Sources:

Web:

Gidon Eshel's excellent data analysis class notes – the main "text book" for this course (note there are two courses on data analysis).

A manual for EOF and SVD analyses of climate data. 1997. Bjornsson Halldor and Venegas, Silvia A. A good introductory web document on EOF and SVD techniques. link to pdf copy.

Bertran Timbal's statistics and climatology presentation were a source for some of my presentation slides.

Books: There are many. I used, for some of the stuff:

Statistical Analysis in Climate Research, by Hans Von Storch, Francis W. Zwiers, Cambridge University Pr, H. V. Storch (Cambridge Univ Pr, 1999)

A good linear algebra book: Strang, G. *Linear Algebra and its Applications* 3rd Ed., Harcourt Brace Jovanovich, San Diego (1988) ISBN 0-15-551005-3.

Goal of this course:

Learn and apply basic statistical concepts and data analysis procedures, using synthetic and real data products.

In the process, you will hopefully gain a sense of:

Data analysis and it’s limitations

Its applications to:

- Examining interrelationships between various climate variables, and testing theoretical models.
- Observational analysis and forecast production (if time permits)
- Forecast verification (if time permits)

Various aspects of observed atmospheric/climate variability

And some practical experience with:

Downloading and processing climate data (using Ingrid or another platform)

Matrix manipulation and data analysis in MATLAB

Atmospheric/climate variables are best viewed as stochastic.

A stochastic process is one whose behavior is non-deterministic in that a state's next state is determined both by the process's predictable actions and by a random element. (Wikipedia)

But isn't the atmosphere governed by known, deterministic dynamics?

Yes, but -

- Very complex system, with numerous controlling factors, not all of which can be taken into account. We usually examine a few and treat the rest as "background noise".

- Observations are not perfect. Not all space and time scales are resolved – some processes need to be parameterized.

- The system is nonlinear

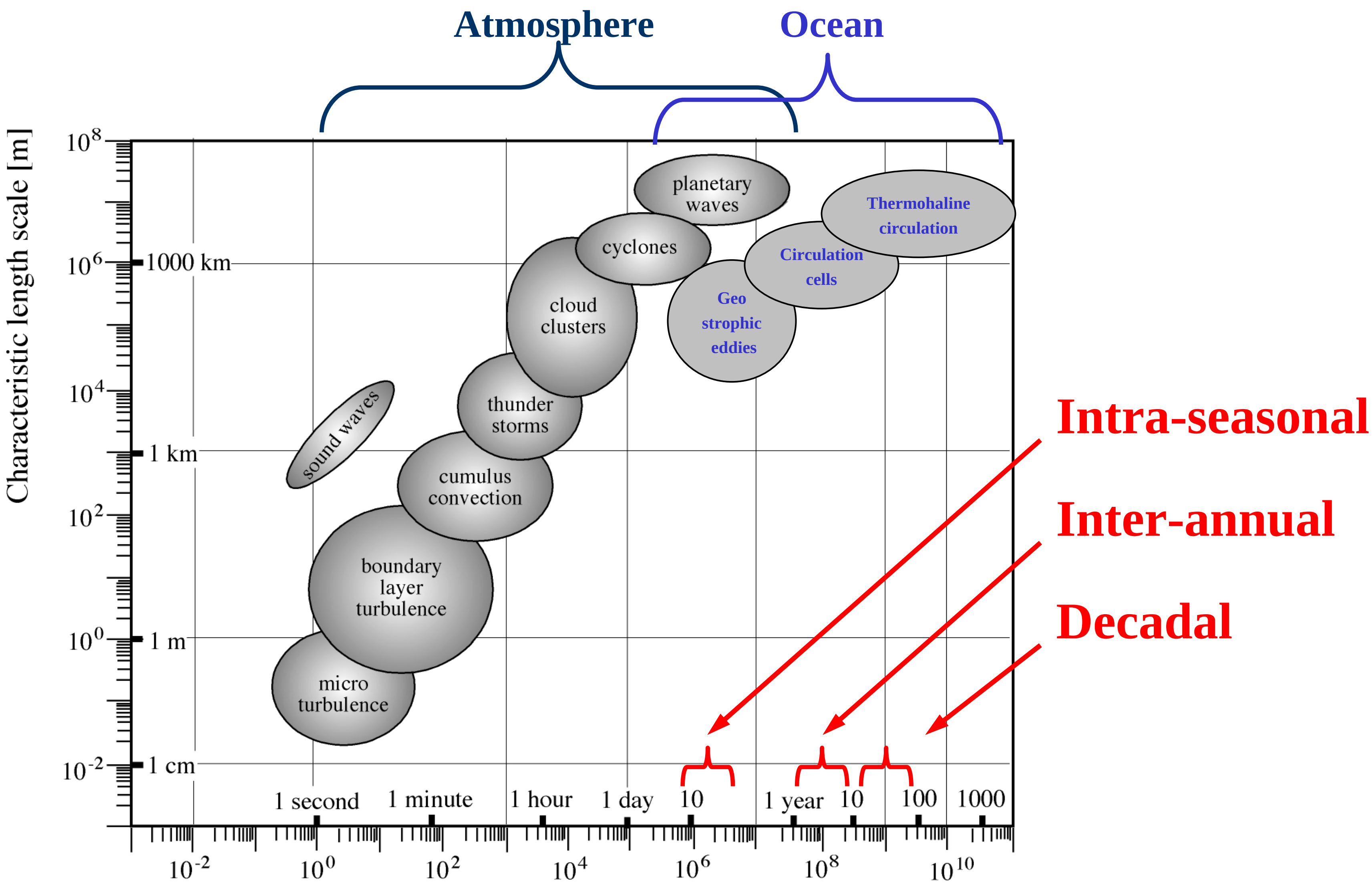
- Instabilities play a major role

→ Sensitivity to initial conditions (chaos)

→ The system is unpredictable

- The system is also dissipative – things cascade to smaller and smaller scales. This means the system does not blow up – can talk about stationary statistics.

A continuum of time and space scales:

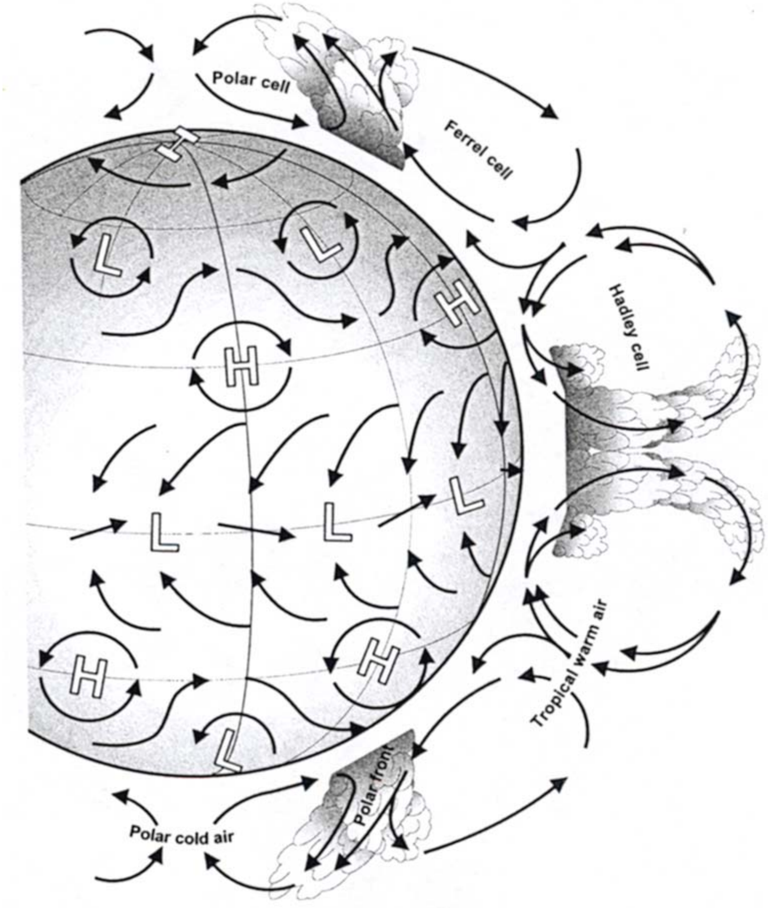


Courtesy of H. von Storch

Atmospheric dynamics → synoptic time scales

Tropical and midlatitude circulations are governed different types of motion. The dominant time scale, at least in midlatitudes, is synoptic (baroclinic storms, Rossby waves).

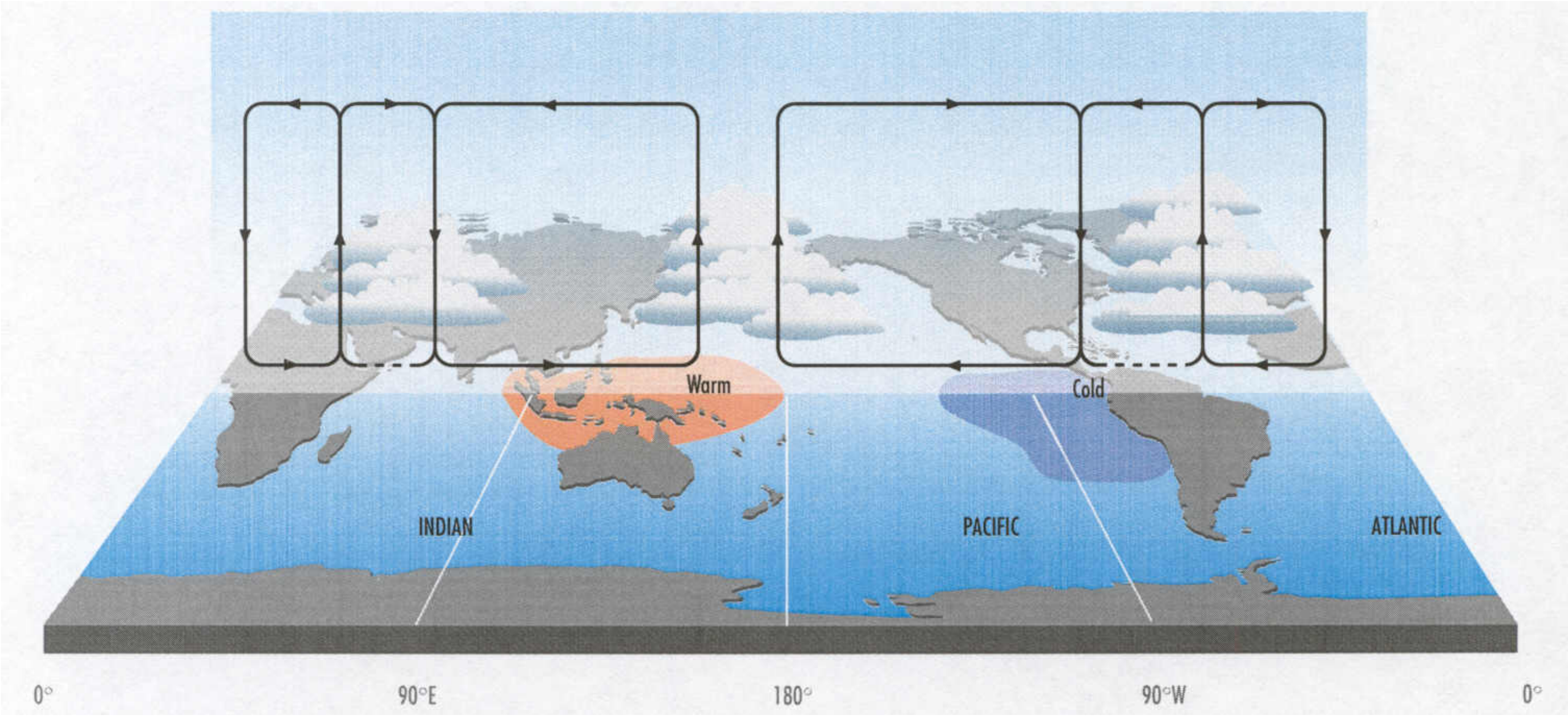
Yields a dynamic, variable state, with steady statistics



Tropical atmosphere-ocean interaction → interannual variability

An interaction between the atmosphere and ocean:
 - Atmospheric walker circulation and tropical Rossby/gravity waves
 - Oceanic wind driven and thermohaline circulations, Rossby waves on the thermocline lower boundary

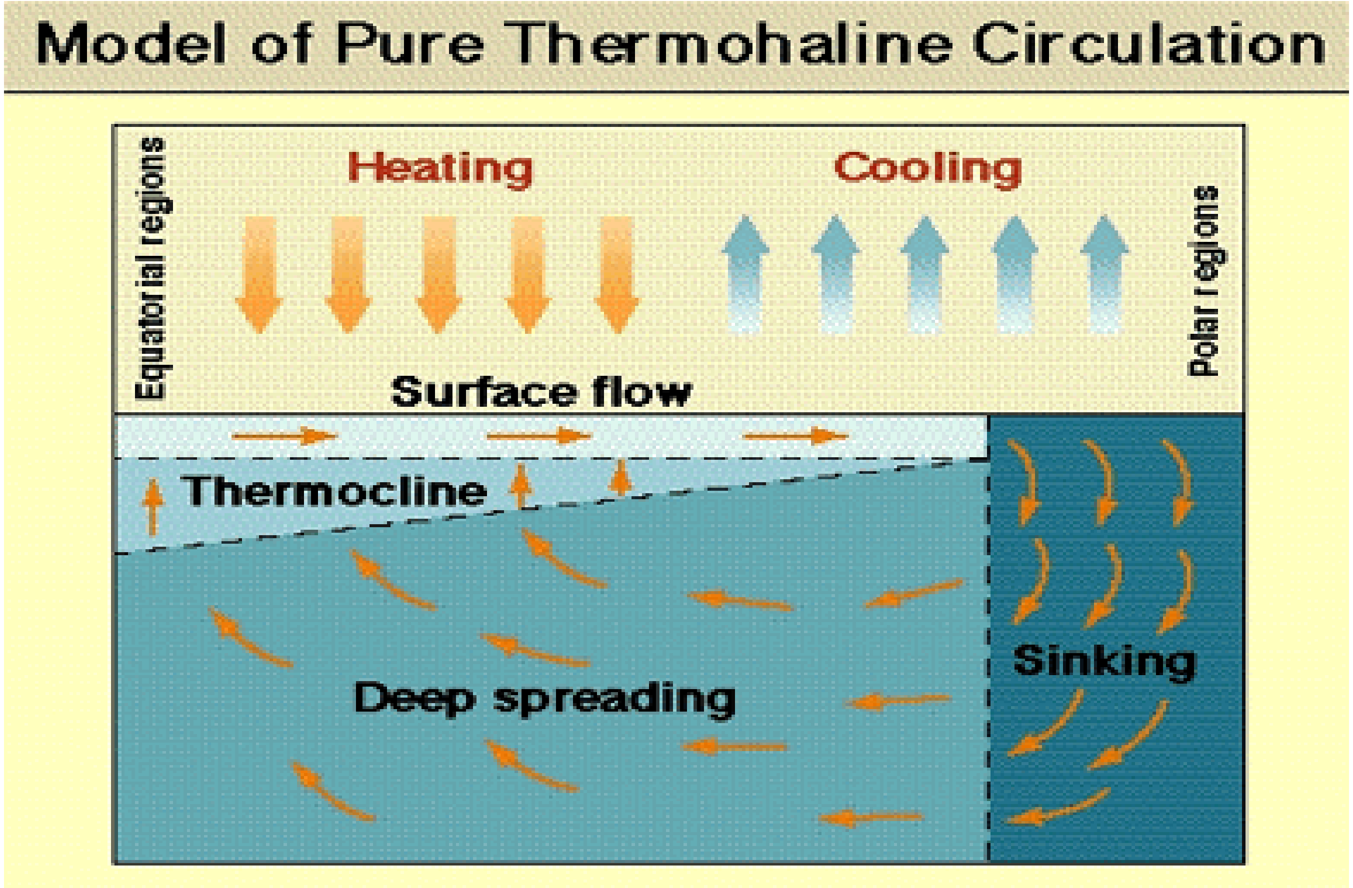
→ yields climate oscillations, e.g. ENSO



The oceanic circulation → inerannual to decadal variability:

- Exchanges between the atmosphere and the ocean surface
    - Transfer of water, energy and momentum
    - Fluxes couple the ocean and the atmosphere
- Vertical mixing between the surface and the deep ocean

A *"dynamic"* mean climatic state

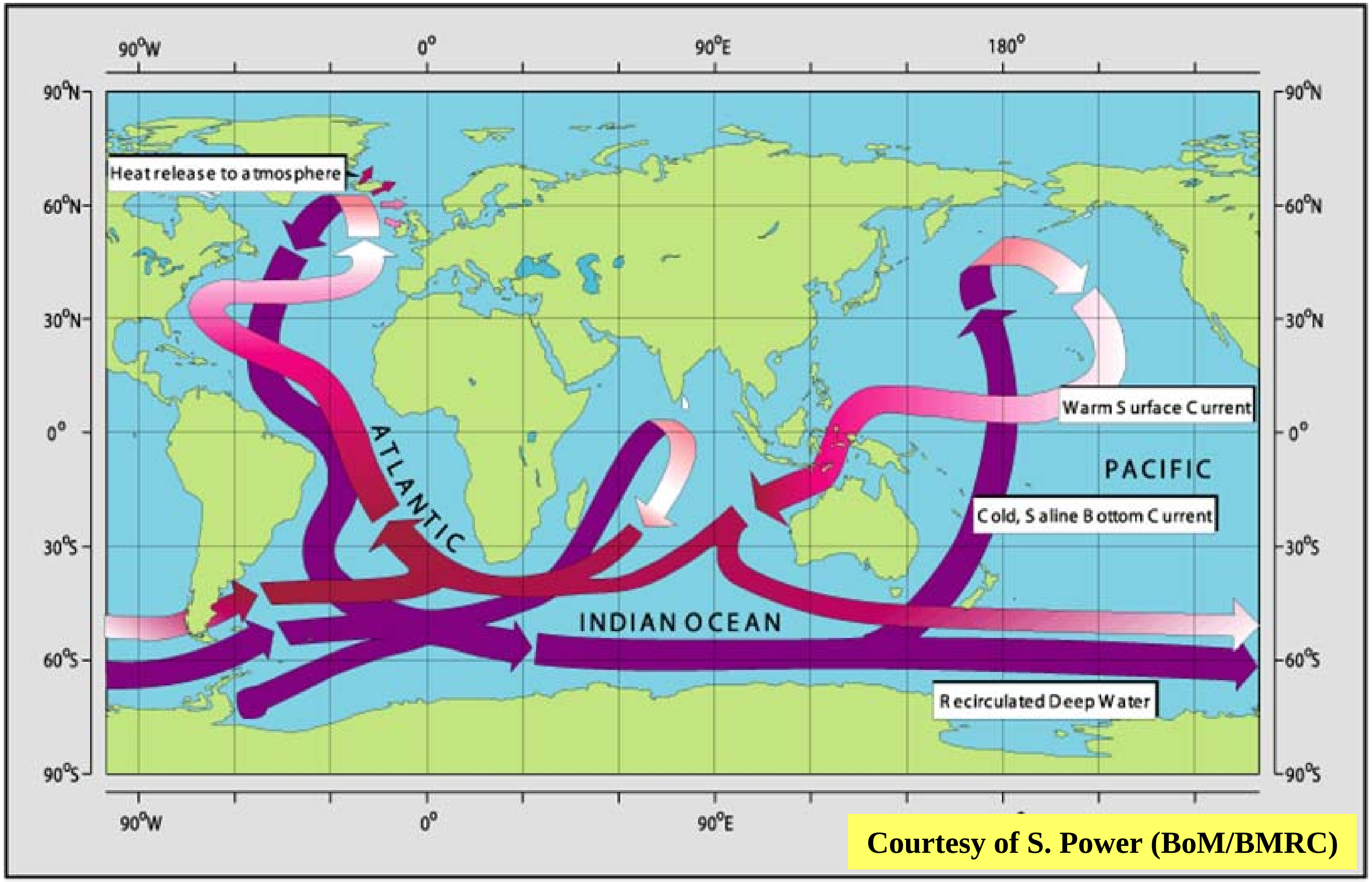


Courtesy of
S. Power
(BoM/BMRC)

The deep oceanic circulation → decadal to centennial variability

- Transfer between the surface layer and the deep ocean
  - *Thermohaline* circulation or *"conveyor belt"*
  - Oceanic mixing: sink of energy and CO_2

This is the "*memory*" of the climate system



Courtesy of S. Power (BoM/BMRC)

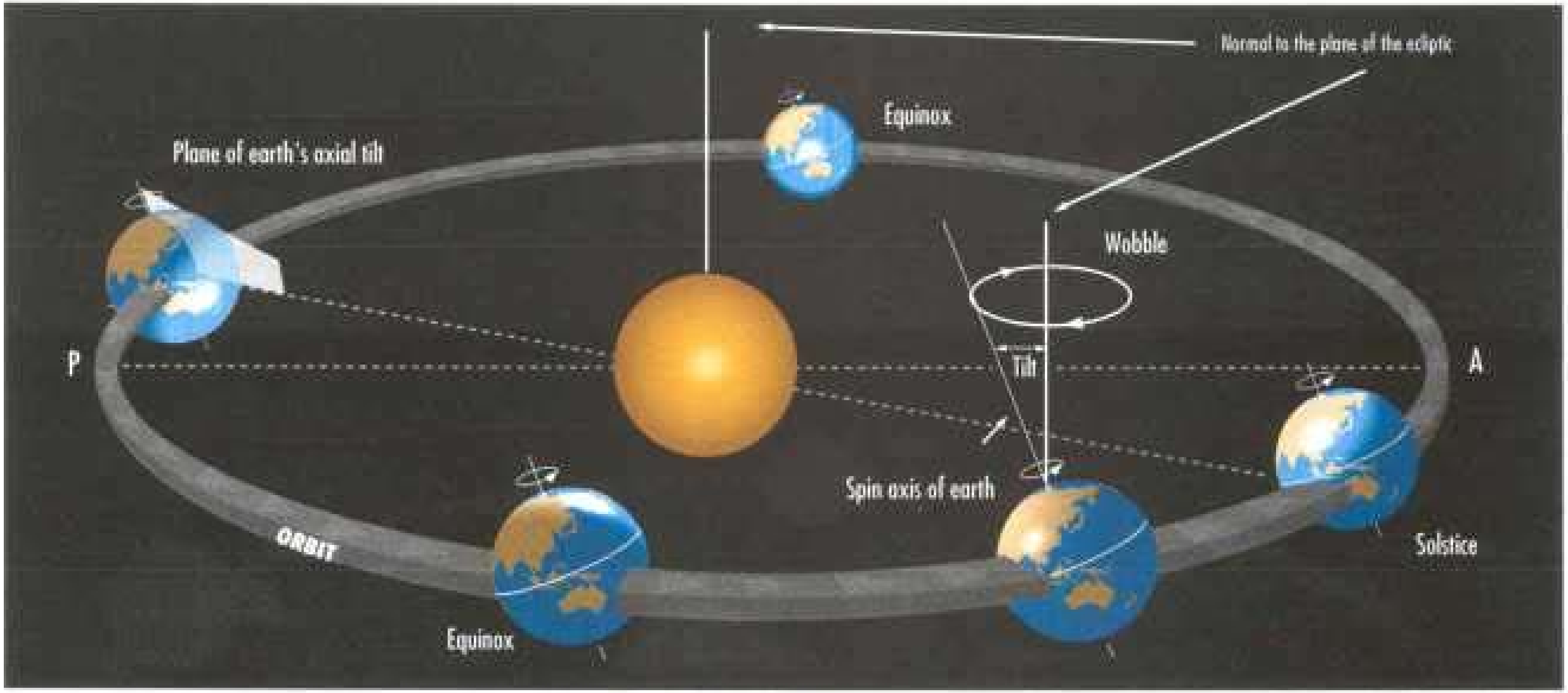
Externally forced variability:

Changes in the distribution of solar energy:

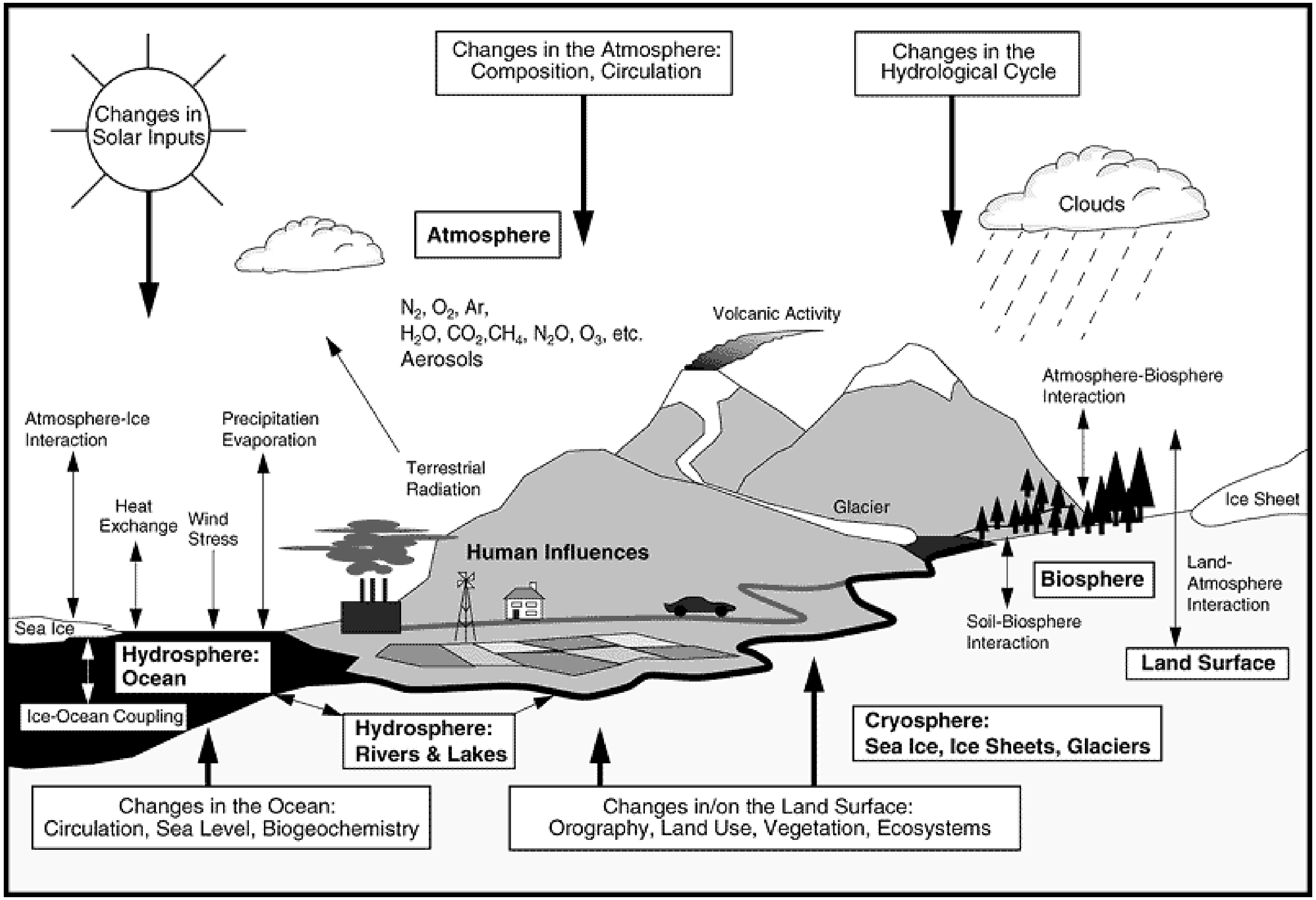
- •Earth spinning on its axes (24 hours)
- • Earth rotation around the sun (365.25 days)
- • Milankovitch: variations in orbital parameters:
  - • Eccentricity: Distance Perihelion-Aphelion varies (100,000 years)
  - • Axial Tilt varies (21.5 to 24.5º)- seasonal contrast (41,000 years)
  - • Precession wobbling around axis – seasonal contrast (23,000 years)

Slight variations in intensity varies (sun spots) (11 years)

Diurnal
Seasonal
Geological (i.e. glaciation)
Solar



Human influence on the climate system → trends



- Affects the entire climate system
- Impacts on Atmo. & Ocean. circulations
- Large enough to be significant & detected
- But hard to differentiate from natural long term variability

Intergovernmental Panel on Climate Change (IPCC), Third Assessment Report (TAR), 2001, Fig. 1.1, pp81.

This complex system requires a sophisticated observing system:

- Basic observations are carried by national agencies
- Manual observations (volunteers, weather offices)
- Automatic Weather stations (AWS)
  - → Rainfall, Temperature, Pressure, Wind and other parameters

- Upper-air soundings are routinely performed (sonde, profiler, radar)
- Satellites measurements (Fluxes, Clouds, Wind, Surface properties
- Scientific campaigns for specific measurements

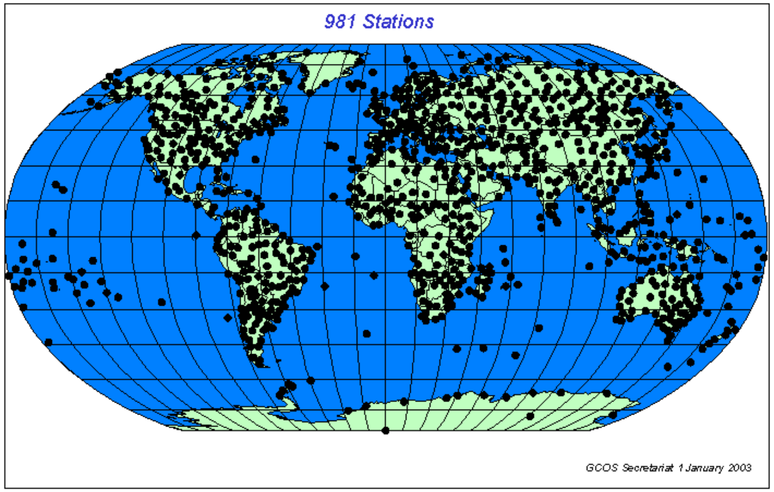
↳ The resulting product is used for weather prediction and archived for climate studies

The Global Climate Observing System (GCOS)

An international effort, sponsored by WMO, UNESCO, UNEP, ICSU

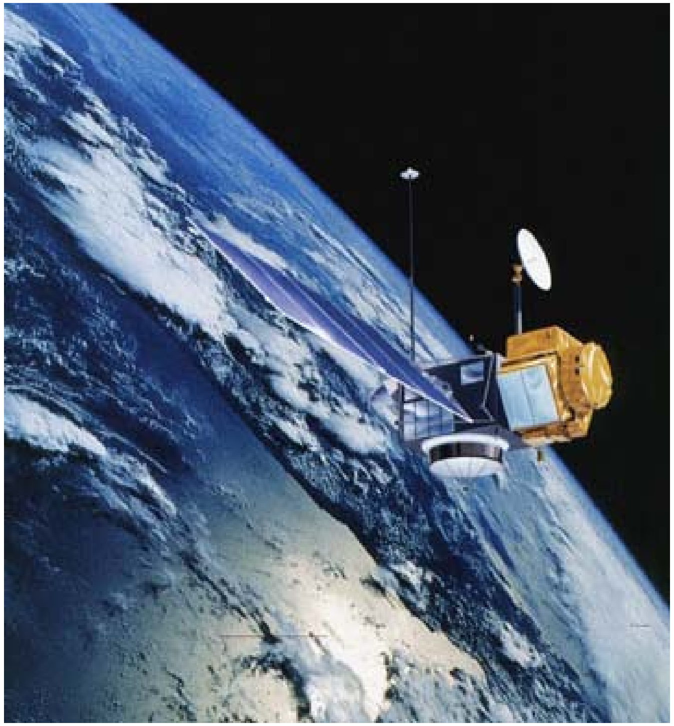
- Data collection and archiving from international observing networks
- Web site: http://gosic.org

GCOS Surface Network (GSN)

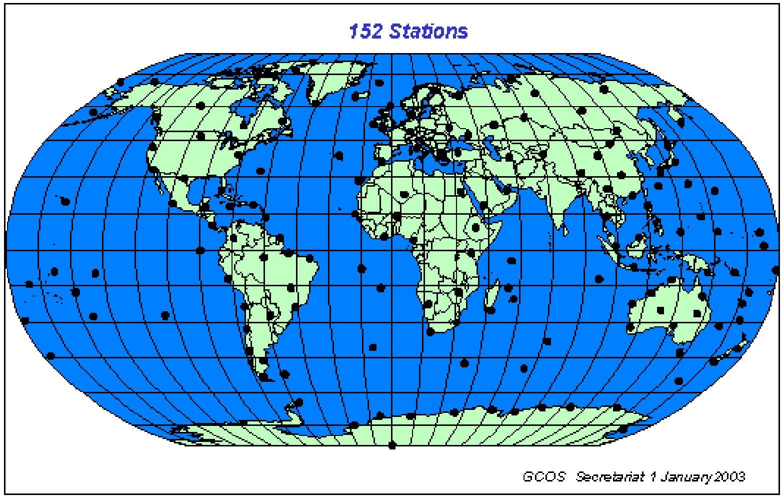


Atmospheric observations:

- Surface network
- Upper air network
- Completed by satellite observations



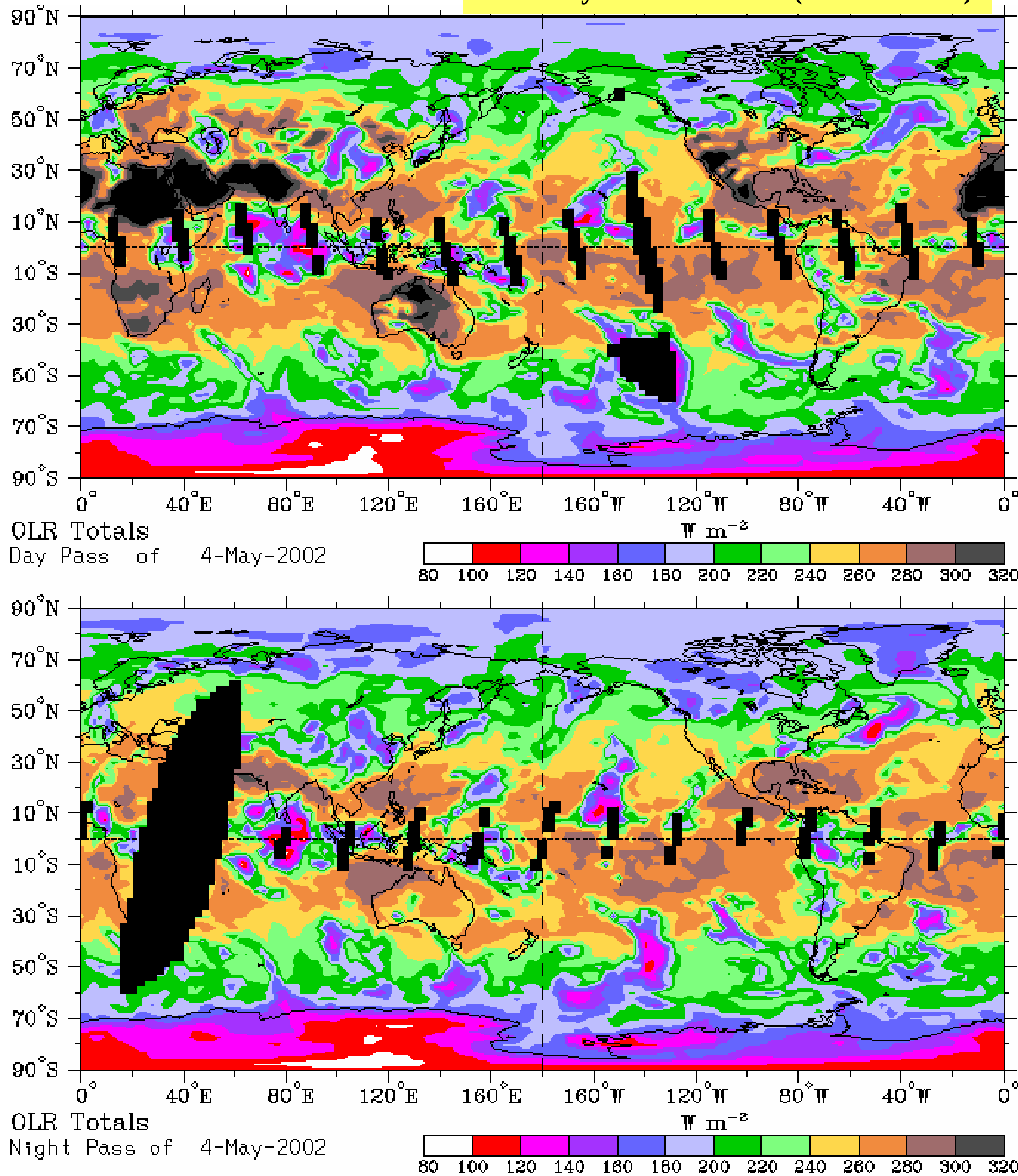
GCOS Upper Air Network (GUAN)



An example of satellite observations:

Courtesy of M. Wheeler (BoM/BMRC)

Outgoing long wave radiation (OLR) from the NOAA polar-orbiting satellites



Issues:

- **“Proxy” only for rainfall**
- **Calibration**
- **Consistency (instruments)**

Daily averages are made from the two satellite passes made each day.

Available daily since 1978.

Chemical composition of the Atmosphere:

- Understanding the chemical cycles themselves (GHG, pollution)
- Input into coupled atmosphere/ocean/chemistry and models.

The Global Ocean Observing System (GOOS):

Ships

Autonomous floats

ARGO

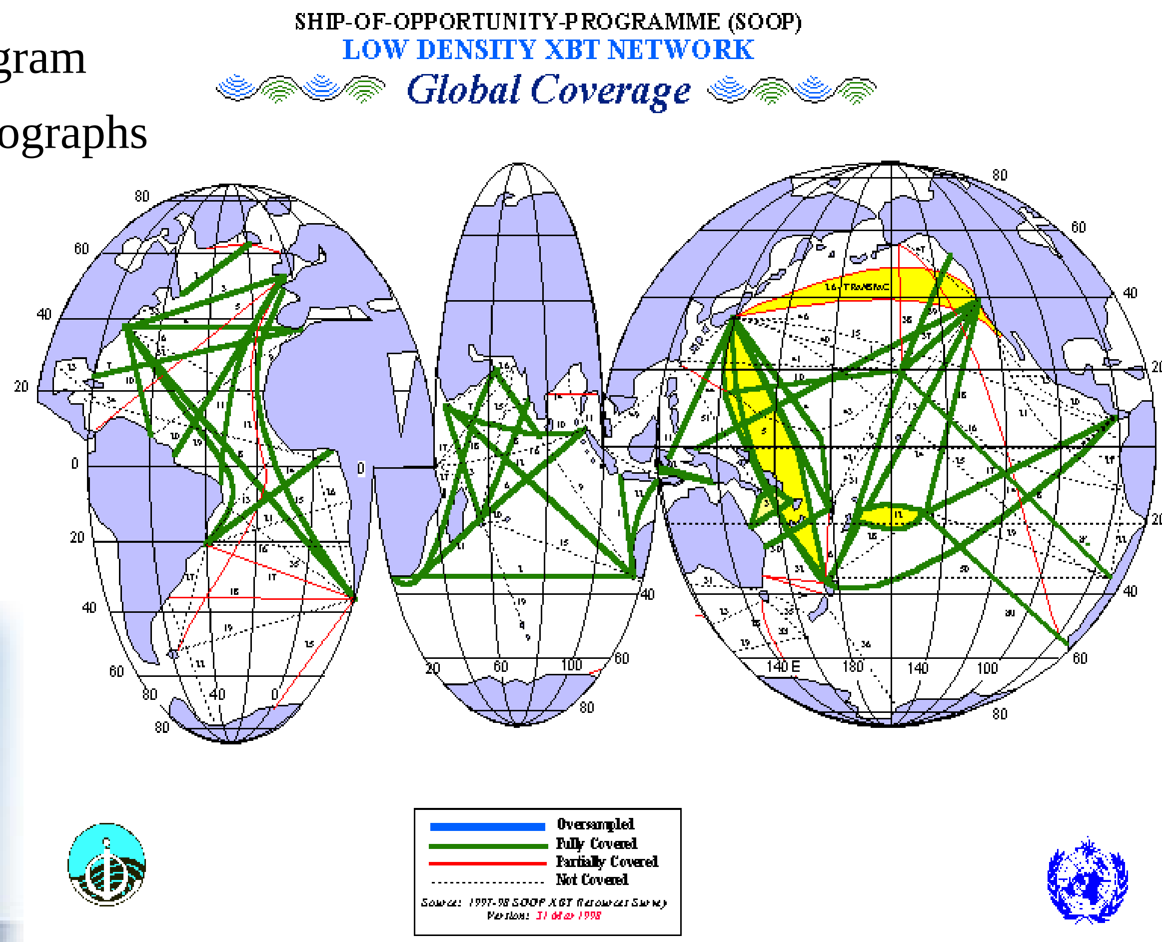
Buoys

Satellites

Courtesy of O. Alves (BoM/BMRC)

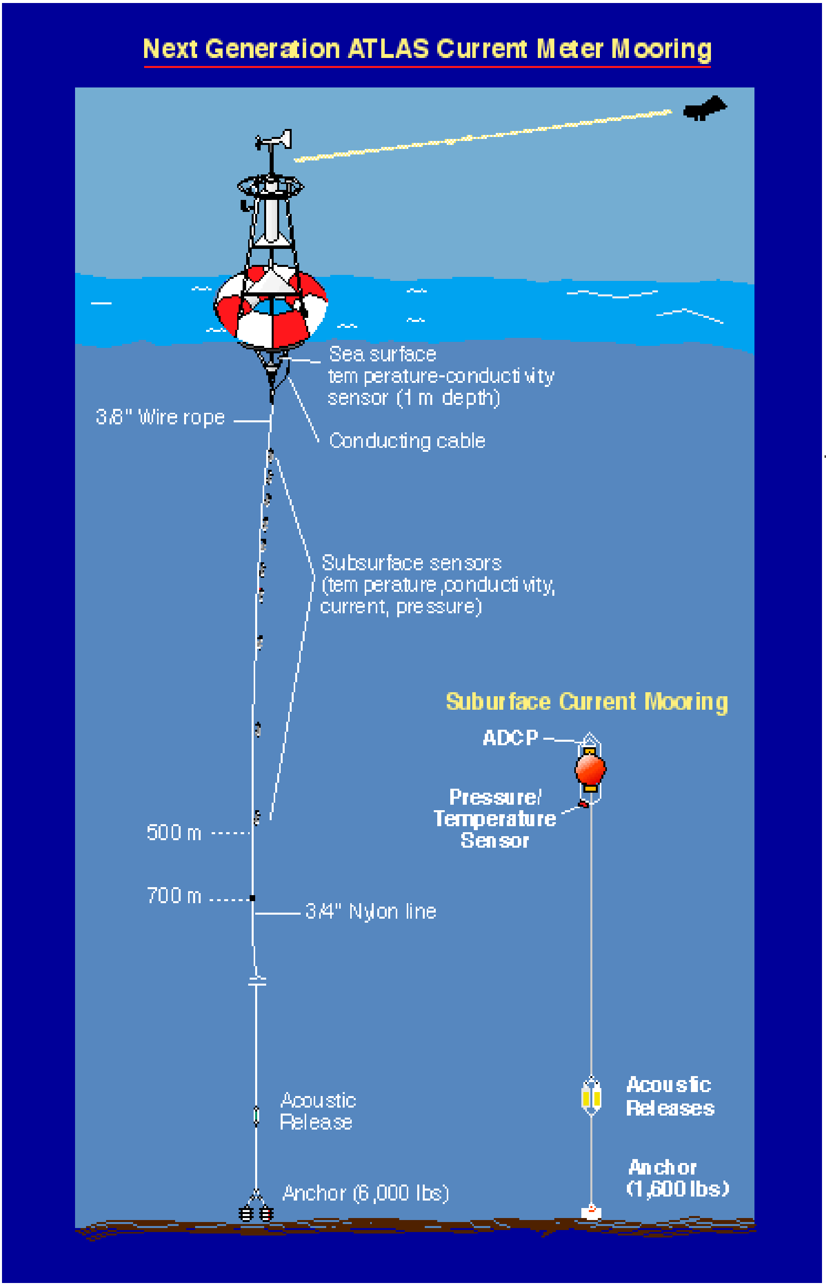
Ship based observations:

- Mainly along shipping lanes
- SOOP - Ship of opportunity program
- XBTs eXpendable Bathy Thermographs
- Research Vessels
- Military

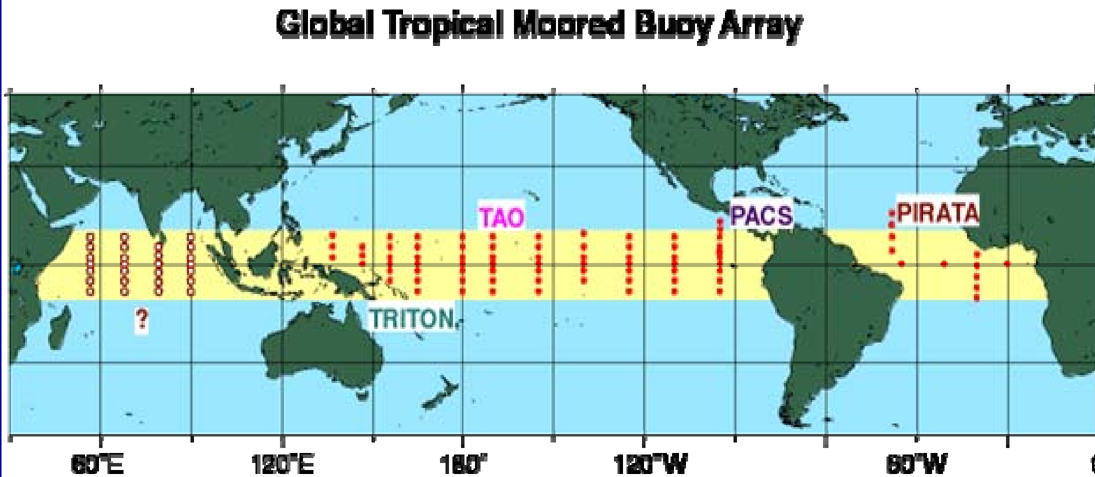


Courtesy of O. Alves (BoM/BMRC)

Moored Buoys in the Tropical Oceans:



- TOGA - Tropical Ocean Global Atmosphere
- Measure Temperature (Salinity) profiles
- Atmospheric Variables
- Daily in real time via satellite



Pacific: TAO array started late 80's, complete in 1992

Atlantic: PIRATA array Started late 1990's

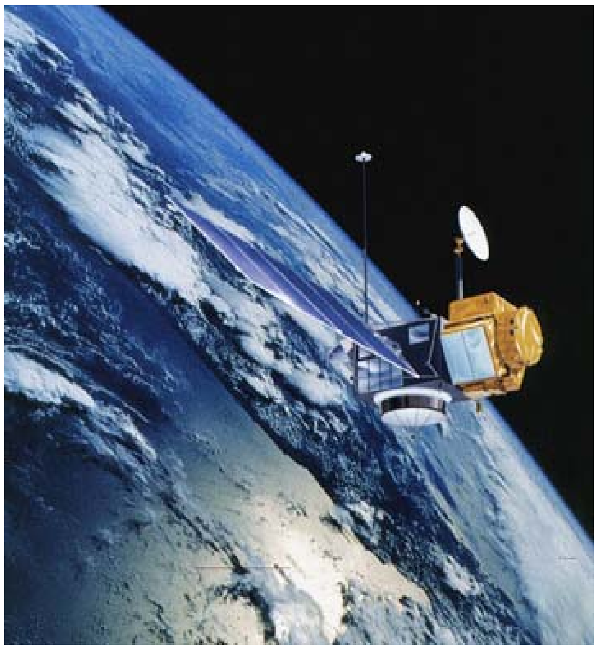
Indian: Future ?

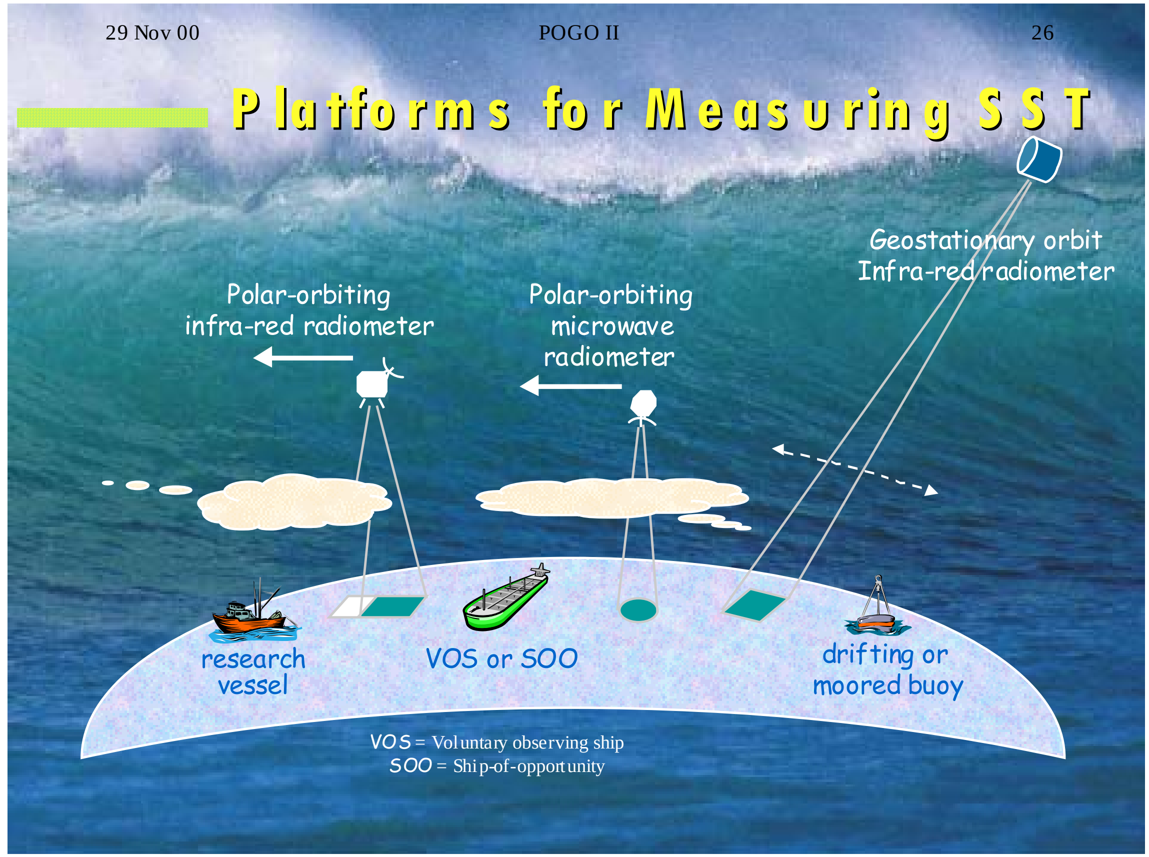
Real time monitoring: **http://www.pmel.noaa.gov/tao**

Courtesy of O. Alves (BoM/BMRC)

Satellite observation of the Sea Surface :

Satellites use electromagnetic radiation which cannot penetrate much below the surface:





- Radiometer - AVHRR Advance Very High Resolution Radiometer - Polar orbiting satellites - cannot see through clouds
- Microwave sensor - sees through clouds - lows resolution - Geo-stationary satellites

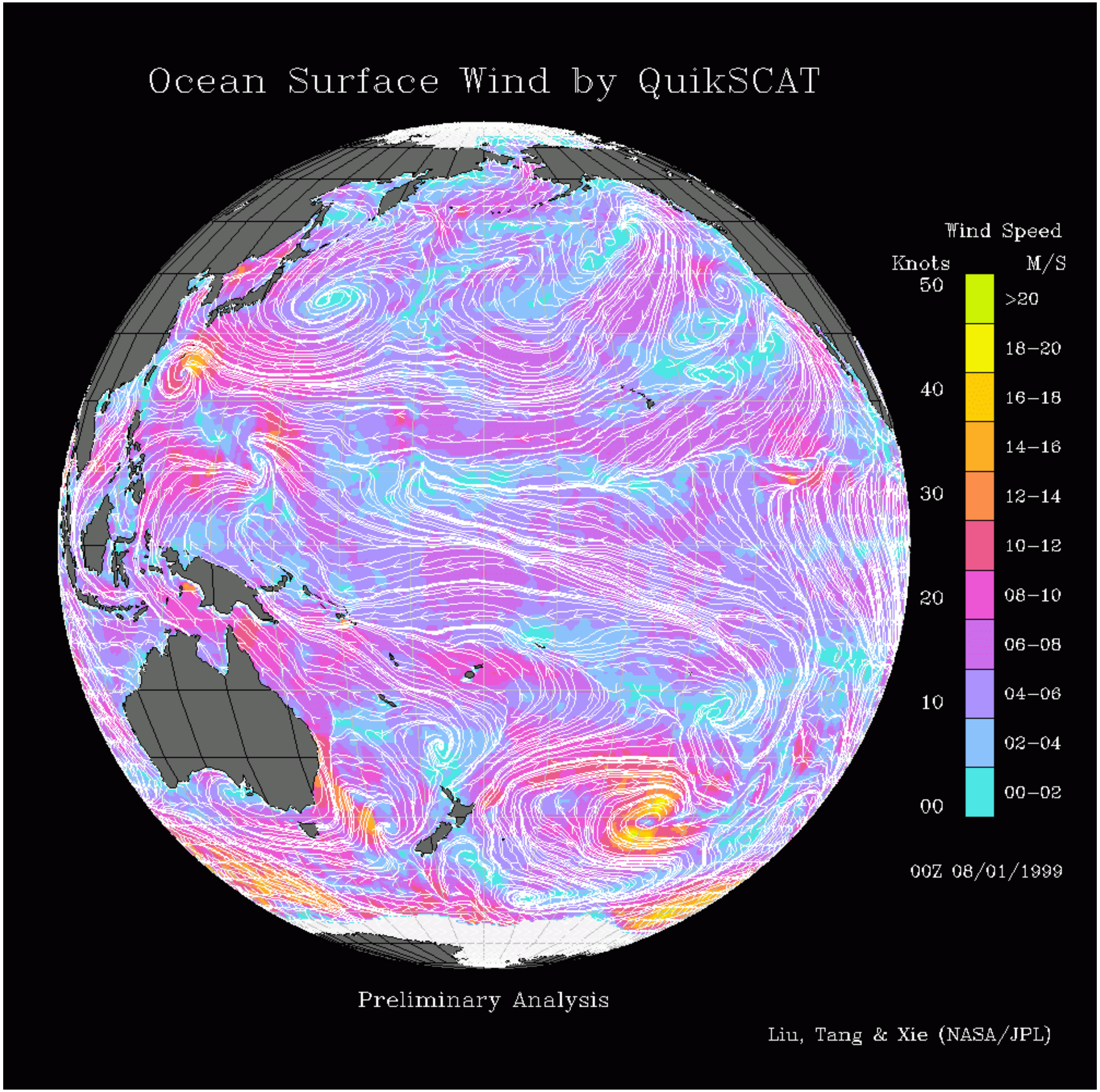
Courtesy of O. Alves (BoM/BMRC)

Observations:

• Sea Surface Temperatures (SSTs)

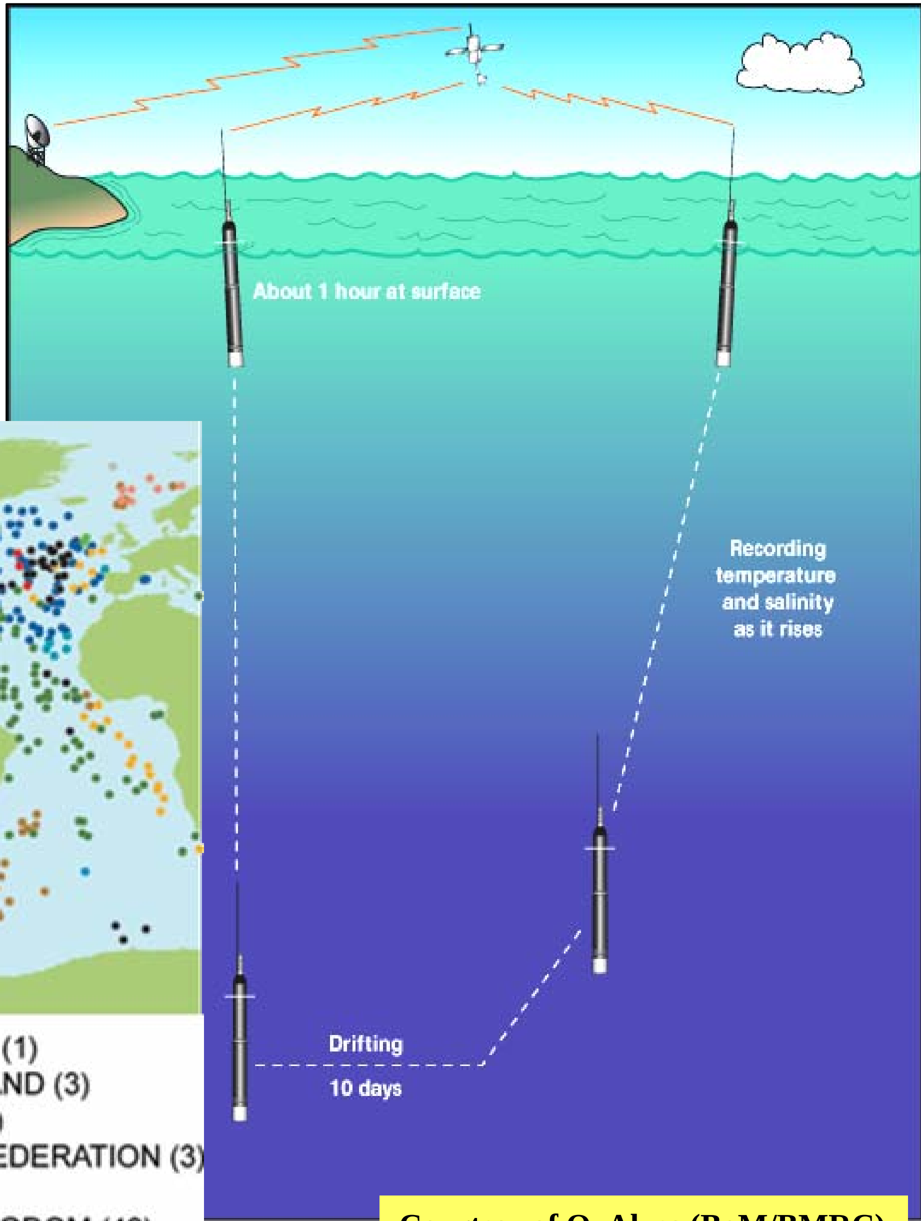
• Sea level (upper level thermal content)

• Surface wind (i.e. Quikscat)

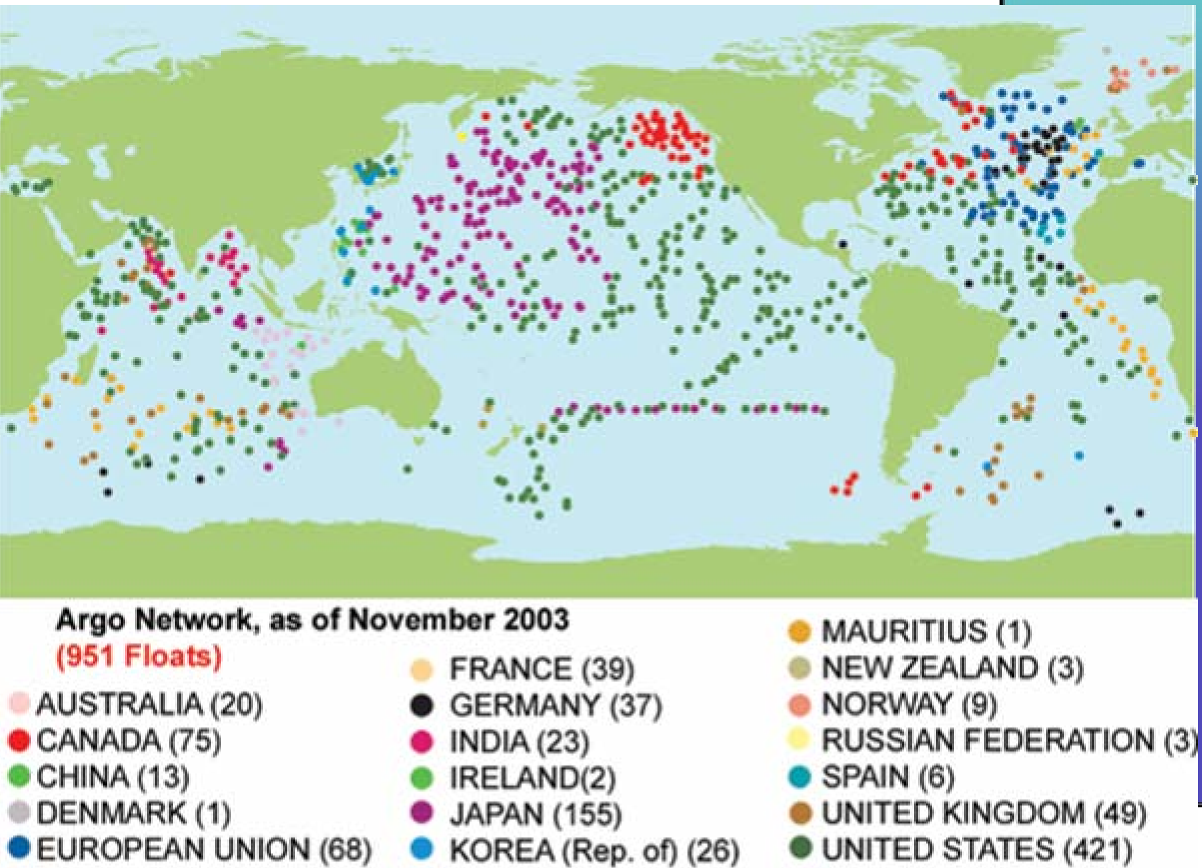


Vertical sounding: the Argo float

- Autonomous profiling floats
- Parked at about 2-3 km depth
- Surface every 10 days
- Measure temperature/salinity
- Transmit measurements in real time via satellite

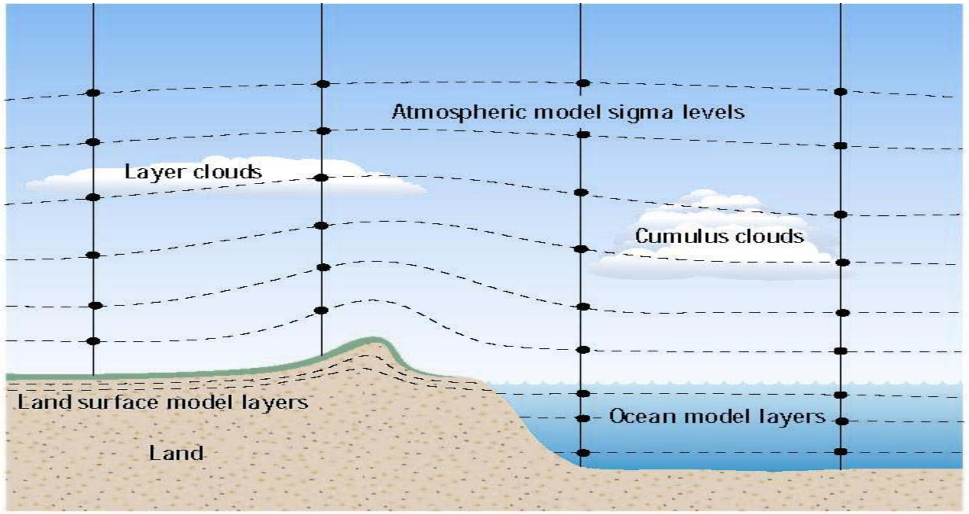


Courtesy of O. Alves (BoM/BMRC)



To get regularly spaced and sampled observations from all this data, we need an *analysis* system.

Use numerical models



Numerical models are used for:

- Numerical Weather Prediction (NWP)
- Climate simulation

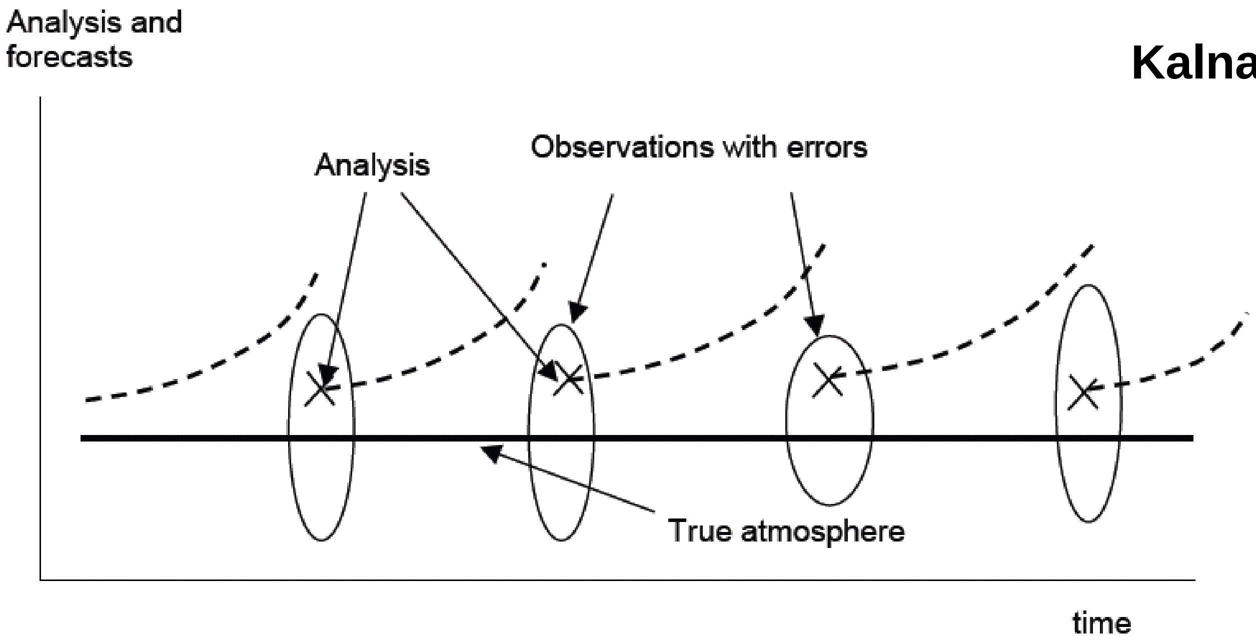
- Analysis maps are derived as part of routine numerical prediction.

- For long term analysis, the data is reanalyzed using the same procedures for the entire data set → Re-analyses: (e.g. NCEP-NCAR, ECMWF …)

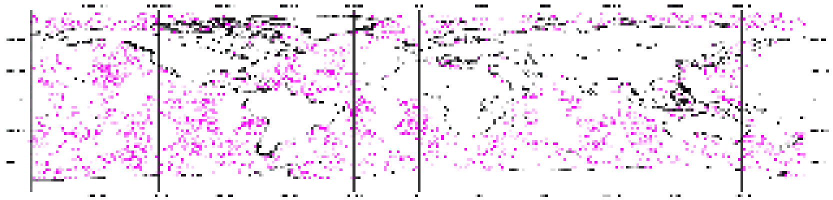
Analysis maps are derived as part of routine numerical prediction

Fig. 6.14: Schematic of the 6-hour analysis cycle. Indicated on the vertical axis are differences between the true state of the atmosphere (or its observational measurements, burdened with observational errors). The difference between the forecast and the true atmosphere (or to the observations) increases with time in the 6 hour forecast because of the presence of growing errors in the analysis (adapted from Kalnay and Toth, 1996).

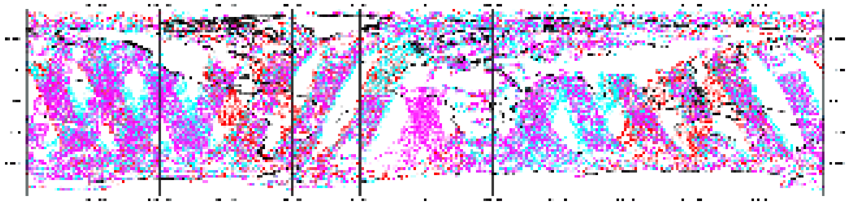
Kalnay, 2003



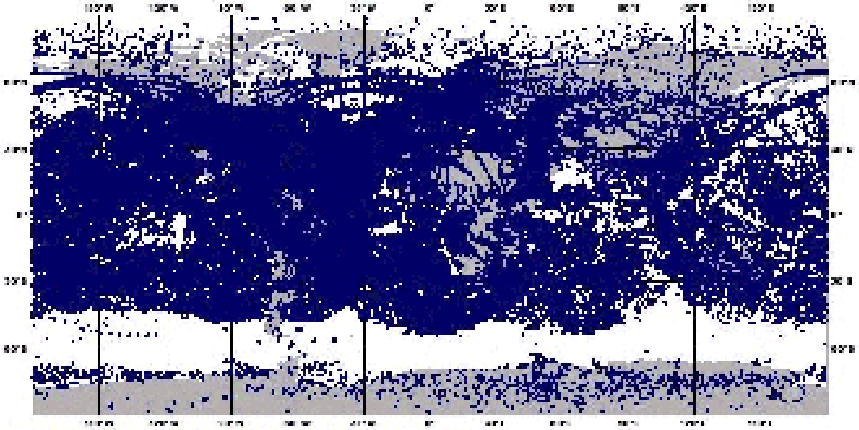
Data coverage for one day, used for ECMWF analyses



Lower tropospheric IR from AQUA satellite



Lower tropospheric microwave channel from 3 satellites (AMSU-A ch.5, NOAA-15, NOAA16m AQUA)



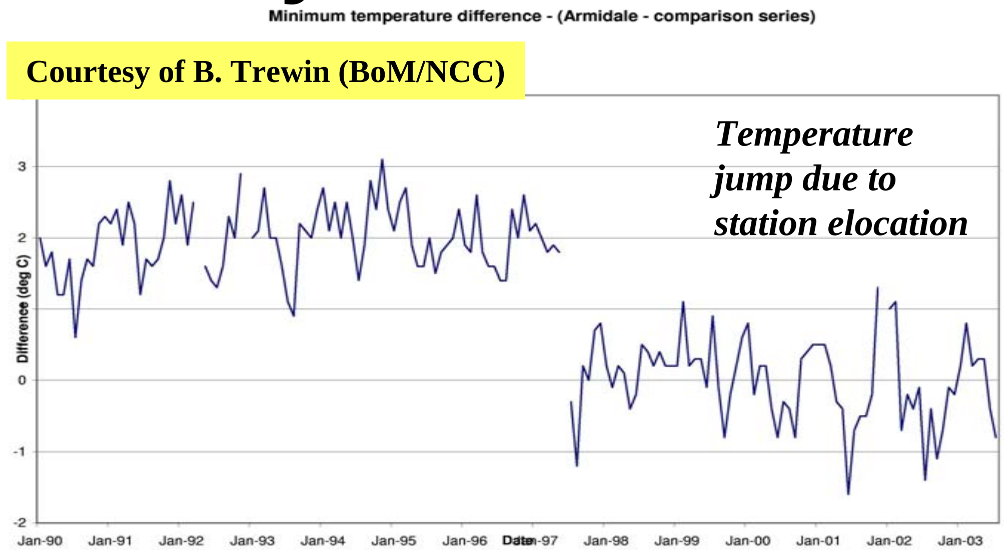
Conventional observations (surface , radiosindes, aircrafts) and satellite dervied winds.

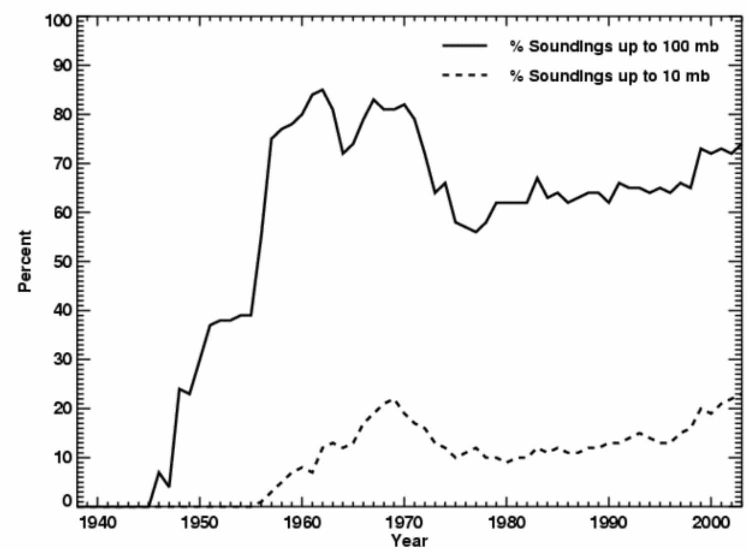
Note: there are still data sparse land areas

Climate data consistency

Long climate records are negatively impacted by:

- Station closure
- Shift of observation site
- Change of instruments
    - In situ observations
    - Upper-air sounding
    - Satellite measurement
- Change of surrounding
    - Urban growth
    - Land use





In this class we will use reanalysis data, which can be obtained online with a great software called ingrid:

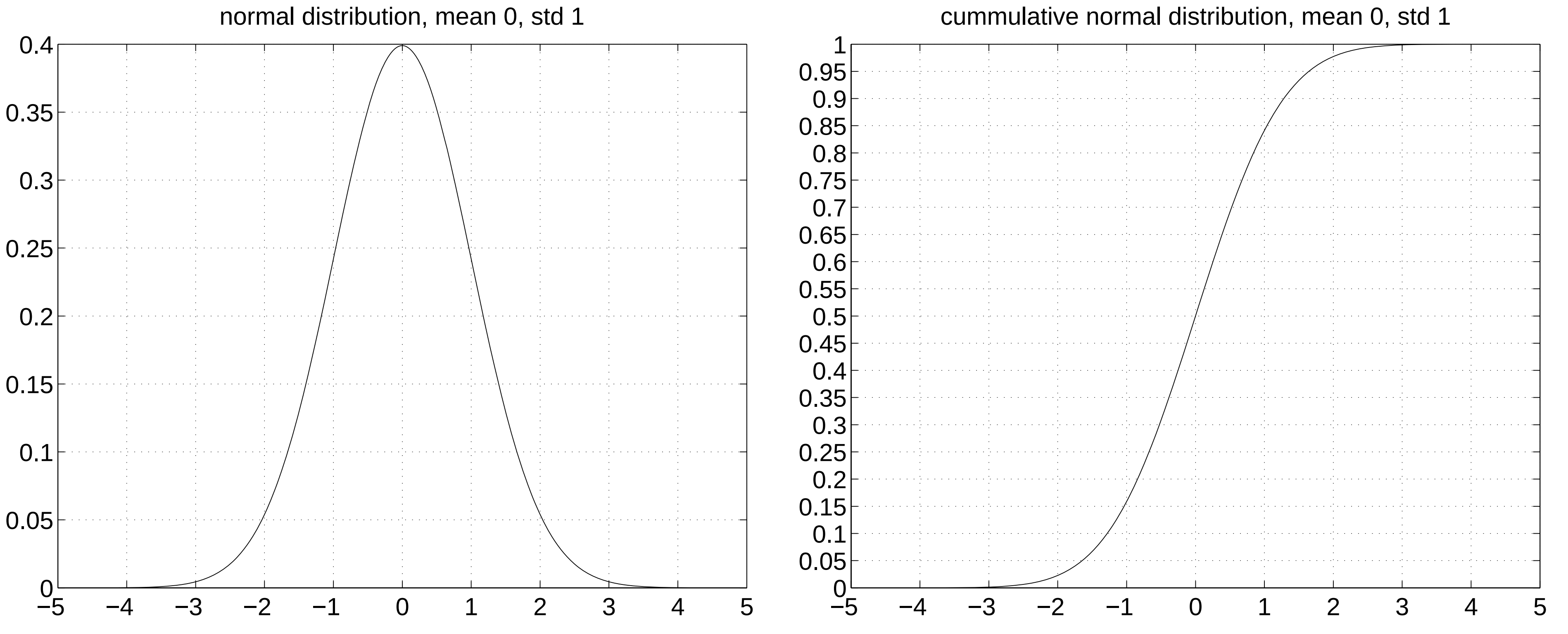
http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/

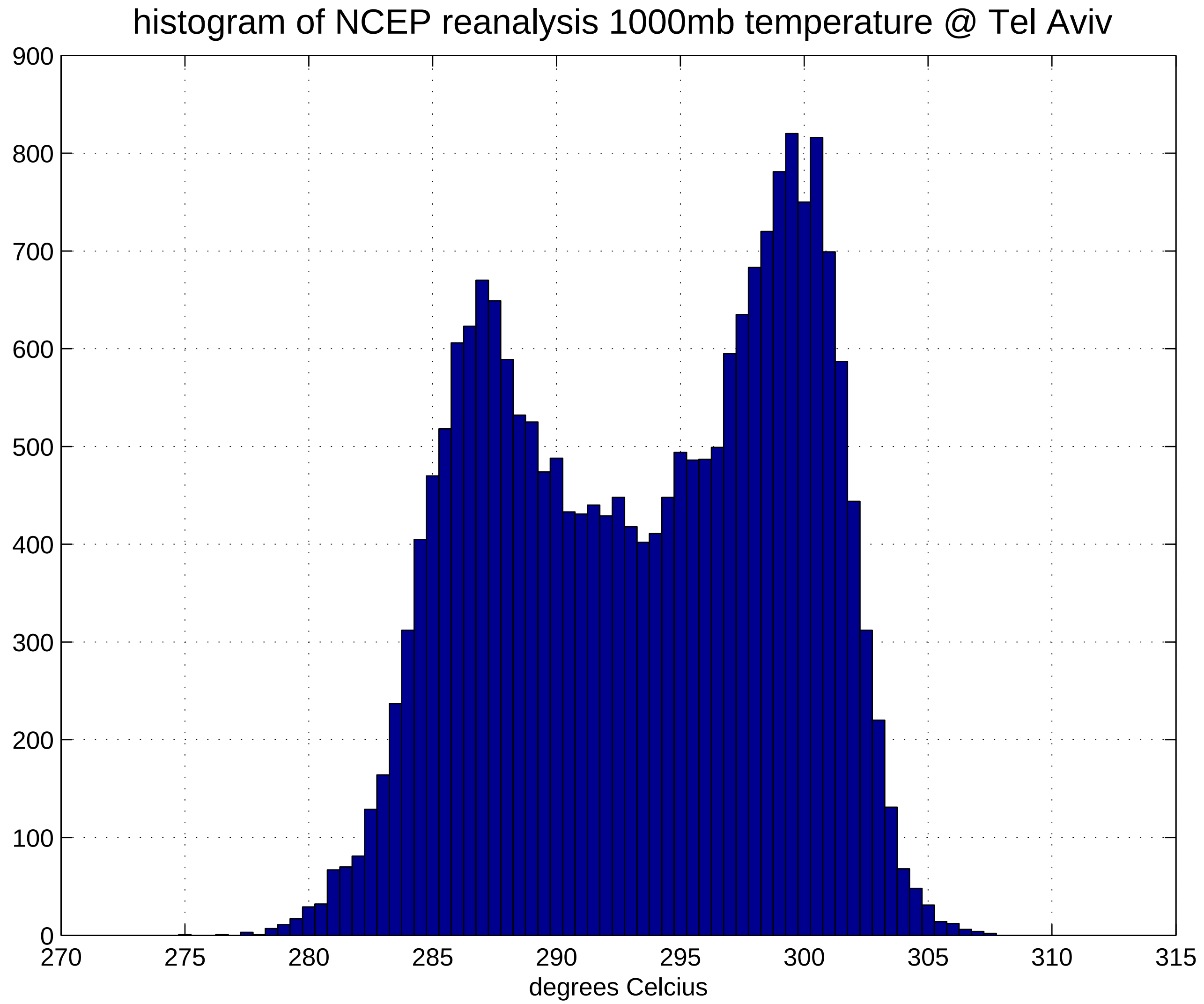
See also:

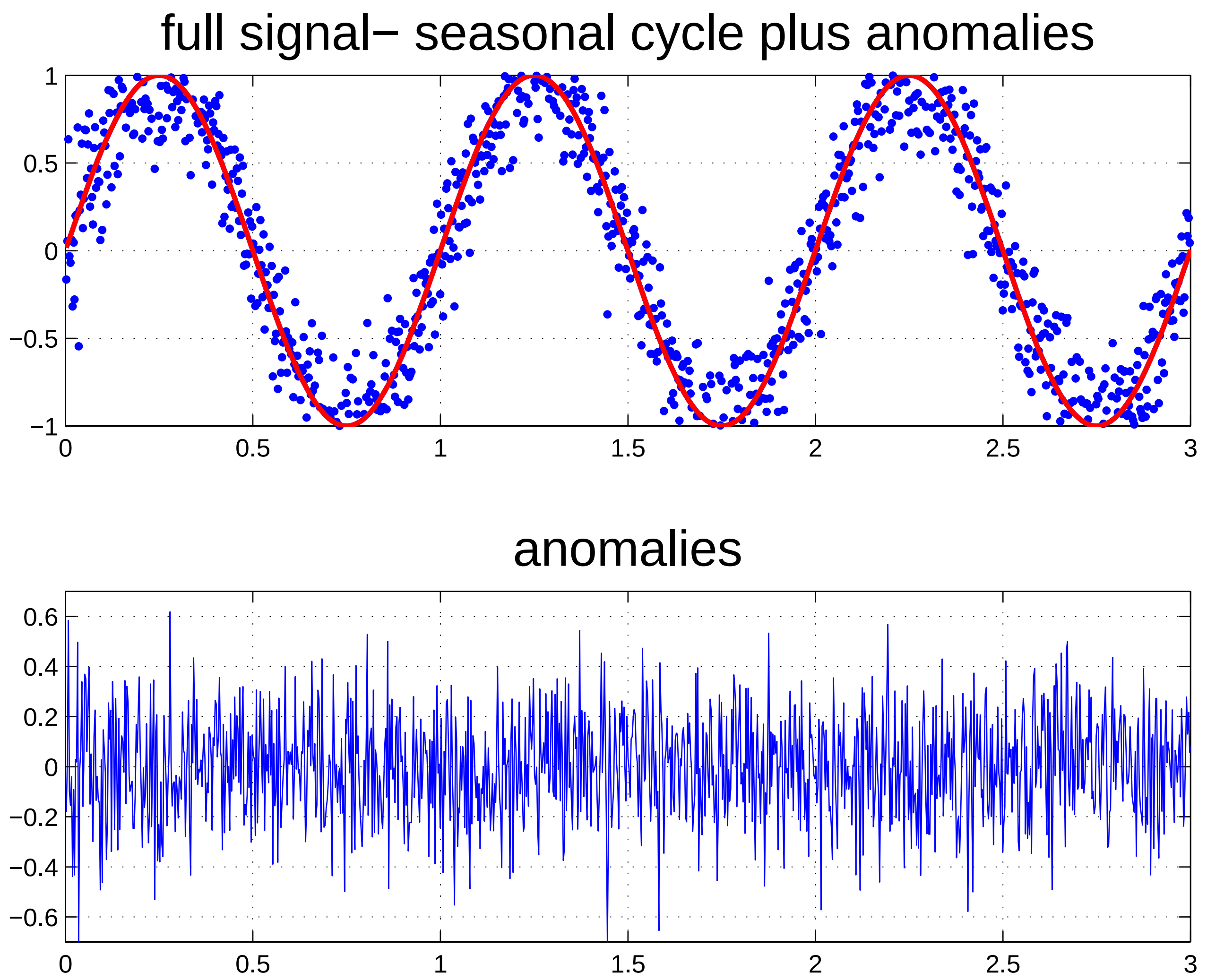
http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis.html

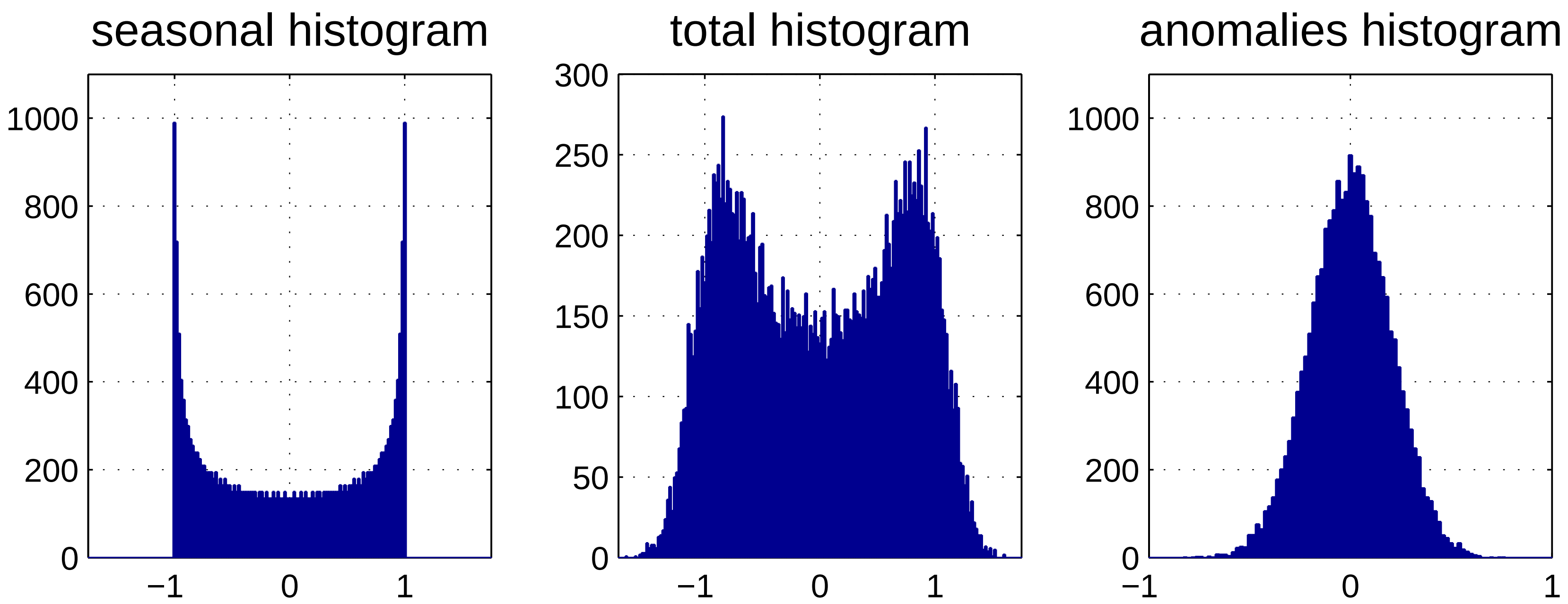
INGRID can also be used to download many types of data

http://ingrid.ldeo.columbia.edu/


normal distribution, mean 0, std 1
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−5
−4
−3
−2
−1
0
1
2
3
4
5
cummulative normal distribution, mean 0, std 1
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−5
−4
−3
−2
−1
0
1
2
3
4
5


histogram of NCEP reanalysis 1000mb temperature @ Tel Aviv
900
800
700
600
500
400
300
200
100
0
270
275
280
285
290
295
300
305
310
315
degrees Celcius


full signal− seasonal cycle plus anomalies
1
0.5
0
−0.5
−1
0
0.5
1
1.5
2
2.5
3
anomalies
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
0
0.5
1
1.5
2
2.5
3


seasonal histogram
total histogram
anomalies histogram
1000
800
600
400
200
0
−1
0
1
300
250
200
150
100
50
0
−1
0
1
1000
800
600
400
200
0
−1
0
1

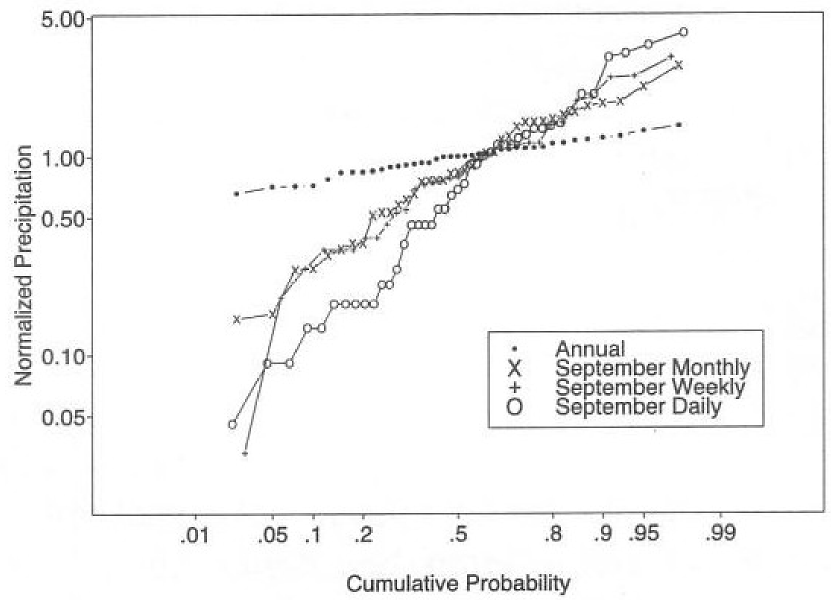
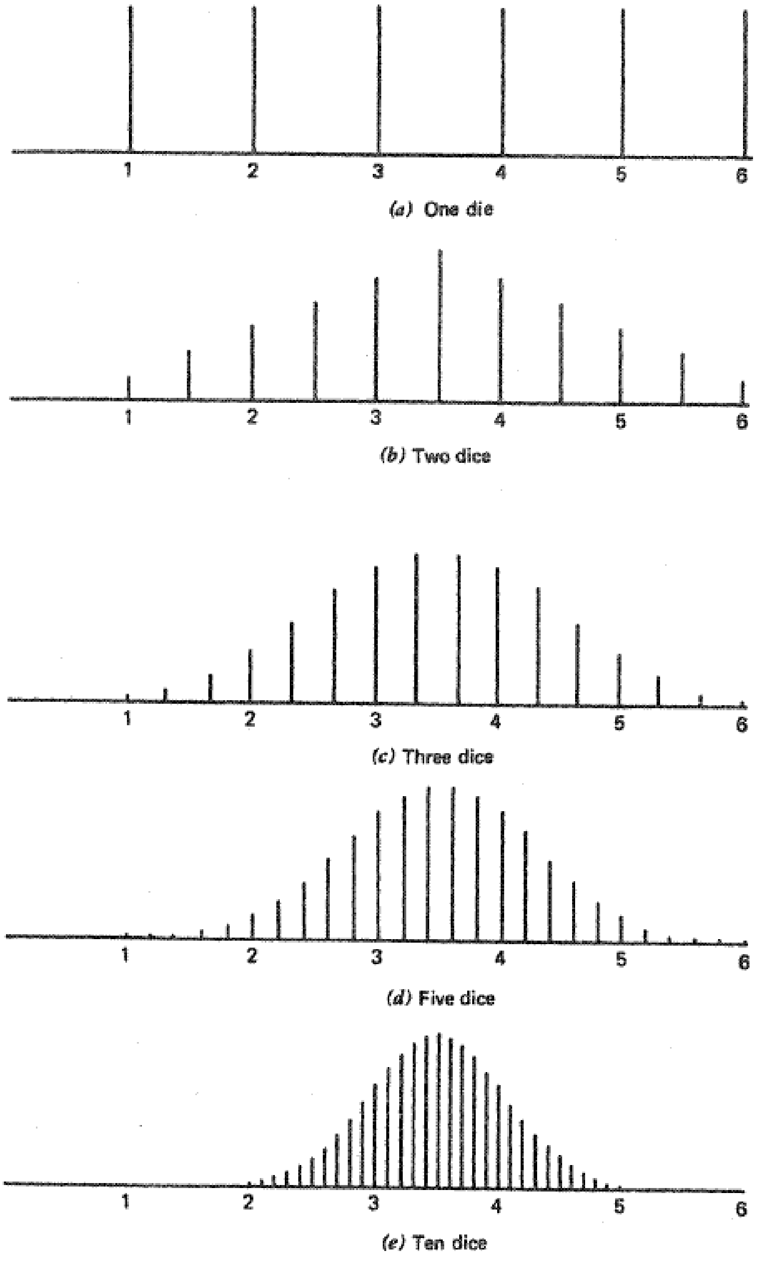
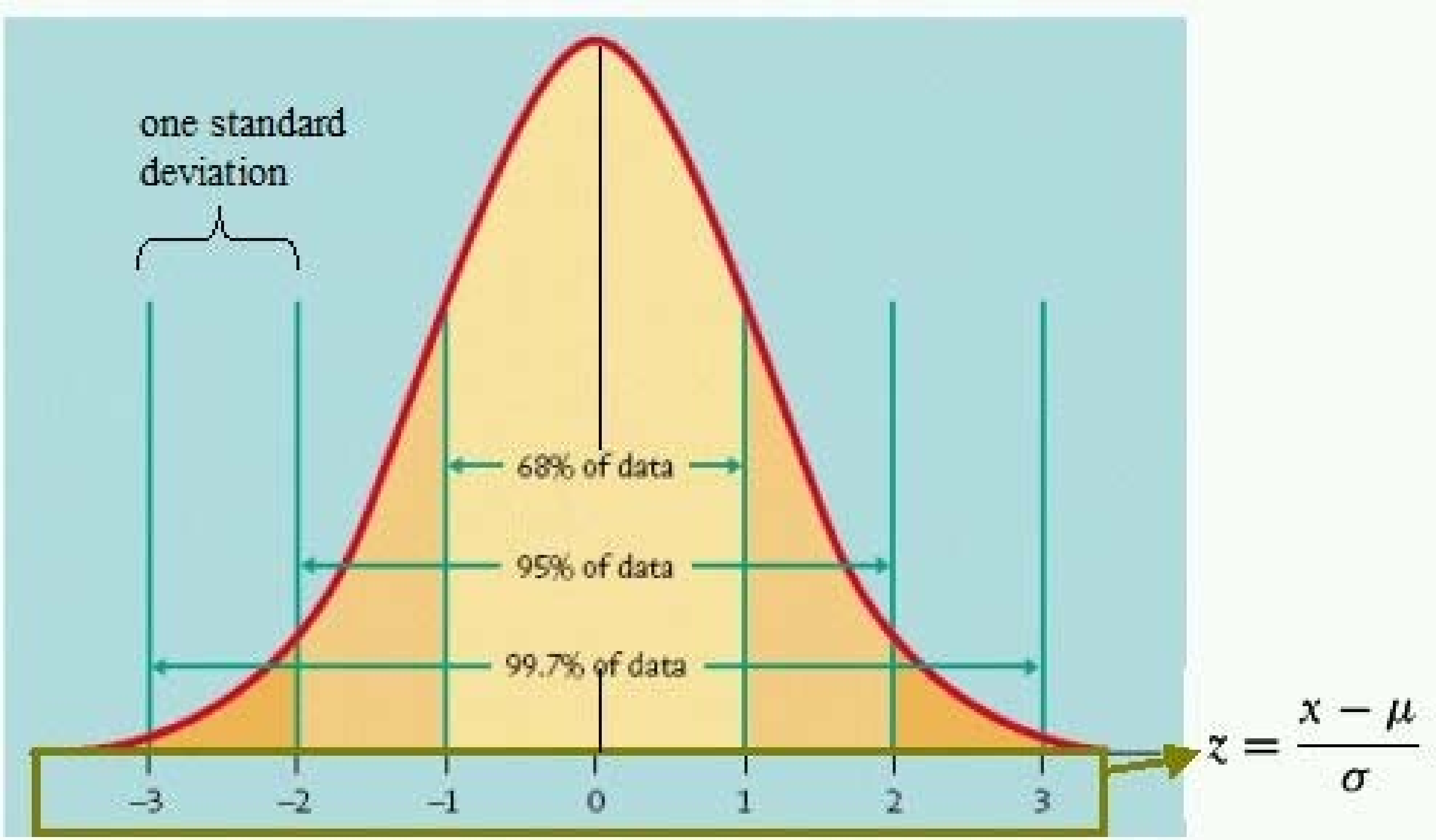


Figure 3.2: *Empirical distribution functions of the amount of precipitation, summed over a day, a week, a month, or a year, at West Glacier, Montana, USA. The amounts have been normalized by the respective means, and are plotted on a probability scale so that a normal distribution appears as a straight line. For further explanations see [3.1.3]. From Lettenmaier [252].*

von Storch and Zwiers, (1999)

Illustration of the
Central limit theorem





one standard deviation
68% of data
95% of data
99.7% of data
−3
−2
−1
0
1
2
3
$z = \frac{x - \mu}{\sigma}$

Student's t distribution