

EUROCLIMA+

Generación de escenarios climáticos en Centroamérica

Curso de capacitación sobre extracción y utilización de datos CORDEX

Georreferenciación, rejillas y NetCDF

Antonio Ángel Serrano de la Torre

25 de febrero al 1 de marzo de 2019

San Salvador



Financiado por la Unión Europea

0 Kilometers
100 300 Miles



Índice

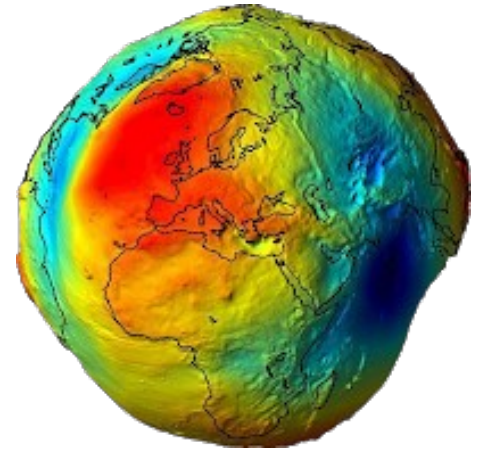
- Georreferenciación:
 - Definición
 - Datum
 - Proyección
- Rejillas:
 - Definición
 - Ejemplos
- El formato NetCDF
 - Definición
 - Ejes espaciales
 - Eje temporal
 - Un vistazo a la estructura interna (cdl)

Georreferenciación

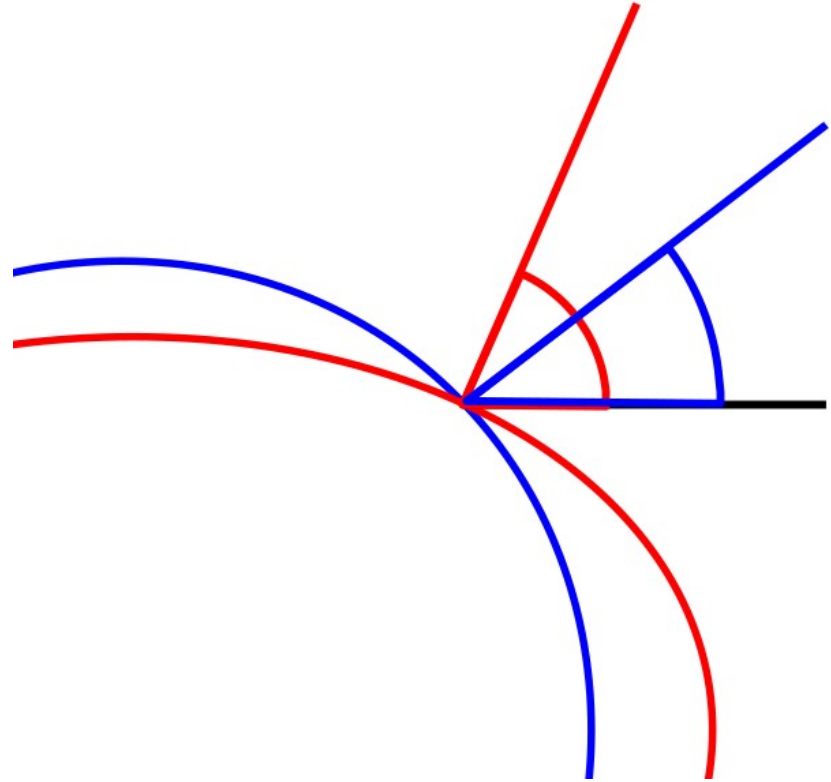
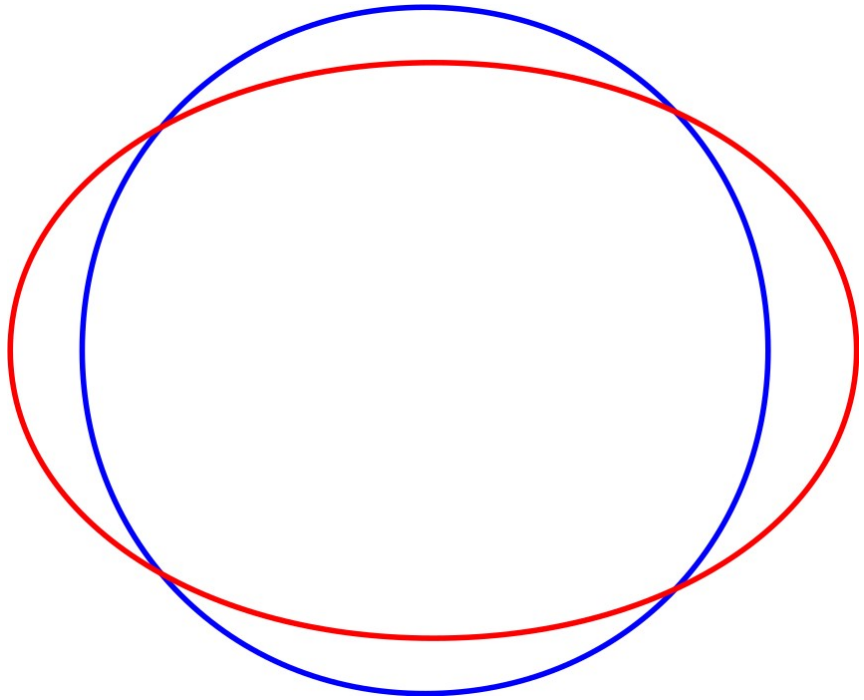
- Técnica que relaciona una posición en la superficie terrestre con una posición en un mapa.
- Puede decirse que cumple dos objetivos:
 - Primer objetivo: asignar longitud y latitud a un punto de la Tierra.
 - Segundo objetivo. Es opcional. Proyecta la esfera terrestre (o una parte de la misma) sobre un plano y asigna las coordenadas como distancias en metros a partir de un origen.

Georreferenciación: asignación de lon y lat (1/4)

- Distintas técnicas de georreferenciación pueden asignar valores de longitud y latitud ligeramente diferentes.
- Los valores asignados dependerán de la forma que consideremos para la Tierra.
- Aunque la forma de la Tierra es irregular, se aproxima mejor mediante un elipsoide que mediante una esfera.



Georreferenciación: asignación de lon y lat (2/4)



Georreferenciación: asignación de lon y lat (3/4)

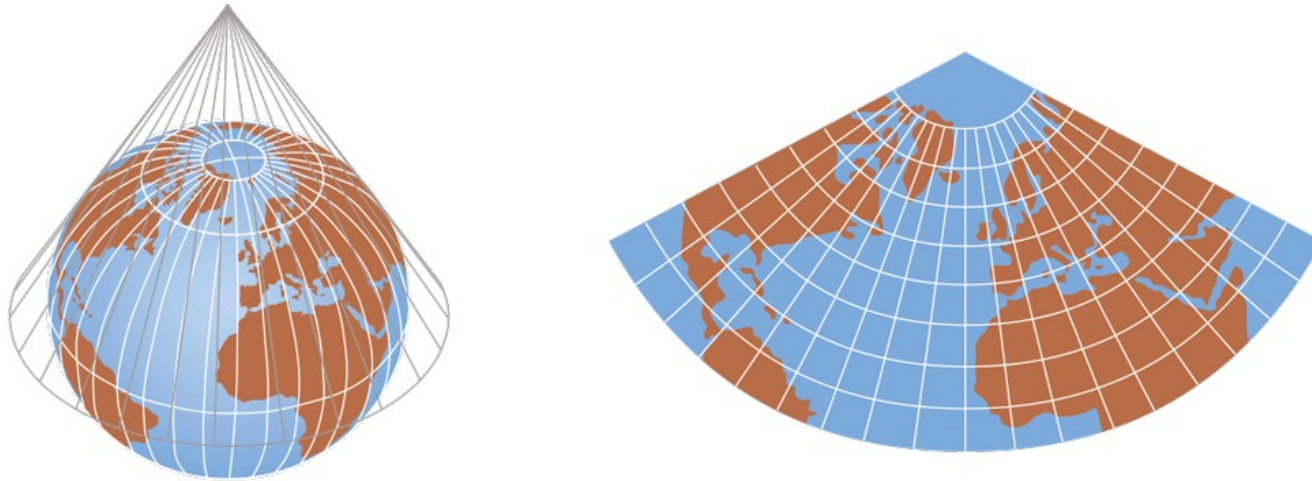
- El elipsoide queda definido mediante sus:
 - Eje mayor
 - Eje menor
- La longitud y latitud de un punto de la Tierra quedan definidas unívocamente al definir los siguientes elementos:
 - Elipsoide de referencia
 - Un punto de contacto del elipsoide con la Tierra
 - Dirección de referencia para definir el Norte
- Estos tres elementos conforman lo que llamamos *datum*.
- Un *datum*: → una (longitud, latitud) de un punto sobre la Tierra

Georreferenciación: asignación de lon y lat (4/4)

- Debido a la forma irregular de la Tierra, en cada zona que se considere se adaptará mejor un elipsoide que otro, un datum que otro.
- Tipos de datums:
 - Locales. Se adaptan a una zona específica de la Tierra. P.ej.:ETSR89, ligado a la parte estable de la placa continental europea.
 - Globales. Tratan de adaptarse a toda la Tierra en conjunto. Ejemplo: WGS84 (*World Geodetic System 1984*). Tanto los modelos numéricos de predicción del clima como los dos modelos de CORDEX para Centroamérica utilizan el WGS84.

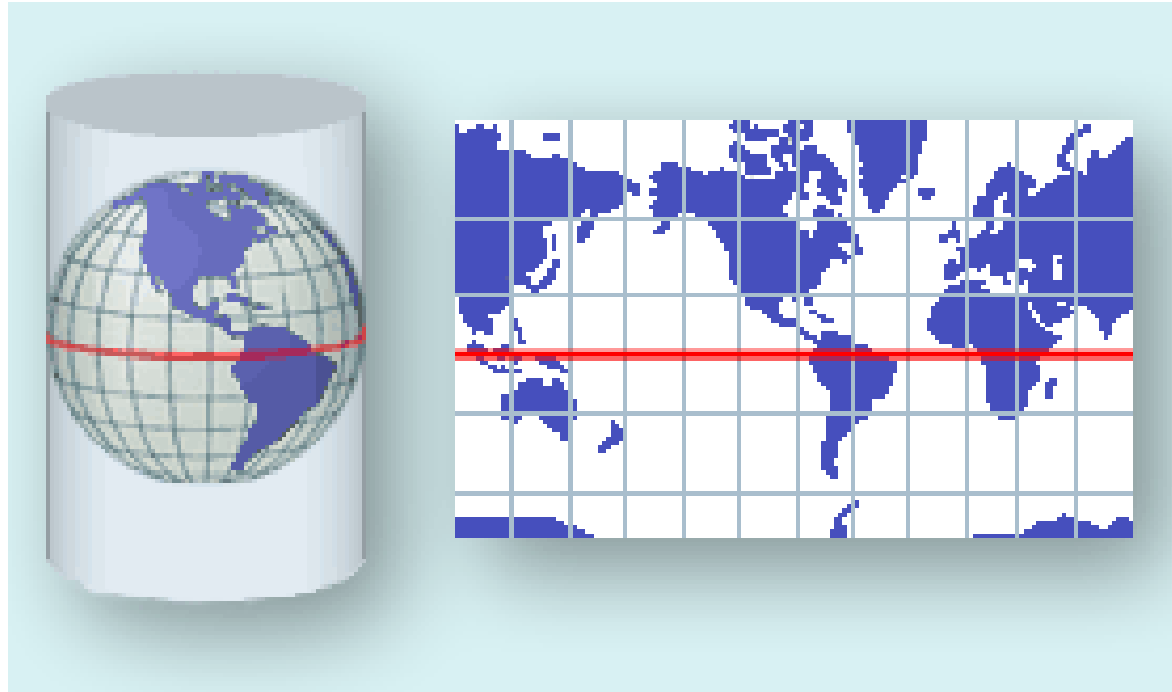
Georreferenciación: proyección

- La superficie de la Tierra se puede proyectar sobre un plano.



Proyección cónica

Georreferenciación: proyección



Proyección cilíndrica (Mercator)

Georreferenciación: proyección

- Existe una línea de contacto entre la Tierra y el plano.
- Existen distorsiones que aumentan a medida que nos alejamos de dicha línea.
- Sin embargo, una proyección se puede construir de modo que conserve alguna de las siguientes propiedades:
 - Conformidad (proyección conforme). Conserva el ángulo entre dos líneas. Es útil para la navegación.
 - Equivalencia (proyección equivalente). Conserva las superficies. Sirve para comparar la extensión de varias zonas sobre el mapa.
 - Equidistancia (proyección equidistante). Conserva la distancia entre dos puntos. Sirve para calcular distancias sobre el mapa.

Georreferenciación: proyección



Proyección conforme (Mercator)

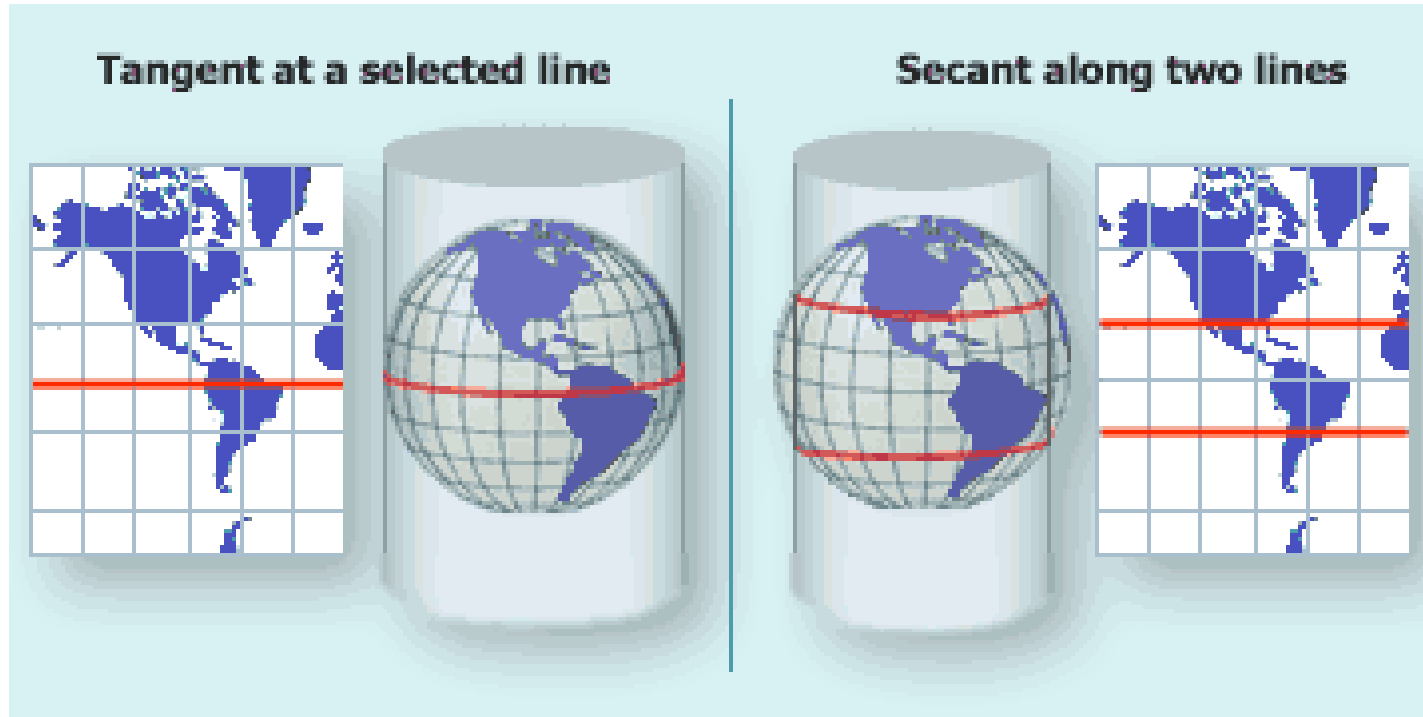


Proyección no conforme
(azimutal equidistante polo norte)

Georreferenciación: proyección

- La proyección introduce una distorsión que aumenta a medida que nos alejamos de la línea de contacto.
- En la proyección de Mercator la distorsión es despreciable hasta $\pm 15^\circ$ alrededor del ecuador (línea de contacto).
- Para minimizar distorsiones, se hace que el plano, en vez de ser tangente, sea secante a la superficie de la Tierra.
- Así, en la proyección cónica de Mercator, podemos hacer que el cilindro corte a la Tierra en los paralelos $\pm 10^\circ$ y la distorsión será despreciable hasta $\pm 25^\circ$ alrededor del ecuador.

Georreferenciación: proyección



Proyección de Mercator
tangente

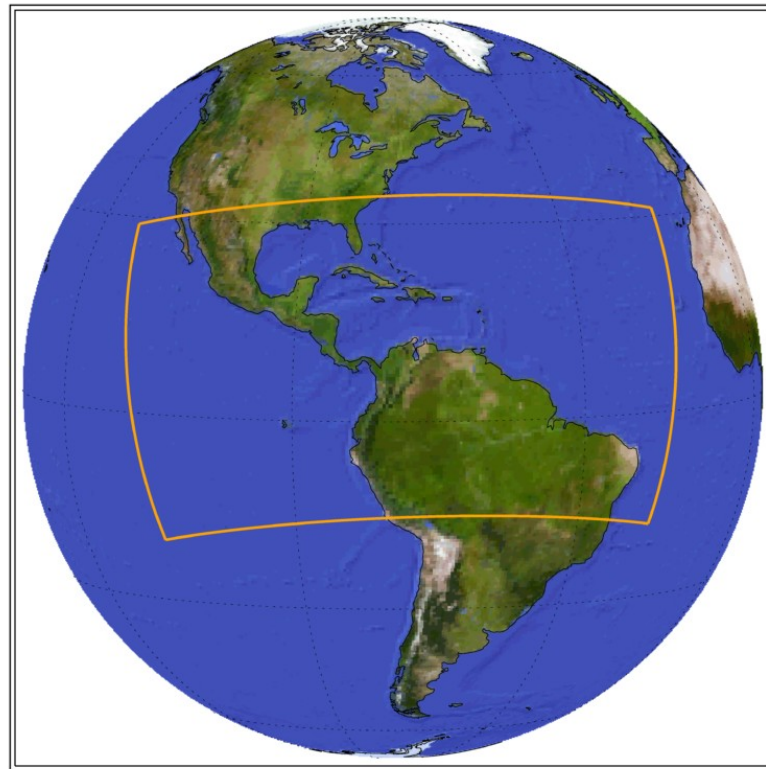
Proyección de Mercator
secante

Georreferenciación: datos CORDEX

- El proyecto CORDEX recoge regionalizaciones dinámicas realizadas por dos centros con sus respectivos modelos:
 - SMHI. (*Swedish Meteorological and Hydrological Institute*) Centro sueco:
 - Datum WGS84.
 - Coordenadas rotadas. Las coordenadas se suelen rotar para situar el ecuador en el centro de la zona en estudio.
 - No realiza ninguna proyección.
 - **Proporciona, para cada punto, la coordenada rotata y la desrotada.**
 - ICTP. (*The Abdus Salam International Centre for Theoretical Physics*) Centro internacional con sede en Trieste, Italia.
 - Datum WGS84.
 - Proyección de Mercator secante en los paralelos $\pm 10^\circ$.
 - Las coordenadas se dan en metros hacia el este y hacia el norte desde el llamado *origen falso*, que se establece como el punto que está a 25 km al oeste y 25 km al sur del punto 10°N , 80°W .
 - **También proporciona la longitud y latitud de cada punto.**

Georreferenciación: datos CORDEX

Dominio CORDEX
para Centroamérica
(dominio CAM)

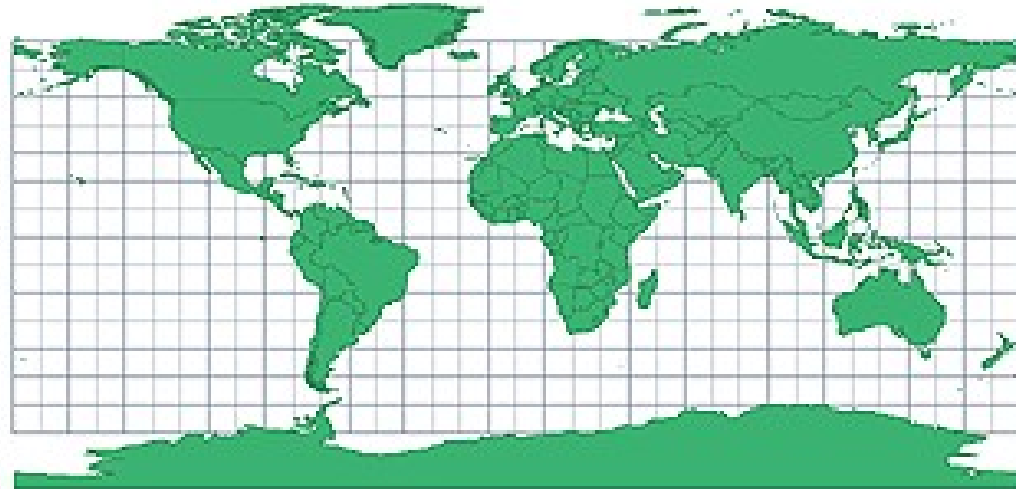


Tipos de rejilla: definición

- Rejilla = malla = *grid* (en inglés)
- Suele designar a:
 - Un tipo de disposición espacial de los datos
 - Un conjunto mínimo de datos organizados según esa disposición

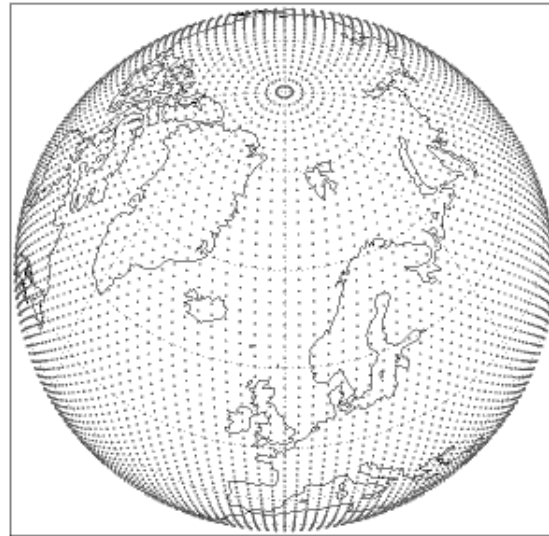
Tipos de rejilla: ejemplos

- Rejilla regular en longitud/latitud.
 - Lejos de los polos: puntos *casi* equiespaciados.
 - Cerca de los polos: puntos no equiespaciados → rotación de coordenadas.



Tipos de rejilla: ejemplos

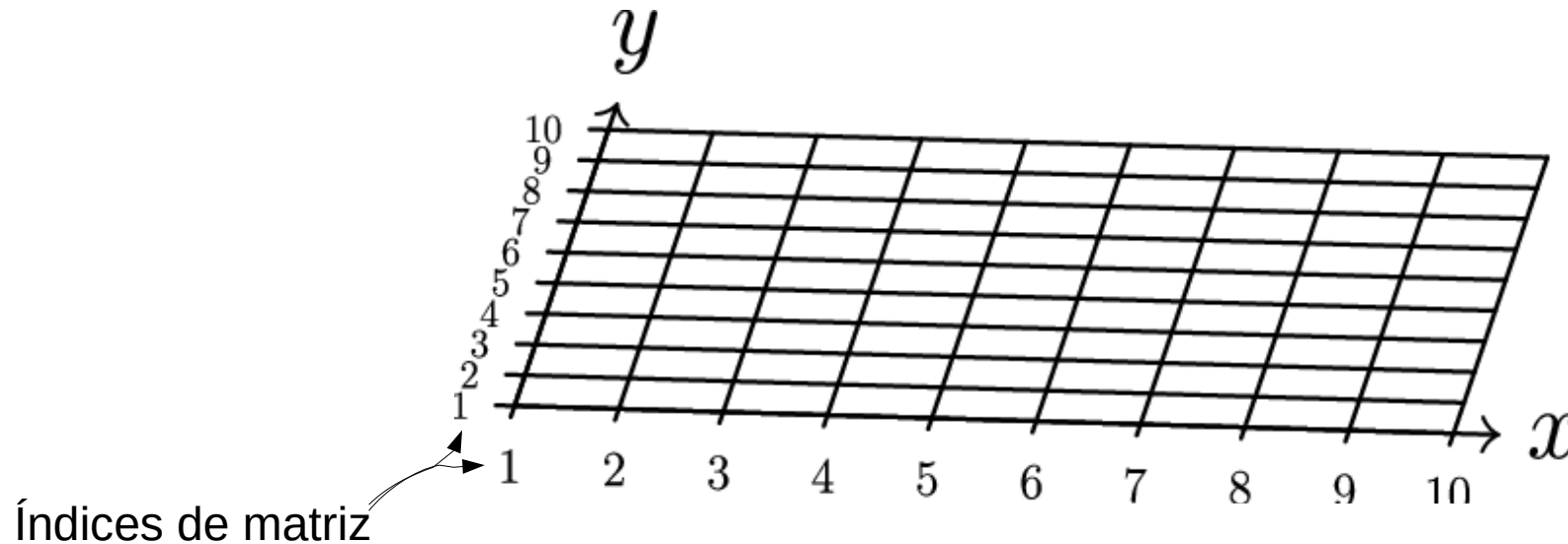
- Rejilla Gaussiana reducida octaédrica (Octahedral reduced Gaussian grids).
 - Densidad de puntos *más* uniforme que la regular lat/lon.



El formato NetCDF

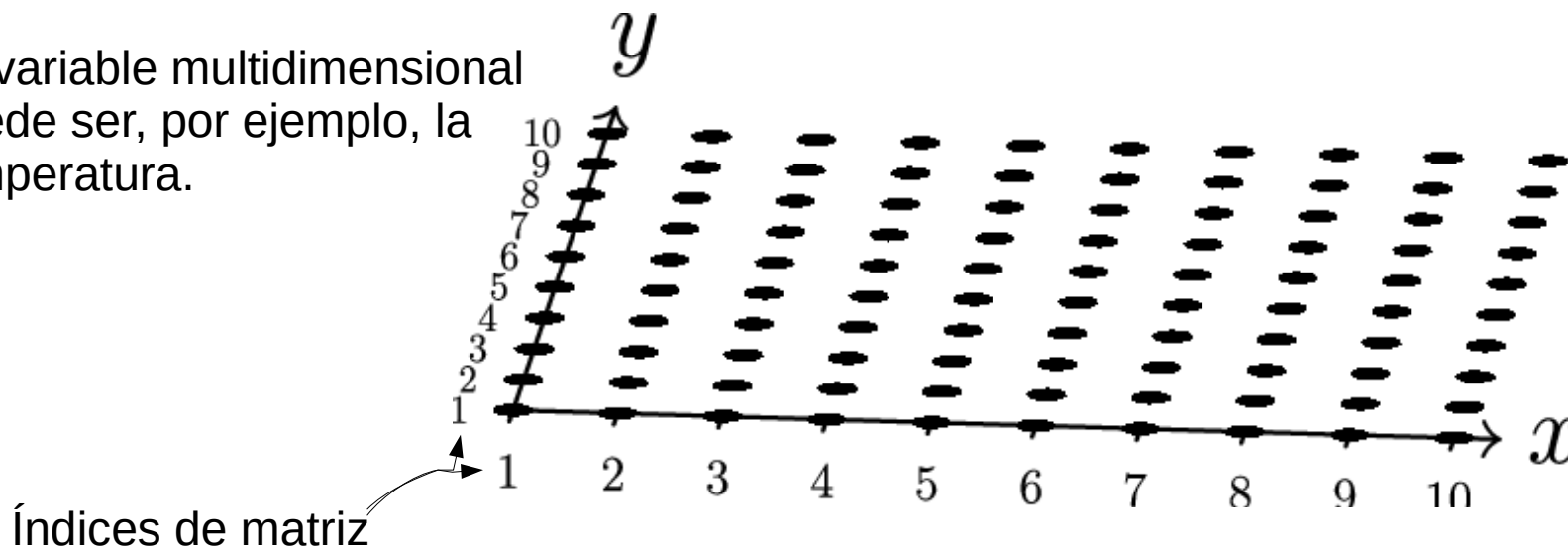
- Creado por la Corporación Universitaria para la Investigación Atmosférica de EE.UU. (*University Corporation for Atmospheric Research, UCAR*).
- Propósito: intercambio de datos científicos que puedan organizarse mediante ejes de coordenadas espaciales y temporales.

El formato NetCDF: ejes espaciales



El formato NetCDF: ejes espaciales

La variable multidimensional puede ser, por ejemplo, la temperatura.

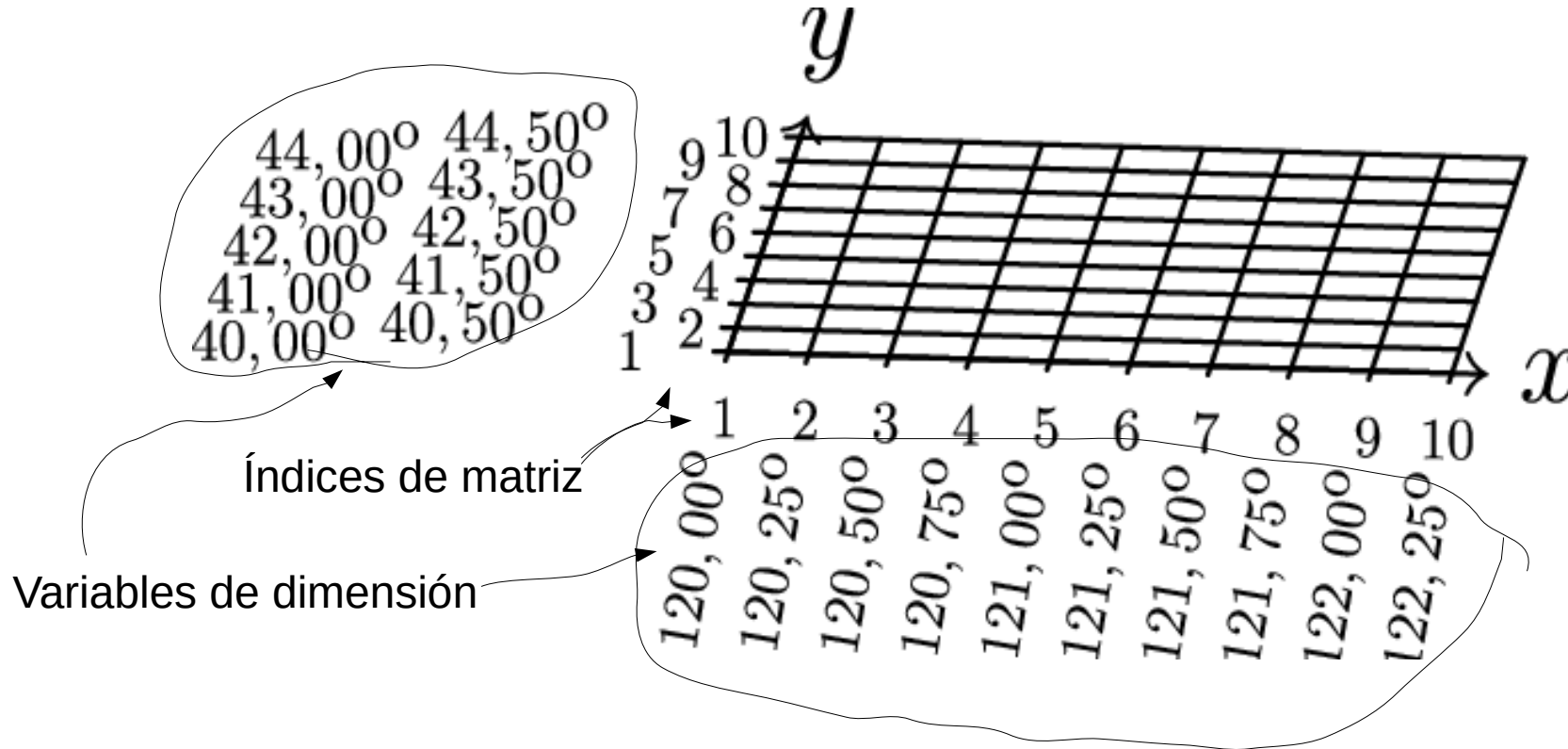


El formato NetCDF: ejes espaciales

- Los índices de la matriz no tienen significado físico
- A cada dimensión de la matriz se puede asociar una variable (en forma de vector) con significado físico. Este tipo de variables se llaman *variables de coordenadas* (*coordinate variables*, en inglés). *
- La asociación se hace llamando a dichas variables (vectores) por el mismo nombre con el que llamamos a la dimensión correspondiente.
- Así, a cada índice del eje X se puede asignar un valor de un vector de longitudes, y a cada índice del eje Y se puede asignar un valor de un vector de latitudes.
- Los ejes X e Y son las dimensiones, y hay que darles un nombre. Por ejemplo, longitud y latitud.

* https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_data_set_components.html#coordinate_variables

El formato NetCDF: ejes espaciales



El formato NetCDF: ejes espaciales

- En una malla regular en longitud/latitud, la longitud (latitud) es una matriz unidimensional (vector). La longitud y latitud de cada punto es un elemento del producto cartesiano de las longitudes y latitudes.
- En una malla sin estructura, la longitud (latitud) es una matriz bidimensional. La longitud (latitud) de un punto cuyos índices de matriz son “i” y “j”, es la longitud que tiene como índices de matriz, el mismo “i” y “j”.
- Existen muchos más tipos de mallas.

El formato NetCDF: ejes espaciales

Resumiendo: Un NetCDF (muy simple) consta de:

- Dimensiones. Por ejemplo: longitud, latitud
- Variables de coordenadas. Por ejemplo, longitud y latitud (mismo nombre que la dimensión correspondiente)
- La variable multidimensional que se representa. Por ejemplo, precipitación, temperatura, etc.

El formato NetCDF: ejes espaciales

Para cada variable se definen:

- Nombre estándar.
- Unidades. Deben ser aceptadas por el paquete *udunits**.
- Las dimensiones de las que depende. Si por ejemplo hemos definido las dimensiones longitud y la latitud, la temperatura podría definirse como “temp(longitud, latitud)”.

Opcionalmente se puede añadir más información:

- “_FillValue” Valor faltante. Los valores iguales a este, se considerarán como faltantes (*missing* en inglés).
- “missing_value”. Opcional. Valor faltante. Debe establecerse con el mismo valor que “_FillValue”. Sólo tiene valor informativo, el que realmente se usa es “_FillValue”.
- Nombre largo
- ...Atributos...

* <https://www.unidata.ucar.edu/software/udunits/>

El formato NetCDF: ejes espaciales

Aunque podemos usar cualesquiera valores para nombre, etc. conviene seguir las convenciones CF (<http://cfconventions.org/cf-conventions/cf-conventions.html>). Concretamente, los nombres (estándar) de las variables y sus unidades deberían ser los recogidos en el listado <http://cfconventions.org/Data/cf-standard-names/63/build/cf-standard-name-table.html>

Por tanto, las convenciones a seguir al definir una variable, son:

- Nombre estándar. Debe ser un nombre de la lista.
- Unidades. Deben usarse las unidades que se dan en la lista para esa variable, u otras que puedan ser convertidas a dichas unidades por la librería *udunits*.
- Nombre largo. Puede ser un nombre en cualquier idioma (no está restringido por las convenciones CF).
- ...Atributos... Los que queramos (no restringido por las convenciones CF).

El formato NetCDF: ejes espaciales

Ejemplo para la precipitación acumulada:

- Nombre estándar: "precipitation_amount"
- Unidades: "kg m-2"
- Nombre largo: "Precipitación acumulada" (el que queramos)
- Definición (especificando sus dimensiones):
"precipitation_amount(longitude, latitude)"
- Nótese que las variables de coordenadas también aparecen en el listado como *longitude* y *latitude*. Por tanto, las dimensiones correspondientes tienen que tener el mismo nombre.

El formato NetCDF: ejes espaciales

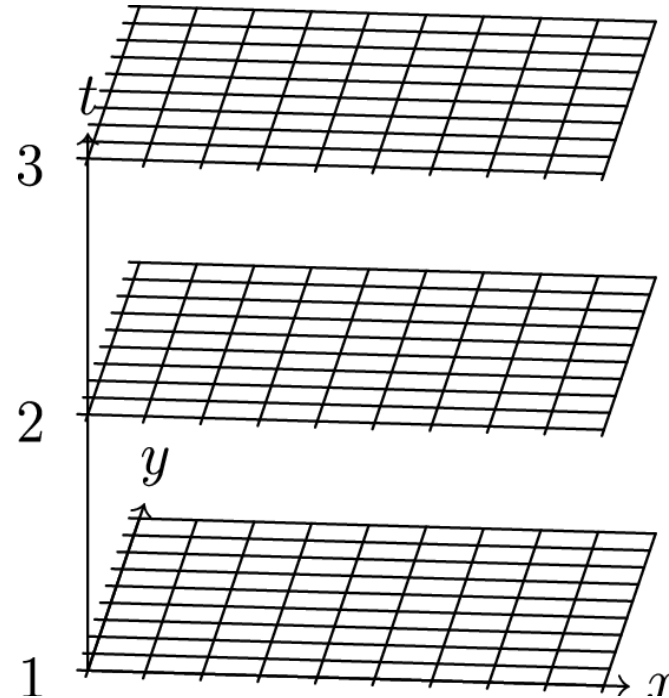
Ejemplo para longitud:

- Nombre estándar: "longitude"
- Unidades: "degree_east"
- Nombre largo: "Longitud (positiva hacia el este)" (el que queramos)
- Definición (especificando sus dimensiones):
"longitude(longitude)"

El formato NetCDF: eje temporal

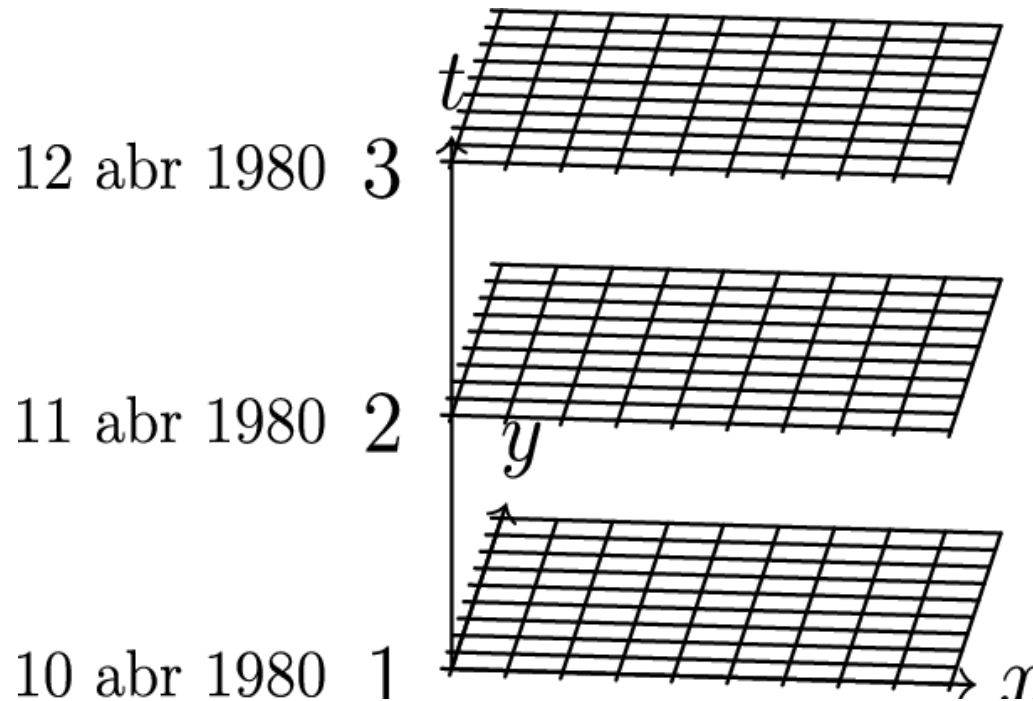
La estructura espacial se repite en el tiempo

Resulta una matriz tridimensional.



El formato NetCDF: eje temporal

A la dimensión temporal también se le asigna una variable de coordenada.

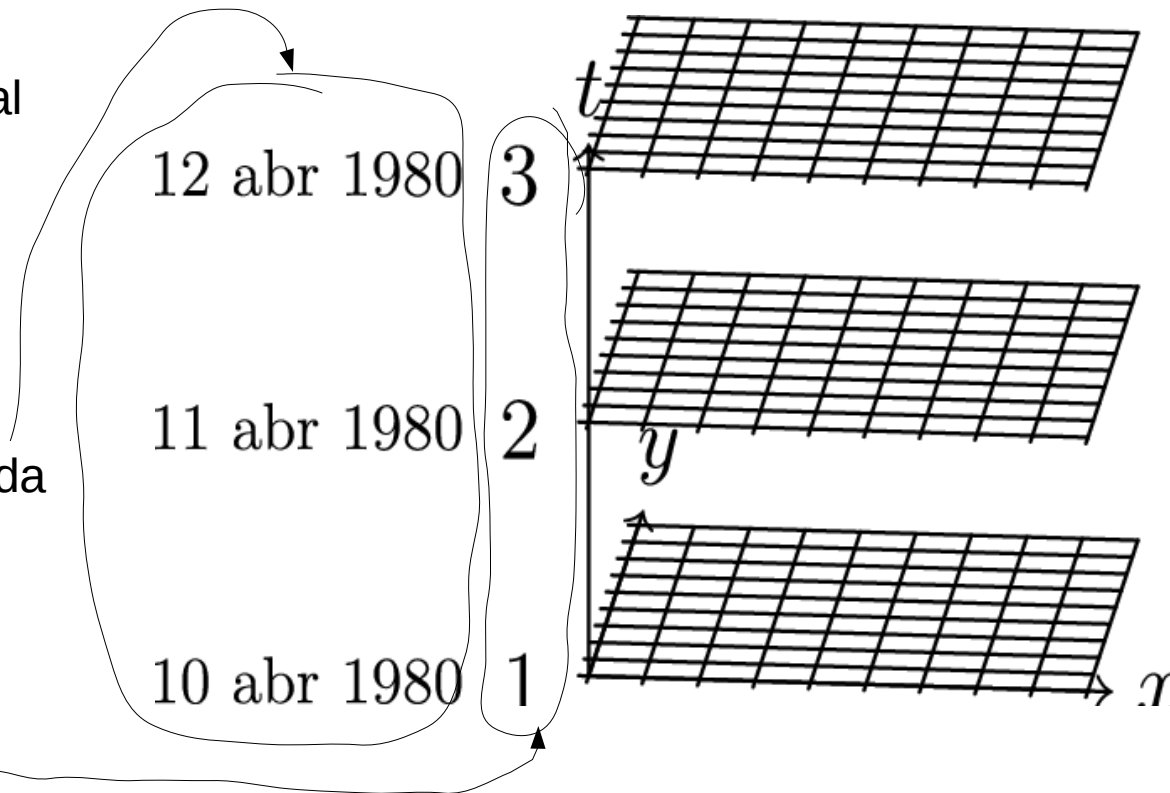


El formato NetCDF: eje temporal

A la dimensión temporal también se le asigna una variable de coordenada.

Variable de coordenada

Índice de la matriz



El formato NetCDF: eje temporal

- Las unidades del eje temporal son las admitidas por el paquete *udunits*, aunque su interpretación dependerá del tipo de alendario.
- Por ejemplo: “days since 1949-12-01 00:00:00”.
De este modo, si el eje temporal tiene los valores “0, 1, 2”, se interpretarán como las fechas “1949-12-01”, “1949-12-02”, “1949-12-03”.
- También hay que definir el tipo de calendario.

El formato NetCDF: eje temporal

Los modelos numéricos suelen simplificar su coordenada temporal adoptando distintas simplificaciones del calendario. Por ejemplo:

- *gregorian* o *standard*. Si no especificamos ninguno, se asume éste. Es el calendario normal, con un salto del 4 al 15 de octubre de 1582. No es una simplificación.
- *proleptic_gregorian*. Como si el gregoriano hubiese regido antes del 15 de octubre de 1582.
- *noleap* o *365_day*. Como el gregoriano pero sin días bisiestos.
- *360_day*. Todos los meses tienen 30 días.
- Más tipos. Ver convenciones CF.

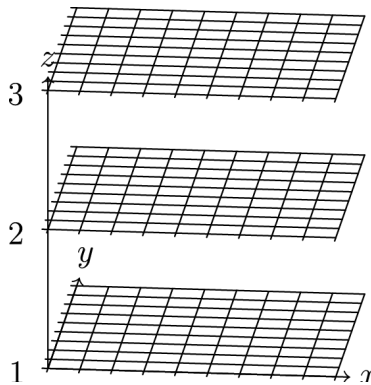
El formato NetCDF: eje temporal

Ejemplos

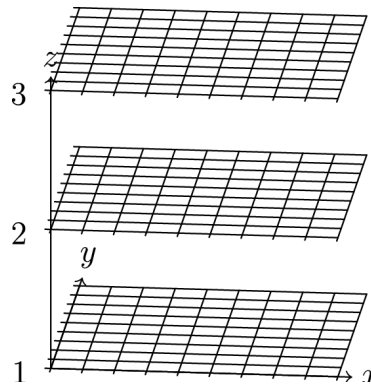
- Calendar = "proleptic_gregorian"
units = "days since 2018-12-29"
Valores de la variable tiempo: 0, 1, 2, 3
Se corresponden con las fechas: "2018-12-29", "2018-12-30", "2018-12-31",
"2019-01-01"
- Calendar = "360_day" ← Sólo cambiamos esto
units = "days since 2019-12-29" ← Igual que antes
Valores de la variable tiempo: 0, 1, 2, 3 ← Igual que antes
Se corresponden con las fechas: "2018-12-29", "2018-12-30", "2019-01-01",
"2019-01-02"

El formato NetCDF: más ejes

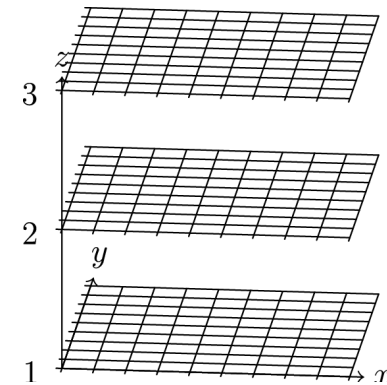
- Hay más posibilidades para los ejes. Por ejemplo, se puede añadir un eje para el nivel (o altura), y cada paso de tiempo contendría una matriz tridimensional, lo que en conjunto hace una matriz tetradimensional.



t = 1



t = 2



t = 3

El formato NetCDF: estructura interna (cdl)

- Un fichero NetCDF se puede especificar en el lenguaje llamado cdl (*Common Data form Language*).
- Podemos hacerlo en un fichero de texto que luego convertiremos a NetCDF mediante la utilidad *ncgen*.
- Se divide en tres secciones:
 - dimensiones: nombre y longitud
 - variables: tipos, nombres, dimensiones de las que depende, y atributos
 - data: los valores que toma la variable
- ";" al final de cada sentencia
- "/" para comenzar un comentario
- Las variables se asocian con un atributo mediante ":"
- Los tipos de los atributos se pueden indicar de modo implícito
- Orden de los datos: el último índice de la variable es el que varía más rápidamente (ordenación por fila).

El formato NetCDF: estructura interna (cdl)

- NetCDF es un formato muy versátil, de modo que puedo poner, por ejemplo, las unidades que yo quiera.
- Sin embargo, para poner orden en lo que se hace, se han confeccionado una convenciones: las convenciones CF (o CF conventions: conventins for Climate and Forecast): <http://cfconventions.org/>
- Las convenciones generales están en: Quick Links → CF Conventions Document → Single HTML
- Los nombres de variables y sus unidades están en: Quick Links → CF Standard Name Table → HTML
- Las unidades soportadas por las *udunits* se pueden consultar en: <https://www.unidata.ucar.edu/software/udunits/udunits-current/doc/udunits/udunits2.html#Database> → Non-SI units

El formato NetCDF: estructura interna (cdl)

- Las especificaciones más básicas de las convenciones CF se refieren a tres atributos que debe tener cada variable:
 - *standard_name*: Debe seguir las convenciones CF en **CF Standard Name Table**.
 - *units*: Deben ser las unidades *canónicas* correspondientes a la variable, u otras que puedan ser convertidas a dichas unidades mediante la librería *udunits*. Las unidades canónicas están también en **CF Standard Name Table**.
 - *_FillValue*: contiene el valor asignado al missing.
- También se puede poner un atributo *missing_value*, pero el que realmente se usa es *_FillValue*.

El formato NetCDF: estructura interna (cdl)

```
netcdf NomFichNC { //ejemplo
```

dimensions:

```
lat = 4 ;
```

```
lon = 3 ;
```

```
time = UNLIMITED ; // currently 2
```

variables:

```
float lat(lat) ;
```

```
lat:units = "degree_north" ;
```

```
float lon(lon) ;
```

```
lon:units = "degree_east" ;
```


El formato NetCDF: estructura interna (cdl)

```
double time(time) ;  
    time:units = "days since 2009-01-01" ;
```

```
float temp(time, lat, lon) ;  
    temp:standard_name = "surface_temperature" ;  
    temp:_FillValue = 1.e+20f ;  
    temp:missing_value = 1.e+20f ;  
    temp:units = "deg_K" ;  
    temp:long_name = "Temperatura a 2m sobre la superficie" ;  
    temp:comentario = "Contiene sesgos positivos"  
:descripcion="Estos datos son un ejemplo"; // Atributo global
```

El formato NetCDF: estructura interna (cdl)

data:

```
lat = 40.0, 40.5, 50.0, 50.5 ;
```

```
lon = -30.0, -29.0, -28.0 ;
```

```
time = 0, 1 ;
```

```
temp =
```

```
297.3, 298.6, 295.7, __, 297.17, 297.3, 298.6, 295.7,  
293.38, 297.17, 297.3, 298.6, 295.7, 293.38, 297.17, 297.3,  
298.6, 295.7, 293.38, 297.17, 295.7, 293.38, 297.17, 297.3 ;
```

```
} // Cierre del fichero
```

Bibliografía

- Map Projections: A Working Manual; 1987; PP; 1395; Snyder, John P.
<https://doi.org/10.3133/pp1395>
- NetCDF Climate and Forecast (CF) Metadata Conventions:
<http://cfconventions.org/>
- CF Standard Name Table (listado estandarizado de nombres de variables con sus unidades):
<http://cfconventions.org/Data/cf-standard-names/63/build/cf-standard-name-table.html>
- Paquete *udunits*: <https://www.unidata.ucar.edu/software/udunits/>
- Lenguaje *cfd*:
https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_utilities_guide.html