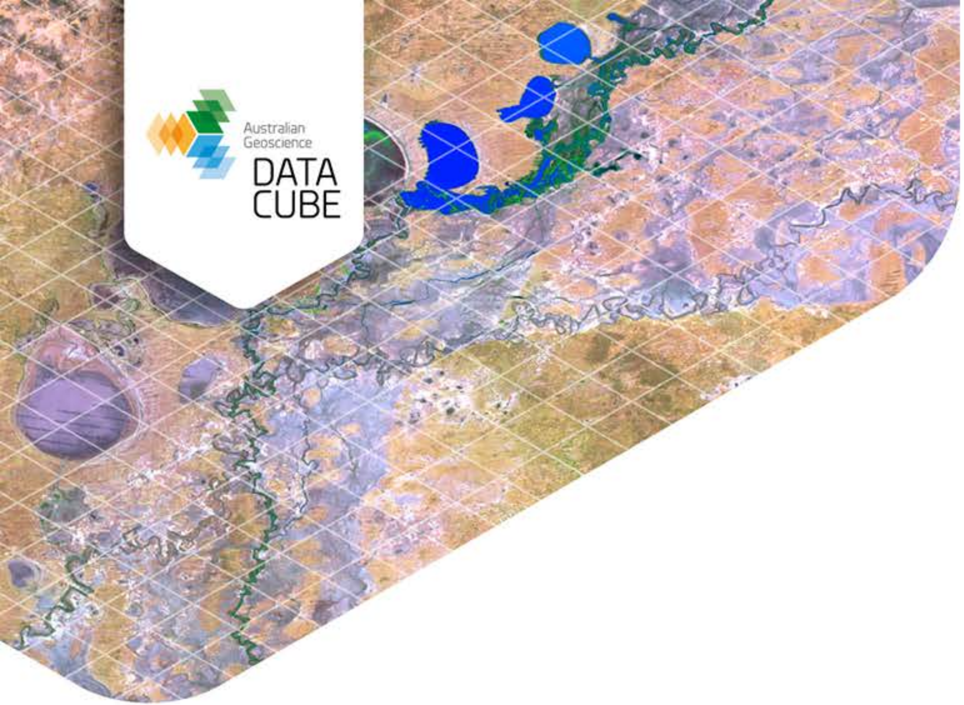


Australian Government
Geoscience Australia

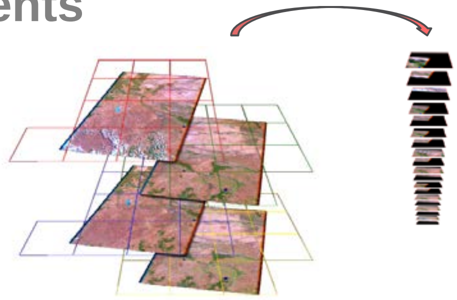
CSIRO

NCI
NATIONAL COMPUTATIONAL INFRASTRUCTURE

AGDC Update – OzEWEX Workshop 2016

AGDC Overview

A series of data structures and tools to enable efficient analysis of large earth observation archives in HPC environments

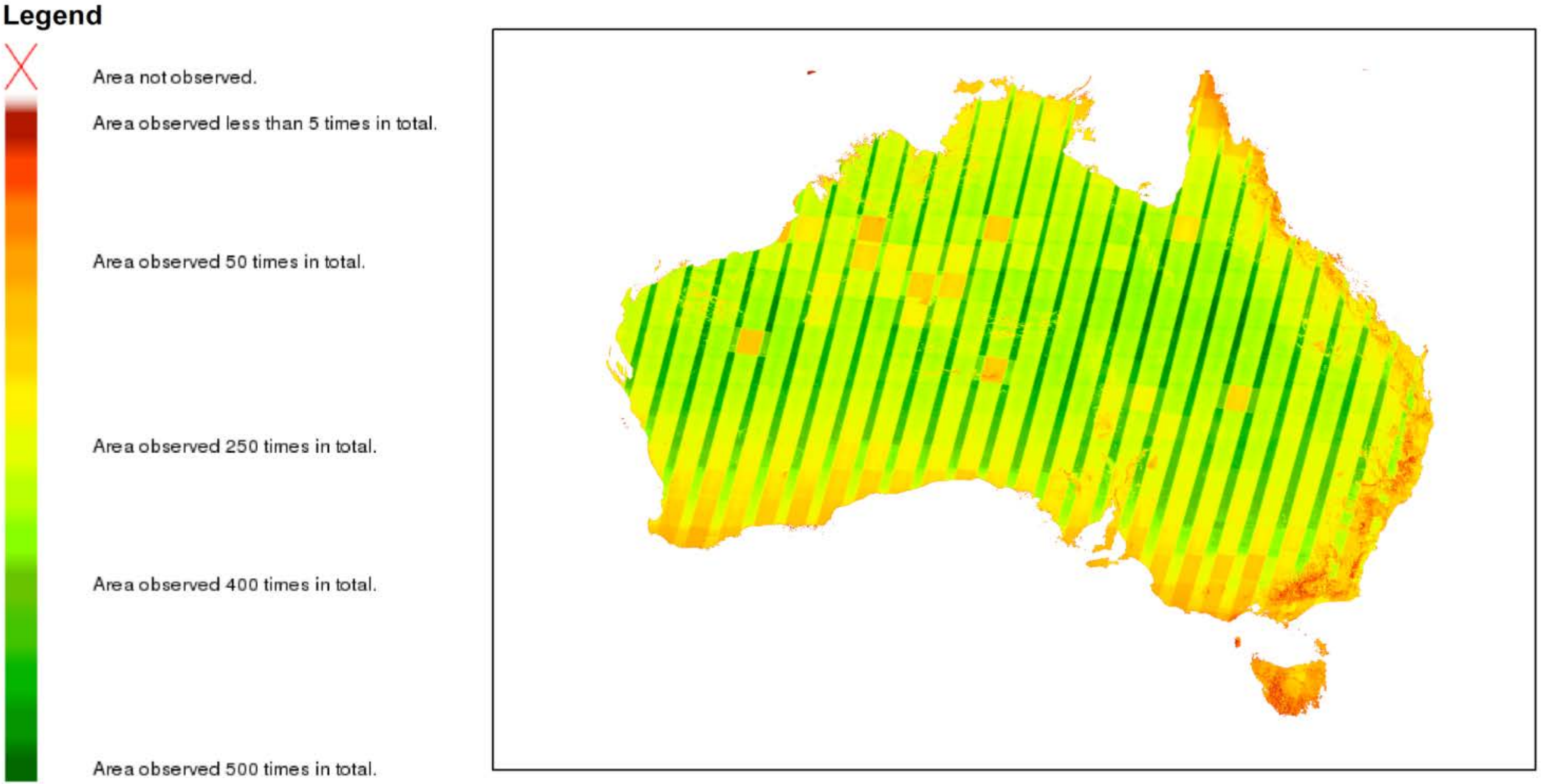


- Simple Data Structures
  - Spatially regular tiles
  - Managed by a relational database
- Calibrated and Standardised Unique Observations
  - Surface Reflectance Observations
- Quality Assured Observations
  - Flagged for cloud, cloud shadow, saturation and other quality indicators
- Open source software

Australian Geoscience Data Cube

- Data collection is dynamic: growing in time, and also subject to modification (existing data) and insertion (new data). The challenge is to enable:
- Attribution of exact observation time for key applications
- Analysis of each observation in the time-series
- Reliable comparison of observations over long periods of time, e.g. change detection, pattern analysis
- Iteration and refinement of processes at continental scale
- Rapid generation of results
- Hosts the Australian archive of Landsat data (LS5, LS7 and LS8) and other Earth observation and environmental data
- Live access analysis system in a HPC architecture.

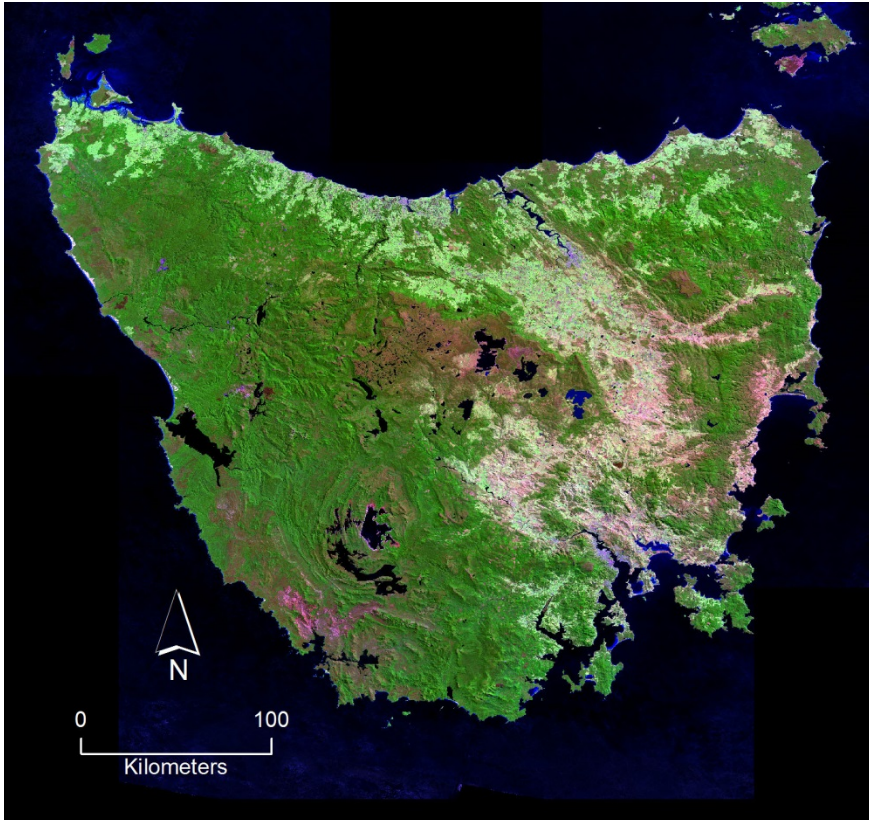
Quality Assured Observations



ve Sum
1
3
7
15
31
63
127
255
511
1023
2047
4095
8191
16383
32767
65535

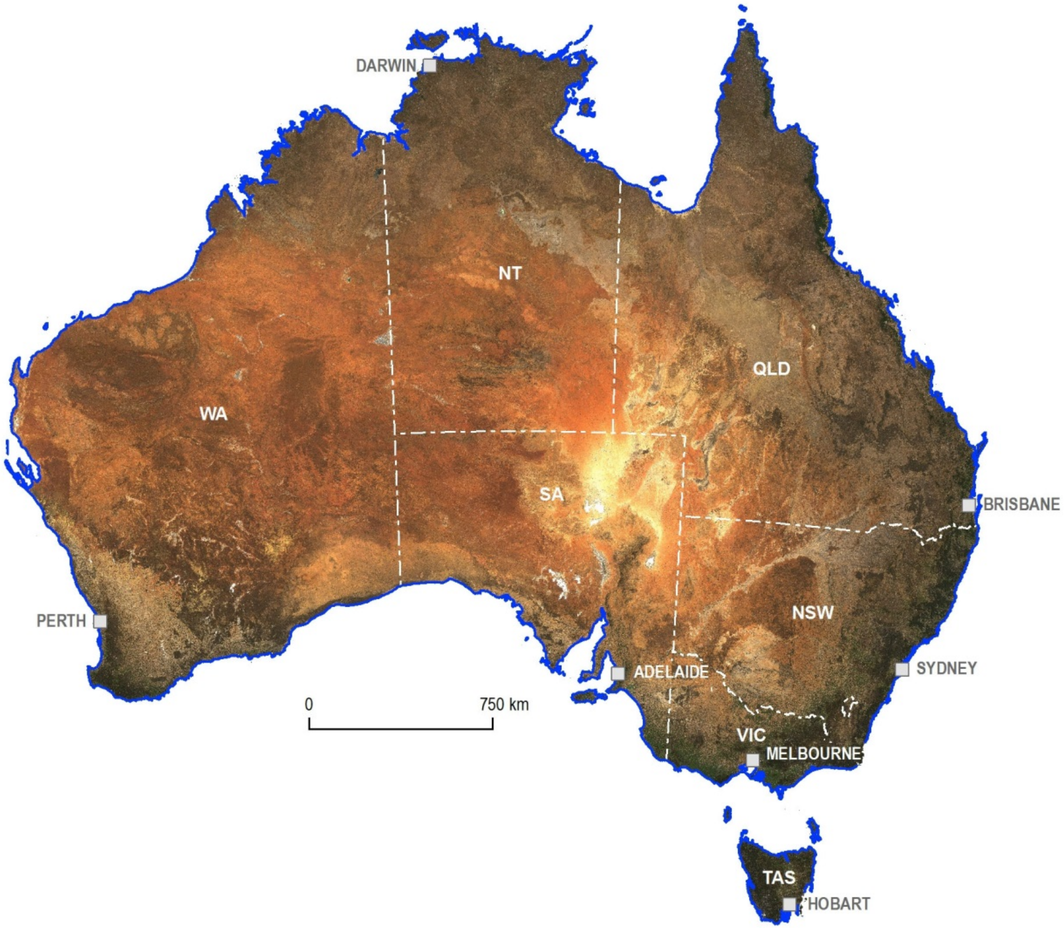
Statistical Composite Mosaics

- Using statistical analyses of surface reflectance over a set time period (say a year or season), create consistent continental scale mosaics.
- This example of Tasmania is created using a high dimensional median of Landsat 8 data from 2014.
- Similarly, by creating annual mosaics for the entire time series we examine the change in Australia over the last 30 years.

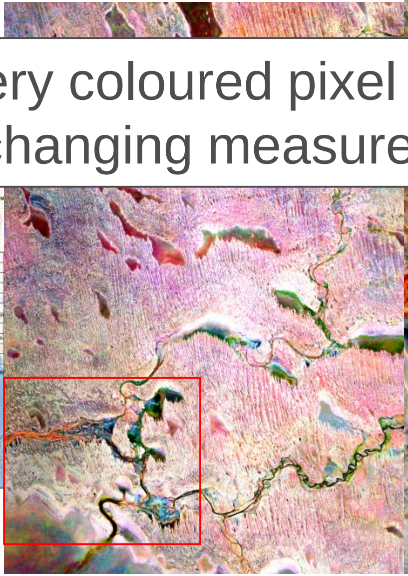


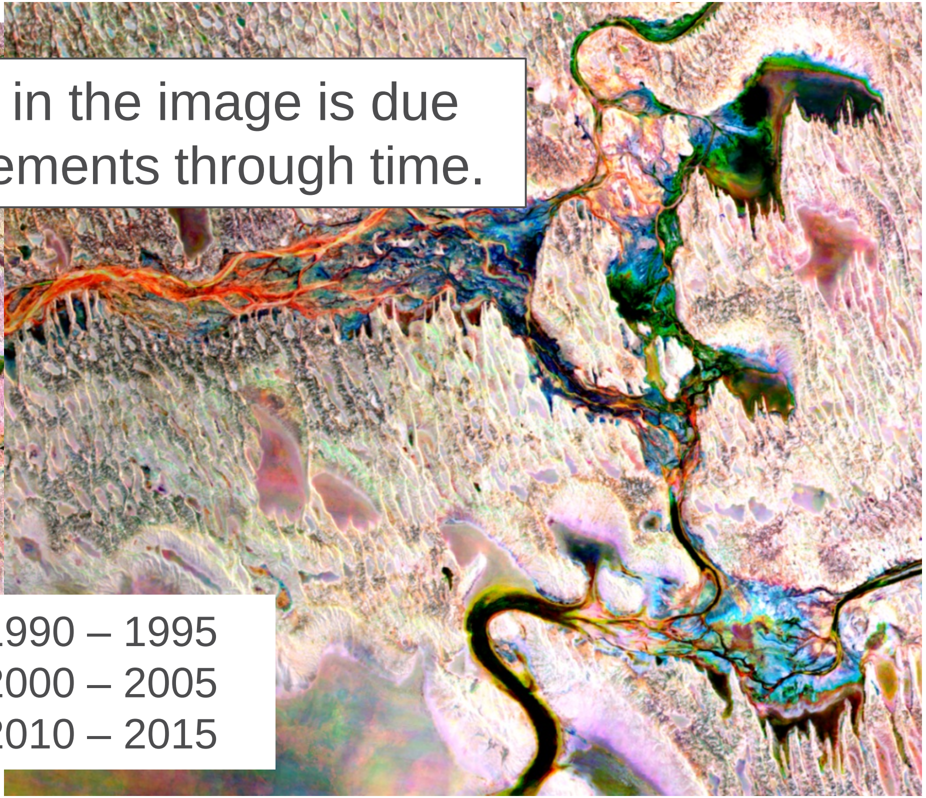
Exposed Earth Mosaic

- By constraining the data to provide the barest pixels through time, a statistical mosaic of the barest pixels can be produced.



Example output - Landsat statistical product

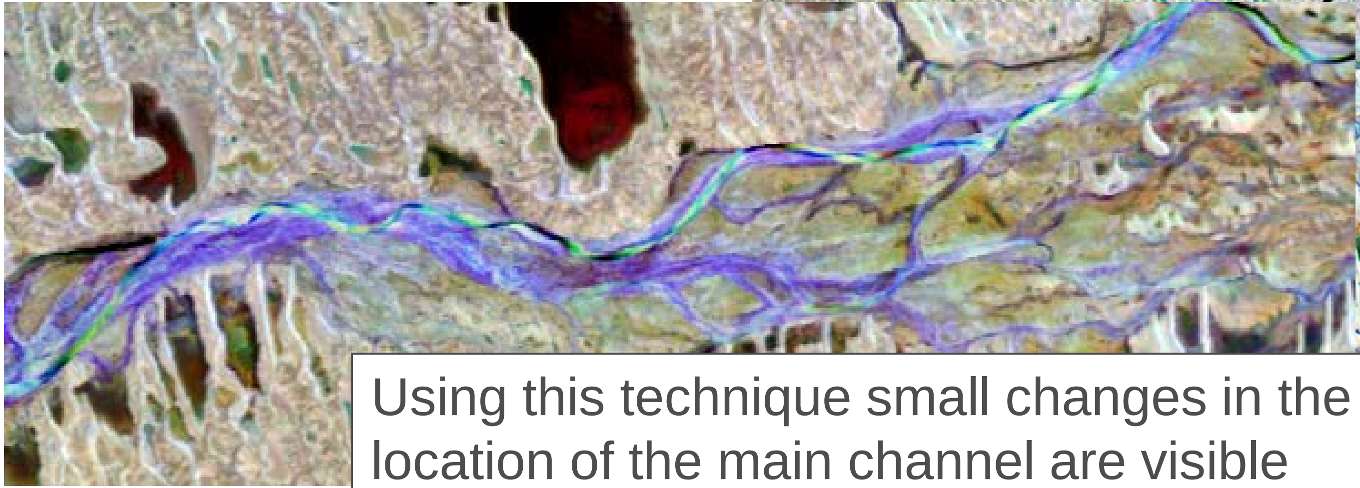


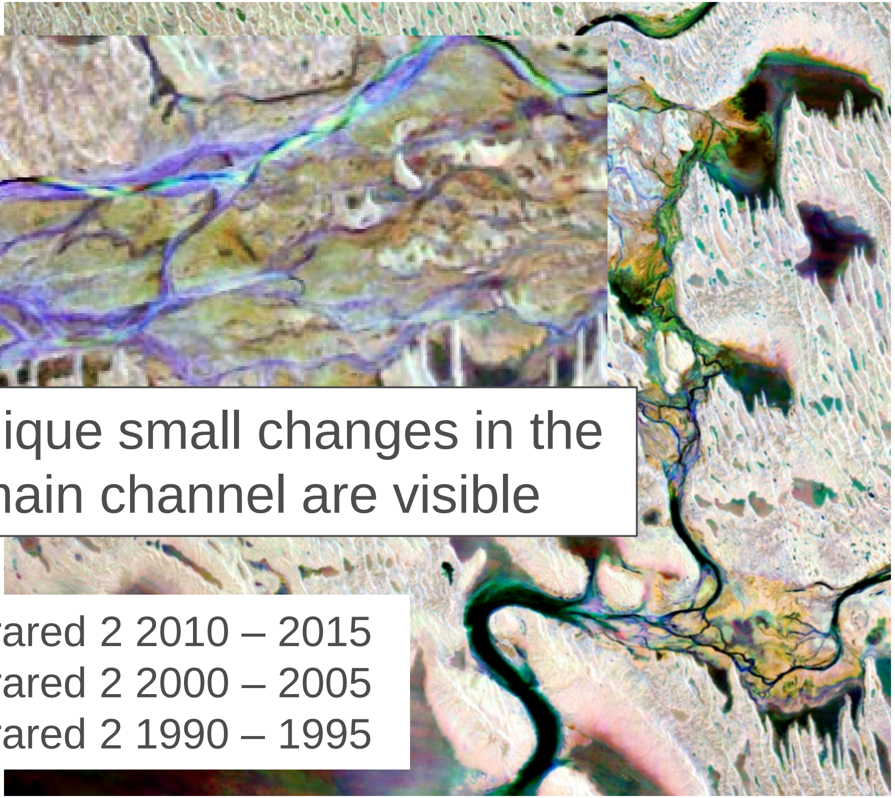


Every coloured pixel in the image is due to changing measurements through time.

Red:	Median summer near-infrared 1990 – 1995
Green:	Median summer near-infrared 2000 – 2005
Blue:	Median summer near-infrared 2010 – 2015

Example output Landsat statistical product



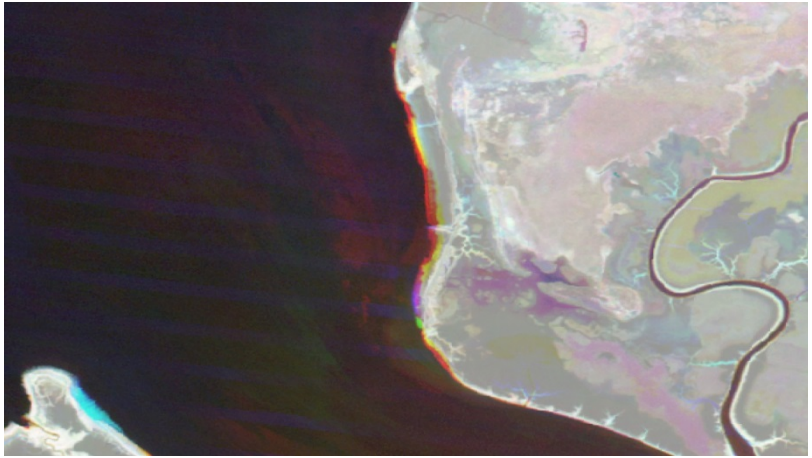


Using this technique small changes in the location of the main channel are visible

Red: Median summer short wave infrared 2 2010 – 2015
Green: Median summer short wave infrared 2 2000 – 2005
Blue: Median summer short wave infrared 2 1990 – 1995

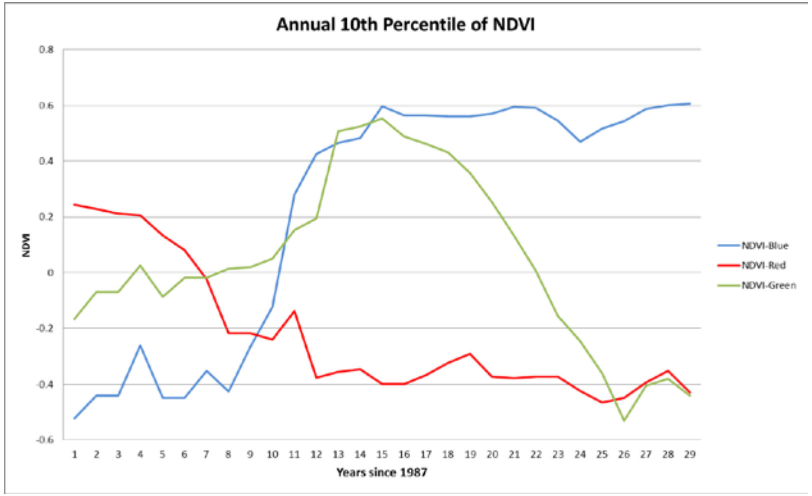
Mangrove example

The RGB image shows the 10th percentile of NDVI for 1987 (red), 1992 (green) and 2016 (blue) for the entrance of the Mary River, NT.

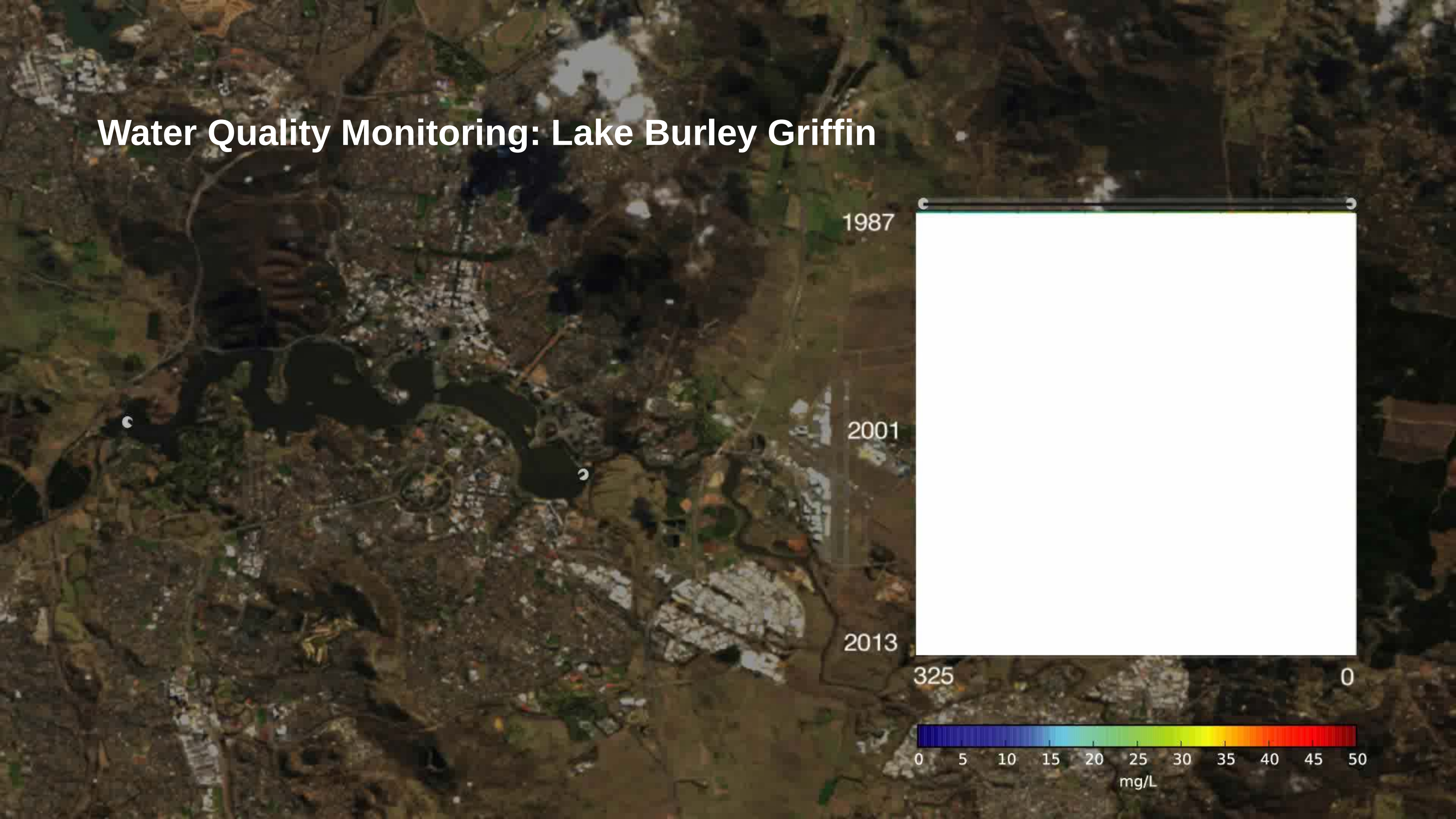


The 10th percentile of NDVI for each coloured area is shown in the graph, indicating the behaviour through time of each area:

- Red areas changed from vegetated to unvegetated
- Green areas changed from unvegetated to vegetated and back again
- Blue areas were unvegetated and changed to vegetated



Water Quality Monitoring: Lake Burley Griffin

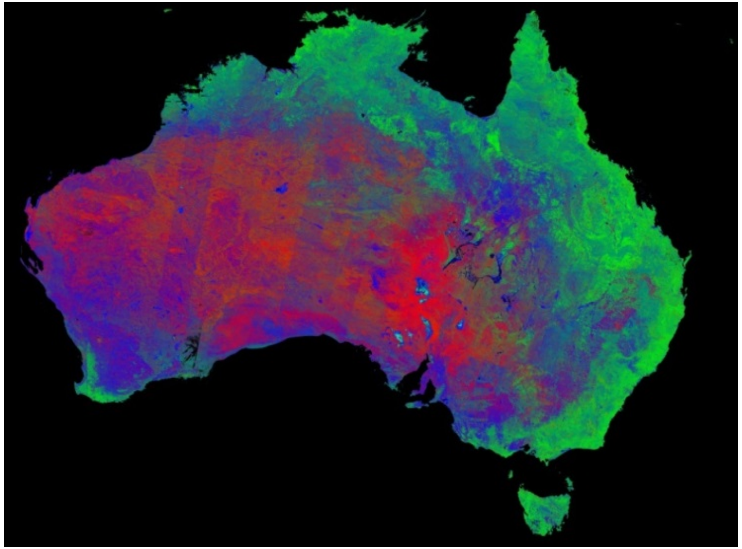


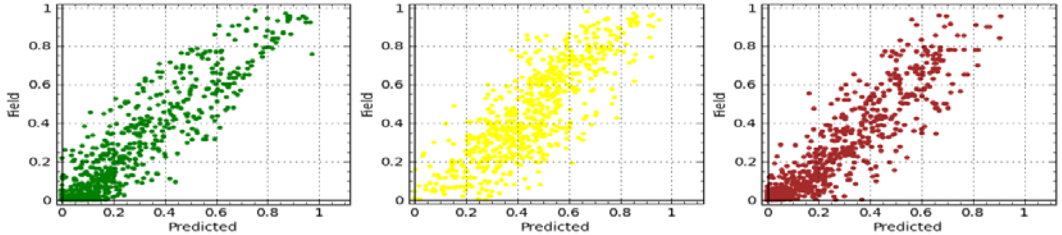
National Fractional Cover – Joint Remote Sensing Research Program

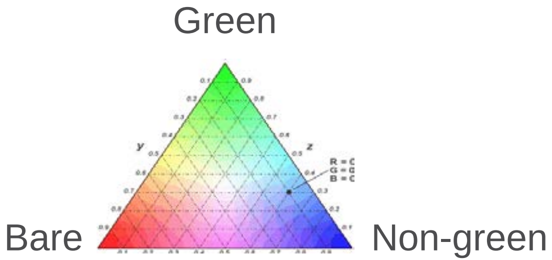
Fractional cover uses a constrained un-mixing model with end-members derived from field sampling

Creates an image with the percentage of bare, green and non-green fractions

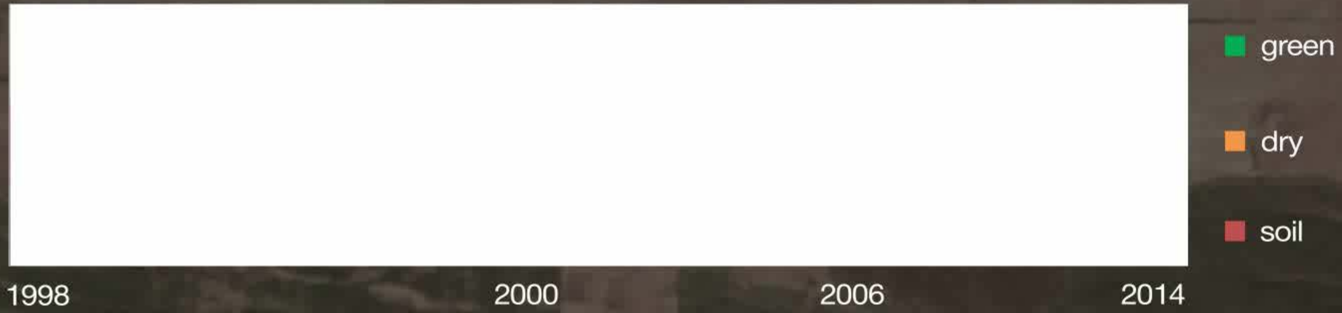
Captures cover dynamics at nominal 25m resolution





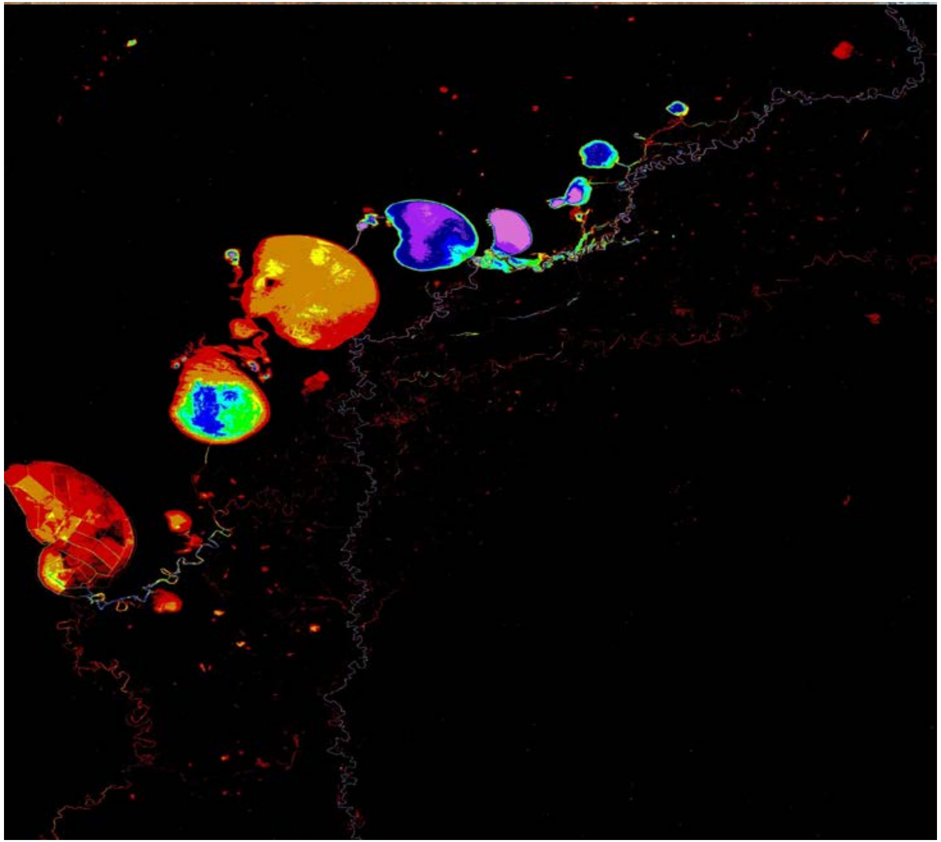


Tracking agricultural change



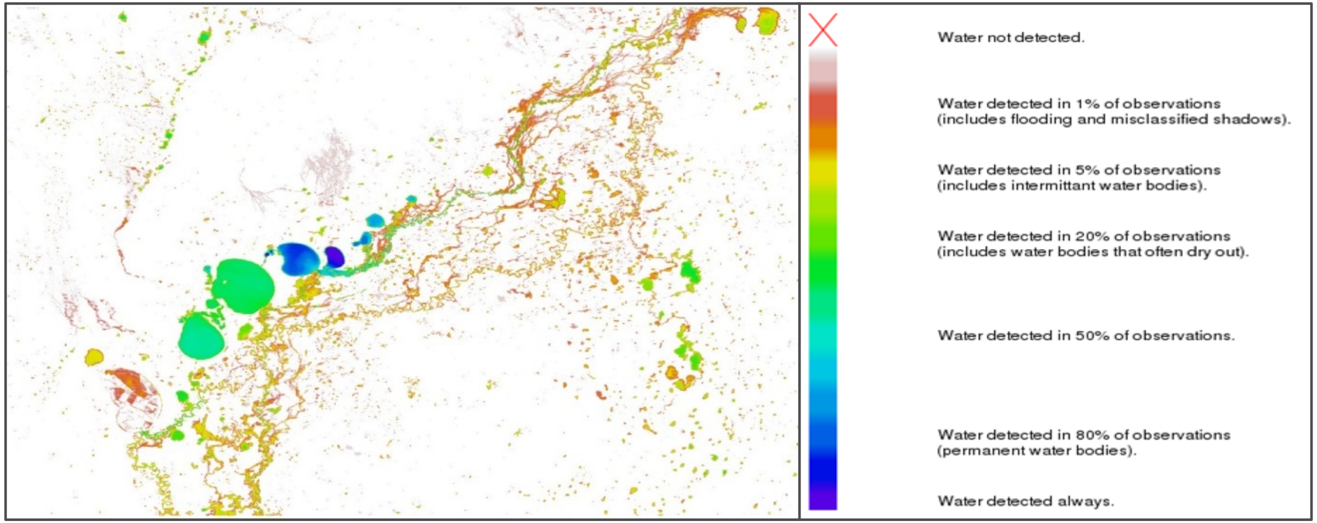
Water Observations from Space

- An ongoing time series of observed surface water over the continent of Australia.
- The time series is derived from the satellite imagery available to Geoscience Australia, in particular the archive of Landsat-5, Landsat-7 and Landsat-8 imagery acquired since 1986.
- The outcome is a product that allows the analysis of the time-series of observed water to generate map products and web feature services on surface water character across Australia.
- Mueller et al 2016.



WOfS Summary Product

- Sum the derived temporal water stack: number of water observations per pixel
- Sum the derived “real” observations for every pixel from the Pixel Quality
- Produce the ratio as a percentage for display
- WMS: http://geoserver.nci.org.au/geoserver Online viewer: www.ga.gov.au/wofs



Menindee Lakes as shown in WOfS, with associated legend

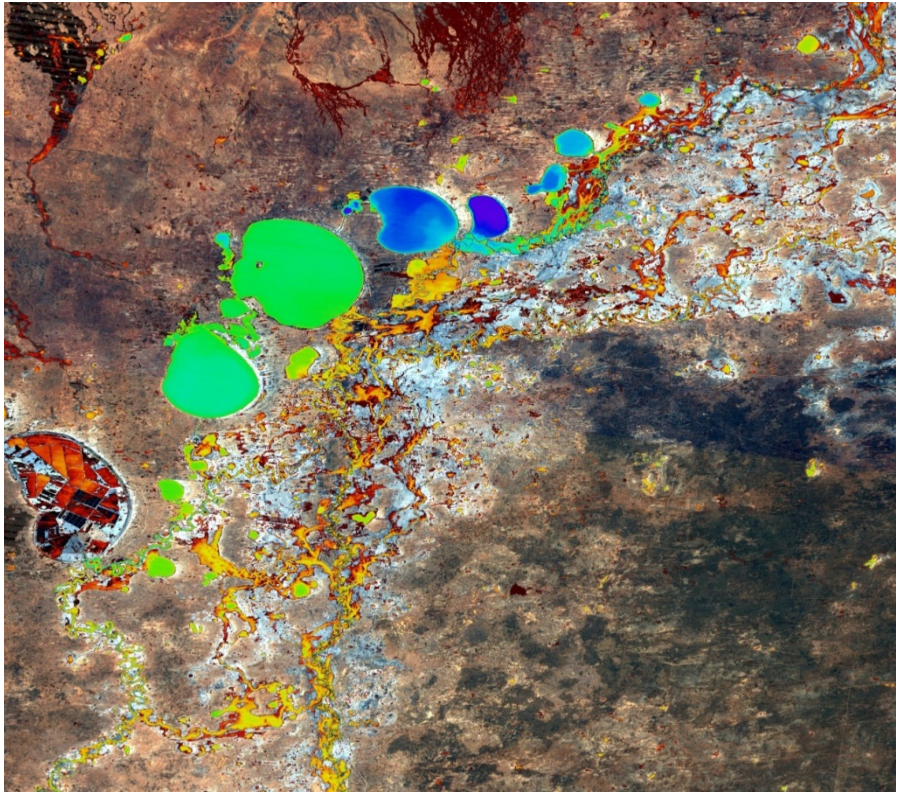
Water body perenniality – Menindee Lakes NSW

The arid areas of Australia have poor information on how permanent the water bodies are.

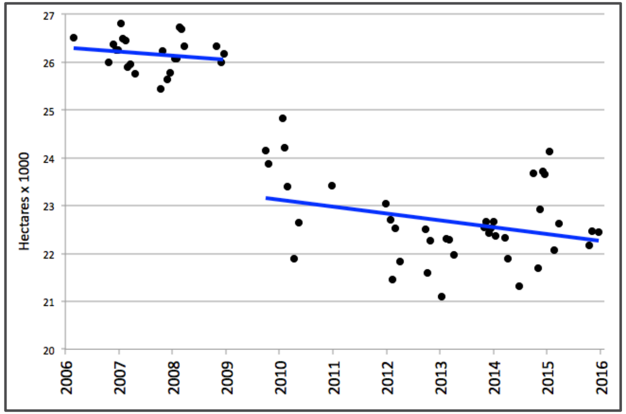
WOfS provides a mechanism to assess the perenniality of all detectable water bodies across Australia.

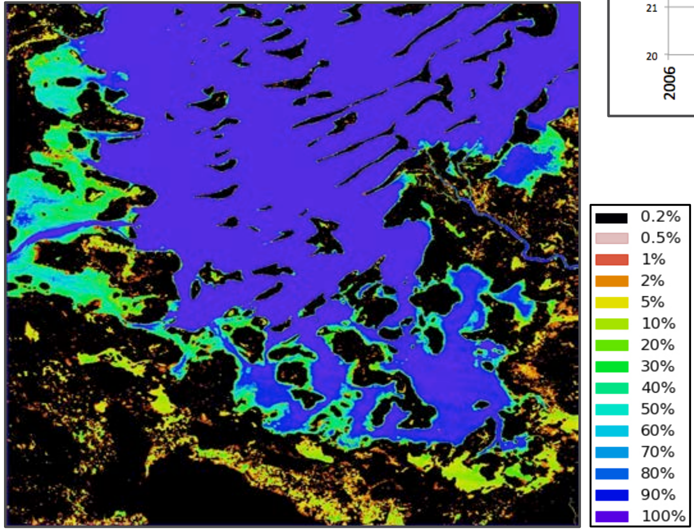
Previously the Menindee Lakes system was mapped as “non-perennial” with no detail on what the real character of these lakes was.

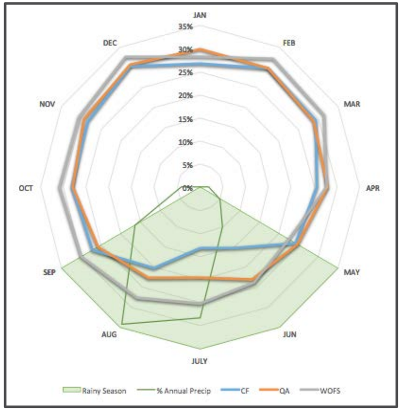
WOfS shows the variation in perenniality for the system.



WOfS Used Internationally – CEOS Data Cube in Cameroon







WOfS implemented on USGS Landsat surface reflectance data over Lake Chad in Cameroon.

Further CEOS Cubes in Kenya and Colombia with WOfS analyses produced over important water bodies.

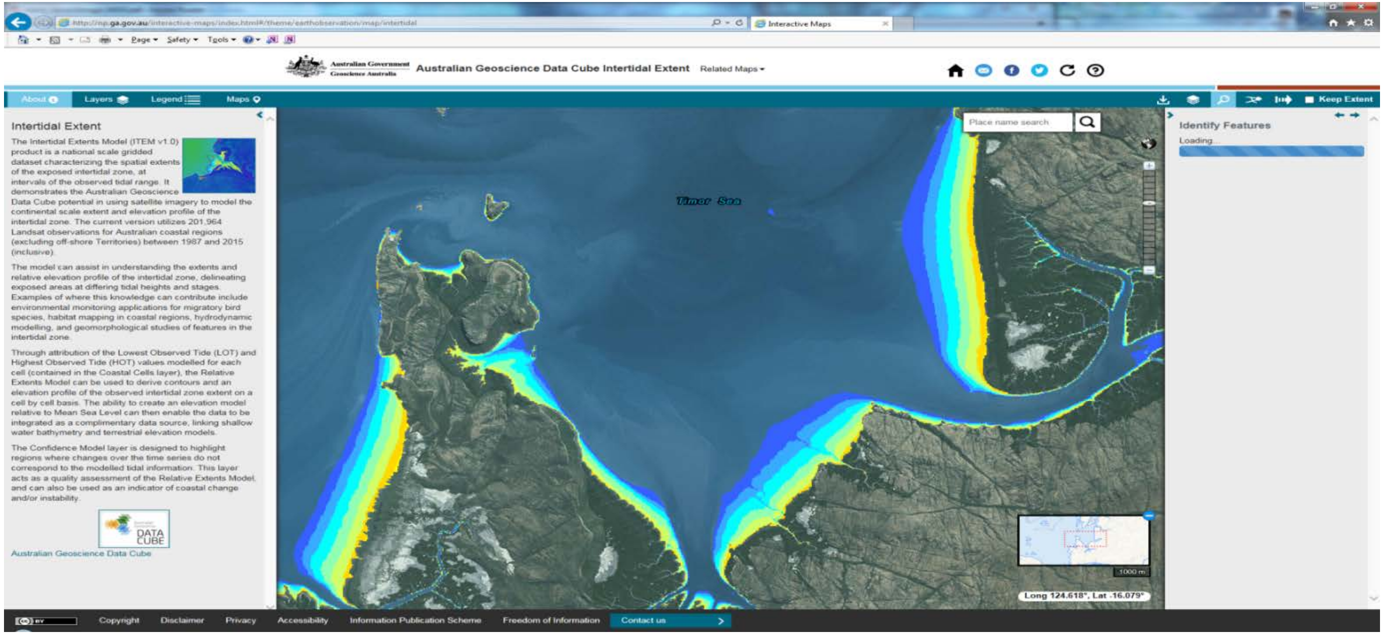
Lake Chad example shows an average reduction in water surface area of 1.4ha per day over a 10 year period.

Images courtesy of Brian Killough of CEOS / NASA

The ITEM Relative Extents Model

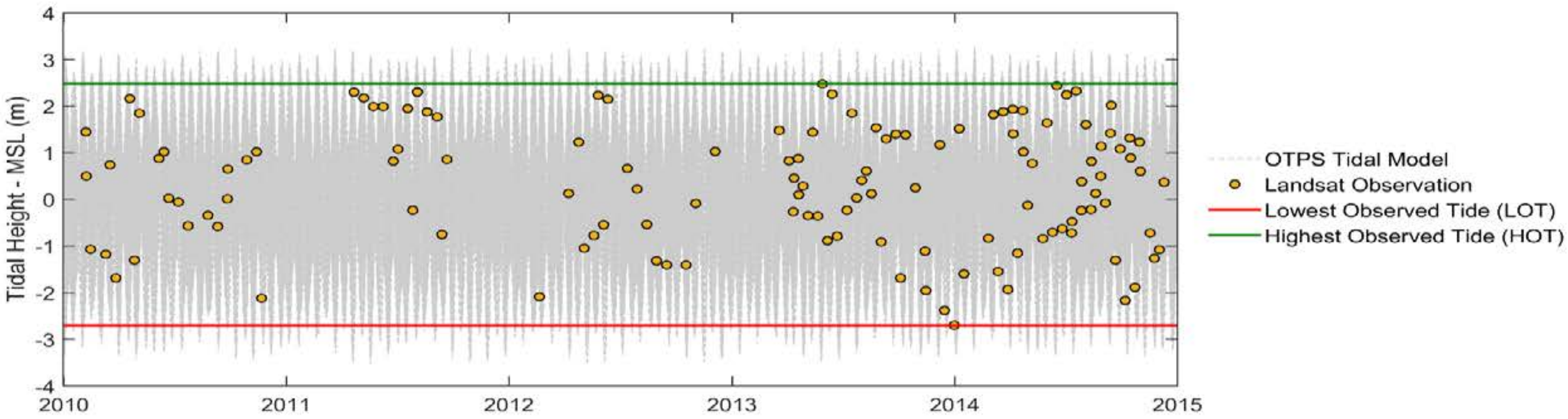
http://www.ga.gov.au/interactivemaps/#/theme/earthobservation/map/intertidal

Sagar et al. in press



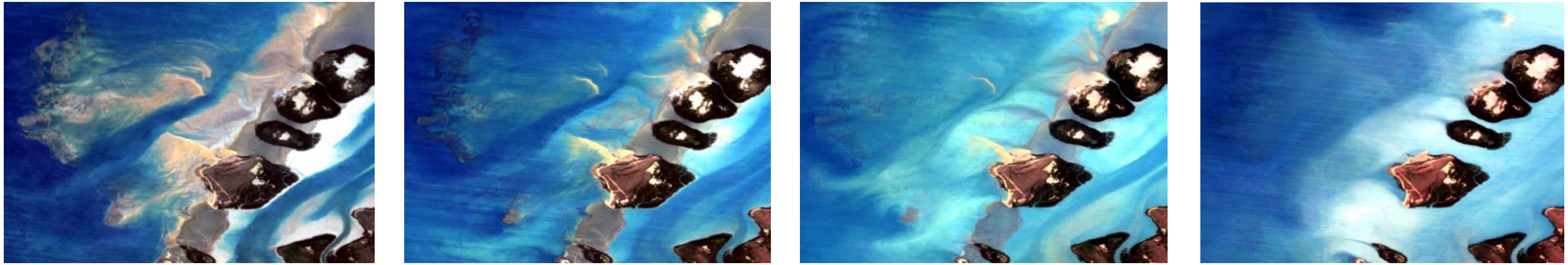
Sun-synchronous sensors and the Observed Tidal Range

- A sun-synchronous sensor – observes at the around the same time of the day for each observation
- This means that even with tidal variations, we most likely will only observe a portion of the full tidal range
- We can characterise this as highest (HOT) and lowest (LOT) observed tide



The Intertidal Extents Model (ITEM) Process

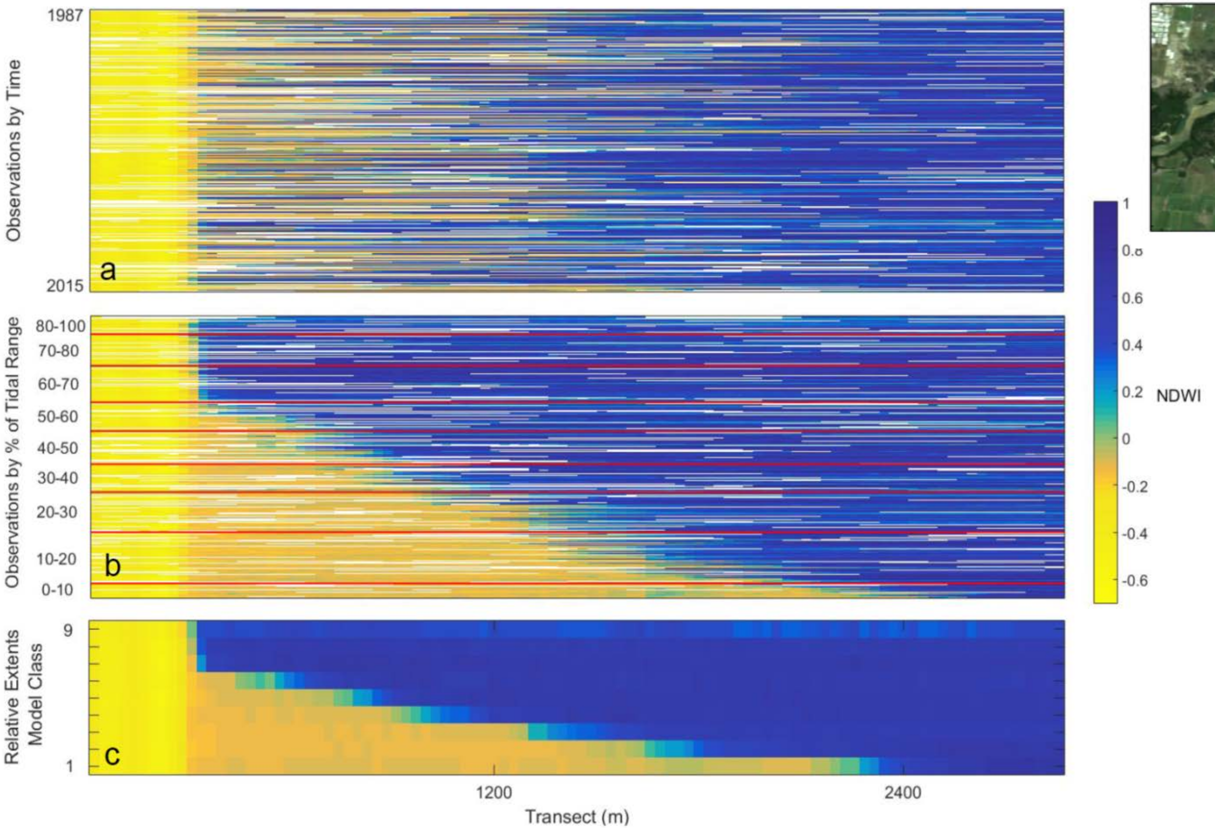
- Each tile stack of observations is attributed with a tidal height utilising the OTPS model
- Observations are the reordered based on tidal height rather than time

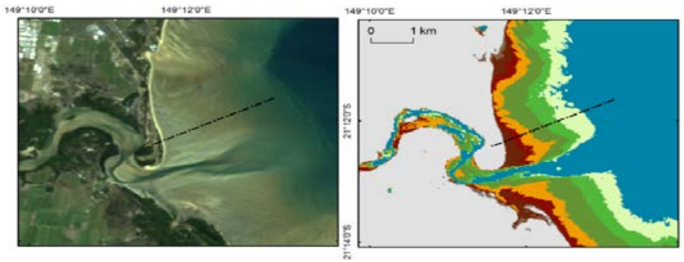


Lowest Observed Tide → Highest Observed Tide

- The Observed Tidal Range is divided into 10 equal interval buckets to create ensemble stacks of observations for each 10% of the range

Dealing with incomplete and noisy data

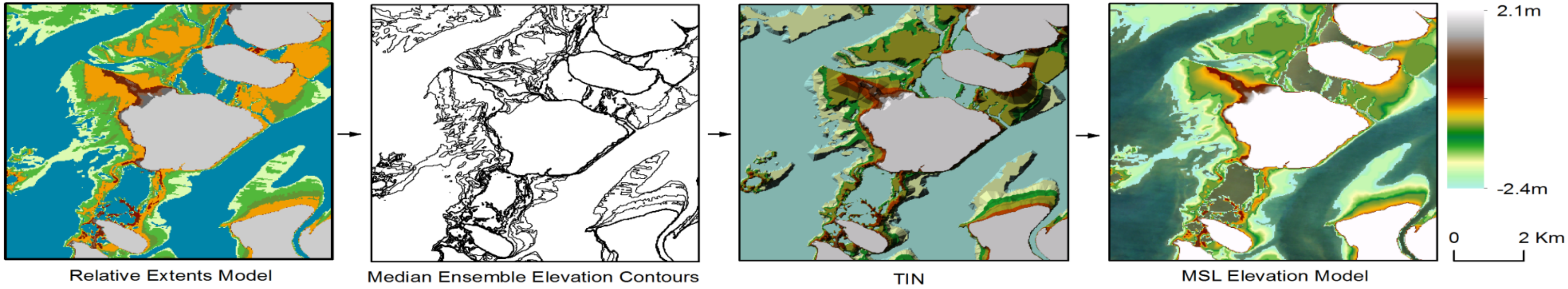




Deriving a Digital Elevation Model

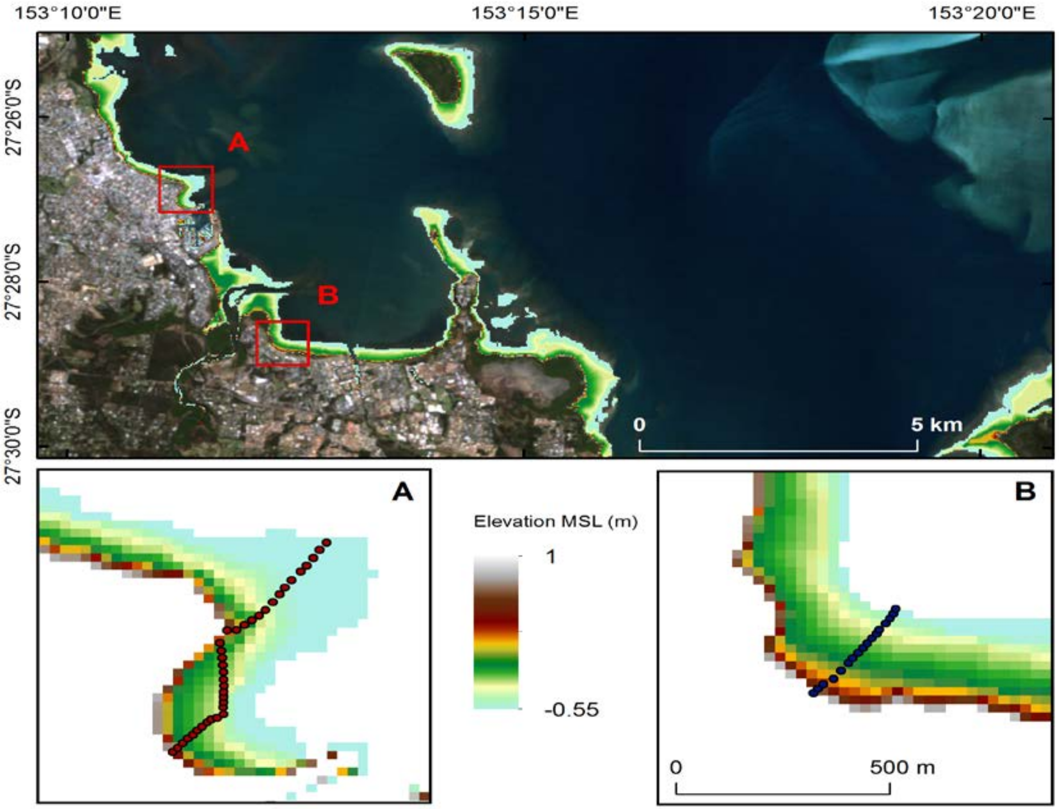
Tidal height attribution to each individual observation enable us to evaluate the distribution of the heights within each interval

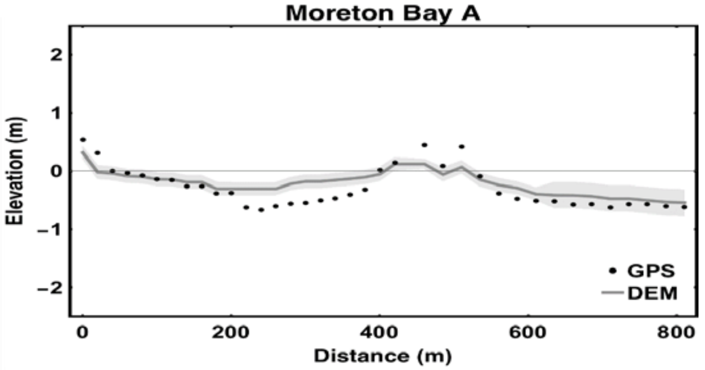
Each distribution corresponds to the extents of the interval in the Relative Extents model



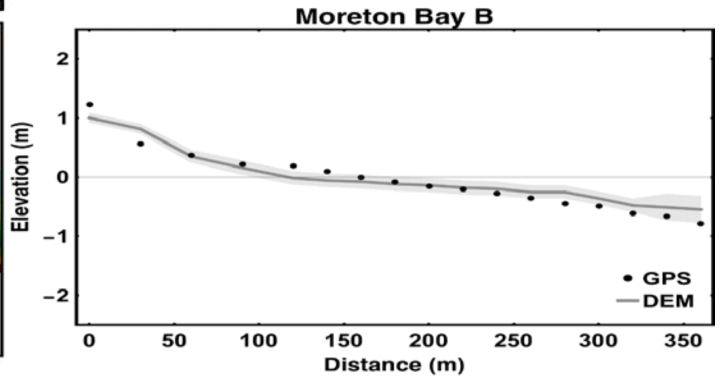
Median heights are used to attribute waterline contours at the boundaries of the intervals extents. The uncertainty of the extent area is reflected in the standard deviation of the heights used to model the interval.

Model Validation - Moreton Bay, Queensland

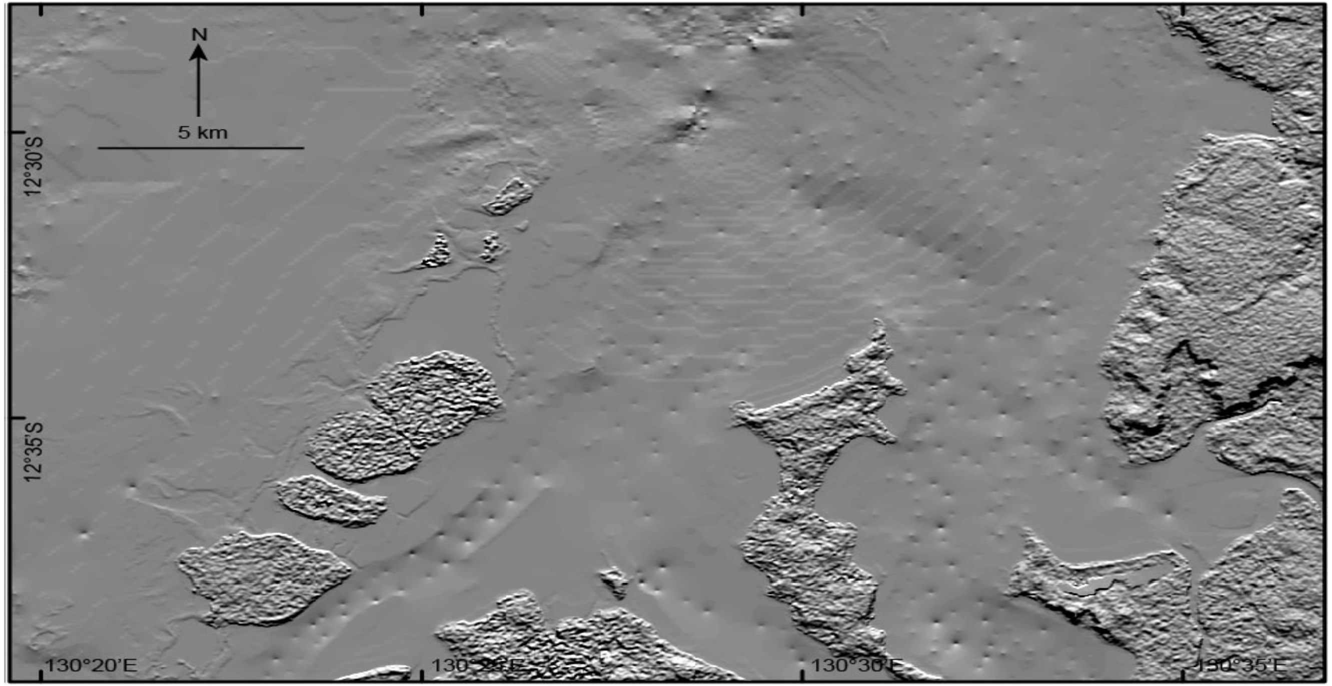




Mean Absolute Height Residual = 0.15m



Integrating the Intertidal DEM in the Northern Australian 100m Bathymetry Grid



Open Data, Code and Getting Access

The tools for use are available by requesting access to project wd8 on the NCI. Compute and storage resources need to be arranged separately.

Data is available as individual files on the NCI THREDDS catalogue:

http://dap.nci.org.au/thredds/catalog.html

Open source code repositories are available through GitHub:

https://github.com/data-cube

Datacube web page: www.datacube.org.au

AEOCCG Technical Capacity Building Webinar - Accessing the Australian Geoscience Data Cube: AEOCCG webinars and presentations page.

Questions?