

Random Forest Result #144 — Analysis Report

Generated 2026-09-21 02:36:57 · 1 run(s) · CAT-G Data Analyzer

1. random forest (#144)

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

The model explains 53.9% of the variance in the target under cross-validation.

task_type	regression
r2_mean	0.538756
r2_std	0.214189
mse_mean	14.986814
mse_std	12.519635

permutation importance

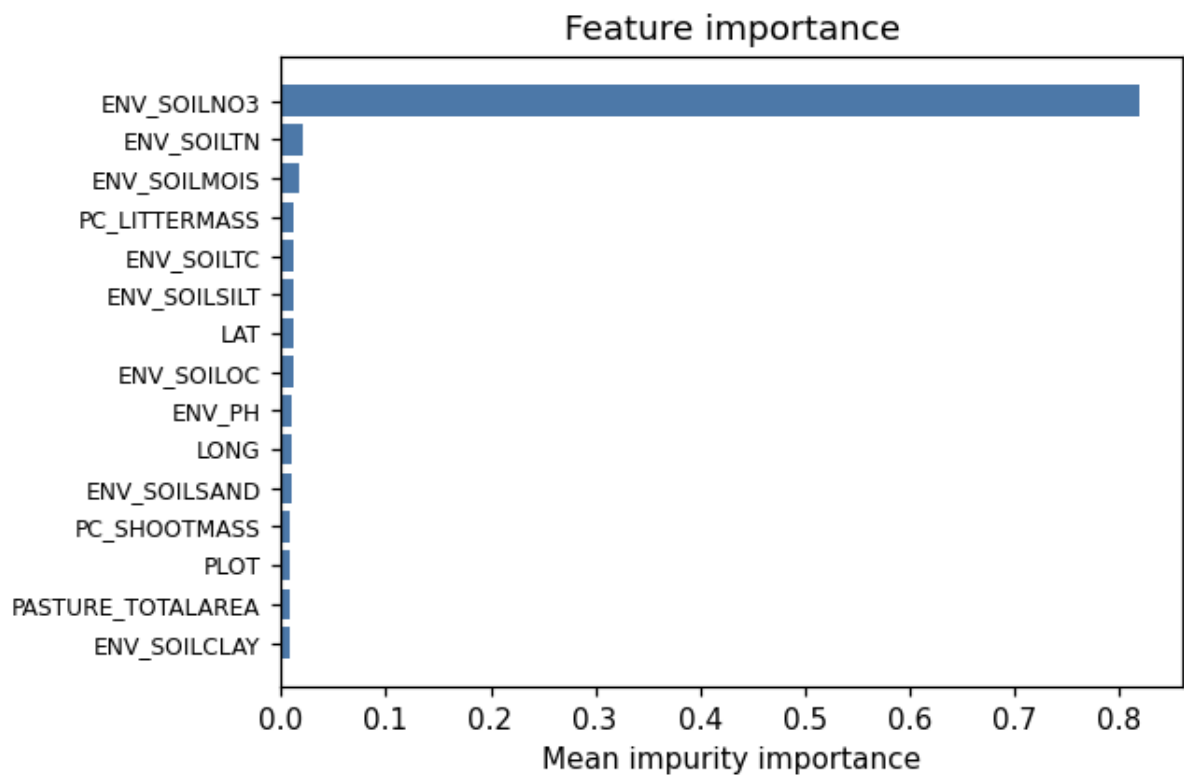
name	mean	std
PLOT	0.011272	0.008572
ENV_SOILOC	0.001235	0.007825
ENV_SOILTC	0.002577	0.011296
ENV_SOILTN	0.102671	0.070685
ENV_PH	-0.002838	0.002858
ENV_SOILCLAY	0.01445	0.009704
ENV_SOILSILT	0.006406	0.023445
ENV_SOILSAND	0.010353	0.00754
ENV_SOILMOIS	0.021319	0.018914
ENV_SOILP	0.000114	0.000991

shap importance

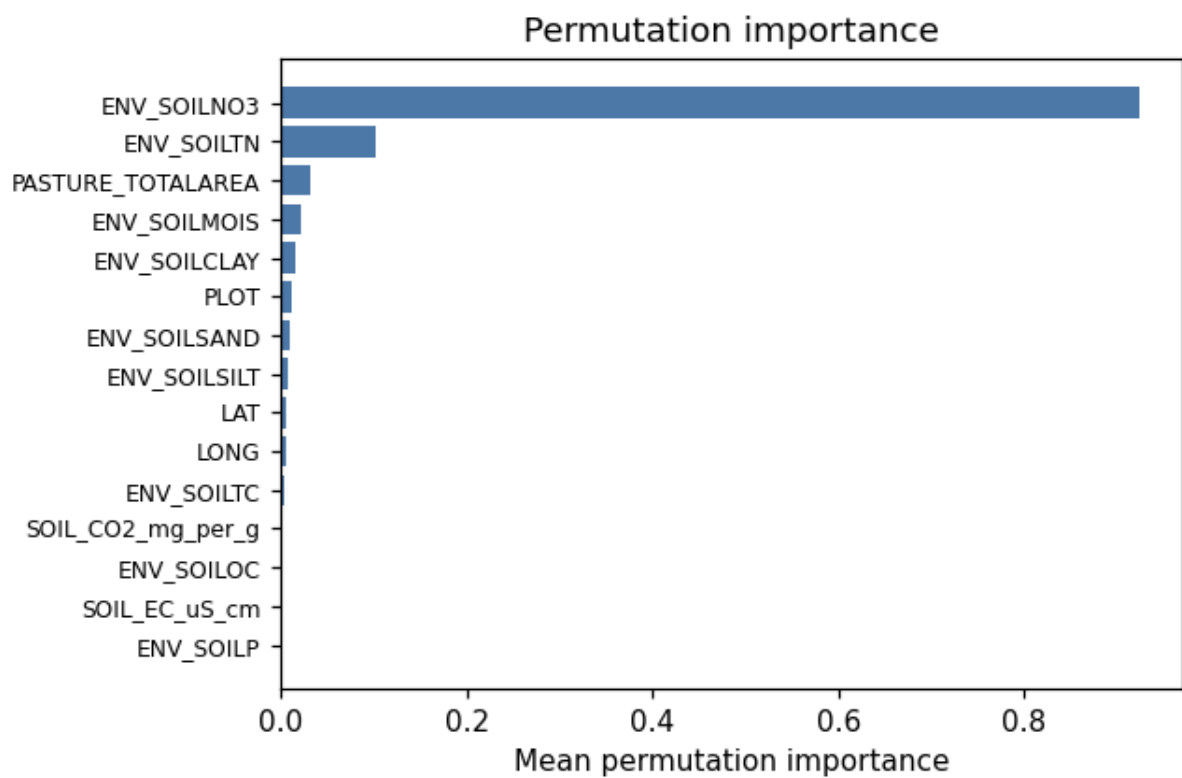
name	mean_abs
PLOT	0.1131
ENV_SOILOC	0.183659
ENV_SOILTC	0.126229
ENV_SOILTN	0.327007
ENV_PH	0.109797
ENV_SOILCLAY	0.061339
ENV_SOILSILT	0.167248
ENV_SOILSAND	0.065622
ENV_SOILMOIS	0.205682
ENV_SOILP	0.03121

split metrics

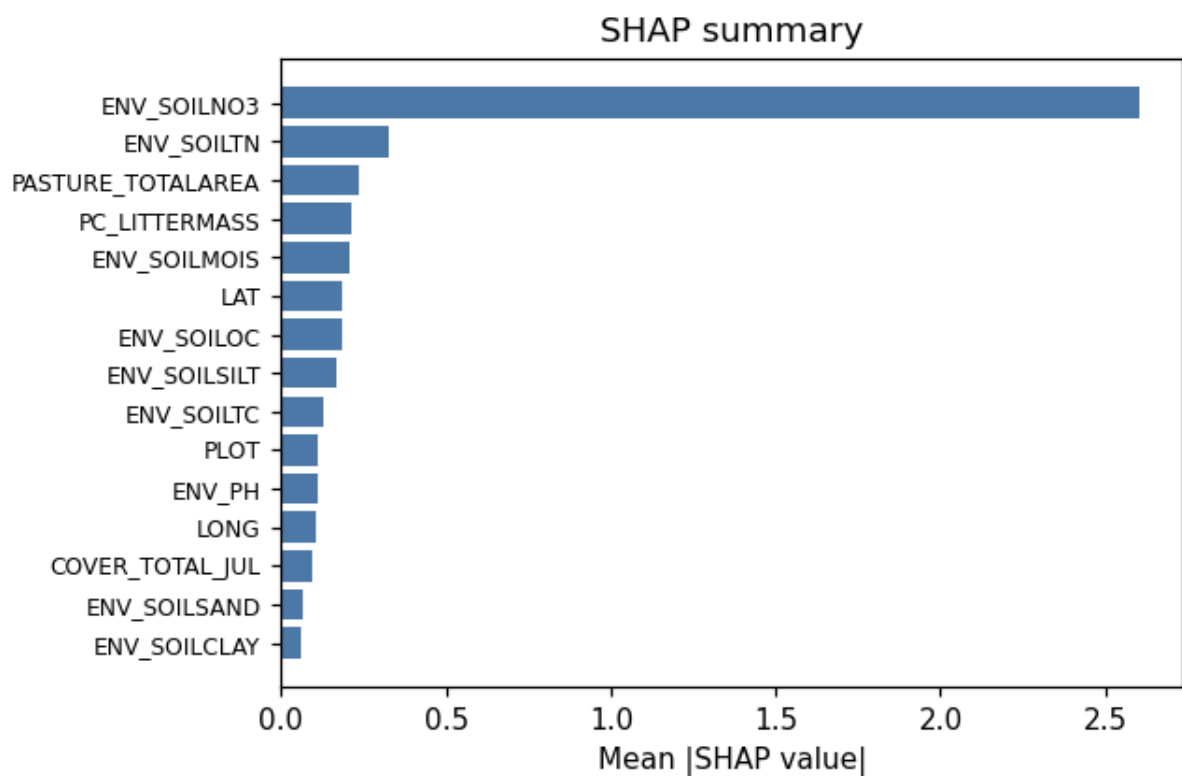
r2	mse
0.308397	5.958141
0.483595	6.311223
0.824277	32.691079



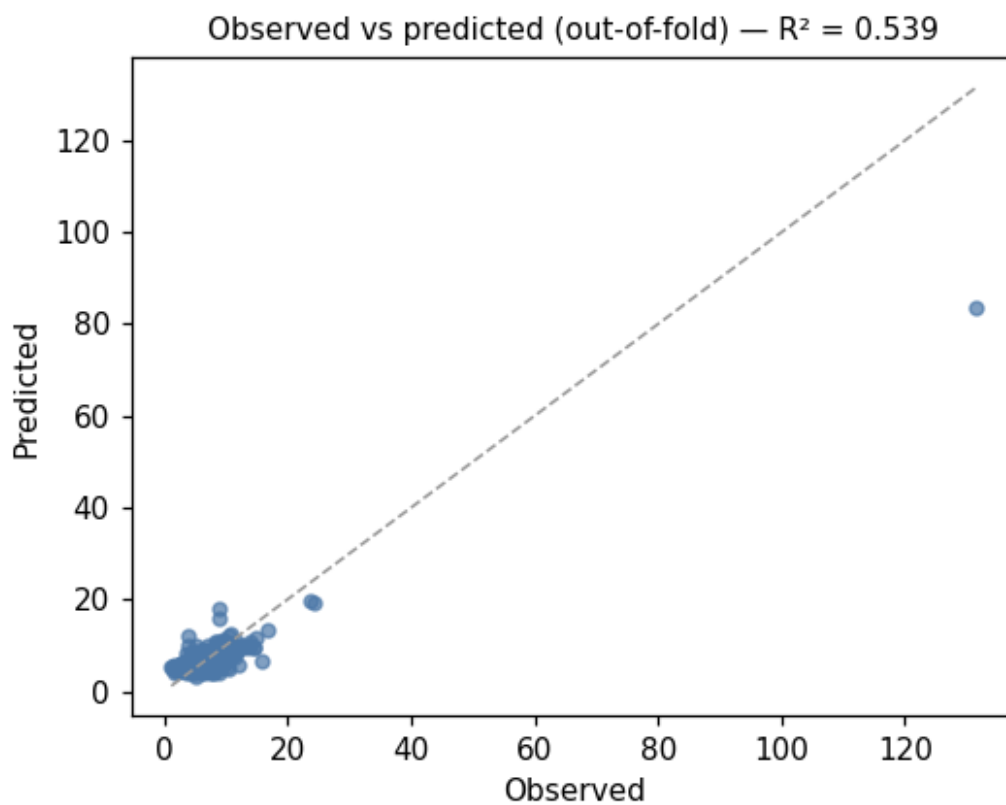
importance bar



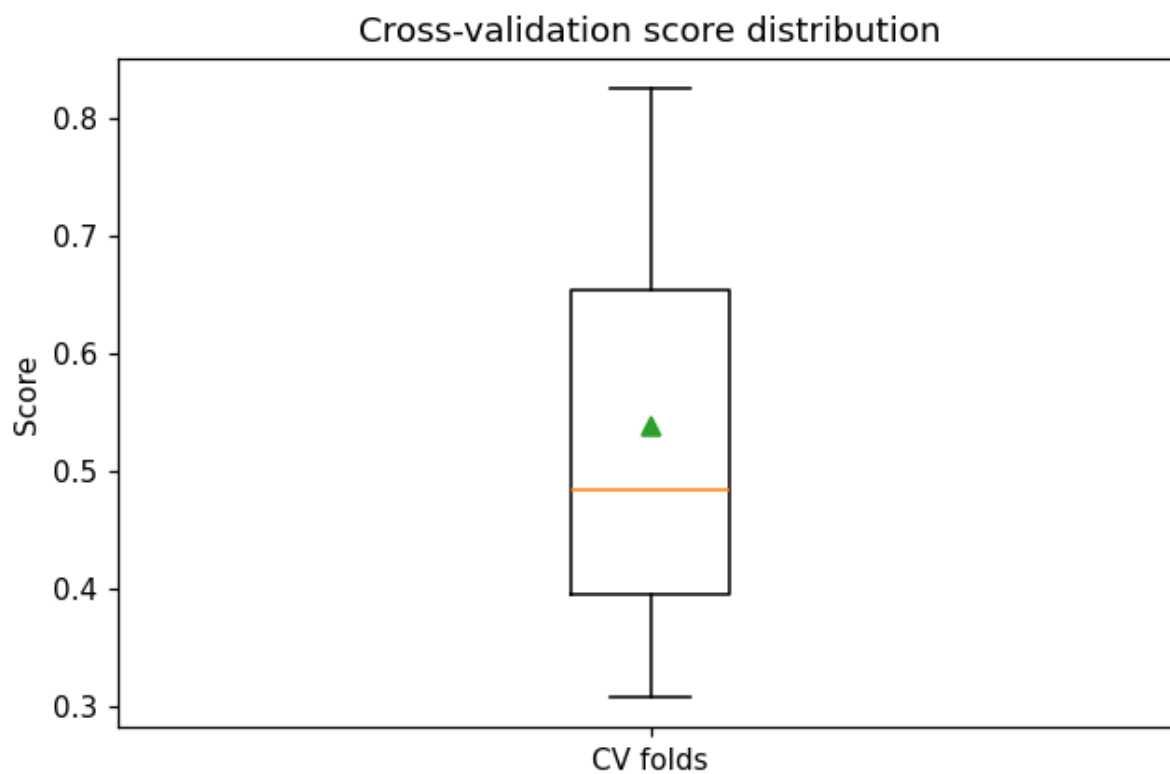
permutation bar



shap summary



pred vs obs



cv boxplot

Method & Provenance

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

engine	Python · scikit-learn 1.8.0
datasets	gallery — catg_example_03_soil_nitrogen_availability.csv
started	2026-09-21 02:14:14.068246
permanent link	https://cat-g-omxpert.wishartlab.com/results/by-token/413ee895cd9742ed95bd1502567

Parameter	Value
task_type	regression
model_type	regression
target_column	ENV_SOILNH4
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)
```