

# EO for agriculture, environment and oceans: Developing data science methods and capability

Kerrie Mengersen  
Jacinta Holloway Brown

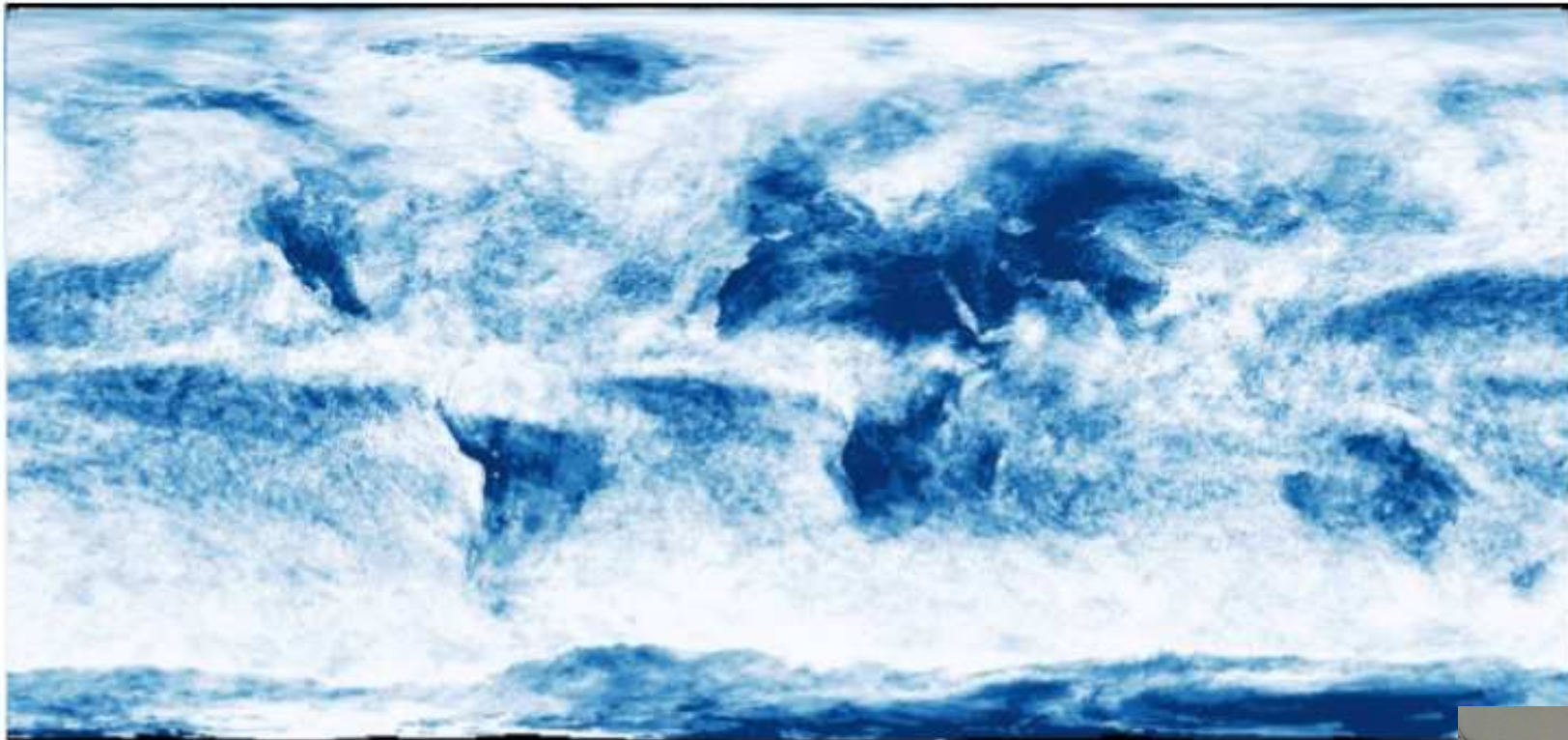
2 September 2020



# Examples of Australian contributors



# Missing data



NASA Cloud Fraction product, which shows average cloud cover for 25 May-1 June 2020.

Source: [https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2\\_M\\_CLD\\_FR](https://neo.sci.gsfc.nasa.gov/view.php?datasetId=MODAL2_M_CLD_FR)



# Satellite images, clouds and missing data

Need faster statistical ways to fill in these gaps.

Why not use radar images?

➡ These images are black and white, and can be very noisy.

Why not wait for the next clear image?

➡ In tropical regions there can be weeks or months of persistent cloud cover (Zhu, 2018, Asner 2001).



**METHODOLOGY**

**Open Access**

# Stochastic spatial random forest (SS-RF) for interpolating probabilities of missing land cover data



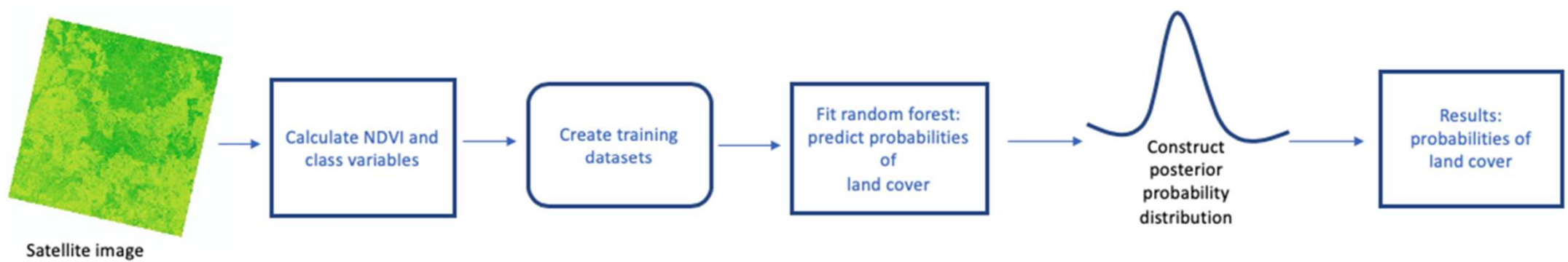
Jacinta Holloway-Brown<sup>1,2\*</sup> , Kate J Helmstedt<sup>1,2</sup> and Kerrie L Mengersen<sup>1,2</sup>

- Fast, scalable and accurate for classifying satellite images.
- Uses machine learning algorithm: random forest.
- Fills missing data gaps and has comparable performance with remote sensing approaches.
- Quantifies uncertainty of classifications.

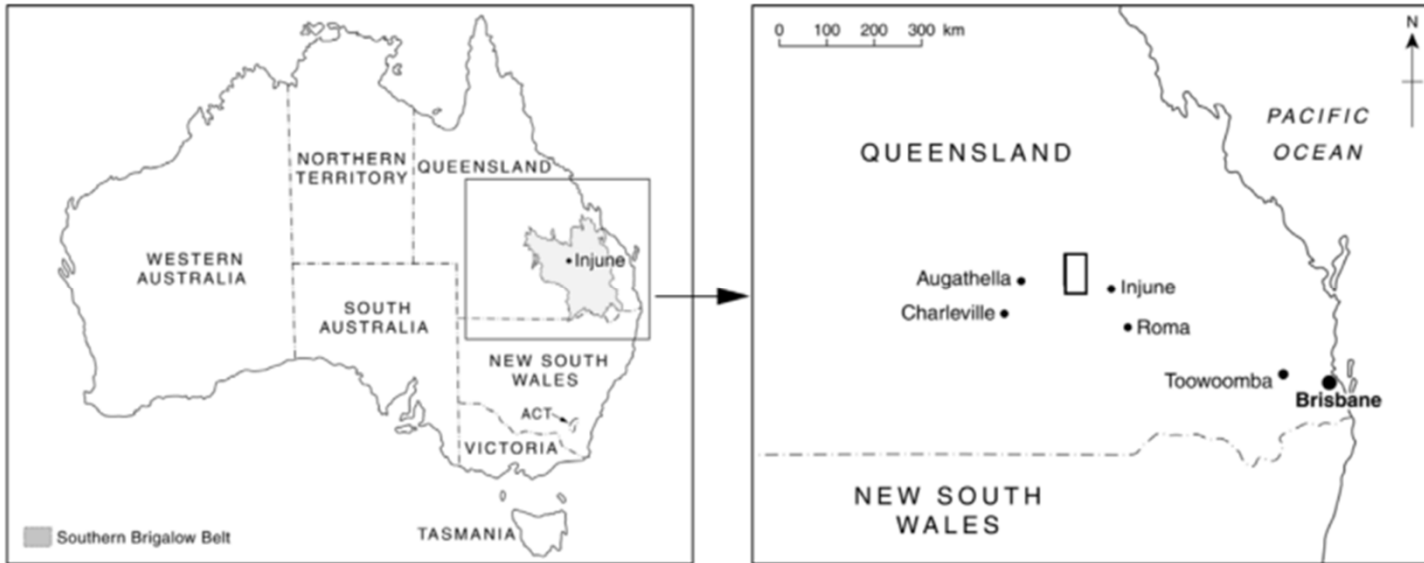




# SS-RF Method Overview



# Application: interpolating missing image data



Landsat 7 satellite image of Injune site.

Location of the Injune study area in Queensland.

37 x 60km area.

Mainly forest and agricultural land.



# Binary forest and not forest cover at pixel scale

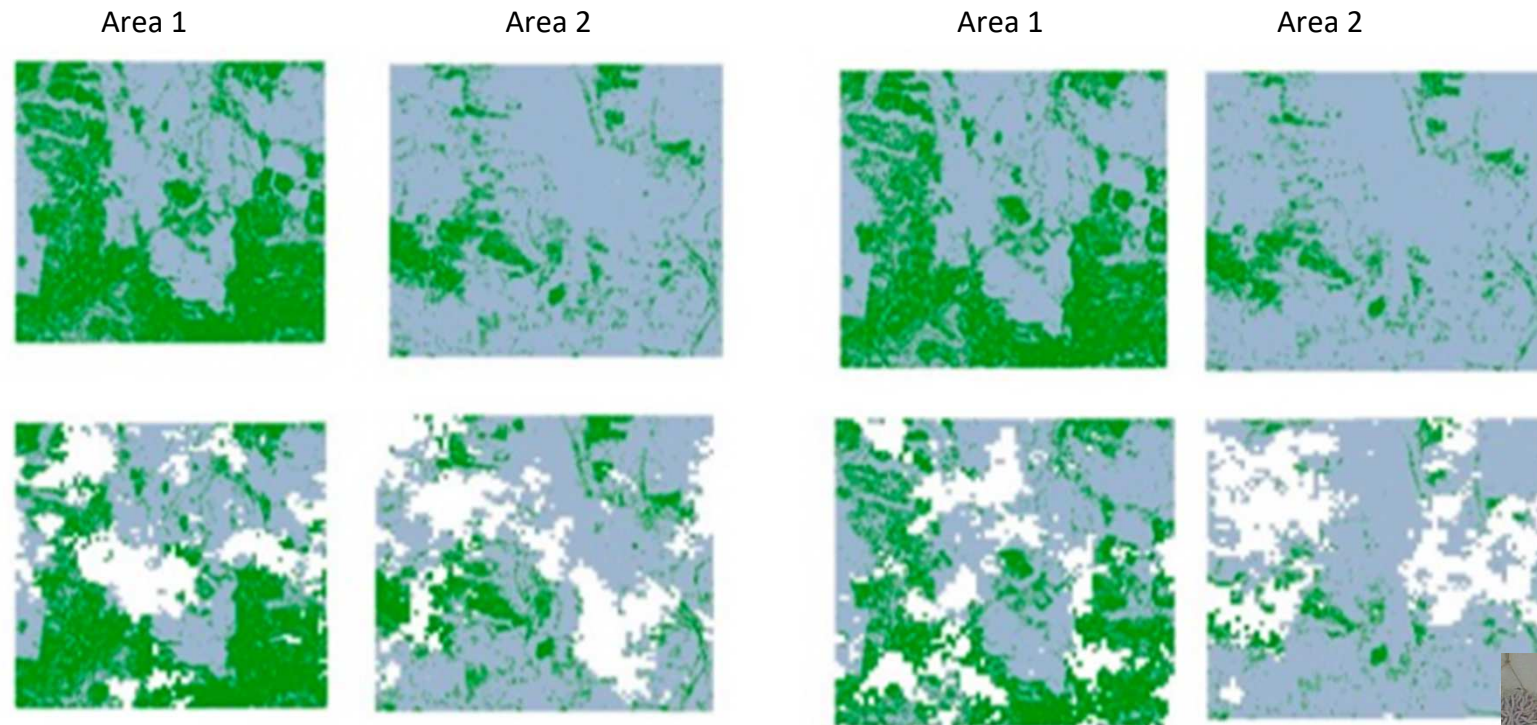


Image date: 25 August 2017

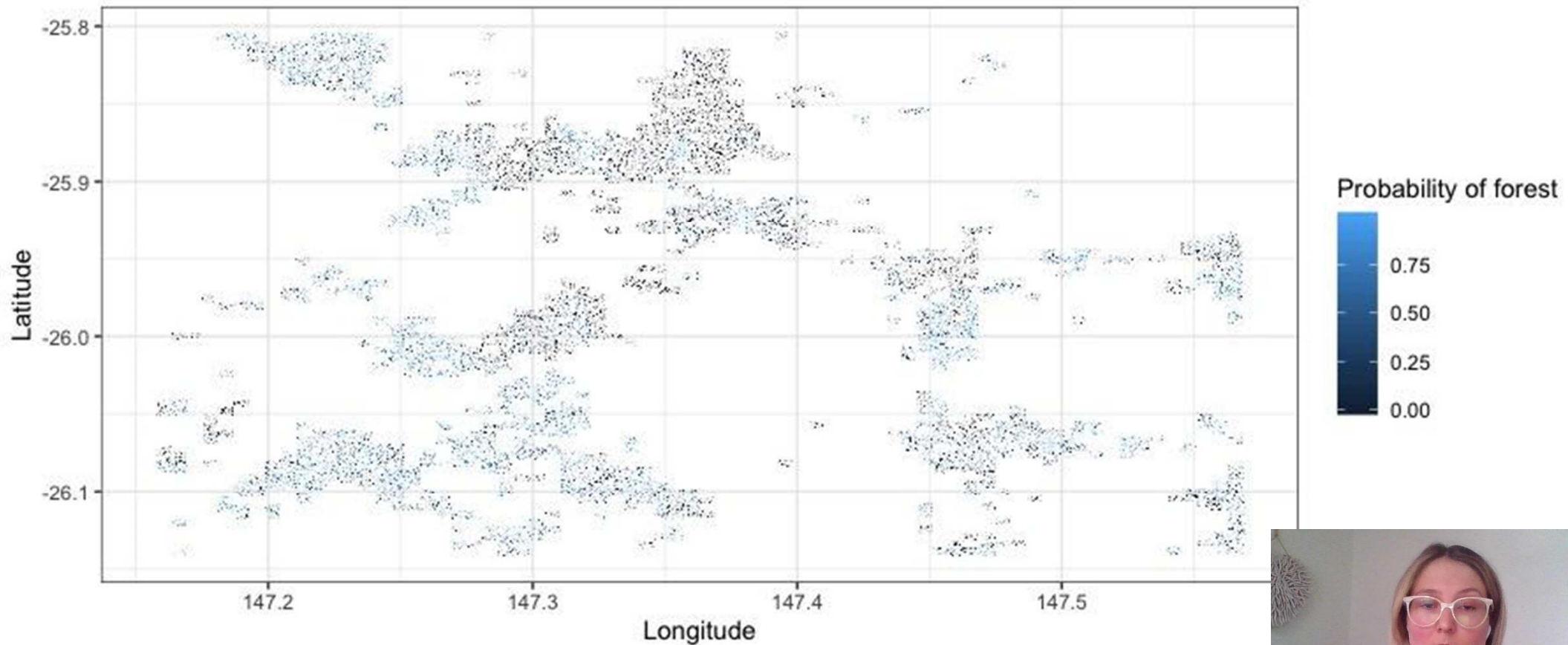
Image date: 10 September 2017





# Example of results

Random forest probability values for sample 5 at t+2 10 September 2017 Area 1



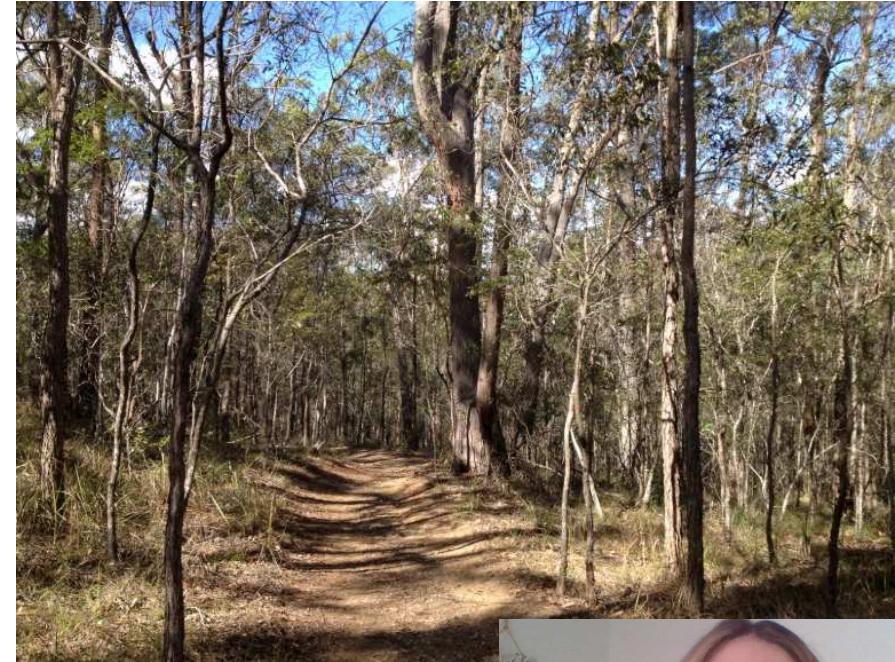
# Summary

- The SS-RF method can accurately interpolate missing data when land cover is relatively stable over time ( $>0.90$  overall accuracy).
- When significant change occurs between images, the SS-RF method tends to have low accuracy.
- The measurement of uncertainty (probability) for the interpolated values is an advantage over other interpolation methods.
- SS-RF can be implemented as a Bayesian updating scheme, as new images are available.



# Extensions

- We have extended the SS-RF approach to accurately detect change events (simulated forest clearing).
- This could extend to other challenges in environment, agriculture and oceans.
- Our method is one example of new developments in EO analysis. There is opportunity for many different types of new methods and technologies to be developed in this field.





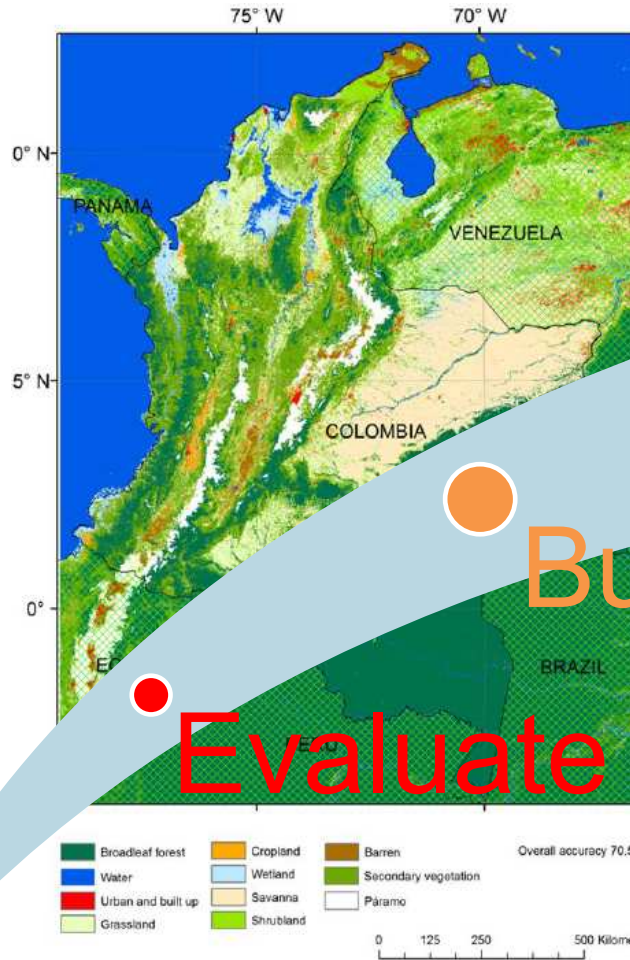
Earth Observations for Official Statistics

Satellite Imagery and Geospatial Data Task Team report

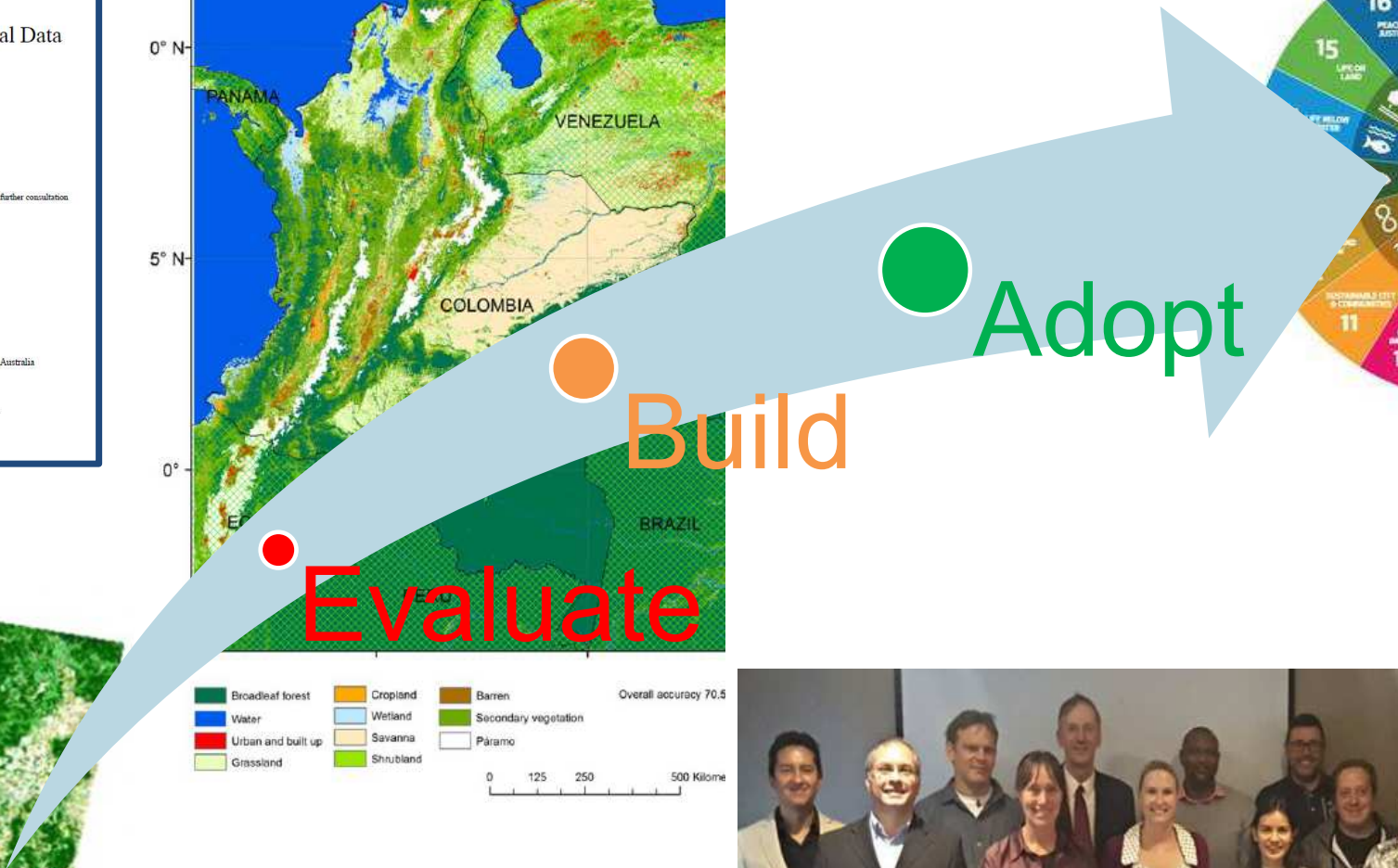
5<sup>th</sup> December 2017

White cover publication, pre-edited text subject to official editing and subject to further consultation

United Nations  
 Australian Bureau of Statistics  
 Queensland University of Technology, Australia  
 Queensland Government, Australia  
 Commonwealth Scientific and Industrial Research Organisation, Australia  
 European Commission – DG Eurostat  
 National Institute of Statistics and Geography, Mexico  
 Statistics Canada



# The EO journey



# The EO analysis journey



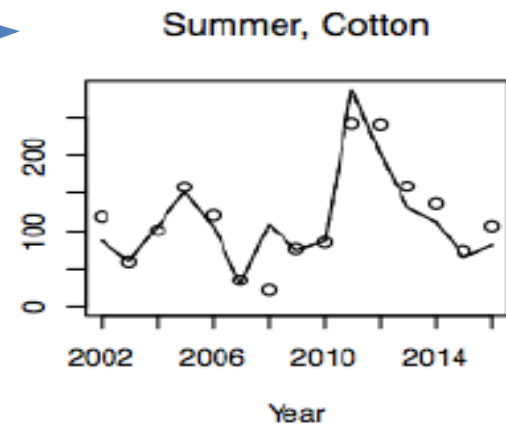
Describe

Estimate



Prescribe

Predict



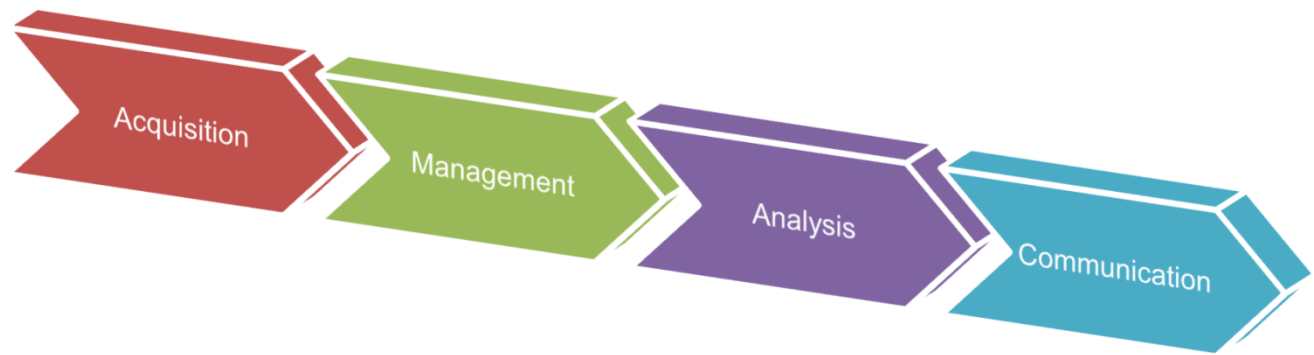


# Research Horizons in EO Data Analysis



# Next Steps – a “call to arms”

1. Collaborative research
2. Success stories
3. Training material
4. Early adopters



[k.mengersen@qut.edu.au](mailto:k.mengersen@qut.edu.au)

[j1.holloway@qut.edu.au](mailto:j1.holloway@qut.edu.au)

