

Random Forest Result #146 — Analysis Report

Generated 2026-09-21 02:37:19 · 1 run(s) · CAT-G Data Analyzer

1. random forest (#146)

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

The model explains 77.2% of the variance in the target under cross-validation.

task_type	regression
r2_mean	0.77179
r2_std	0.041005
mse_mean	199.71057
mse_std	27.668687

permutation importance

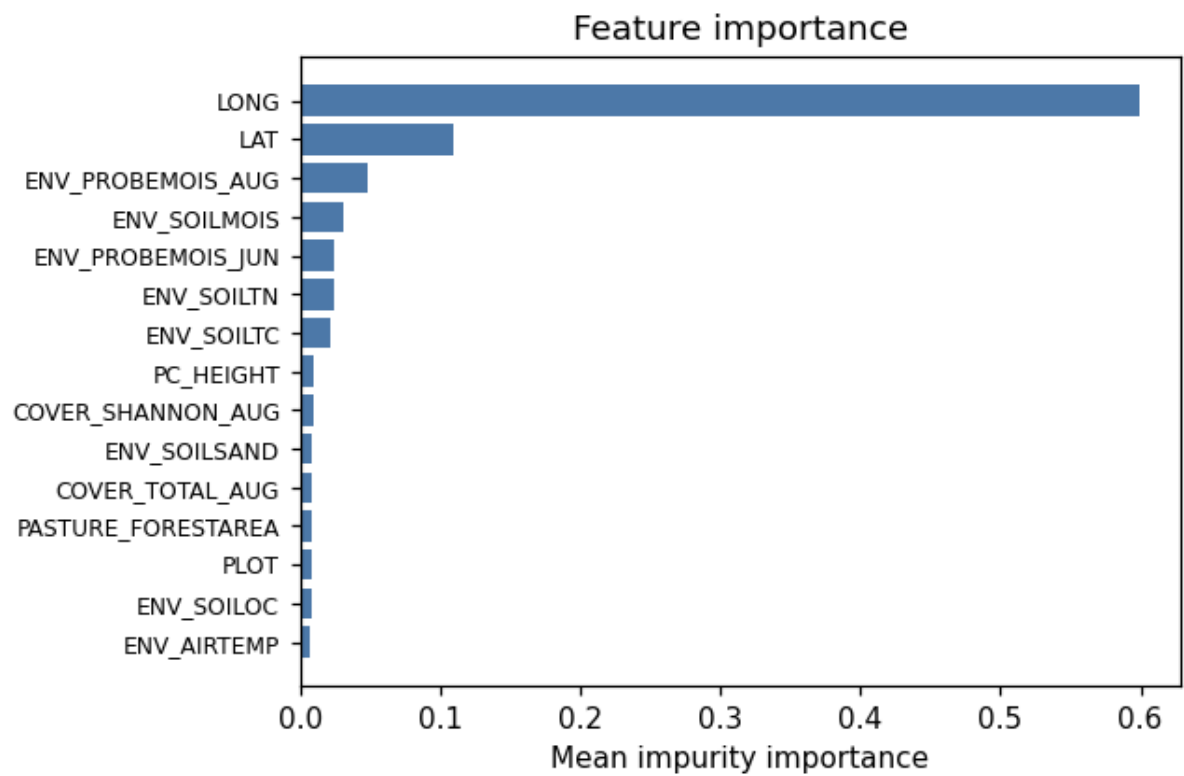
name	mean	std
PLOT	0.000762	0.008948
ENV_SOILNH4	-0.00042	0.002022
ENV_SOILNO3	-0.000454	0.004077
ENV_SOILP	0.004555	0.00288
ENV_SOILOC	0.002502	0.004561
ENV_SOILTN	0.016942	0.011589
ENV_SOILTC	0.018122	0.00958
ENV_PH	-0.000352	0.000364
ENV_SOILCLAY	-0.000108	0.000876
ENV_SOILSILT	0.005832	0.001824

shap importance

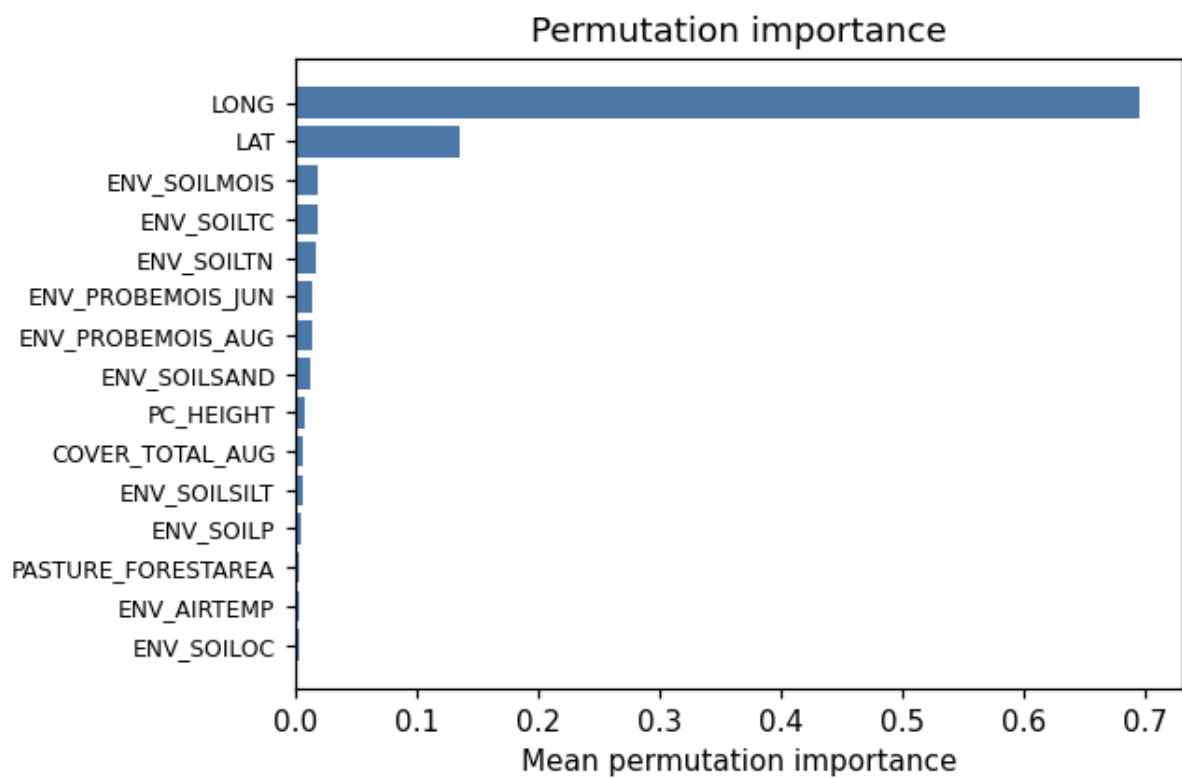
name	mean_abs
PLOT	0.47108
ENV_SOILNH4	0.215687
ENV_SOILNO3	0.136184
ENV_SOILP	0.350946
ENV_SOILOC	0.854207
ENV_SOILTN	0.358174
ENV_SOILTC	0.571677
ENV_PH	0.214236
ENV_SOILCLAY	0.299414
ENV_SOILSILT	0.290186

split metrics

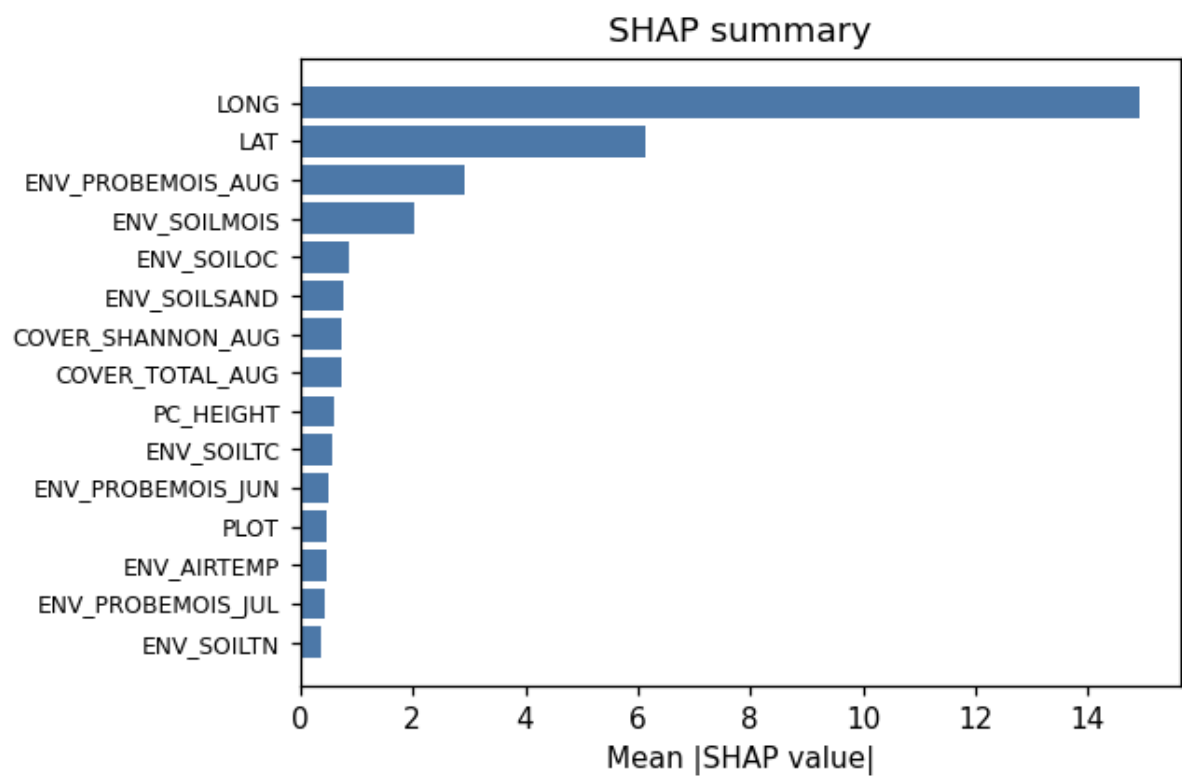
r2	mse
0.745856	234.197297
0.739838	198.477631
0.829675	166.456784



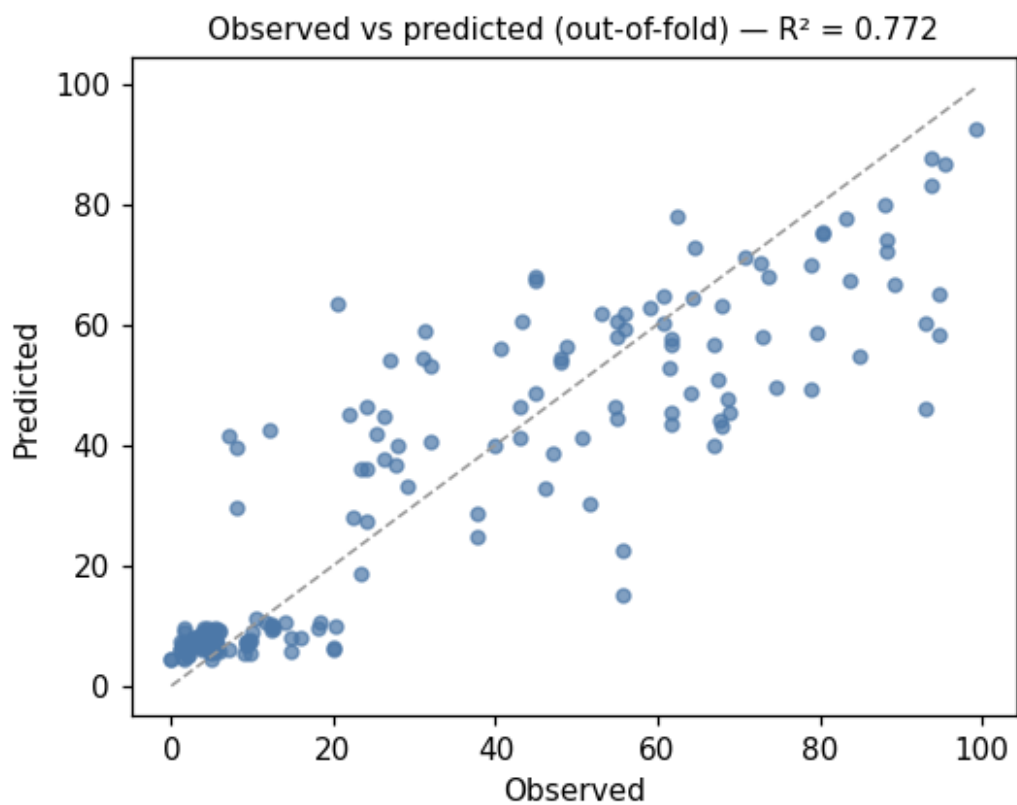
importance bar



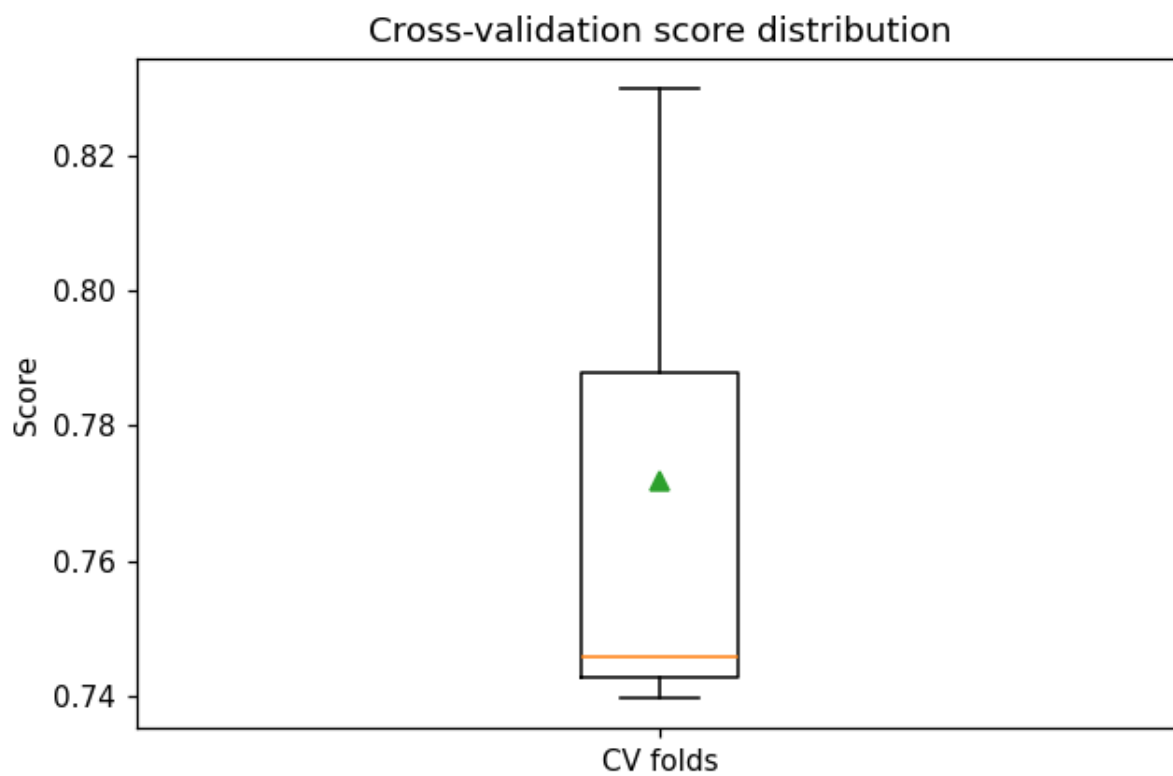
permutation bar



shap summary



pred vs obs



cv boxplot

Method & Provenance

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

engine	Python · scikit-learn 1.8.0
datasets	gallery — catg_example_05_grasshopper_pest_pressure.csv
started	2026-09-21 02:15:02.446843
permanent link	https://cat-g-omxpert.wishartlab.com/results/by-token/41da97babc1945769dd1bb43f3a1

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