GiGa infosystems

GeoHub: Sustainable Geomodelling

GST User Meeting 2022

Georg Semmler

10.03.2022

Problem description

Facts:

1. Geoscientists use digital models
2. Models getting larger and more complex every day

But:

- How was a certain model built?
- Which data was used?
- Can someone else built the same model given the same data?

Requirements

Understandable: Model construction must be comprehensible

Repeatable: Model construction must be repeatable based on the given data and description

Reproducible: A repeated model constructing needs to yield the same results

Existing Software

GST

- Database system to store 3d-geomodels
- Supports several geometry types
- Stores data
- Data visualization

→ **Stores data, but no construction workflows**

Existing Software

GOCAD

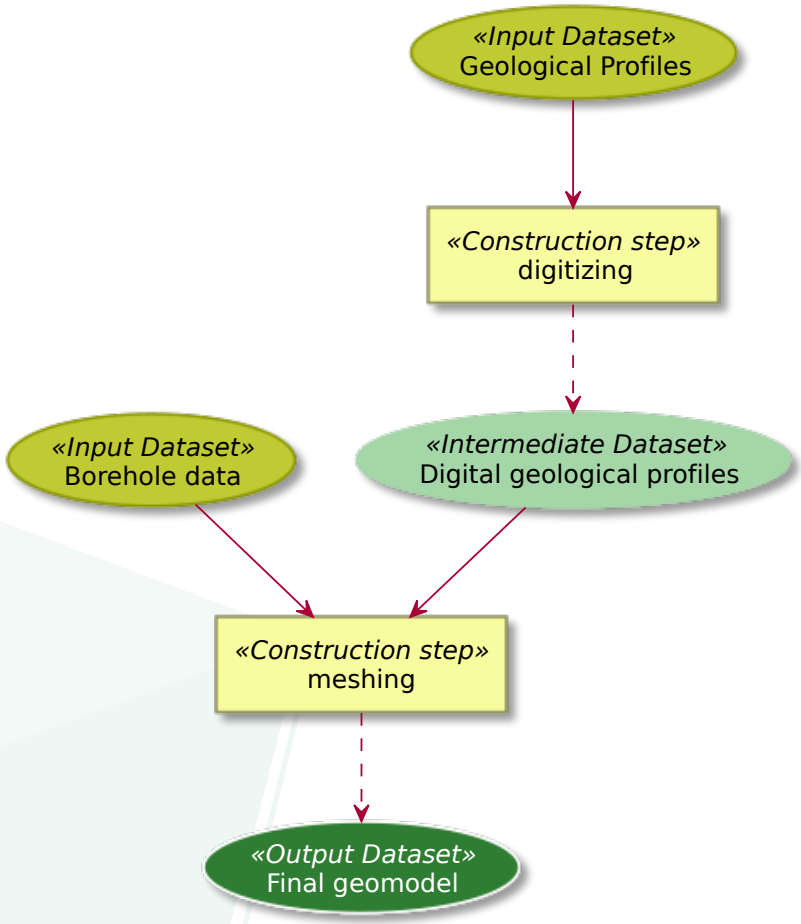
- Software suite to create geomodels
- Construction process is only partially **Understandable**
- Construction process is only manually **Repeatable**

→ **Creates models, but no construction processes**

Approach

We can represent model constructions as hypergraph:

- Nodes ≙ Datasets
- Hyperedges ≙ Transformations



Requirements

Understandable: Structure of the construction hypergraph helps to understand how datasets are combined

Repeatable: Computer executable construction steps are critical to repeat the computer aided construction at a later time

Reproducible: Comparison of two construction realisations allows to reason about reproducibility

What does equal mean?

Technical: The files are identical at bit/byte level

Structural: Both datasets contain the same parts

Mathematical: A distance between the two models is smaller than a given threshold

Geological: Both datasets support the same geological interpretation

$\rightarrow$ There is no approach that always fits.

Current state

Prototype:

- ✓ Frontend
- ✓ Backend
- ✓ Some executors

Exemplary Workflows:

- Geophysical data processing
- Subsurface 3D modelling
- Hydrological simulations

Further work needed:

- Improve user interface
- Integration with existing data storage solutions
- Parameterize construction steps

Workflow demonstration

Subsurface 3D modelling

Preprocessing ⟺ **Conversion** ⟺ **3D modelling**

- Extract data in specific areas
- Seperate GIS layers into different files

- Convert between different file formats
- Fix data inconsistencies

- Construct a 3D model in GoCAD
- Combine, project and mesh different datasets

[Datasets provided by LfULG Saxony]

Workflow demonstration

Automated data processing

Requirements:

1. Have access to the tools used in construction steps
2. Know what to do with data + tools
3. Know where to expect results

Solution:

1. Bundle tools in a docker image
2. Describe construction steps as script
3. Define which files are used as output dataset

Workflow demonstration

Workflow description:

```
image: gocad-docker-image
operation:
  - command: |
    cat > project_map.js << EOF
    skua.run('MapPointsOnPoints', {
     'targets':"/gobj:ts_DEM",
     'to_map':["/gobj:lin_Q"]
    });
    // export objects here
    EOF
    ./SKUA -run-js-script \
        project_map.js project.sprj
output:
  - file:  pl_lin_q.pl
```

Figure 1: Input

Figure 2: Output

Workflow demonstration

Current state

The construction process is:

✓ **Understandable:** via the construction hypergraph

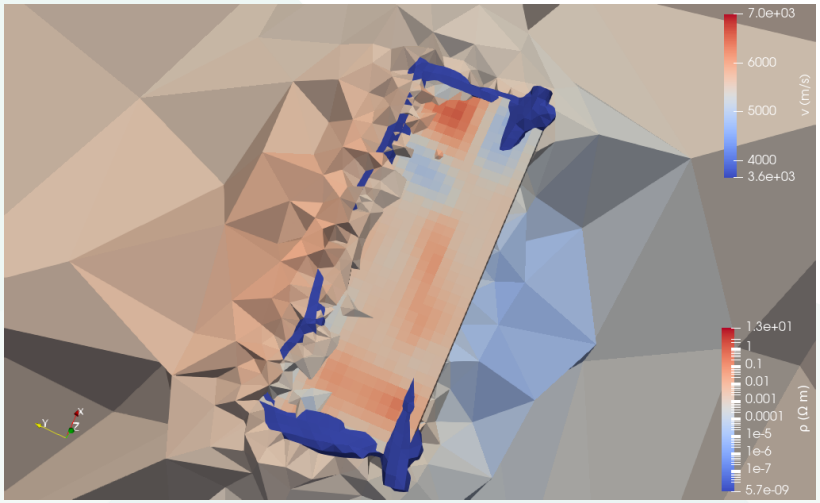
✓ **Repeatable:** by describing construction steps as scripts

[✓] **Reproducible:** for some causal definitions of equality[1]

[1] GoCAD operation may produce different results for triangle based operations

Workflow demonstration

Geophysical data processing

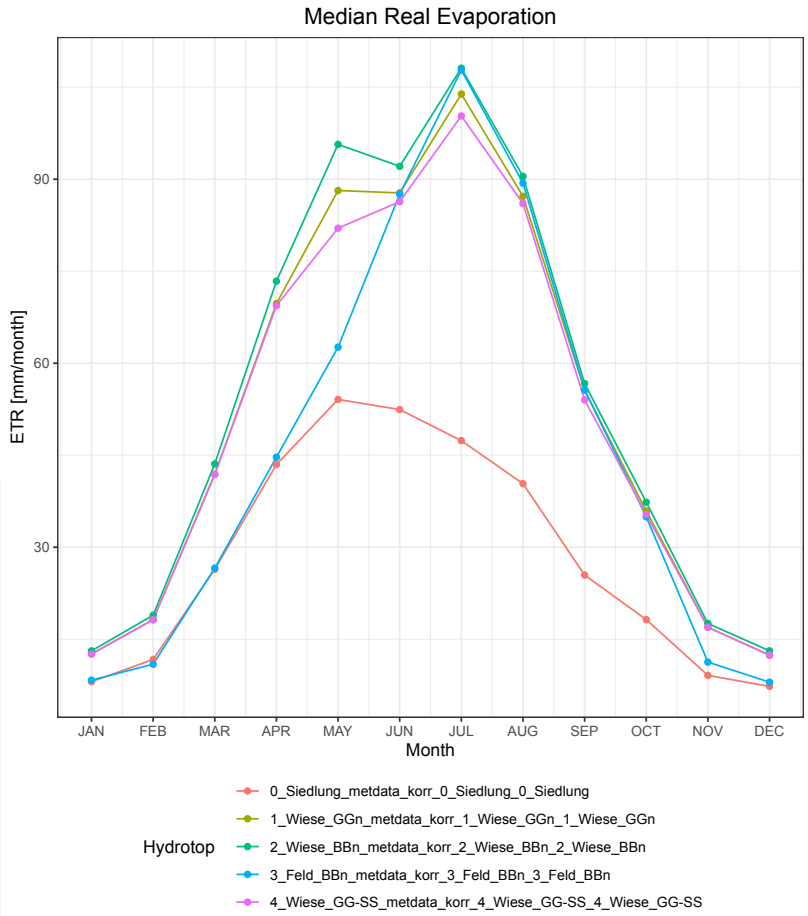


- Process geophysical measurement data
- Combined seismic + geoelectric survey
- Combines Matlab and Python scripts and data visualization
- Workflow fully reproducible using bit-to-bit equality

Workflow demonstration

Hydrological simulations

- Workflow to create a hydrological balance model
- Using Q-GIS, Python and R scripts and BOWAHALD for data processing
- Focus on workflow automation
- Reduces needed time to create a new model for a different study area from **approx. 1 week to < 1 day**



Summary

- Geohub makes model **constructions understandable**
- Construction processes are recorded as construction hypergraph
- Geohub is designed to **interact with existing software package**
- Any construction hypergraph stored in Geohub can be used to:
  - **Repeat** the same construction workflow with different data
  - **Verify** that a model construction can be reproduced later by others
  - Recognize problematic construction steps

Survey

Contact me via georg.semmler@giga-infosystems.com if you want to try it with your dataset or to discuss a possible use within your institute