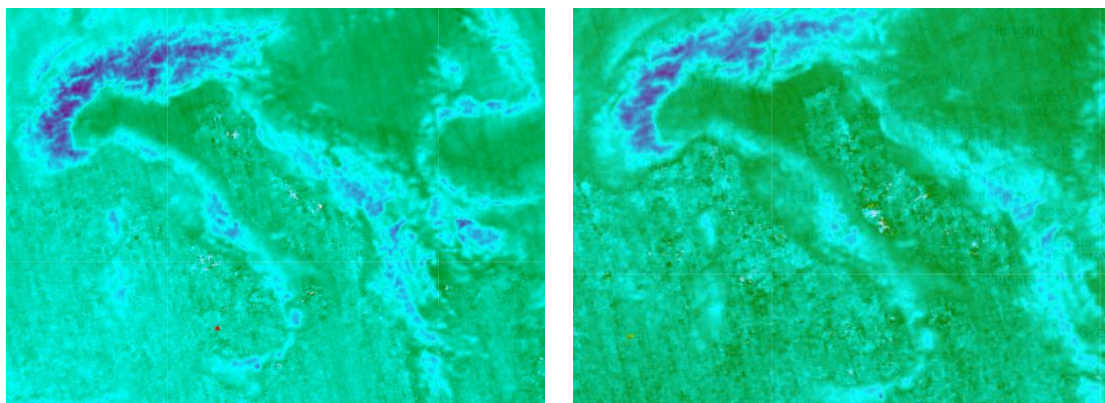
gran concentración de dicho gas, alcanzando unos niveles bastante altos como muestra la gran mancha roja de la imagen de la izquierda. Esta gran concentración se debe a los Alpes italianos, situados justamente al norte del país, haciendo de barrera y acumulando todo tipo de gases.

A continuación, se muestran los resultados obtenidos a partir del tratamiento de la base de datos del gas CO.



**Figura 5.21:** Resultados obtenidos para el gas CO en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19

Como se puede observar en las imágenes anteriores, los niveles de CO han bajado a lo largo de los meses de la cuarentena, sobre todo en la zona de Castilla y León. La zona en la que parece no haber descendido demasiado la concentración de CO es en la zona noreste del país, por los Alpes.



**Figura 5.22:** Resultados obtenidos para el gas CO en España antes (enero) y durante (marzo) de la pandemia provocada por el COVID-19

Como se podía observar en el anterior caso de Italia, con el gas NO<sub>2</sub>, los Alpes Italianos hacen de barrera para el gas de CO también, acumulando en la zona norte del país la mayor concentración del gas CO.

## 6. Conclusiones y Trabajo a futuro

### 6.1. Conclusiones

Este Trabajo Fin de Grado ha tratado de presentar dos de los grandes problemas a los que la sociedad tiene que hacer frente, la deforestación y la contaminación, a través de la pionera plataforma proporcionada por Google, Google Earth Engine. Se ha ofrecido una visión general del uso de la plataforma, así como de las posibilidades que ofrece de cara a poder llevar a cabo estudios más exhaustivos y poder obtener conclusiones de cara a denunciar abusos sobre la tierra, así como de su patrimonio.

Se ha expuesto la problemática de la deforestación para cuatro de las regiones más afectadas de la Amazonia brasileña, siendo éstas: Pará, Mato Grosso, Rondonia y Amazonas. Se han aportado los datos obtenidos a partir del estudio sobre cada una de las cuatro zonas y se han comparado con los datos proporcionados por el Instituto Nacional de Pesquisas Espaciais (INPE), obteniendo y comparando los resultados obtenidos para el estudio propio. Así mismo, y con el ánimo de poder aportar un valor añadido al presente trabajo, se ha implementado una sencilla red neuronal, aprovechando las posibilidades que ofrece la plataforma de Google Earth Engine para trabajar en conjunto con otras plataformas también ofrecidas por Google, como son: Google Collaboratory, Google Cloud y la librería de Tensorflow.

Por otra parte, y de manera adicional por la situación en la que este estudio se ha desarrollado, por la pandemia provocada por el COVID-19, se ha llevado a cabo el estudio de la contaminación en los países de Italia y España, coincidiendo éstos con los países más afectados de Europa por dicha pandemia. A partir de los resultados, se puede afirmar que, en ambos países, y en especial en Italia, se han reducido de manera notable los niveles de los principales gases contaminantes: NO<sub>2</sub> y CO.

Como conclusión, y a modo de reflexión personal, este trabajo permite exponer las grandes problemáticas a las que tenemos que hacer frente en la actualidad, así como denunciar los abusos que se están acometiendo sobre el mayor patrimonio que tenemos actualmente, la Tierra.

### 6.2. Trabajo a futuro

Como trabajo a futuro, se presenta la mejora en el estudio de la deforestación a partir de la inclusión de más bases de datos que permitan llevar a cabo un estudio más exhaustivo de la pérdida de cobertura forestal, así como incluir en el estudio los años 2019 y 2020, ya que durante el año 2019 [35] el INPE registró un aumento de la deforestación entre el año 2018 y 2019 de un 35%, lo que supone un aumento significativo, siendo éste un importante objeto de estudio. Así como expandir el estudio a otras zonas del planeta o profundizar más en las zonas expuestas en este estudio.

También se propone el estudio en una menor escala en las zonas propuestas, viendo así si la deforestación se está produciendo de manera aislada o se está comenzando una deforestación de una gran zona.

En relación con la implementación de una red neuronal, como trabajo a futuro se propone la ampliación de las clases para entrenar la misma, diferencian entre deforestación más o menos severa. Así como la implementación de otro tipo de redes neuronales con el objetivo de predecir posibles zonas que serán deforestadas en un futuro.

Así mismo, se propone profundizar más en el uso de la plataforma de Google Earth Engine, explorando sus múltiples opciones y conexión con el resto de los productos que ofrece Google.

## Bibliografía

- [1] Molina, D. O. (2012). Determinantes de la deforestación en la Amazonia. *Revista Interamericana de Ambiente y Turismo-RIAT*, 8(1), 17-25.
- [2] Subdirección de Información Ganadera y De Análisis y Política del Sector Dirección de Producción y Sanidad Animal - Políticas pecuarias 03 Subdirección de Información Ganadera y De Análisis y Política del Sector Dirección de Producción y Sanidad Animal - Políticas pecuarias 03. (2020). Recuperado 25 de junio de 2020, de <http://www.fao.org/>: <http://www.fao.org/3/a-a0262s.pdf>
- [3] Mertens, B., & E. Lambin. (1997). Spatial modelling of deforestation in southern Cameroon. Spatial disaggregation of diverse deforestation processes. *Applied Geography*, 17, 143-162.
- [4] Duarte, V., Martini, PR, Shimabukuro, YE, Freitas, RM y Arai, E. (2007). Monitoreo de la deforestación a escala global: una propuesta basada en los proyectos PRODES Digital y DETER. Simposio brasileño sobre teledetección, 13 , 6687-6694.
- [5] REAL ACADEMIA ESPAÑOLA: *Diccionario de la lengua española*, 23.<sup>a</sup> ed., [versión 23.3 en línea]. <<https://dle.rae.es>> [25 de junio de 2020].
- [6] Ballester, F. (2005). Contaminación atmosférica, cambio climático y salud. *Revista Española de Salud Pública*, 79(2), 159-175.
- [7] Téllez, J., & Rodríguez, A. (2006). Contaminación por monóxido de carbono: un problema de salud ambiental. *Revista de salud pública*, 8, 108-117.
- [8] Subdirección de Información Ganadera y De Análisis y Política del Sector Dirección de Producción y Sanidad Animal - Políticas pecuarias 03 Subdirección de Información Ganadera y De Análisis y Política del Sector Dirección de Producción y Sanidad Animal - Políticas pecuarias 03. (2020). Recuperado 25 de junio de 2020, de <https://www.saludgeoambiental.org/dioxido-nitrogeno-no2>
- [9] Chaolin Huang; Yeming Wang; Xingwang Li; Lili Ren; Jianping Zhao; Yi Hu; Li Zhang; Guohui Fan; Jiuyang Xu; Xiaoying Gu (30 de enero de 2020). «Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China» [Características clínicas de los pacientes infectados con la novela 2019 coronavirus en Wuhan, China] (pdf). *The Lancet* (en inglés). p. 3.
- [10] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*.
- [11] Omar Orellana Díaz. (octubre 2018). Introducción Básica a Google Earth Engine (GEE). Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, 1, 24. 25

de junio 2020, De Programa Regional Reducción de Emisiones por Deforestación y Degradación de Bosques en Centroamérica y República Dominicana (REDD/CCAD-GIZ) Base de datos.

[12] Verma, A., Pedrosa, L., Korupolu, M., Oppenheimer, D., Tune, E., & Wilkes, J. (2015, April). Large-scale cluster management at Google with Borg. In *Proceedings of the Tenth European Conference on Computer Systems* (pp. 1-17).

[13] Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., ... & Gruber, R. E. (2008). Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems (TOCS)*, 26(2), 1-26.

[14] Corbett, J. C., Dean, J., Epstein, M., Fikes, A., Frost, C., Furman, J. J., ... & Hsieh, W. (2013). Spanner: Google's globally distributed database. *ACM Transactions on Computer Systems (TOCS)*, 31(3), 1-22.

[15] Ghemawat, S., Gobiuff, H., & Leung, S. T. (2003, October). The Google file system. In *Proceedings of the nineteenth ACM symposium on Operating systems principles* (pp. 29-43).

[16] Chambers, C., Raniwala, A., Perry, F., Adams, S., Henry, R. R., Bradshaw, R., & Weizenbaum, N. (2010). FlumeJava: easy, efficient data-parallel pipelines. *ACM Sigplan Notices*, 45(6), 363-375.

[17] Gonzalez, H., Halevy, A. Y., Jensen, C. S., Langen, A., Madhavan, J., Shapley, R., ... & Goldberg-Kidon, J. (2010, June). Google fusion tables: web-centered data management and collaboration. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data* (pp. 1061-1066).

[18] Bisong, E. (2019). Google Colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (pp. 59-64). Apress, Berkeley, CA.

[19] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., ... & Kommareddy, A. (2013). High-resolution global maps of 21st-century forest cover change. *science*, 342(6160), 850-853.

[20] Hammer, D., Kraft, R., & Wheeler, D. (2009). Forma: Forest monitoring for action-rapid identification of pan-tropical deforestation using moderate-resolution remotely sensed data. *Center for Global Development Working Paper*, (192).

[21] Hammer, Dan, Robin Kraft, and David Wheeler. 2013. FORMA Alerts. World Resources Institute and Center for Global Development. [www.globalforestwatch.org](http://www.globalforestwatch.org).

[22] PRODES (Deforestation). (2020). Recuperado 25 de abril de 2020, de TerraBrasilis website:

[http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/biomes/legal\\_amazon/rates](http://terrabrasilis.dpi.inpe.br/app/dashboard/deforestation/biomes/legal_amazon/rates)



- [23] Loh, W. Y. (2011). Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1), 14-23.
- [24] De'ath, G., & Fabricius, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), 3178-3192.
- [25] Goldsborough, P. (2016). A tour of tensorflow. arXiv preprint arXiv:1610.01178.
- [26] Abadi, M. (2016, September). TensorFlow: learning functions at scale. In *Proceedings of the 21st ACM SIGPLAN International Conference on Functional Programming* (pp. 1-1).
- [27] Loveland, T. R., & Irons, J. R. (2016). Landsat 8: The plans, the reality, and the legacy. *Remote Sensing of Environment*, 185, 1-6.
- [28] Veefkind, J. P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., ... & Van Weele, M. (2012). TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, 120, 70-83.
- [29] Instituto Brasileiro de Geografia e Estatística. (2011). State Boundary: Pará, Brasil, 2010 IBGE. Available at: <http://purl.stanford.edu/jv851jq5632>
- [30] Instituto Brasileiro de Geografia e Estatística. (2011). State Boundary: Mato Grosso, Brasil, 2010 IBGE. Available at: <http://purl.stanford.edu/mt656st7052>
- [31] Instituto Brasileiro de Geografia e Estatística. (2011). State Boundary: Rondônia, Brasil, 2010 IBGE. Available at: <http://purl.stanford.edu/jd972mq7065>
- [32] Instituto Brasileiro de Geografia e Estatística. (2011). State Boundary: Amazonas, Brasil, 2010 IBGE. Available at: <http://purl.stanford.edu/hs472ch1790>
- [33] Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (15 November): 850–53. Data available on-line at: <http://earthenginepartners.appspot.com/science-2013-global-forest>.
- [34] Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the artificial neural networks. *Artificial Neural Networks: Methodological Advances and Biomedical Applications*. InTech, 1-18.
- [35] La tasa consolidada de deforestación clara para los nueve estados de la Amazonía Legal (AC, AM, AP, MA, MT, PA, RO, RR y TO) en 2019 es de 10,129 km<sup>2</sup>. (2020). Recuperado 25 de junio de 2020, de <http://www.inpe.br/> website: [http://www.inpe.br/noticias/noticia.php?Cod\\_Noticia=5465](http://www.inpe.br/noticias/noticia.php?Cod_Noticia=5465)