Navigating connected data & models through semantics for SEEA

Stefano Balbi, Ken Bagstad, Alessio Bulckaen, Marcel Buchhorn, Bruno Smets, Marc Paganini, Ferdinando Villa

London Group, October 3, 2024

esa

vito remote sensing

bc3 BASQUE CENTRE FOR CLIMATE CHANGE Klima Aldaketa Ikergai Sustainability, that's it!

USGS science for a changing world

Ecounting

UNIVERSITY OF PATRAS ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

https://esa-people-ea.org/en

The challenge

- Global implementation of SEEA EA & spatially explicit SEEA CF
- Incomplete integration of geospatial/Earth Observation (EO) & statistical data
- INCA models meet initial EU SEEA EA needs, but would benefit from:
  - Reduced latency/improved quality offered by EO
  - Better bridging of EU-wide & member state data & models
  - Applications beyond the EU to support global implementation

Wish list based on our experience

1. Flexibly incorporate best-available knowledge
   - From global public to user-provided data
2. Rescale smartly across spatial & temporal scales
   - From local to global & vice versa
3. Open, transparent, well-documented models
   - Simple, modular coding, mandatory encapsulated documentation
   - Individually documented models & computational workflows for collection & processing
4. Data & models alive on the web
   - Non-static
   - Computational workflows for account readiness
   - DOI, peer reviewed, metadata, etc.

PEOPLE-EA components

ARIES-OpenEO-UDP/UDF catalog
(public repository to access INCA functionality & advanced OpenEO functionality)

Access via k.LAB client API in UDF

OpenEO adapter

models & methods

data pre-processing

Data and Processing Engine

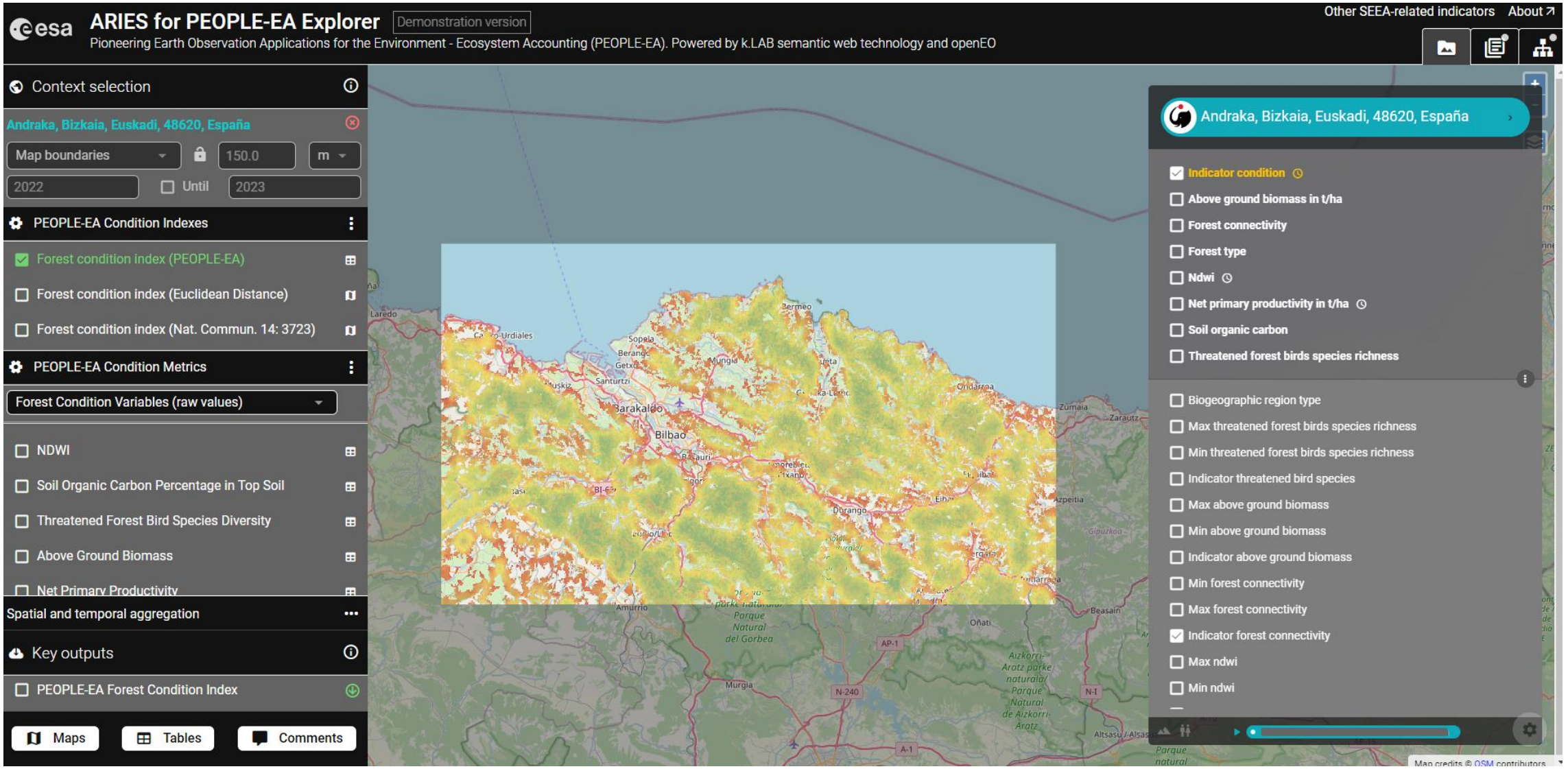
k.LAB
Knowledge Integrated Modelling
SEEA

Access via k.LAB client API

Semantic front-end

NCA standard for EU

What you see


esa
ARIES for PEOPLE-EA Explorer
Demonstration version
Pioneering Earth Observation Applications for the Environment - Ecosystem Accounting (PEOPLE-EA). Powered by k.LAB semantic web technology and openEO
Other SEEA-related indicators
About
Context selection
Andraka, Bizkaia, Euskadi, 48620, España
Map boundaries
150.0
m
2022
Until
2023
PEOPLE-EA Condition Indexes
Forest condition index (PEOPLE-EA)
Forest condition index (Euclidean Distance)
Forest condition index (Nat. Commun. 14: 3723)
PEOPLE-EA Condition Metrics
Forest Condition Variables (raw values)
NDWI
Soil Organic Carbon Percentage in Top Soil
Threatened Forest Bird Species Diversity
Above Ground Biomass
Net Primary Productivity
Spatial and temporal aggregation
Key outputs
PEOPLE-EA Forest Condition Index
Maps
Tables
Comments
Andraka, Bizkaia, Euskadi, 48620, España
Indicator condition
Above ground biomass in t/ha
Forest connectivity
Forest type
Ndwi
Net primary productivity in t/ha
Soil organic carbon
Threatened forest birds species richness
Biogeographic region type
Max threatened forest birds species richness
Min threatened forest birds species richness
Indicator threatened bird species
Max above ground biomass
Min above ground biomass
Indicator above ground biomass
Min forest connectivity
Max forest connectivity
Indicator forest connectivity
Max ndwi
Min ndwi
Bilbao
Barakaldo
Durango
Laredo
Mungia
Amurrio
Murgia
Oñati
Beasain
Zumaia
Zarautz
Azpeitia
Parque Natural del Gorbea
Aizkorri-Aratz parke naturala/ Parque Natural de Aizkorri-Aratz
AP-1
N-240
N-1
A-1
Map credits © OSM contributors

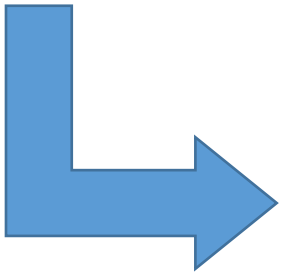
Behind the curtains

1. EO retrieval & processing (e.g., Sentinel, Landsat)
2. Data & models used on the fly on the web
3. Distributed system, two different computational architectures:
   - OpenEO - Terrascope
   - ARIES for SEEA Sector Hub of the UN Global Platform
     - Link to online data services, including OGC-compliant & STAC catalogues
4. Open source, openly documented code
5. API Networked Economy (virtual credits): single authentication point

Machine actionable, semantically annotated data & models

- Resources (data & models) & semantics clearly delineated:
  - Resources: online, unique URNs, metadata
  - Semantics: orchestrator to connect the pieces

```
/**
 * Global UDP parameters can be set in the URN
 */
model 'local:ferdinando.villa:im.openeo.sandbox:vito.corine.arablemask#output_warp=true'
    as presence of landcover:ArableLand earth:Region;
```



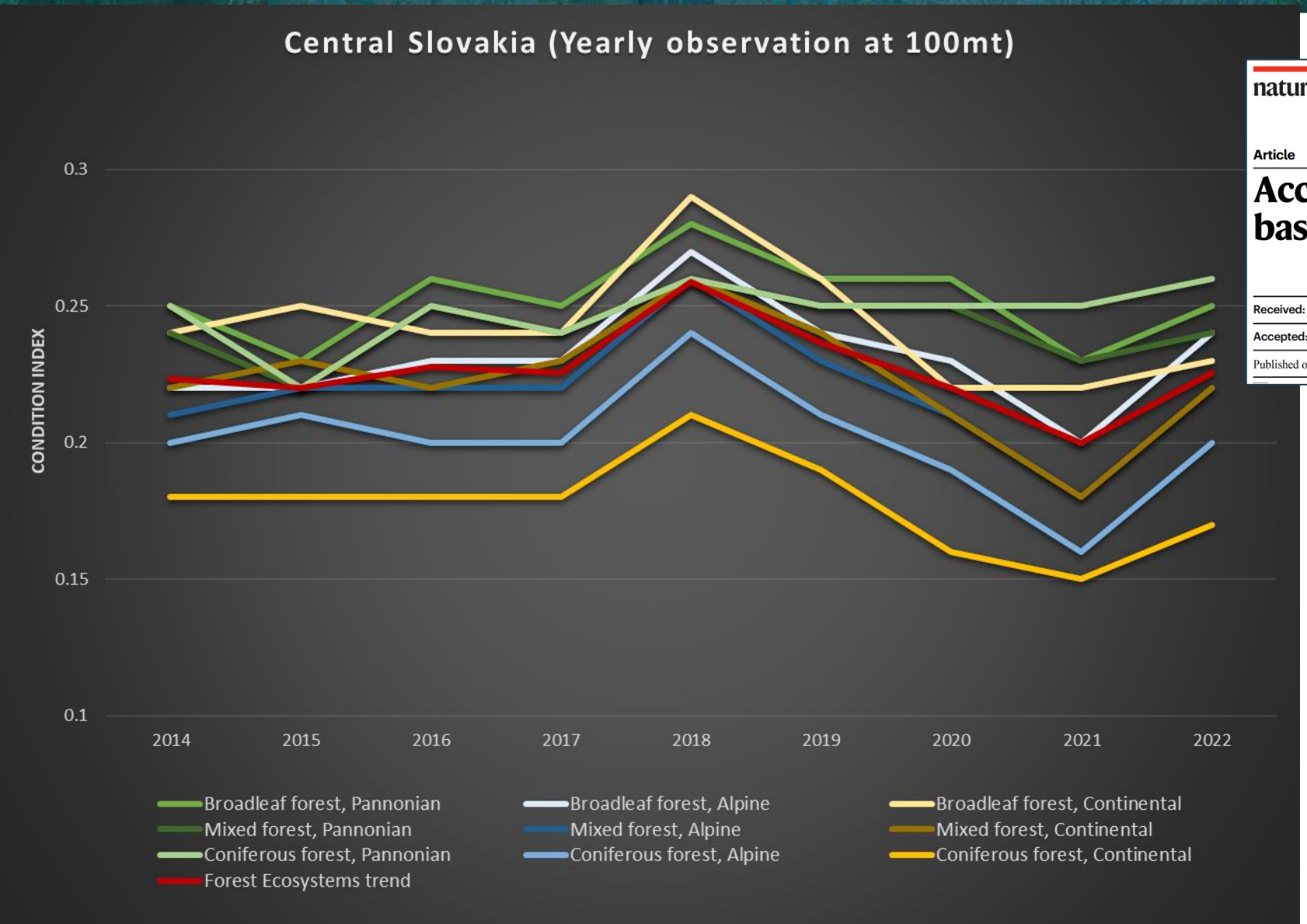
"observing" the **presence of ArableLand Region** concept triggers execution of UDP/UDF on OpenEO server

Smart resource prioritization:

Soil erosion control accounting, Central Slovakia, 2020 (KT soil retention)

Land cover/ecosystem type	Global C-factor	INCA/OpenEO C-factor	% difference
Agricultural land with natural vegetation*	3,079.0	5,444.1	76.8%
Annual cropland associated with permanent*	0.4	0.4	0%
Bare rock**	0	2.8	-
Broadleaf forest**	30,635.3	30,011.7	-2.0%
Burned land**	0	60.8	-
Complex cultivation patterned land*	430.8	710.3	64.9%
Coniferous forest**	34,666.5	33,965.6	-2.0%
…	…	…	
Transitional woodland scrub**	7,868.0	9,552.0	21.4%
Vineyard*	34.3	15.6	-54.5%
Total	**114,245.6**	**121,832.7**	**6.6%**
All agricultural ecosystem types	**12,248.1**	**20,016.7**	**63.4%**
All non-agricultural ecosystem types	**101,997.5**	**101,816.0**	**-0.2%**

Annual, low-latency accounting:
EO for SEEA EA Forest Condition Accounting



nature communications

Article

https://doi.org/10.1038/s41467-023-39434-0

Accounting for forest condition in Europe based on an international statistical standard

Received: 30 September 2022

Accepted: 12 June 2023

Published online: 22 June 2023

Joachim Maes[1,2], Adrián G. Bruzón[3], José I. Barredo[2], Sara Vallecillo[2], Peter Vogt[2], Inés Marí Rivero[2] & Fernando Santos-Martín[3]

Covering 35% of Europe's land area, forest ecosystems play a crucial role in

Questions for the London Group

1. What barriers exist to using EO & interoperable data/models by the SEEA community? Could their wider use imply major changes in production of accounts? Is the statistical community ready for such changes? If not, how could its readiness be improved?
2. How can we make this work more accessible to & widely used by NSOs? E.g., demonstration projects:
    a. Linking public & private data while ensuring compliance with data privacy requirements
    b. Quantifying improvements in accounts, e.g.:
        i. Reduced latency
        ii. Incorporation of higher-quality data & models
        iii. Reuse of data & models from one country that improve accounts in another - improving global SEEA uptake
    c. Other ideas?

Final considerations

- Time is ripe to build & maintain a common knowledge base for **SEEA & beyond***
  - Good practices, Standards, Datasets, Algorithms, Protocols, Platform APIs
- Precondition for deep integration & interoperability
  - Beyond (just) “open science”
  - Online resources independent of their semantic orchestration
  - Distributed, autonomously produced, peer-reviewed scientific products on the web, especially geospatial & EO products

*ecosystem service assessments, biodiversity monitoring, other global reporting frameworks

Video: https://www.youtube.com/watch?v=fvChjWO5IN8
Documentation: https://confluence.integratedmodelling.org/display/AFP
(register here: https://integratedmodelling.org/hub)