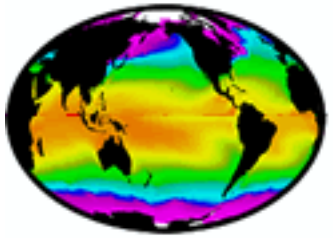


GHRSST
GROUP FOR HIGH RESOLUTION
SEA SURFACE TEMPERATURE

CDAF Task Team

Jonathan Mittaz

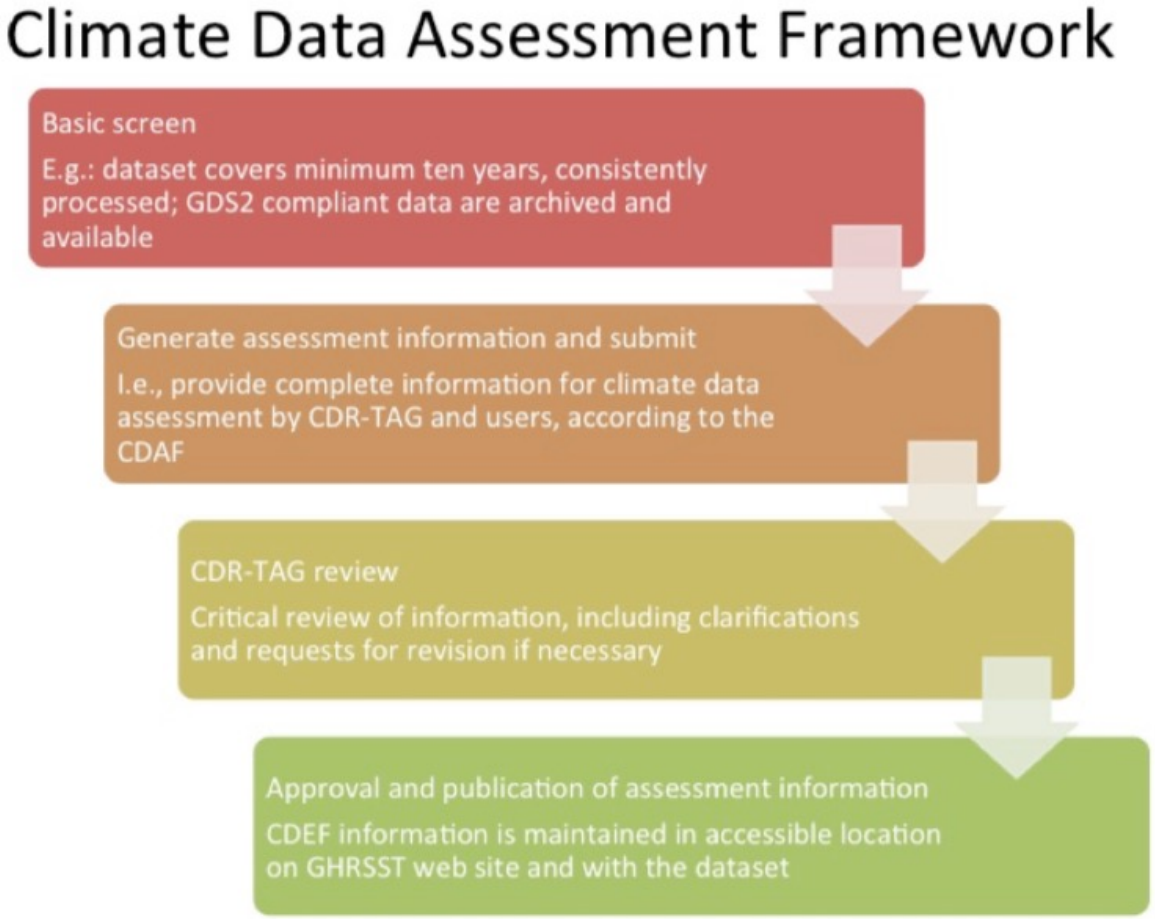
Prasanjit Dash, Jean-Francois Piolle

The CDAF

- Climate Data Assessment Framework
  - A CDR is "a time series of measurements of sufficient length, consistency and continuity to determine climate variability and change" (NRC, 2004), ideally traceable to SI standards.
  - The CDAF is intended to be used to support users of sea surface temperature (SST) datasets to understand the suitability of GHRSST datasets for use as Climate Data Records (CDRs).
    - Will provide authoritative, comparable information about GHRSST datasets that will allow users to make their own judgment about the use of the datasets as CDRs for their application.
- Task Team
  - Has been in existence for several years but other pressures/projects have meant that progress has been slow to date
  - Progress has been made this year, particularly on the MMD side

Overview of the CDAF

- There are a number of sources of information that are needed



CDAF Tool metrics

- Primarily to be used to provide information for section 2
  - Generate assessment information (Quantitative measures)
    - Systematic differences referenced to drifting buoys
      - *Global median difference*
      - *Bin SST to buoy differences on a 10^{o}x10^{o} scale and estimate standard deviation of subsets (filter out bins with too few matches first)*
    - Systematic differences referenced to Argo measurements
      - *Global median difference*
      - *Bin SST to buoy differences on a 20^{o}(lat)x90^{o}(lon) scale and estimate standard deviation of subsets (filter out bins with too few matches first)*
    - Non-Systematic effects
      - *Calculate robust standard deviations after median values (including geographic variations) have been removed using above statistics*
    - Stability
      - Use the GTMBA using simplified method from Merchant et al. (2012)
    - SST Sensitivity
      - Calculated SST sensitivities based on retrieval algorithm characteristics (provided by data provider)

Example of current output form

KEY DESCRIPTIVE FEATURES	INFORMATION
Period covered	
Geographic range	
Spatial resolution	
Temporal resolution	
Timeliness of new data	
Dataset volume	
Valid data fraction	
Data level / grid	
Observation technology	
Dependence on other data	
Type(s) of SST	
Traceability	
Uncertainty info in product	

QUANTITATIVE MEASURES	VALUE	COMMENTS
Difference relative to drifting buoys		*Global median difference of satellite minus drifting buoy SST, across full dataset. The satellite SSTs are SST_{skin} with no skin-effect adjustment, so a skin-effect difference of order -0.2 K is to be expected.*
Difference relative to Argo		*Global median difference of satellite minus upper Argo float SST, across full dataset. The satellite SSTs are SST_{skin} with no skin-effect adjustment, so a skin-effect difference of order -0.2 K is to be expected.*
Geographical variation in difference relative to drifting buoys		*Geographical variation in difference, as described by the standard deviation of median satellite minus drifting buoy SST differences on space scales of ~1000 km, across the full dataset.*
Geographical variation in difference relative to Argo measurements		*Geographical variation in bias, as described by the standard deviation of median satellite minus upper Argo float SST differences on space scales of 20° latitude by 90° longitude, across the full dataset.*
Dispersion relative to drifting buoys		*systematic effects as quantified by a robust standard deviation of differences of satellite and drifting buoy data, after removing the geographical variations in differences quantified above*
Stability		*95% confidence interval for the relative multi-year trend between satellite SSTs and the Global Tropical Moored Buoy Array*
Sensitivity to true SST		*Average weight of the satellite observations in determining SSTs in the dataset, the difference from 100% representing the weight of prior information in the SSTs*

AVAILABILITY, DOC'N, FEEDBACK	
Data URL / ftp / DOI	
Primary peer reviewed reference	
Source of technical documents	
Dataset restrictions	
Facility for user feedback	
Other documentation	

OTHER PRINCIPLES (GCOS)	COMMENTS
2. and 12. Overlaps between sensors exist and are exploited to harmonize the dataset	
3. Detailed history of methods/ algorithms is available	
11. Constant sampling within diurnal cycle	

Workplan

- **Get reference data**
- **Produce Matchups**

Together with Matchup Task Team (Jean-Francois)
Matchup production coordinate with Matchup TT
Initial data produced

- **Create Statistics**

As much as possible this will use already existing code/processes.
  - Have looked at Meteo-France code (**Thank you!**) as a starting point

The current set of metrics are fairly simple
- Means/Medians
- Standard Deviation/Robust Standard Deviations
- Geographic binning code

- **Web/Form output**

Initial designs created (Prasanjit)

CDAF Match-Up Generation

Preliminary step of the CDAF Tool Workflow

felyx MMDB generation system is being upgraded to run both:

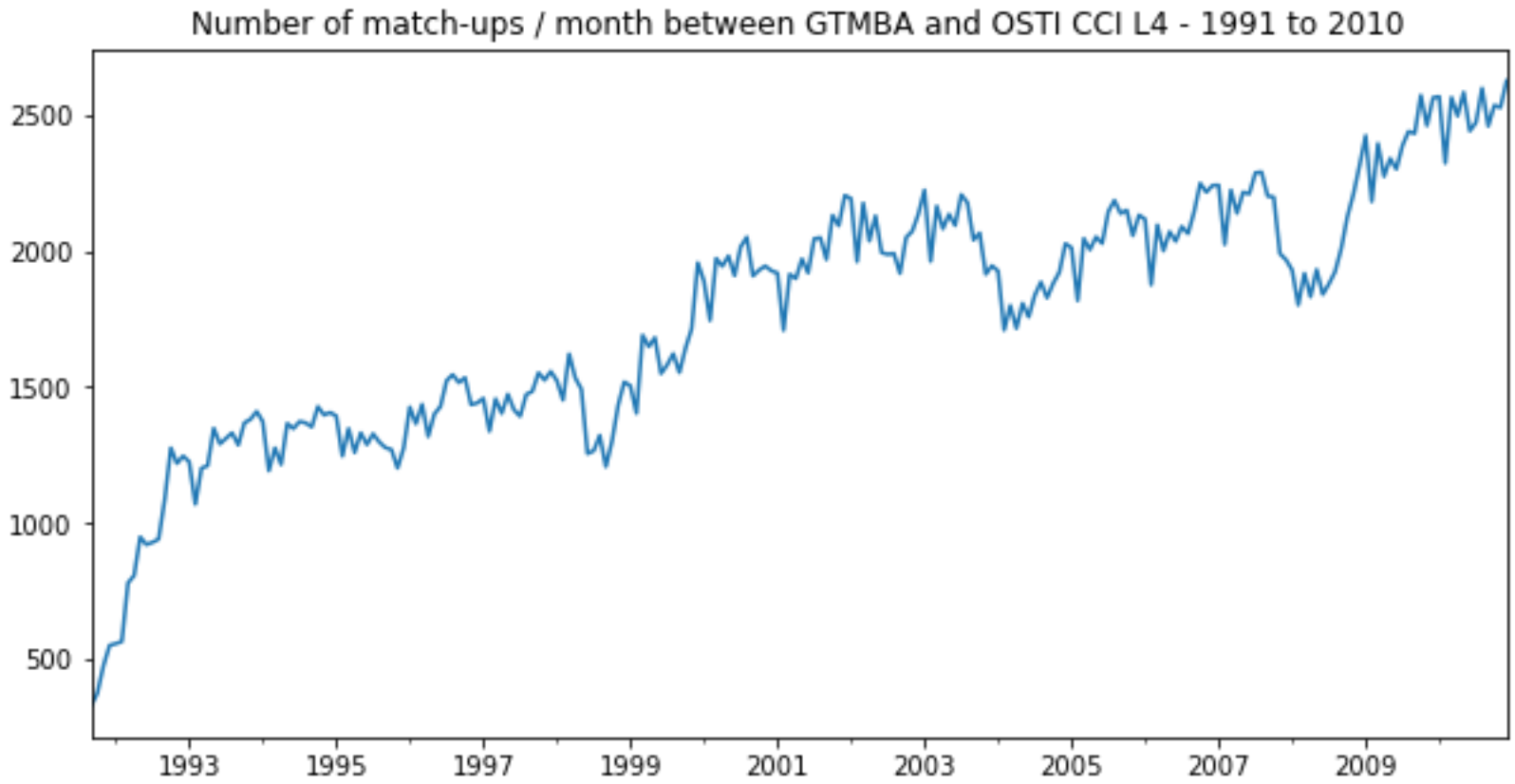
- in distributed mode (as currently) : runs on a cluster, intended for large scale and operational processing
- as an easily installable standalone running on a single machine: for local processing, testing and evaluation => easier to run and use in the context of the CDAF toolbox (but slower as using less computing resources)

Two sample datasets generated:

- using GTMBA moored buoy dataset from CCI SST project as in situ reference
- two candidate SST long time series
  - AVHRR L3C from CCI Project
  - OSTIA L4 from CMEMS
- available for testing and integration with the next steps in the workflow (metrics & report generation)

Matchup statistics

- An example of matchup data



Web Workflow Design

Sitemap: Climate Data Assessment Framework (CDAF) web

ABOUT

- What
- Definitions
  - What may be a CDR?
  - What is a reference?
  - Match-up criteria
  - Diagnostic metric
- Who
- Links

COOKBOOK

- Generating Summary Statistics
  (can be used by a data producer or an independent user)
  - Perform basic screening
  - Generate assessment report
  - Get CDR TT review per list
  - Ad Hoc Approval & publish
  - OK ? → NO
  - YES
- Status & Summary report
- CDR candidature
- Downloadable/fillable Forms
  - Assessment form (12 points)

CDAF RESOURCES

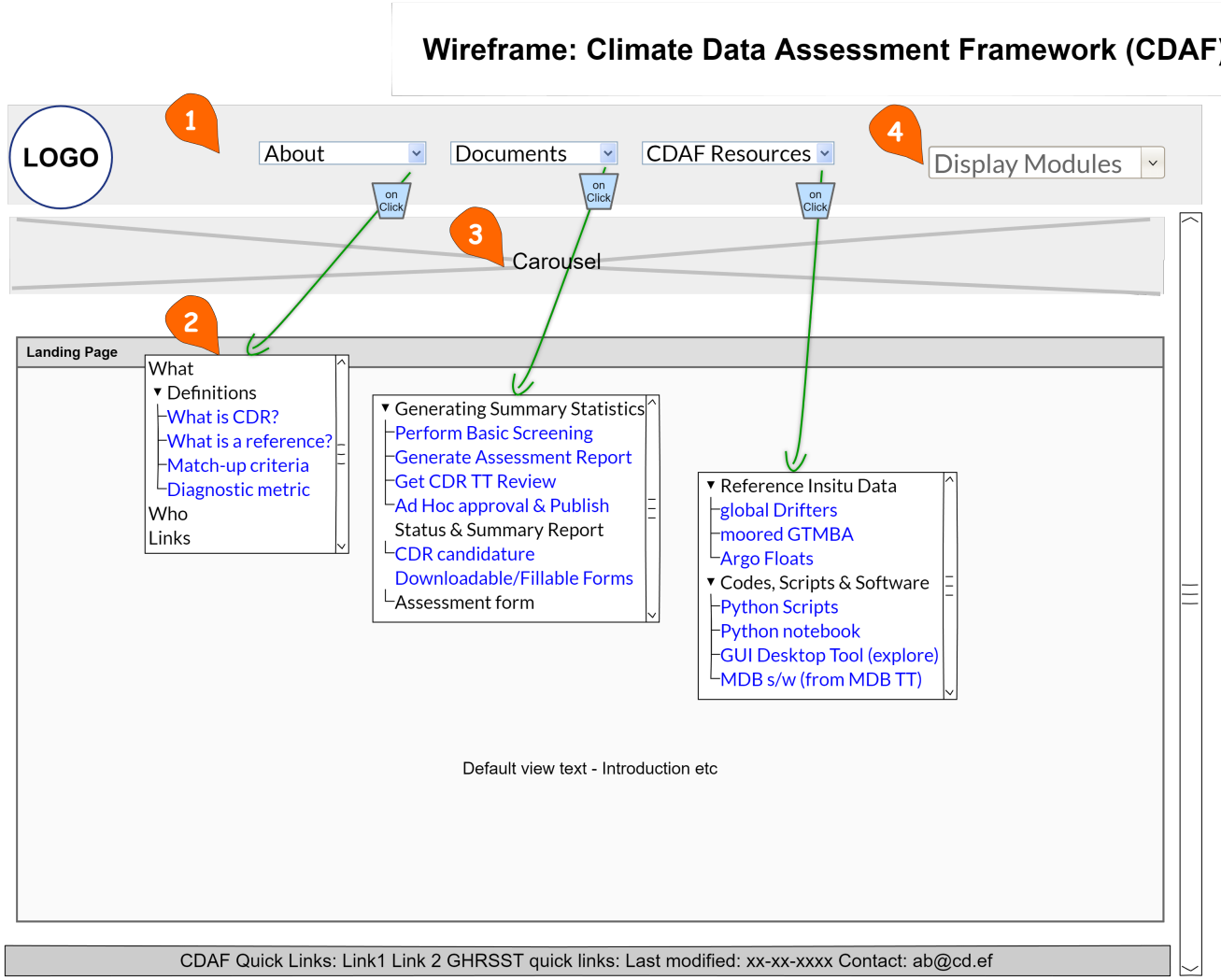
- Reference Insitu Data
  - global Drifters
  - moored GTMBA
  - Argo floats
- Codes, scripts and software
  - Python scripts
  - Python notebook
  - *consider GUI desktop tool*
  - MDB s/w - from MDB TT

DISPLAY MODULES

Choice of SST product | Choice of SST reference | Robust stat parameter

- Robust Statistics
  - Maps
  - Histograms
  - Time-series
- Stability (tropical)
  - Deseasonalized maps
  - Time-series
- Sensitivity
  - Table for all (SST producer dependent)

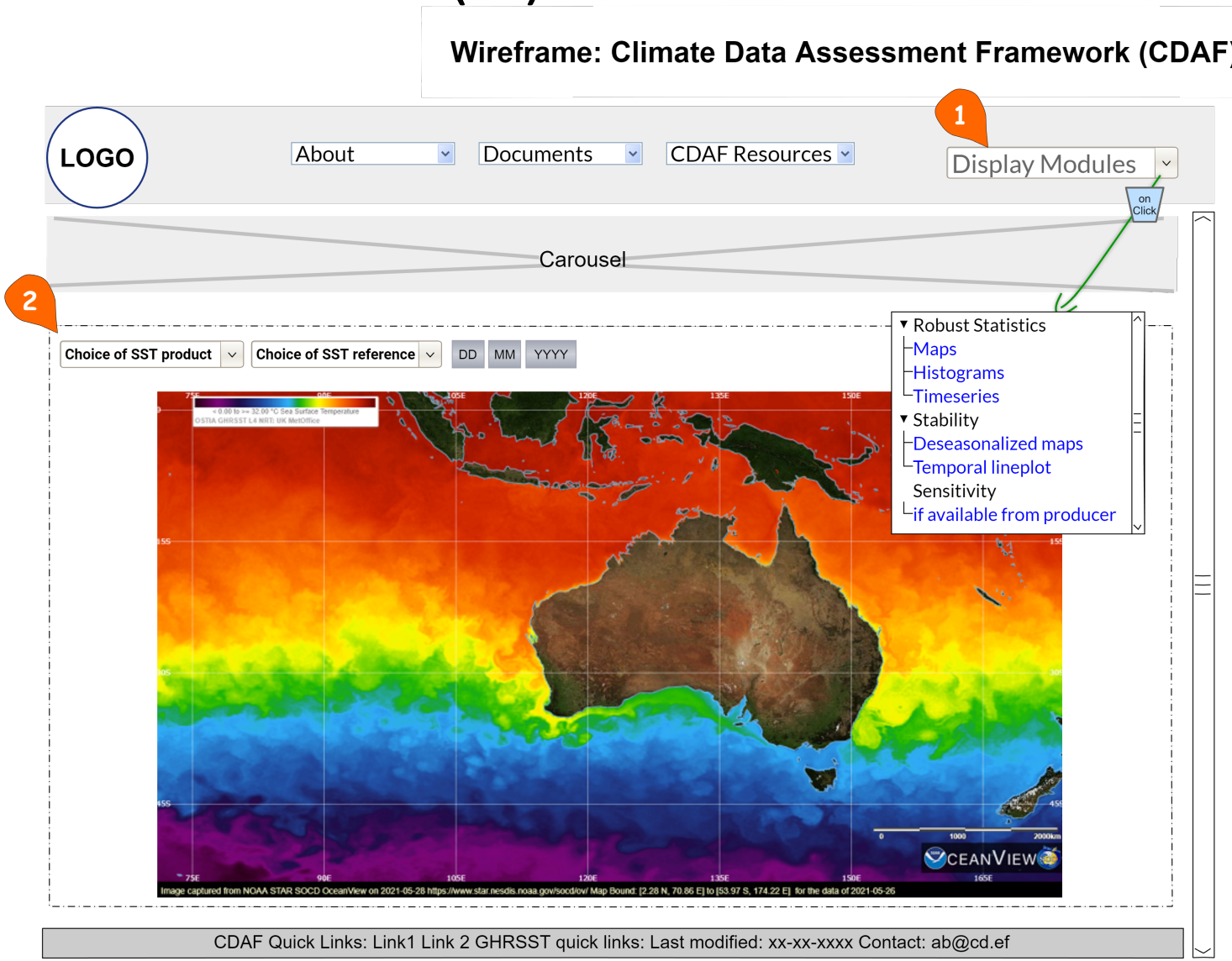
Action of menu items



simple Single Page Application (SPA)
(mobile device compatible)

Itemized meaning	
1	Dropdown Menu, cf. sitemap
2	Dropdown sub-menu choice
3	Rolling items: highlight new additions, call to action
4	Diagnostics display option: Statistics, Stability, Sensitivity

Action of menus (2)



simple Single Page Application (SPA)
(mobile device compatible)

Itemized meaning

1 Diagnostics display option: Statistics, Stability, Sensitivity

2 Dropdowns for product and reference, Date picker

Next steps

- Continue tracking updates to felyx regarding processing mode
  - Data already available for next steps
  - standalone processing
  - Setup scripts to create matchups
- Modify/Write code to generate statistics
  - Leverage of other code where possible
    - Started thinking about this
  - Straightforward for many cases
  - Include mapping software for regional analysis
  - Stability analysis
- Setup initial interface
  - Design already in progress
  - Start working on interfaces/code