

our template

SciVeri-Bench

Scientific Task Proposal

1

Task Objective

2

Why it matters

3

Data, software & environment

4

Evaluation: rubric + metric

5

Deliverables

6

Can a metric capture quality?

paper.tex

· Overleaf or XeLaTeX

a real proposal

SciVeri-Bench

Scientific Task Proposal

What does community context add to predicting future local occurrence?

Task Proposer (domain scientist)

1 Task Objective

1.1 Scientific question and objective

Given a focal plant species' identity, cover, location and observation history, what additional information, if any, does the surrounding assemblage provide about its occurrence in the same plot next year? Local plant communities reflect environmental selection, dispersal and stochastic change (Vellend, 2010). Neighbouring species may therefore reveal conditions relevant to a focal species' continued local occurrence that its own measurements miss. This project asks what that information adds to prediction and where its contribution is most informative.

Baseline vegetation plot

Information compared

Next year

What does the baseline community add to the forecast?

The National Plant Monitoring Scheme (NPMS) supplies repeated vegetation inventories with species cover and observation histories (National Plant Monitoring Scheme, 2026). The measured outcome is whether a focal species observed at baseline is detected on either of two qualifying surveys of the same plot next year. This is an observational proxy for short-term local persistence: imperfect detection and possible turnover between surveys prevent it from identifying continuous occupancy or individual survival (Chen et al., 2009). The predictive target is this observed outcome.

1.2 Work to be performed and scope

Develop predictions with and without neighbouring-species information and explain their differences. A strong focal-only comparison should account for species identity, cover, location, recorded habitat, management, season and previous records, alongside inventory richness. The scientific objective is to establish what composition adds beyond these predictors, including convincing evidence of little or no added value. Model families, community representations, hypotheses and diagnostic analyses remain open.

Each eligible species in a baseline plot defines its own prediction case: given species S and its cover in plot p in year t , predict its detection there in year $t + 1$. A plot therefore supplies several focal cases. NPMS inventory surveys record vascular flora throughout the plot. The last valid baseline-year visit defines a single observed assemblage; earlier visits contribute history. The first and last qualifying visits next year, at least 30 days apart, define the outcome. The temporal design is:

• Training: observations from 2015–2022, with next-year outcomes available through 2023.

• Validation: 2023 observations with 2024 outcomes, used to develop and choose methods.

• Test: 2024 observations with 2025 outcomes. The 2025 outcomes remain hidden until all methodological choices and predictions are frozen.

The test contains 3,228 species–plot cases from 234 plots in 59 survey squares. It tests forecasting at monitored locations: plots may also occur in earlier years.

1

What does community context add to predicting future local occurrence?

plant ecology

public NPMS data

4 pages

what the scientist pins down

1 An open question

“Given a focal plant species’ identity, cover, location and observation history, what additional information, if any, does the surrounding assemblage provide about its occurrence in the same plot next year?”

Baseline vegetation plot

Information compared

Next year

What does the baseline community add to the forecast?

p.1

4 A rubric, fixed before any experiment

Validity

25

Insight

25

Interpretation

25

Reproducibility

25

leakage or fabrication = invalid

other faults –25 each · AUROC reported as a diagnostic

p.3

6 Can one scalar judge the science?

☐ Yes

☒ **Partly**

☐ No

“AUROC measures prediction, not what community context contributes or why.”

p.4