# Extracción y procesamiento de datos de lluvia

CHIRPS Daily via Google Earth Engine

Isaac Arroyo

20 de mayo de 2024

# Tabla de contenidos

1	Introducción		2	
2	Sob	Sobre los datos  Varibles y constantes  4		
3	Vari			
4	Reducción a valores mensuales			
	4.1	Etiquetado de año y mes del año	6	
	4.2	Precipitación mensual del año de interés	7	
	4.3	Meses como bandas de una ee.ImageCollection	8	
5	Métricas a extraer			
	5.1	Acumulación normal	9	
	5.2	Anomalía en milimetros	11	
	5.3	Anomalía en porcentaje	11	
6	De raster a CSV			
	6.1	Información de ee.Image a ee.FeatureCollection	13	
	6.2	Exportar ee.FeatureCollection a CSV	14	
7	Código final			
	7.1	Funciones escenciales	16	
	7.2	Función de extracción, procesamiento y exportación de datos	16	
	7.3	Extracción	19	

## Introducción

En este documento se encuentran documentados los pasos y el código para la extracción mensual de variables derivadas de la precipitación, tales como: precipitación mensual promedio, anomalía de la precipitación en porcentaje con respecto de la normal y anomalía de la precipitación en milímetros con respecto de la normal.

Cada aspecto del código se documenta en diferentes capítulos, donde el último capítulo estará la función final, con la que se resume y concluye el proceso de extracción.

## Sobre los datos

Los fuente de los datos se llama CHIRPS (Climate Hazards Group InfraRed Precipitation With Station Data) Daily, se puede encontrar en diferentes lugares, uno de estos la plataforma Google Earth Engine.

De acuerdo con la descripción:

Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) is a 30+ year quasi-global rainfall dataset. CHIRPS incorporates 0.05° resolution satellite imagery with in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring.

Estos datos cuentan con la **precipitación diaria** medida en milímetros (mm) desde Enero 01, de 1981 hasta el mes inmediato anterior a la fecha actual<sup>1</sup>

El procesamiento de texto es similar al realizado en el proyecto "Desplazamiento climático: La migración que no vemos".

```
import ee

try:
    ee.Initialize()
    print("Se ha inicializado correctamente")
except:
    print("Error en la inicialización")
```

- (1) Importar API de (Google) Earth Engine
- (2) Inicializar API

<sup>&</sup>lt;sup>1</sup>Esto quiere decir, que si la fecha actual es Abril 2024, entonces los datos cubren hasta Marzo 2024

# Varibles y constantes

La extracción de datos se planea que sea periódica a niveles estatales y municipales, por lo que se dejan declarados variables que se mantendrán constantes (como el rango del promedio *normal* o histórico) o (valga la redundancia) cambiarán dependiendo de los datos que se quieran. La Tabla 3.1 entra a mayor detalle de lo que se esta haciendo

Tabla 3.1: Variables y constantes

Variable	Tipo	Notas	
date_year_interes	Cambiante Cambiante	Año del que se van a extraer las métricas de precipitación ee.FeatureCollection de las geomtrías e información de los Estados o Municipios de México	
fc			
chirps	Constante	ee.ImageCollection de CHIRPS Daily	
year_base_inicio	Constante	Inicio del periodo historico para el cálculo de la normal, Enero 01 de 1981	
<pre>year_base_fin</pre>	Constante	Fin del periodo historico para el cálculo de la normal, Diciembre 31 de 2010	
geom_mex	Constante	Geometría del perímetro de México, usada para delimitar espacialmente la información	

date\_year\_interes = 2023
select\_fc = "mun"

2

```
dict fc = dict(
                                                                                    (3)
       ent = "projects/ee-unisaacarroyov/assets/GEOM-MX/MX ENT 2022",
       mun = "projects/ee-unisaacarroyov/assets/GEOM-MX/MX_MUN_2022")
   fc = ee.FeatureCollection(dict fc[select fc])
                                                                                    (4)
   year_base_inicio = 1981
                                                                                    (5)
   year base fin = 2010
   geom mex = (ee.FeatureCollection("USDOS/LSIB/2017")
                                                                                    (6)
10
                .filter(ee.Filter.eq("COUNTRY NA", "Mexico"))
                                                                                    7
11
                .first()
                                                                                    (8)
12
                .geometry())
                                                                                    9
13
14
   chirps = (ee.ImageCollection('UCSB-CHG/CHIRPS/DAILY')
                                                                                   (10)
15
              .select("precipitation")
16
              .filter(ee.Filter.bounds(geom mex)))
                                                                                   (11)
17
```

- (1) Selección del año de interés
- ② Seleccion de ee.FeatureCollection, sea de Entidades (ent), Municipios (mun) o Cuencas Hidrológicas (ch¹)
- (3) Diccionario con los paths hacia la ee.FeatureCollection de elección
- (4) Carga de ee.FeatureCollection de interés
- (5) Fechas de inicio y fin del periodo historico para el cálculo de la normal (30 años)
- (6) ee. FeatureCollection de división politica de los países del mundo
- (7) Filtro donde la propiedad COUNTRY NA sea igual a **Mexico**
- (8) Selección de la primera ee.Feature
- (9) Extracción de únicamente la geometría
- (10) ee.ImageCollection de CHIRPS Daily
- (11) Limitar el raster a la geometría de México

<sup>&</sup>lt;sup>1</sup>Pendiente subir a Google Earth Engine

## Reducción a valores mensuales

chirps es una ee.ImageCollection de más de 15 mil imágenes (ee.Image). El procesamiento y tratado de las imagenes es *pesado*, y aunque la computadora no se encargue de hacer el trabajo, esto puede hacer que el servidor de Google Earth demore en hacer los cálculos y la transformación de los datos raster.

Es por eso que en el procesamiento se incluye un primer filtro: la creación de una ee.ImageCollection de 365 imágenes<sup>1</sup>.

```
chirps_year_interes = (chirps
.filter(ee.Filter.calendarRange(start = date_year_interes,
field = "year")))
```

## 4.1 Etiquetado de año y mes del año

Para ir agrupando y sumando la precipitación mensual, hay que tener las imágenes etiquetadas con el mes para poder agruparlas y sumar la precipitación. Para ello se crea una función que haga ese etiquetado en cada una de las imágenes

- (1) La función toma una sola ee. Image
- ② Obtener la fecha de la imagen, como esta en formato UNIX, se tiene que transformar a fecha con ee.Date

<sup>&</sup>lt;sup>1</sup>366 si es año bisiesto

- (3) De la fecha se obtiene el valor numérico de la semana del año
- (4) Asignación de año y semana del año como propiedades de la ee. Image
- (5) Crear nueva ee. ImageCollection con el etiquetado

## 4.2 Precipitación mensual del año de interés

Para una fácil extracción de los valores del año es necesario tener la información de la precipitación (o la métrica de interés) como una ee. Image de 12 bandas, donde cada banda es el valor mensual de la región.

Para crear una imagen de 12 bandas se necesita primero una ee.ImageColletion de 12 imágenes.

Para ello, se va a crear una lista de 12 imágenes.

```
list_months = ee.List.sequence(1, 12)
```

La transformación conlleva múltiples iteraciones, y mientras que en JavaScript se puedan declarar funciones dentro de map, para el caso se Python se tendran que crear las funciones a parte, y después serán llamadas a su respectivo map.

- ① Función para sumar todas las imágenes que pertenezcan a un mes en específico. Esta función itera sobre elementos de una ee.List, estos elementos son los los meses que ocupa la ee.ImageCollection
- (2) Se filtran aquellas imágenes que sean del mes de interés
- (3) Se reduce la colección a una imagen a través de la suma
- ④ Se le asigna la propiedad del mes que representa.

- ① Se aplica la función para reducir el número de imagenes en la colección, como resultado da una lista de 12 imágenes
- ② Se crea una colección de imágenes a partir de una lista.

## 4.3 Meses como bandas de una ee. ImageCollection

Ya que se logró tener una colección de 12 imágenes, entonces se crea la imagen de 12 bandas

```
img_monthly_pr = (img_coll_monthly_pr
    .toBands()
    .rename([f"0{i}" if i < 10 else str(i) for i in range(1,13)]))
2</pre>
```

- ① Pasar ee. ImageCollection de n imágenes a una ee. Image de n bandas
- (2) Renombrar el nombre de las bandas a los números de los meses

## Métricas a extraer

El objetivo es poder extraer información mensual sobre las lluvias de cada año, a las que se les prestará atención son:

- Precipitación
- · Anomalía de precipitación en mm
- Anomalía de precipitación en porcentaje

Para las últimas dos hace falta tener el valor de la acumulación normal.

#### 5.1 Acumulación normal

De acuerdo a con el glosario de la NOAA una anomalía es la desviación de una unidad dentro de un periodo en una región en específico con respecto a su promedio histórico o normal. Este promedio es usualmente de 30 años.

Para el caso del CHIRPS, es de 1981 hasta el 2010 (incluyendo a diciembre).

Esta tarea se tiene que hacer con ayuda de dos funciones. La primera etiquetará por año y mes la colección *base*. La segunda función tendrá que reducir a 12 imagenes cada año de esa base.

```
def func_tag_year_month_base_period(img):
    full_date = ee.Date(ee.Number(img.get("system:time_start")))
    n_year = ee.Number(full_date.get("year"))
    n_month = ee.Number(full_date.get("month"))
    return img.set({"n_month": n_month, "n_year": n_year})

def func_reduce2yearmonths_base_period(n_year):
    imgcoll_year_interes = (base_period_tagged
```

```
.filter(ee.Filter.eq("n year", n year)))
10
       def func_reduce2months_base_period(n_month):
11
           return (imgcoll_year_interes
12
                    .filter(ee.Filter.eq("n_month", n_month))
13
                    .set({"n_year": n_year, "n_month": n_month}))
15
       list monthly pr per year = (list months
16
                                     .map(func reduce2months base period))
17
18
       return list monthly pr per year
19
```

- Función para etiquetar año y mes
- ② Función para crear una lista de listas de precipitaciones mensuales para cada año del periodo base

Con estas dos funciones se creará una colección de imagenes de  $\approx$  360 imágenes

```
base period = (chirps
                                                                                    (1)
     .filter(ee.Filter.calendarRange(start = year_base_inicio,
2
                                        end = year base fin, field = "year")))
3
4
   base period tagged = base period.map(func tag year month base period)
                                                                                    (2)
   base_period_tagged_reduced_year_month = (ee.ImageCollection.fromImages(
                                                                                    (3)
       (ee.List
8
                                                                                    (4)
       .sequence(year base inicio, year base fin)
9
        .map(func reduce2yearmonths base period)
                                                                                    (5)
10
       .flatten())))
                                                                                    (6)
11
12
   base_pr_monthly_accumulation = (ee.ImageCollection.fromImages(
                                                                                    7
13
       list months.map(lambda n month: (
                                                                                    (8)
14
                           base_period_tagged_reduced_year_month
                                                                                    (9)
15
                           .filter(ee.Filter.eq("n month", n month))
16
                           .mean()
                                                                                    (10)
17
                           .set({"n_month": n_month}))))
                                                                                    (11)
18
                                                                                    12
     .toBands()
19
     .rename([f"0{i}" if i < 10 else str(i) for i in range(1,13)])
20
```

- 1) Limitar la colección los años del periodo de referencia
- ② Etiquetar el año y mes al que pertenece cada imagen
- 3 Crear una ee.ImageCollection a partir de una lista
- 4 La es una secuencia de números que representan los años del periodo base

- ⑤ A cada elemento (número) se le aplica una función. Esta función regresa una lista de 12 imágenes por año, es decir, el resultado es una lista de 30 elementos, donde cada elemento es una lista de 12 imágenes.
- 6 Se cambia la lista de sublistas a una lista, es decir, se *desempacan* los elementos de las sublistas.
- (7) A partir de una lista se crea una colección
- 8 Esta lista es de 12 elementos, la lista del promedio historico de precipitación por cada mes del periodo base.
- (9) Se filtra por el mes indicado
- (10) Obtener la media de los 30 años
- (11) Marcar como propiedad el año del mes
- (12) De una coleccción de 12 imágenes, se crea una imagen de 12 bandas
- (13) Renombramiento de las bandas (número del mes)

#### 5.2 Anomalía en milimetros

Es la diferencia en milimetros, de la precipitación de un determinado mes  $(\overline{x}_i)$  y el promedio histórico o la normal  $(\mu_{\text{normal}})$  de ese mes

$$anom_{mm} = \overline{x}_i - \mu_{normal}$$

- (1) Restar el promedio histórico
- ② Copiar todas las propiedades en la nueva imagen

## 5.3 Anomalía en porcentaje

Es el resultado de dividir la diferencia de la precipitación de un determinado mes  $(\overline{x}_i)$  y el promedio histórico o la normal  $(\mu_{\text{normal}})$  entre la normal de ese mismo mes.

$$\mathrm{anom}_{\%} = \frac{\overline{x}_i - \mu_{\mathrm{normal}}}{\mu_{\mathrm{normal}}}$$

- Restar el promedio histórico
   Dividir entre el promedio histórico
- ③ Copiar todas las propiedades en la nueva imagen

## De raster a CSV

## 6.1 Información de ee. Image a ee. Feature Collection

Este apartado se hará la demostración con img\_monthly\_pr pero puede ser aplicado a cualquiera de las imágenes que se crearon

Para poder exportar la información como una tabla de CSV, primero se tiene que almacenar o reducir la información a las geometrias de las regiones del país (sean entidades, municipios o cualquier otro tipo de división).

```
img2fc_monthly_pr = (img_monthly_pr
     .reduceRegions(
                                                                                     (1)
2
         collection = fc,
                                                                                     (2)
3
         reducer = ee.Reducer.mean(),
                                                                                     (3)
4
         scale = 5566)
                                                                                     4
5
     .map(lambda feature: (ee.Feature(feature)
                                                                                     (5)
6
                              .set({'n_year': date_year_interes})
                              .setGeometry(None))))
8
   fc_monthly_pr = ee.FeatureCollection(
                                                                                     6
10
     (img2fc monthly pr
11
      .toList(3000)
                                                                                     (7)
12
      .flatten()))
```

- ① Se crea una ee. FeatureCollection a partir de la información de la imagen de 12 bandas
- ② La información que se extraerá vendrá de las geometrías de México (sean entidades, municipios o cualquier otro tipo de división de interés)
- 3 Se extraerá el promedio de la región
- 4 La escala a la que se hará la reducción, debe ser la misma a la que se encuentra la imagen. Esta puede encontrarse en la página de información de la imagen o colección

- (5) Cuando se crea la nueva ee.FeatureCollection, se itera por cada ee.Feature para poder asignar la propiedad (columna) del año de la información. Las 12 bandas se transforman en tambien en columnas, entonces se tiene la información mensual. Finalmente se elimina la geometría asignada porque de esta manera la exporación es más fácil y no demora mucho.
- (6) Se crea una ee.FeatureCollection a partir de una lista de features
- (7) Transformar la ee. Feature a una lista de máximo 3000 elementos
- (8) Se eliminan sublistas (de existir).

## 6.2 Exportar ee.FeatureCollection a CSV

Dentro del editor de código de Earth Engine existe la función para exportar una tabla, pero para el caso de la API de Python se usa la librería de geemap a través de la función ee\_export\_vector\_to\_drive.

```
from geemap import ee_export_vector_to_drive

description_task = f"{select_fc}_monthly_pr_{date_year_interes}"

ee_export_vector_to_drive(
    collection= fc_monthly_pr,
    description= description_task,
    fileFormat= "CSV",
    folder= "pruebas_ee")
```

# Código final

Tras explicar cada aspecto del procesamiento y extracción de los datos se concluye el documento con la función para pasar los datos raster de CHIRPS a un archivo CSV.

La función toma como argumentos:

- 1. El año de interés
- 2. La métrica de interes:
- Precipitación  $\rightarrow$  pr
- Anomalía de precipitación en mm o anomaly\_pr\_mm
- Anomalía de precipitación en porcentaje → anomaly\_pr\_prop
- 3. Tipo de ee.FeatureCollection:
- Entidades  $\rightarrow$  ent
- Municipios  $\rightarrow$  mun
- Cuencas Hidrológicas → ch

```
import ee
from geemap import ee_export_vector_to_drive

try:
    ee.Initialize()
    print("Se ha inicializado correctamente")
except:
    print("Error en la inicialización")
```

- (1) Cargar librerías y funciones necesarias
- 2 Inicializar sesion de Earth Engine

#### 7.1 Funciones escenciales

```
def func_tag_month(img):
    full_date = ee.Date(ee.Number(img.get("system:time_start")))
    n_month = ee.Number(full_date.get("month"))
    return img.set({"n_month": n_month})

def func_tag_year_month_hist_pr(img):
    full_date = ee.Date(ee.Number(img.get("system:time_start")))
    n_year = ee.Number(full_date.get("year"))
    n_month = ee.Number(full_date.get("month"))
    return img.set({"n_month": n_month, "n_year": n_year})
```

- ① Función para taggear únicamente el mes
- ② Función para taggear año y mes. Usada únicamente para la ee.ImageCollection que cubre la normal de 30 años (1981-2010)

# 7.2 Función de extracción, procesamiento y exportación de datos

```
def extract from chirps daily(
                                                                                   (1)
           year = 2024,
           metrica interes = "pr",
3
           tipo fc = 'ent'):
5
       dict fc = dict(
                                                                                   (2)
6
           ent = "projects/ee-unisaacarroyov/assets/GEOM-MX/MX ENT 2022",
           mun = "projects/ee-unisaacarroyov/assets/GEOM-MX/MX MUN 2022")
8
       fc = ee.FeatureCollection(dict fc[tipo fc])
9
10
       geom_mex = (ee.FeatureCollection("USDOS/LSIB/2017")
11
                    .filter(ee.Filter.eq("COUNTRY NA", "Mexico"))
12
                    .first()
13
                    .geometry())
15
       chirps = (ee.ImageCollection('UCSB-CHG/CHIRPS/DAILY')
                                                                                   (3)
16
                  .select("precipitation")
17
                  .filter(ee.Filter.bounds(geom mex)))
18
19
```

```
chirps year = (chirps.filter( # 4>
20
            ee.Filter.calendarRange(start = year, field = "year")))
                                                                                   4
21
22
       chirps year tagged = chirps year.map(func tag month)
                                                                                   (5)
23
24
       list_months = ee.List.sequence(1, 12)
25
26
       def func reduce2months(n month):
27
           return (chirps_year_tagged
28
                    .filter(ee.Filter.eq("n month", n month))
29
30
                    .set({"n month": n month}))
31
32
       list_month_pr = (list_months.map(func_reduce2months))
                                                                                   6
33
34
       imgcoll month pr = ee.ImageCollection.fromImages(list month pr)
35
36
       if year > 2023:
                                                                                   (7)
37
            img_bands = ["01", "02", "03", "04"]
38
       else:
39
            img bands = [f'''] if i < 10 else str(i) for i in range(1,13)
40
       img month pr = imgcoll month pr.toBands().rename(img bands)
                                                                                   (8)
42
43
       if metrica interes != "pr":
                                                                                   (9)
44
           hist pr = (chirps
45
                .filter(ee.Filter.calendarRange(1981, 2010, field = "year")))
46
47
           hist_pr_tagged = hist_pr.map(func_tag_year_month_hist pr)
48
49
           def func reduce2yearmonths_hist_pr(n_year):
50
                imgcoll interes = (hist pr tagged
51
                                      .filter(ee.Filter.eq("n year", n year)))
52
                def func_reduce2months_hist_pr(n_month):
53
                    return (imgcoll interes
54
                             .filter(ee.Filter.eq("n month", n month))
55
56
                             .set({"n_year": n_year, "n_month": n_month}))
57
                list month pr per year = (list months
58
                                              .map(func reduce2months hist pr))
59
                return list_month_pr_per_year
61
           hist_pr_tagged_reduced_year_month = (ee.ImageCollection
62
                .fromImages((ee.List.sequence(1981, 2010)
63
```

```
.map(func reduce2yearmonths hist pr)
                               .flatten())))
65
            img hist pr = (ee.ImageCollection.fromImages(
67
                 list months.map(lambda n month: (
68
                                  hist_pr_tagged_reduced_year_month
69
                                  .filter(ee.Filter.eq("n month", n month))
70
71
                                  .set({"n_month": n_month}))))
72
                 .toBands()
73
                 .rename([f"0{i}" if i < 10 else str(i) for i in range(1,13)])
75
            if metrica interes == "anomaly pr mm":
                                                                                    (10)
76
                 img_metrica_interes = ee.Image(
                     (img month pr
78
                     .subtract(img hist pr.select(img bands))
                     .copyProperties(img_hist_pr, img_hist_pr.propertyNames())))
80
            else:
81
                 img_metrica_interes = ee.Image(
                                                                                    (11)
82
                     (img month pr
                     .subtract(img hist pr.select(img bands))
84
                     .divide(img_hist_pr.select(img_bands))
                     .copyProperties(img hist pr, img hist pr.propertyNames())))
86
        else:
87
            img_metrica_interes = img_month_pr
                                                                                    (12)
88
89
        img2fc_metrica_interes = (img metrica interes
90
            .reduceRegions(collection = fc,
91
                            reducer = ee.Reducer.mean(),
92
                            scale = 5566)
93
            .map(lambda feature: (ee.Feature(feature)
                                    .set({'n year': year})
95
                                    .setGeometry(None))))
96
97
        fc metrica interes = ee.FeatureCollection(
98
            img2fc metrica interes.toList(3000).flatten())
99
100
        # Guardar este pedo
101
        descr task = f"chirps daily {metrica interes} {tipo fc} {year}"
                                                                                    (14)
102
        folder name = f"gee chirps daily {metrica interes}"
103
        print(f"Va al servidor: '{descr_task}' y se gurda en {folder_name}")
105
        ee_export_vector_to_drive(
106
            collection = fc metrica interes,
107
```

```
description= descr_task,

fileFormat= "CSV",

folder= folder_name)

return None
```

- (1) La **precipitación** del **2024** en las **entidades** de México es lo que por *default* se extraerá
- ② Carga de geometrías y ee.FeatureCollections
- (3) Carga de CHIRPS Daily
- (4) Selección del año del cual se obtendrán las metricas
- 5 Etiquetado de los meses a la ee.ImageCollection de interés
- 6 Reducción de una ee. Image<br/>Collection de  $\approx 365$  imágenes a una de (máximo) 12 imágenes.
- Si el año de interés es menor o igual que el 2023, entonces se tiene información de todos los meses (12 bandas), de lo contrario son menos
- (8) Renombramiento de las bandas a el número de los meses del año
- Si la métrica no es la precipitación ('pr'), es decir es anomalía de la precipitación en porcentaje ('anomaly\_pr\_prop') o en milimetros ('anomaly\_pr\_mm'), entonces se hace el cálculo del promedio histórico de la precipitación de cada uno de los meses (img hist pr)
- (10) Identificar si es anomalía de la precipitación en milimetros ('anomaly\_pr\_mm')
- (i) Si no es 'anomaly\_pr\_mm' entonces se hace la división del promedio histórico para el cálculo de la anomalía de la precipitación en porcentaje ('anomaly\_pr\_prop')
- (12) Si la métrica de interés **es la precipitación** ('pr'), entonces no se hace el cálculo del promedio histórico.
- (13) Se crea ee.FeatureCollection de la ee.Image
- (14) Se crean las variables para exportar los datos
- (15) Se exportan los resultados

#### 7.3 Extracción

Con el codigo creado en esta Sección lo único que queda por hacer es iterar o seleccionar el año, división política y tipo de métrica a extraer

```
for anio in range(1981, 2025):
                                                                                   1
       extract from chirps daily(year = anio,
                                                                                   2
3
                                   metrica_interes= "pr",
                                   tipo fc= "ent")
5
6
       extract_from_chirps_daily(year = anio,
                                                                                   (3)
7
                                   metrica interes= "pr",
8
                                   tipo fc= "mun")
9
```

```
10
       extract_from_chirps_daily(year = anio,
                                                                                     4
11
                                    metrica interes= "anomaly_pr_prop",
12
                                    tipo fc= "ent")
13
14
                                                                                    (5)
       extract_from_chirps_daily(year = anio,
15
                                    metrica_interes= "anomaly_pr_prop",
16
                                    tipo fc= "mun")
17
```

- 1 Extraer información de todos los años disponibles
- 2 Precipitación en los estados de México
- 3 Precipitación en los municipios de México
- 4 Anomalía de precipitación en porcentaje con respecto a la normal en los estados de México
- (5) Anomalía de precipitación en porcentaje con respecto a la normal en los municipios de México