Online Training Course
On

Visualization of Marine Met Data (Using FERRET)

Organised by

International Training Centre for Operational Oceanography (ITCOO)
ESSO-INCOIS, MoES, Hyderabad, India

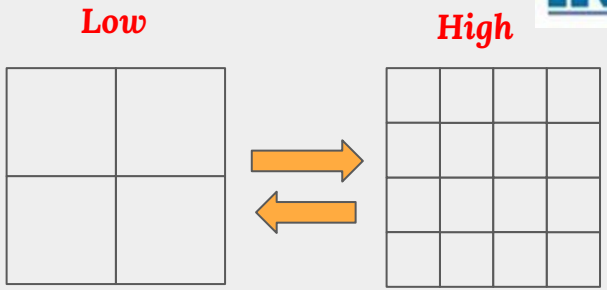
July 24 - 28, 2023

Sivaiah B & PSN Acharyulu
Project Scientist, s.borra-p@incois.gov.in

Contents

- Regridding Data
  - Space (Horizontal - XY; Vertical - Z)
  - Time (Time Series - T)

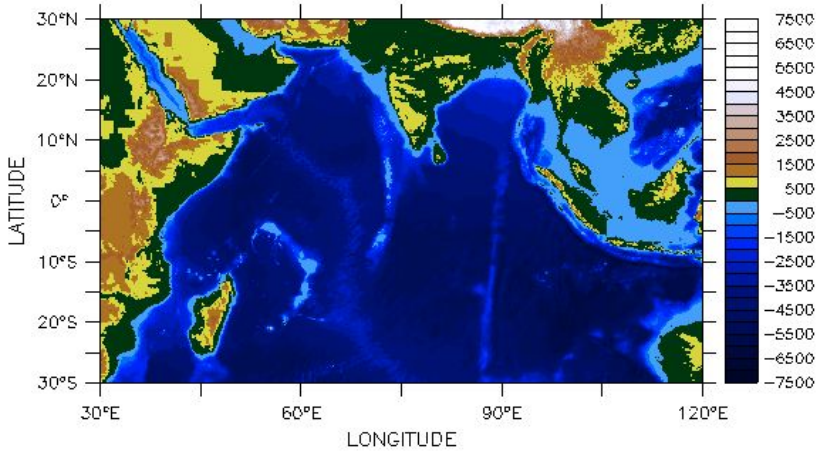
- Basic Statistics
  - Bias
  - Mean
  - Variance
  - Standard deviation
  - RMSE
  - Scatter Index
  - Correlation
  - Climatology
  - Anomaly
  - Regression Analysis
  - FFT
  - Low pass filter



Regridding Data

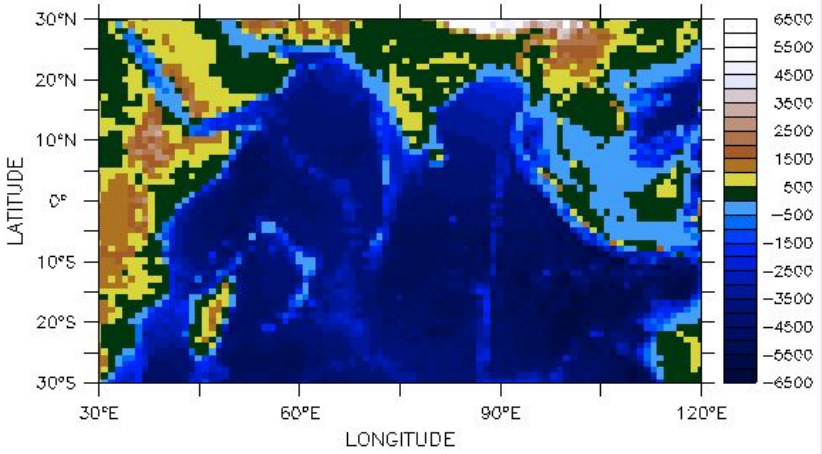
Space - Horizontal (XY)

High Spatial Resolution



Relief Of the Surface of the Earth (meters)

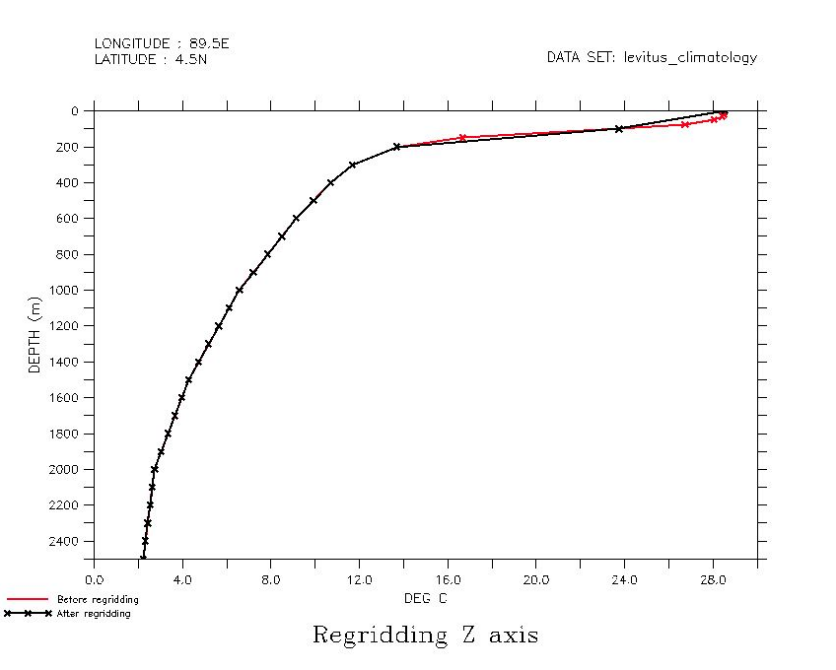
Low Spatial Resolution



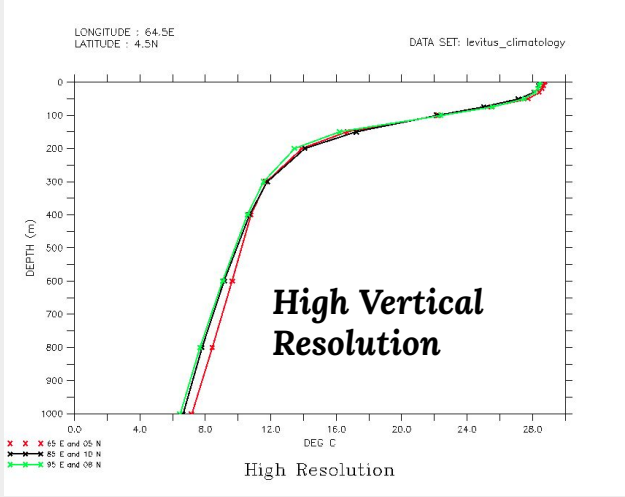
ROSE[G=NGRD]

Regridding Data

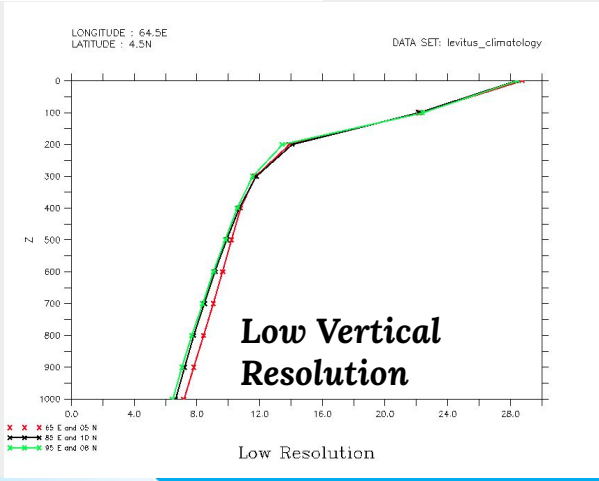
Space - Vertical (Z)



Regridding Z axis



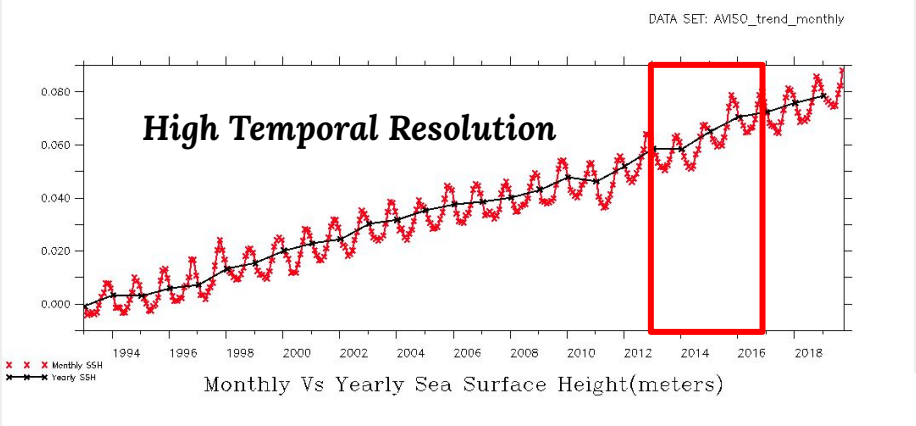
High Resolution

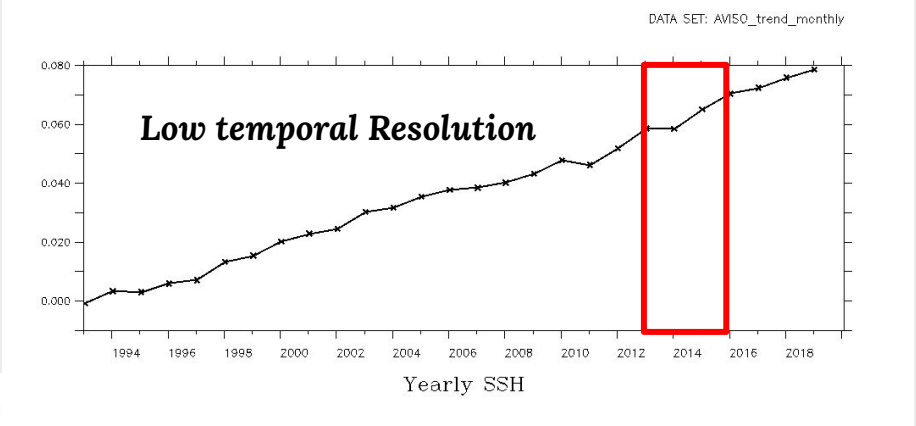


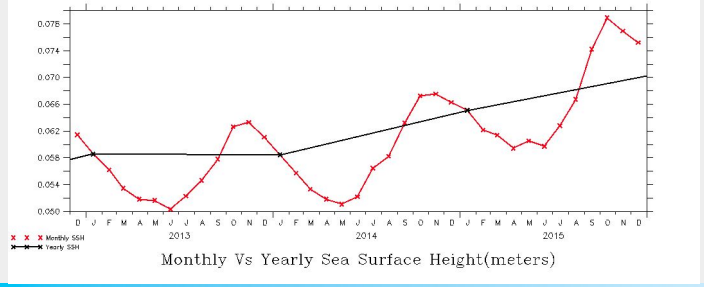
Low Resolution

Regridding Data

Time - Time Series (T)







Basic Statistics

Default Ferret functions for Statistics

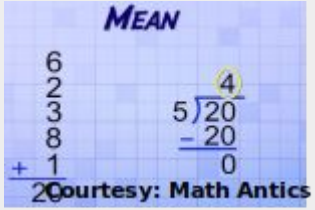
- Average (variable[l=@AVE])
- Standard Deviation (variable[l=@STD])
- Monthly Climatology (variable[GT=month_reg@MOD])
- Anomaly (variable - variable[GT=month_reg@MOD])
- Regression (go regresst or regressx or regressy)
- Correlation (go variance)
- FFT (ffta(variance) (https://ferret.pmel.noaa.gov/static/Demos/ef_fft_demo/ef_fft_demo.html)
- Lowpass (lsl_lowpass (variable, cutoff_period, filter_span))

Basic Statistics

Bias and Mean

Mean (or) Average:

For dataset, mean/average is the **summation/addition of the values divided by the number of values in sample**. i.e. In other words, it is the sum divided by the count.



$$\overline{O}=\frac{\sum_{i=1}^{N} O_i}{N} \qquad \overline{X}=\frac{\sum_{i=1}^{N} X_i}{N} \qquad \overline{X}=\frac{\sum X}{N}$$

Bias (or) Error:

It is **difference between the forecasted and observed values in dataset.** i.e. In other words, bias/error is the tendency of a statistic to overestimate or underestimate a parameter

$$¿\,Bias=¿\frac{\sum_{i=1}^{N} F_i - O_i}{N}$$

variable[l=@AVE]

https://www.youtube.com/watch?v=QqibyNJdyDU&t=769s

My Videos for more details on Basic Statistics

Variance and Standard Deviation

Standard Deviation :

Standard deviation is the **square root of the variance**. i.e. and it can never be negative.

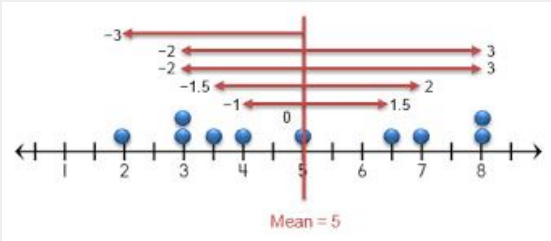
It is a **measure of how far each observed value from the mean** and can be derived from variance by taking square root of variance.

Values in data are widely spread (far) from mean if high and narrow spread (near) from mean if low

It's denoted by and units are same as parameter itself.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (O_i - \overline{O})^2}{N}}$$

variable[l=@STD]



Variance :

It is calculated as the **average squared difference of each number from the mean of the data set.**

Variance measures the average degree to which each point differs from the mean and it can be derived from standard deviation by taking the square of the standard deviation.

It is denoted by and units are squared units of parameter itself.

$$\sigma^2 = \frac{\sum_{i=1}^{N} (O_i - \overline{O})^2}{N}$$

RMSE

Root Mean Square Error (RMSE):

RMSE is the standard deviation of the residuals (prediction errors or difference between forecast and observation .i.e. [F-O]).

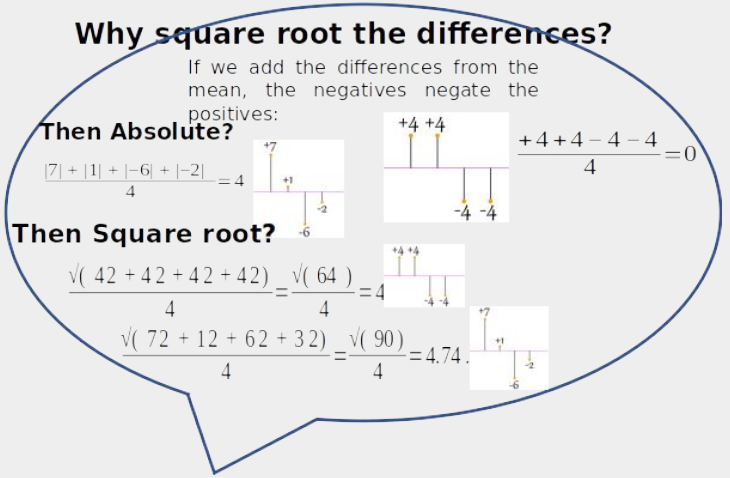
RMSE is a **measure of how spread out these residuals are.** In other words, it tells you how concentrated the data is around the line of best fit.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (F_i - O_i)^2}{N}}$$

Difference between and Standard deviation actually shows how spread out a group of values are from the mean by looking at the square root of variance whereas, variance measures the average degree to which each point differs from the mean.

Basic Statistics

Scatter Index



Scatter Index (SI):

It is the **RMSE normalized with the averaged observed value** and calculated by dividing RMSE with mean of the observations. i.e. SI = RMSE / mean of observations

It represents the percentage of RMSE with respect to mean observation and gives the expected error for the parameter.

It can also be represented in percentage (%) by multiplying SI with 100. i.e. SI*100

$$SI = \frac{RMSE}{\overline{O}} \qquad SI = \frac{\sqrt{\frac{\sum_{i=1}^{N}(F_i - O_i)^2}{N}}}{\overline{O}}$$

Basic Statistics

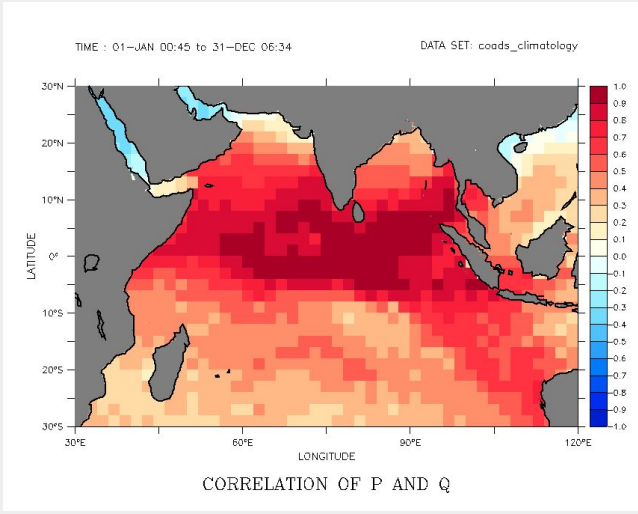
Correlation

Correlation (R)

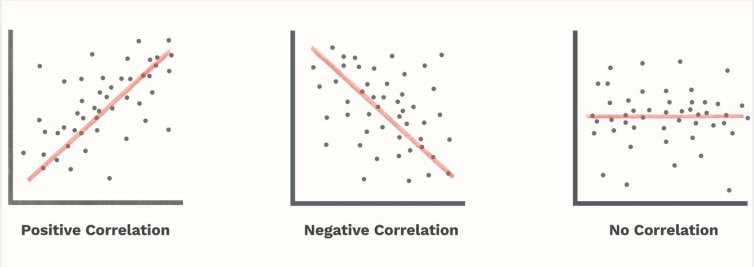
It is the **measure of the degree of linear relationship between forecasted and observed data** and it can be calculated by using formula given below.

Its value ranges from +1 to -1

It is Positive (R=1), Negative(R=-1) and No correlation(R=0)



CORRELATION OF P AND Q



$$R_{(o,f)} = \frac{\sum_{i=1}^{N} (O_i - \overline{O})(F_i - \overline{F})}{\sqrt{\sum_{i=1}^{N} (O_i - \overline{O})^2 \sum_{i=1}^{N} (F_i - \overline{F})^2}}$$

$$R_{(o,f)} = \frac{Cov(O,F)}{Std_O \cdot Std_F}$$

$$R_{(o,f)} = \frac{Covaraince(O,F)}{\sqrt{varianceO \cdot varianceF}}$$

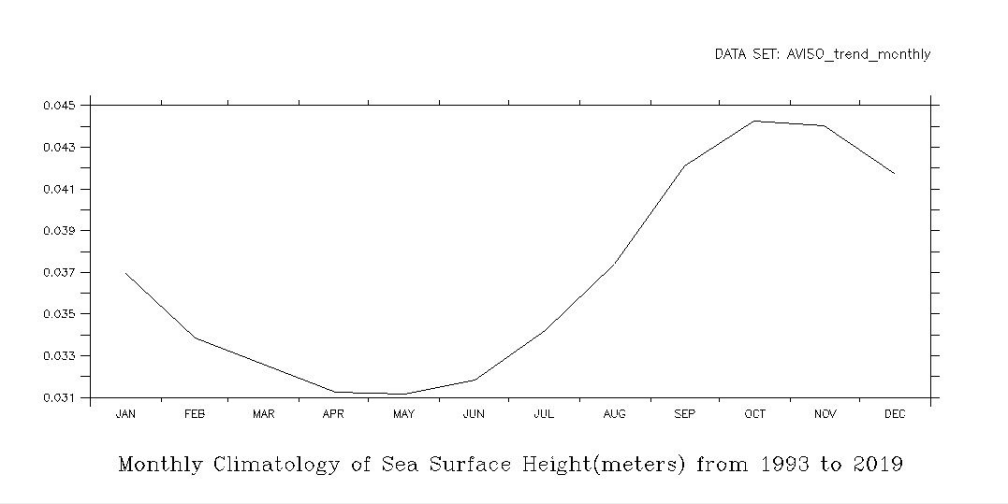
Go variance

Basic Statistics

Climatology

Climatology:
Long term average
Annual and Inter annual
Monthly and Seasonal etc.

Used mostly for climatological long spatial and temporal scale studies



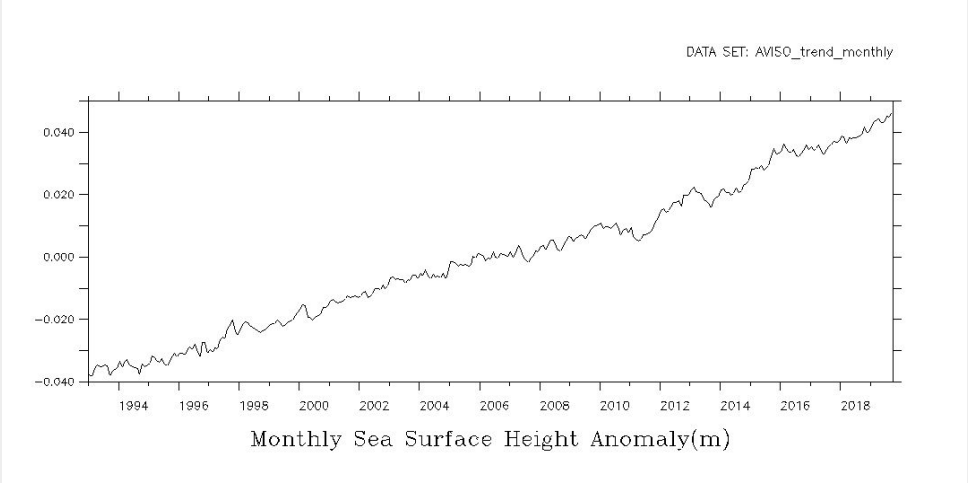
variable[GT=month_reg@MOD]

Basic Statistics

Anomaly

Anomaly:
Deviation from long term climatological mean

Used to various variabilities and corresponding processes



Variable - Variable[GT=month_reg@MOD]

Basic Statistics

Regression Analysis

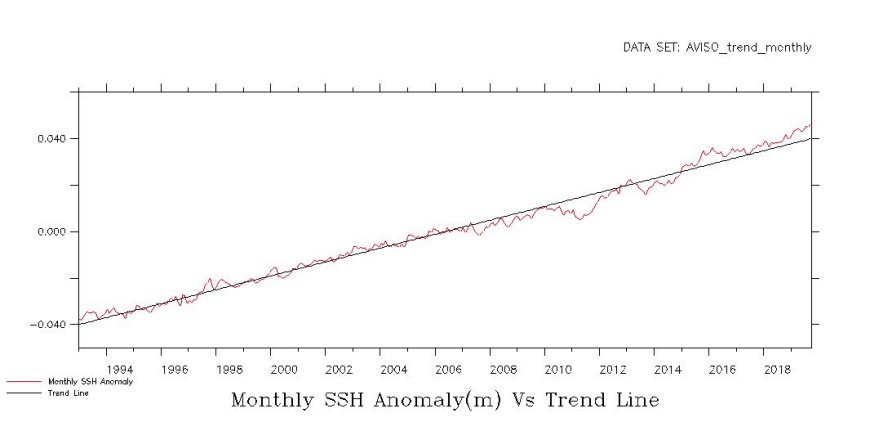
Regression Analysis:
Degree of association between two or more parameters

Regression analysis is a statistical method that shows the relationship between two or more variables.

It generally contains, independent, dependent variables, slope, intercept and constants.

Various regression analysis are there such as
- Linear
- Multiple
- Logistic etc.

Uses for future predictions and forecast



Go regresst

Basic Statistics

FFT

ffta(variable[l=])

Basic Statistics

Low Pass Filter

lsl_lowpass (variable, cutoff_period, filter_span)

Thank You for your kind patience

For more details, you can approach me through

INCOIS: - s.borra-p@incois.gov.in

Quora: https://www.quora.com/profile/Sivaiah-Borra

YouTube: https://www.youtube.com/@guidemycareer-meteorologya8816/featured