Geo874 | FS25
University of Zürich

Databases in the Wild – The Swiss National Forest Inventory NFI

Rolf Meile

Eidg. Forschungsanstalt für Wald, Schnee und Landschaft (WSL)

Swiss Federal Institute for Forest, Snow and Landscape Research

Mission

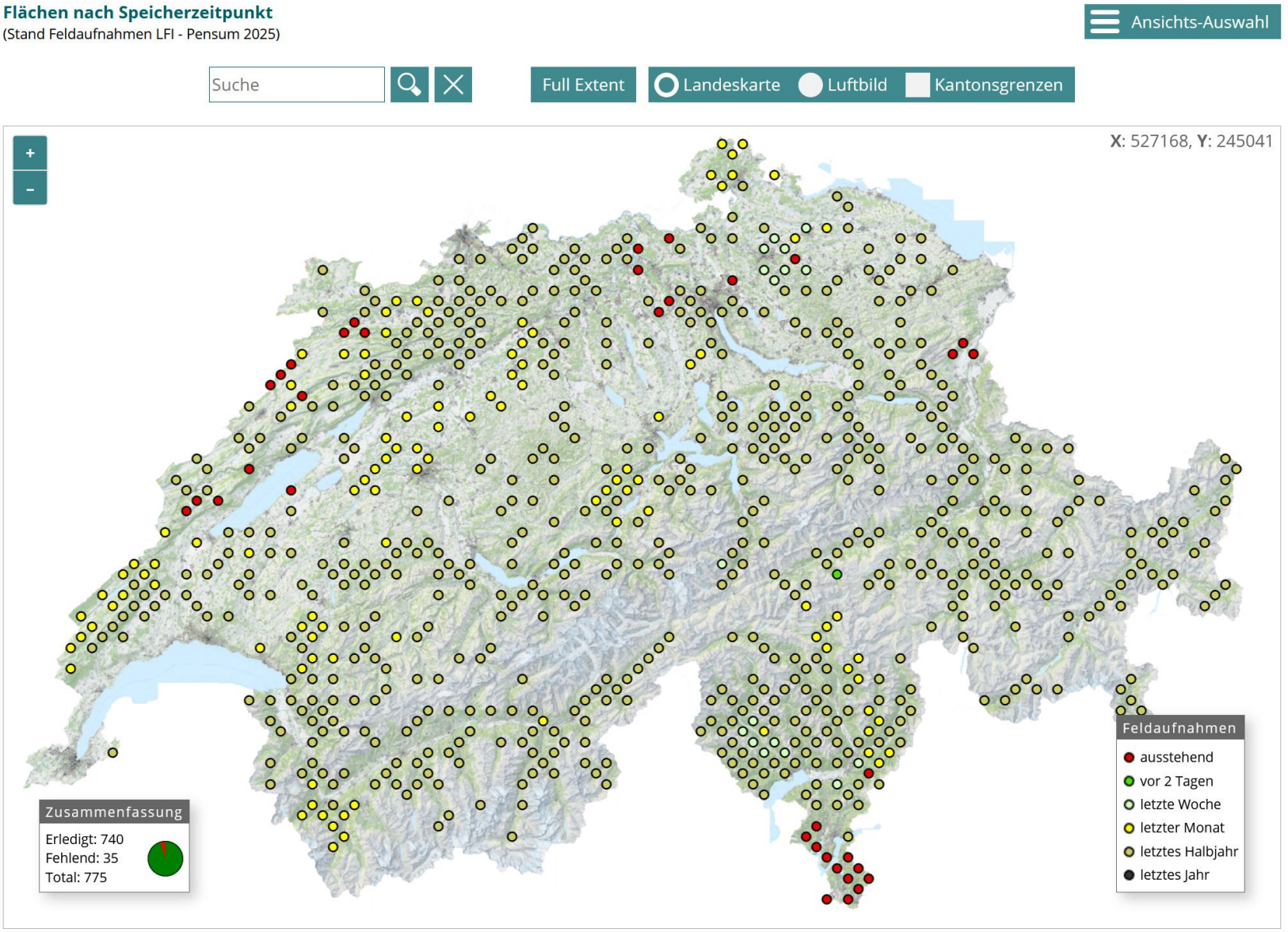
Overview

- Recording objective information at the national level about forests
- **Political**, social, economic and **research** purposes
- Data collection is mainly based on a **systematic sample inventory**
- Aim: Provide results for Switzerland as a whole and for larger regions and individual cantons
- Collaboration project of Federal Office of Environment FOEN and WSL

Topics

- Forest area, number of stems, **growing stock**, increment, fellings, biodiversity, protective forest quality, and **socio-cultural services** of the forest
- International climate and forest reporting (e.g. **greenhouse gas** inventory)
- Forest research in general as well as bringing together remote sensing data with field data (ground truth, calibrating models)

Systematic Sample...



In the Wild...

Field work

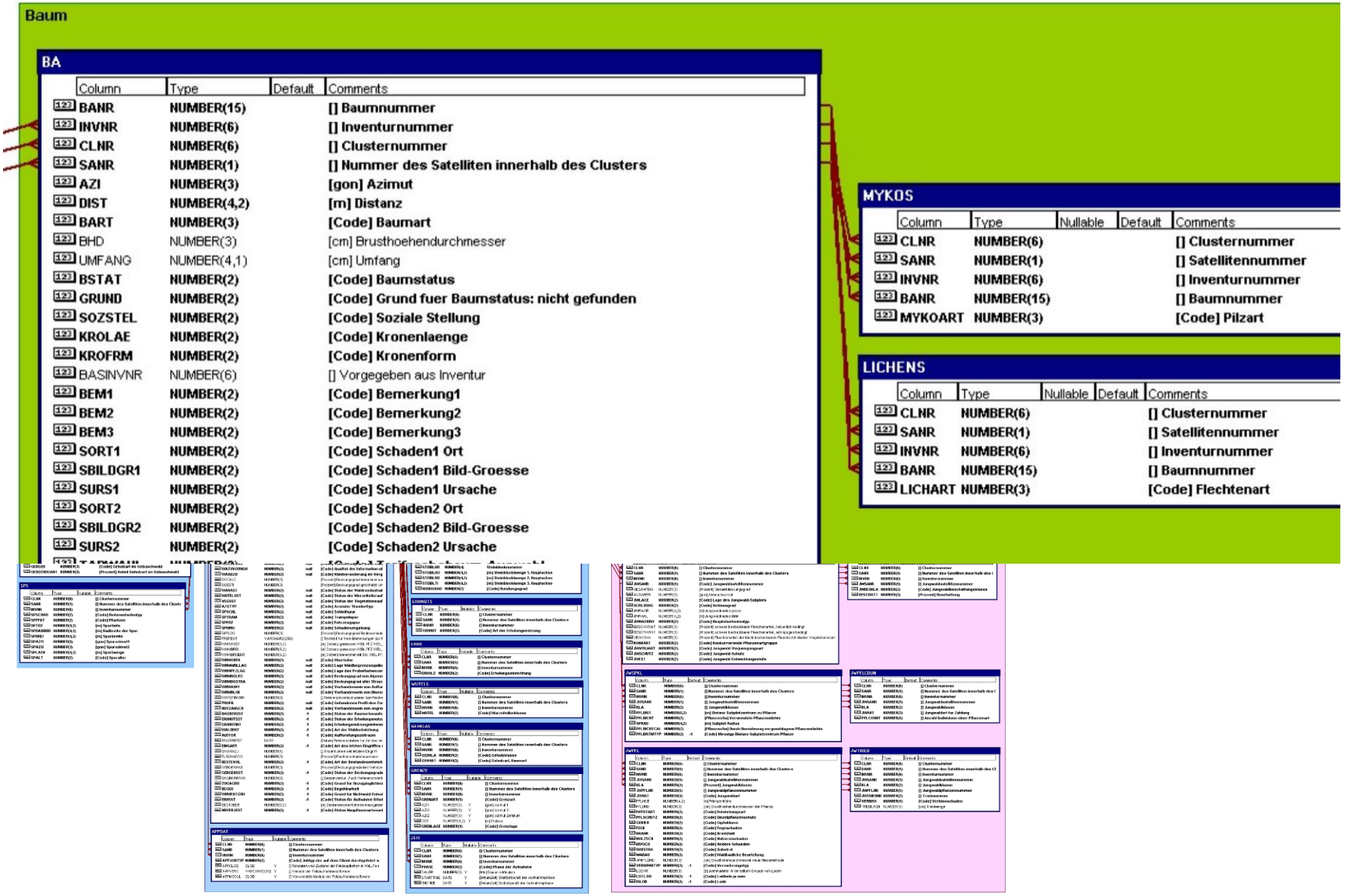
- 6'500 sites and 80'000 trees measured along with 300 attributes over 9 years
- Why? Get the base data for further processing.
- Ruggedized Tablets with specialized Java Application (business logic)
- Each plot 2-4 hours of work (includes plot searching)
- Field manual with rules and definitions for the field teams

The essence

- Data transfered to **raw database** at WSL;
- Formalized files/pics being uploaded; then decomposed and spread to the correct tables;
- All databases spatially enabled

➔ https://www.lfi.ch/en/forest-photos

In the Raw Database...



In the Office I ...

Deriving

- Raw data is just the beginning; we freeze their state and transform them in a formalized way to a **derived database**;
- Processes to do so are all documented and reproducible; Each value in each formalized table column is defined by a formula; the values can be recalculated at any time;
- The derivations (formulas) itself are stored in the database with helping **metadata**; automatic calculations possible;
- We maintain 2000 derivations sitting as columns and filling values on ~30 tables; examples: wood volume of a tree (column on the derivation-tree table), number giant trees on a plot (column on the derivation-plot table);
- **Dependencies** for each derivation are stored; this leads to a directed acyclic graph (DAG) pattern; the more complex a derivation/formula is the deeper goes the tree;

➔ https://www.lfi.ch/en/about-the-nfi/purpose-and-aims/movie

In the Derivation Database...

Tabelle SQL UPDATE:

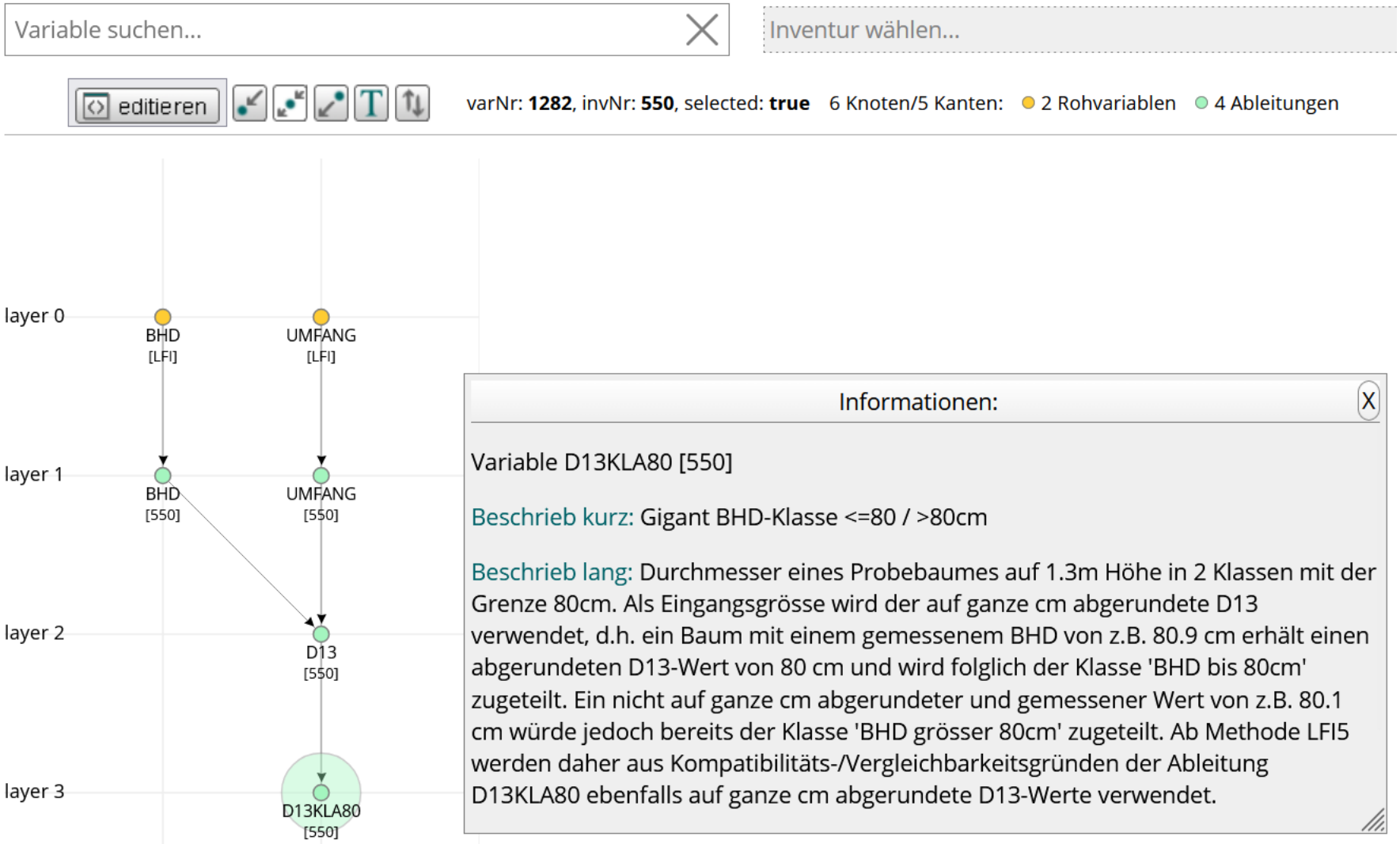
Ctrl+Z Rückgängig, Ctrl+Y Wiederholen, Ctrl+F: Suchen/Ersetzen, Ctrl+G: findNext, Ctrl+Shift+G: findPrevious, Alt+G: gotoLine, mehr

```
/*
Abhaengigkeit:
-------------------------------
Variable              Inventur
D13                   550
*/

UPDATE lfa.baa u
SET    u.D13KLA80 =
       (
       SELECT CASE WHEN a.D13 = -1 OR a.D13 IS NULL THEN -1
                   WHEN a.D13 > 0 AND a.D13 <= 80   THEN 1
                   WHEN a.D13 > 80                  THEN 2
              ELSE -1
              END
       FROM lfa.baa a
       WHERE a.invnr = 550
       AND   u.clnr  = a.clnr
       AND   u.invnr = a.invnr
       AND   u.banr  = a.banr
       AND   u.invnr = 550
       )
WHERE u.invnr = 550;
```

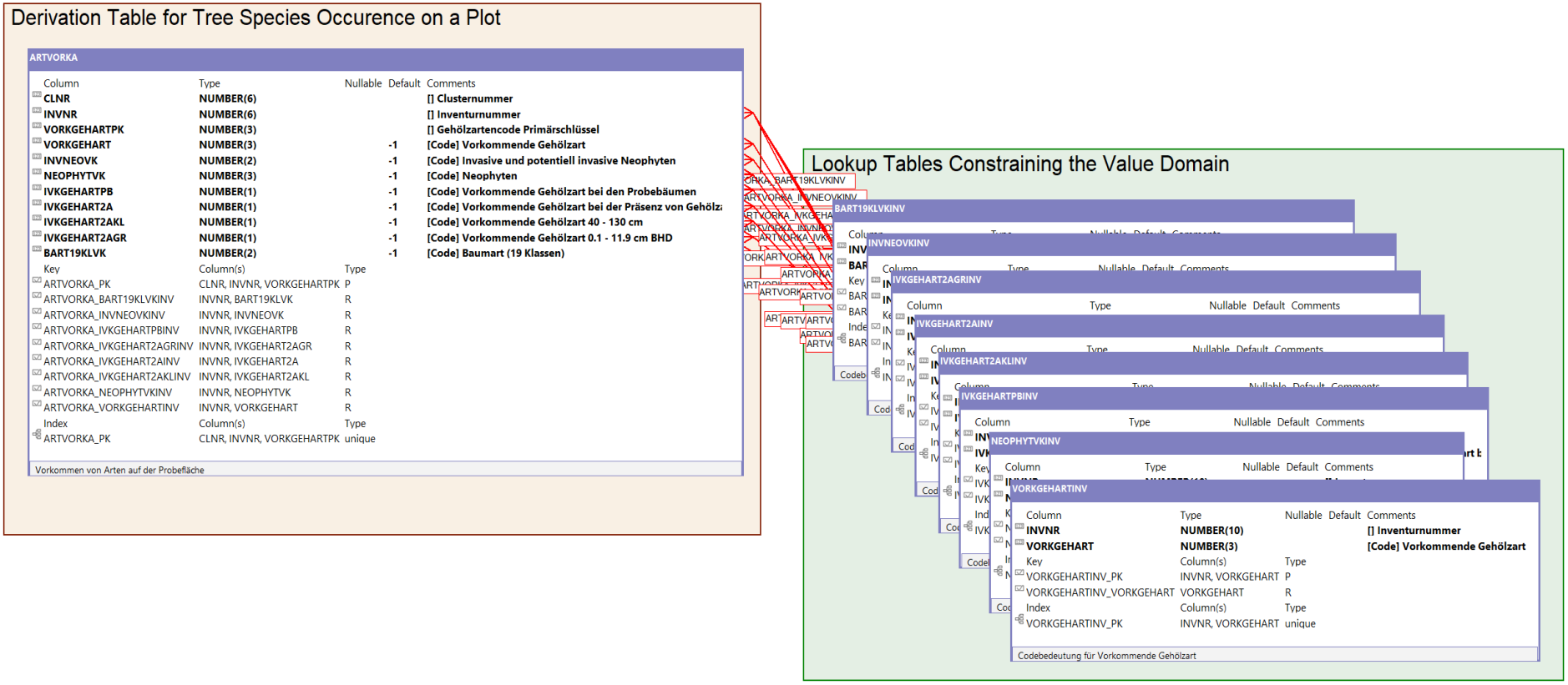
Derivation D13KLA80 – categorizes trees to be a giant tree or not

In the Derivation Database...



Derivation D13KLA80 is dependent of D13 and BHD and UMFANG (= calculation chain)

In the Derivation Database...



One derivation table with eight derivation columns; all columns constrained with a lookup table;

In the Office II ...

Statistics

- Only derived data and their specific metadata serves as a primary source for the statistical calculations;
- Analysis of the thematic and spatial characteristics of the forest as well as geostatistic;
- Aerea based results (cantons, production regions, ...) based on plot based information coming from systematic sampling in the field
- In addition: more dense grid with aerial photo interpretation for two phase sampling;
- Technology stack: database (oracle, postgresql), python server (flask, gunicorn), webserver (nginx), webapplication (apache, php)

Results of Statistics ...

definitions interpretation help download display map

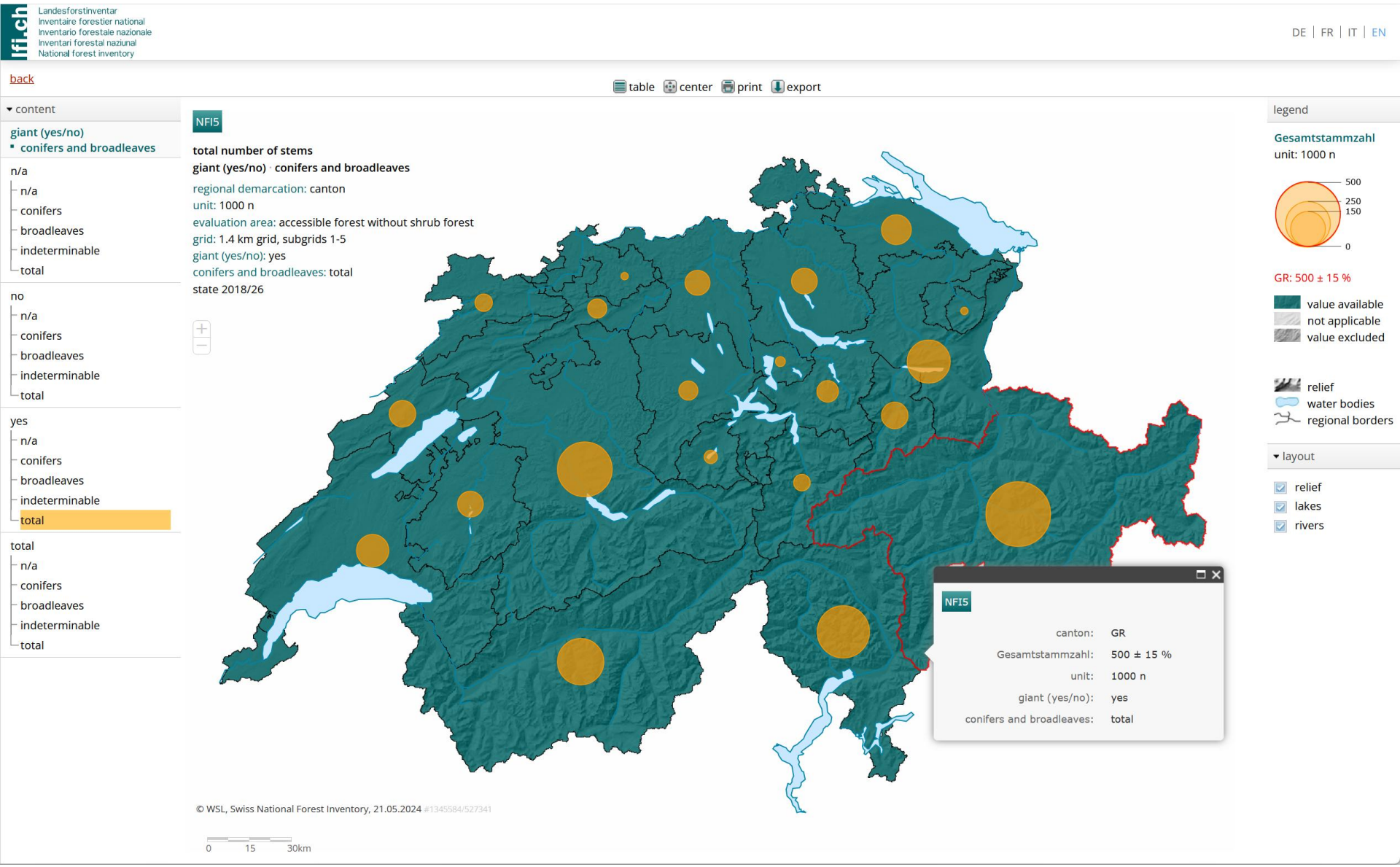
NFI5

total number of stems
giant (yes/no) · conifers and broadleaves

regional demarcation: production region
unit: 1000 n
evaluation area: accessible forest without shrub forest
grid: 1.4 km grid, subgrids 1-5
state 2018/26

		production region											
		Jura		Plateau		Pre-Alps		Alps		Southern Alps		Switzerland	
giant (yes/no)	conifers and broadleaves	1000 n	±%	1000 n	±%	1000 n	±%	1000 n	±%	1000 n	±%	1000 n	±%
no	conifers	39081	5	41206	5	71021	4	143278	3	27163	8	321749	2
	broadleaves	45808	4	44826	4	37553	6	48198	5	56633	5	233018	2
	indeterminable	54	74	17	100	17	100	49	58	0	.	137	40
	total	**84943**	3	**86049**	3	**108591**	3	**191525**	2	**83796**	4	**554904**	1
yes	conifers	195	23	318	16	407	14	860	10	286	22	2066	7
	broadleaves	15	71	191	24	48	41	54	38	185	29	493	16
	total	**210**	22	**508**	13	**455**	14	**913**	10	**472**	17	**2558**	6
total	conifers	39277	5	41524	5	71428	4	144138	3	27449	8	323815	2
	broadleaves	45823	4	45017	4	37600	6	48252	5	56818	5	233510	2
	indeterminable	54	74	17	100	17	100	49	58	0	.	137	40
	total	**85153**	3	**86557**	3	**109045**	3	**192439**	2	**84268**	4	**557462**	1

Webapplication for Mapping ...



Out of the Office ...

Products

- Internal calculations and web applications to support processes;
- External webapplication with 100‘000+ official result tables; four languages; https://www.lfi.ch/en/
- Map producing directly from database with vector-tile technology
- Providing **linked data** for Swiss Open Government Data https://visualize.admin.ch/en
- Base data for scientific papers and applied publications for stakeholders (cantons, FOEN, communities, foresters, public)
- Providing high precise **ground truth** data for calibrating all kinds of models, training machines (KI-models) and remote sensing products;

The Full Overview of NFI Dataflow ...

a) collecting data according to the formula used in the derivation
b) saving the results of each derivation in the respective analysis table and saving all metadata about the derivation
c) importing collected parameter set which defines a specific analysis
d) importing data for a desired statistical analysis
e) generating web result based on statistical results delivered

Database - Raw Data
- Field Assessment
- Aerial Imagery
- Spatial Datasets
- Administation &

Web Application -
- Derivation
- Results
- Content Maintenance
- System Information
- Administration Tools

Statistical Analysis -
- Import
- Data Preparation
- Statistical Analysis
- Statistical Results

Database - Derived Data
- Analysis Tables
- Imported External Data
- Results & Parametersets
- Administation & Metadata

Web Application - Public
- Result Tables & Maps
- Products
- Filter & Search

www.lfi.ch

f) saving the results including the parameter set
g) visualizing saved result tables
h) collecting saved meta information for displaying on web pages

-> see Traub, Meile, Speich, Rösler: https://dx.doi.org/10.1016/j.compag.2016.11.016