

Random Forest Result #148 — Analysis Report

Generated 2026-09-21 02:37:40 · 1 run(s) · CAT-G Data Analyzer

1. random forest (#148)

2026-09-21 02:15:49 · random_forest · status: completed

The model explains 93.6% of the variance in the target under cross-validation.

task_type	regression
r2_mean	0.935626
r2_std	0.055255
mse_mean	32.913705
mse_std	29.804256

permutation importance

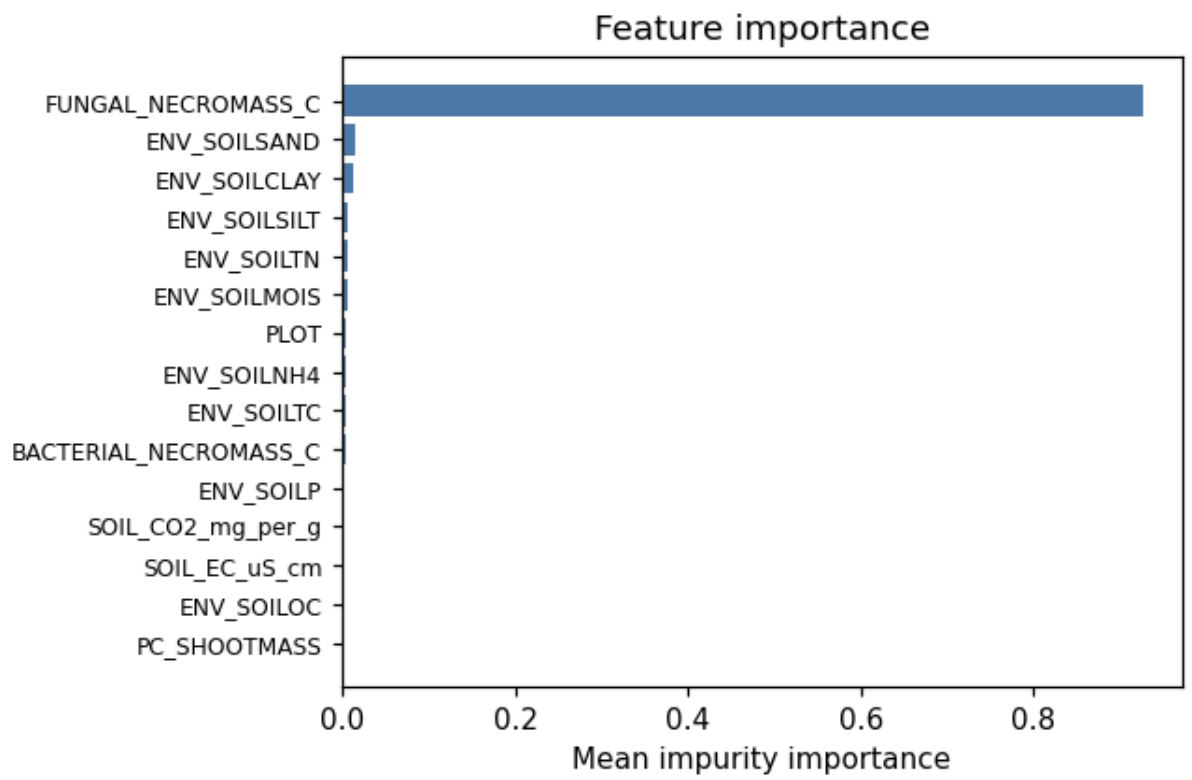
name	mean	std
PLOT	0.000829	0.001282
ENV_SOILNH4	-0.003147	0.004054
ENV_SOILNO3	8.9e-05	0.000354
ENV_SOILP	-0.000501	0.000819
ENV_SOILOC	2.4e-05	0.000269
ENV_SOILTN	0.00232	0.002203
ENV_SOILTC	0.000416	0.000181
ENV_PH	-0.000126	0.000154
ENV_SOILCLAY	0.001992	0.000425
ENV_SOILSILT	0.000718	0.001379

shap importance

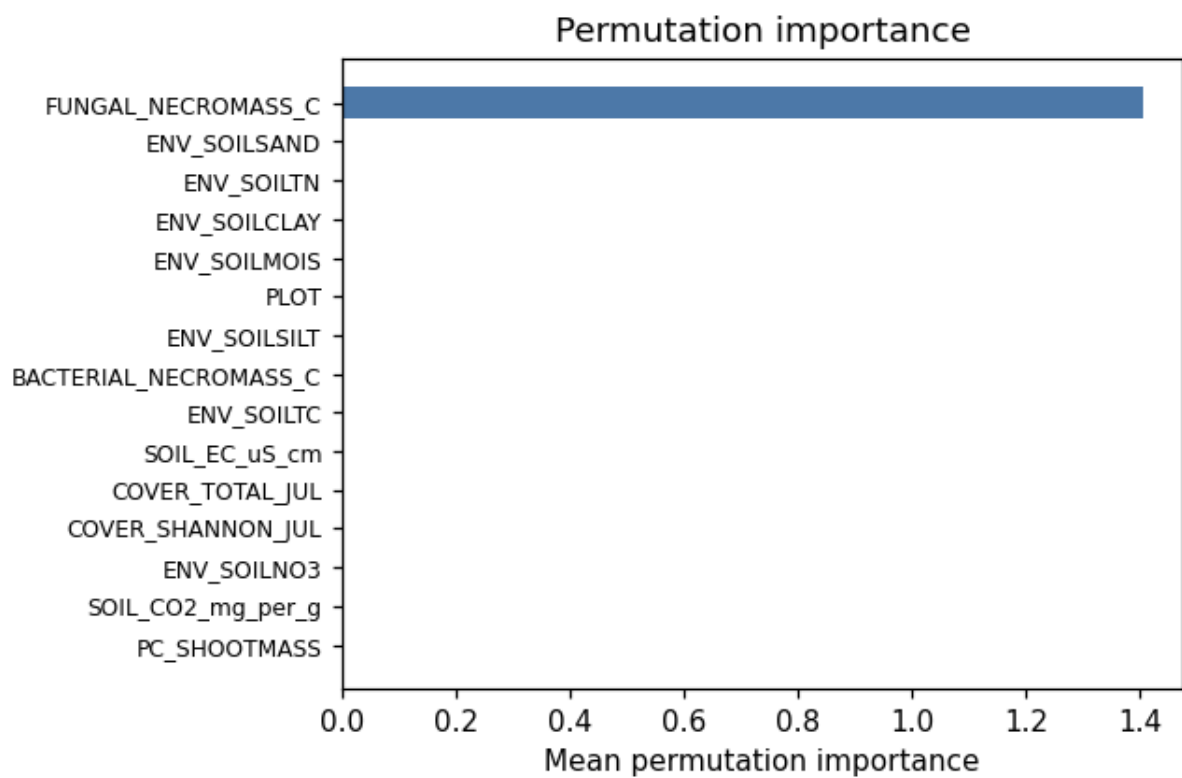
name	mean_abs
PLOT	0.195952
ENV_SOILNH4	0.10979
ENV_SOILNO3	0.064227
ENV_SOILP	0.100928
ENV_SOILOC	0.017297
ENV_SOILTN	0.077718
ENV_SOILTC	0.185985
ENV_PH	0.02226
ENV_SOILCLAY	0.558899
ENV_SOILSILT	0.415885

split metrics

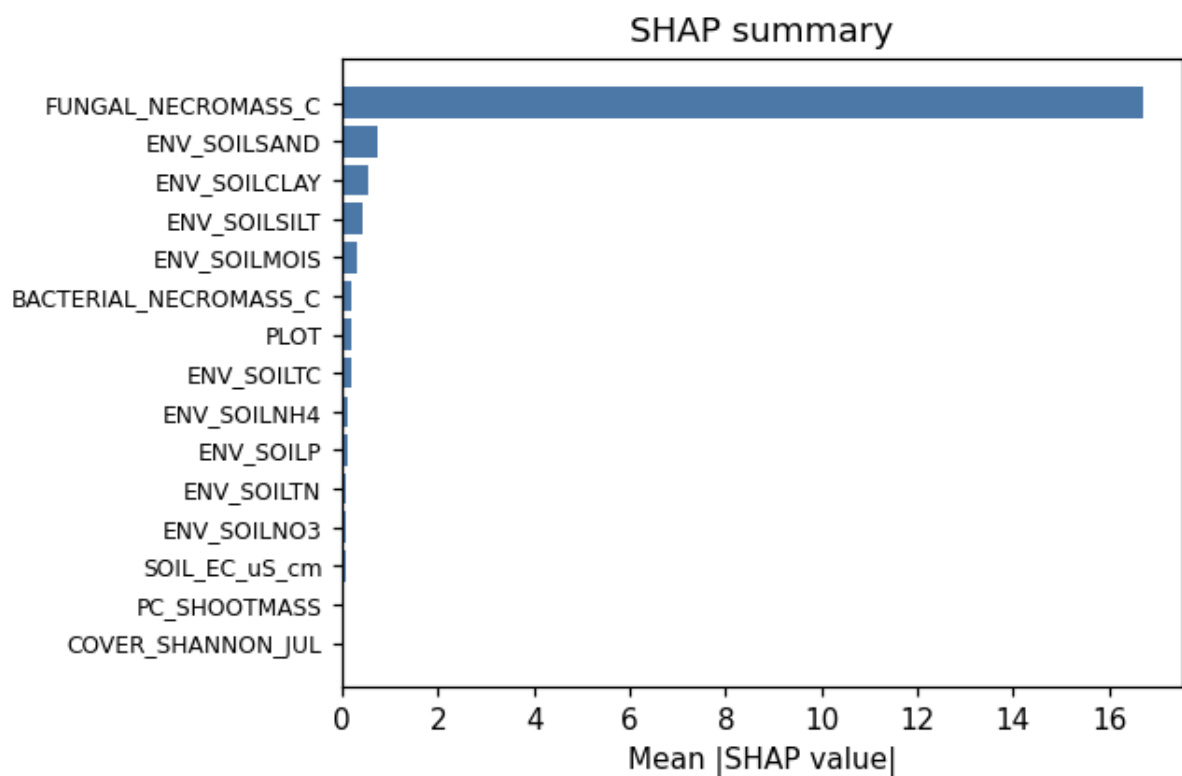
r2	mse
0.990608	3.079392
0.860047	73.616049
0.956222	22.045674



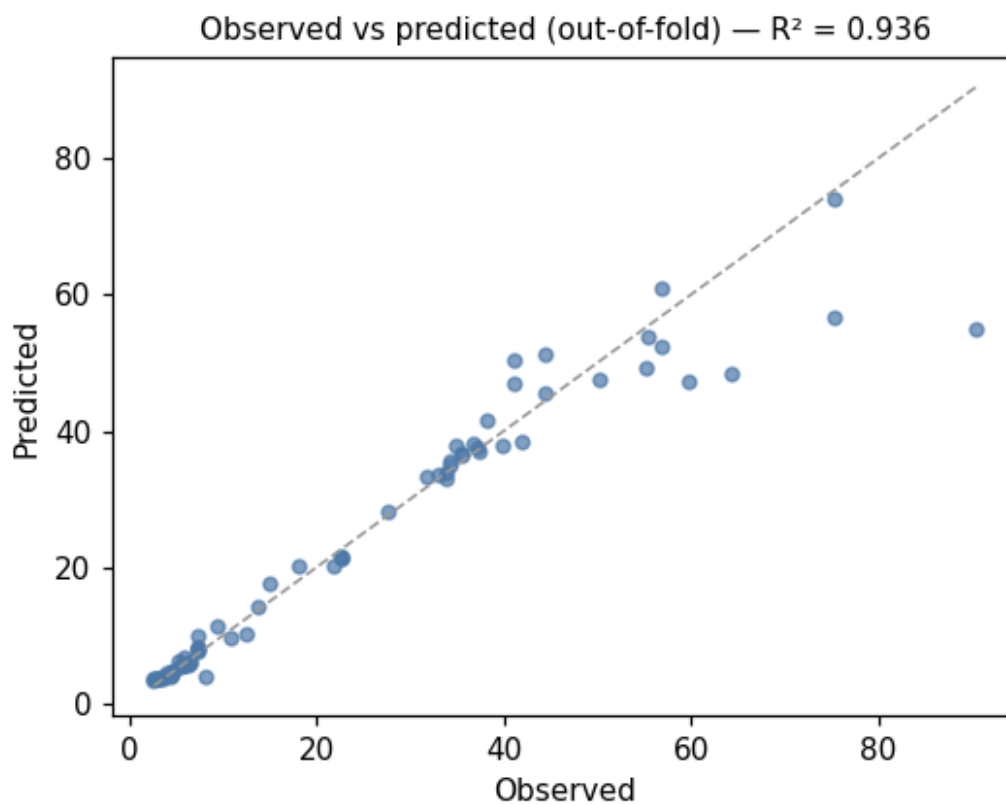
importance bar



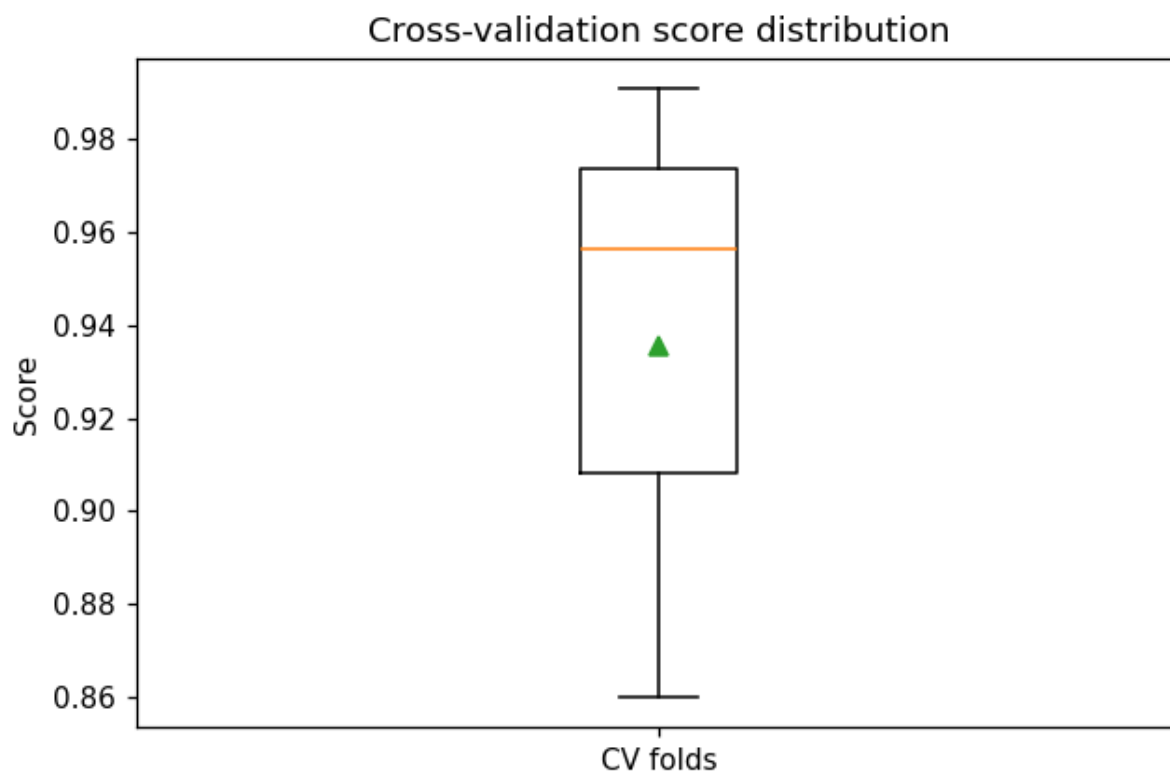
permutation bar



shap summary



pred vs obs



cv boxplot

Method & Provenance

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

engine	Python · scikit-learn 1.8.0
datasets	gallery — catg_example_07_microbial_soil_biology.csv
started	2026-09-21 02:15:49.155216
permanent link	https://cat-g-omxpert.wishartlab.com/results/by-token/8705c62fa577496199d0d6e6fc4f4

Parameter	Value
task_type	regression
model_type	regression
target_column	MICROBIAL_NECROMASS_C
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)
```