Discrete Global Grid Systems to integrate statistical and geospatial information

Adam Lewis, A/Chief Scientist, Geoscience Australia (Adam.Lewis@ga.gov.au)

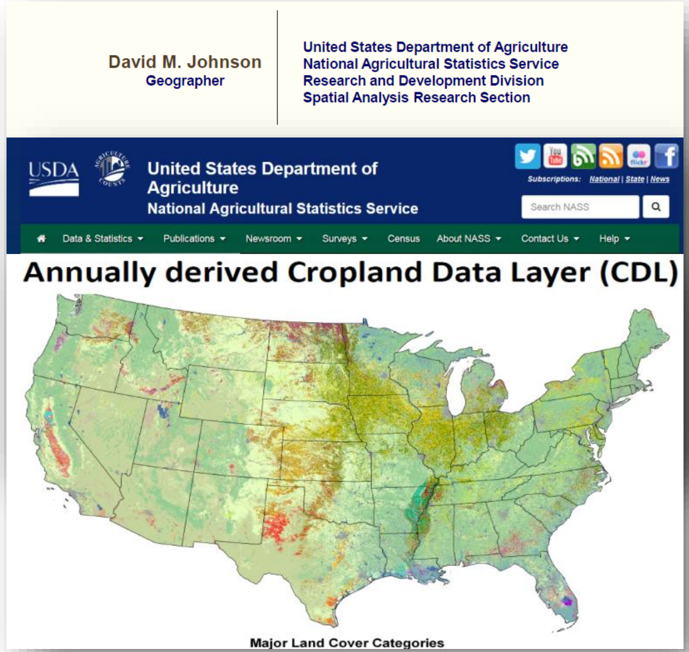
with contributions from Matthew Purss, Stuart Minchin, Robert Gibb, Faramarz Samavati, Perry Peterson, Clinton Dow, Jin Ben, Jonathan Ross, Trevor Dhu, Martin Brady

UNECE – Workshop on Integrating Geospatial and Statistical Standards
Stockholm, Sweden, 6-8 November 2017

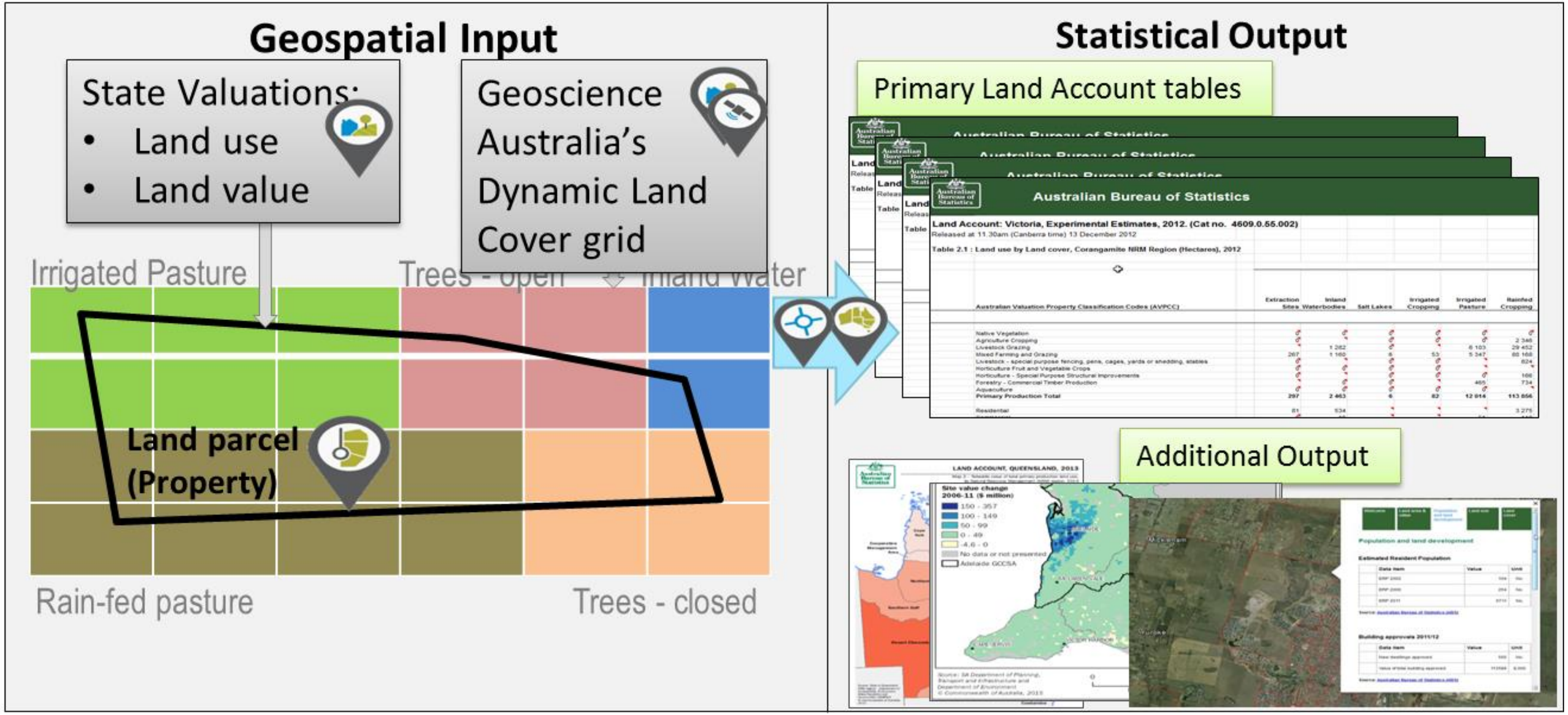
Geographical data in official statistics

Geographical data has an established role in official statistics, playing a greater or lesser part in some countries and applications than others

- US Department of Agriculture and US National Agricultural Statistical Services may be the exemplar
- Applies remote sensing and field data in a rigorous process
- Appropriate care is needed; e.g., ‘pixel counting’ is not sufficient - estimation biases need to be addressed (which is straightforward)



Geographical data in official statistics



Martin Brady, Australian Bureau of Statistics

The role of geospatial data is increasing

Frequent observation is the most important factor in useful data "*heavy volumes of time series imagery important*" (D M Johnson)

Weekly, high-quality, operational observation is now a reality through multiple government and commercial systems

E.g., over Australia:

- Landsat data 1987-2017 ~ 400TB
- Sentinel-2 data 2015-2017 ~ 200TB

Globally, Sentinel 1 & 2 ~ 10 PB per annum

Crop area mapping lessons learned

- Heavy volumes of time-series imagery important
  - Agriculture is a dynamic land cover
- Fine spatial resolution is somewhat important
  - Particularly if field sizes are relatively large
- Multi-spectral resolution least important
  - The time component reigns supreme

The role of geospatial data is increasing

Global reporting is increasingly looking to geographical data

- UNFCCC – Global Forest Observing Initiative
  From Sept 2017 every country in the world has access to free satellite data sufficient to support forest cover monitoring and reporting
- Sustainable Development Goals:
  Geographic measurements from satellites will support reporting on themes such as water and land degradation
- The Global Partnership for Sustainable Development Data is one initiative drawing capabilities together

Digital Earth Australia - DEA

Australian Government
Geoscience Australia

Digital Earth
AUSTRALIA

BIG DATA FOR A BIG COUNTRY

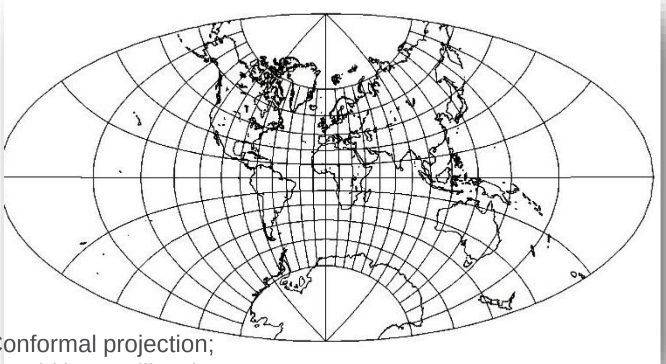
ga.gov.au/dea

Problems with traditional GIS approaches

"*It's a curious thing that integrating spatial datasets for analysis is still a problematic task*"
Position Magazine, 86, Dec-Jan 2017 (Simon Chester)

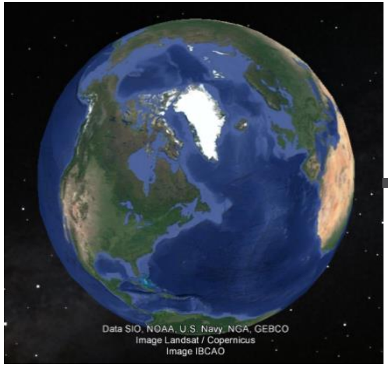
Traditional GIS/RS approaches have built from 'automated maps' and 'image analysis'. These bring deep legacy issues which have to be 'worked through', e.g.:

- Consistent approaches to gridding of data
- Conversion of vectors to grids
- Grids are single-resolution representations
- Map projections
    - Create anomalies
    - Are 'tuned' to certain areas
    - *Should be redundant in a digital world*

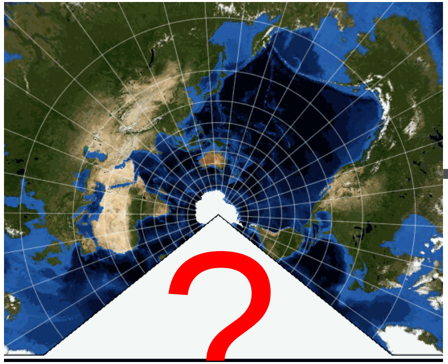


Conformal projection;
(World in an Ellipse);
Miscellaneous; Conformal;
Oscar S. Adams; 1925
With thanks to Paul B Anderson for materials

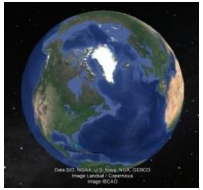
The spatial analysis workflow



Earth

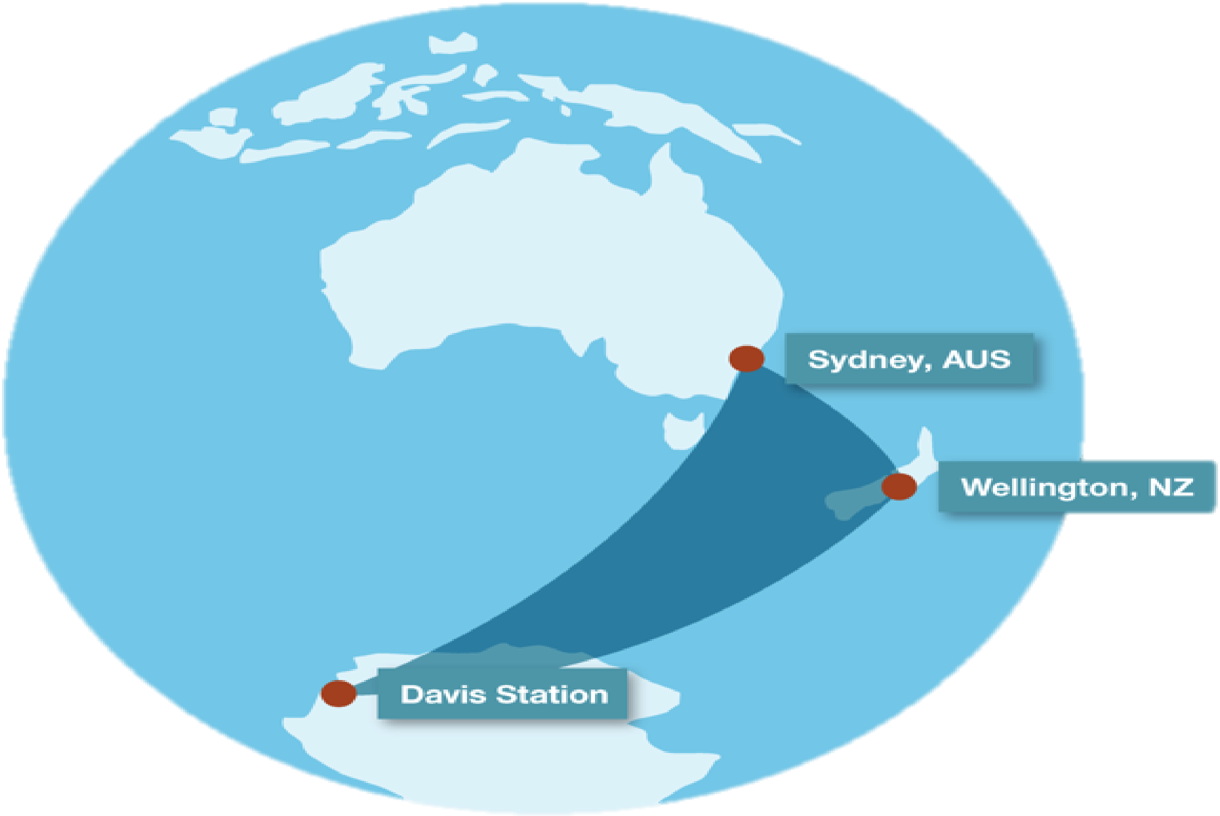


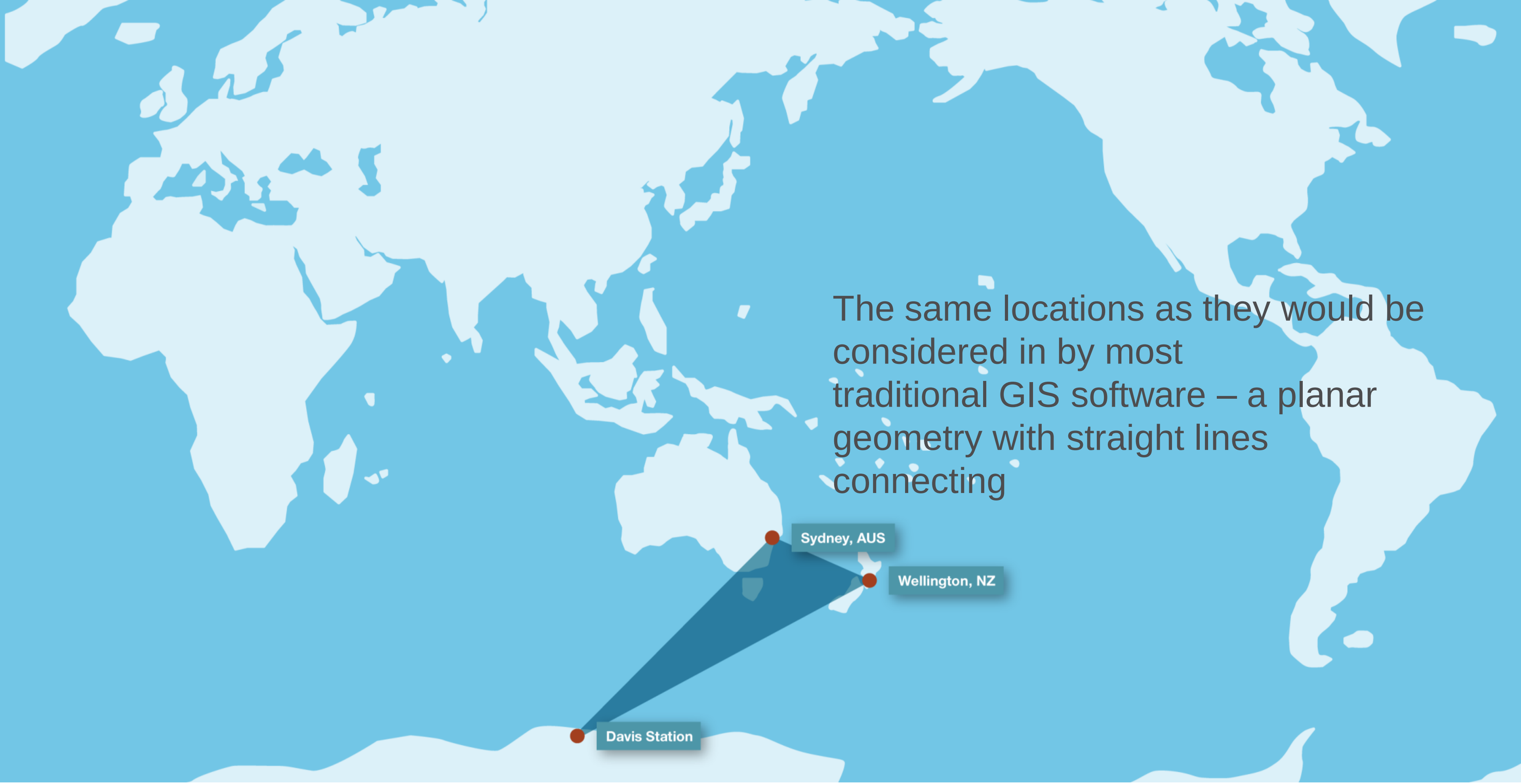
- Project to a planar surface
- Measure, compare, etc. in a (convenient) planar coordinate system
- Report results



Visualise, Report

A logical path between three important places (shortest distance)





The same locations as they would be considered in by most traditional GIS software – a planar geometry with straight lines connecting

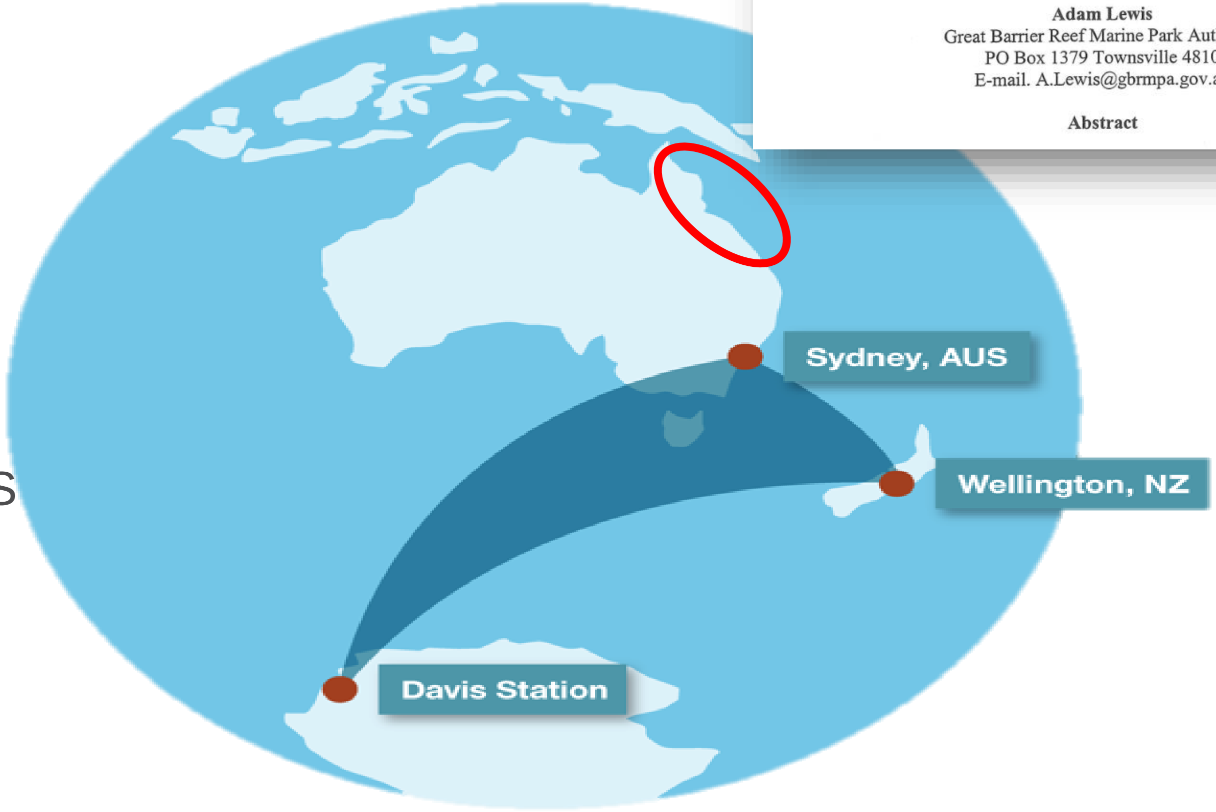
Accuracy 2002 Symposium, 10-12 July, Melbourne, Australia

ACCURATE MAPPING OF MARINE PARK BOUNDARIES: THE GEODESIC PROBLEM

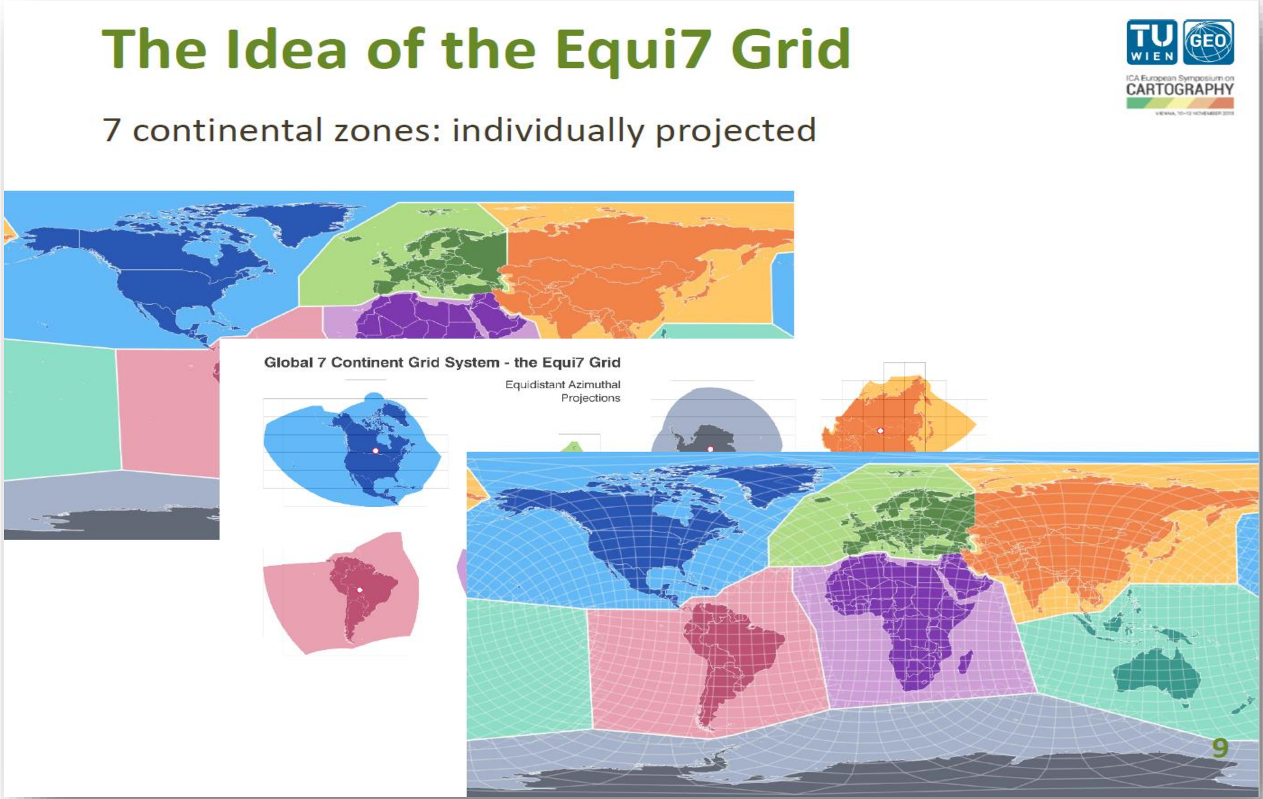
Adam Lewis
Great Barrier Reef Marine Park Authority
PO Box 1379 Townsville 4810
E-mail. A.Lewis@gbrmpa.gov.au

Abstract

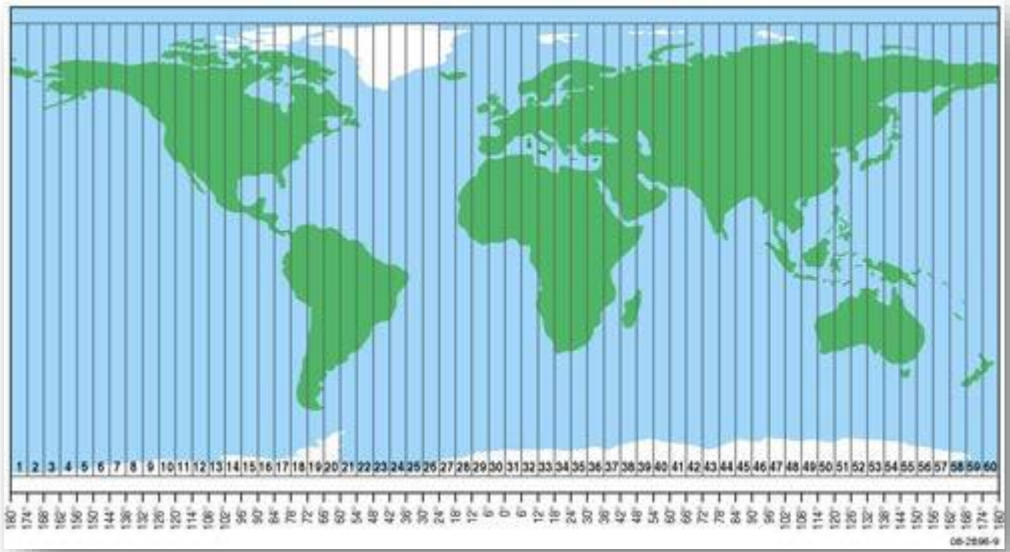
The ‘straight lines’ in the planar representation are in reality the *wrong paths;* Going from the GIS to the globe gives the *wrong results*



Global work-arounds: Landsat Global data



A (different) work-around for Sentinel-2



The Universal Transverse Mercator Projections (UTM is a *system* of 60 projections)

We are "*…shackled to a cartographic past*": "*GIS, an inherently digital system – relies on an inherently analogue spatial paradigm based on a flattened, scale-dependent 'map view' of the Earth that was designed for navigation with paper maps*" **Position Magazine**, 86, Dec-Jan 2017

Discrete Global Grid Systems are a new approach

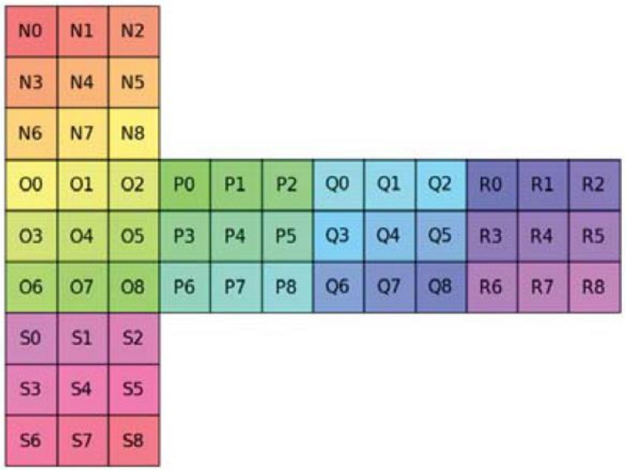
To bring data together we need a robust common approach that allows us to confidently integrate spatial data locally, regionally and globally.

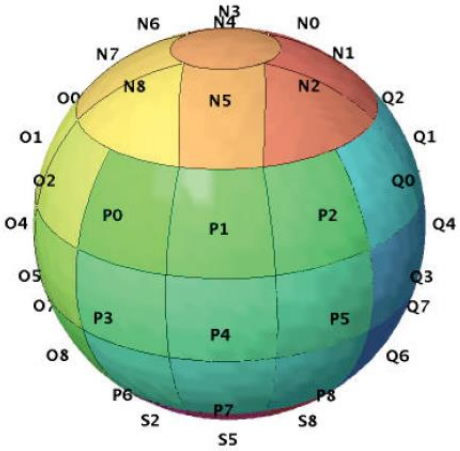
The core of the problem is how to address *areas* of the globe

- Globally – covering the poles etc. without loss
- At multiple resolutions – to arbitrary levels of detail

Discrete Global Grid Systems are a new approach

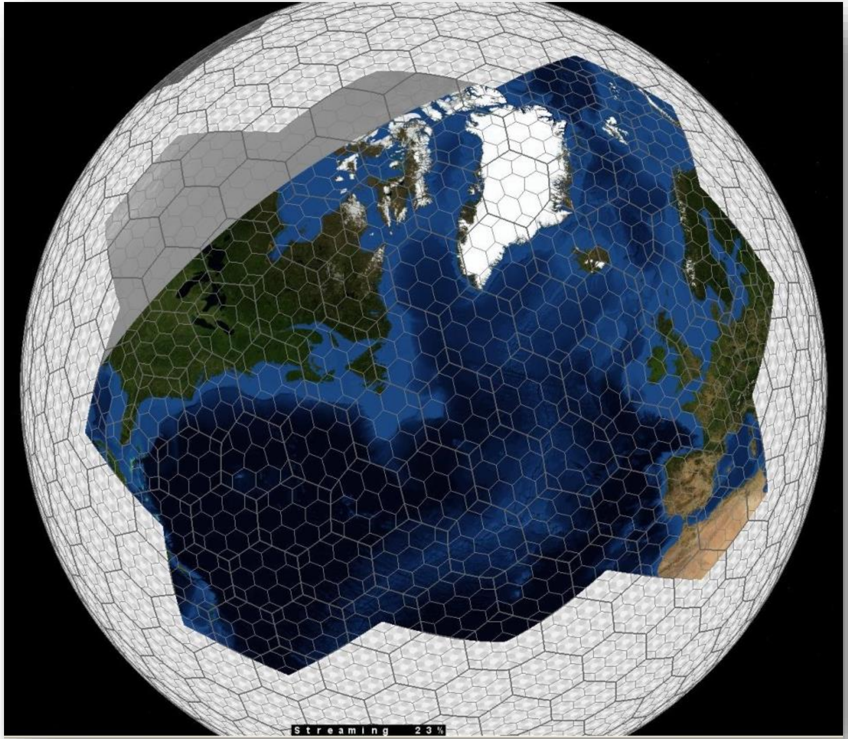
- DGGS directly address the geoid rather than going through a set of projections
- Work at multiple resolutions
- Build DGGS addresses of text strings *S0,S1,S2*, etc.

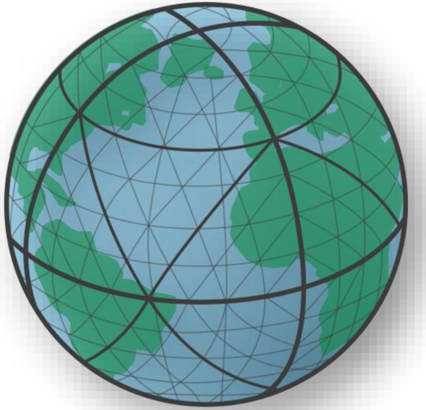




rHEALPix DGGS
Robert Gibb
Landcare Research
New Zealand

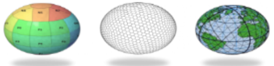
DGGS ‘flavours’



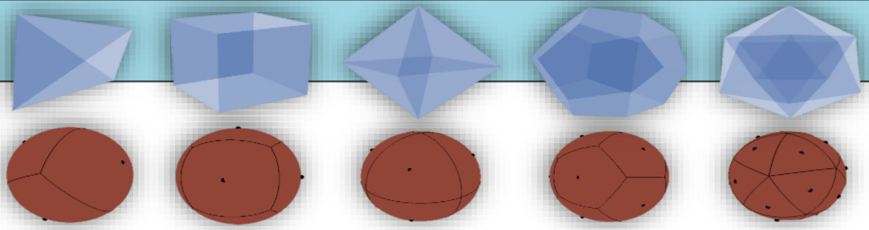


ISEA3H, Pyxis Innovation

Icosahedral Snyder Equal Area Aperture 3 Hexagonal Grid

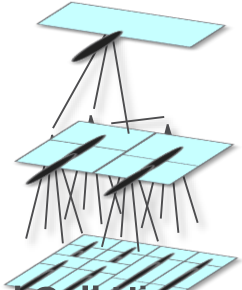


Discrete Global Grid Systems – Key Features

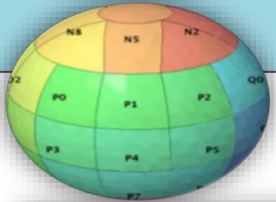


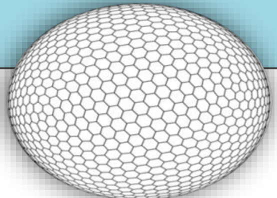
Tetrahedron Cube Octahedron Dodecahedron Icosahedron

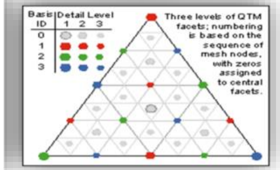
Equal Area Tessellations of the Earth's Surface

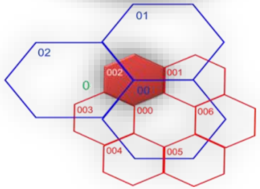


Fixed Cell Hierarchy





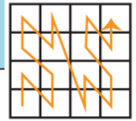




Simple cell Types:

- Rectangular
- Triangular
- Hexagonal

Massively Scalable



11	13	31	33
10	12	30	32
01	03	21	23
00	02	20	22

nD Spatial Analyses

↓

1D Array Processes

Globally Unique Cell Indices

00	01	02	03	10	11	12	13	20	21	22	23	30	31	32	33

Open Geospatial Consortium (OGC) standards

Open Geospatial Consortium (OGC) standards

OGC defines a DGGS as:

"*…a spatial reference system that uses a hierarchical tessellation of cells to partition and address the globe.*

DGGS are characterized by the properties of their cell structure, geo-encoding, quantization strategy and associated mathematical functions"

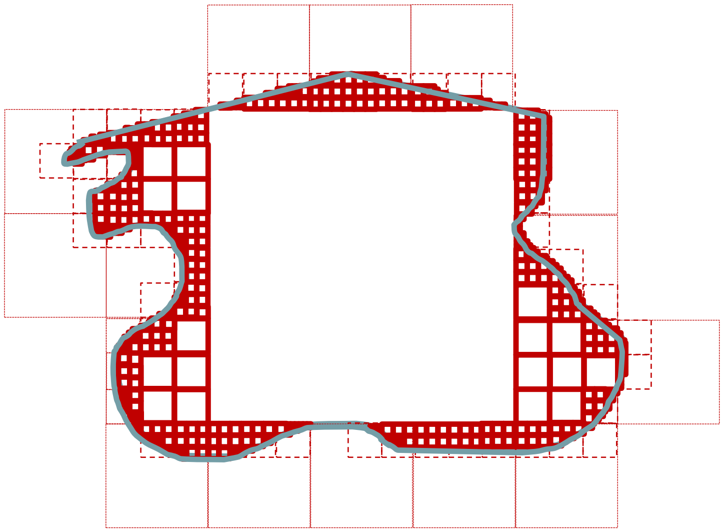
Purss, et. al. 2017, OGC Discrete Global Grid Systems Abstract Specification – Topic 21 [OGC 15-104r5]

Discrete Global Grid Systems DWG

Chair(s):

Sabeur, Zoheir (University of Southampton)
Peterson, Perry (the PYXIS innovation)
Purss, Matthew (Geoscience Australia)
Strobl, Peter (Joint Research Centre (JRC))

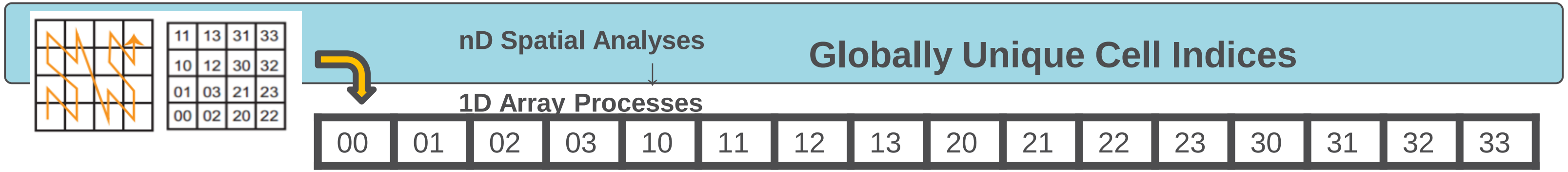
Hierarchical addressing / tessellation



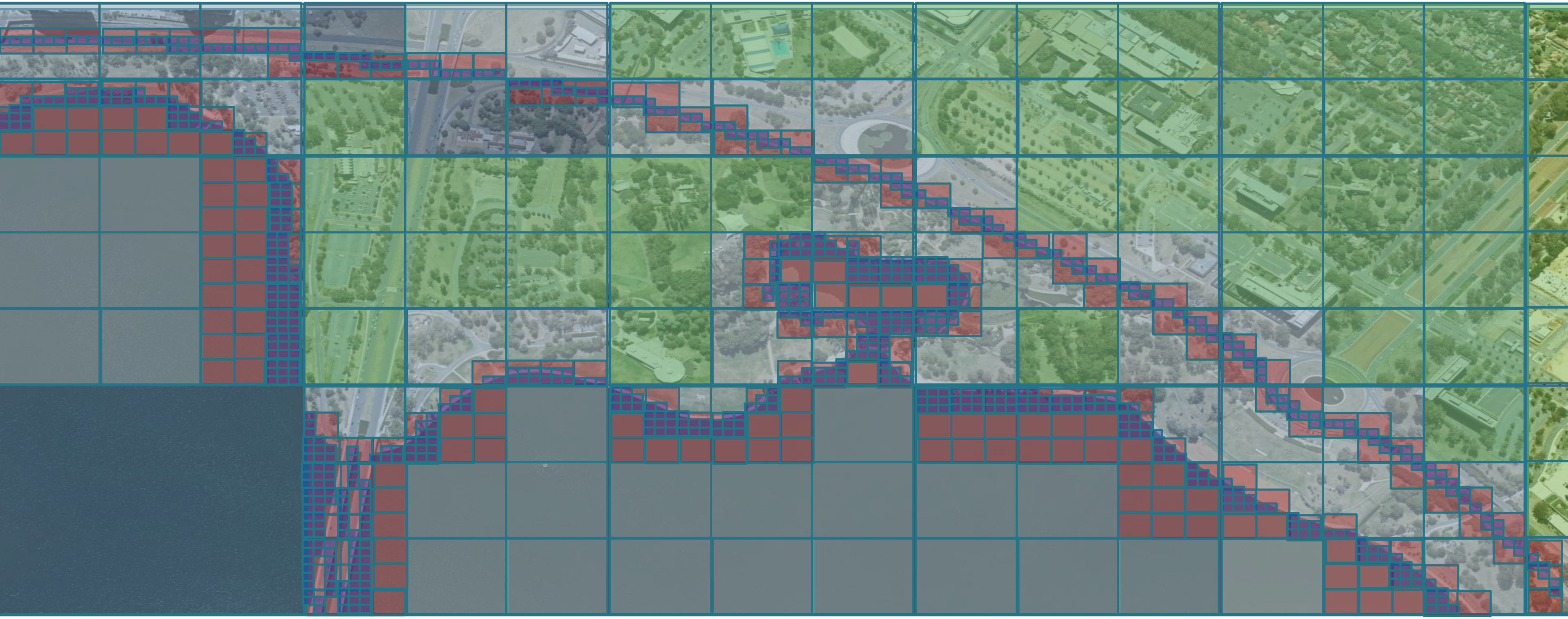
Encoding a feature into a DGGS reduces the geography to a series of 1-dimensional addresses suited to large scale computation

The length of the address string will be longer if the resolution is greater

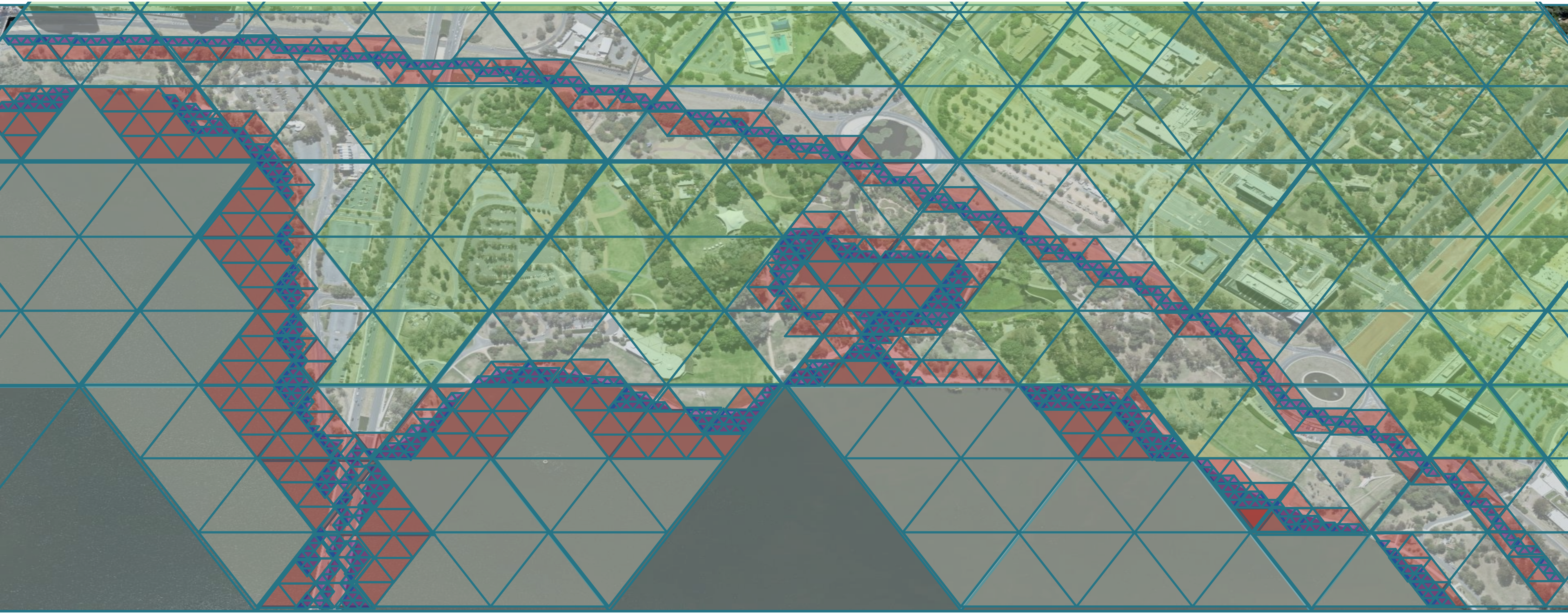
The address strings have strong logical relationships inherited from the geography



Hierarchical addressing / tessellation

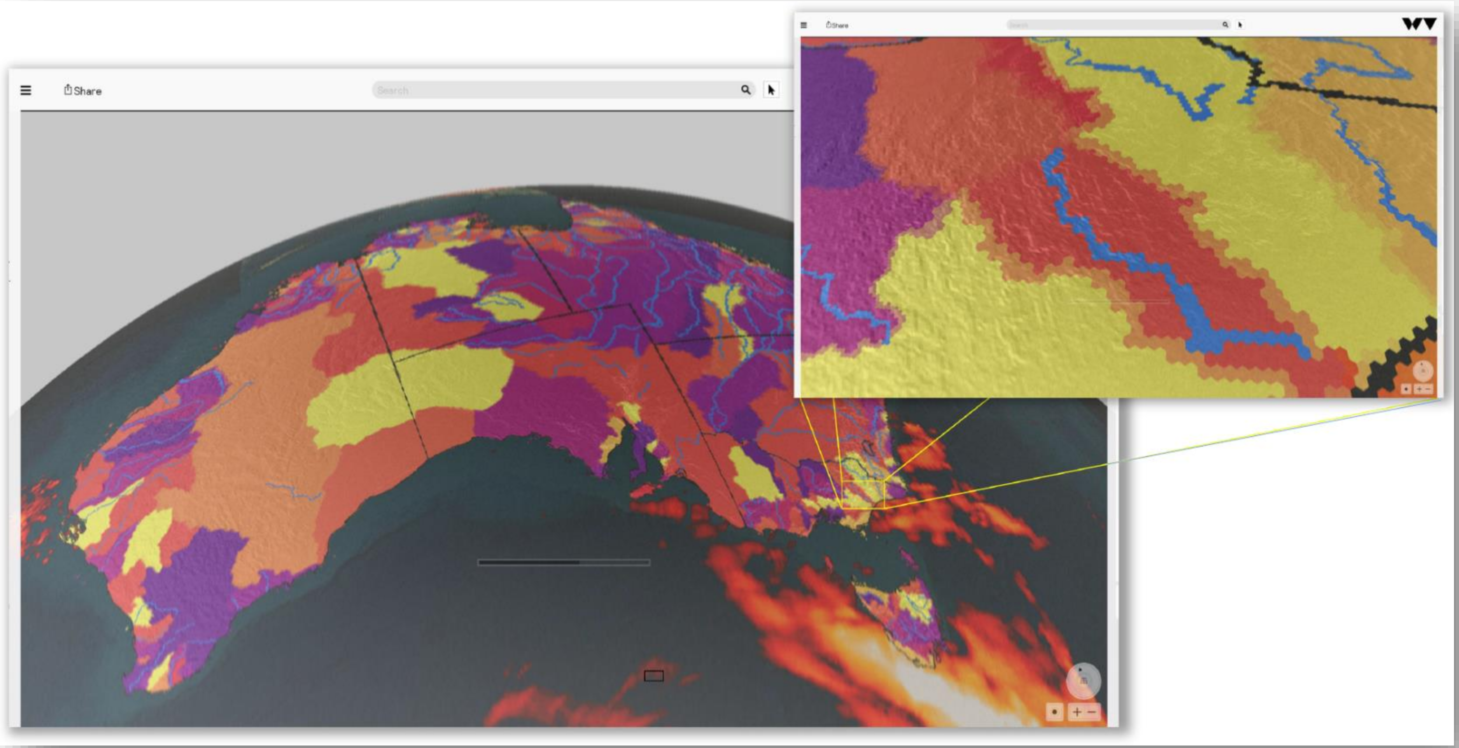


Hierarchical addressing / tessellation



Hierarchical addressing / tessellation

DGGS covering Australia



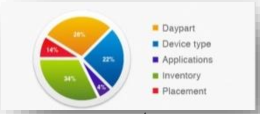
Pyxis Innovations

Integrating statistical and geospatial information

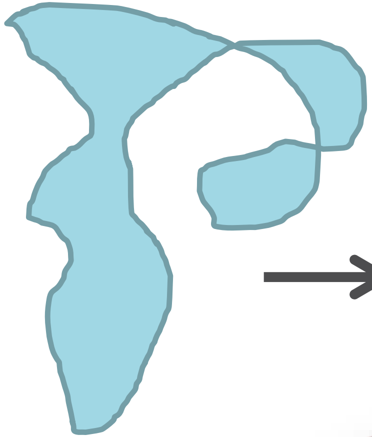
DGGS becomes a common addressing system that links geospatial and statistical communities via standard computing environments

- Computation is simplified and scalable
- Problems are no longer ‘spatial’

DGGS as a common addressing system

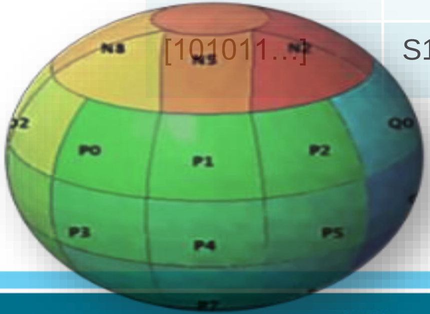


Integrated reports



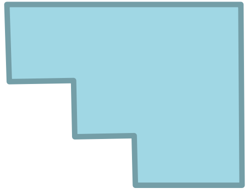
Data provider (geophysical)

Geospatial Data	DGGS-address
[101011…]	R0087333
[101011…]	R0087345
[101011…]	R0089773
…….	……
[101011…]	S1400555



query tools and protocols

DGGS-address	Statistical Data
R0083	[101011…]
S008757	[101011…]
S00873	[101011…]
……	…….
T1400555	[101011…]



Data provider (statistical)

Integrating statistical and geospatial information

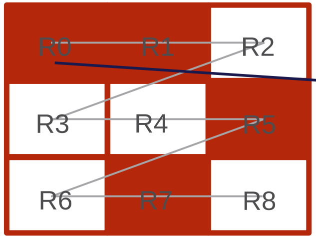
DGGS becomes a common addressing system that links geospatial and statistical communities via standard computing environments

- Computation is simplified and scalable
- Problems are no longer ‘spatial’

DGGS can support multiple geographies and allow sensitive information to be used ‘behind the scenes’

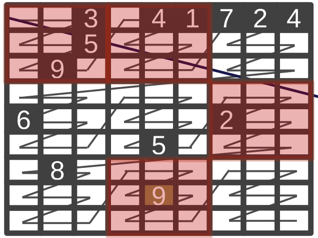
- ‘hidden layers’ of computing are supported much more readily – e.g., no need to ‘exchange GIS data’
- Interrogate sensitive data at full resolution & return results at aggregate level

Bushfire Impact Zone



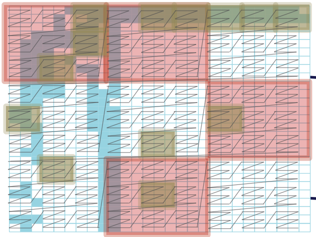
R0	R1	R2	R3	R4	R5	R6	R7	R8
Fire	Fire	Nil	Nil	Nil	Fire	Nil	Fire	Nil

Population (per dwelling)



R00	R01	R02	R03	R04	R05	R06	R07	R08
0	0	3	0	0	5	0	9	0

Access to Water



R070	R071	R072	R073	R074	R075	R076	R077	R078
water	Nil	water	water	Nil	Nil	water	Nil	Nil

R740	R741	R742	R743	R744	R745	R746	R747	R748
Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil	Nil

Using Statistical Boundary only

Fire Impacting whole area

65 People Impacted

65 People at risk but water is available

Using DGGS

Fire Impacting ~44% of area

33 People Impacted

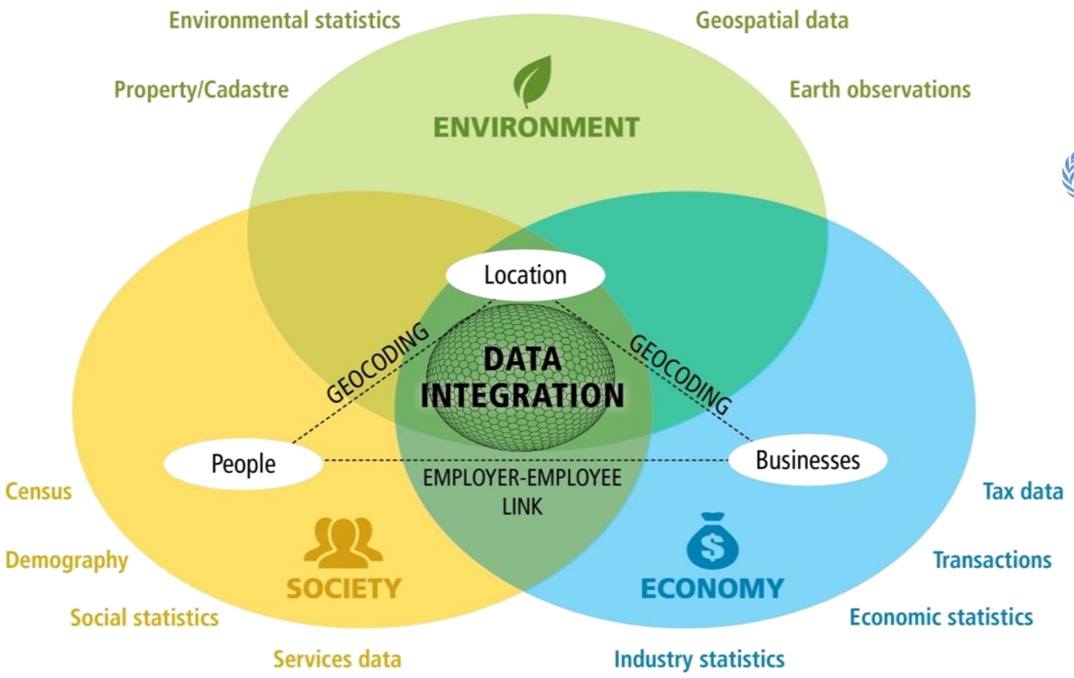
22 People at risk
11 People in critical danger

Integrating statistical and geospatial information

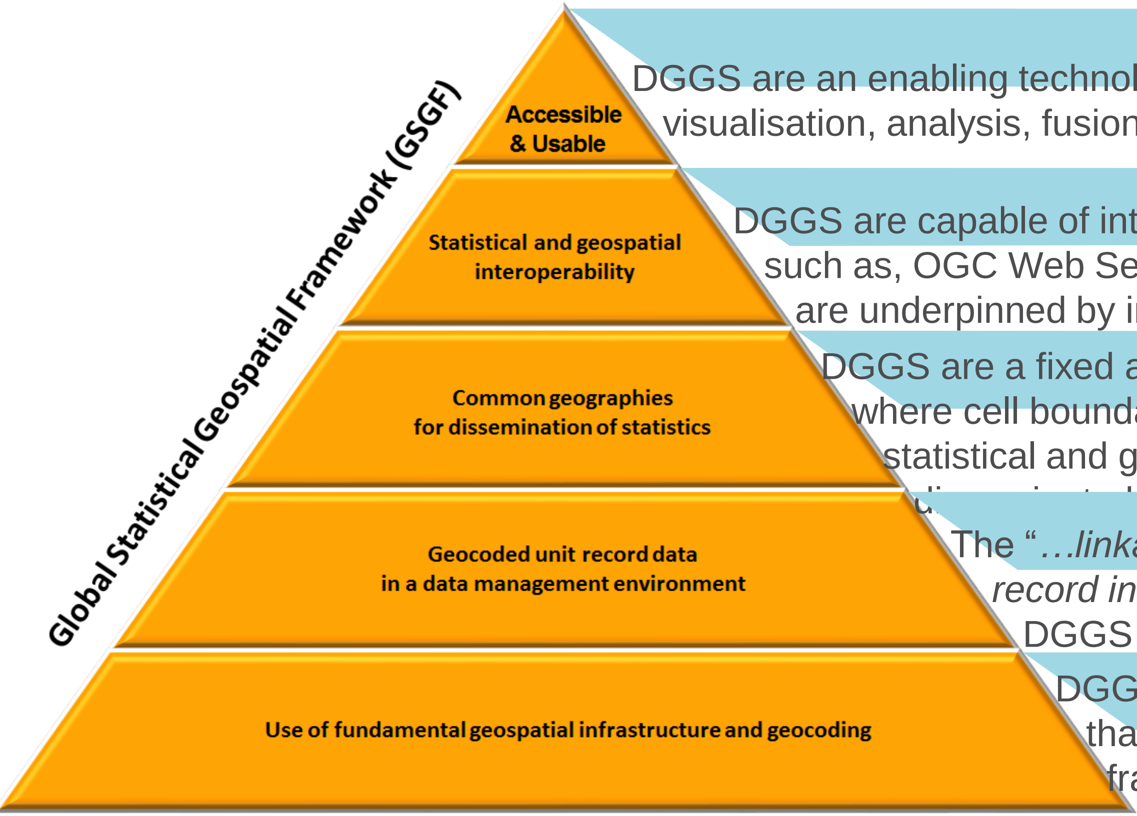
DGGS uses discrete addresses at multiple resolutions – it is not tied to pre-set geographies and can allow data to be integrated between geographies without necessarily disclosing detailed (potentially sensitive) information

DGGS can be discreet, as well as discrete ☺

Integrating statistical and geospatial information



DGGS within the GSGF



DGGS are an enabling technology that can facilitate rapid discovery, visualisation, analysis, fusion and dissemination of geospatial data.

DGGS are capable of interoperating with data exchange mechanisms such as, OGC Web Services and Linked Data architectures, and are underpinned by international data and metadata standards.

DGGS are a fixed and globally consistent spatial framework where cell boundary relationships are constant – allowing statistical and geospatial data to be aggregated and

The "*…linkage of a geocode for each statistical unit record in a dataset…"* is an inherent property of DGGS via their cell indexing schemas.

DGGS are a fundamental geospatial infrastructure that provides a common and consistent framework to store, reference and integrate spatial data to inform better decision making.

Some challenges to implementing DGGS

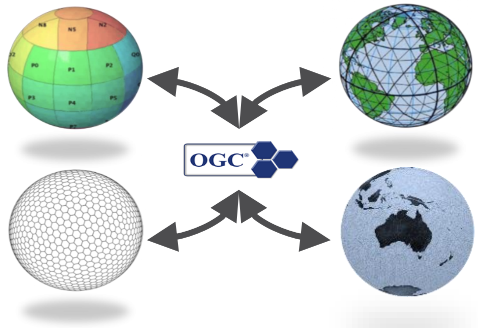
- Limited (but growing) awareness of DGGS throughout the geospatial community
- Limited (but growing) number of DGGS implementations currently available for use
- Building/Implementing a DGGS implementation from scratch requires specialist software development and geospatial skills
- Current spatial analytics software lacks interfaces to DGGS
- Big Earth Data providers (e.g. NASA/USGS/GA/ESA/Airbus/Digital Globe etc…) will need to re-engineer their data storage architectures and data collections in order to best take advantage of DGGS.

How is Australia supporting this change?

- Engaging with the Australian Bureau of Statistics (ABS) and other government and inter-governmental organisations to develop pilot DGGS implementations
- Envisioning an inter-agency spatial data infrastructure to integrate National Statistics and Geoscientific Information
- Building “AusPIX”, an Australian implementation of the “rHealPIX” DGGS under the Digital Earth Australia Program
- Continuing to support the development of international standards for DGGS, and build awareness of DGGS concepts, technologies and implementations across the broader geospatial and statistics communities.

DGGS in summary:

- Robust spatial framework that has many synergies with the proposed UN-GGIM Global Statistical Geospatial Framework.
- Will facilitate integration of statistical and geographical data from multiple communities
- A geographical addressing system – replaces 'maps' with common addressing. Avoids map-projections; is multi-resolution
- Development is being driven by academic, government, commercial and standards interests and expertise, and is underpinned by a growing awareness of DGGS as a necessary spatial information infrastructure

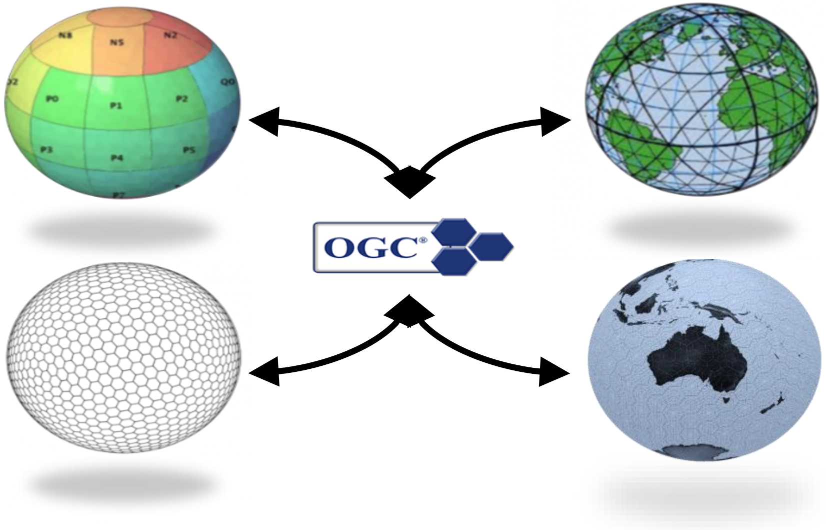


Conclusion – what should this community do?

- Anticipate DGGS as *Disruptive*, *Inevitable*, *Challenging* and *Rewarding*
- Engage with and support the emerging DGGS discussion. DGGS as a standing item in relevant meetings?
- Seek to support and be involved with test-beds, prototypes and workshops
- Encourage this work to progress as quickly as possible. With the exponential growth in data volumes, the longer it takes the more difficult it will become to change legacy systems.

Australian Government
Geoscience Australia

Questions?



Phone: +61 2 6249 9111

Web: www.ga.gov.au

Email: matt.purse@ga.gov.au; adam.lewis@ga.gov.au

Address: Cnr Jerrabomberra Avenue and Hindmarsh Drive, Symonston ACT 2609

Postal Address: GPO Box 378, Canberra ACT 2601