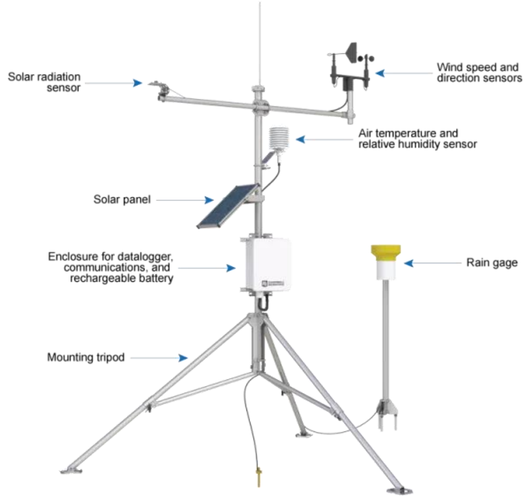
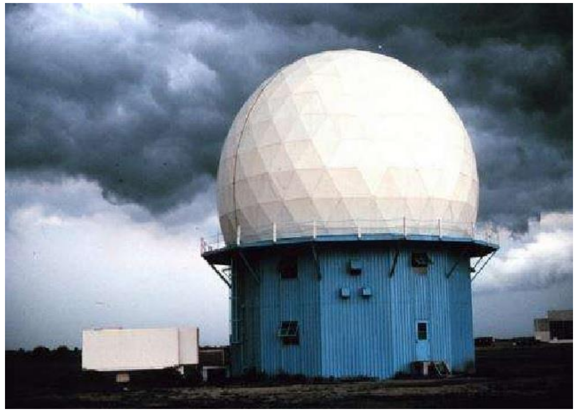
Introduction to climate data types and various sources

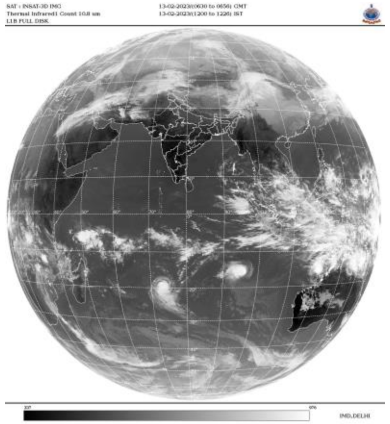


Dr. K.Bhuvaneswari, Ph.D.,

Tamil Nadu Agricultural University

Coimbatore





Introduction

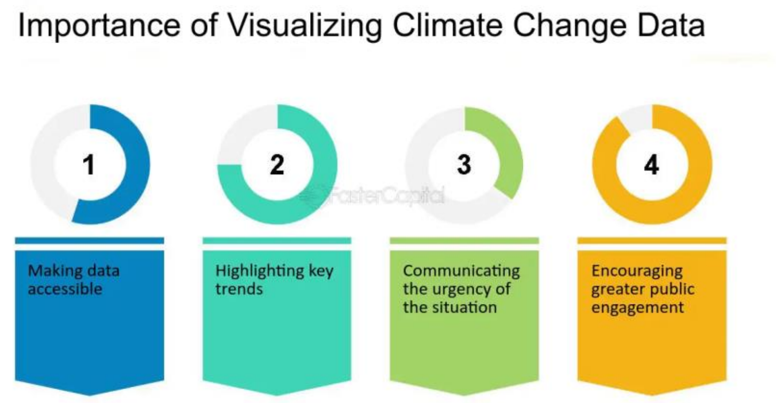
- Climate is one of the major drivers of agricultural productivity, water availability, disaster occurrence and ecosystem balance.
- Understanding climate through data is not optional—it is a necessity for planning, preparedness and sustainability.
- Historical data alone is not enough; real-time and forecast data must be analyzed to anticipate risks.

Why Climate Data Analysis is Needed

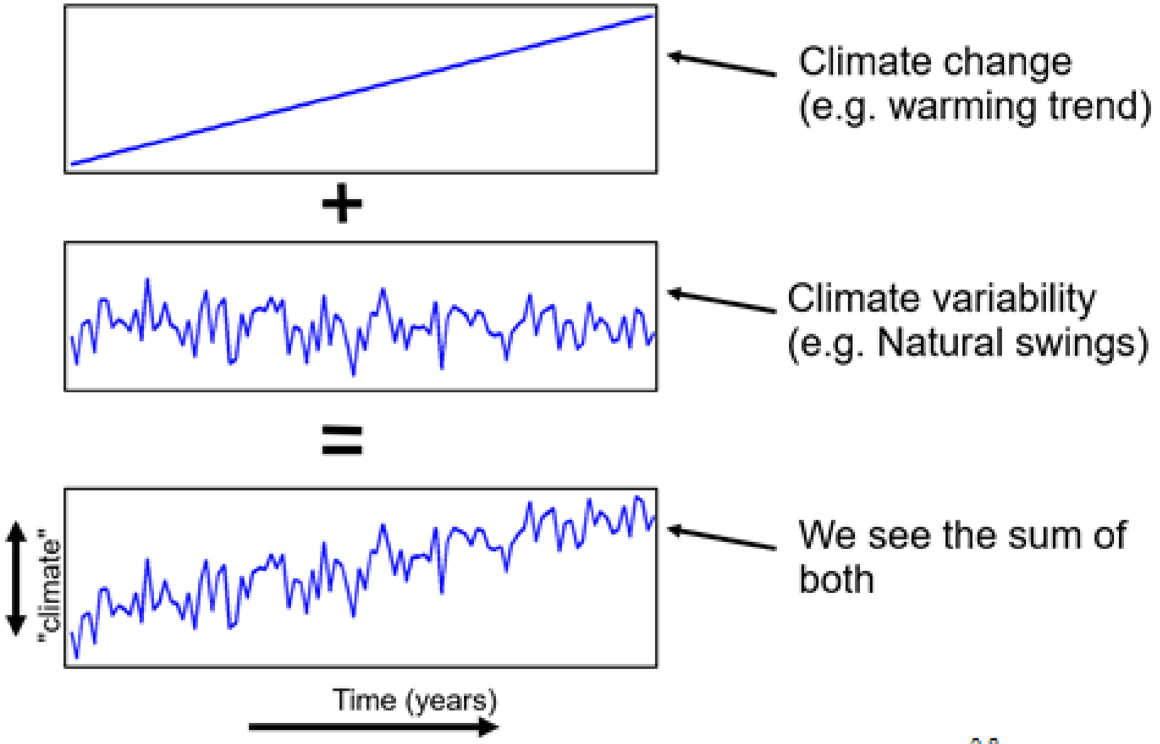
- To assess Regional Climate variability and Change
- Frequency and intensity of drought and flood years
- Extreme weather events (heat, cold waves, heavy precipitation events etc.,)
- Understand trends in weather and climate variables
- Rainfall probability analysis, crop planning, LGP estimation
- For making contingency plans
- Water availability for agriculture
- Crop-weather relationship and pest-weather relationship
- Crop yield prediction
- Pest and disease forewarning
- Crop insurance products
- Making policy decisions

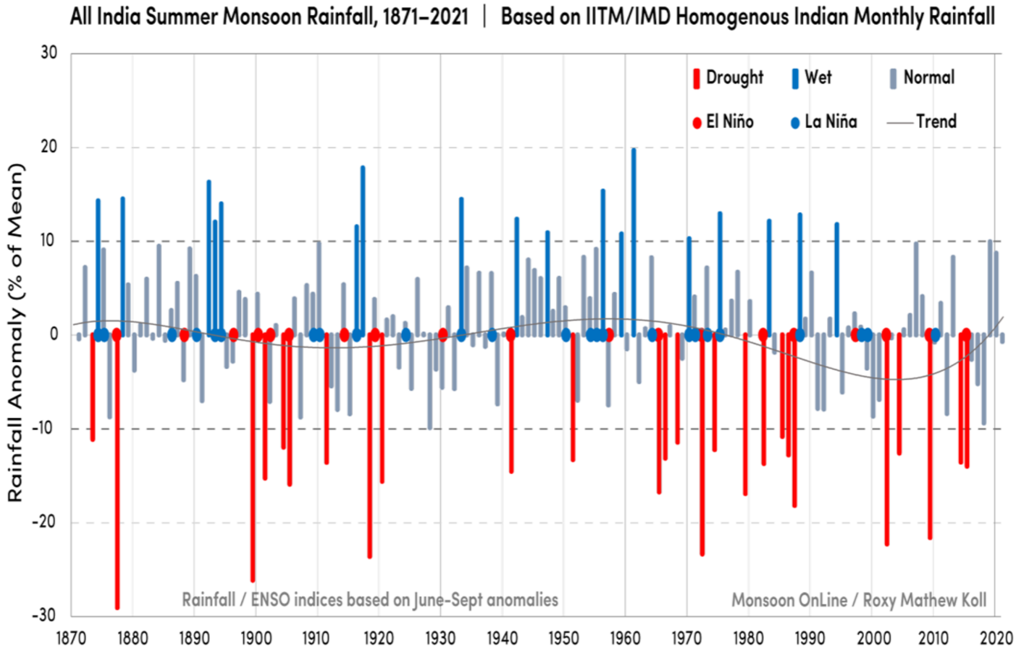
Why Climate Data Analysis is Needed

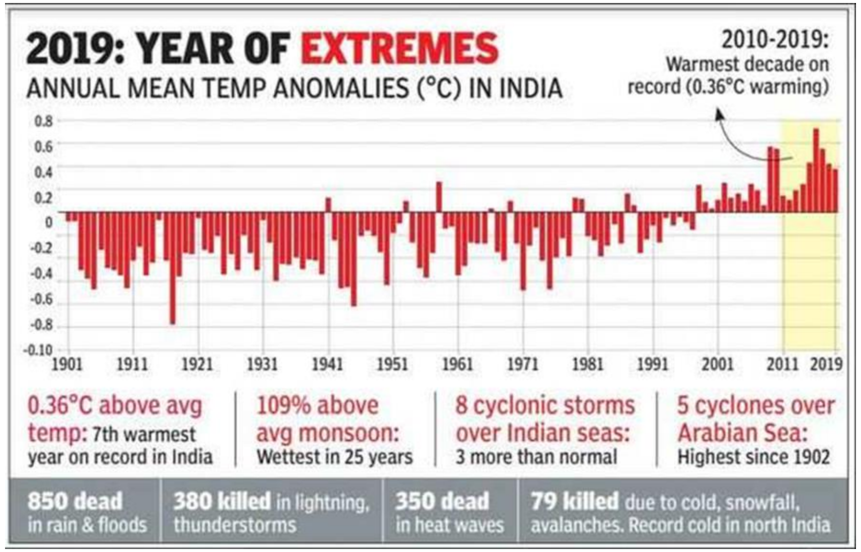
- It supports risk reduction and sustainable decision-making in agriculture, water and disaster management
- Infrastructure Design
  - Design of canals, check dams, and urban drainage depends on rainfall intensity and return periods. Long-term data helps estimate severe events
- Agro-advisories use processed climate data to give farmers timely information on rainfall, temperature trends, and recommended actions.

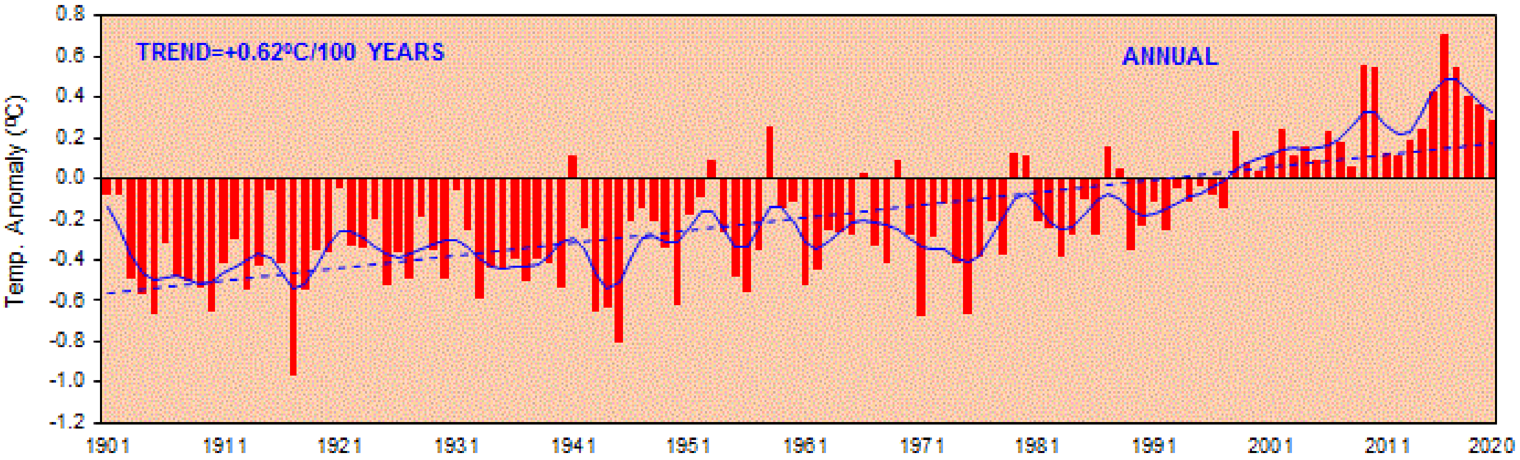


Climate Change v. Climate Variability









TYPES OF DATA

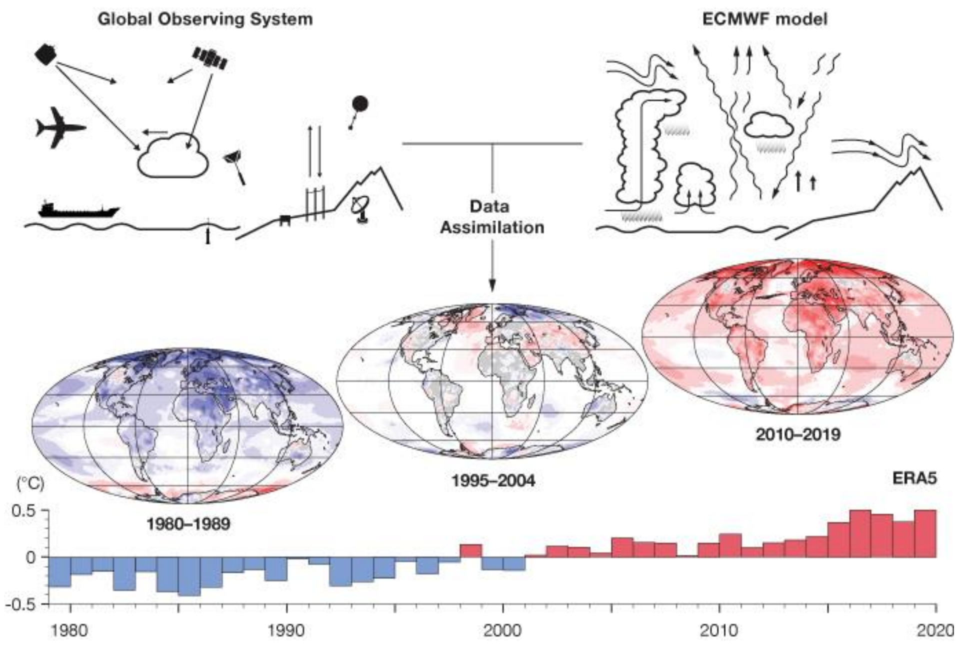
Main Types of Climate Data

1. Observational Data – Ground-based measurements from weather stations

2. Satellite Data – Remote sensing information over large areas

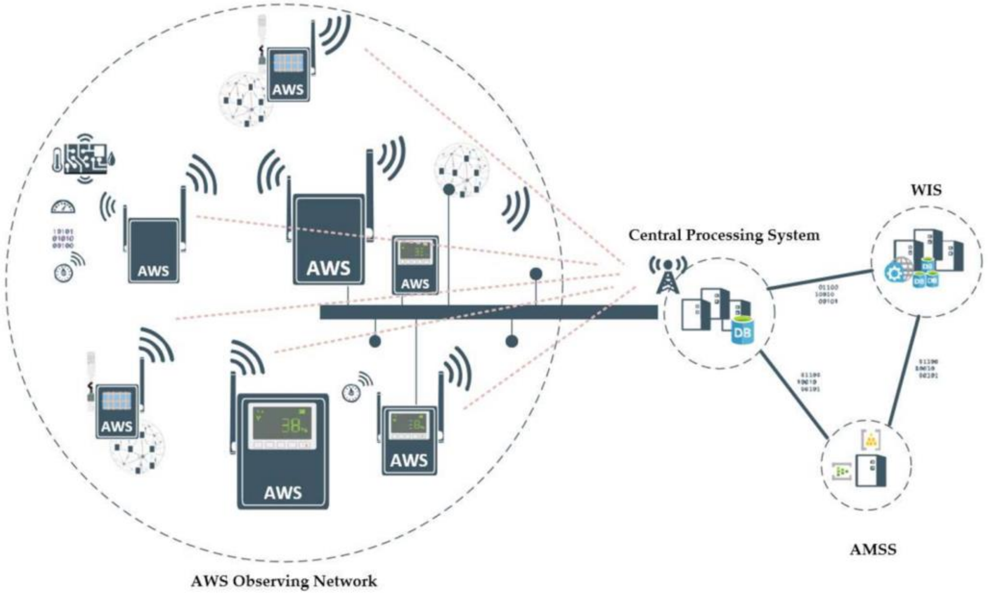
3. Model-Based Data – Reanalysis and projections from climate models

4. Derived Indices – Computed variables like SPI, NDVI, and GDD



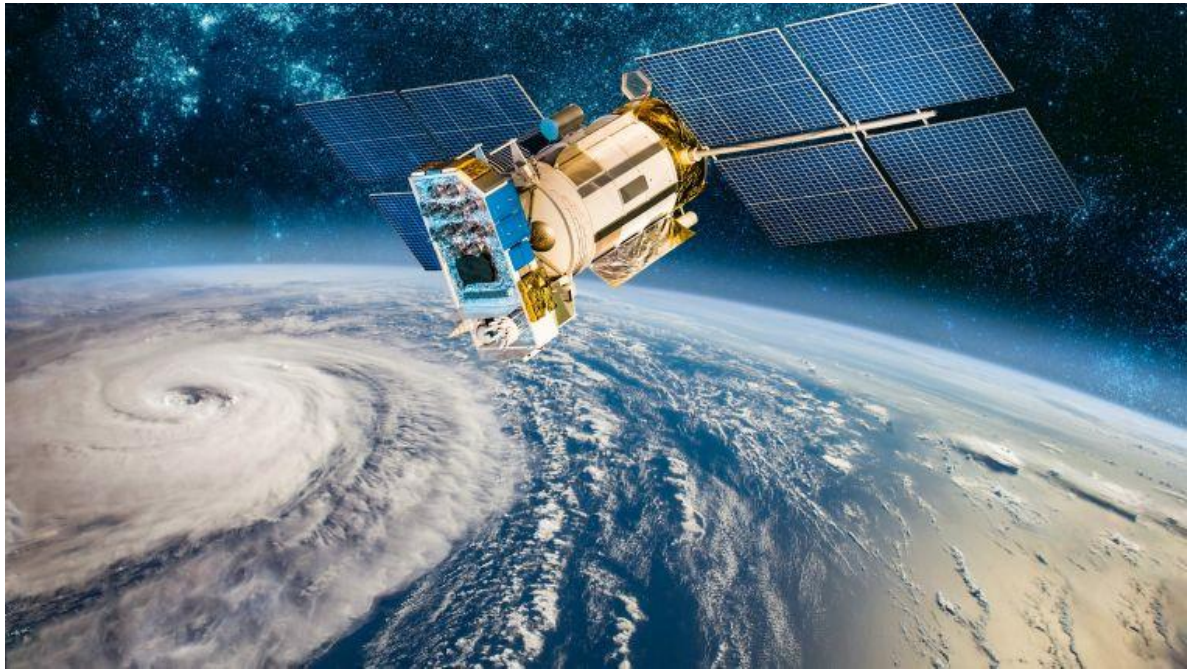
Observational Data

- Collected through ground stations and instruments such as rain gauges, thermometers, and anemometers.
- High accuracy at specific points but may have spatial gaps.
- Example: India Meteorological Department (IMD), SAUs, ICAR-KVKs.



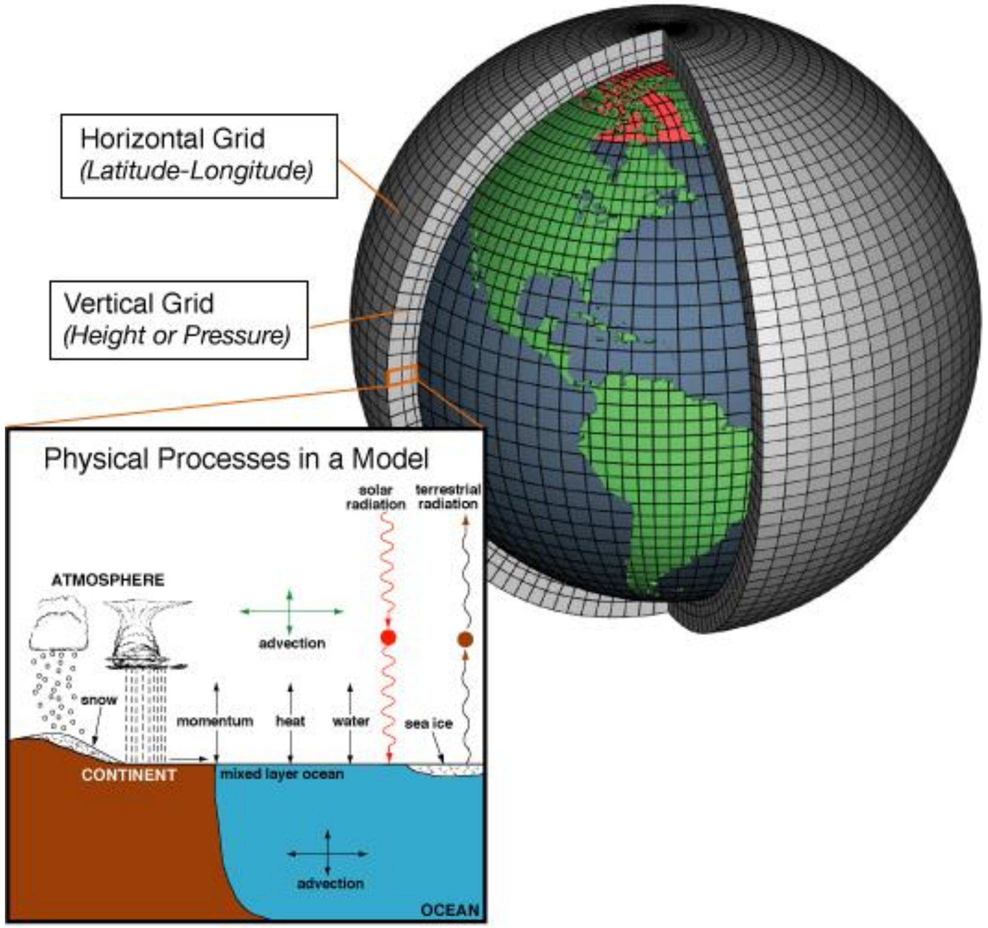
Satellite Data

- Covers large geographic areas using remote sensing instruments onboard satellites.
- Provides information on rainfall estimates, land surface temperature, vegetation health, and soil moisture.
- Sources: NASA (MODIS, CHIRPS), ISRO, ESA etc



Model-Based Climate Data

- 1. Reanalysis Data – Combines past observations with model outputs (e.g., ERA5, NCEP).
- 2. Climate Projections – Simulates future climate under different emission scenarios (e.g., CMIP6, CORDEX).
- Used in long-term planning and vulnerability assessments.



DATA FORMAT

Climate data format

- Climate data is produced in different forms depending on the instrument and purpose.
- Understanding these formats helps in storing, analyzing, sharing, and visualizing climate information.
- Each format has specific strengths and tools for access and analysis.

CSV – Comma-Separated Values

- Common for weather station data: daily temperature, rainfall, wind speed
- Easily opened in Excel, Python, and R

Example :

Date, Tmax(°C), Tmin(°C), Rainfall(mm)

2025-06-01, 34.8, 25.2, 15.0

NetCDF – Network Common Data Form

- Stores multidimensional data (e.g., time × latitude × longitude)
- Common in climate models, reanalysis, and satellite data
- Includes metadata (units, descriptions, etc.)
- File extension: .nc
- Tools: Panoply, CDO, Python (xarray, netCDF4)

GRIB – GRIdded Binary

- Binary format for operational meteorological data
- Used by ECMWF, IMD, NOAA, etc.
- Compact, fast to transmit and process
- File extension: .grib, .grb
- Tools: wgrib2, cfgrib (Python), ecCodes

HDF – Hierarchical Data Format

- Designed for managing complex data structures
- Widely used in satellite products (e.g., MODIS)
- Allows storage of different data types and metadata in one file
- Extensions: .hdf, .h5
- Tools: HDFView, Python (h5py), Panoply

GeoTIFF – Georeferenced TIFF

- Raster format with embedded geographic coordinates
- Used in remote sensing and GIS (e.g., NDVI, LST, CHIRPS rainfall)
- Supports spatial analysis
- File extension: .tif (Tagged Image File Format)
- Tools: QGIS, ArcGIS, GDAL (Geospatial Data Abstraction Library), Python (rasterio)

JSON & XML – Structured Data for APIs

- Lightweight formats for structured data exchange
- Used in climate web services and dashboards
- Easily parsed using web technologies and scripts
- File extensions: .json (JavaScript Object Notation), .xml(Extensible Markup Language)
- Tools: Browsers, Python, R, JavaScript

Summary of Climate Data Formats

Format	Common Use	Tools
CSV	Station data	Excel, R, Python
NetCDF	Gridded model/reanalysis data/forecast/Satellite products/GIS	Panoply, Python, CDO
GRIB	Weather forecasts	wgrib2, cfgrib, ecCodes
HDF	Satellite products	HDFView, Python, Panoply
GeoTIFF	Remote sensing, GIS	QGIS, ArcGIS, GDAL
JSON/XML	API (Application Programming Interface)/web data	Web, Python, JavaScript

DATA SOURCES

1. National Meteorological Observations

- IMD (India Meteorological Department):

 - Gridded rainfall (1901–2024), temperature datasets

 - Format: NetCDF

spatial resolution (0.25°×0.25°)

Link : https://imdpune.gov.in/cmpg/Griddata/Rainfall_25_NetCDF.html

Link: https://www.imdpune.gov.in/cmpg/Griddata/Rainfall_25_Bin.html

- **IMD - Data Service Portal Version 5.0 -** for getting historical meteorological observation data from IMD

Link: https://dsp.imdpune.gov.in/

State level: TNAU: AWS network data at block level

2. Reanalysis and Satellite data

- **NASA POWER**: Combines data from MERRA-2 and CERES to provide meteorological variables (rainfall, solar radiation, temperature, humidity, etc.) power.larc.nasa.gov.
  - MERRA-2, or the Modern-Era Retrospective analysis for Research and Applications, Version 2, is a comprehensive, long-term global atmospheric reanalysis dataset produced by NASA's Global Modeling and Assimilation Office (GMAO)
  - MERRA-2 (NASA): Atmospheric data from 1981 to present

Satellite data

MSWEP - https://www.gloh2o.org/mswep/

IMERG - https://disc.gsfc.nasa.gov/

CMORPH - https://www.ncei.noaa.gov/products/climate-data-records/precipitation-cmorph

CHIRPS - Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) - USAID, NASA, and NOAA - https://www.chc.ucsb.edu/data/chirps3

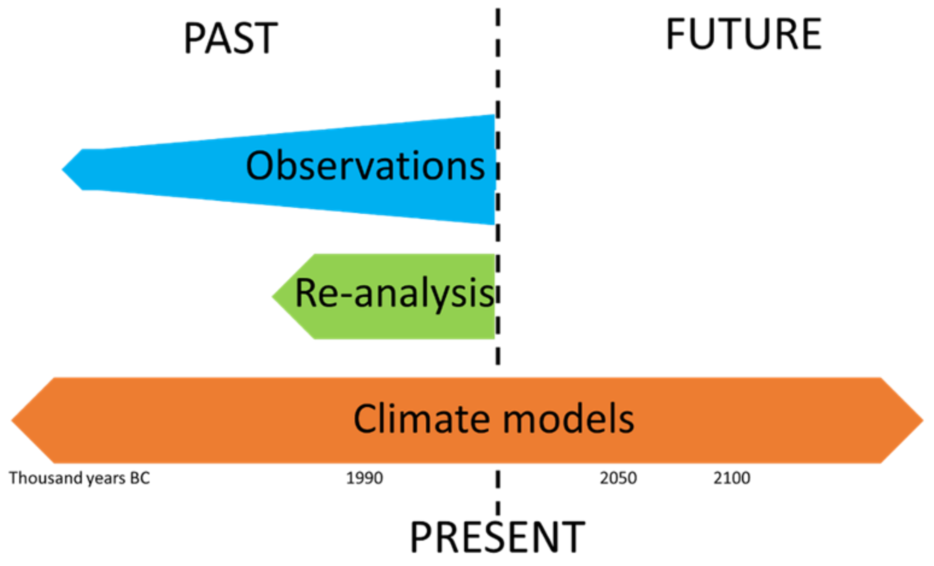
Sentinel and MODIS: land surface temperature

3. Reanalysis Datasets

- ERA5 (ECMWF): Hourly global data since 1940
  - Formats: NetCDF, GRIB
- NCEP/NCAR, CFSR: Long-term reanalyses

Link: https://climatedataguide.ucar.edu/climate-data/reanalysis

Link: https://cds.climate.copernicus.eu/



4. Climate Model Projections

- CMIP6 – Global climate simulations (multiple scenarios)
- CORDEX – Regional downscaled projections
- Formats: NetCDF

Link: https://esgf-node.llnl.gov/projects/cmip6/

CMIP6 data from Copernicus
https://cds.climate.copernicus.eu/datasets/projections-cmip6?tab=download

CMIP6 NASA NEX-GDDP: https://www.nccs.nasa.gov/services/data-collections/land-based-products/nex-gddp-cmip6

Open Platforms & Tools

• Google Earth Engine – Large-scale climate and remote sensing analysis

• Copernicus CDS – Free access to climate datasets

• Climate Data Guide (NCAR) – Resource hub

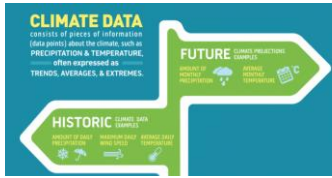
- Link: https://earthengine.google.com/

Summary of Data Sources

- • National and state: IMD, SAUs
- • Satellite: NASA, ISRO, ESA
- • Reanalysis: ERA5, MERRA-2, CFSR
- • Historical: GHCN, CRU, HadISD
- • Model: CMIP6, CORDEX
- • Platforms: Google Earth Engine, Copernicus CDS

Sources of climate data

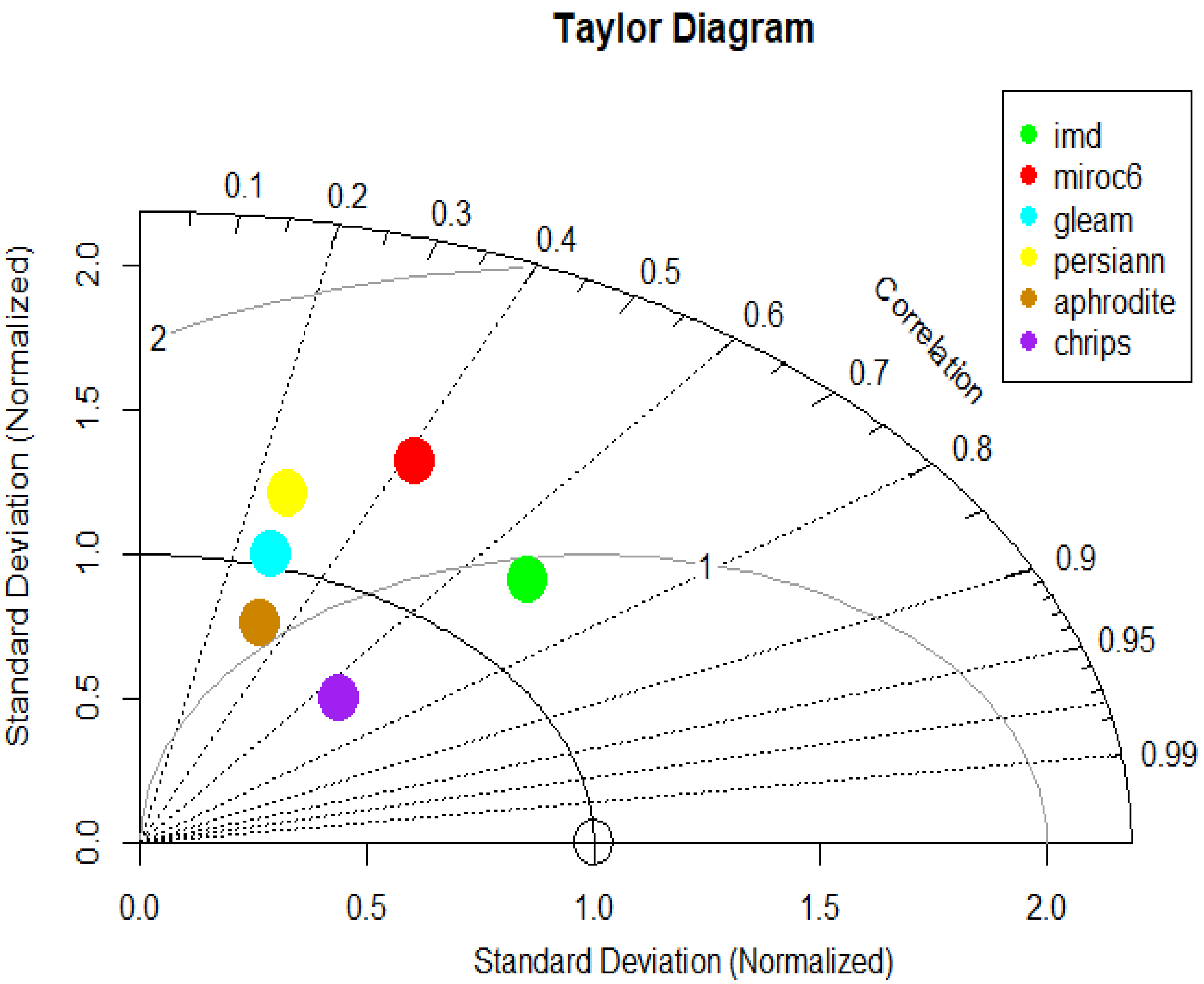
- APHRODITE (Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources)- https://climatedataguide.ucar.edu/climate-data/aphrodite-asian-precipitation-highly-resolved-observational-data-integration-towards
- Global Historical Climatology Network (GHCN) data – Climate data online-National Climatic Data Center – NOAA -https://psl.noaa.gov/data/gridded/
- CPC (Climate prediction centre) - NOAA- https://psl.noaa.gov/data/gridded/
- PERSIANN-CDR(Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record)- NCAR- https://climatedataguide.ucar.edu/climate-data/persiann-cdr-precipitation-estimation-remotely-sensed-information-using-artificial
- TRMM (Tropical Rainfall Measuring Mission)- NASA- https://gpm.nasa.gov/missions/trmm
- ERA-Interim (European Environment Agency)- ECMWF - https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim
- ERA-5 –ECMWF- https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5
- IMDAA-Like –NEMRWF- https://rds.ncmrwf.gov.in/



Performance evaluation of precipitation data for Tamil Nadu

- **Observed data**
  - TNAU
- **Gauge-based gridded**
  - IMD, APHRODITE,
- **Gauge- Satellite-based gridded**
  - CHRIPS, GLEAM
- **Satellite precipitation products**
  - PERSIANN-CDR
- **Climate model**
  - MIROC6

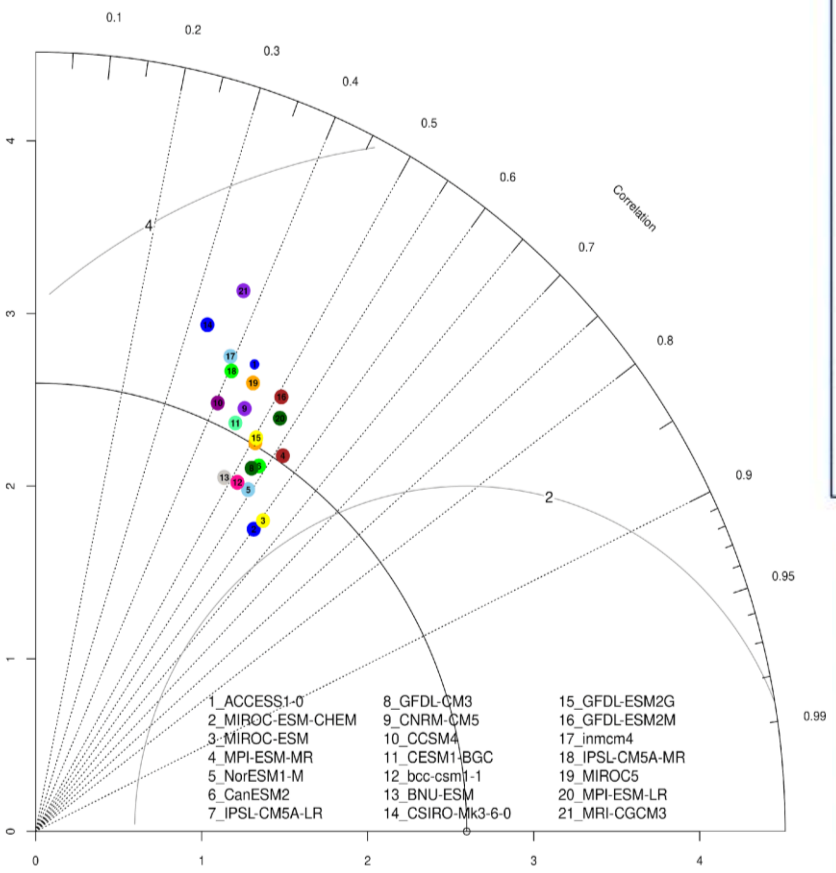
	observed	miroc6	imd	chrips	glem	perisian	aphrodite
mean	964.1	1194.5	962.1	995.6	844.3	1421.5	813.1
std	199.9	291.3	249.9	133.8	209.0	250.8	161.8
corr	-	0.42	0.68	0.66	0.27	0.26	0.33
RMSE	-	355.5	181.0	151.1	269.3	532.0	256.6
MAPE	-	30.4	15.5	14.4	19.7	52.5	16.5
NSE	-	-2.30	0.14	0.40	-0.89	-6.39	-0.72



Reference: https://waterprogramming.wordpress.com/2020/12/22/taylor-diagram/

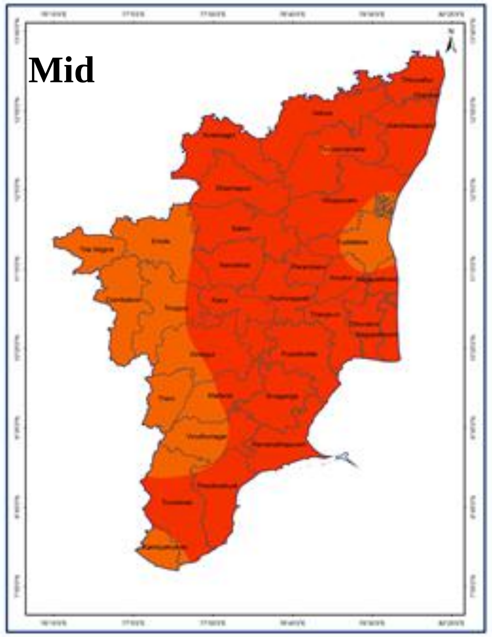
Climate change projections

CanESM2

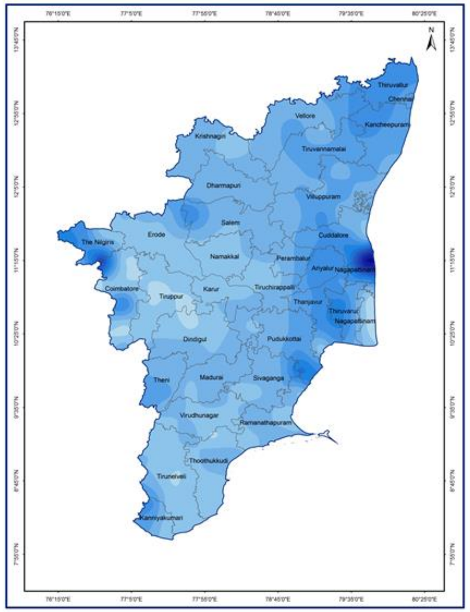


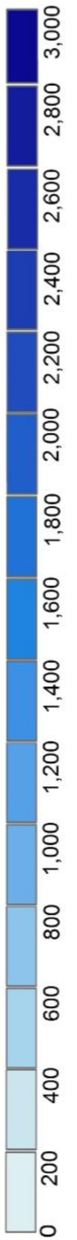
1	MPI-ESM-MR
2	NORESM-M
3	CANESM2
4	IPSL-CM5A-LR

Tmax -annual

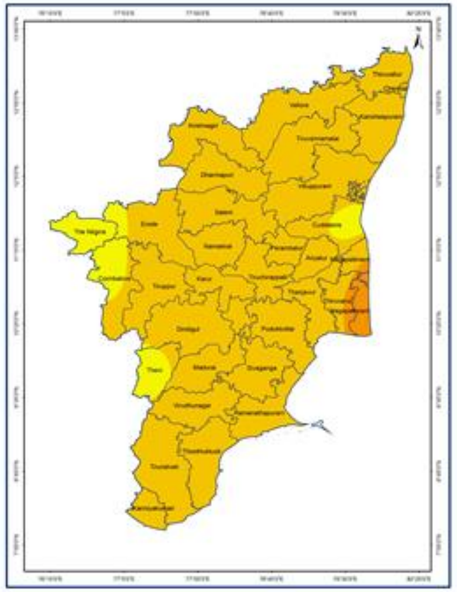


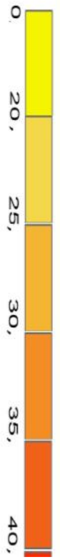
Rainfall -annual

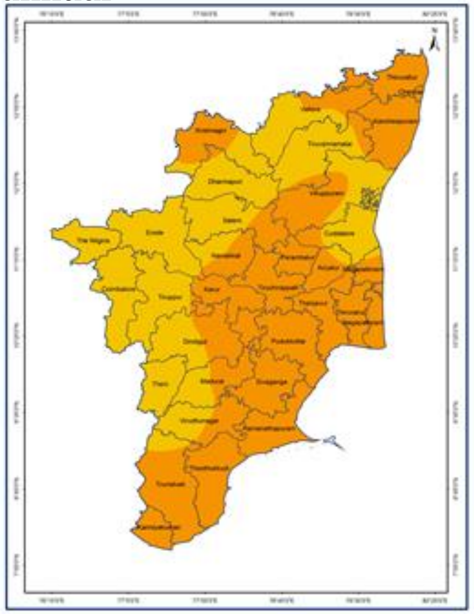


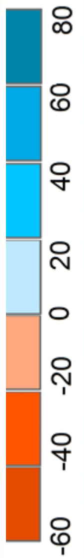


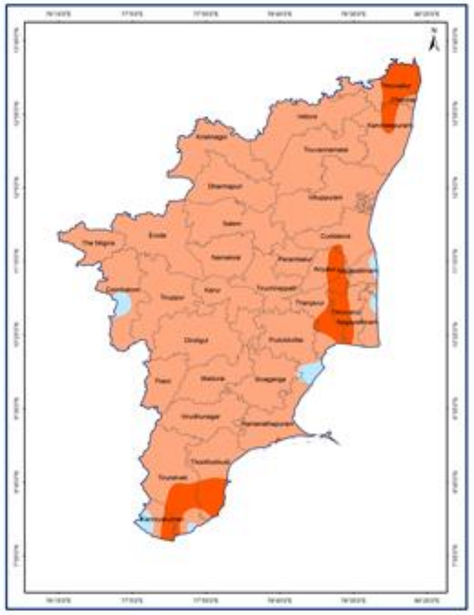
Tmin-annual











Climate data analysis

- Climate data analysis refers to examining, processing and interpreting the records to understand patterns, trends, risks etc

1. Examining

This step involves checking the raw data:

- Looking at daily, monthly, or seasonal records
- Spotting missing data, errors, or inconsistencies
- Ensuring the quality of the data before any calculations
- **Example:** Checking 30 years of rainfall data from a district to identify missing values or outliers.

2. Processing

In this step, the data are cleaned, formatted, and organized for analysis:

- Aggregating daily data into monthly or seasonal totals
- Converting units (e.g., from Fahrenheit to Celsius)
- Handling large files like NetCDF using tools like Python or CDO
- Using statistical techniques like averaging, standard deviation, and smoothing
- **Example:** Converting raw rainfall data into seasonal averages to compare monsoon strength across years.

3. Interpreting

- Once the data are processed, the results must be made meaningful:
- Identifying **trends**: Is rainfall decreasing over decades?
- Spotting **patterns**: Are heatwaves happening more frequently in April than before?
- Assessing **risks**: Is a region more prone to drought or flood based on past events?
- **Example:** Graph shows a rising trend in maximum temperature in a district over 30 years - indicate heat stress.

Tools for climate analysis

- Climpact2

https://climpact-sci.org/
https://climpact-sci.org/assets/climpact2-user-guide.pdf

- weathercock

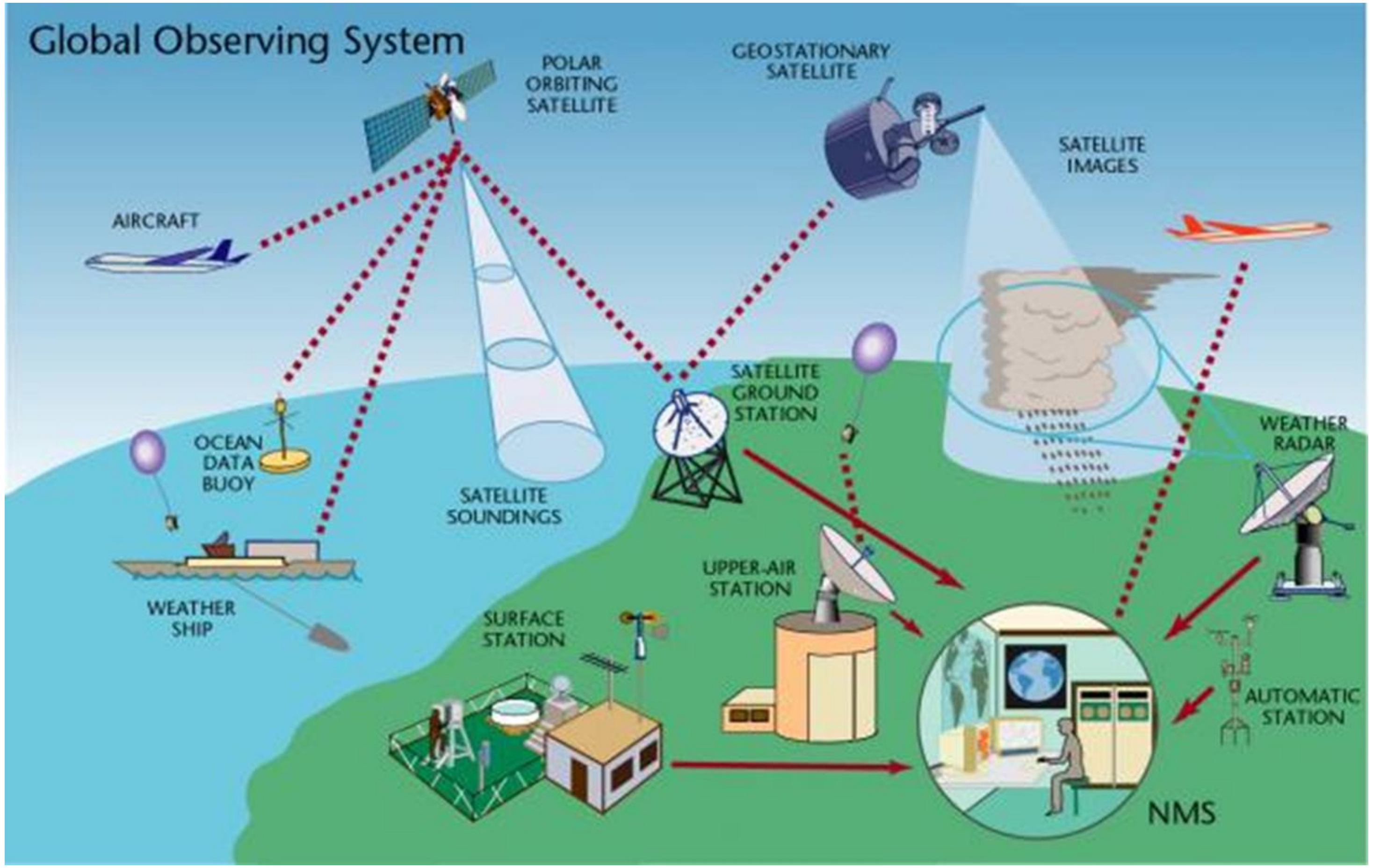
Indices with description

Index	Description
CDD	Annual maximum length of dry spell: maximum number of consecutive dry days with RR < 1mm
CWD	Annual maximum length of wet spell: maximum number of consecutive wet days with RR < 1mm
PRCPTOT	Annual total precipitation on wet days
SDII	Simple precipitation intensity index
r-10 mm	Annual count of days when PRCP < 10mm
r-20 mm	Annual count of days when PRCP < 20 mm
r-30 mm	Annual count of days when PRCP < 30 mm, where nn is a user-defined threshold
rx 1 day	Annual maximum 1-day precipitation
rx 3 day	Annual maximum 3-day precipitation
rx 5 day	Annual maximum consecutive 5-day precipitation
tmm	Annual the mean daily mean temperature.
txm	Annual mean daily maximum temperature.
tnm	Annual mean daily minimum temperature.
Spi	Standard precipitation index
Spei	Standardized precipitation evapotranspiration index

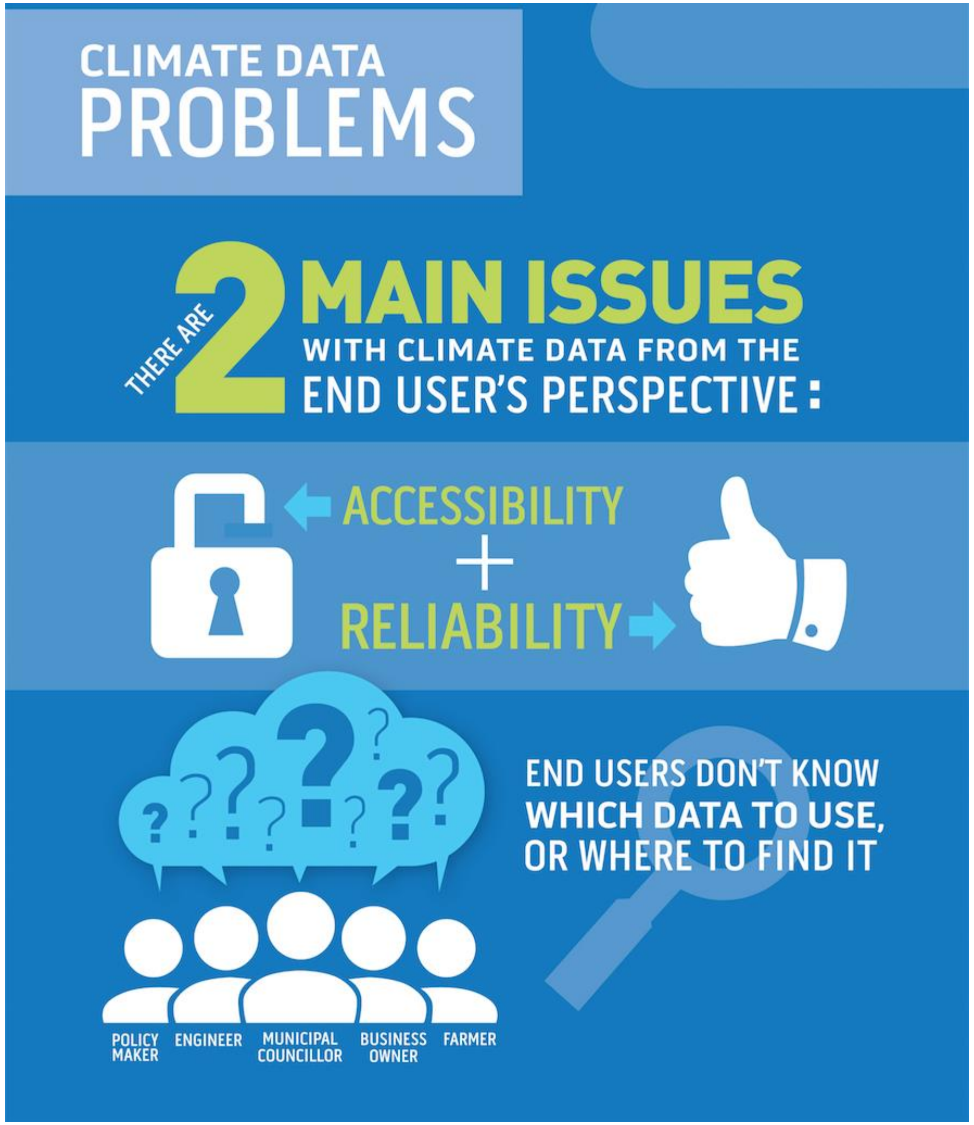
Conclusion

- Understanding different types of climate data - such as observations, satellite products, reanalysis datasets, and model projections - is essential for accurate climate analysis and informed decision-making.
- Each data type has unique strengths and is suited for specific applications in agriculture, water management, disaster planning and policy.
- Accessing reliable sources like IMD, NASA, ECMWF and Google Earth Engine ensures data quality and relevance.
- Familiarity with these formats and platforms enables better analysis, visualization and climate-resilient planning.

Thank you

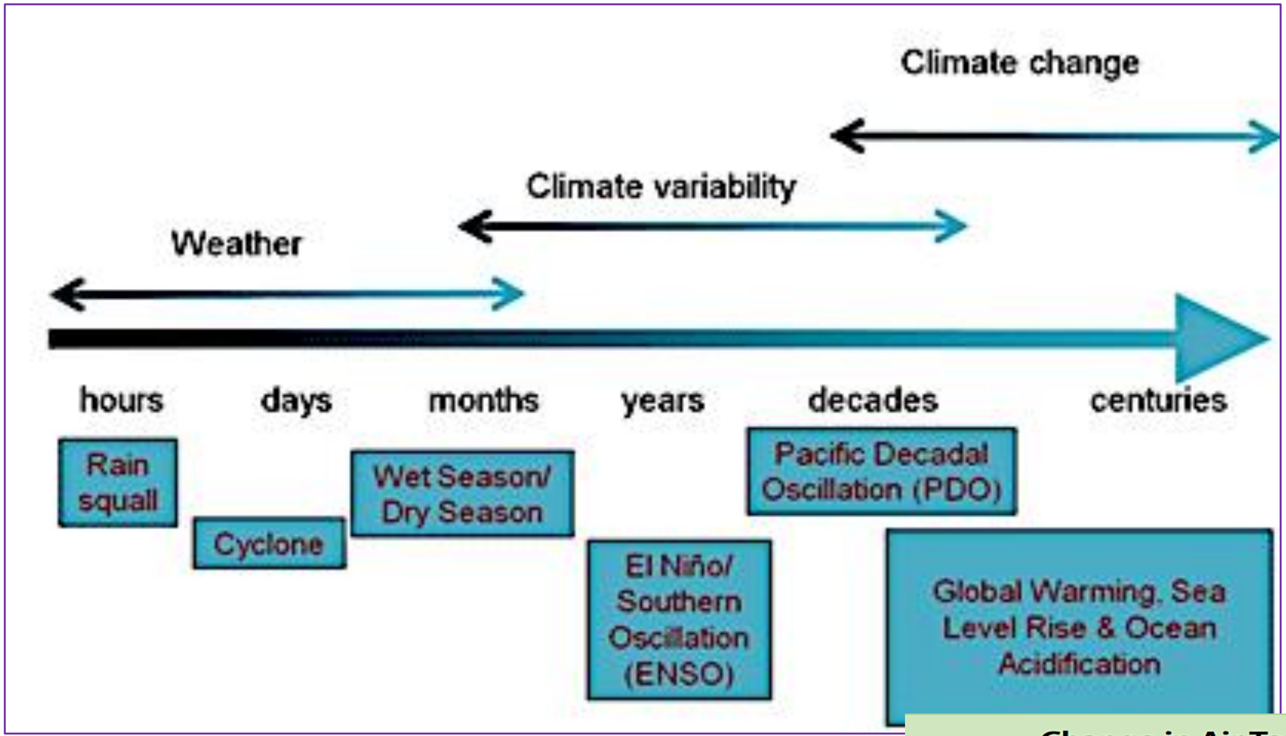

Global Observing System
POLAR ORBITING SATELLITE
GEOSTATIONARY SATELLITE
SATELLITE IMAGES
AIRCRAFT
OCEAN DATA BUOY
SATELLITE SOUNDINGS
SATELLITE GROUND STATION
WEATHER RADAR
WEATHER SHIP
UPPER-AIR STATION
SURFACE STATION
AUTOMATIC STATION
NMS

Climate data problems and solutions



GRIB (**GRIdded Binary or General Regularly-distributed Information in Binary form**) , NetCDF (network Common Data Form), .grd file,

Environmental commissioner of Ontario



Climate Change and Climate Variability

Climate variability: Short-term change

Deviations of climate statistics over a given period of time (such as a specific month, season or year) from the long-term climate statistics relating to the corresponding calendar period.

Climate Change -long term change in weather statistics over a period of time that ranges from **decades to millions of years.**

Weather: current state of the atmosphere

day-to-day state of the atmosphere and its short-term (from hours to a few weeks) variations such as temperature, humidity, precipitation, cloudiness, visibility or wind.

Climate : average weather conditions over a long period of time – statistical information -30 years