



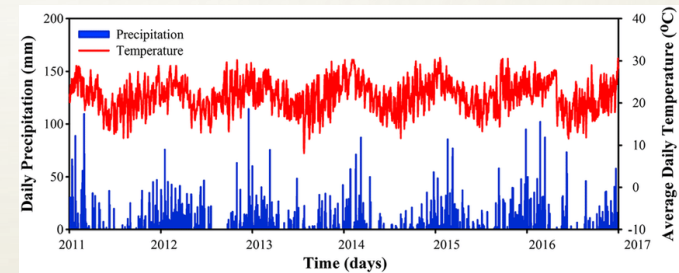
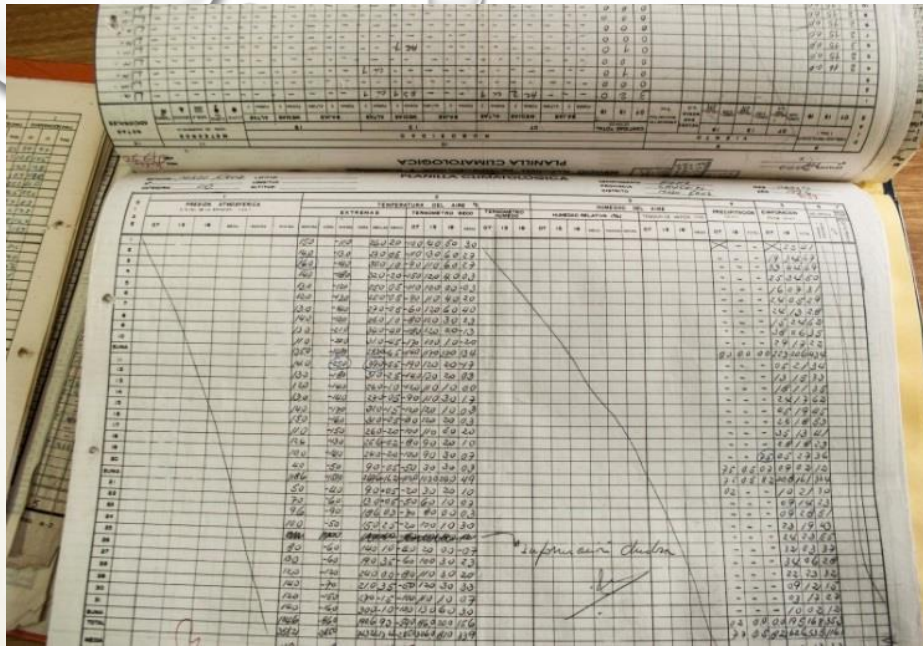
TECNICAS DE PRONOSTICOS

M.Sc. Christian W. Barreto Schuler



CONTENIDO

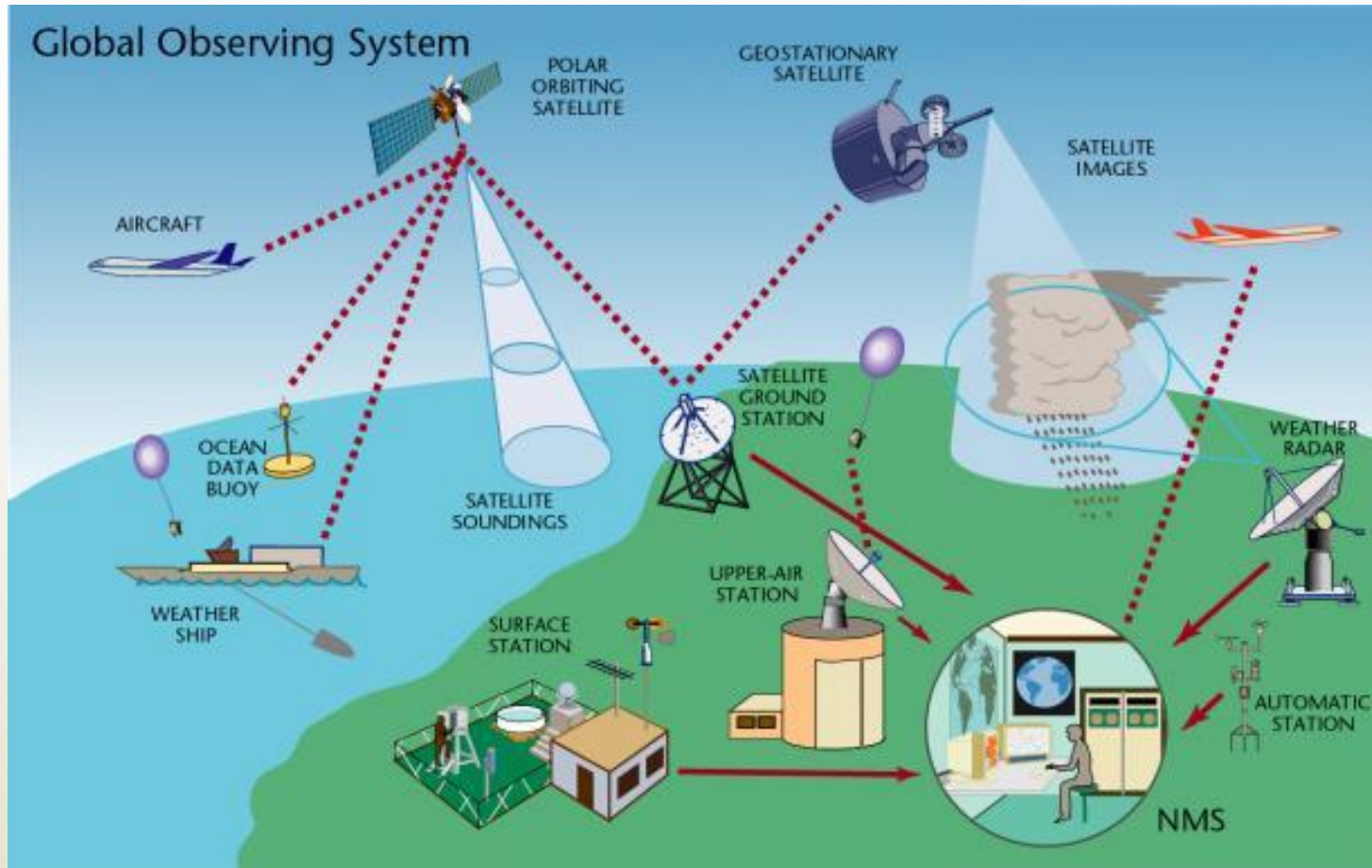
IV. ANALISIS EXPLORATORIO DE DATOS, DISTRIBUCIONES EMPIRICAS Y PARAMETRICAS



Ogashawara et al., 2019*

*<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019GH000186>

BACKGROUND



www.ncdc.noaa.gov,
www.ecmwf.int/en/forecasts/datasets,
www.data.gov, y
www.data.gov.uk.

BACKGROUND

Robustez y resistencia

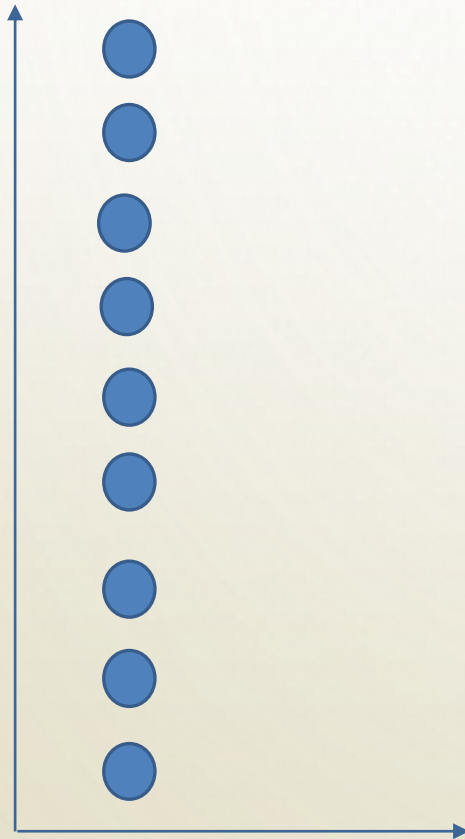
$\{11, 12, 13, 14, 15, 16, 17, 18, 19\}$

$\{11, 12, 13, 14, 15, 16, 17, 18, 91\}$

CUANTILES

Cuantiles y fractiles son esencialmente equivalentes a su termino mas familiar, percentiles

La determinación de cuantiles de muestra requiere que un lote de datos se organice primero en orden



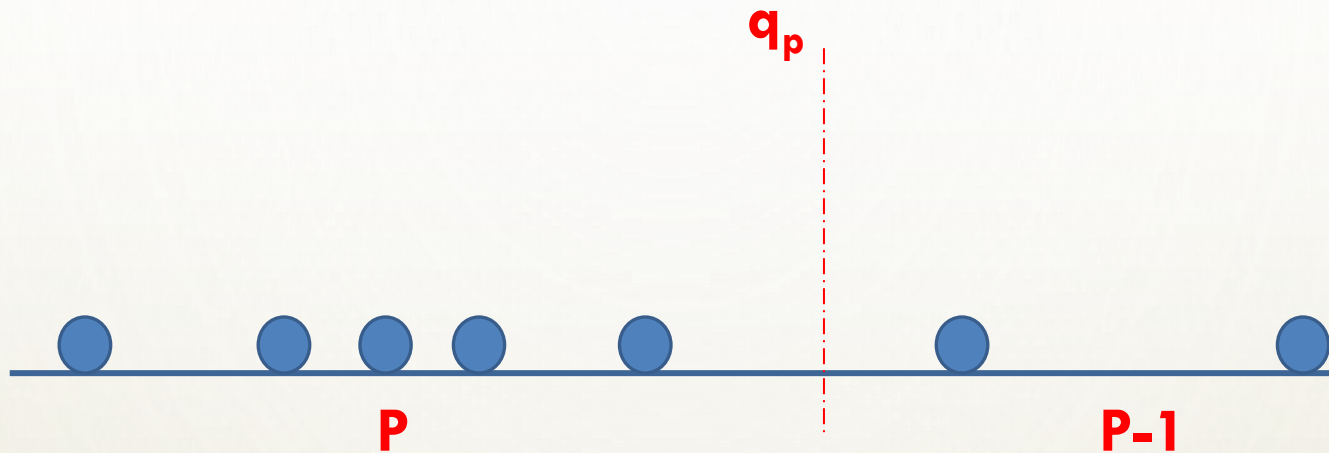
$$q_{0.5} = \begin{cases} x_{([n+1]/2)} & , n \text{ odd} \\ \frac{x_{(n/2)} + x_{([n/2]+1)}}{2} & , n \text{ even} \end{cases}$$

{-2 0 3 -8 4 8 1}

{-8 -2 0 1 3 4 8}

CUANTILES

Percentiles, en cierto sentido, dan una idea de probabilidad



MEDIDAS NUMÉRICAS DE RESUMEN

Ubicación: se refiere la *tendencia central* o la magnitud general de los valores en la data

La media

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

La mediana $q_{0.5}$

Trimean

$$\text{Trimean} = \frac{q_{0.25} + 2q_{0.5} + q_{0.75}}{4}.$$

Otros: media geométrica, media ponderada

MEDIDAS NUMÉRICAS DE RESUMEN

Dispersión: se refiere a qué tan ampliamente se distribuyen los valores de la data alrededor del valor central

Rango intercuartil (IQR)

$$IQR = q_{0.75} - q_{0.25}.$$

Es la medida mas común, simple, robusta y resistente de las medidas de dispersión

Varianza (S^2)

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Desviación estándar (S)

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

Coeficiente de variación (C_v)

$$C_v = \frac{S}{\bar{x}}$$

Desviación mediana absoluta

$$MAD = \text{median} |x_i - q_{0.5}|.$$

MEDIDAS NUMÉRICAS DE RESUMEN

Simetría: Simetría con respecto a la media.

Coeficiente de simetría de la muestra
(*sample skewness coefficient*)

$$\gamma = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^3}{s^3}.$$

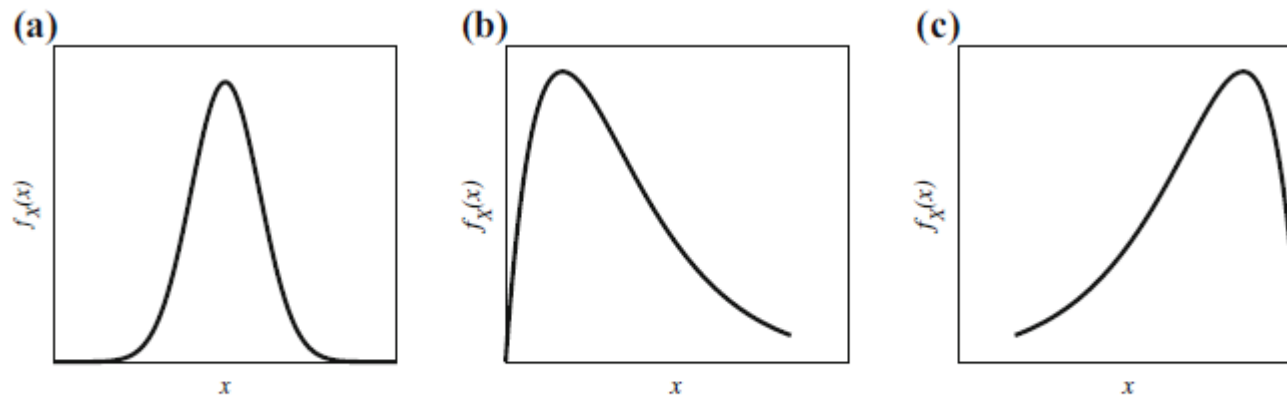


Fig. 3.3 Typical *pdf* plots of a **a** symmetric, **b** positively skewed distribution, and **c** negatively skewed distribution

Yule-Kendall index

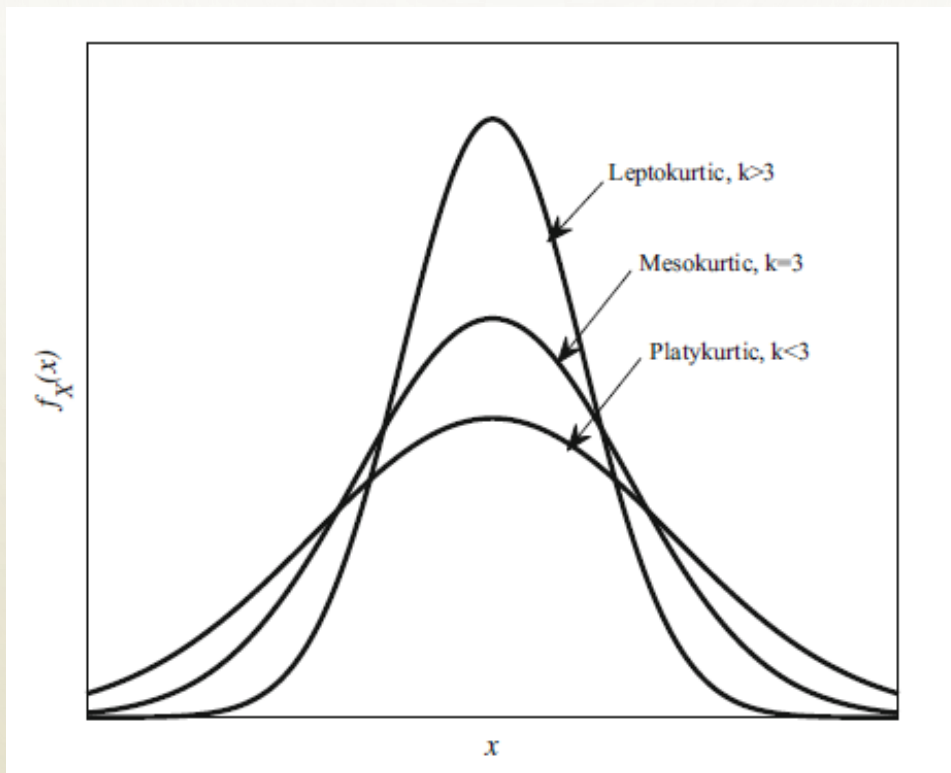
$$\gamma_{YK} = \frac{(q_{0.75} - q_{0.5}) - (q_{0.5} - q_{0.25})}{\text{IQR}} = \frac{q_{0.25} - 2q_{0.5} + q_{0.75}}{\text{IQR}},$$

MEDIDAS NUMÉRICAS DE RESUMEN

Curtosis: medida de los pesos de las colas superior e inferior de una distribución en relación con el centro de distribución

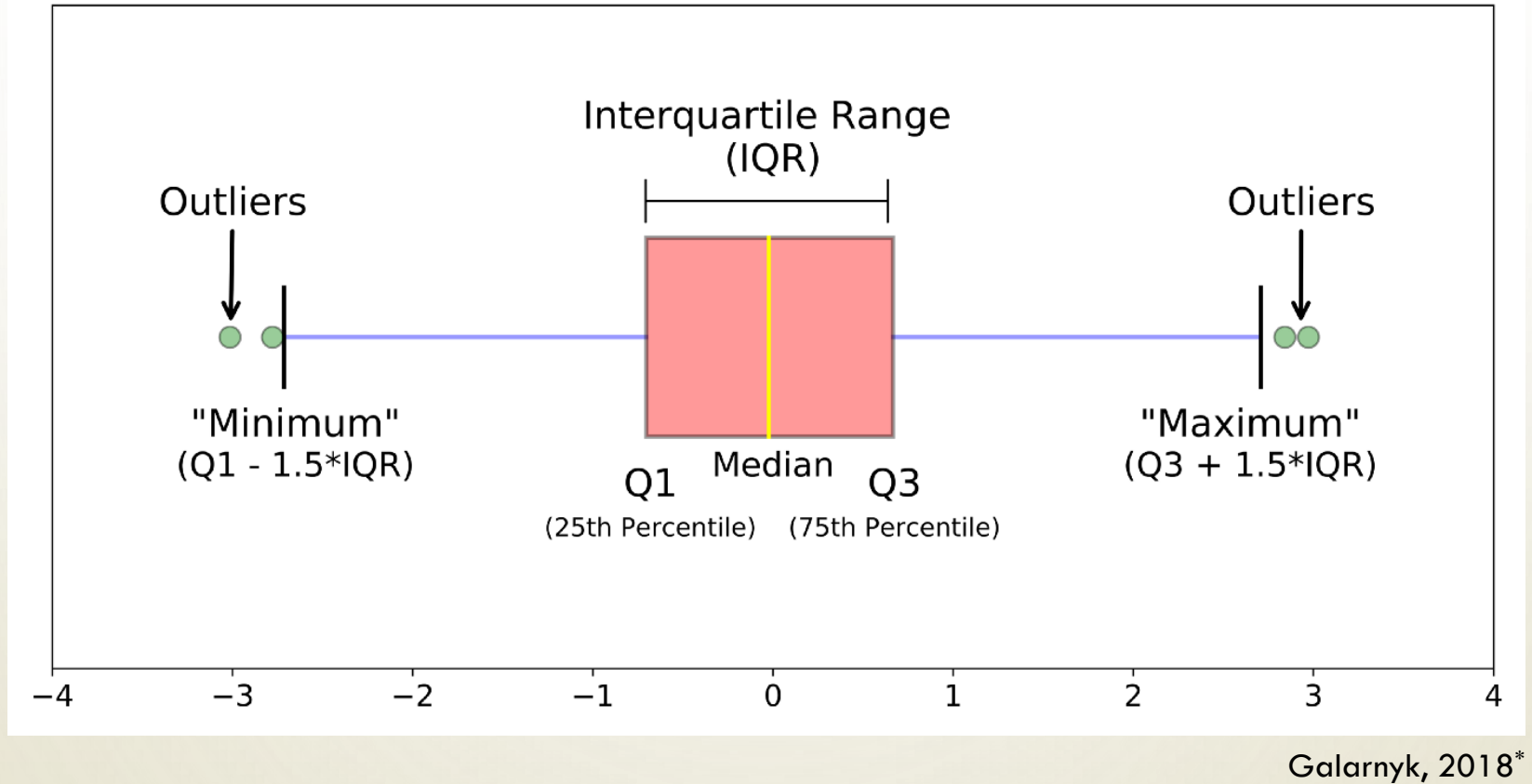
Coeficiente de curtosis

$$\kappa = \frac{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^4}{s^4} - 3.$$



VISUALIZACION DE DATOS

Boxplot

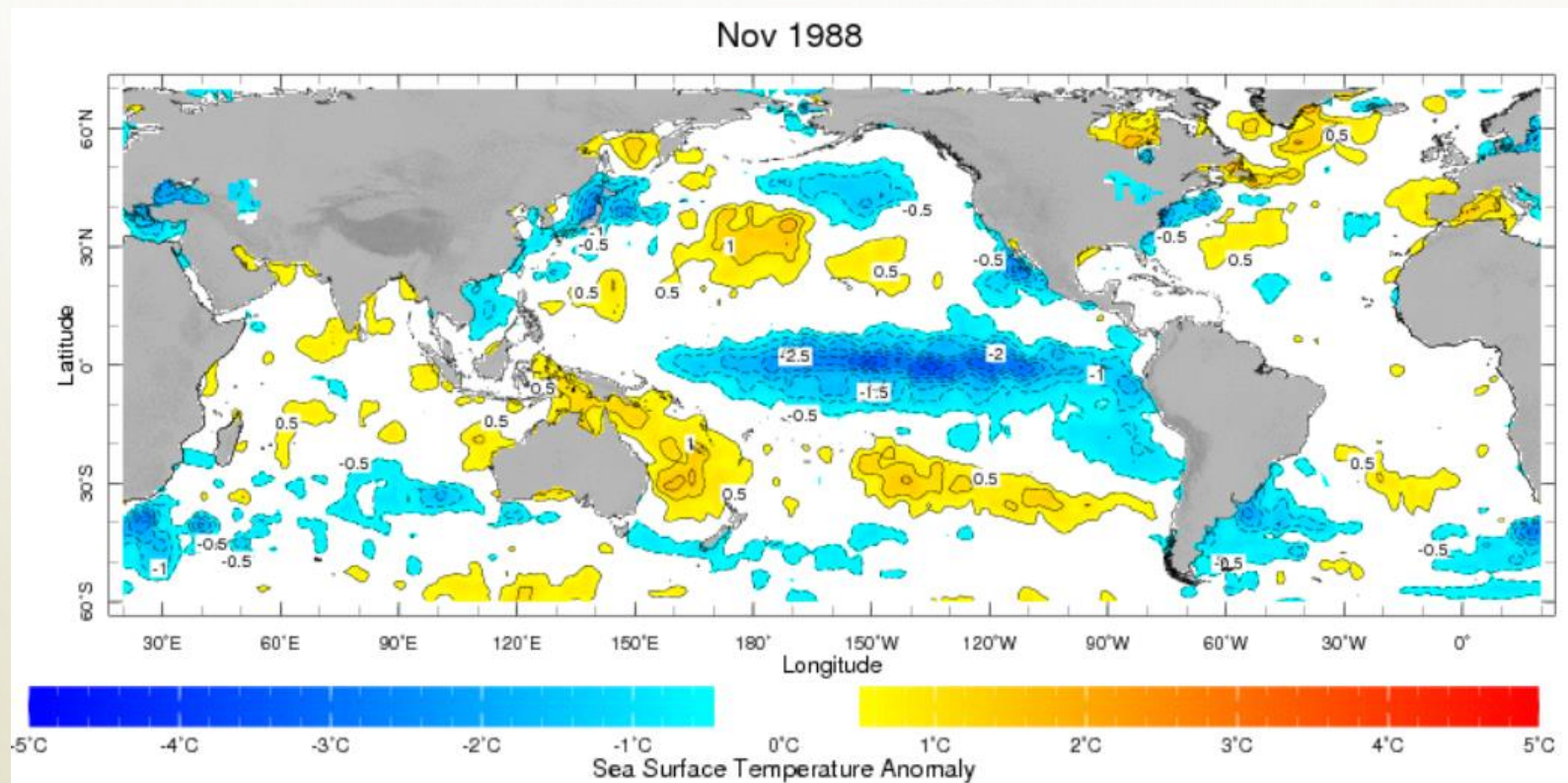


>>> GO TO JUPYTER

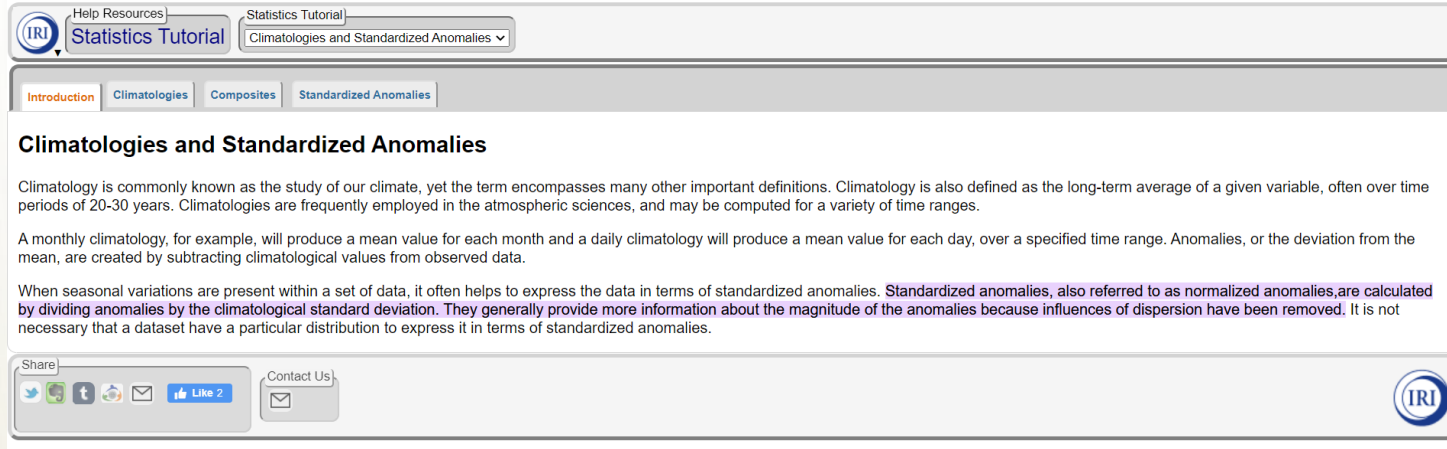
>>> *CONTINUARA*

ANOMALIAS

$OBS_i - \text{PROMEDIO}$



ANOMALÍA ESTANDARIZADA



The screenshot shows the IRI Statistics Tutorial interface. At the top, there's a navigation bar with 'Help Resources', 'Statistics Tutorial', and a dropdown menu for 'Climatologies and Standardized Anomalies'. Below this is a sub-navigation bar with tabs for 'Introduction', 'Climatologies', 'Composites', and 'Standardized Anomalies'. The main content area is titled 'Climatologies and Standardized Anomalies' and contains three paragraphs of text. The first paragraph defines climatology as the long-term average of a given variable. The second paragraph explains how monthly and daily climatologies are produced. The third paragraph discusses standardized anomalies, noting they are calculated by dividing anomalies by the climatological standard deviation. At the bottom, there's a 'Share' section with social media icons and a 'Contact Us' button. The IRI logo is in the bottom right corner.

Help Resources | Statistics Tutorial | Climatologies and Standardized Anomalies

Introduction | Climatologies | Composites | Standardized Anomalies

Climatologies and Standardized Anomalies

Climatology is commonly known as the study of our climate, yet the term encompasses many other important definitions. Climatology is also defined as the long-term average of a given variable, often over time periods of 20-30 years. Climatologies are frequently employed in the atmospheric sciences, and may be computed for a variety of time ranges.

A monthly climatology, for example, will produce a mean value for each month and a daily climatology will produce a mean value for each day, over a specified time range. Anomalies, or the deviation from the mean, are created by subtracting climatological values from observed data.

When seasonal variations are present within a set of data, it often helps to express the data in terms of standardized anomalies. Standardized anomalies, also referred to as normalized anomalies, are calculated by dividing anomalies by the climatological standard deviation. They generally provide more information about the magnitude of the anomalies because influences of dispersion have been removed. It is not necessary that a dataset have a particular distribution to express it in terms of standardized anomalies.

Share | Contact Us

Like 2

IRI

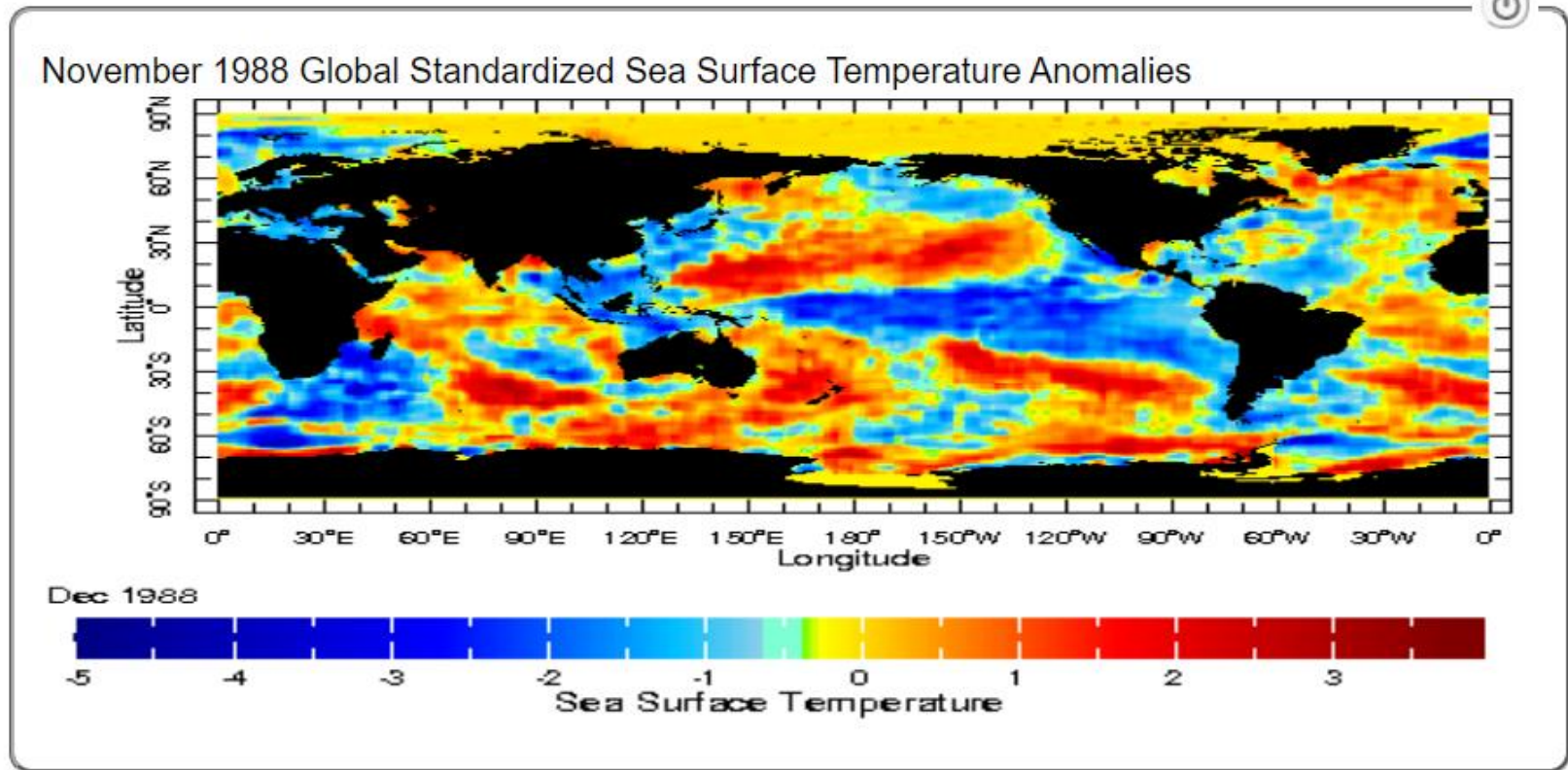
$$z = \frac{x - \bar{x}}{s_x} = \frac{x'}{s_x}$$

Anomalías Estandarizadas

- Una medida de distancia, en unidades estándar, entre un valor de datos y su media.
- Elimina las influencias de la ubicación y la dispersión de los datos.
- Más fácil de discernir entre valores normales e inusuales.
- ***Se calcula restando la media (climatológica) a cada observación, y luego dividiendo por la desviación estándar (climatológica).***
- Tienen las siguientes características: media = 0 y desviación estándar = 1.
- Adimensional

IRI*

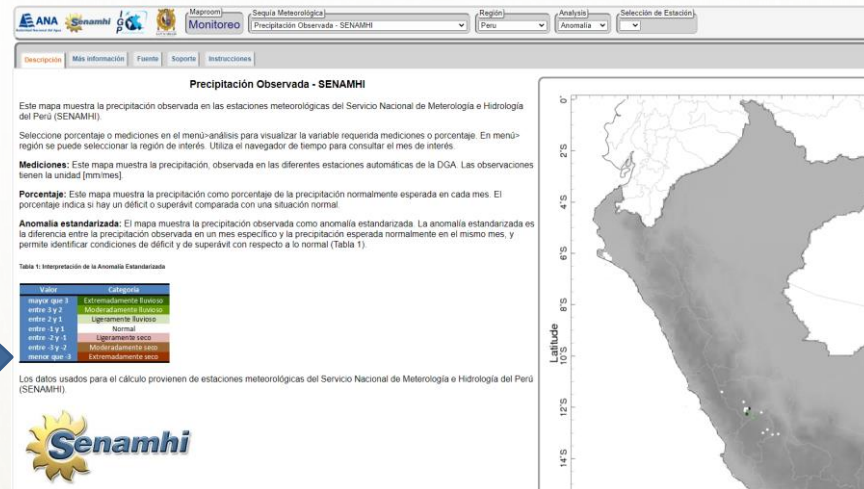
ANOMALÍA ESTANDARIZADA



IRI*

La estandarización de anomalías facilita la comparación de observaciones en dos ubicaciones diferentes. Por ejemplo, una anomalía de 2 ° Celsius puede ser **insignificante** en la costa de Perú y muy **significativa** en el Océano Atlántico central. Una anomalía estandarizada de 2 tiene el mismo **significado** relativo en ambas ubicaciones. Tenga en cuenta que las temperaturas de la superficie del mar por debajo del promedio en el Pacífico central y oriental son indicativas del evento La Niña de 1988.

ANOMALÍA ESTANDARIZADA

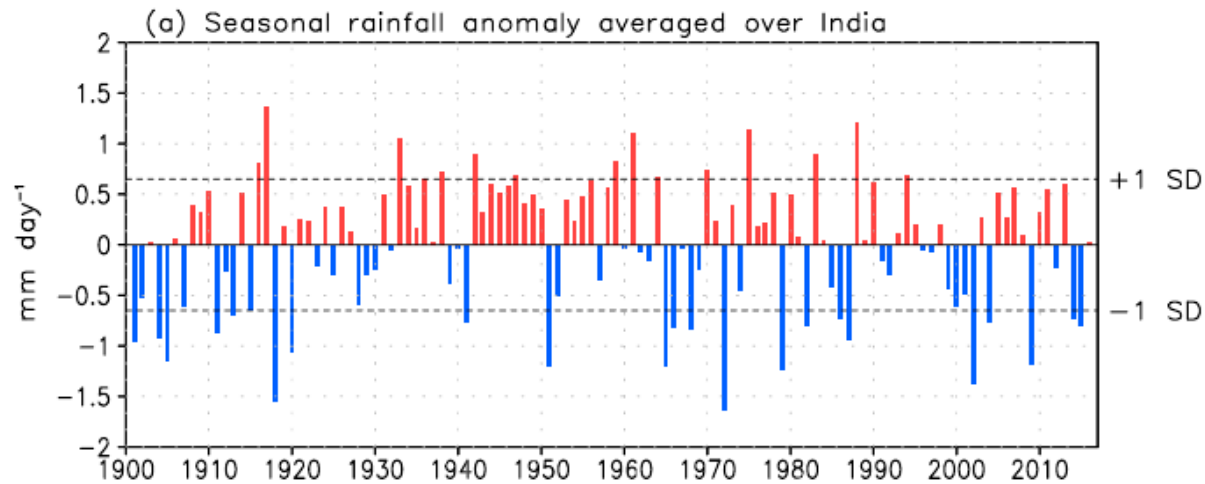


http://www.climatedatalibrary.cl/Peru/maproom/Monitoring/Meteorological/Precipitation_SENAMHI.html?anal=Anomaly

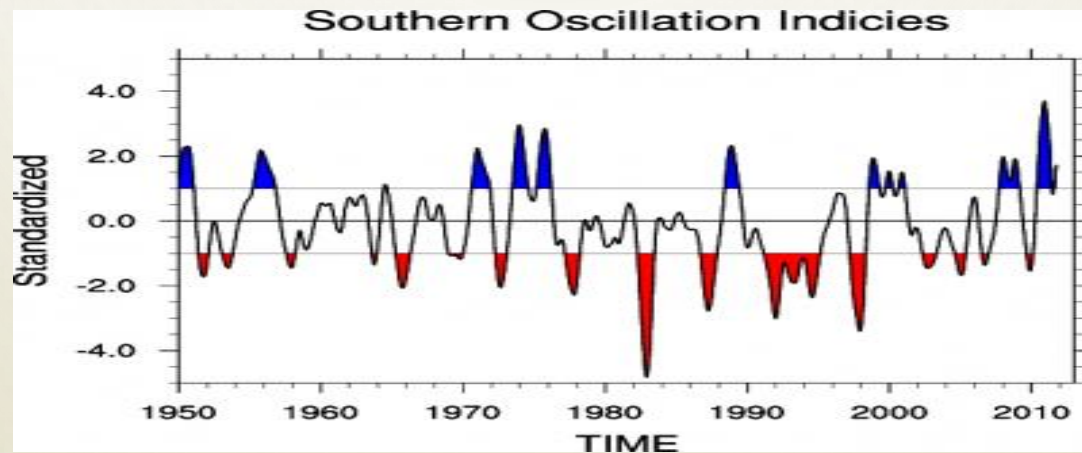
Anomalia estandarizada: Permite identificar condiciones de déficit y de superávit con respecto a lo normal

Valor	Categoría
mayor que 3	Extremadamente lluvioso
entre 3 y 2	Moderadamente lluvioso
entre 2 y 1	Ligeramente lluvioso
entre -1 y 1	Normal
entre -2 y -1	Ligeramente seco
entre -3 y -2	Moderadamente seco
menor que -3	Extremadamente seco

ANOMALÍA ESTANDARIZADA



Krishnamurthy, 2019*

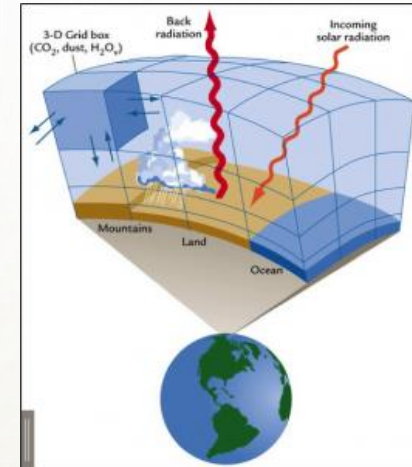
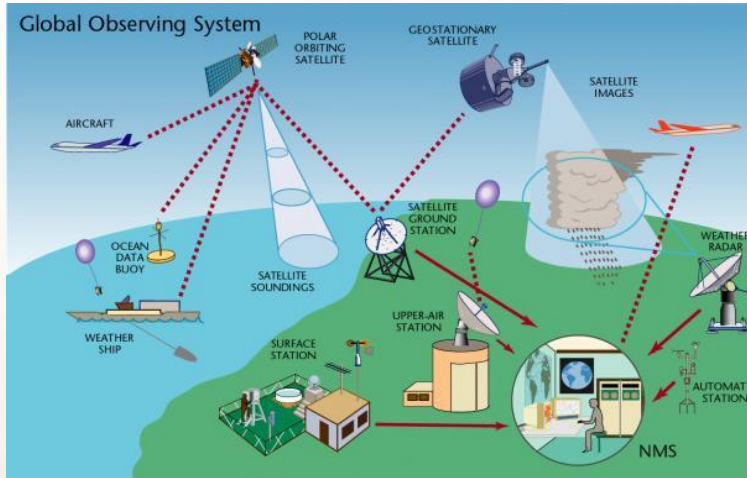


NCAR**

*"The Climate Data Guide: Overview: Climate Indices." Retrieved from <https://climatedataguide.ucar.edu/climate-data/overview-climate-indices>.

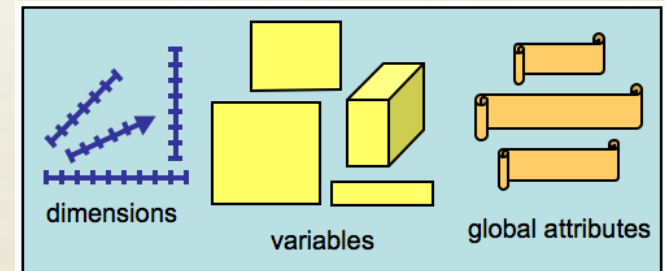
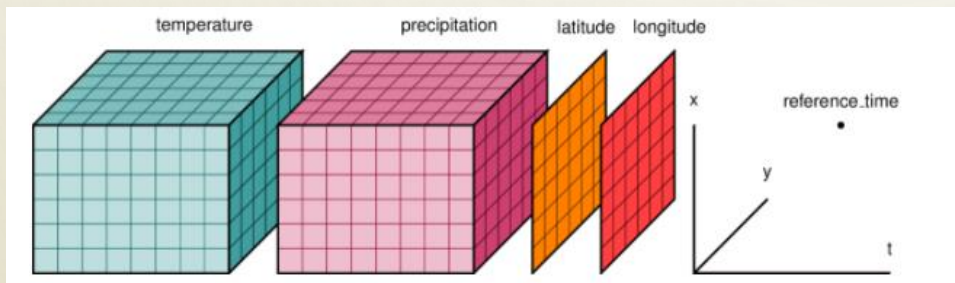
**<https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2019EA000586>

PRACTICA CON DATOS OBSERVADOS Y GRILLADOS



Harper, 2018

NetCDF (Network Common Data Form)



WEBS CON DATOS GRILLADOS

<https://psl.noaa.gov/data/gridded/tables/precipitation.html>

<https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html>

[https://cds.climate.copernicus.eu/cdsapp#!/search?text=ERA5
%20back%20extension&type=dataset](https://cds.climate.copernicus.eu/cdsapp#!/search?text=ERA5%20back%20extension&type=dataset)

<https://climatedataguide.ucar.edu/climate-data/jra-55>

>>> GO TO JUPYTER

