

Regional vegetation validation in the Gaillard synthesis

Citation. Gaillard, M.-J. et al. (2010), “[Holocene land-cover reconstructions for studies on land cover-climate feedbacks](#),” *Climate of the Past* **6**, 483–499. The paper argues for quantitative land-cover histories as inputs to climate research and describes the LANDCLIM program. Its Figure 6 summarizes southern Swedish REVEALS validation reported by Hellman et al. (2008a, b). It illustrates that pollen abundance and plant cover differ, and that productivity/dispersal corrections can bring regional estimates closer to observed cover. We examine that figure as published validation evidence, while keeping it separate from a test of our implementation. [Primary article and Figure 6](#); [our vegetation-reference decision](#).

Our source-bound calculation finds **5.11 percentage points** mean absolute disagreement for the published REVEALS values versus **17.33 points** for raw pollen, on nine rounded region/taxon labels visible in all three figure panels. This is an interpretable reading of an already published figure. It is not reproduction of the underlying 26-taxon analysis, nine independent landscapes or accuracy of any repository prediction. [Retained calculation](#); [nine exact label rows](#).

What is observed, modeled and compared

Figure 6 contains two regions, **Skåne and Småland**, with panels for pollen percentage, REVEALS-estimated cover and observed vegetation percentage. The original comparison covers 26 taxa and names labels of at least 2%; its caption identifies southern Swedish productivity estimates and the underlying Hellman studies. These regional percentages have different observation routes. A pollen percentage divides counted grains; observed cover refers to vegetation area; modeled cover follows a pollen-production and transport calculation. None is a taxon count of physical samples. [Figure 6, p. 490](#).

REVEALS means Regional Estimates of VEgetation Abundance from Large Sites. Pollen productivity estimates (**PPE**) account for taxa releasing different numbers of grains, while fall speed and transport affect how far grains reach a basin. Basin and spatial support determine which landscape an estimate describes. A modeled vegetation fraction cannot be validated by comparing it with another modeled fraction from the same source observations. The [LandClim paper account](#) explains the related local **LOVE** calculation; this Figure 6 result is a regional REVEALS comparison.

Our retained [producer](#) authenticates the exact original PDF bytes and calculates from a declared manual transcription. It selects a row only when that taxon’s name and whole-percentage value are printed in all three panels for the same region. There are five Skåne and four Småland rows. The remaining taxa are excluded from this calculation, rather than assigned zero. The selected rows therefore do not constitute a complete composition and must not be normalized to 100% or interpreted as an unbiased selection of model errors. Manual reading is a source-interpretation decision; source hashing authenticates the document, not the transcription’s independent correctness.

A worked label-to-result example

The Skåne **Cerealia-t** row has printed percentages **4 pollen**, **32 REVEALS** and **41 observed vegetation**. Its raw-pollen absolute difference is $\text{abs}(4 - 41) = 37$ percentage points; its REVEALS difference is $\text{abs}(32 - 41) = 9$ points. Those units are percentage points, not a relative percent change or an error on a 0–1 fraction scale. The row illustrates why a small grain share can coexist with a much larger land-cover share. It does not establish a cereal observation at one historical core or a present-day predictor’s error. [Source Figure 6](#); [exact retained row](#).

Selected published labels	Rows	Pollen MAE, percentage points	REVEALS MAE, percentage points
Skåne	5	14.80	4.20
Småland	4	20.50	6.25
Pooled named subset	9	17.33	5.11

Every selected region/taxon cell has equal weight in the pooled mean, so the regions contribute five and four weights respectively. The [machine result](#) retains unrounded arithmetic; the figure inputs themselves are whole percentages. No standard error, confidence interval, significance test or between-landscape uncertainty is calculated from these nine labels. The original figure's explanatory role and our selected-label summary must not be extended to all 26 taxa.

What this evidence leaves open

The primary paper supplies an existing regional validation illustration. We reproduce the arithmetic of its selected printed labels, rather than rerunning its source REVEALS model. Exact original vegetation survey areas, full numerical 26-taxon vectors, pollen source counts, basin/site membership, productivity uncertainty and the sampling design belong to the underlying Hellman studies and need their own authenticated inputs. A clear comparison in a review figure does not supply all of those records. [Original-paper figure and references](#); [current source review](#).

For this repository, the [EMPD2–NFI pairing decision](#) still admits **zero of 16** nearby vegetation leads. That denominator concerns different sources and predictions. The published southern Swedish agreement neither admits those pairs nor scores our present engine. The [source/model lineage account](#) also prevents treating LandClim modeled cover as independent measured truth for related input lakes.

Reproduce the bounded calculation

The original PDF has SHA-256 46bd92497d03169735b9a4665d71214c5dd8ff52ceddf39c577f0a4495fe73bd. With those lawful source bytes available, run from the repository root:

```
sh
```

```
PYTHONDONTWRITEBYTECODE=1 .venv/bin/python \
  reports/science/vegetation-reference/southern-sweden-figure/compare.py \
  --source-pdf /path/to/source/gaillard-2010.pdf \
  --output artifacts/jobs/southern-sweden-figure/reproduction
cmp artifacts/jobs/southern-sweden-figure/reproduction/printed-labels.csv \
  reports/science/vegetation-reference/southern-sweden-figure/printed-labels.csv
cmp artifacts/jobs/southern-sweden-figure/reproduction/report.json \
  reports/science/vegetation-reference/southern-sweden-figure/report.json
```

This produces the nine-label ledger and summary, without changing the existing results or supplying missing raw observations. A reader can inspect the retained rows and open article without the local source capsule.

The precise **unsent** request is for the underlying study's full unrounded pollen, REVEALS and independently observed cover vectors, landscape/taxon denominators, original survey and sample collection times, physical lake memberships and basin settings, PPE/fall-speed versions and uncertainty inputs, plus reuse terms. A qualified reviewer should verify the nine printed-label transcriptions and decide whether a proposed new landscape comparison shares the original estimand and independent observation route. No reviewer feedback or collaborator comprehension exercise is recorded here.