Rapid analysis of extreme coastal events

Through enhanced data and tools

Claire Trenham
18 October 2018

Coastal extremes eResearch project

- Background
- Problem
- Project
- Solutions and impacts
- Delivery
- Outcome

Background: Coastal extremes

- Risks from coastal hazards are from storm waves and flooding
  - Sea level rise
  - Local risk factors
    - Astronomical tides
    - Severe weather-induced storm surge (low pressure + strong wind)
    - Wind and swell waves
    - Intraseasonal and interannual variations

Image credit: Fairfax Media, Peter Rae

- Want to understand and predict coastal sea levels
  - Proliferation of observed and modelled data
- For more information see https://research.csiro.au/slrwavescoast

Problem: Data management and use

- The Sea Level, Waves & Coastal Extremes team maintain and archive many relevant datasets…
  - Some stored on tape
  - Some data structures make analysis prohibitively slow
  - Analysis code held by individual researchers
  - Datasets not regularly updated

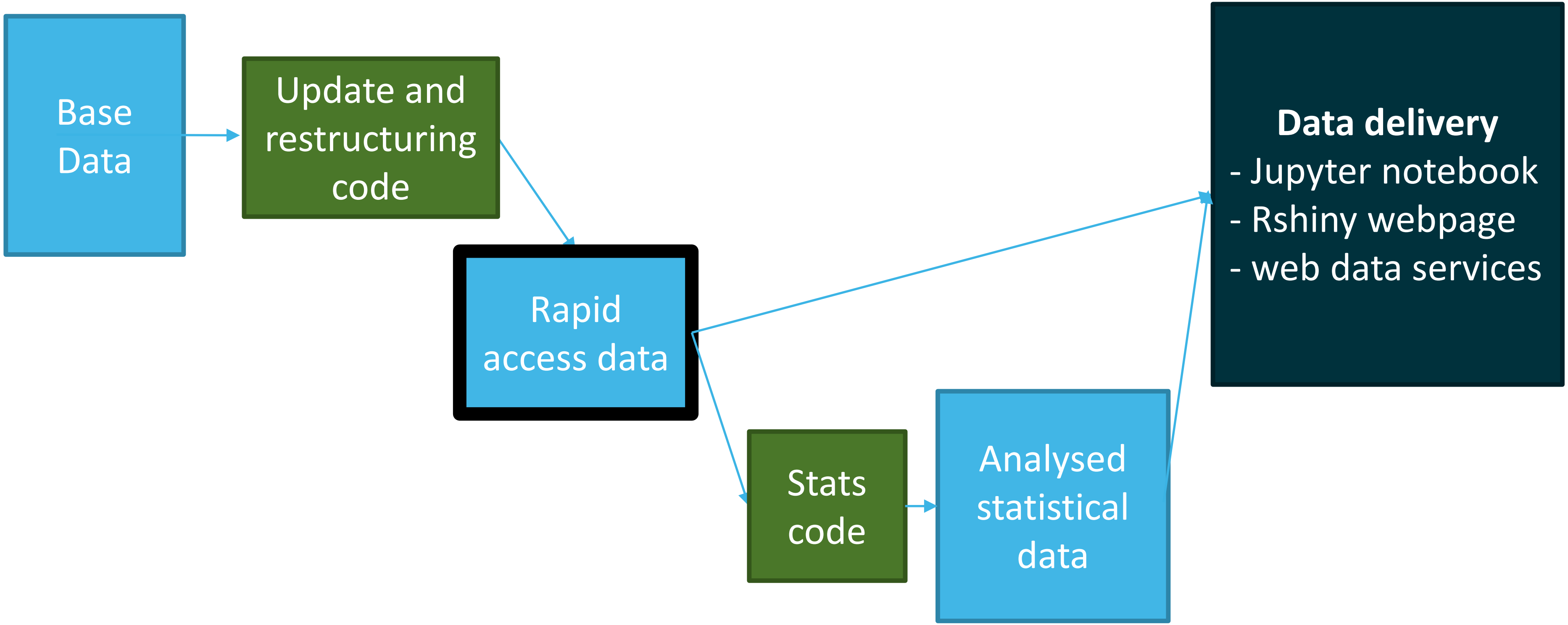
- No systematic and computationally efficient means to access the data to understand their contributions to extreme coastal sea levels for multiple locations or large gridded regions

eResearch Coastal & Maritime Extremes project

- Set up eResearch project between NESP Earth Systems & Climate Change Hub Coastal Hazards project and IM&T to fix these problems!
  - Collate list of datasets
  - Co-locate data on fast-access media
  - Make accessible via web services where possible
  - Optimise data organisation and extraction (particularly wave hindcast)
  - Establish version control around code to process and analyse data
  - Create software to perform analyses
  - Develop web tools to deliver data products

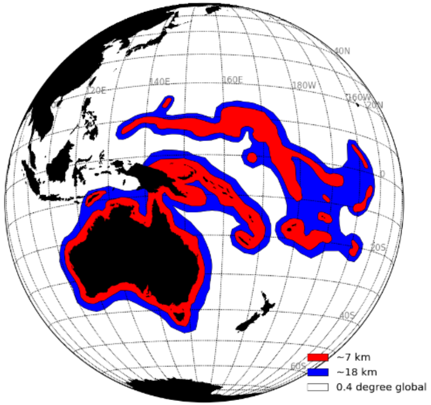
Earth Systems and Climate Change Hub
National Environmental Science Programme

Flowchart: Rapid analysis of coastal extremes



Solutions: Data restructuring

- CAWCR Wave Hindcast
  - Published in CSIRO DAP
  - 1979-present on 5 grids
  - ~20TB of monthly files
  - Not good for timeseries or extreme values analysis
- Restructure
  - Spatial tiles
  - Full timeseries (concatenate in time)
  - NetCDF chunking
  - Needs a lot of memory so run on CSIRO HPC (Ruby)



Solutions: Data restructuring 2

- Run automatically
  - monthly updates
  - data produced at NCI

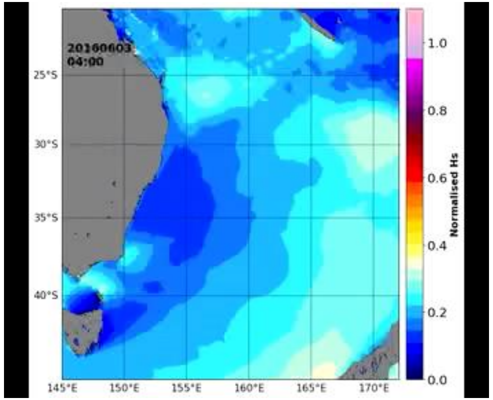
Domain	# Tiles	# Grid points	Time before tiling	Time after tiling
1 point	-	1	90min	0.1sec
glob_24m	60	196k	Not possible	20min
pac_4m	31	138k	Not possible	24min
aus_4m	25	67k	Not possible	10min

- Computational impacts:
  - Time to extract time series for a single grid point cut from 90min to 0.1sec.
  - Whole of Australia extreme values can now be calculated in parallel on CSIRO HPC (Pearcey) in ~10min

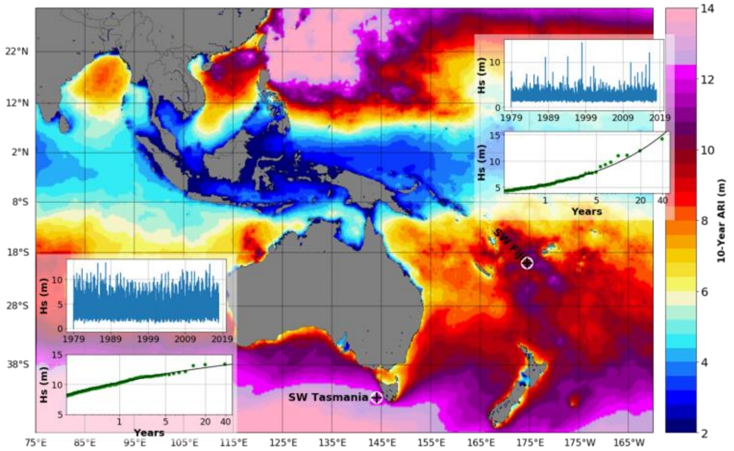
- Work done by Robert Davy, IM&T, presented at C3DIS 2018.

Solutions: Data restructuring 3

- Science impacts:
  - Rapid analysis of extreme coastal events
    - Extract area/time of interest
    - Interpolate grids
    - Calculate return period values
    - Analyse
    - Animate



June 2016 East Coast storm wave height plotted on a scale of 50-year return height



10-year average return interval (ARI) significant wave height (Hs), calculated from a generalized pareto (GPD) fit of all hourly data for years 1979 –2018 tiles in the map region. Two example locations, south-west (SW) Tasmania and south-west (SW) Fiji archipelago illustrate the full hourly timeseries (upper plots) and empirical and fitted ARI values (green circles and black lines, respectively, in lower plots), both sites are plotted on the same scale.

Image credits: Ron Hoeke

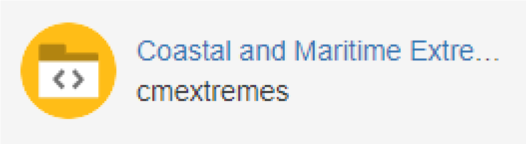
Solutions: Data aggregation

- Identify datasets needed for coastal extremes analysis
- Co-locate datasets on CSIRO Bowen disk storage
- Automatic ingest/updates (as approp.)
- Impacts:
  - Data organised in an OSM
  - No tape recall delays
  - Easy to use for analyses
  - Can make available via data services

Dataset	Status
Sea level	
CSIRO sea level products (reconstructions and combined altimetery)	partial
Copernicus Marine environment monitoring service altimeter sea level anomaly (NRT and REP).	NA
Global Extreme Sea Level Analysis GESLA-2 database (Tide gauge based).	partial
Sea level rise projections, e.g. Zhang, et al. 2017	NA
Waves	
CAWCR Wave Hindcast (gridded products)	complete
CAWCR Wave Hindcast (spectral products)	partial
Coordinated Ocean Wave Climate Project (COWCLIP)	partial
BOM AUSWave Forecast	partial
NOAA NCEP WaveWatch III Forecast	NA
Tides and short-term variables:	
TPXO8.0 global astronomic tidal inversion	partial
FES2014 global tidal model	NA
CFSR/CFSR2 surface winds and sea level pressure (SLP)	Near-complete
Australia storm surge ROMS model (Colberg et al. 2018)	partial
Australia storm surge Mike21 model (Haigh et al. 2012)	partial

Solutions: Code development

- Establish bitbucket project for code sharing and version control
  - Includes repos for data downloads, tiling, extremes analysis, web app
  - Improve our team's use of version control
  - collaborative coding

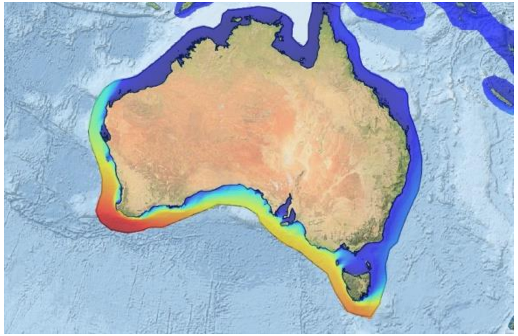


- HPC scripts for data crunching
  - Data tiling
  - Extreme value and return period calculation
  - Tide prediction
- Jupyter Notebooks to study storms interactively
- Impacts:
  - Reproducibility

Delivery: Data services

Data Access Portal

- Publish data through CSIRO's Data Access Portal
  - monthly CAWCR Wave Hindcast
  - Australian Coastal Sealevel Simulations (storm surge)
- Data accessible via web services
- Inclusion in tools like National Map
  - Australian Wave Energy Atlas



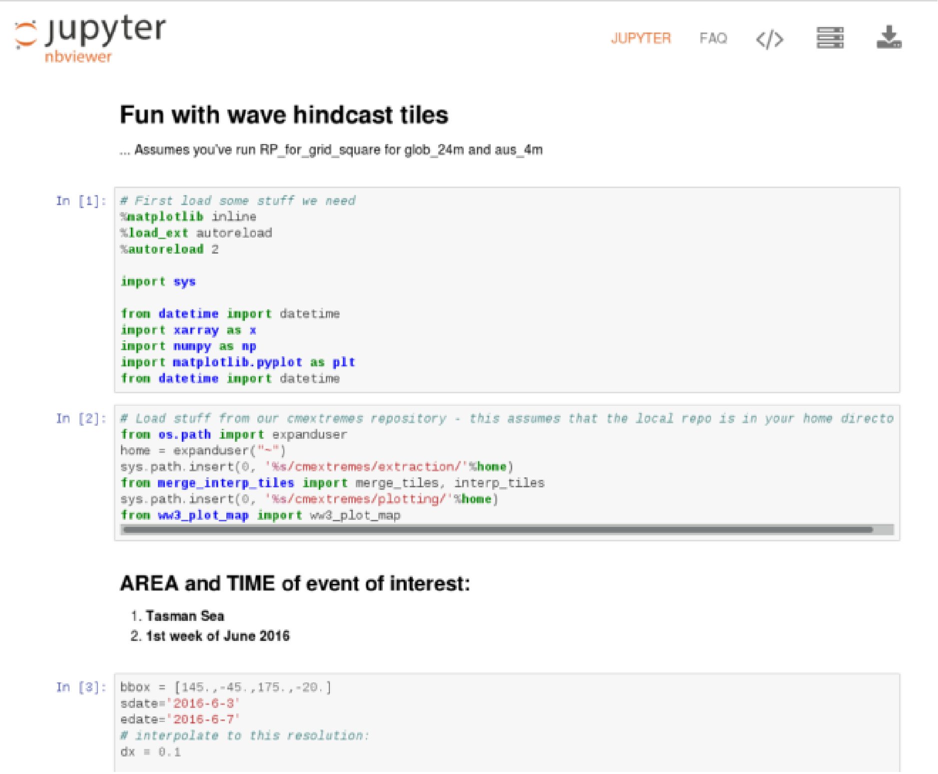
- Impacts:
  - Open science is enabled through open data

THREDDS OPeNDAP

- VM hosting data and code for specialised applications

Delivery: Jupyter notebooks

- Excellent for exploratory development, data analysis, and visualisation
- Notebooks can be rendered on the web to share results
- External peers can run notebooks where they use OPeNDAP URLs rather than file paths



jupyter nbviewer

JUPYTER FAQ </>

Fun with wave hindcast tiles

... Assumes you've run RP_for_grid_square for glob_24m and aus_4m

In [1]:

```
# First load some stuff we need
%matplotlib inline
%load_ext autoreload
%autoreload 2

import sys

from datetime import datetime
import xarray as x
import numpy as np
import matplotlib.pyplot as plt
from datetime import datetime
```

In [2]:

```
# Load stuff from our cmextremes repository - this assumes that the local repo is in your home directo
from os.path import expanduser
home = expanduser("~")
sys.path.insert(0, '%s/cmextremes/extraction/'%home)
from merge_interp_tiles import merge_tiles, interp_tiles
sys.path.insert(0, '%s/cmextremes/plotting/'%home)
from ww3_plot_map import ww3_plot_map
```

AREA and TIME of event of interest:

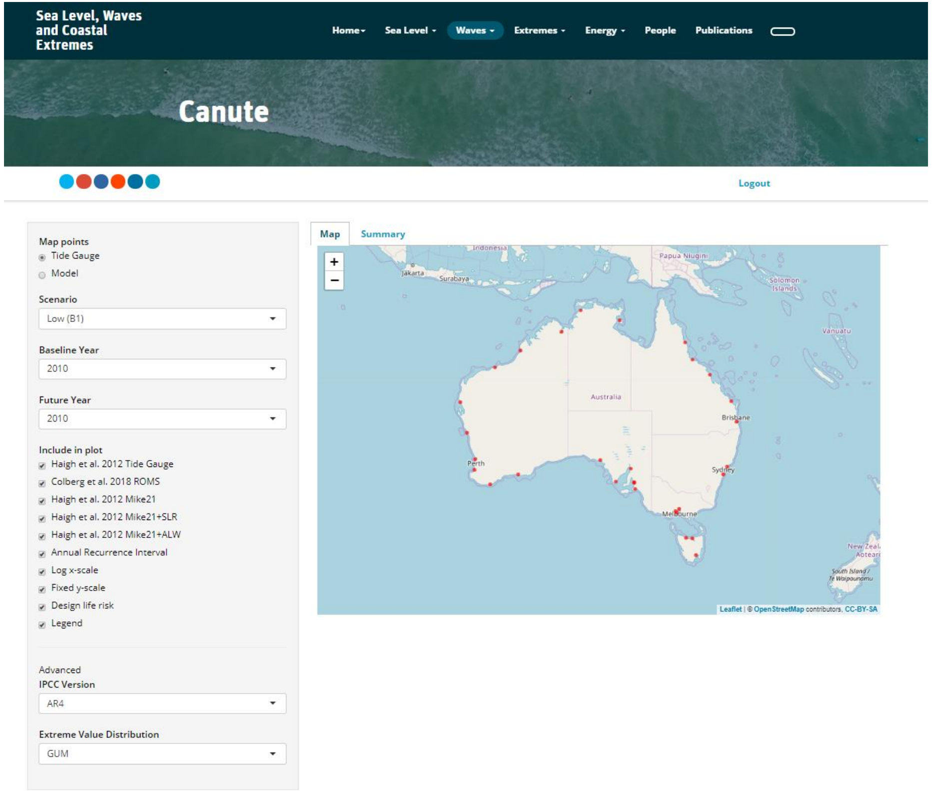
1. Tasman Sea
2. 1st week of June 2016

In [3]:

```
bbox = [145.,-45.,175.,-20.]
sdate='2016-6-3'
edate='2016-6-7'
# interpolate to this resolution:
dx = 0.1
```

Delivery: Rshiny web app

- Canute 3.0 Sea level projections
  - Builds on collaborators' work
  - Incorporates additional datasets
  - R used to produce statistics
    - Annual recurrence intervals (1-in-X year sea levels)
  - Rshiny for web delivery

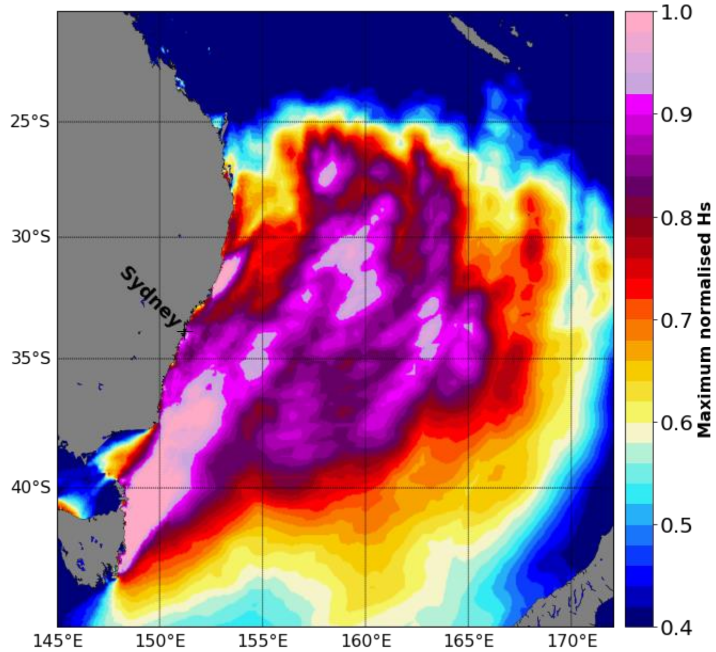


Credit: Julian O'Grady & Tim Erwin

See talk at FOSS4G-Oceania, Nov 2018

Outcome

- Maintenance of up-to-date scientific datasets
- Rapid assessment of output for various applications
  - Now have the capability to perform extreme value analysis of gridded wave hindcast output in parallel.
- Data ready for integration into other coastal hazard assessment tools



Extent of waves during East Coast Low in June 2016 normalised by 1-in-50 year wave height. Extract data, calculate statistics, combine grids and make plots in ~20min.

Thank you

Oceans & Atmosphere
Claire Trenham, Ron Hoeke, Julian O'Grady, Rebecca Gregory, Kathy McInnes & Mark Hemer

t +61 3 9239 4442
e Claire.Trenham@csiro.au
w research.csiro.au/slrwavescoast

Information Management & Technology
Robert Davy