DEMIX

Digital Elevation Model Intercomparsion eXercise - overview and results

Peter Strobl for the DEMIX team

EC-JRC

VH-RODA, ESA/ESRIN, Italy

29th November 2023

DEMIX origins from the JRC DEM benchmarking workshop (Jan 2019):

- **new data sets** are coming up (“Copernicus DEM“), which might change the DEM ‘landscape‘
- EO platforms and ‘data cubes’ make data increasingly available also at **continental to global scales**
- literature is rich in **DEM validation and comparisons** of (almost) everything with everything else in many different places
- methodologies vary and results are not always **representative or comparable** between studies and locations

➢ a coordinated approach is desirable!

➢ bring CEOS TMSG and the International Society for Geomorphometry (ISG) together!

The Committee on Earth Observation Satellites (CEOS) was established in 1984 under aegis of the G7 Economic Summit of Industrial Nations Working Group on Growth, Technology, and Employment

Now in its fourth decade, CEOS comprises

- **34 Members**
  (*Space Agencies*) and
- **29 Associates**
  (*UN Agencies, Phase A programmes or supporting ground facility programmes*)

All of whom contribute to CEOS on a 'best efforts' and voluntary basis.

Mission: CEOS ensures international coordination of civil space-based Earth observation programs and promotes exchange of data to optimize societal benefit and inform decision making for securing a prosperous and sustainable future for humankind.

CEOS WGCV mandate for DEMIX:

- perform a state-of-the-art comparison of the major global (free&open) DEMs
- provide recommendations on best available DEM options depending on domain and area to allow informed choices

decision in 2020

Expected Outcomes

- Consistent and comprehensive DEM definitions and terminology (t)
- Base (t) and extended (g) set of benchmarking metrics and respective algorithms (t) and open source tools (g)
- Detailed comparison results on test areas (t) and aggregated wall to wall benchmarking results (g)
- Recommendations regarding reference DEMs (t) and consistent orthoimage (g)
- Final report (t) and peer-reviewed publication (g)

(t) threshold; (g) goal

- The CEOS Working Group on Calibration and Validation (WGCV) dedicates itself to ensure long-term confidence in the accuracy and quality of Earth Observation data and products and is the forum for exchange of information on calibration, validation, and associated cooperative activities.

- The WGCV has six Subgroups that operate as individual entities and focus on specific technical areas related to calibration and validation:
  - **Atmospheric Composition** (ACSG)
  - **Land Product Validation** (LPV)
  - **Infrared Visible Optical Sensors** (IVOS)
  - **Microwave Sensors** (MSSG)
  - **Synthetic Aperture Radar** (SAR)
  - **Terrain Mapping** (TMSG)

2023-2024 WGCV Chair
Philippe Goryl

2023-2024 WGCV Vice-Chair
Cody Anderson

https://ceos.org/ourwork/workinggroups/wgcv/

Status of the CEOS/WGCV Terrain Mapping SubGroup (TMSG)

- Re-activated early 2020
- as of Nov 2023:
  - ~60 subscriptions
  - 15 countries
  - ~50% with CEOS background
  - ~30% Geomorphometry.org
  - ~35 expressed interest in the intercomparison exercise DEMIX (incl. industry!)
- main activity so far is DEMIX, but more is coming
- DEMIX workshop & TMSG plenary held on 12/13 July 2023 at Geomorphometry23 in Iasi, Romania

Geomorphometry23 web page: https://geomorphometry.org/geomorphometry-2023/
TMSG Subscription page: https://ec.europa.eu/eusurvey/runner/WGCV-TMSG_membership

CEOS WGCV mandate for DEMIX:

- perform a **state-of-the-art comparison** of the major global (free&open) DEMs
- provide **recommendations** on best available DEM options depending on domain and area to allow informed choices

decision in 2020

Expected Outcomes

- Consistent and comprehensive DEM definitions and terminology (t)
- Base (t) and extended (g) set of benchmarking metrics and respective algorithms (t) and open source tools (g)
- Detailed comparison results on test areas (t) and aggregated wall to wall benchmarking results (g)
- Recommendations regarding reference DEMs (t) and consistent orthoimage (g)
- Final report (t) and peer-reviewed publication (g)

(t) threshold; (g) goal

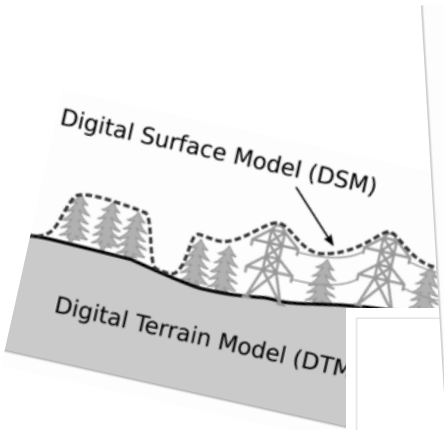
after:

- 3 years,
- 3 plenaries,
- wrangling down Teams
- 3 subgroups, each with 5-15 active members,
- 130+ subgroup meetings, each with at least 4 participants
- a conference paper and video,
- 2+ peer-reviewed publications,
- a new ‘DEMIX tiling’ system,
- a processing platform, …

we are almost there…

- Revised terminology and comprehensive definitions (glossary) finished
- Peer reviewed paper published:

[Guth et. al. 2021](https://doi.org/10.3390/rs13183581)

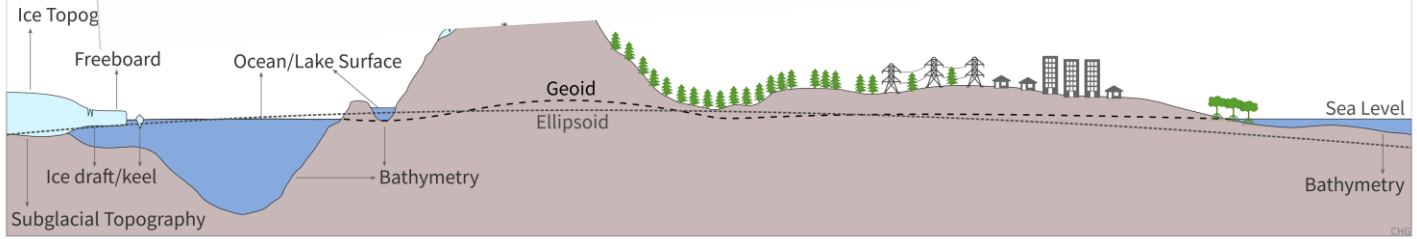


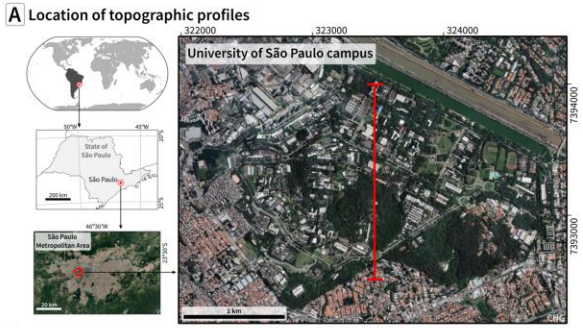
remote sensing
https://doi.org/10.3390/rs13183581
MDPI

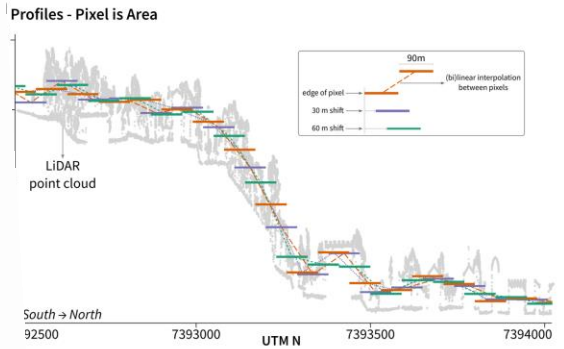
Artide

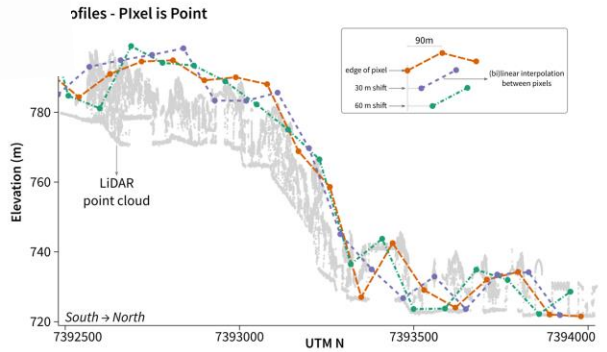
Digital Elevation Models: Terminology and Definitions

Peter L. Guth [1,*], Adriaan Van Niekerk [2], Carlos H. Grohmann [3], Jan-Peter Muller [4], Laurence Hawker [5], Igor V. Florinsky [6], Dean Gesch [7], Hannes I. Reuter [8], Virginia Herrera-Cruz [9], Serge Riazanoff [10], Carlos López-Vázquez [11], Claudia C. Carabajal [12], Clément Albinet [13] and Peter Strobl [14]









Before talking about quality we must define criteria and metrics which characterize a DEM and which can be used to compare them

Three main groups were identified:

1. Linear difference (or error) measures such as RMSE, LE90, CE90, Median and normalized median absolute difference (NMAD),
   separately for horizontal displacement and (vertical) elevation difference,
   distinguished by e.g. slope, land cover, and not generalized spatially over more than 10^6-10^7 values
2. Morphological descriptors, e.g. slope, aspect, roughness. Complex morphological metrics like number of peaks and pits, length of ridges and troughs, number of outliers (spikes), consistency of stream networks
3. Other: Autocorrelation length, SNR

+non-quantitative*:

Completeness, availability and reliability of Metadata, visual appearance …

*https://zenodo.org/records/8030735

- Intercomparison is only useful for (non-expert) users if in the end they get a **recommendation** (or **ranking)**:

  Are there significant differences and if so
  which options are better and which are worse?

- Not every user might want to apply the same criteria and even the same criteria could result in different rankings depending e.g. on location

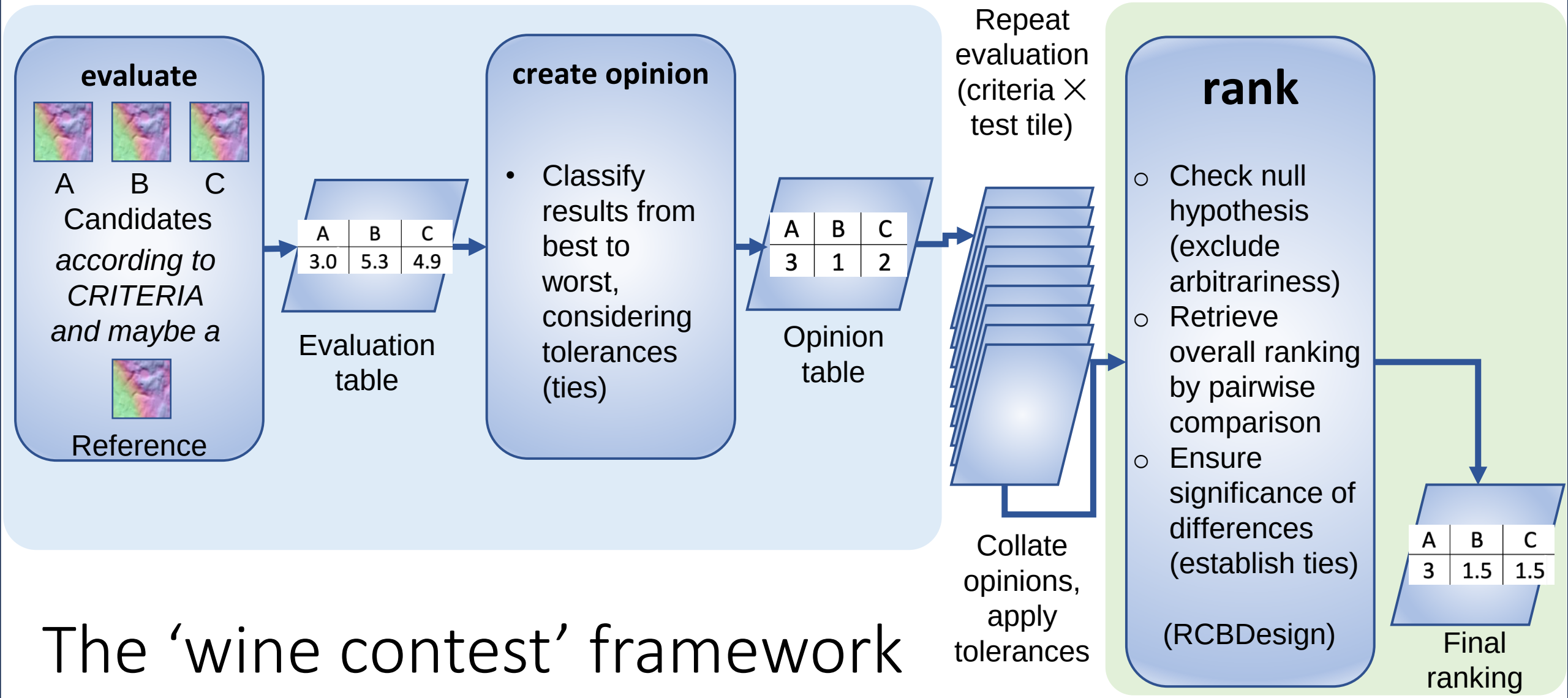
  We are less interested in an 'overall winner',
  but the best solution in a given context!

➢ We need a configurable, re-usable and expandible test environment!

DEM ranking

- A major challenge is how to combine different criteria and metrics to arrive at an overall ranking (if justifyable) of different tested DEMs
- Statisticians use the ‘wine contest’ method (or RCBD design) in which k different wines (candidate DEMs) are assessed according to C different criteria (metrics) by N different judges (test sites)
- allows
  - establishing of an overall ranking
  - testing the significance of the differences
  - adding, removing, and filtering candidates, metrics, and test sites

The ‘wine contest’ framework



Scope and Products to be included: All datasets which have an at least continental coverage and are available under a free & open data policy, including:

- **SRTM** (v3 NASA/CGIAR, the de-facto reference for more than a decade)
- **NASADEM** (NASA, JPL, most recent reprocessing of the SRTM product line)
- **ASTER-GDEM v3**, (METI, NASA, from ASTER stereo imagery)
- **ALOS World 3D 'AW3D30'** (JAXA, based on the PRISM stereo scanner)
- **Copernicus DEM GLO30 'COPDEM'** (EC/ESA, f&o version of WorldDEM™ procured by Airbus, the commercial version of DLR's TanDEM-X mission)
- **FABDEM**, (L. Hawker et al., UoB), DTM based on COPDEM

WORKFLOW

MICRODEM

Step 1 — Obtain high quality source elevation data

- **Obtain and mosaic DEMs covering test area**
  - 6 candidate global DEMs, 1", geographic
  - Source reference DTM: 1-5m UTM projection
  - Source reference DSM: if available

Step 2 — Prepare reference DEMs from the source data

- **Aggregate reference DEMs to 1" spacing**
  - Pixel-is-point
  - Pixel-is-area
  - For high lat COPDEM, separate Pixel-is-point
  - For high lat ALOS, separate Pixel-is-area
- **Convert to WGS84/EGM2008 if needed**
  - Convert to ellipsoid, then to geoid
  - Use grids from NGA via PROJ and national/regional mapping agency
  - Use GDAL for areas in the USA

Step 3 — Evaluate the reference and test DEMs

- **Process test areas and create GIS database**
  - Compute test area statistics
  - Create difference outputs
  - Compute metrics
  - Produce Evaluations table with initial tolerances

Jupyter Notebook

Step 4 — Rank the global DEMs

- **Produce Wine Contest**
  - Produce Opinions table
  - Compute confidence levels
  - Produce final rankings
  - Adjust tolerances if required
  - Filter database
  - Produce graphics

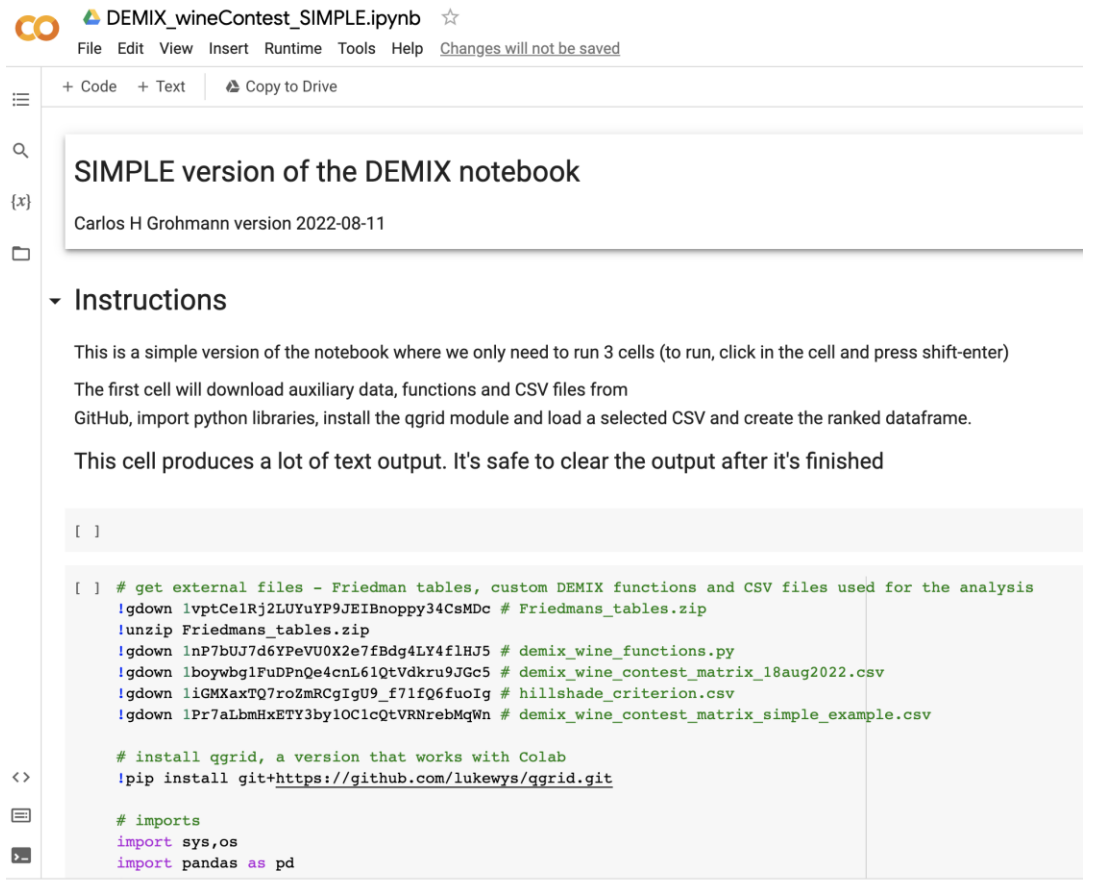
Courtesy: C. Grohman

The DEMIX wine contest:

- 24 test areas, 236 DEMIX tiles (10x10km^2) on four different continents with reference data
- Reference data preparation tool
- All major geomorphological landforms and landcover types represented incl. coastal areas (partial water)
- 15 different criteria in 3 classes
- Pixel by pixel comparison against reference data
- >55.000 individual test scenarios (rows in opinions database)

Courtesy: C. Grohman

- Idea is to have a simple to use ranking tool allowing users to filter for
  - Criteria type (elevation, slope, …)
  - Spatial characteristics (geomorphology, landcov,
  - Reference (DTM or DSM)
- Ranking is recomputed according to user's purpose
- Python notebook as base (portable and cloud compatible)
- For details see: https://zenodo.org/records/7779256



AWARDS

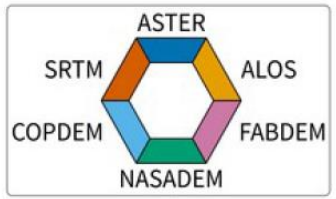
COPDEM | Overall best DSM

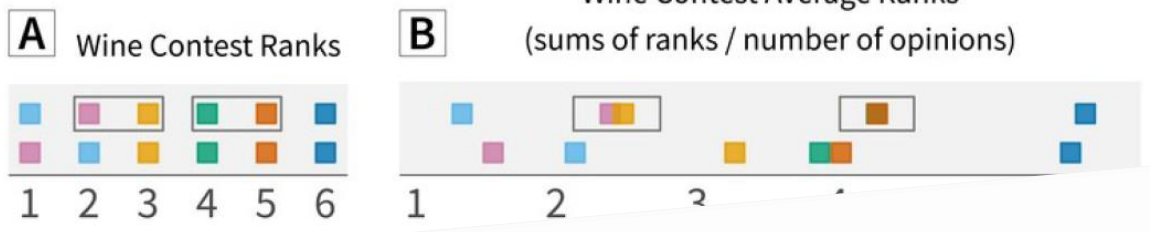
FABDEM | Best DTM (except for steep terrain)

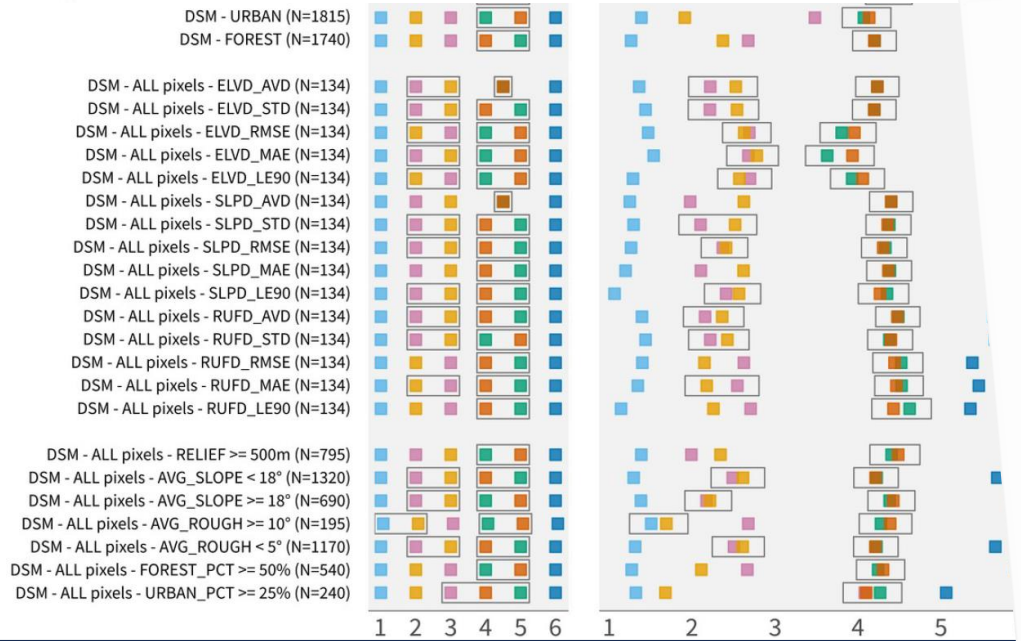
ALOS AW3D30 | Sometimes 2nd place might be better in steep terrain than FABDEM

Courtesy: C. Grohman

Wine contest implemented and comprehensive intercomparison undertaken
Peer reviewed paper submitted: Bielski et al 2023







IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, VOL. XX, 2023

Novel approach for ranking DEMs: Copernicus DEMs improve open global topography

Conrad Bielski *Member, IEEE*, Carlos López-Vázquez *Senior Member, IEEE*, Carlos H. Grohmann *Member, IEEE*, Peter L. Guth, Laurence Hawker, Dean Gesch, Sebastiano Trevisani, Virginia Herrera-Cruz, Serge Riazanoff, Hannes I. Reuter, and Peter Strobl

https://arxiv.org/abs/2302.08425

***Abstract*—We present a practical approach to inter-compare a range of candidate digital elevation models (DEMs) based on pre-defined criteria and statistically sound ranking approach. The presented approach integrates the randomized complete block design (RCBD) into a novel framework for DEMs comparison. The method presented provides a flexible, statistically sound and customizable tool for evaluating the quality of any raster - in this case a DEM - by means of a ranking approach, which takes into account a confidence level, and can use both quantitative and qualitative criteria. The users can design their own criteria for the quality evaluation in relation to their specific needs. The application of the RCBD method to rank six 1" global DEMs, considering a wide set of study sites, covering different morphological and landcover settings, highlights the potentialities of the approach. We used a suite of criteria relating to the differences in the elevation, slope, and roughness distributions compared to reference DEMs aggregated from 1-5 m lidar-**

OVER the past two decades, several Earth observation missions have resulted in finer than 100 m resolution global digital elevation models (DEMs), most of which are shared freely and openly worldwide. These data revolutionized earth sciences and spurred many applications that require accurate information about the shape of Earth's surface. Consequently, the demand for high quality DEMs across multiple disciplines continues to grow and users around the world are faced with the challenge to understand the DEMs' intrinsic characteristics and to make an informed choice for their particular application.

At this time, at least six different global medium resolution (i.e., 10-100 m) DEMs have been produced using a variety of techniques. We expect more to come in the near future

Food for discussion

- Lessons learned from DEMIX?
  - ❑Clarifying terminology and concepts is worth the time, but we have just begun!
  - ❑Outside-CEOS partners are indispensable, as are sponsors!
  - ❑Careless use of grid conventions, metadata, and vertical datums are a real nightmare!

Diversity (often) kills interoperability!

- ‘Wine contest’ for everyone?
  - (Inter-)Comparisons are academic without ranking!
  - Ranking is only sound if based on proper statistics!
  - **BUT**, is ‘wine contest’ an appropriate name after all?

More food…

- Geolocation is a pre-requisite for spatial data interoperability
- DEMs are a key input for geolocating any non-nadir remote observation
- Co-registration (precision) is more important than absolute accuracy
- Consistency requires global references, in x, y, and z

- Ideas for future activities:
  - DEMIX reloaded: more criteria, more reference tiles, fully in the cloud
  - GCPIX: intercomparison of GCP libraries
  - GDMIX: spatial matching and comparison of global GCPs with (shaded) DEMs

Thank you!

Big thanks to all active volunteers!
In particular the **sub-group leaders**:
Peter Guth, Carlos Grohman,
Conrad Bielski, Serge Riazanoff,
and **Carlos López-Vázquez**, the wine contest mastermind!

as well as ESA (Clement Albinet) and USGS (Dean Gesch) for their support!

any questions?

Peter.Strobl@ec.europa.eu

CEOS
Committee on
Earth Observation Satellites