What is Landcover?

Thoughts about the abstraction of landscape in a Digital Earth.

The European Commission's science and knowledge service

Joint Research Centre (JRC)

Peter Strobl
EC-JRC, D.6

1st MAESTRIA User Workshop
IGN, 3 Dec 2019

Outline

- The Digital Earth and the role of discretisation
- What is Land Cover?
- Hierarchical surface categorisation
- Examples and Discussion

The Digital Earth

- Digital Earth is an abstraction (or virtual representation) of the 'real' analogue environment in an IT infrastructure.
- One major enabling step being discretisation of otherwise continuous phenomena.
- *The* **International Society for Digital Earth (ISDE)** in its 2019 Florence declaration recommends: *"that in developing Digital Earth concept, efforts must be made to take full advantage of existing and novel technologies and explore new avenues to* ***facilitate data access and re-use****."*

Why a Digital Earth?

- Data deluge – “Big Data”
- Objective (measure) vs. subjective (interpret) approach
- Enhanced interoperability
- “Data Democracy”
- Building of “Data Cubes”
- All information, all the time ...

The FAIR data promise

- **Findable**
  metadata, terminology, cataloguing, standardisation
- **Accessible**
  free&openness, efficiency, standardisation
- **Interoperable**
  interfacing, standardisation, harmonisation, unification
- **Reusable**
  licensing, scalability, modularity, hierarchy

Enablers of interoperability

- **Standardisation**: compare to something established by authority, custom, or general consent as a model or example
- **Harmonisation**: be compatible, similar or consistent; coincide in their characteristics
- **Unification**: combine and coordinate diverse elements so that they can be considered a whole

→ What is appropriate depends on the reversibility of the interaction!

The role of discretisation

Discretisation is necessary to

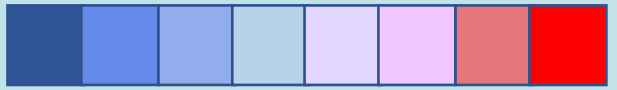
- digitise
- focus
- reduce
- simplify
- decide
- …

Discretisation in practice

Continous,

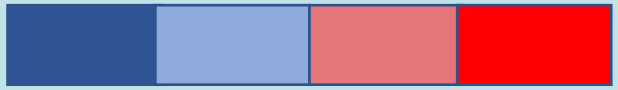
2 categories,

8 categories,



- Discretisation is too coarse if within-interval differences are too large to treat all class members the same
- Discretisation is too fine if cross-interval differences are too small to distinguish different intervals
- Colours are easy – but how about temperature?

or 4 categories



Discretisation prerequisites

- a clearly defined and limited feature space:
  What is being discretised?
- an information/data base:
  Which characteristics (or measures) are available?
- A set of criteria:
  How can the different samples according to their characteristics be assigned to intervals?

Good practise in discretisation

- **assessable:** all properties used for distinguishing intervals must be observable/measurable,
- **unambiguous:** intervals mutually exclude each other,
- **gap free:** each valid observation/measurement sample can be assigned – depending on noise with a certain probability - to a specific interval,

Good practise in discretisation cont'd

- **intrinsic:** assignment of an observation/measurement to an interval is independent of other, non observed/measured properties,
- **instantaneous:** (assignment of an observation/measurement to an interval is independent of that of other observations/measurements).
- **hierarchical:** different granularity in intervals is achieved by hierarchical nesting

Why is ”Land””Cover” important?

Because it is the main ***interface*** between:

- the **Geo(bio)-shere** and the **Atmosphere**
- the **Humans** and their **Environment**

➡ Keep in mind that these are two very different roles!

What is “Land”“Cover”?

Terminology is important: it triggers ‘imagination‘

The term suggests that the feature in question is something solid (*“land”)* with a layer of something else on top (*“cover”)*.

This is very often **NOT** the case!

How to observe “Land”“Cover”?

Humans are predestined to capture their environment with their visual sense.

- Therefore our perception of “Land”“Cover” is largely determined by the visible optical properties of the surfaces that surround us.
- PLUS a great deal of interpretation which the brain does based on an individual and years long training.

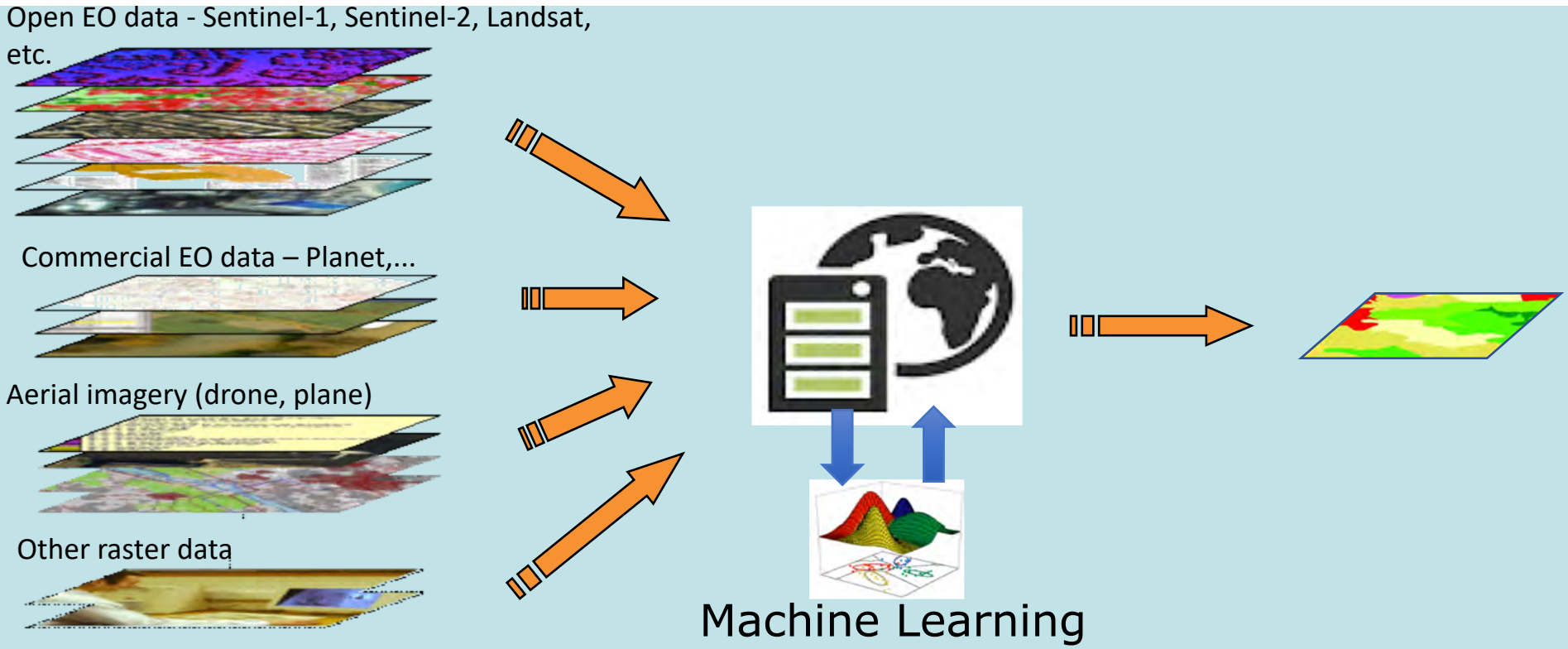
How to digitise “Land”“Cover”?

- So far humans categorised the visual appearance according to contextual properties and –quite a lot of- ‘experience’.
- Visual sensing is replaced by digital imaging devices, so the observations are discrete (and if calibrated can be called measurements)
- These measurements are synthesised into discrete categories using a complex process mimicking the ‘experience’ of the interpreter.

Image to map

- Until recently LC maps were mostly the result of an interpretation of one image, or a few similar ones, usually based on training samples
- Human interaction was still feasible (and often mandatory)

Big Data to map



- Processes need to be automated, human interaction minimised

Adapted from "The European Data Cube Facility Service (DCFS)", G. Landgraf, ESA

System Engineering approach

- How to retain control over an increasingly complex process?

➢More complex problems become easier to address if broken down into smaller pieces.

➢System engineering: ‘separation of concerns’

Refocus “Land”“Cover”

- 'land cover' as a term should be revised as it implicitly adds a complexity which distracts from the actual problem.
- What is important as a **first step**, is to characterise the **bio-geo-physical properties of the planetary surface** (regardless on whether it is 'land' or whether it is 'covered') by **means of a number of discrete classes** (“intervals”).

Planetary Surface

- instead of 'cover' (which implies some difference to what is below) the term 'surface' should be used as the uppermost (in gravitational direction) **interface** between the solid or liquid planetary body and its atmosphere (if any) or the space vacuum.
- In case of the Earth it is then at the same time the lower bound of the atmosphere.
- The **depth extent** of what is still regarded as 'surface' depends intrinsically on the measurement method (or sensing technology)

Good practise in discretisation of “surface”

- **assessable:** classes distinguishable by radiometric (optical?) properties
- **unambiguous:** each surface element is (at least by majority and with a specific uncertainty) assigned to only one specific class
- **gap free:** for any possible surface always one class applies
- **intrinsic:** class assignment purely based on radiometric properties
- **instantaneous:** each surface element (independent of size) can be classified based on a single observation of its radiometric properties
- **hierarchical:** start with most important properties and refine gradually

Hierarchical Surface Categories: an attempt

bit 1 substance
- 0 other than water
- 1 water

(because water/non water is the primary characteristic of Earth's surface)

bit 2 aggregation state
- x0 liquid
- x1 solid

(because aggregation state is the primary characteristic wrt to the property of the substance)

results in 4 principle surface categories:
- 00 : magma (other non-water liquids on the surface e.g. carbohydrates or mercury are very rare)
- 01 : land
- 10 : water
- 11 : snow, ice

HSC cont'd

bit 3 photosynthesis

- xx0 photosynthetically active
- xx1 photosynthetically in-active

because photosynthesis is the primary visible (optically detectable) evidence of live

allows in total 8 (2^3) classes, of which the following four, two on land and two on water, are widespread:

- 011 : water without apparent photosynthetic activity
- 010 : water with apparent photosynthetic activity (i.e. algae)
- 101 : dry or non-vegetated land surfaces
- 100 : (living) vegetation

Another two classes are more or less irrelevant because (bio-)physically impossible or very rare

- 000, 110 : i.e. photosynthetic activity in non-water liquids (magma) or ice

and therefore of the remaining two:

- 001: is synonymous to 00 (magma)
- 111: is synonymous to 11 (solid water, i.e. snow, ice)

strictly speaking one could stop here as a further distinction applicable to all classes (generic biogeophysical criterium) is difficult to define.

Hierarchical Surface Categories: the don’ts

It is of utmost importance to keep in mind that for surface classification the following attributes must be avoided:

- **‘permanent’, ‘periodical’, ‘temporal**’, as they are not instantaneous
- **'sealed'** is unsuitable as it is contextual and scale-dependent
- **'anthropogenic'** or **'artificial'** is generally not a physical property and such not measurable (in many if not most cases the very same surface can be either or), in addition it lacks an unambiguous definition
- reference to **size** or **shape**, as these are not scale invariant

From Surface to Cover

Surface categories cannot replace Land Cover or Land Use definition and classes, but they are an important step towards a **modular, consistent and transparent** definition of Land Cover and Land Use.

"Land Cover" could be built based on three dimensions as follows:

- ✓**surface** characteristics (as described above)
- ✓**spatial** context (in all three dimensions, i.e. size, shape, composition of objects, which at the smallest scale are pixels or voxels)
- ✓**temporal** context (the change of surface characteristics and/or spatial context over time)

… and Use

to obtain **Land Use** further dimensions could be added:

- ✓ non-surface related properties (population, traffic, sealing, …)
- ✓ immaterial or unobservable properties (administrative affiliation, origin(e.g. ‘artificial’), potential(e.g. arable), etc.)

BUT: the different schemes must remain strictly separated!

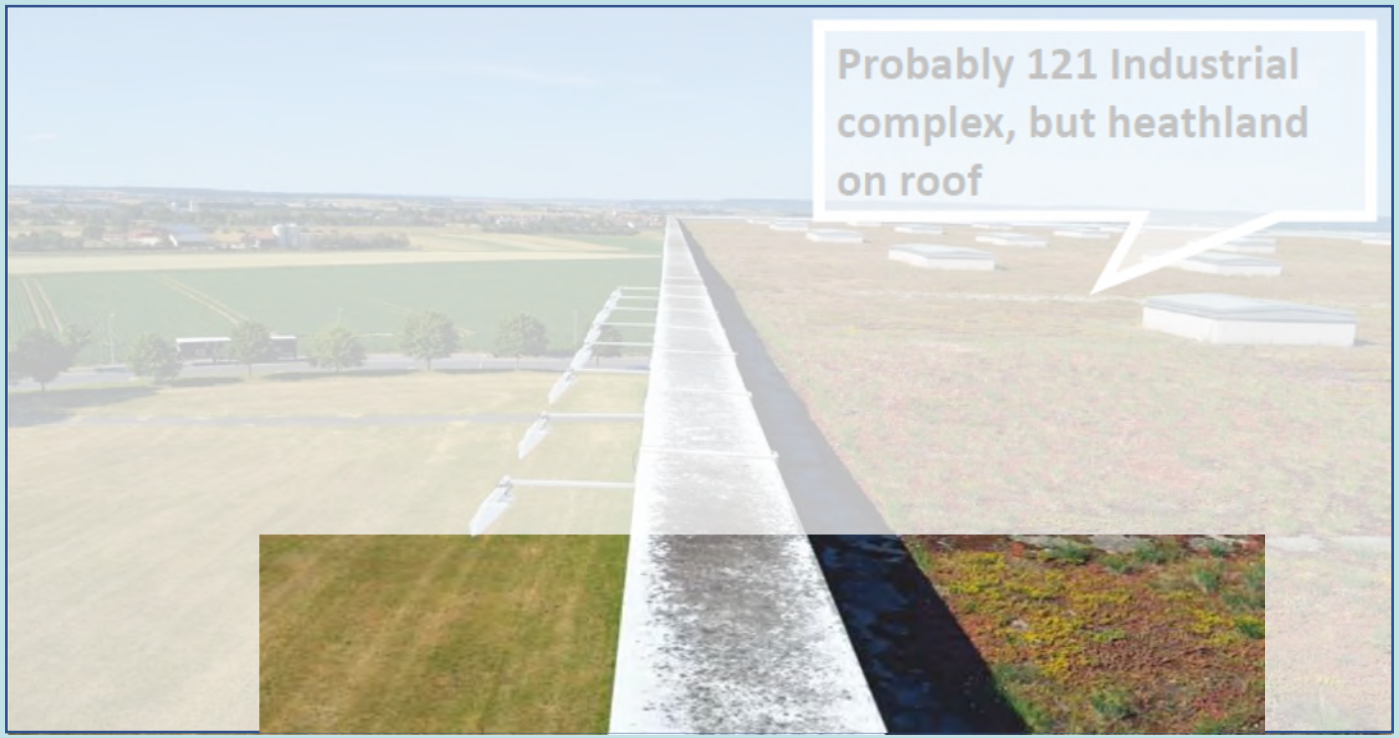
Objects

are inevitable wherever **shape** or **size** are part of the definition of a LC or LU class, they can either be:

- intrinsic: defined by their content and therefore dynamic
- external: determined by an additional data source

And they will always render the scheme scale-dependent!

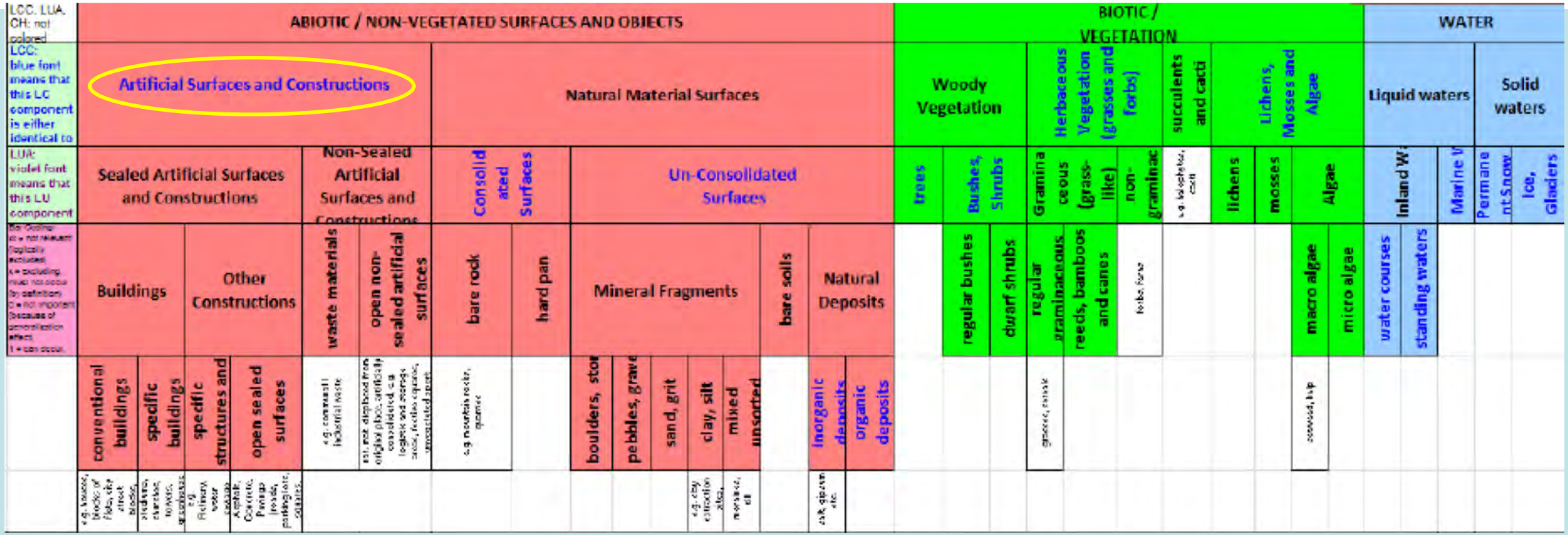
Food for discussion



What is ”land”, and what is “cover”?

Adapted from “EAGLE Matrix and Data Model”, S. Arnold, EEA

Food for discussion



EAGLE: What is ”artificial”, and what is “construction”?

Adapted from “EAGLE Matrix and Data Model”, S. Arnold, EEA

Any questions?

You can find me at @username & user@mail.me