

A path towards reproducible magnetotelluric (MT) time series processing on HPC

The Australasian Leadership Computing Symposium 2023

Nigel Rees, Lesley Wyborn, Rui Yang, Hannes Hollmann, Jo Croucher, Rebecca Farrington, Yue Sun, Yiling Liu, Andrew Robinson, Ben Evans

15 June 2023

We acknowledge and celebrate the First Australians on whose traditional lands we meet and pay our respect to the Elders past and present.



Artist: Lynnice Letty Church – Tribes: Ngunnawal, Wiradjuri & Kamilaroi (ACT and NSW)
Gadi - "to search for" in Ngunnawal language - January 2020 for NCI Gadi Supercomputer

The 2030 Geophysics Collection Project

- 2030 is an R&D project funded through a collaboration between AuScope, NCI, TERN and ARDC (<https://ardc.edu.au/project/2030-geophysics-collections/>)

- It seeks to:
 - a. Make national-scale high-resolution geophysics datasets suitable for programmatic access in HPC environments;
 - b. Lay the foundations for more rapid data processing by 2030 next-generation scalable, data intensive computation including Artificial Intelligence (AI) / Machine Learning (ML) and data assimilation.

- The project is about positioning Australian geophysical data collections to be capable of taking advantage of next generation technologies and computational infrastructures by 2030.

What do we know about 2030 computing?

1. High-end computational power will be at exascale
2. Today's emerging collaborative platforms will continue to evolve as a mix of HPC and cloud
3. Data volumes will be measured in Zettabytes (10^{21} bytes), which is about 10 times more than today
4. **It will be mandatory for data discovery, accessibility, interoperability and reusability to be fully machine-to-machine as envisaged by the FAIR principles in 2016**



What are the opportunities of 2030 computing?

1. So often today's research is undertaken on pre-canned, analysis-ready datasets (ARD) that are tuned towards the highest common denominator as determined by the data owner/publisher.
2. By 2030:
 - a. Increased computational power co-located with fast-access storage systems will mean that geophysicists will be able to work on less processed data levels and then transparently develop their own derivative products.
 - b. Researchers will be able to see the quality of their algorithms more rapidly: there will be multiple versions of open source software used as researchers fine tune individual algorithms to suit their specific requirements.
 - c. We will be capable of more precise solutions and in hazards space and other relevant areas, analytics will be done in faster-than-real-time.



Landsat-8 image courtesy of the U.S. Geological Survey

Wyborn, L., Rees, N., Klump, J., Evans, B., Rawling, T., and Druken, K.: The Known Knowns, the Known Unknowns and the Unknown Unknowns of Geophysics Data Processing in 2030, EGU General Assembly 2022, Vienna, Austria, 23–27 May 2022, EGU22-11012, <https://doi.org/10.5194/egusphere-egu22-11012>, 2022.

Expanding our geophysics data collections on HPC

 NCI Data Catalogue

[< Previous](#) [Next >](#)

Geoscience Australia Geophysics Reference Data Collection

This collection has been compiled by Geoscience Australia from an extensive archive of over 2,200 geophysical surveys dating back to 1947. The datasets have been acquired by Geoscience Australia and its State and Territory Government partners. The collection includes datasets acquired using potential field (magnetics and gravity) and radiometric (gamma-ray spectrometry) geophysical techniques. Ground elevation datasets (digital elevation models), which were acquired incidentally during many of the geophysical surveys, are also included. The datasets have been organised into four sub-collections based on the individual surveys for each of the three survey types: ground gravity surveys; airborne magnetic, radiometric and elevation surveys; and marine potential field surveys. The fourth sub-collection consists of the national scale compilations of data for each geophysical method. The data are processed (non-raw) and are stored in both their fundamental/original form as point-located measurements, as well as in interpolated regular grid (raster) form. The collection has potential applications in a wide range of disciplines including geological mapping, mineral and petroleum resource exploration, groundwater and environmental resource management, soil science, geodesy, and infrastructure planning.

Data Access

Register for local access <https://my.nci.org.au/mancini/project/v65/join>





Link to NCI THREDDS <http://dapds00.nci.org.au/thredds/catalogs/v65/catalog.html>



Geoscience Australia Source Record <http://dx.doi.org/10.4225/25/5a68071878b93>

Expanding our geophysics data collections on HPC

 NCI Data Catalogue

 NCI Data Catalogue

< P

< Previous Next page >

Geoscience Australia

National ASTER Map of Australia

This collection has been derived from geophysical surveys of its State and Territory (magnetics and gravity elevation datasets) dig on the individual surveys, a radiometric and elevation national scale com and are stored in both interpolated regular grid disciplines including ge environmental resource

This is the parent datafile of a dataset that comprises a set of 14+ geoscience products made up of mosaiced ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) scenes across Australia. The individual geoscience products are a combination of bands and band ratios to highlight different mineral groups and parameters including: False colour composite CSIRO Landsat TM Regolith Ratios Green vegetation content Ferric oxide content Ferric oxide composition Ferrous iron index Opaque Index AIOH group composition Kaolin group index FeOH group content MgOH group content MgOH group composition Ferrous iron content in MgOH/carbonate Surface mineral group distribution (relative abundance and composition)

--

This record represents data published by NCI. It uses catalogue metadata modified from Geoscience Australia. See our Record Lineage Information for more details.

Data Access

Data Access



Register for local access

Register for local access <https://my.nci.org.au/mancini/project/wx7/oln>

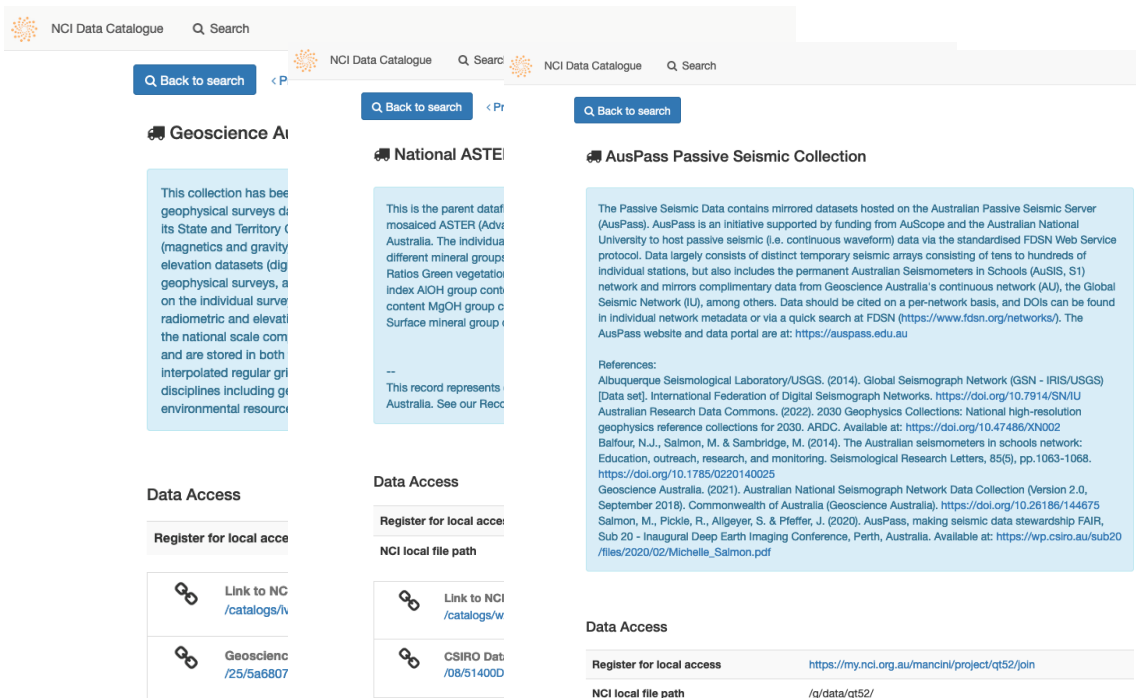
Link to NCI /catalogs/v

Geoscience /25/5a6807

NCI local file path /g/data/wx7/Australian_ASTER_Geosciences_Map_Project

	Link to NCI THREDDS https://dapds00.nci.org.au/thredds/catalogs/wx7/catalog.html	<input type="button" value="Open link"/>
	CSIRO Data Access Portal Record https://doi.org/10.4225/08/51400D6F7B335	<input type="button" value="Open link"/>

Expanding our geophysics data collections on HPC



The screenshot displays the NCI Data Catalogue interface with three data collection entries:

- Geoscience Australia:** This collection has been compiled from geophysical surveys of its State and Territory (magnetics and gravity elevation datasets (digital geophysical surveys, a on the individual survey radiometric and elevation national scale com and are stored in both interpolated regular grid disciplines including ge environmental resource
- National ASTER:** This is the parent dataset mosaiced ASTER (Advanced Australia. The individual different mineral groups: Ratios Green vegetation Index AIOH group content MgOH group c Surface mineral group i
- AusPass Passive Seismic Collection:** The Passive Seismic Data contains mirrored datasets hosted on the Australian Passive Seismic Server (AusPass). AusPass is an initiative supported by funding from AuScope and the Australian National University to host passive seismic (i.e. continuous waveform) data via the standardised FDSN Web Service protocol. Data largely consists of distinct temporary seismic arrays consisting of tens to hundreds of individual stations, but also includes the permanent Australian Seismometers in Schools (AuSIS, S1) network and mirrors complementary data from Geoscience Australia's continuous network (AU), the Global Seismic Network (IU), among others. Data should be cited on a per-network basis, and DOIs can be found in individual network metadata or via a quick search at FDSN (<https://www.fdsn.org/networks/>). The AusPass website and data portal are at: <https://auspass.edu.au>

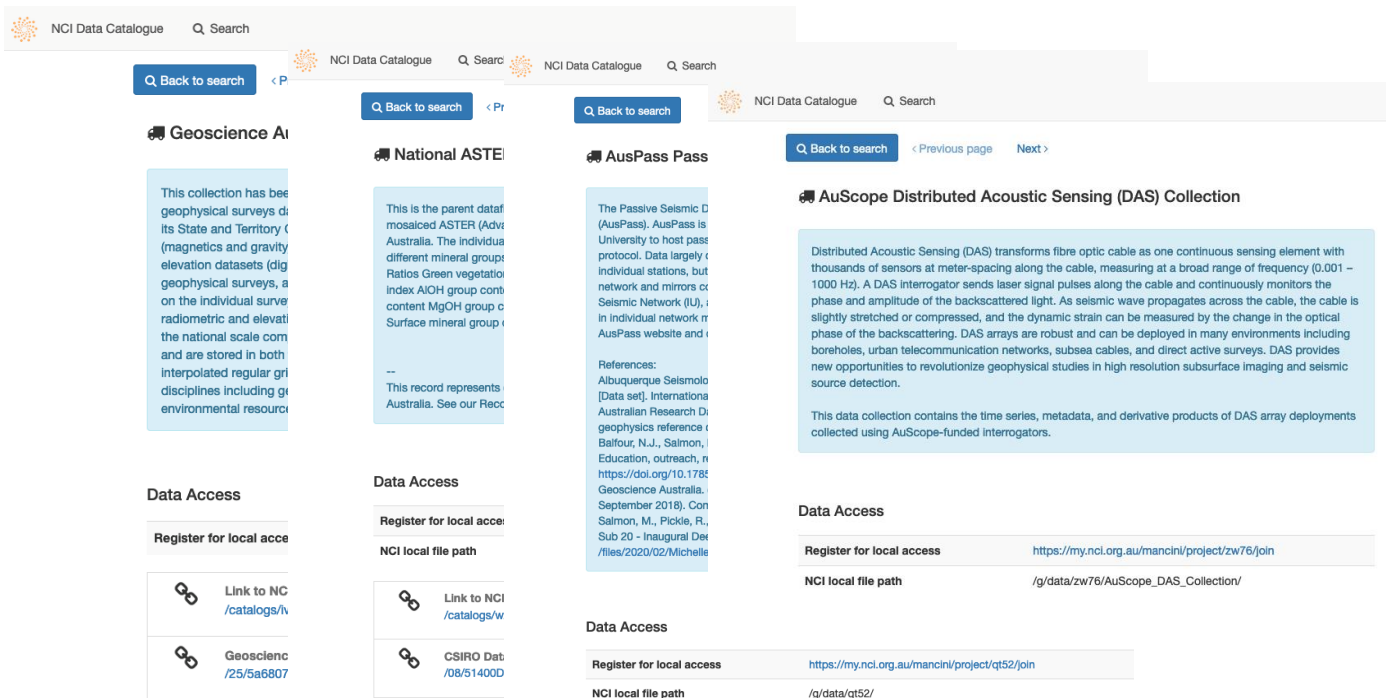
References for AusPass:

- Albuquerque Seismological Laboratory/USGS. (2014). Global Seismograph Network (GSN - IRIS/USGS) [Data set]. International Federation of Digital Seismograph Networks. <https://doi.org/10.7914/SN/IU>
- Australian Research Data Commons. (2022). 2030 Geophysics Collections: National high-resolution geophysics reference collections for 2030. ARDC. Available at: <https://doi.org/10.47486/XN002>
- Balfour, N.J., Salmon, M. & Sambridge, M. (2014). The Australian seismometers in schools network: Education, outreach, research, and monitoring. *Seismological Research Letters*, 85(5), pp.1063-1068. <https://doi.org/10.1785/0220140025>
- Geoscience Australia. (2021). Australian National Seismograph Network Data Collection (Version 2.0, September 2018). Commonwealth of Australia (Geoscience Australia). <https://doi.org/10.26186/144675>
- Salmon, M., Pickle, R., Algeyer, S. & Pfeffer, J. (2020). AusPass, making seismic data stewardship FAIR, Sub 20 - Inaugural Deep Earth Imaging Conference, Perth, Australia. Available at: https://wp.csiro.au/sub20/files/2020/02/Michelle_Salmon.pdf

Data Access for each collection:

- Geoscience Australia:**
 - Register for local access
 - Link to NCI /catalogs/v
 - Geoscienc /25/5a6807
- National ASTER:**
 - Register for local access
 - Link to NCI /catalogs/w
 - CSIRO Dat /08/51400D
- AusPass Passive Seismic Collection:**
 - Register for local access <https://my.nci.org.au/mancini/project/qt52/join>
 - NCI local file path </g/data/qt52/>

Expanding our geophysics data collections on HPC



The screenshot displays the NCI Data Catalogue interface with four data collection entries. Each entry includes a title, a description, a 'Data Access' section with options to register for local access or view the local file path, and a 'Data Access' section with options to register for local access or view the local file path.

Geoscience Australia

This collection has been derived from geophysical surveys of its State and Territory (magnetics and gravity elevation datasets (dig geophysical surveys, a on the individual survey radiometric and elevati the national scale com and are stored in both interpolated regular gri disciplines including ge environmental resource

National ASTER

This is the parent dataf mosaiced ASTER (Adv Australia. The individua different mineral group: Ratios Green vegetatio Index AIOH group cont content MgOH group c Surface mineral group c

--
This record represents Australia. See our Rec

AusPass Pass

The Passive Seismic D (AusPass). AusPass is University to host pass protocol. Data largely i individual stations, but network and mirrors c Seismic Network (IU), i In individual network r AusPass website and i

References:
Albuquerque Seismolo [Data set]. Internatio Australian Research Di geophysics reference c Balfour, N.J., Salmon, Education, outreach, n https://doi.org/10.178 Geoscience Australia, September 2018). Con Salmon, M., Pickle, R., Sub 20 - Inaugural De /files/2020/02/Michelle

AuScope Distributed Acoustic Sensing (DAS) Collection

Distributed Acoustic Sensing (DAS) transforms fibre optic cable as one continuous sensing element with thousands of sensors at meter-spacing along the cable, measuring at a broad range of frequency (0.001 – 1000 Hz). A DAS interrogator sends laser signal pulses along the cable and continuously monitors the phase and amplitude of the backscattered light. As seismic wave propagates across the cable, the cable is slightly stretched or compressed, and the dynamic strain can be measured by the change in the optical phase of the backscattering. DAS arrays are robust and can be deployed in many environments including boreholes, urban telecommunication networks, subsea cables, and direct active surveys. DAS provides new opportunities to revolutionize geophysical studies in high resolution subsurface imaging and seismic source detection.

This data collection contains the time series, metadata, and derivative products of DAS array deployments collected using AuScope-funded interrogators.

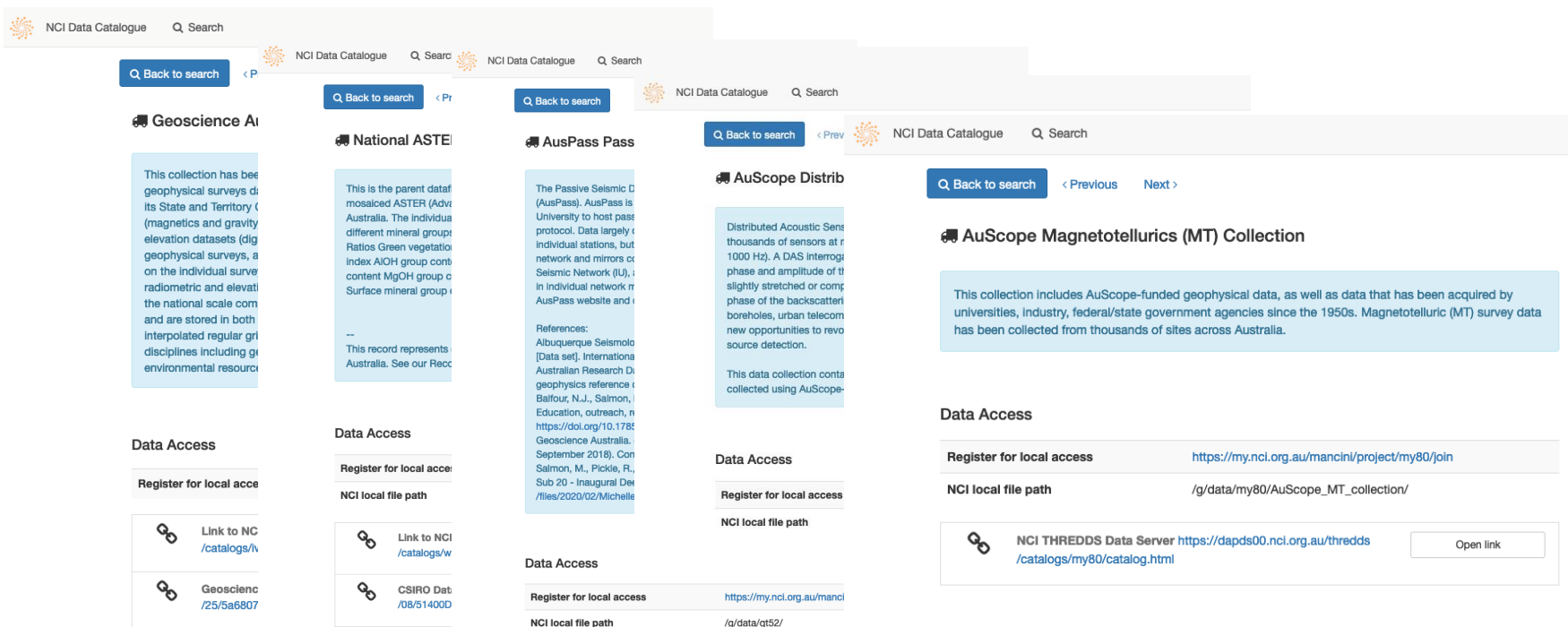
Data Access

Register for local access	https://my.nci.org.au/mancini/project/ztw76/join
NCI local file path	/g/data/ztw76/AuScope_DAS_Collection/

Data Access

Register for local access	https://my.nci.org.au/mancini/project/qt52/join
NCI local file path	/g/data/qt52/

Expanding our geophysics data collections on HPC



The image shows four overlapping screenshots of the NCI Data Catalogue interface. Each screenshot displays a different data collection page with a title, a description, and data access information.

- Geoscience Australia:** This collection has been derived from geophysical surveys of its State and Territory (magnetics and gravity elevation datasets (dig geophysical surveys, a on the individual survey radiometric and elevati the national scale com and are stored in both interpolated regular gri disciplines including ge environmental resource
- National ASTER:** This is the parent dataf mosaiced ASTER (Adv Australia. The individua different mineral group: Ratios Green vegetatio Index AIOH group cont content MgOH group c Surface mineral group c
- AusPass Pass:** The Passive Seismic D (AusPass). AusPass is University to host pass protocol. Data largely c individual stations, but network and mirrors of Seismic Network (IU), i In individual network r AusPass website and c
- AuScope Magnetotellurics (MT) Collection:** This collection includes AuScope-funded geophysical data, as well as data that has been acquired by universities, industry, federal/state government agencies since the 1950s. Magnetotelluric (MT) survey data has been collected from thousands of sites across Australia.

Each screenshot also shows a 'Data Access' section with options to 'Register for local access' and 'Link to NCI /catalogs/w'.

Building up our geophysical software libraries on HPC

The following geophysical software are available on the NCI platforms mostly under [NCI project up99](#):

BIRRP: The Bounded Influence Remote Reference Processing program

EMTF: Oregon State University robust single station, remote reference and multiple station MT timeseries data processing program

OCCAM1DCSEM / DIPOLE1D: Smooth one-dimensional models from CSEM and MT data

OCCAM2DMT: Occam's inversion for 2D MT modeling

MARE2DEM: parallel adaptive finite element code for 2D forward and inverse modeling for electromagnetic geophysics

ModEM: a flexible electromagnetic modelling and inversion program for 2D and 3D MT problems

<https://my.nci.org.au/mancini/project/ModEM-geophys>

FEMTIC: A 3-D Finite Element MagnetoTelluric Inversion code

esys-escript: a programming environment for implementing mathematical models in python using the finite element method (FEM).

Fire Drake: an automated system for solving partial differential equations using the finite element method (FEM).

jif3D: a framework for joint inversion of different types of geophysical data in 3D.

OpenQuake: an open-source application that allows users to compute seismic hazard and seismic risk of earthquakes on a global scale.

<https://opus.nci.org.au/x/KAIWC>

Developing community driven specialised software environments

NCI Geophysics Environment

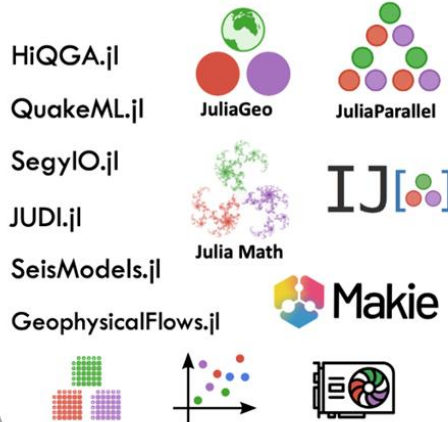


Python



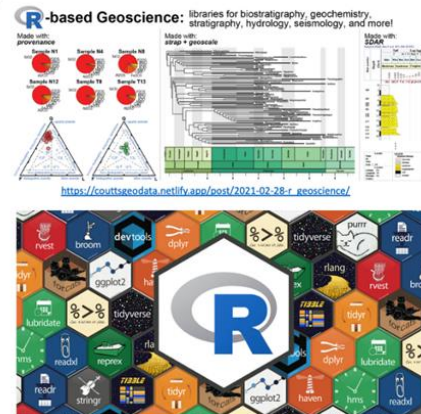
650+ Python packages

Julia



500+ Julia packages

R

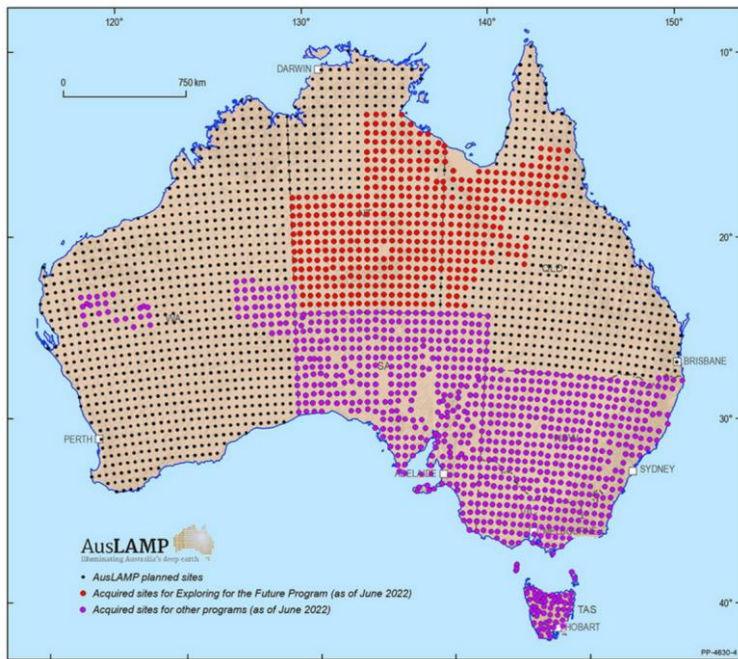


540+ R packages

```
$ module use /g/data/up99/modulefiles
$ module load NCI-geophys/<version>
```

<https://doi.org/10.6084/m9.figshare.21919584.v1>

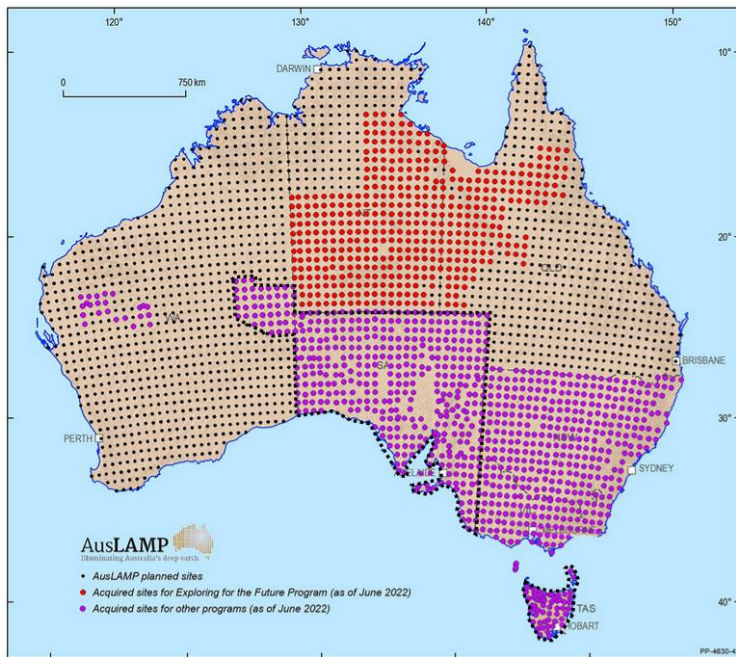
National Scale Analysis: The Australian Lithospheric Architecture Magnetotelluric Project (AusLAMP)



- A collaborative project between Geoscience Australia, the state and Northern Territory geological surveys, AuScope, universities and other research organisations.
- Aims to acquire long-period MT data at approximately 3000 sites across Australia.
- Dataset will be used to map electrical conductivity structure of the crust and upper mantle to improve the understanding of the geology and tectonic evolution of the Australian plate.
- Results may provide new information for identifying regions with mineral and energy resource potential.

<https://www.ga.gov.au/about/projects/resources/auslamp>

National Scale Analysis: The Australian Lithospheric Architecture Magnetotelluric Project (AusLAMP)

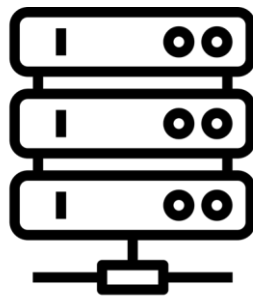
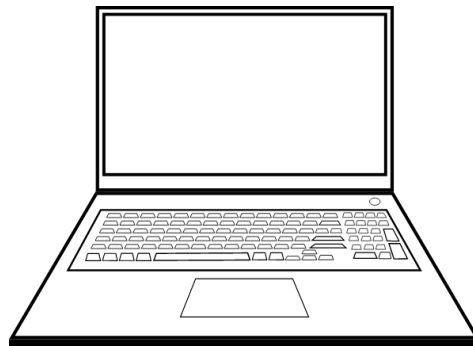


- A collaborative project between Geoscience Australia, the state and Northern Territory geological surveys, AuScope, universities and other research organisations.
- Aims to acquire long-period MT data at approximately 3000 sites across Australia.
- Dataset will be used to map electrical conductivity structure of the crust and upper mantle to improve the understanding of the geology and tectonic evolution of the Australian plate.
- Results may provide new information for identifying regions with mineral and energy resource potential.

<https://www.ga.gov.au/about/projects/resources/auslamp>

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series data?



Local Server



Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series data?

PR6-24 Earth Data Logger



LEMI 424



University of Adelaide Orange Boxes



<https://www.gfz-potsdam.de/en/section/geophysical-imaging/infrastructure/geophysical-instrument-pool-potsdam-gipp/pool-components/depas-pool/recorder/-earthdata-pr6-24>

<https://www.isr.lviv.ua/lemi424.htm>

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series data?

PR6-24 Earth Data Logger

```
SA252_170818000000.ambientTemperature SA252_170818050000.EY
SA252_170818000000.BX SA252_170818050000.TP
SA252_170818000000.BY SA252_170818060000.ambientTemperature
SA252_170818000000.BZ SA252_170818060000.BX
SA252_170818000000.EK SA252_170818060000.BY
SA252_170818000000.EY SA252_170818060000.BZ
SA252_170818000000.gps SA252_170818060000.EK
SA252_170818000000.gst SA252_170818060000.EY
SA252_170818000000.p11 SA252_170818060000.TP
SA252_170818000000.TP SA252_170818070000.ambientTemperature
SA252_170818010000.ambientTemperature SA252_170818070000.BX
SA252_170818010000.BX SA252_170818070000.BY
SA252_170818010000.BY SA252_170818070000.BZ
SA252_170818010000.BZ SA252_170818070000.EK
SA252_170818010000.EK SA252_170818070000.EY
SA252_170818010000.EY SA252_170818070000.TP
SA252_170818010000.TP SA252_170818080000.ambientTemperature
SA252_170818020000.ambientTemperature SA252_170818080000.BX
SA252_170818020000.BX SA252_170818080000.BY
SA252_170818020000.BY SA252_170818080000.BZ
SA252_170818020000.BZ SA252_170818080000.EK
SA252_170818020000.EK SA252_170818080000.EY
SA252_170818020000.EY SA252_170818080000.TP
SA252_170818020000.TP SA252_170818090000.ambientTemperature
SA252_170818030000.ambientTemperature SA252_170818090000.BX
```

LEMI 424

```
201610260243.INF 201610310000.TXT 201611060000.TXT
201610260244.TXT 201611010000.TXT 201611070000.TXT
201610270000.TXT 201611020000.TXT 201611080000.TXT
201610280000.TXT 201611030000.TXT 201611090000.TXT
201610290000.TXT 201611040000.TXT 201611100000.TXT
201610300000.TXT 201611050000.TXT 201611110000.TXT
```

University of Adelaide Orange Boxes

```
HFMS-000.BIN HFMS-023.BIN HFMS-046.BIN HFMS-069.BIN
HFMS-001.BIN HFMS-024.BIN HFMS-047.BIN HFMS-070.BIN
HFMS-002.BIN HFMS-025.BIN HFMS-048.BIN HFMS-071.BIN
HFMS-003.BIN HFMS-026.BIN HFMS-049.BIN HFMS-072.BIN
HFMS-004.BIN HFMS-027.BIN HFMS-050.BIN HFMS-073.BIN
HFMS-005.BIN HFMS-028.BIN HFMS-051.BIN HFMS-074.BIN
HFMS-006.BIN HFMS-029.BIN HFMS-052.BIN HFMS-075.BIN
HFMS-007.BIN HFMS-030.BIN HFMS-053.BIN HFMS-076.BIN
HFMS-008.BIN HFMS-031.BIN HFMS-054.BIN HFMS-077.BIN
HFMS-009.BIN HFMS-032.BIN HFMS-055.BIN HFMS-078.BIN
HFMS-010.BIN HFMS-033.BIN HFMS-056.BIN HFMS-079.BIN
HFMS-011.BIN HFMS-034.BIN HFMS-057.BIN HFMS-080.BIN
HFMS-012.BIN HFMS-035.BIN HFMS-058.BIN HFMS-081.BIN
HFMS-013.BIN HFMS-036.BIN HFMS-059.BIN HFMS-082.BIN
HFMS-014.BIN HFMS-037.BIN HFMS-060.BIN HFMS-083.BIN
HFMS-015.BIN HFMS-038.BIN HFMS-061.BIN HFMS-084.BIN
HFMS-016.BIN HFMS-039.BIN HFMS-062.BIN HFMS-085.BIN
HFMS-017.BIN HFMS-040.BIN HFMS-063.BIN HFMS-086.BIN
HFMS-018.BIN HFMS-041.BIN HFMS-064.BIN HFMS-087.BIN
HFMS-019.BIN HFMS-042.BIN HFMS-065.BIN HFMS-088.BIN
HFMS-020.BIN HFMS-043.BIN HFMS-066.BIN HFMS-089.BIN
HFMS-021.BIN HFMS-044.BIN HFMS-067.BIN HFMS-090.BIN
HFMS-022.BIN HFMS-045.BIN HFMS-068.BIN HFMS-091.BIN
```

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series data?

PR6-24 Earth Data Logger

3813265
3813266
3813275
3813281
3813275
3813277
3813277
3813272
3813277
3813281
3813278
3813281
3813287
3813284
3813282
3813282
3813284
3813290
3813290
3813288
3813290

LEMI 424

2016	10	27	00	00	00	27747.359	-1985.590	-47045.008	28.59	31.76	164.456	-45.318	-2.309	-2.227	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	01	27747.361	-1985.599	-47045.035	28.62	31.76	164.470	-45.330	-2.373	-2.357	14.23	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	02	27747.330	-1985.631	-47045.031	28.61	31.76	164.496	-45.316	-2.369	-2.305	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	03	27747.309	-1985.651	-47045.047	28.62	31.76	164.498	-45.307	-2.361	-2.297	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	04	27747.299	-1985.660	-47045.055	28.62	31.76	164.494	-45.273	-2.191	-2.182	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	05	27747.285	-1985.667	-47045.035	28.63	31.76	164.466	-45.232	-2.389	-2.215	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	06	27747.289	-1985.691	-47045.055	28.61	31.76	164.442	-45.215	-2.225	-2.160	14.21	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	07	27747.277	-1985.701	-47045.063	28.59	31.76	164.408	-45.168	-2.250	-2.275	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	08	27747.289	-1985.692	-47045.063	28.61	31.76	164.366	-45.112	-2.285	-2.229	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	09	27747.268	-1985.678	-47045.051	28.62	31.76	164.276	-45.098	-2.359	-2.213	14.25	485.8	2629.7502	S	12826.0778	E	9	2	2
2016	10	27	00	00	10	27747.293	-1985.671	-47045.066	28.59	31.76	164.181	-45.053	-2.285	-2.318	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	11	27747.285	-1985.654	-47045.059	28.59	31.76	164.079	-45.009	-2.443	-2.121	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	12	27747.303	-1985.644	-47045.059	28.60	31.76	164.041	-44.953	-2.338	-2.184	14.25	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	13	27747.307	-1985.628	-47045.055	28.62	31.76	163.971	-44.906	-2.342	-2.229	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	14	27747.320	-1985.614	-47045.055	28.64	31.76	163.931	-44.864	-2.385	-2.191	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	15	27747.336	-1985.617	-47045.039	28.63	31.76	163.912	-44.813	-2.408	-2.119	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	16	27747.340	-1985.622	-47045.023	28.60	31.76	163.919	-44.764	-2.373	-2.227	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	17	27747.363	-1985.599	-47045.023	28.61	31.76	163.886	-44.690	-2.348	-2.242	14.15	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	18	27747.400	-1985.600	-47045.027	28.62	31.76	163.901	-44.647	-2.318	-2.225	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	19	27747.412	-1985.590	-47045.008	28.59	31.76	163.924	-44.619	-2.369	-2.268	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	20	27747.426	-1985.571	-47045.000	28.61	31.76	163.910	-44.611	-2.391	-2.193	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	21	27747.449	-1985.591	-47044.980	28.63	31.76	163.923	-44.597	-2.219	-2.146	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	22	27747.469	-1985.583	-47044.973	28.61	31.76	163.931	-44.560	-2.307	-2.369	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	23	27747.473	-1985.569	-47044.980	28.64	31.76	163.908	-44.550	-2.336	-2.254	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	24	27747.506	-1985.585	-47044.977	28.64	31.76	163.923	-44.510	-2.307	-2.205	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	25	27747.516	-1985.583	-47044.965	28.62	31.76	163.926	-44.525	-2.322	-2.184	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	26	27747.510	-1985.579	-47044.957	28.59	31.76	163.919	-44.593	-2.396	-2.160	14.14	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	27	27747.523	-1985.588	-47044.973	28.63	31.76	163.888	-44.603	-2.188	-2.201	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	28	27747.525	-1985.580	-47044.957	28.62	31.76	163.860	-44.650	-2.266	-2.303	14.17	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	29	27747.527	-1985.580	-47044.969	28.63	31.76	163.838	-44.666	-2.344	-2.326	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series data?

PR6-24 Earth Data Logger

3813265
3813266
3813275
3813281
3813275
3813277
3813277
3813272
3813277
3813281
3813278
3813281
3813287
3813284
3813282
3813282
3813284
3813290
3813290
3813288
3813290

LEMI 424

2016	10	27	00	00	00	27747.359	-1985.590	-47045.008	28.59	31.76	164.456	-45.318	-2.309	-2.227	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	01	27747.361	-1985.599	-47045.035	28.62	31.76	164.470	-45.330	-2.373	-2.357	14.23	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	02	27747.330	-1985.631	-47045.031	28.61	31.76	164.496	-45.316	-2.369	-2.305	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	03	27747.309	-1985.651	-47045.047	28.62	31.76	164.498	-45.307	-2.361	-2.297	14.24	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	04	27747.299	-1985.660	-47045.055	28.62	31.76	164.494	-45.273	-2.191	-2.182	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	05	27747.285	-1985.667	-47045.035	28.63	31.76	164.466	-45.232	-2.309	-2.215	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	06	27747.289	-1985.691	-47045.055	28.61	31.76	164.442	-45.215	-2.225	-2.160	14.21	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	07	27747.277	-1985.701	-47045.063	28.59	31.76	164.408	-45.168	-2.250	-2.275	14.25	485.9	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	08	27747.289	-1985.692	-47045.063	28.61	31.76	164.366	-45.112	-2.285	-2.229	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	09	27747.268	-1985.678	-47045.051	28.62	31.76	164.276	-45.098	-2.359	-2.213	14.25	485.8	2629.7502	S	12826.0778	E	9	2	2
2016	10	27	00	00	10	27747.293	-1985.671	-47045.066	28.59	31.76	164.181	-45.053	-2.285	-2.318	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	11	27747.285	-1985.654	-47045.059	28.59	31.76	164.079	-45.009	-2.443	-2.121	14.25	485.8	2629.7502	S	12826.0778	E	10	2	2
2016	10	27	00	00	12	27747.303	-1985.644	-47045.059	28.60	31.76	164.041	-44.953	-2.338	-2.184	14.25	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	13	27747.307	-1985.628	-47045.055	28.62	31.76	163.971	-44.906	-2.342	-2.229	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	14	27747.320	-1985.614	-47045.055	28.64	31.76	163.931	-44.864	-2.385	-2.191	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	15	27747.336	-1985.617	-47045.039	28.63	31.76	163.912	-44.813	-2.408	-2.119	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	16	27747.340	-1985.622	-47045.023	28.60	31.76	163.919	-44.764	-2.373	-2.227	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	17	27747.363	-1985.599	-47045.023	28.61	31.76	163.886	-44.690	-2.348	-2.242	14.15	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	18	27747.400	-1985.600	-47045.027	28.62	31.76	163.901	-44.647	-2.318	-2.225	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	19	27747.412	-1985.590	-47045.008	28.59	31.76	163.924	-44.619	-2.369	-2.268	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	20	27747.426	-1985.571	-47045.000	28.61	31.76	163.910	-44.611	-2.391	-2.193	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	21	27747.449	-1985.591	-47044.980	28.63	31.76	163.923	-44.597	-2.219	-2.146	14.10	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	22	27747.469	-1985.583	-47044.973	28.61	31.76	163.931	-44.560	-2.307	-2.369	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	23	27747.473	-1985.569	-47044.980	28.64	31.76	163.908	-44.550	-2.336	-2.254	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	24	27747.506	-1985.585	-47044.977	28.64	31.76	163.923	-44.510	-2.307	-2.205	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	25	27747.516	-1985.583	-47044.965	28.62	31.76	163.926	-44.525	-2.322	-2.184	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	26	27747.510	-1985.579	-47044.957	28.59	31.76	163.919	-44.593	-2.396	-2.160	14.14	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	27	27747.523	-1985.588	-47044.973	28.63	31.76	163.888	-44.603	-2.188	-2.201	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	28	27747.525	-1985.580	-47044.957	28.62	31.76	163.860	-44.650	-2.266	-2.303	14.17	485.8	2629.7503	S	12826.0778	E	10	2	2
2016	10	27	00	00	29	27747.527	-1985.580	-47044.969	28.63	31.76	163.838	-44.666	-2.344	-2.326	14.09	485.8	2629.7503	S	12826.0778	E	10	2	2

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series metadata?

1	Not Used	Not Used	Not Used	Not Used	Compulsory	Optional	Optional	Optional	Compulsory	Compulsory	Compulsory	Compulsory	Compulsory	Compulsory	Compulsory	Optional		
2	Site Number	Coords Available	Deployed Date Entered	Pickedup Date Entered	Site Number	UTC Start Time Minute	UTC Start Time Second	Deployment Julian Day	Recording Method	MT Recorder Type/ Model	Magnetometer Electrode Type/ Model			Data Confidentiality	North Reference	Drift calculation end time - Time		
406	SA333	Coords	dep date	Retrival date	SA333	44	12	139	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
407	SA333-2A	Coords	dep date	Retrival date	SA333-2A	23	0	170	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
408	SA333-2B	Coords	dep date	Retrival date	SA333-2B	56	0	170	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
409	SA334	Coords	dep date	Retrival date	SA334	55	0	139	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
410	SA335	Coords	dep date	Retrival date	SA335	53	51	140	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
411	SA336	Coords	dep date	Retrival date	SA336	17	24	140	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
412	SA337	Coords	dep date	Retrival date	SA337	6	48	142	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
413	SA338	Coords	dep date	Retrival date	SA338	3	17	142	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
414	SA339	Coords	dep date	Retrival date	SA339	26	51	143	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
415	SA340	Coords	dep date	Retrival date	SA340	35	0	140	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
416	SA340-2	Coords	dep date	Retrival date	SA340-2	1	0	143	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
417	SA341	Coords	dep date	Retrival date	SA341	10	0	324	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
418	SA342	Coords	dep date	Retrival date	SA342	54	0	324	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
419	SA344	Coords	dep date	Retrival date	SA344	35	0	135	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
420	SA344-2	Coords	dep date	Retrival date	SA344-2	6	16	171	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
421	SA345	Coords	dep date	Retrival date	SA345	59	15	132	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
422	SA346	Coords	dep date	Retrival date	SA346	39	0	132	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
423	SA347	Coords	dep date	Retrival date	SA347	7	0	234	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
424	SA348	Coords	dep date	Retrival date	SA348	21	0	232	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
425	SA349	Coords	dep date	Retrival date	SA349			231	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North		
426	SA350	Coords	dep date	Retrival date	SA350		41	13	231	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North	
427	SA351	Coords	dep date	Retrival date	SA351		52	20	231	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North	
428	SA354	Coords	dep date	Retrival date	SA354		18	52	79	LP	Earth Data Recc	Bartington, MoiSDEC (France), N 12Volt 72 Amp/Hr Battery, Power Supply Charging - Solar Panel, 60Watt				Open	Magnetic North	

Can we process AusLAMP time series data at scale on HPC?

Where is the AusLAMP MT time series metadata?

Instrument Deployment

Site: #R01
 Station: _____
 Name: _____
 Site Location: Longitude: 138° 26.350' E Latitude: 30° 33.020' S Elevation: 251m
 Easting: 254348 m Northing: 6617439 Zone: 54 J
 Declination: _____

Local Adelaide time is 09:30h ahead of UTC time
 (08:30h during summer time)

Instrument: _____
 Instrument Box: 15 Interface box No: 15 Mag Type: Fluxgate
 EDLogger No: 5933 Power box No: 1 Mag No: 1115
 Hard drive No: 5951 Battery No: 1 Coil No: (East) 1
 Coil No: (North) 1

X dipole length: 50m X dipole orientation from North: 0° (N)
 Y dipole length: 50m Y dipole orientation from North: 90° (E)

Sampling rate: 10 Hz File length: 60min Gain: low

START UP		SHUT DOWN	
Date	Time	Date	Time
2013-09-22	09:35 Local	17/10/13	08:00 Local
2013-09-22	00:05 UTC	13/11/13	14:00 UTC
Day No: 265		Day No: 286	

Battery: (start) 13.72 v Battery: (finished) 12.85 v

GPS off time: _____ Date: _____ (Orange Mag box only)
 Instr. off time: _____ Date: _____
 Drift (s) _____ (GPS time - Instrument time)

Notes - Deployment:
 Description of Area: Distance from roads, powerlines, etc., Soil, Vegetation, Outcrop:
 50m to horse paddock

Converting MT time series (meta)data to international standards



Computers & Geosciences

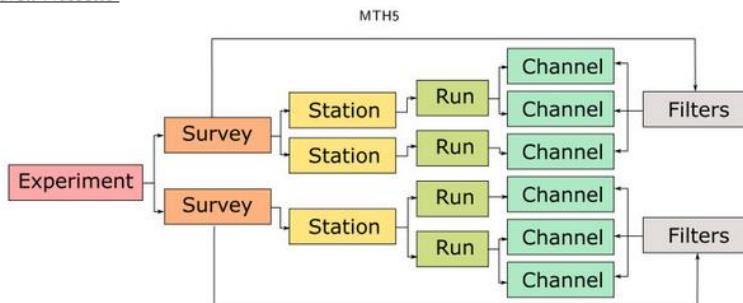
Volume 162, May 2022, 105102



Research paper

MTH5: An archive and exchangeable data format for magnetotelluric time series data

Jared Peacock^a  , Karl Kappler^b, Lindsey Heagy^c, Timothy Ronan^d, Anna Kelbert^e, Andrew Frassetto^d



MT_metadata

- Current version of MT_metadata has 355 attributes
 - 152 required parameters
 - 203 non mandatory parameters (but useful to have)
- For legacy AusLAMP metadata, we could not meet this standard as many of the required parameters were not recorded at the time of acquisition
- We did our best with what we had and worked on a Legacy AusLAMP MT_metadata profile that was a 72 element subset of the full MT_metadata standard
- Held regular meetings with the USGS based lead developer of MT_metadata and added in some revisions to the original standard (e.g., survey.funding_source.organization, survey.state, more options for station.release_license)

Scaling up MT time series processing

The **M**agneto**T**ellurics **t**ime **s**eries **d**ata **p**ublication (**MTtsdp**) codes: <https://github.com/nci/MTtsdp>

Processing Levels	Name	Description
Packed Raw Data	Raw Time Series	Telemetry data streamed from site loggers
Level 0	Edited Time Series	Time ordered instrument recorded data (e.g., raw voltages, counts) at full resolution
Level 1	Transformed Time Series	Level 0 data that have been transformed (e.g., calibrated, resampled, rotated, had noisy data removed, filters applied).
Level 2	Derived frequency domain processed data	Geophysical parameters (e.g., impedance tensors) derived from frequency domain time series processing of Level 1 data
Level 3	Derived modelling inputs and outputs	Level 2 parameters converted into input files for modelling and inversion algorithms with outputs mapped onto space-time grids.

Rees, N., Evans, B., Heinson, G., Conway, D., Yang, R., Thiel, S., Robertson, K., Druken, K., Goleby, B., Wang, J., Wyborn, L. & Seillé, H., 2019. The Geosciences DeVL Experiment: new information generated from old magnetotelluric data of The University of Adelaide on the NCI High Performance Computing Platform, ASEG Extended Abstracts, 2019:1, 1-6, DOI: [10.1080/22020586.2019.12073015](https://doi.org/10.1080/22020586.2019.12073015)

Scaling up MT time series processing

The **M**agneto**T**ellurics **t**ime **s**eries **d**ata **p**ublication (**MTtsdp**) codes: <https://github.com/nci/MTtsdp>

Processing Levels	Name	Description
Packed Raw Data	Raw Time Series	Telemetry data streamed from site loggers
Level 0	Edited Time Series	Time ordered instrument recorded data (e.g., raw voltages, counts) at full resolution
Level 1	Transformed Time Series	Level 0 data that have been transformed (e.g., calibrated, resampled, rotated, had noisy data removed, filters applied).
Level 2	Derived frequency domain processed data	Geophysical parameters (e.g., impedance tensors) derived from frequency domain time series processing of Level 1 data
Level 3	Derived modelling inputs and outputs	Level 2 parameters converted into input files for modelling and inversion algorithms with outputs mapped onto space-time grids.

Rees, N., Evans, B., Heinson, G., Conway, D., Yang, R., Thiel, S., Robertson, K., Druken, K., Goleby, B., Wang, J., Wyborn, L. & Seillé, H., 2019. The Geosciences DeVL Experiment: new information generated from old magnetotelluric data of The University of Adelaide on the NCI High Performance Computing Platform, ASEG Extended Abstracts, 2019:1, 1-6, DOI: [10.1080/22020586.2019.12073015](https://doi.org/10.1080/22020586.2019.12073015)

HPC has reduced processing times significantly

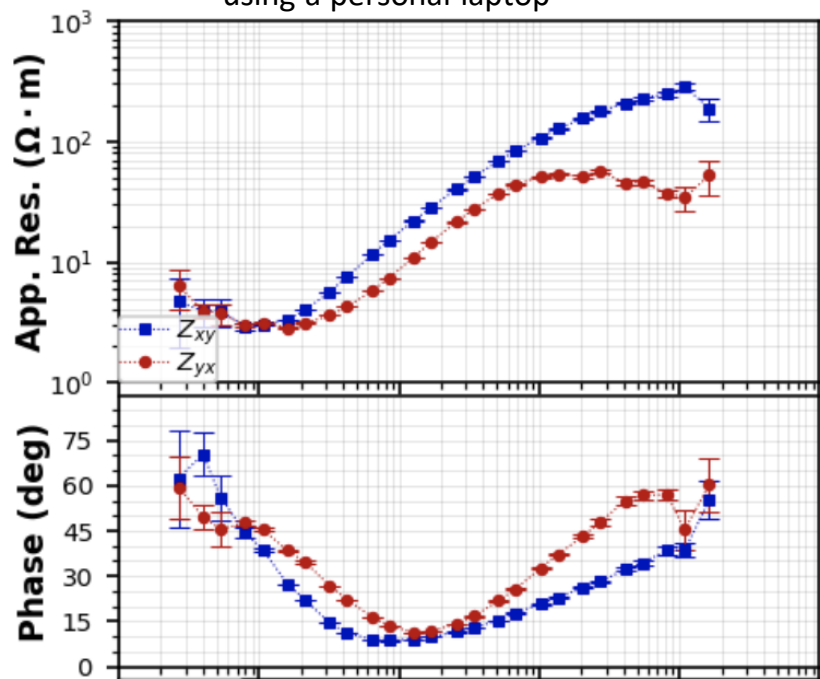
Parallel I/O via NCI-geophys environment

- MTH5: HDF5 data container for magnetotelluric time series data
- MT_metadata: standardised time series and transfer function metadata for magnetotelluric data
- Tests conducted on the AusLAMP Musgraves Province time series data:
<https://dx.doi.org/10.25914/58gr-1550>

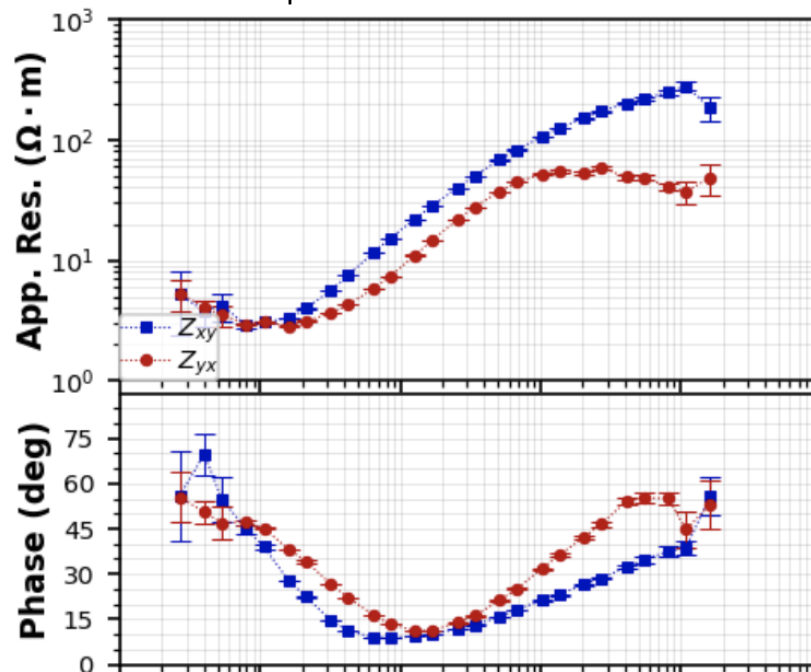
Dataset of 93 stations	Serial I/O	MPI based Parallel I/O (96 cores)
Level 0: one mth5/mt_metadata file per station	~ 5 hours 47 minutes	~ 4 minutes
Level 1: one mth5/mt_metadata file per station	~ 49 minutes	~1 minute 13 seconds
Level 2: one EDI file per station	~ 2 hours 30 minutes	~ 2 minutes

Computational reproducibility

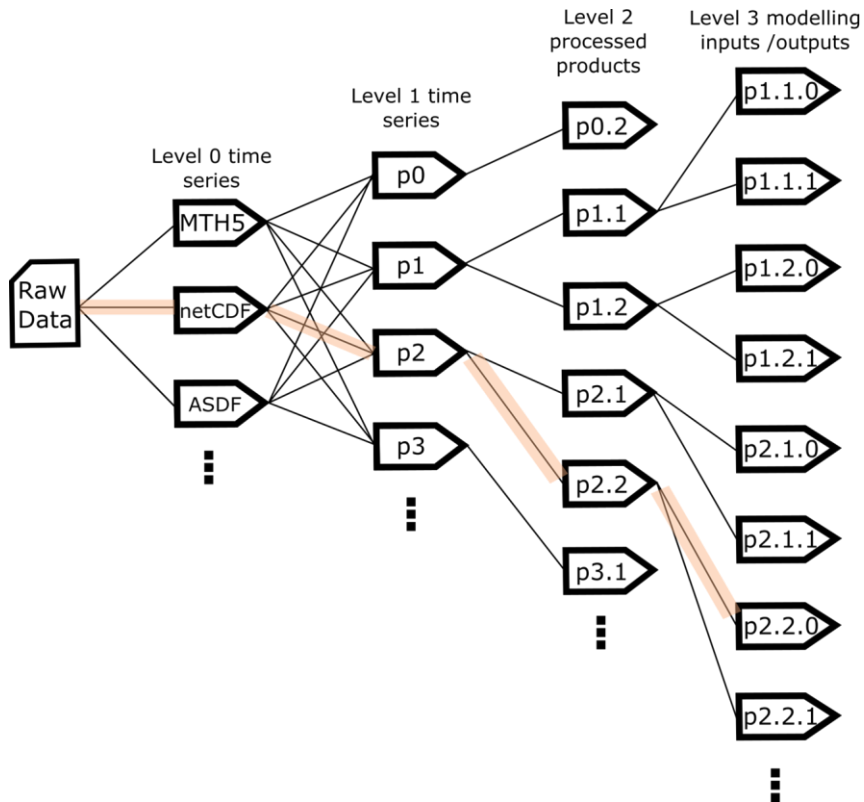
Original station WA72 processed using a personal laptop



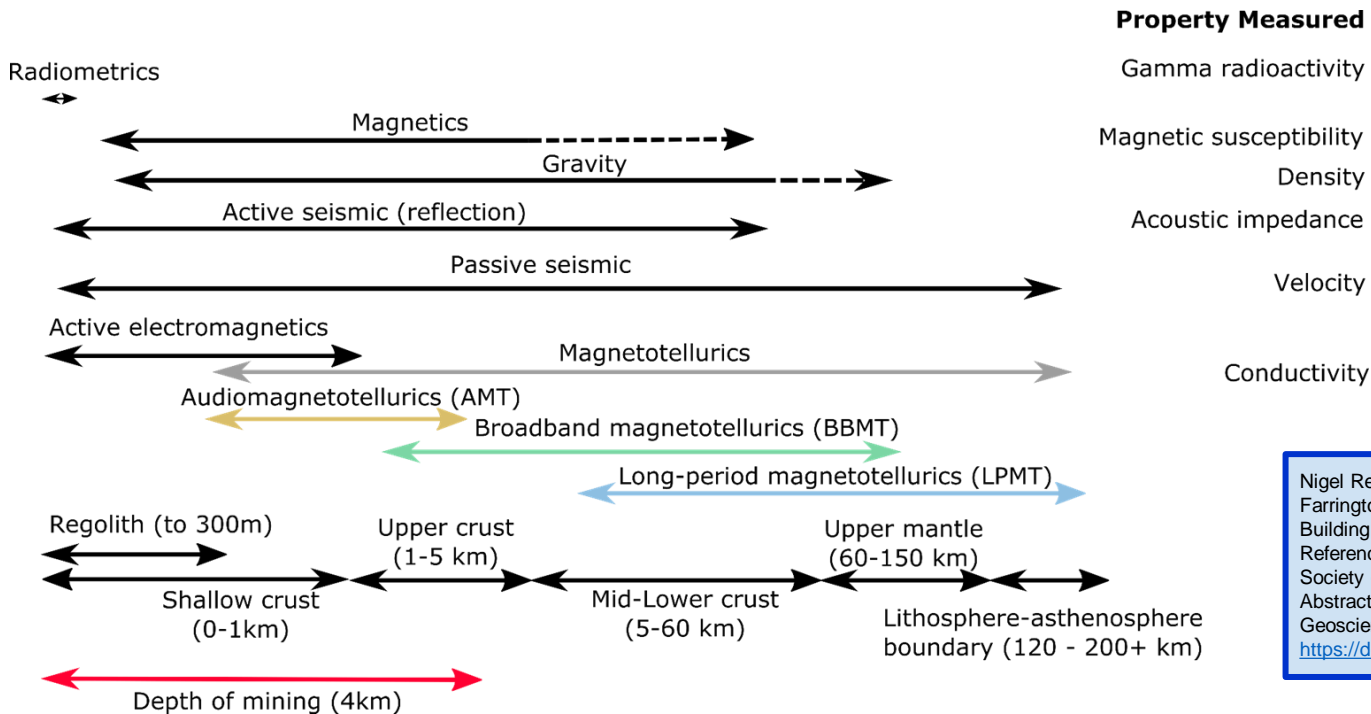
Processing station WA72 using the MTtsdp automations on HPC



Transparent provenance between processing levels



2030 ambition: Multiphysics Analysis but starting with PS and MT



Nigel Rees, Lesley Wyborn, Ben Evans, Rebecca Farrington, Tim Rawling, Rui Yang, & Yue Sun. (2023). Building a National High-Resolution Geophysics Reference Collection for 2030 Computation. Australian Society of Exploration Geophysicists Extended Abstracts, Volume 2023, 4th Australasian Exploration Geoscience Conference, Brisbane, 2023. <https://doi.org/10.5281/zenodo.7980192>

Types of geophysical data collected in Australia, the physical property measured and the depth of the crust that is sampled: also shown is the depth of current mining. Figure modified from original of Richard Chopping (GSWA).

Can we create a collaborative National High Resolution Geophysics Data Platform?



Brian Kennett, Richard Chopping and Richard Blewett, 2018. The Australian Continent, a Geophysical synthesis. Available on <https://press.anu.edu.au/publications/australian-continent#tabanchor>

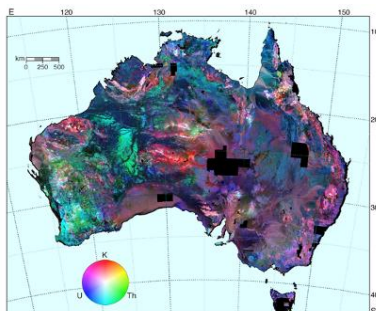


Figure 3.1: Gamma-ray spectrometric map for the continent of Australia created by

Radiometrics

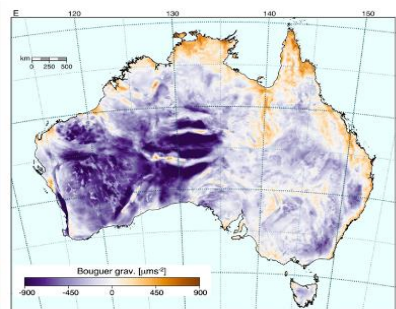


Figure 5.7: Bouguer gravity anomaly.

Bouger Gravity

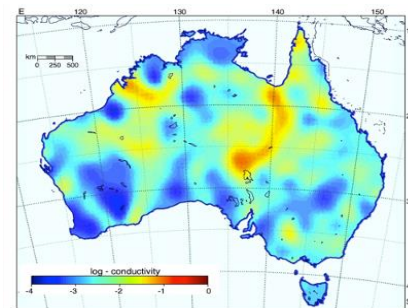


Figure 10.3: Logarithm of electrical conductivity [S/m] at a depth of 52 km.

Electrical Conductivity at 52 km

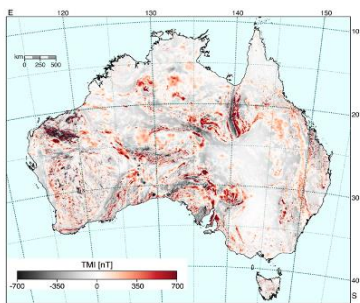


Figure 4.6: Total magnetic intensity across the Australian continent.

Total Magnetic Intensity

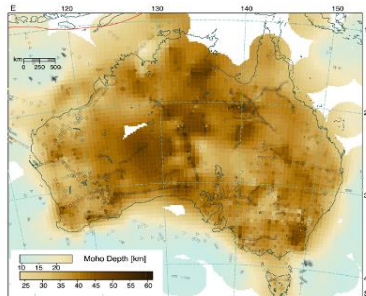


Figure 7.16: Moho surface across Australia utilising the full range of seismic information

Moho Depth

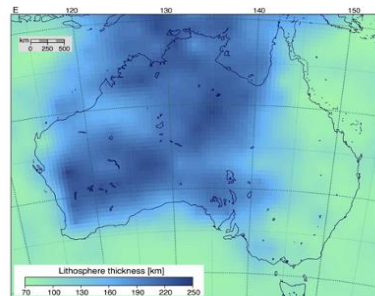


Figure 8.9: Estimate of thickness of the lithosphere across the Australian region.

Thickness of the lithosphere

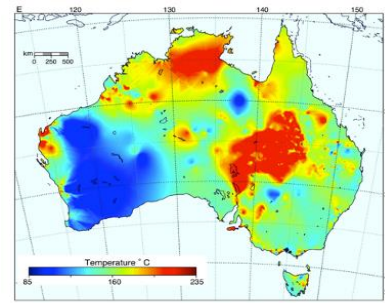


Figure 11.2: O2Temp estimate of the temperature distribution at 5 km depth across the Australian continent.

Temperature at 5km

Preparing for 2030: we no longer have the gift of time

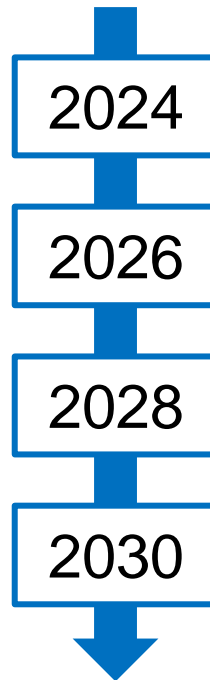


Photo by [Nathan Dumlaio](#) on [Unsplash](#)

We need to make up our minds today if we want to scale up.

If so we need to:

- Find all the rawer forms of critical geophysical datasets (and their collectors!) and work on ensuring all metadata, data and vocabularies are FAIR and machine actionable.
- Start working now on more automated systems that capture provenance through each successive processing level.
- Make less processed forms of data more accessible and able to be aggregated into seamless national high-resolution datasets.
- Ensure that whatever we do, it is always scalable to the future and can maximise benefits from new compute, data and software technologies as they come on line.





NCI
AUSTRALIA

NCI Contacts

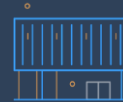


General enquiries: +61 2 6125 9800

Support: help@nci.org.au



Email: nigel.rees@anu.edu.au



Address

NCI, ANU Building 143
143 Ward Road
The Australian National University
Canberra ACT 2601

License

