Global Grids for CARD

indispensable or unnecessary?

Peter Strobl, EC-JRC

CEOS LSI-VC-10 virtual meeting, 20 May 2021

European Commission

Analysis Ready Data (ARD)

An Analysis Ready Data (ARD) product is generated from raw data and processed so that it can be used **without the need for further processing** to be applied by users.

… minimum processing requirement to be an ARD-compliant product: the data **must be processed to a geo-referenced projection** to enable the position identification within the data product. …

- If “geo-referenced projection“ is to be understood as being transformed to a “***georeferenced grid***” than this requires **re-sampling**!

CEOS Analysis Ready Data (ARD)

CEOS Analysis Ready Data (CARD) are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and **interoperability** both through time and **with other datasets**.

- It is fair to assume that “other datasets” means those who also meet the CARD requirements, i.e. they are “gridded” (and thus re-sampled).

CARD Interoperability

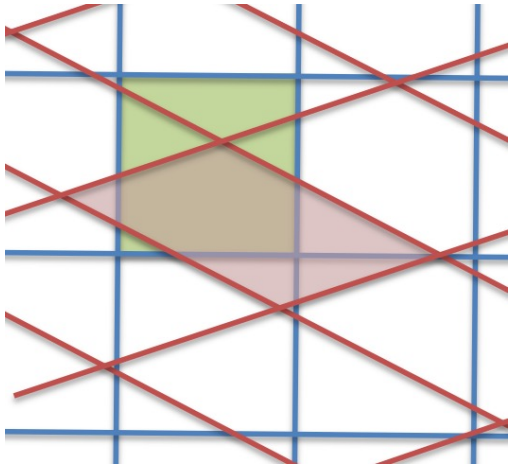
***Interoperable** Products* refers to a set of two or more ARD products which are sufficiently documented **to enable processing** across a continuum of geometric and/or radiometric standards **to permit direct** quantitative **comparison**."

- Not the *interoperable products* themselves are able to 'interoperate' (i.e. be compared or analysed together) but their derivates.
- 'Interoperability' here means a 'can' not an 'is' and, if the underlying references are not the same, it requires adaptation (processing)!
- How ready is 'ready' in ARD?

Changing references

Observations need to relate to the same standards or references to be comparable:

- scaled data (based on points)
  need to be **re-scaled** (e.g. ºF → ºC)
- gridded data (based on intervals)
  need to be **re-gridded** (or resampled)
  (e.g. 1” WGS84 /Pseudo Mercator EPSG:3857
  → 30m Mollweide EPSG:54009)



➢ Sounds doable for continuous parameters which can be interpolated (e.g. radiance, temperature), but categorical data (e.g. masks/flags)?

ARD and Datacubes

- ARD are meant to build data cubes!
- OGC data cube Community Practise* says:

All layers in a data cube need to share the same grid
to allow interoperability between layers

➢ Co-gridding is an important element of interoperability WITHIN a data cube

*https://portal.ogc.org/files/18-095r7

Datacube Interoperability

BUT

Two data cubes (or ARD datasets) do not need to share the same grid to be considered interoperable(?)

- If so, is interoperability restricted to a ‘one way’ road?
  (i.e. a specific cube can only be involved once during an analysis workflow)
- And then, how is reproducibility being secured?
  (e.g. for Cubes A,B routing A→B gives a *slightly* different result as B→A, repeating exchange and involving more Cubes worsens things considerably!)

Consequences of re-sampling for interoperability

(unless the volume of data is largely amplified each time)

- ❑ always entails an interpolation of data
- ❑ always diminishes data accuracy or entail data loss
- ❑ always is irreversible
- ❑ is (more or less) computer-intense
- ❑ **accumulates these effects when repeated!**

➢ How compatible is this with FAIR principles?

Big Geospatial Data Analysis Strategies

To avoid repeated re-sampling in complex multi-source environments there are essentially two options:

- "Point Cloud" approach:
  - Store all observations with their locations (as n-tupels)
  - Resample (all input data) to a user selected grid only at the point of analysis
  - ➢ High processing effort, only end-to-end processing, low re-usability

- "Grid System" approach:
  - Discretise (re-sample) all observations to common grid system (not only in spatial dimension!)
  - ➢ Limited number of (spatial) representations, lack of user acceptance

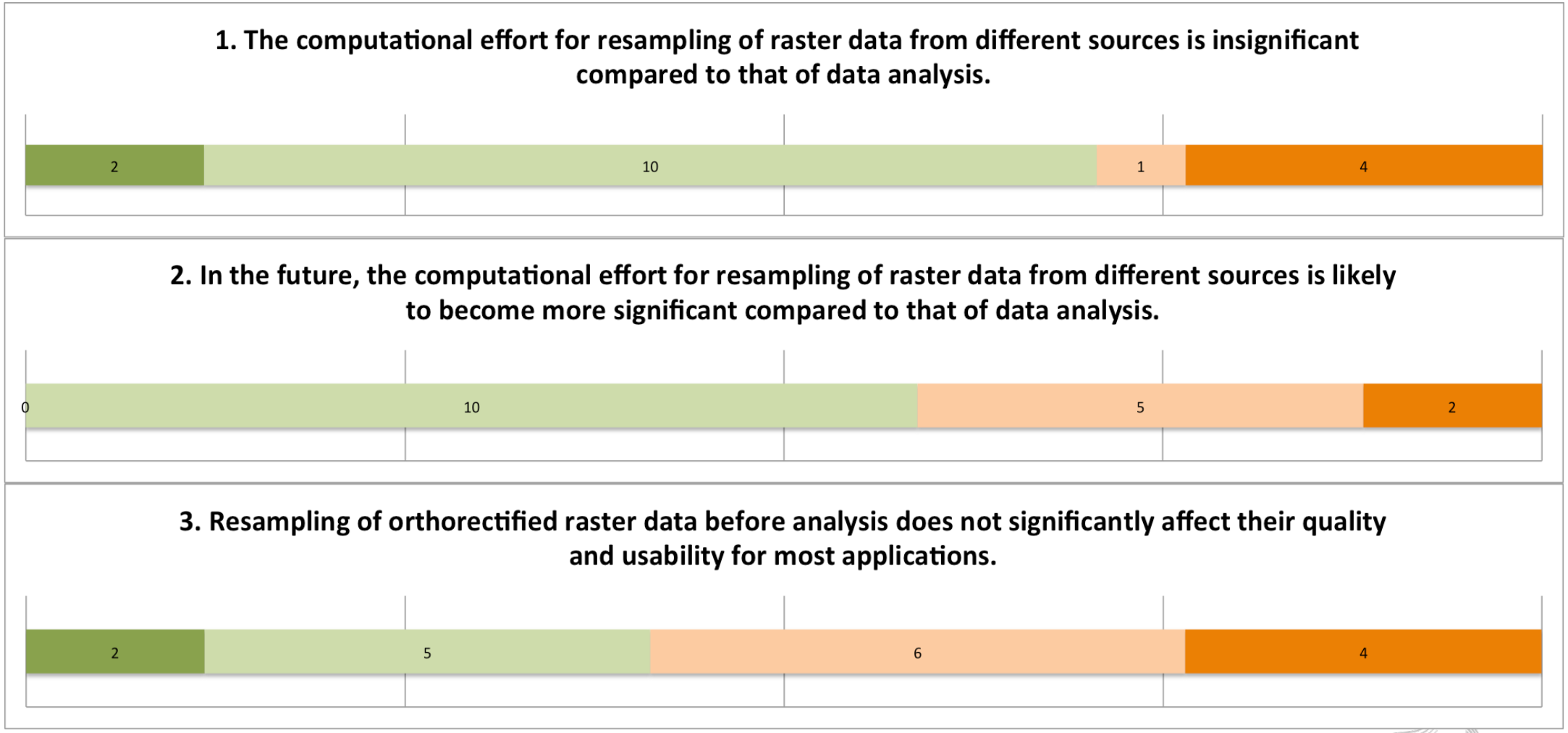
INSPIRE* wisdom

"… it would be highly desirable that all the themes with similar needs make use of the same geographical grid system in order to maintain their coherence."

Source: INSPIRE D2.8.II.1 Data Specification on Elevation – Technical Guidelines (2013)

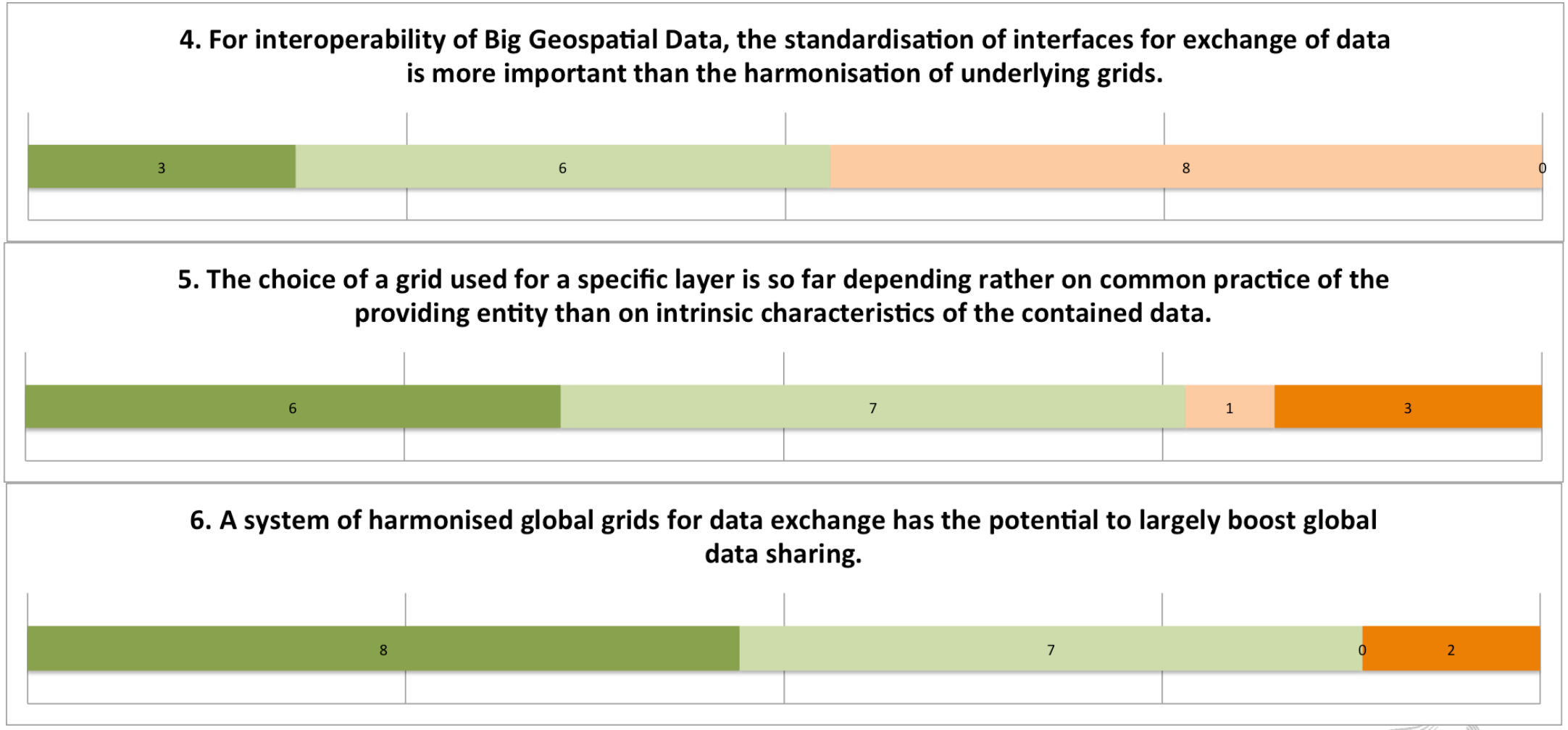
*INSPIRE is the EU initiative to establish an infrastructure for spatial information in Europe that will help to make spatial or geographical information more accessible and interoperable for a wide range of purposes supporting sustainable development. https://inspire.ec.europa.eu/

JRC-INSPIRE GRG Workshop 2017

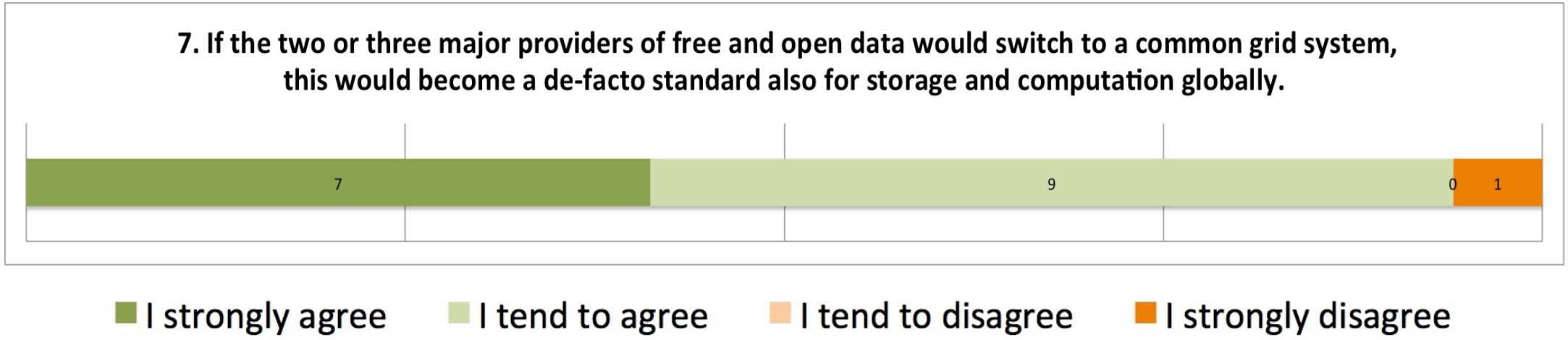


European Commission

JRC-INSPIRE Workshop 2017 Questionnaire



JRC-INSPIRE Workshop 2017 Questionnaire



Geodata representation in the 21st century

Questions for the leading global EO data providers:

Continuous (point-clouds) or discrete (grids)?

If grids then:

Global or continental,
mono-resolution or hierarchical?

Which criteria for 'good' global grids?
(e.g. Goodchild/Kimerling)

Main candidate global grid(system)s?

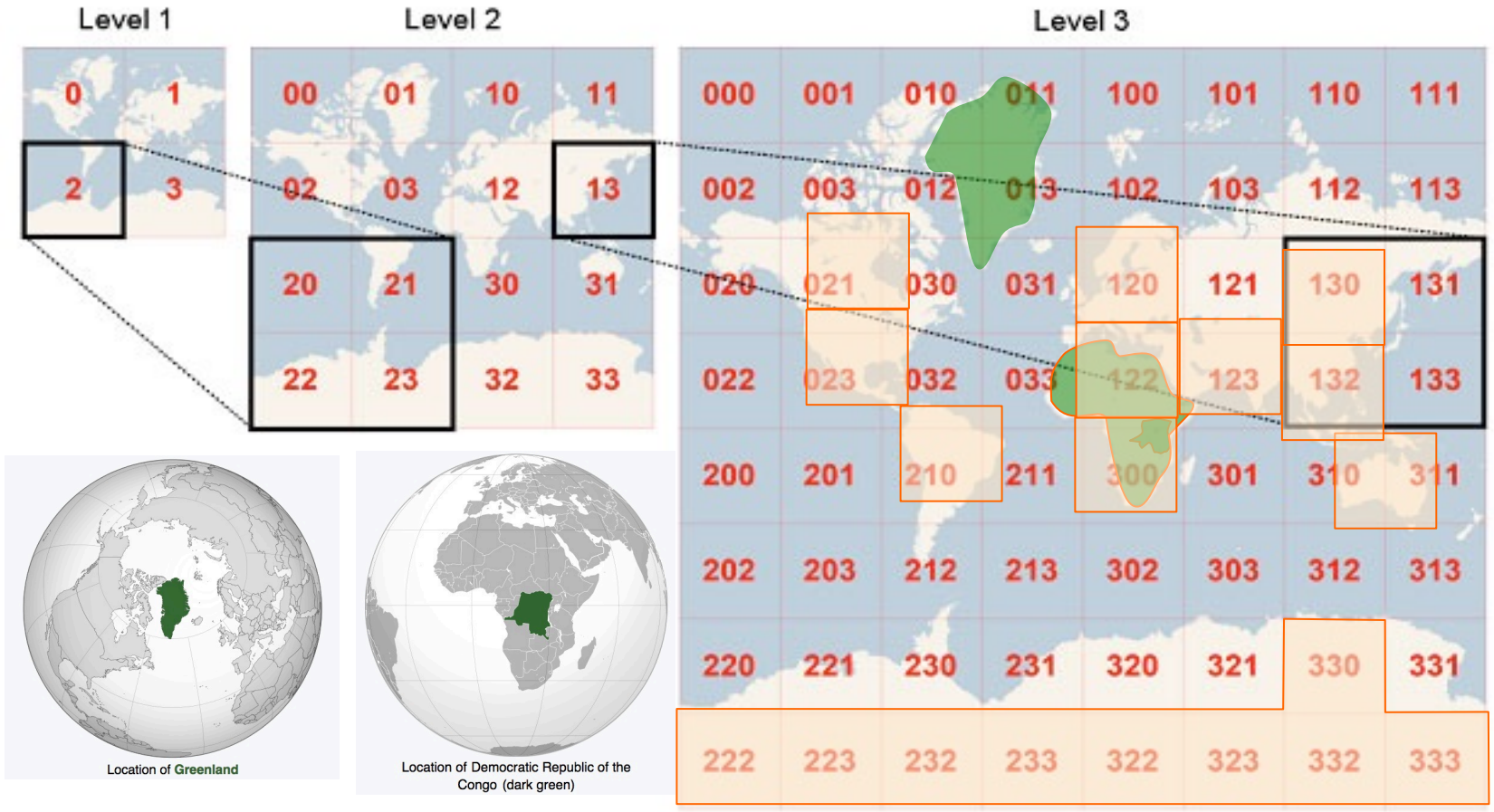
Separate or together?

M. Goodchild (2019):
So, in the final analysis, the big-picture question for DGGSs remains the same as it has been for more than two decades: *how do we use the compelling arguments for these multi-resolution systems to persuade the larger scientific community to adopt them, in preference to the distorted representations of digital maps?*
https://doi.org/10.3138/cart.54.1.preface

Criteria for (spatial) discretisation

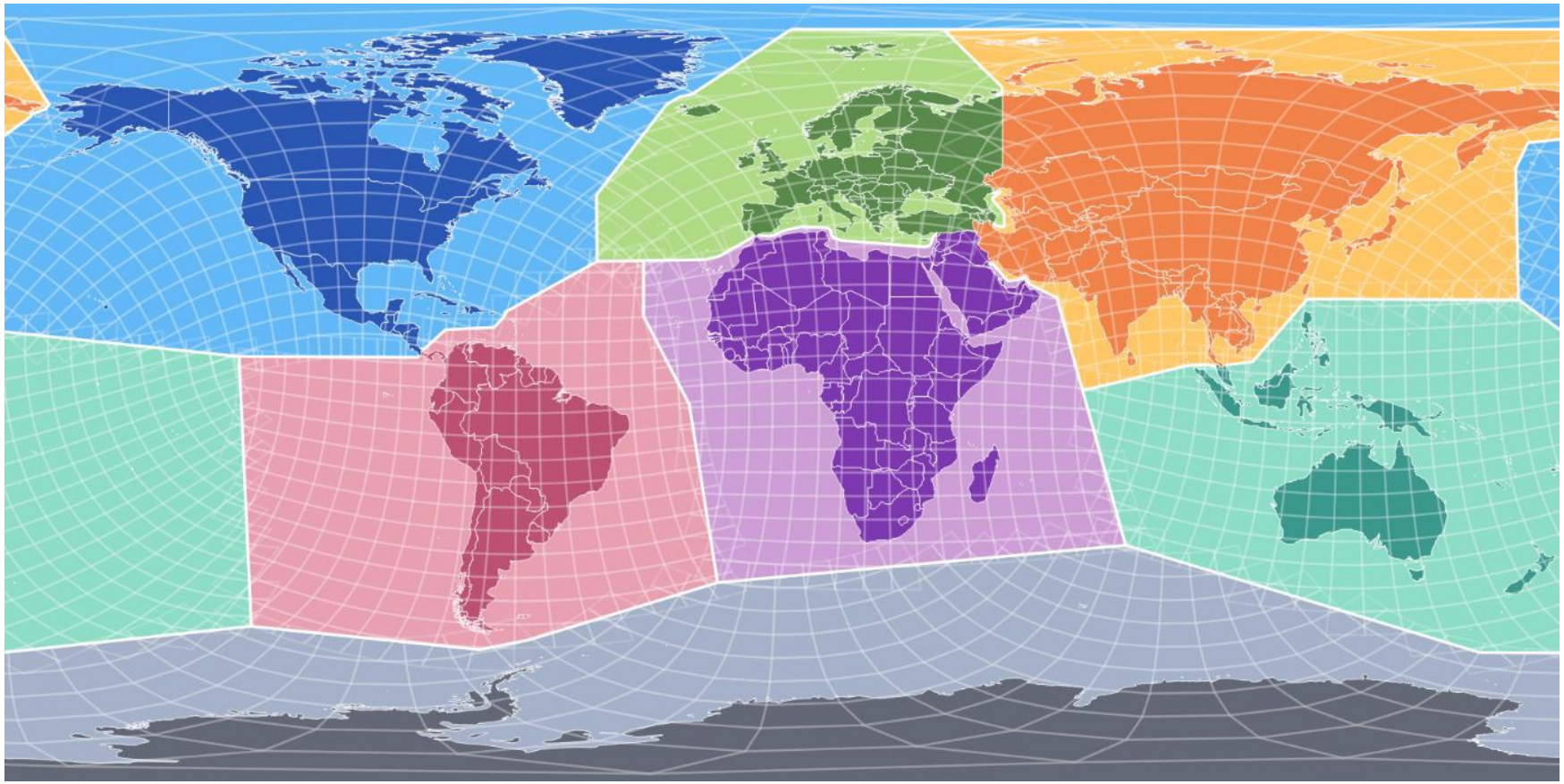
- assessable: based on ellipsoidal Earth model
- unambiguous: every point on the surface belongs to a cell
- gap free: no point on the surface belongs to more than one cell
- hierarchical: grids can be refined from coarser to finer levels following mathematical rules (cell refinement)
- nested: finer level cells do not overlap coarser cells
- intrinsic: the grid is a product of a mathematical tessellation of the ellipsoid, a cell is only determined by location
- instantaneous: the grid is defined for any point in time

The WMTS standard (base EPSG:3857)



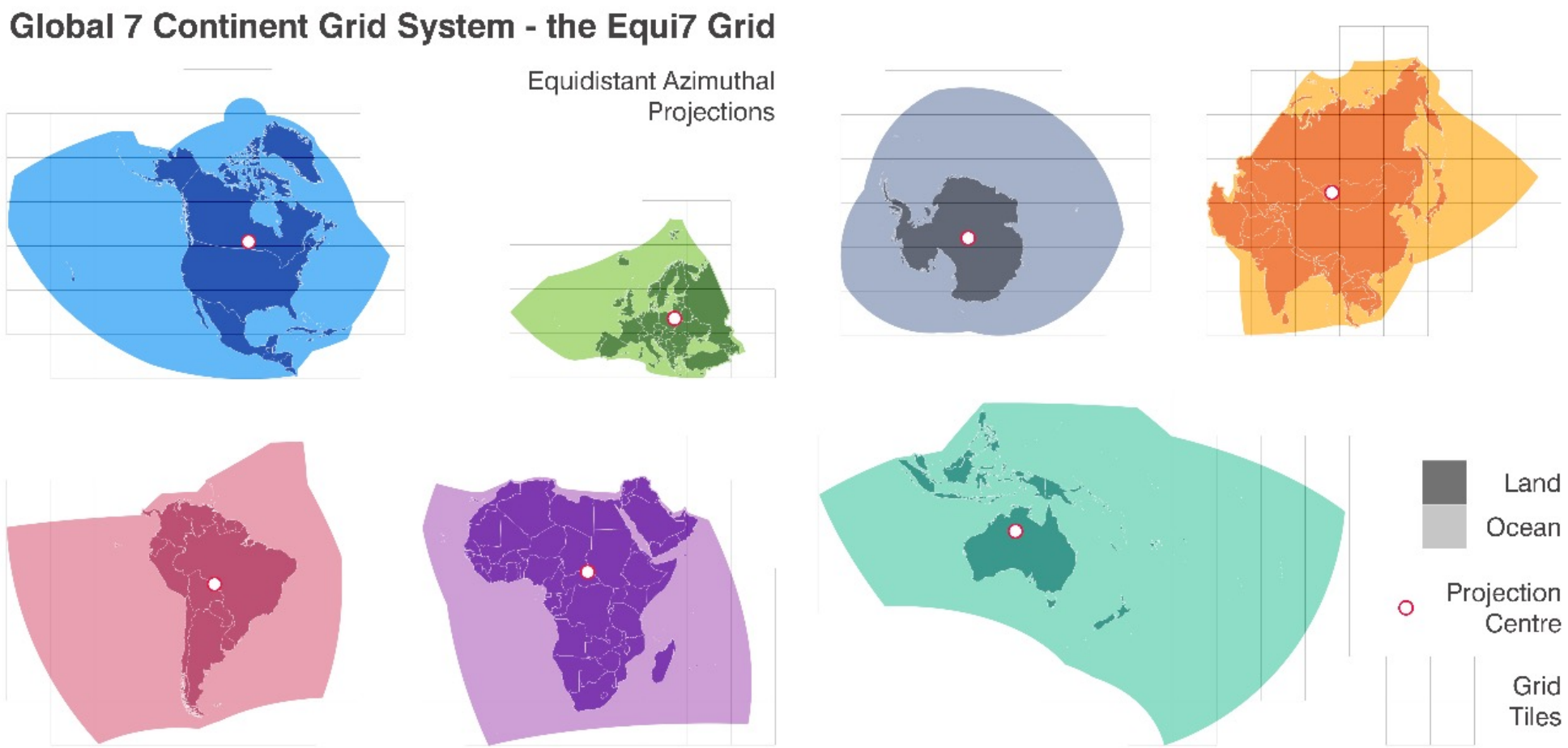
European Commission

The EQUI7 grid (TU Vienna)



B. Bauer-Marschallinger, Optimisationof global grids for high-resolution remote sensing data, Computers & Geosciences, 2014
doi:10.1016/j.cageo.2014.07.005

The EQUI7 grid (TU Vienna)

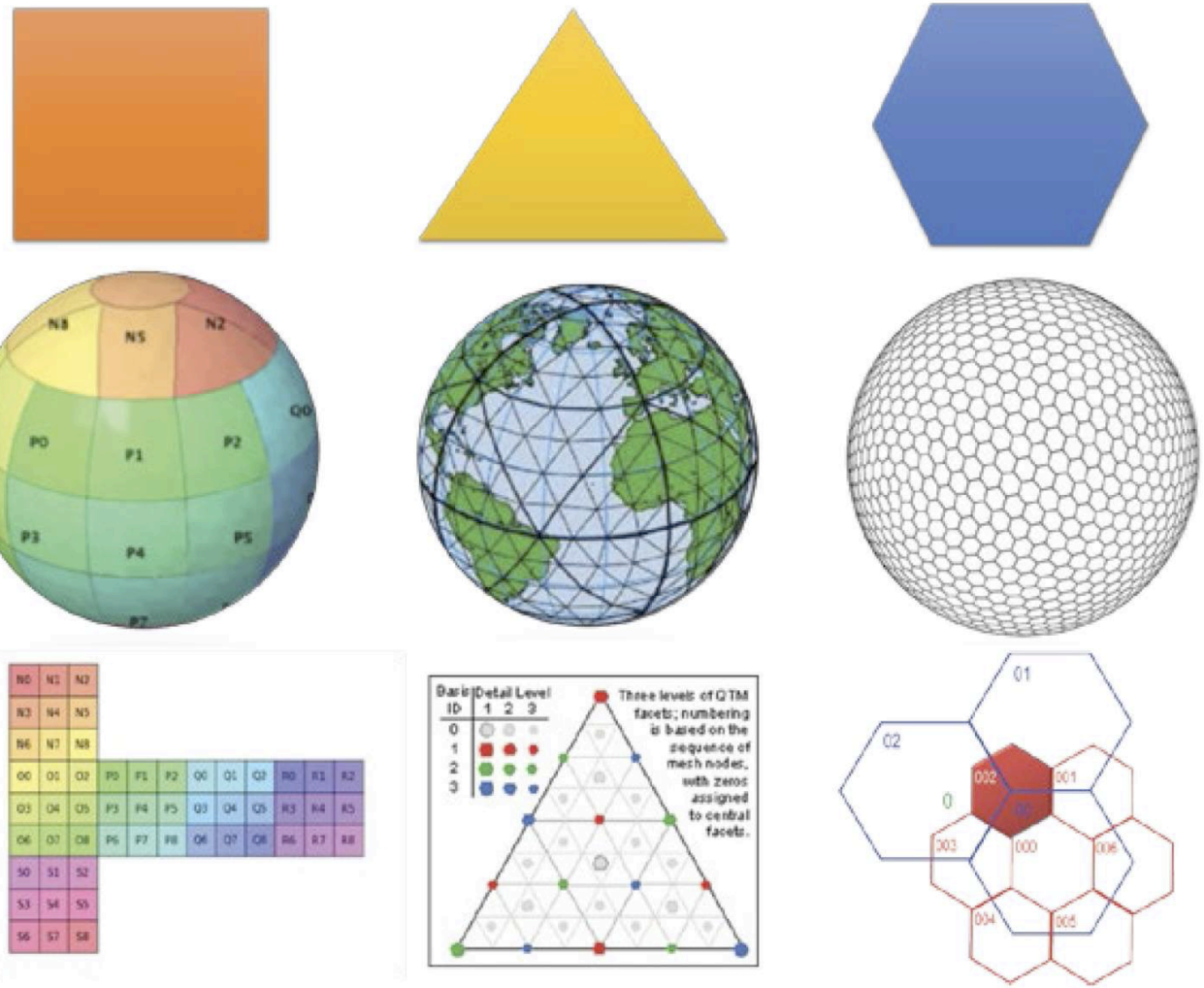


Discrete Global Grid Systems

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

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

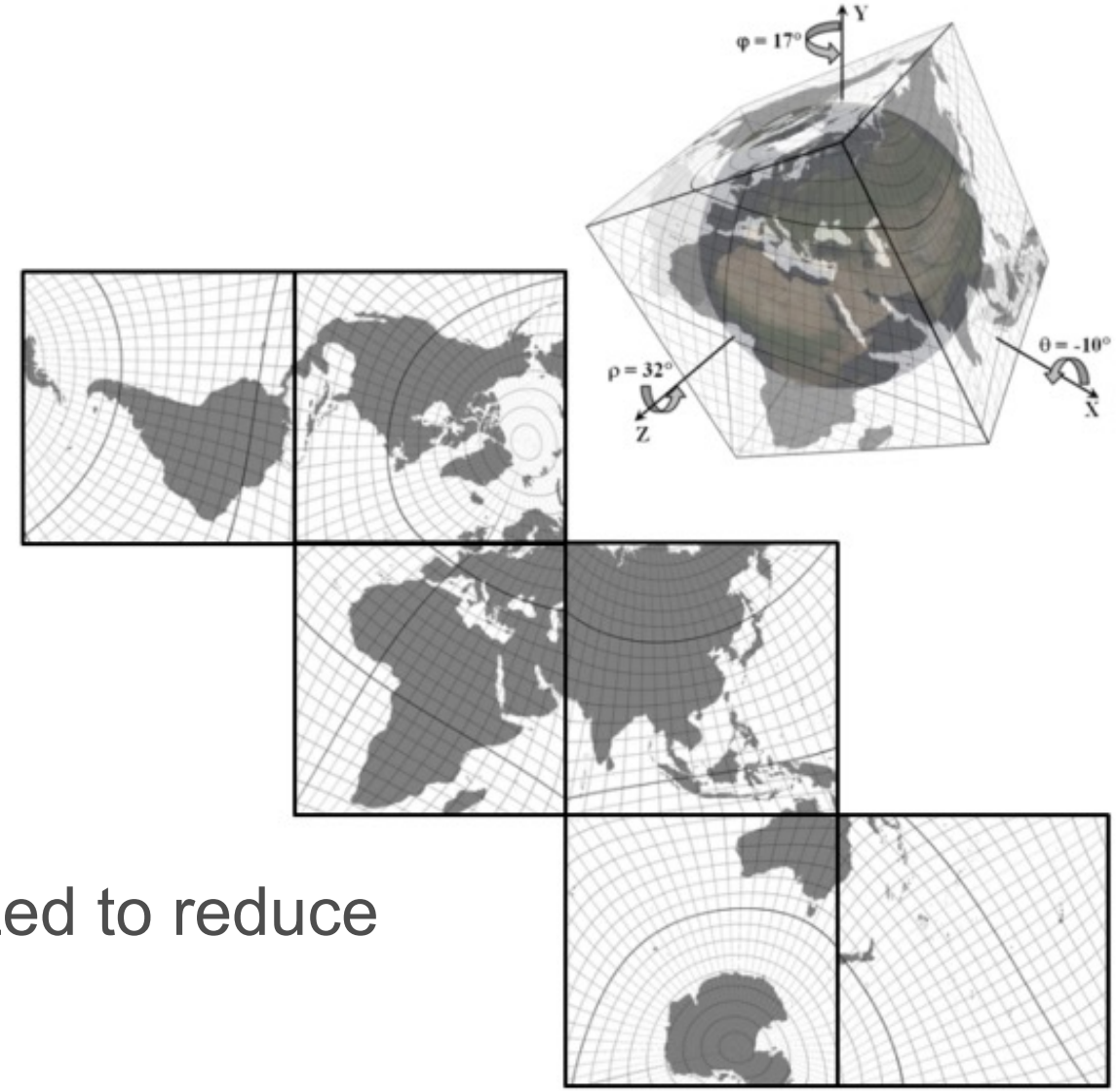
ISO 19170-1 Geographic information — Discrete Global Grid Systems Specifications — Part 1:Core Reference System and Operations, and Equal Area Earth Reference System
https://www.iso.org/standard/32588.html



European Commission

DGGS optimization

- After the main choices, i.e.:
  - shape (”square”)
  - Refinement ratio (4 - quadtree)
  - ➢ “cube-sphere mapping”
- Cube sphere mapping can be optimized to reduce distortions e.g. over land masses



Dimirijevic A., Strobl P., *Continuous 2D Maps Based on Spherical Cube Datasets,* Proc. 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), doi:10.1109/ICEST49890.2020.9232678, 2020

Discussion

- Does CARD require gridding?
- Is interoperability (in a processing chain restricted to a ‘one way’ road?
- How is reproducibility being secured when data are resampled?
- How compatible is repeated resampling with FAIR principles?
- Is there room for a “common global CARD grid system”
- Which are the top priority criteria for a CARD global grid?

Thank you!
Any questions?

peter.strobl@ec.europa.eu

European Commission