

Random Forest Result #145 — Analysis Report

Generated 2026-09-21 02:37:11 · 1 run(s) · CAT-G Data Analyzer

1. random forest (#145)

2026-09-21 02:14:42 · random_forest · status: completed

The model explains 78.4% of the variance in the target under cross-validation.

task_type	regression
r2_mean	0.784445
r2_std	0.027108
mse_mean	0.028559
mse_std	0.001398

permutation importance

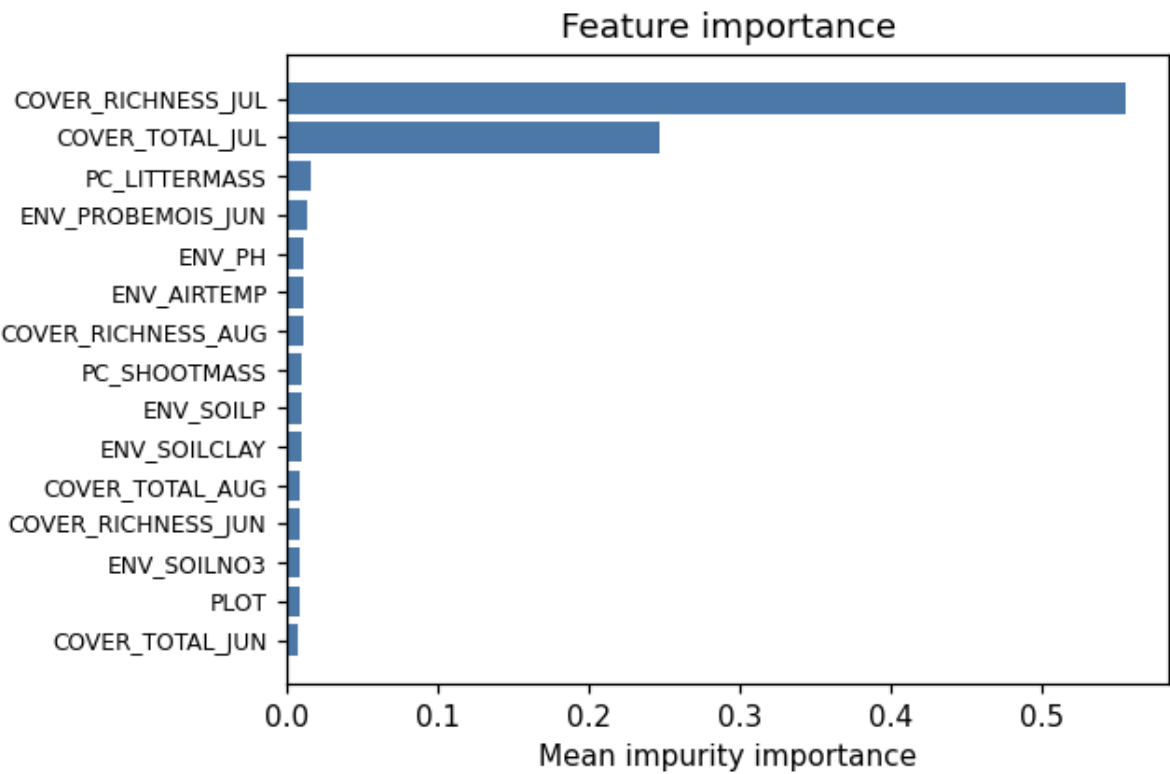
name	mean	std
PLOT	0.001728	0.001922
ENV_SOILNH4	0.001038	0.001454
ENV_SOILNO3	-0.000827	0.001448
ENV_SOILP	-0.001439	0.001833
ENV_SOILOC	0.002019	0.000753
ENV_SOILTN	-0.000309	0.000756
ENV_SOILTC	-0.000263	0.001078
ENV_PH	0.001013	0.000999
ENV_SOILCLAY	-0.001452	0.00374
ENV_SOILSILT	1.2e-05	0.001276

shap importance

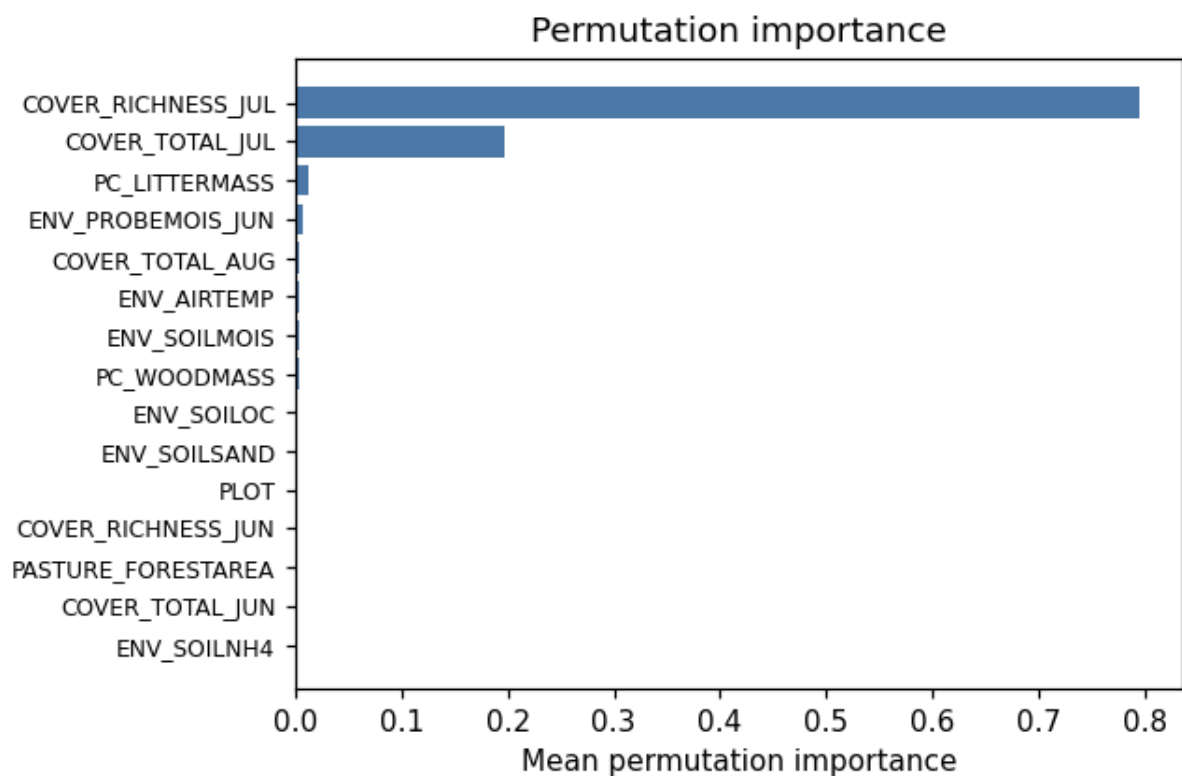
name	mean_abs
PLOT	0.002544
ENV_SOILNH4	0.001869
ENV_SOILNO3	0.002629
ENV_SOILP	0.002738
ENV_SOILOC	0.002754
ENV_SOILTN	0.001206
ENV_SOILTC	0.001763
ENV_PH	0.006021
ENV_SOILCLAY	0.002991
ENV_SOILSILT	0.002607

split metrics

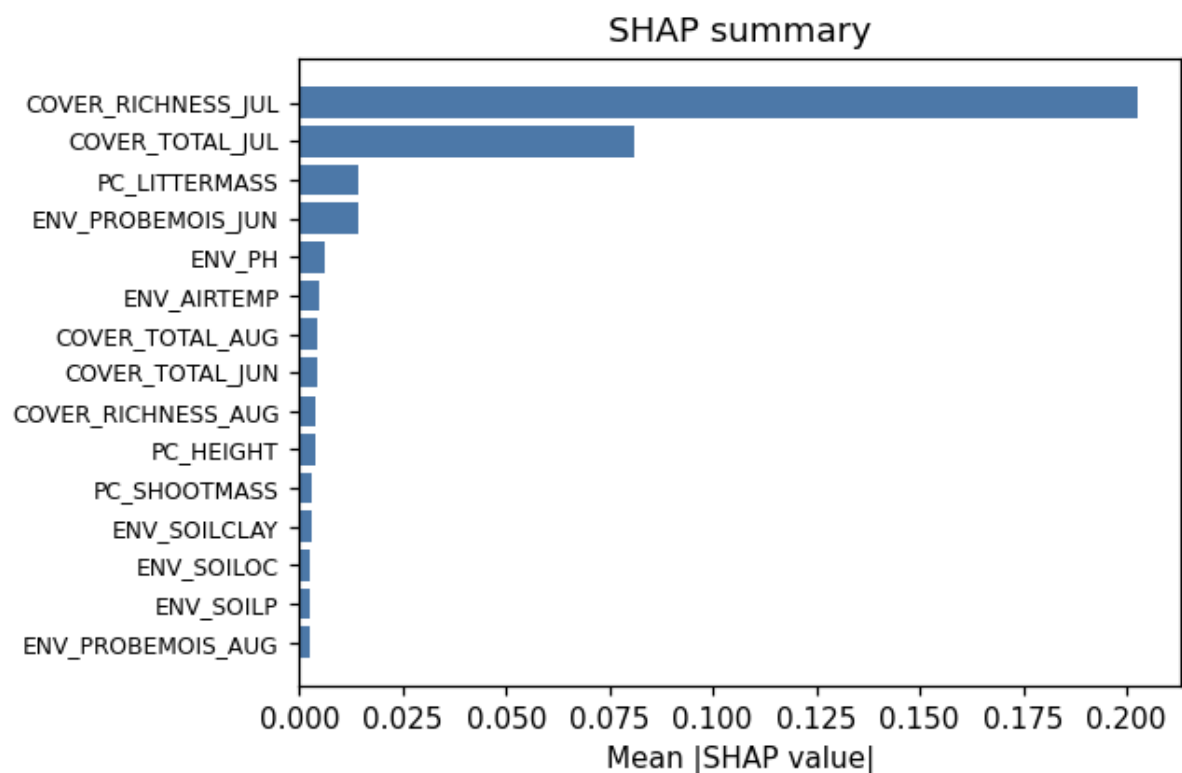
r2	mse
0.746124	0.030228
0.80457	0.026807
0.802639	0.028641



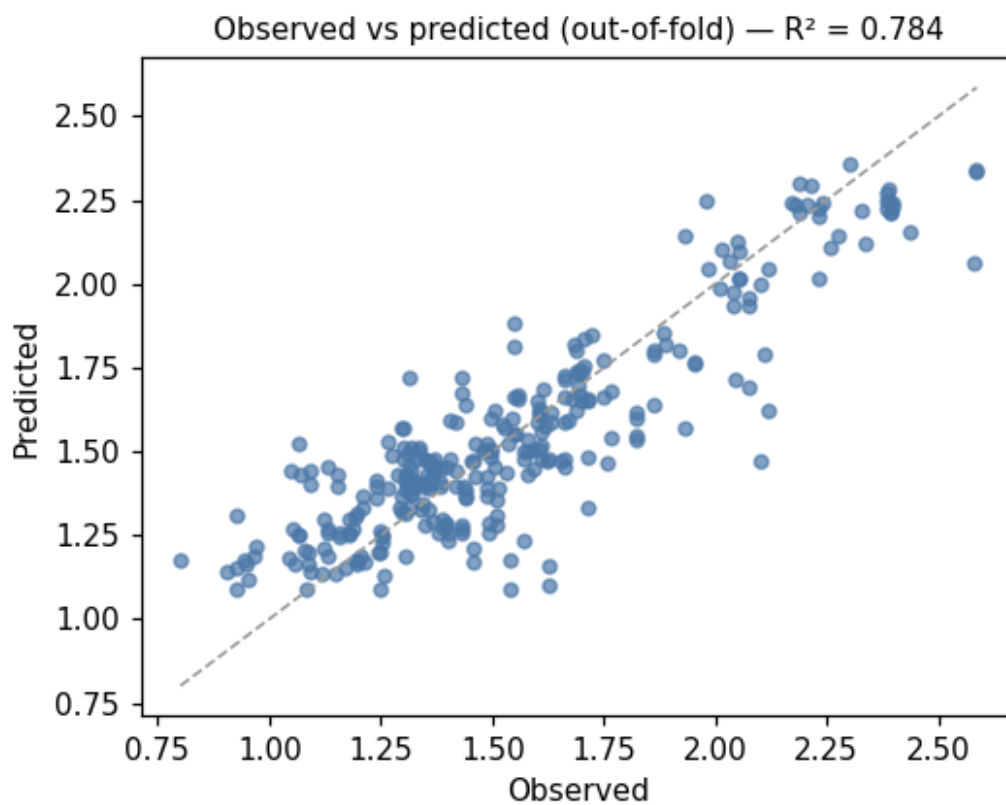
importance bar



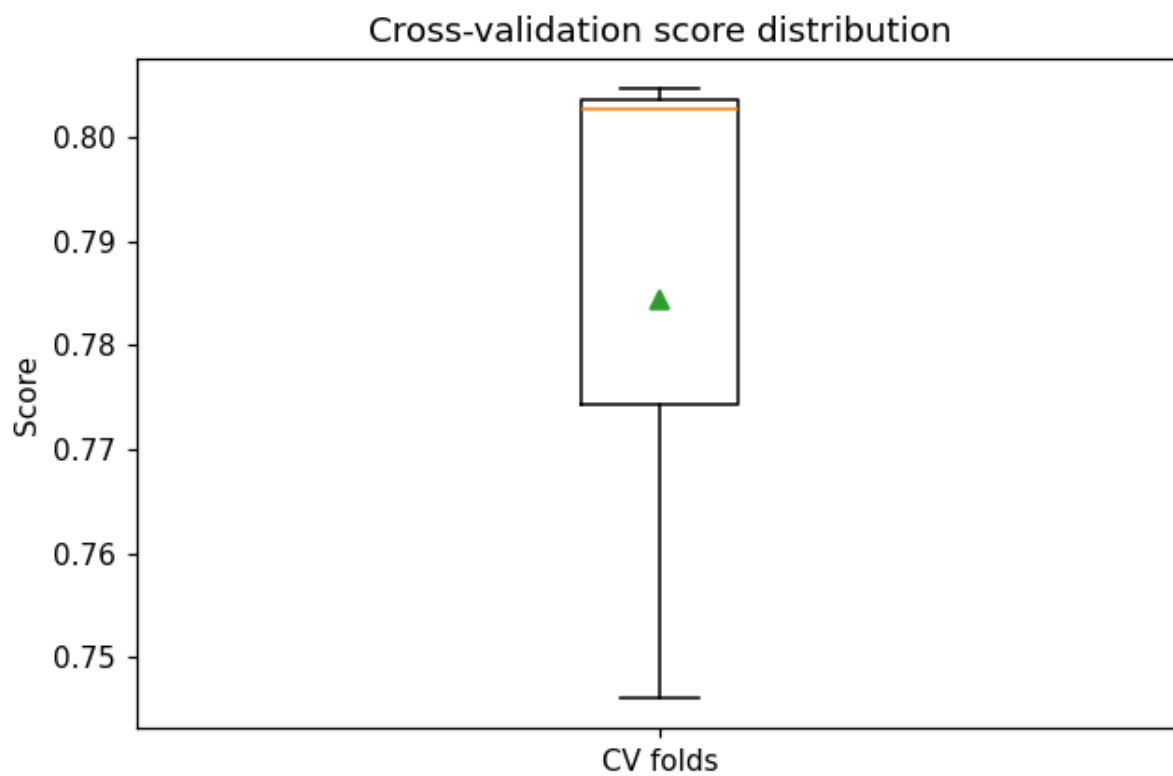
permutation bar



shap summary



pred vs obs



cv boxplot

Method & Provenance

method	Random Forest (Regression)
--------	----------------------------

engine	Python · scikit-learn 1.8.0
datasets	gallery — catg_example_04_vegetation_diversity.csv
started	2026-09-21 02:14:42.722532
permanent link	https://cat-g-omxpert.wishartlab.com/results/by-token/023c0e89aaf94a488763f77c20de

Parameter	Value
task_type	regression
model_type	regression
target_column	COVER_SHANNON_JUL
n_estimators	60
n_splits	3
test_size	0.25
run_univariate	True
run_permutation	True
run_shap	True
permutation_repeats	10
fdr_threshold	0.05

Command history

```
# python
from sklearn.ensemble import RandomForestRegressor

# python
model = RandomForestRegressor(n_estimators=60, random_state=42)

# python
scores = sklearn.model_selection.cross_val_score(model, X, y, cv=3)
```