

Monografía Visualización

Moisés Alfonso Guerrero Jiménez, Andrés Castaño Licona

07-03-2024

Diseño de solución para predicción de radiación solar en diferentes escalas temporales de la comunidad de Castilla y León, España, para la gestión de proyectos de generación fotovoltaica.

Visualización de datos

Integrantes

- Moisés Alfonso Guerrero Jiménez
- Andrés Castaño Licona

Descripción de los datos

Los datos usados hacen parte de un repositorio abierto con licencia Creative Commons Attribution, que contiene mediciones de radiación solar horizontal global (GHI), recopiladas durante un periodo de 18 años, desde el 1 de enero de 2002 hasta el 31 de diciembre de 2019, con una granularidad de 30 minutos, que fueron medidos en 37 estaciones ubicadas en la región española de Castilla y León.

El conjunto de datos no solo cuenta con datos crudos, sino también con datos refinados y etiquetados por estaciones de medición, contando además con variables meteorológicas y geográficas que complementan el valor de los datos de irradiación.

Los datos fueron publicados por integrantes del Departamento de Ingeniería Topográfica y Cartográfica de la Escuela Técnica Superior de Ingenieros en Topografía, Geodesia y Cartografía, Universidad Politécnica de Madrid, y se encuentran disponibles para su uso con fines de investigación en un repositorio público con nombre “CyL_GHI” disponible en el sitio web Zenodo (<https://zenodo.org/records/7404167>), y además cuentan con una documentación detallada con nombre “CyL-GHI: Global Horizontal Irradiance Dataset Containing 18 Years of Refined Data at 30-Min Granularity from 37 Stations Located in Castile and León (Spain)”, que describe de manera precisa mucho del preprocesamiento, estandarización y limpieza realizado por los autores, así como muchas de las características del dataset.

La lectura de los datos requiere de un paso inicial de descarga del repositorio de Zenodo, para lo cual se hace uso del comando `wget` de linux, que nos permite realizar la descarga directamente al directorio de ejecución en Colab.

De acuerdo a la documentación del dataset, las mediciones fueron realizadas en un total de 37 estaciones meteorológicas, por lo cual es necesario realiza la extracción de los datos por cada estación meteorológica en cada día.

```
## | station_code | name | latitude | longitude | height |
## | :----- | :----- | :----- | :----- | :----- |
```

##		AV01		Nava de Arévalo		40.9688		-4.76794		921	
##		BU02		Valle de Valdelucio		42.7463		-4.13287		976	
##		BU03		Lerma		42.0397		-3.7682		840	
##		BU04		Tardajos		42.3461		-3.80258		822	
##		BU05		Vadocondes		41.6371		-3.5768		811	

```
## | 0 |
## |:-----|
## | object |
## | object |
## | float64 |
## | float64 |
## | int64 |
```

Los datos crudos cuentan con un encabezado con 10 columnas separadas por punto y coma (;), correspondientes a:

- Código: Código único de la estación en la cual se realizaron las mediciones.
- Ubicación: Indica el nombre de la estación de medición de acuerdo a su ubicación.
- Fecha (AAAA-MM-DD): Hace referencia a la fecha en la cual se realizaron las mediciones (UTC).
- Hora (HHMM): Indica la hora en la cual se realizó la medición para la fecha correspondiente (UTC).
- Precipitación (mm): Precipitaciones medidas en milímetros en la ubicación de la estación de medición.
- Temperatura (°C): Temperatura en grados celsius medida en la ubicación de la estación meteorológica.
- Humedad relativa (%): Indica el porcentaje de humedad medido.
- Radiación (W/m2): Medición de la radiación solar global horizontal (GHI).
- Vel. viento (m/s): Velocidad del viento medida en m/s
- Dir. viento (°): Dirección del viento en grados.

Sin embargo, se observan caracteres especiales que sugieren que el archivo debe ser leído con una codificación diferente a UTF-8. De acuerdo a la documentación disponible en <https://docs.python.org/3/library/codecs.html#standard-encodings>, hacemos uso de una codificación latin-1, lo que permite visualizar correctamente los datos al ser leídos.

Tal como se observa a continuación al intentar leer el archivo con la codificación por defecto no es posible leer el archivo.

```
## Cargando archivos para el año 2002...
## Cargando archivos para el año 2003...
## Cargando archivos para el año 2004...
```

```
## Dimensiones del dataframe general creado: (2205519, 11)
```

Obtención de radiación solar incidente teórica

Con el objetivo de complementar el conjunto de datos, se busca mediante el desarrollo de una función obtener la radiación solar incidente teórica para una latitud y longitud específicas.

De acuerdo a lo comentado en “Cálculo de la radiación solar extraterrestre en función de la latitud y la declinación solar” es posible obtener dicho valor mediante ecuaciones que describen la radiación solar teórica extraterrestre.

Por medio del dataframe de datos geográficos `df_geo`, podemos hacer uso de las latitudes y longitudes específicas de cada estación de medición y de esta manera complementar los datos para cada estación meteorológica.

El siguiente ejemplo toma los datos de la estación de medición en Nava de Arévalo (AV01) y a partir de la latitud y longitud, así como la longitud del uso horario, permite obtener valores de GHI teórico.

##	station_code	station_name	date	time	precipitation	air_temp	humidity
##	:	:	:	:	:	:	:
##	AV01	Nava de Arévalo	2002-01-01	0030	0	8.67	81.8
##	AV01	Nava de Arévalo	2002-01-01	0100	0	8.54	82.6
##	AV01	Nava de Arévalo	2002-01-01	0130	0	8.6	82.9
##	AV01	Nava de Arévalo	2002-01-01	0200	0	8.49	83.4
##	AV01	Nava de Arévalo	2002-01-01	0230	0	7.91	85.1

##	station_code	station_name	date	time	precipitation	air_temp	humidity
##	:	:	:	:	:	:	:
##	AV01	Nava de Arévalo	2002-01-01	0030	0	8.67	81.8
##	AV01	Nava de Arévalo	2002-01-01	0100	0	8.54	82.6
##	AV01	Nava de Arévalo	2002-01-01	0130	0	8.6	82.9
##	AV01	Nava de Arévalo	2002-01-01	0200	0	8.49	83.4
##	AV01	Nava de Arévalo	2002-01-01	0230	0	7.91	85.1

