The parallelization of geoscience apps at C3L with Azure

Prof J Craig Mudge FTSE

Collaborative Cloud Computing Lab (C3L)

THE UNIVERSITY of ADELAIDE

Outline

1. A no-machines lab
2. Magnetotelluric method for imaging sub-surface
3. Cleaning and processing the data - weeks
4. Results
5. Future work

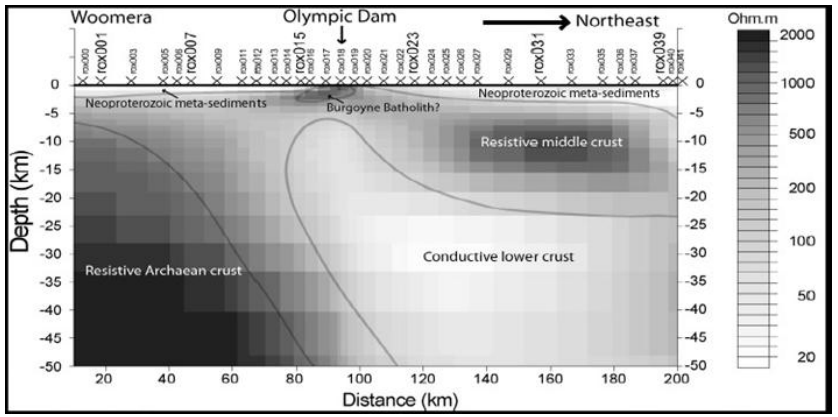
A no-machines Lab

eScience enabled by cloud computing

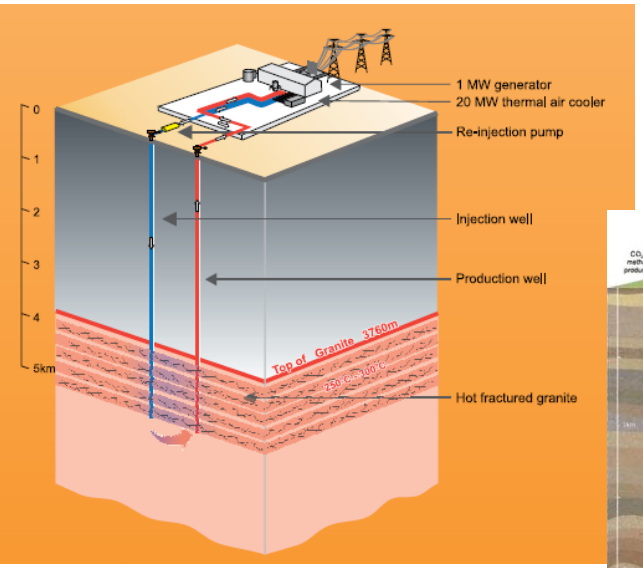
Seed funding from
-- department of mines
www.pir.sa.gov.au
-- MSFT Research Jim Gray Seed Grant

Started June 2010

Magnetotelluric (MT) imaging



1. Using the magnetic and electric fields of the earth, MT imaging determines the resistivity structure of a sub-surface area of interest.
2. It goes deeper (hundred or so Km) than seismic (<2 Km) but does not have the same resolution
3. Applications
    1. mineral exploration,
    2. water management in mining,
    3. geothermal exploration,
    4. carbon storage,
    5. aquifer research and management
    6. earthquake and volcano studies.



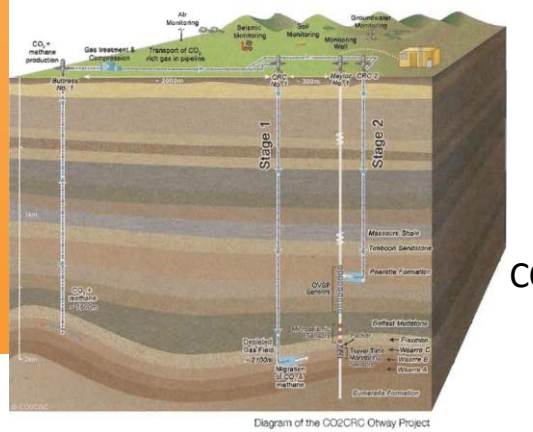


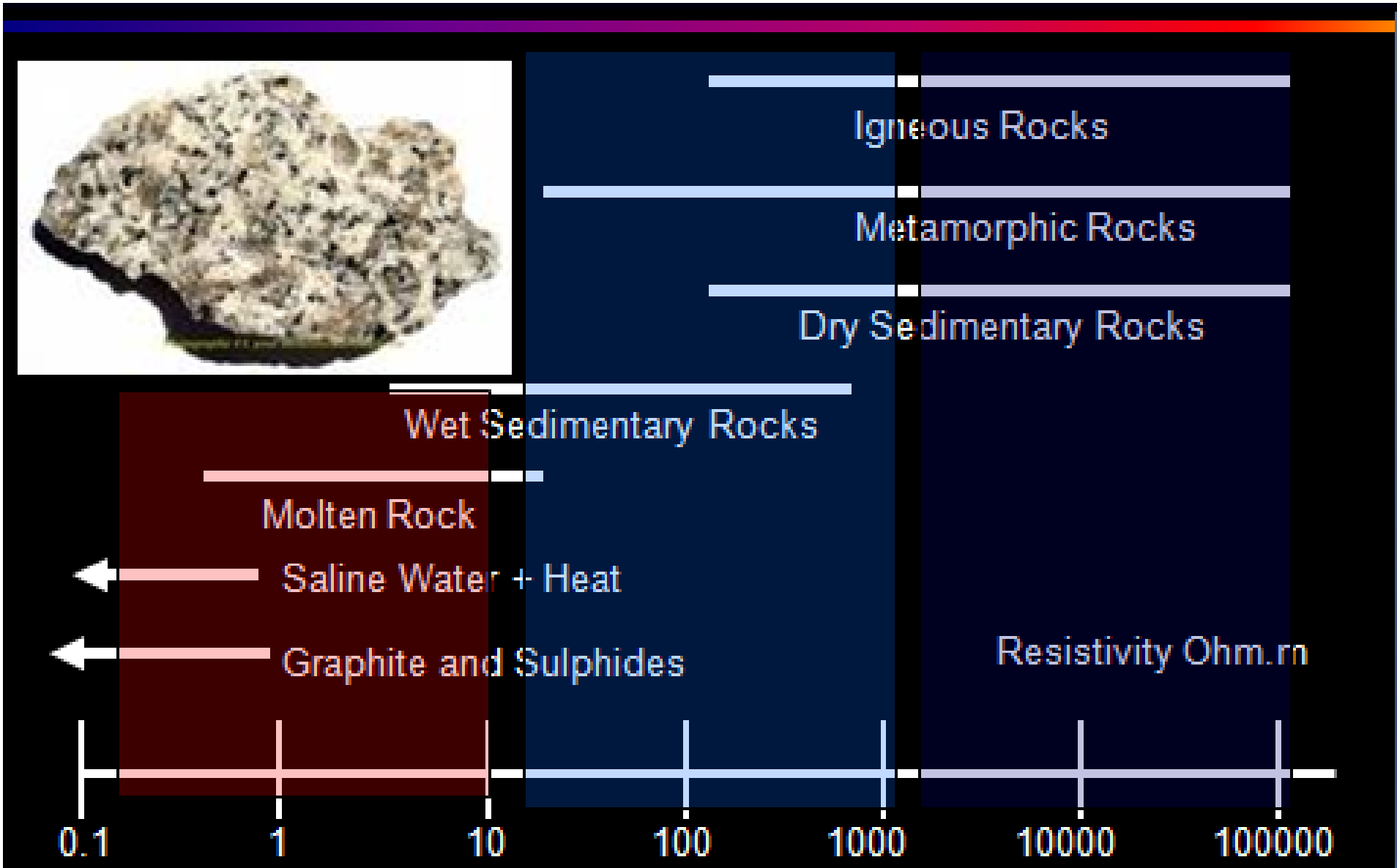
Diagram of the CO2CRC Otway Project

(Heinson and Mudge, 2010)

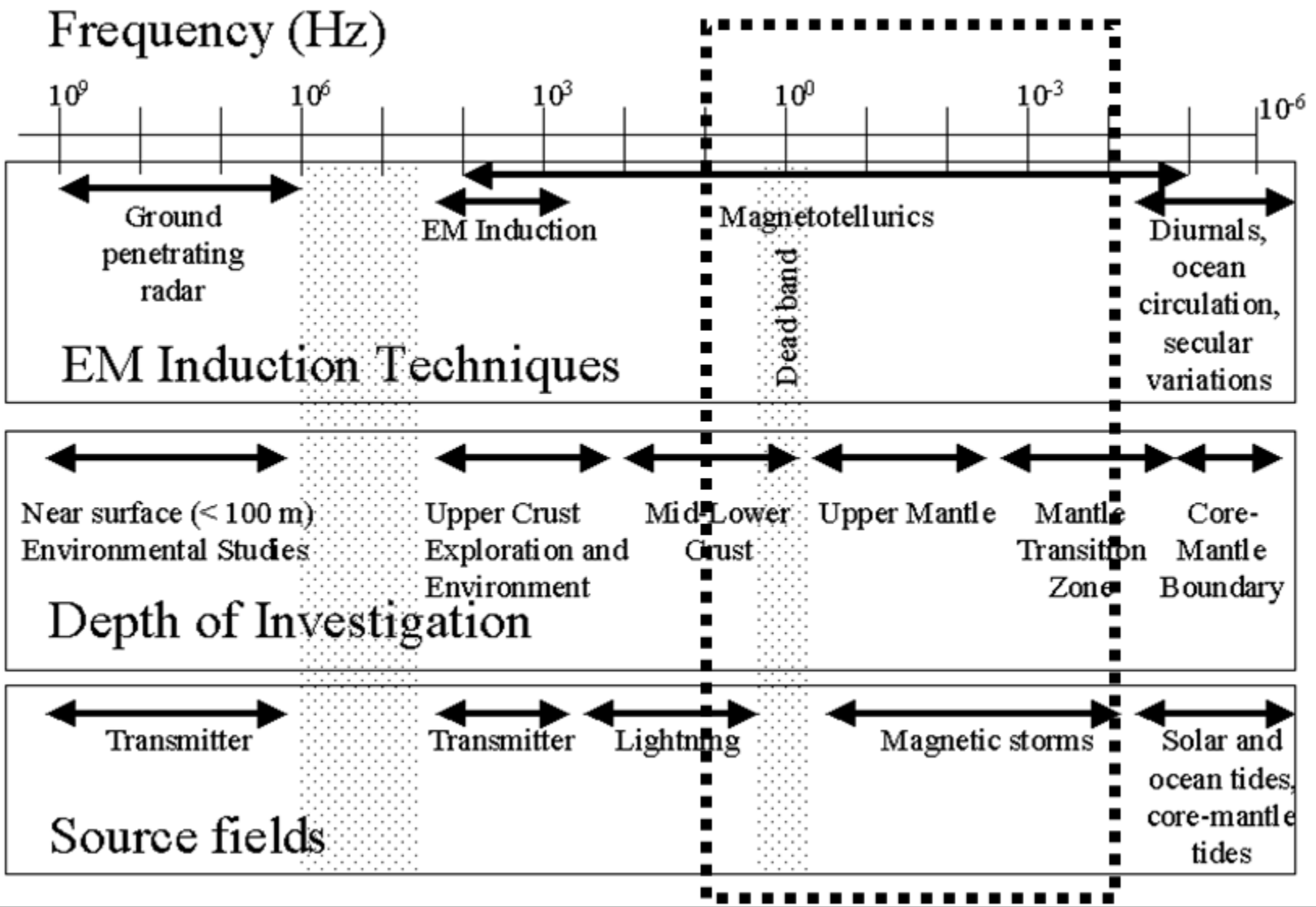
CO_2 in depleted gas field

craig.mudge@adelaide.edu.au
27 sep 2010

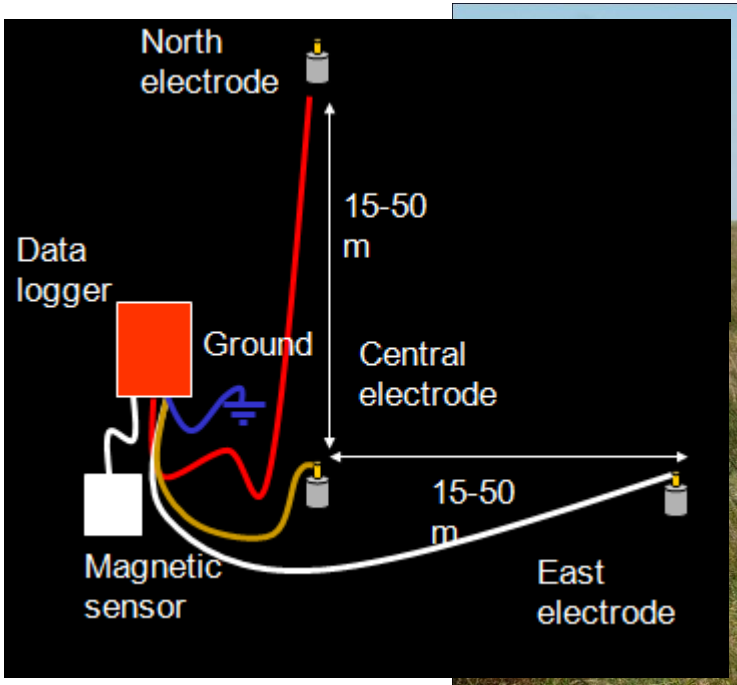
Electrical resistivity



Electromagnetic methods



Data logging by University of Adelaide Geophysics, on a geothermal site – Paralana, SA, Australia



MT Station data from logging in the field

Clean
Broadband processing
E field conversion to standard units

Station 1

BIRRP

Station 2

Station n

Outputs from BIRRP are
(a)impedance Z, where E=ZB
(b)coherence data
(c) Apparent resistivity and phase

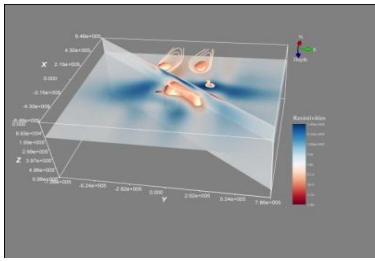
Convert to EDI

Convert to EDI

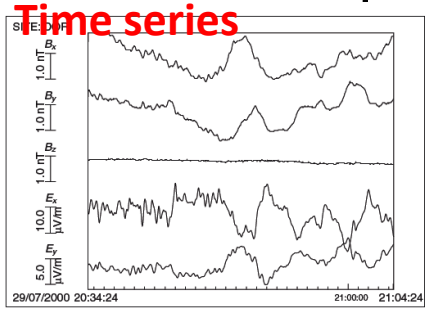
Convert to EDI

3D

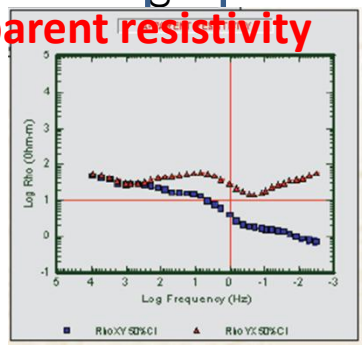
Forward Modeling And Inversion

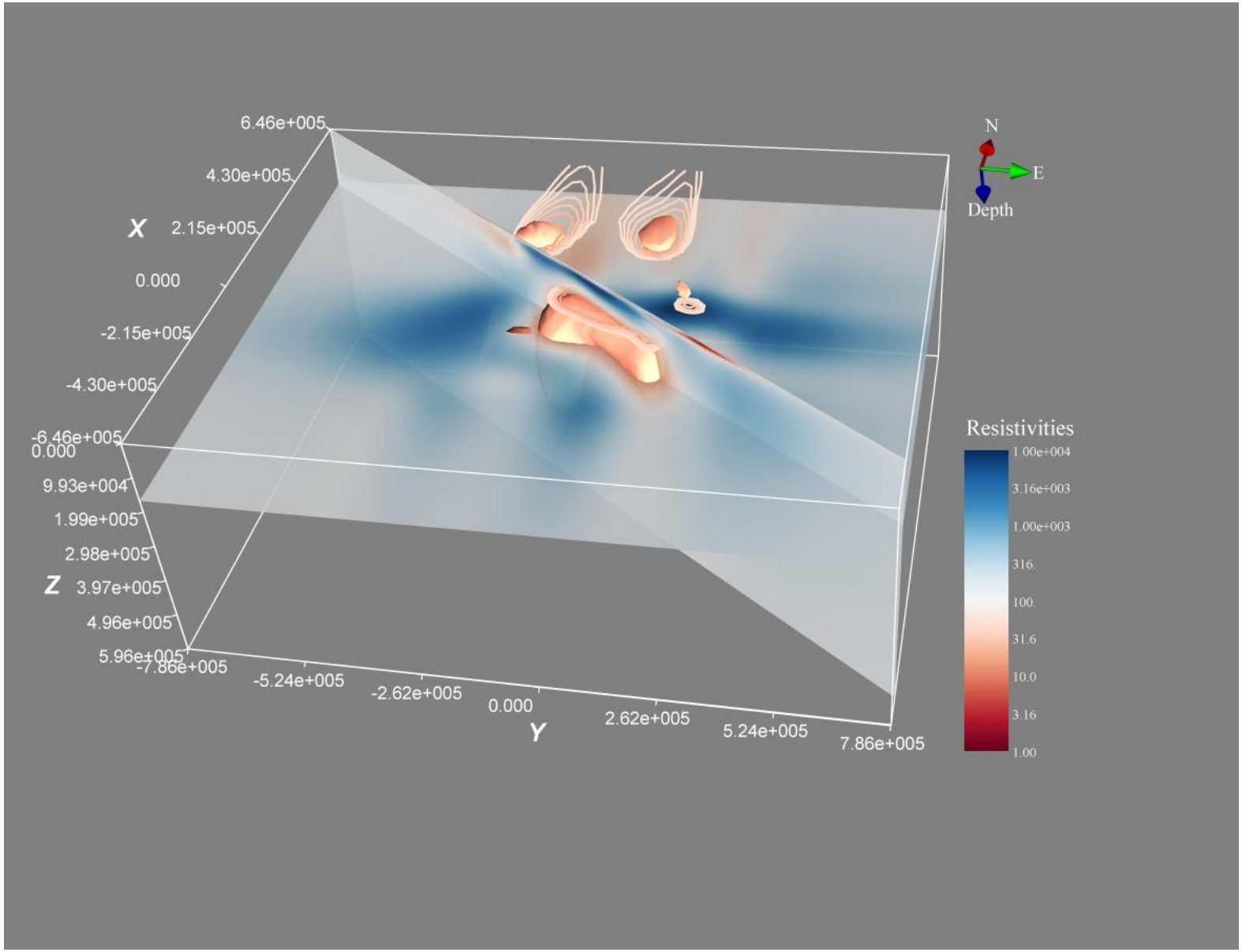


Time series

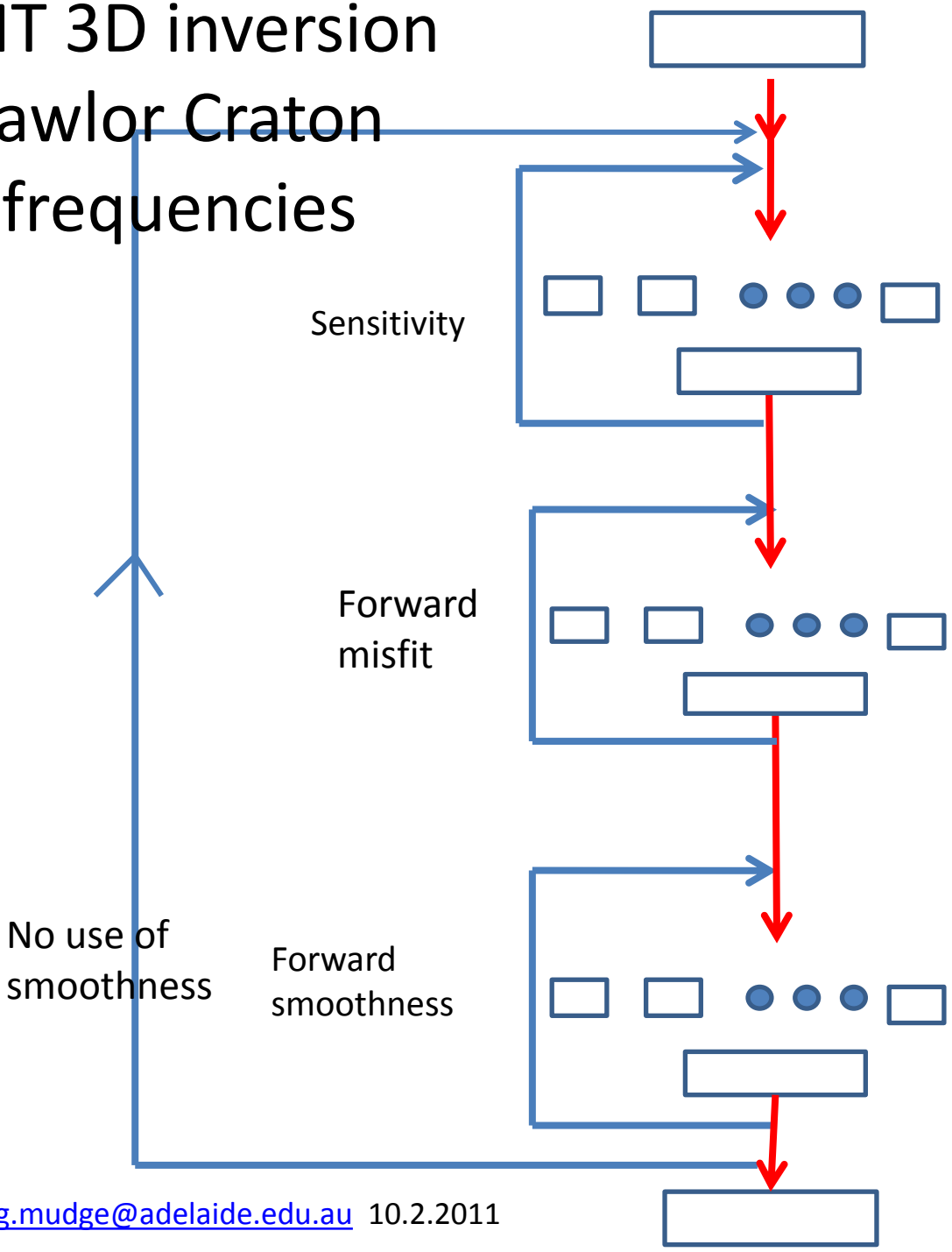


Apparent resistivity




6.46e+005
4.30e+005
X
2.15e+005
0.000
-2.15e+005
-4.30e+005
-6.46e+005
0.000
9.93e+004
1.99e+005
2.98e+005
Z
3.97e+005
4.96e+005
5.96e+005
-7.86e+005
-5.24e+005
-2.62e+005
0.000
Y
2.62e+005
5.24e+005
7.86e+005
N
E
Depth
Resistivities
1.00e+004
3.16e+003
1.00e+003
316.
100.
31.6
10.0
3.16
1.00

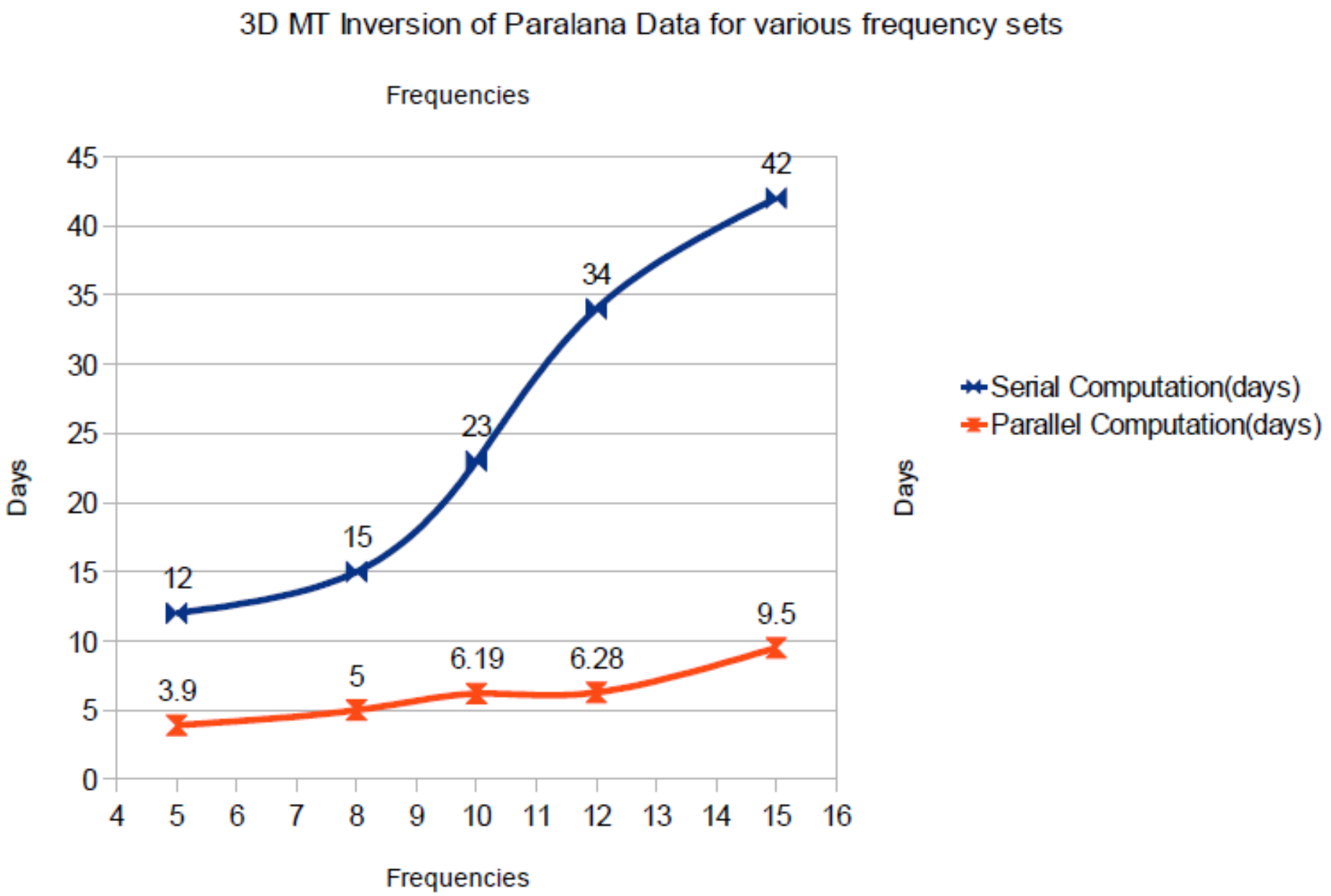
MT 3D inversion
Gawlor Craton
6 frequencies

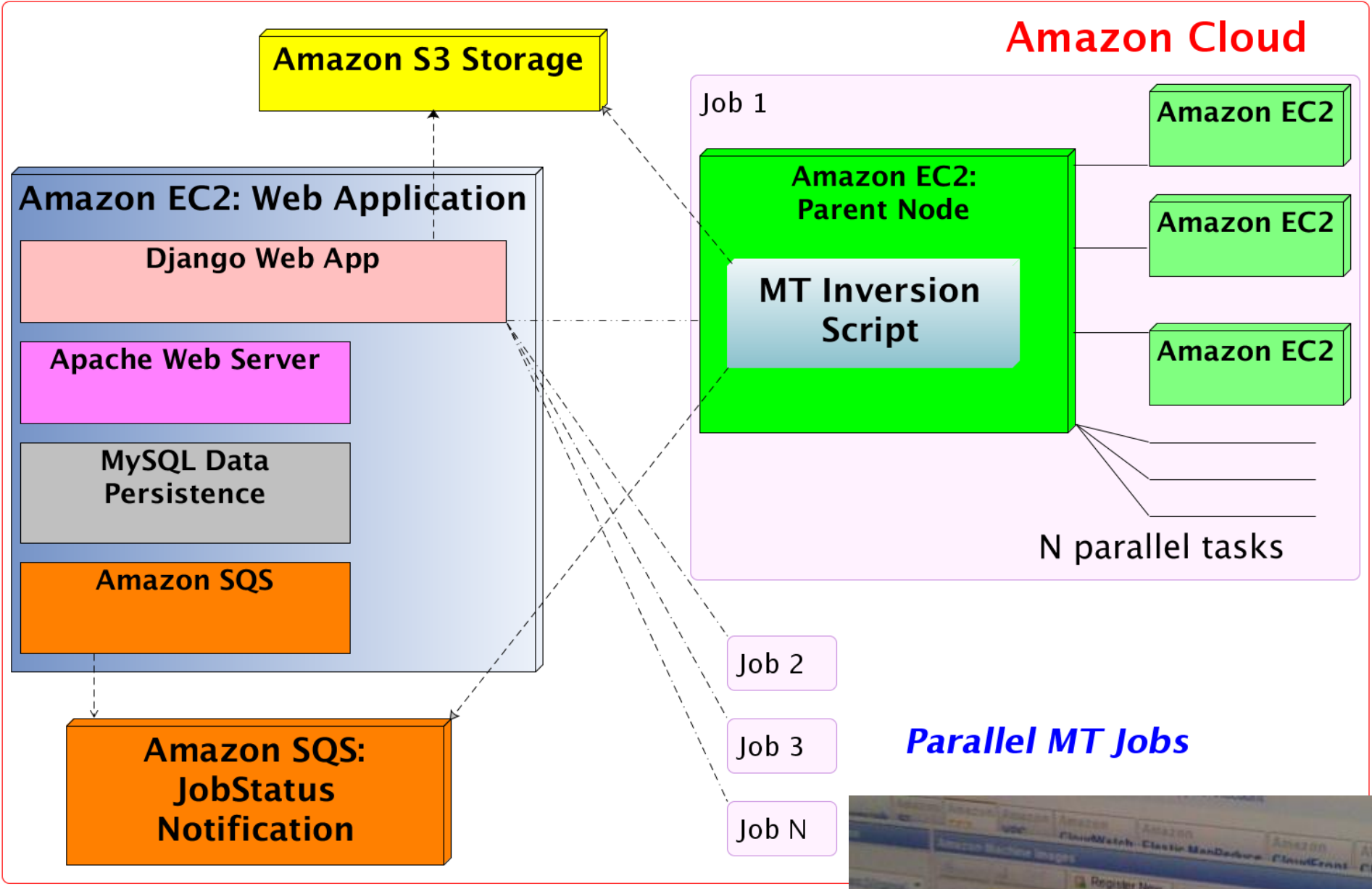


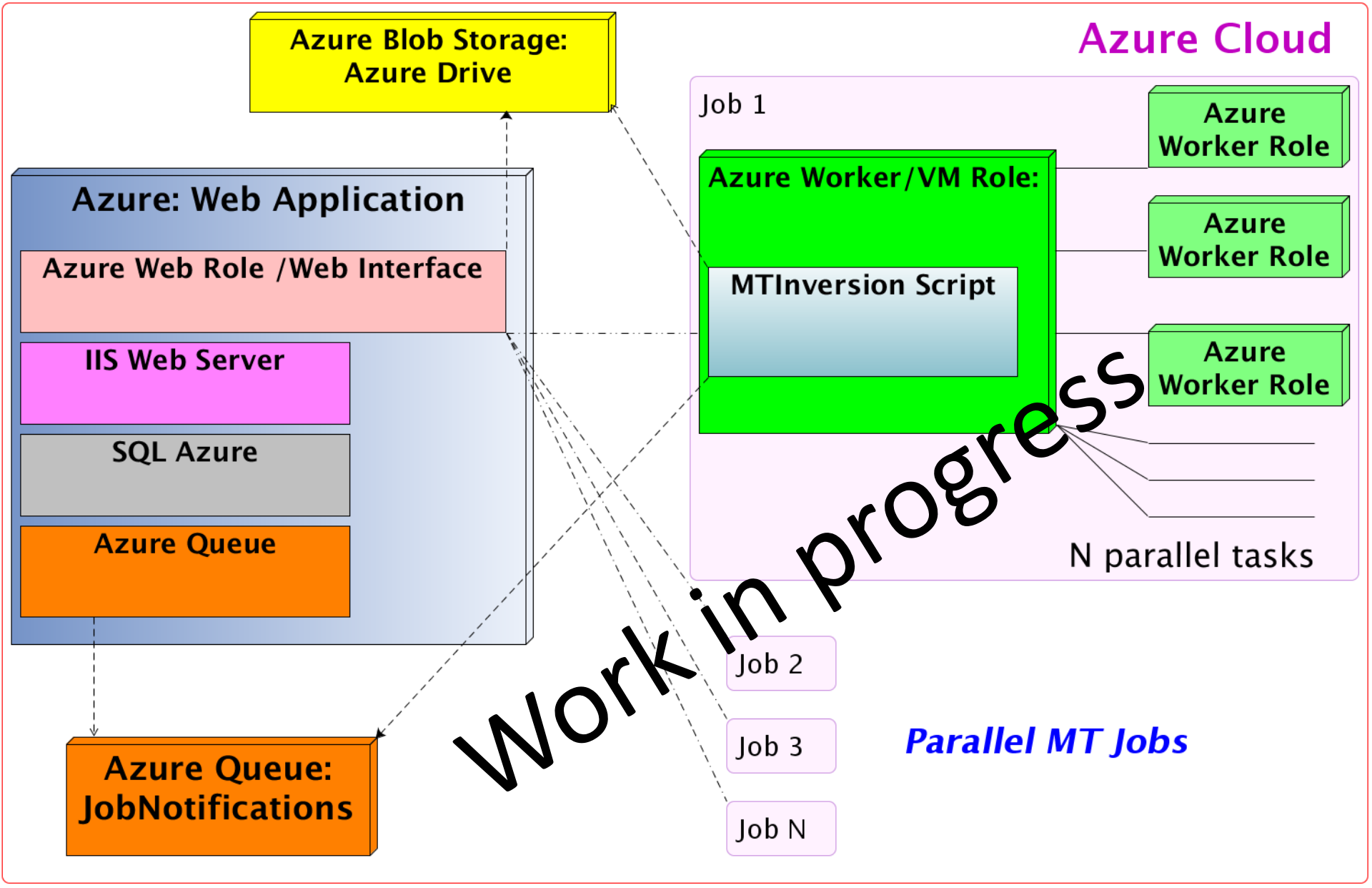
,000 secs

	Sequential	Parallel
	0	0
	82	82
	163	60
	339	46
	0	0
	1	1
Total	584	189

craig.mudge@adelaide.edu.au 10.2.2011


3D MT Inversion of Paralana Data for various frequency sets
Frequencies
Days
45
40
35
30
25
20
15
10
5
0
4
5
6
7
8
9
10
11
12
13
14
15
16
42
34
23
15
12
9.5
6.19
6.28
5
3.9
Frequencies
Days
Serial Computation(days)
Parallel Computation(days)


Amazon Cloud
Amazon S3 Storage
Job 1
Amazon EC2
Amazon EC2: Parent Node
Amazon EC2
MT Inversion Script
Amazon EC2
Amazon EC2: Web Application
Django Web App
Apache Web Server
MySQL Data Persistence
Amazon SQS
N parallel tasks
Job 2
Job 3
Job N
Amazon SQS: JobStatus Notification
Parallel MT Jobs


Azure Blob Storage:
Azure Drive
Azure Cloud
Job 1
Azure
Worker Role
Azure Worker/VM Role:
Azure
Worker Role
MTInversion Script
Azure
Worker Role
Azure: Web Application
Azure Web Role /Web Interface
IIS Web Server
SQL Azure
Azure Queue
N parallel tasks
Work in progress
Job 2
Job 3
Job N
Parallel MT Jobs
Azure Queue:
JobNotifications

MT Inversion Code Reconstruction

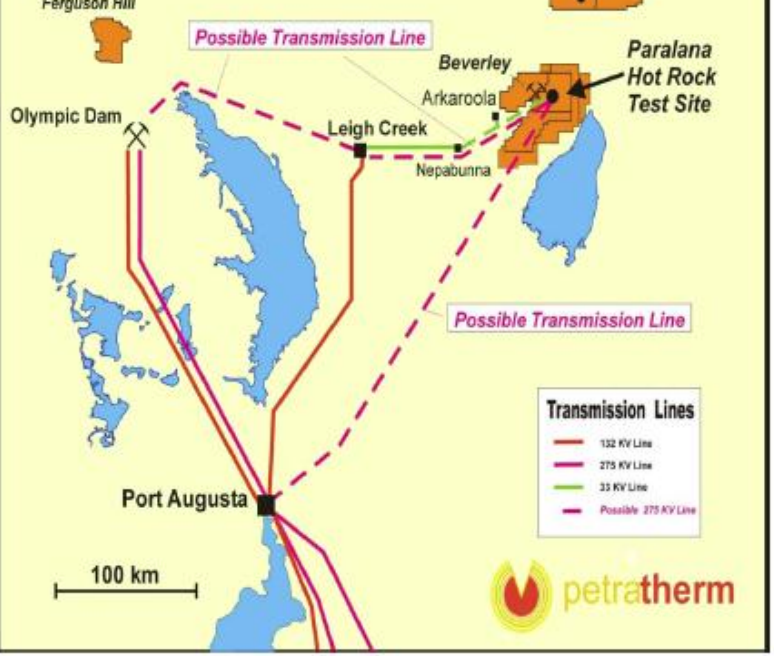
Sequential code: FORTRAN 77 22,000 lines

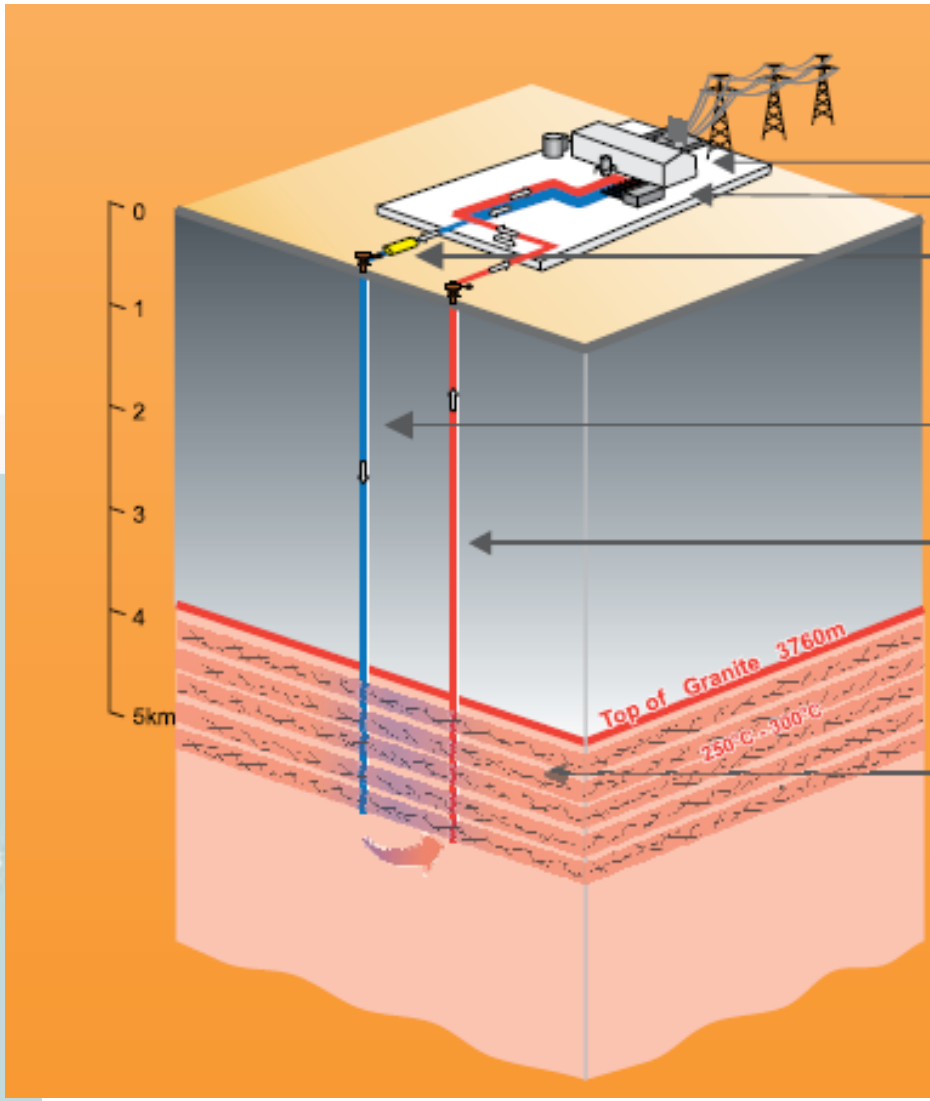
1. Find computational blocks amenable to easy parallelization
   - Sensitivity & Forward modelling computations independent for each frequency.
2. Separate control logic from numerical computations.
3. Optimization taken over by a parallel workflow control script (Python)

In other words: FORTRAN code split into multiple computational sub-processes driven by the control script

Email craig.mudge@adelaide.edu.au for TN 2011-3

Enhanced Geothermal System





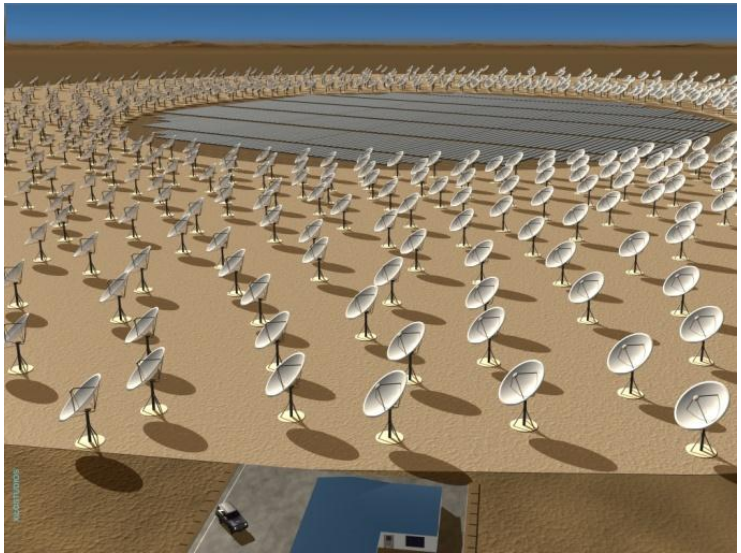
Hot rocks

Current Highlights

Perth Geothermal

Direct heat – urban

Potential: Extraction of deep heat; perhaps part of Square Kilometre Array



Groundwater

Nepabunna, Flinders Ranges

Geoscientists Without Borders
Society of Exploration Geophysicists
Founding Supporter Schlumberger

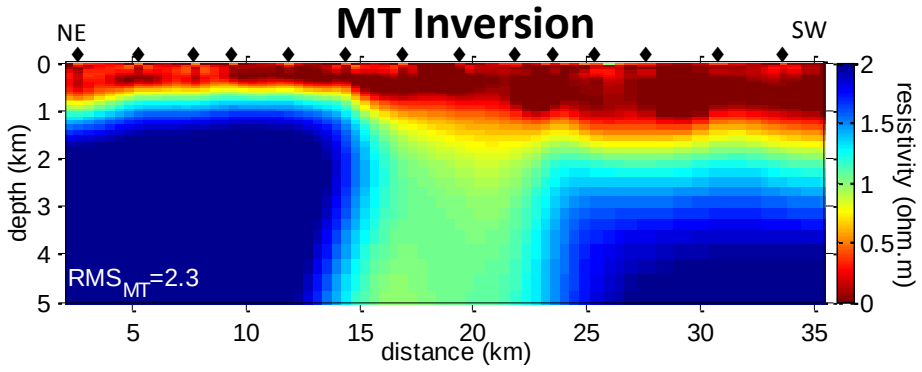
Future areas

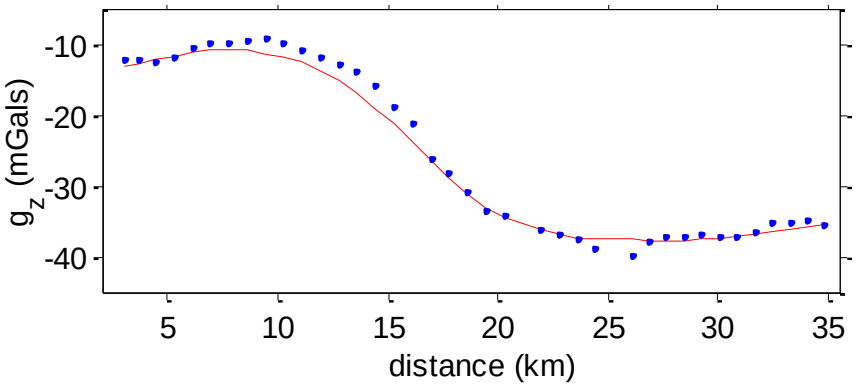
1. Near real-time feedback during data logging
2. Improvements to current MT FORTRAN program

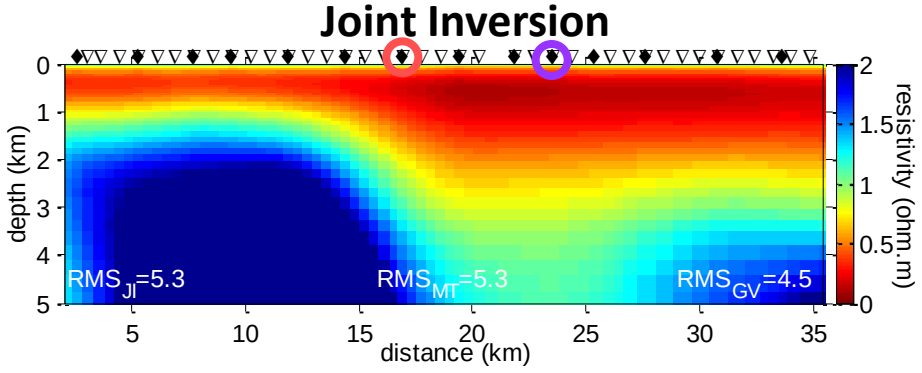
   e.g., Parallelise the matrix operations
3. Genetic algorithm – alternative to conjugate gradient
4. Joint inversion – seismic + MT + gravity
5. Our modelling accessible as a cloud application available to explorers world-wide

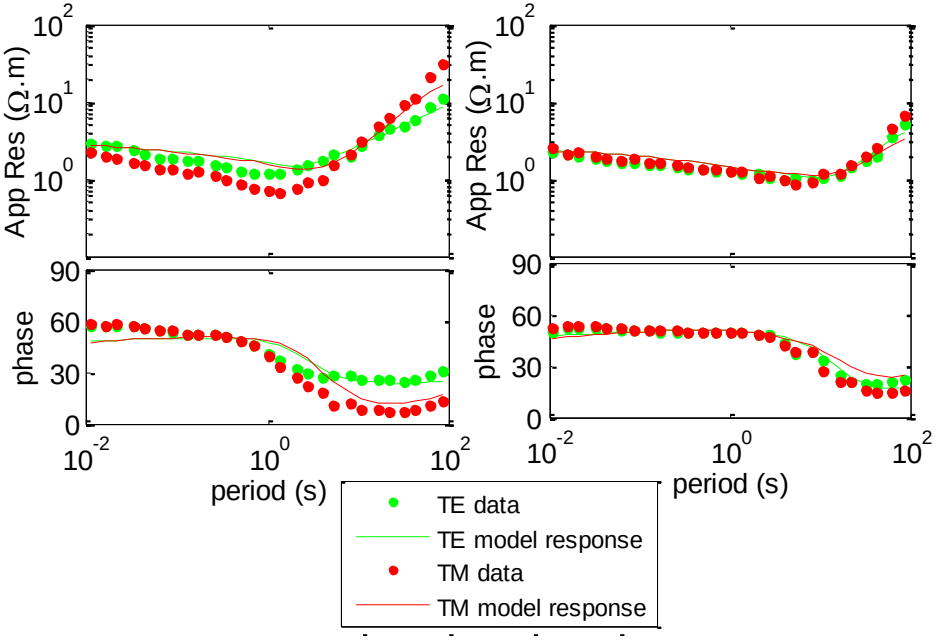
Rachel Maier, 2010)

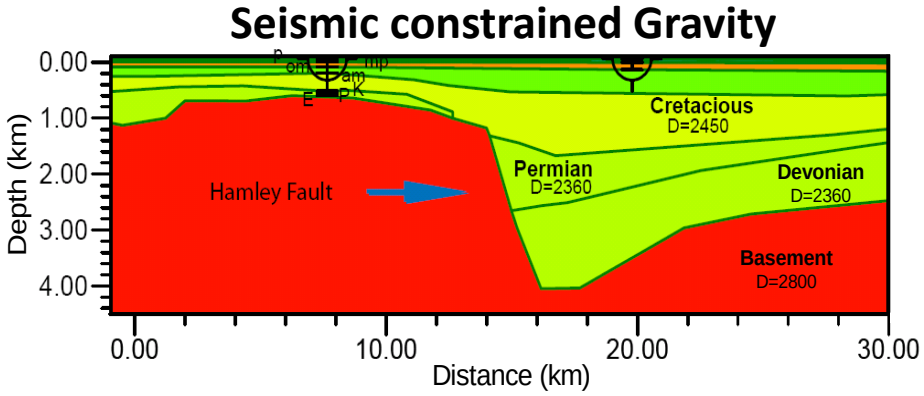
Renmark Trough











Future areas

1. Near real-time feedback during data logging
2. Improvements to current MT FORTRAN program

   e.g., Parallelise the matrix operations
3. Genetic algorithm – alternative to conjugate gradient
4. Joint inversion – seismic + MT + gravity
5. Our modelling accessible as a cloud application available to explorers world-wide

Acknowledgements

Ed Lazowska

Gordon Bell

Dennis Gannon

Catharine van Ingen

John Warren

Acknowledgements

Graham Heinson

Pinaki Chandrasekhar

Wei Wang

Stephan Thiel

National geophysics facility - magnetotellurics

Summary

1. A no-machines lab - June 2010
2. Magnetotelluric method
3. Processing the data
4. Results -- weeks down to days
5. Future work

   Production and development: almost 100% Azure

   eScience – geothermal, aquifer, as well as minerals

   Modelling on demand for explorers world-wide

Thanks
and
questions

craig.mudge@adelaide.edu.au

www.cloudinnovation.com.au

+61 417 679 266
+1 650 224 2111