

Random Forest Result #139 — Analysis Report

Generated 2026-09-21 02:35:50 · 1 run(s) · CAT-G Data Analyzer

1. random forest (#139)

2026-09-21 02:11:43 · random_forest · status: completed

The model explains 53.4% of the variance in the target under cross-validation.

task_type	regression
r2_mean	0.534495
r2_std	0.020216
mse_mean	40.124996
mse_std	3.35536

permutation importance

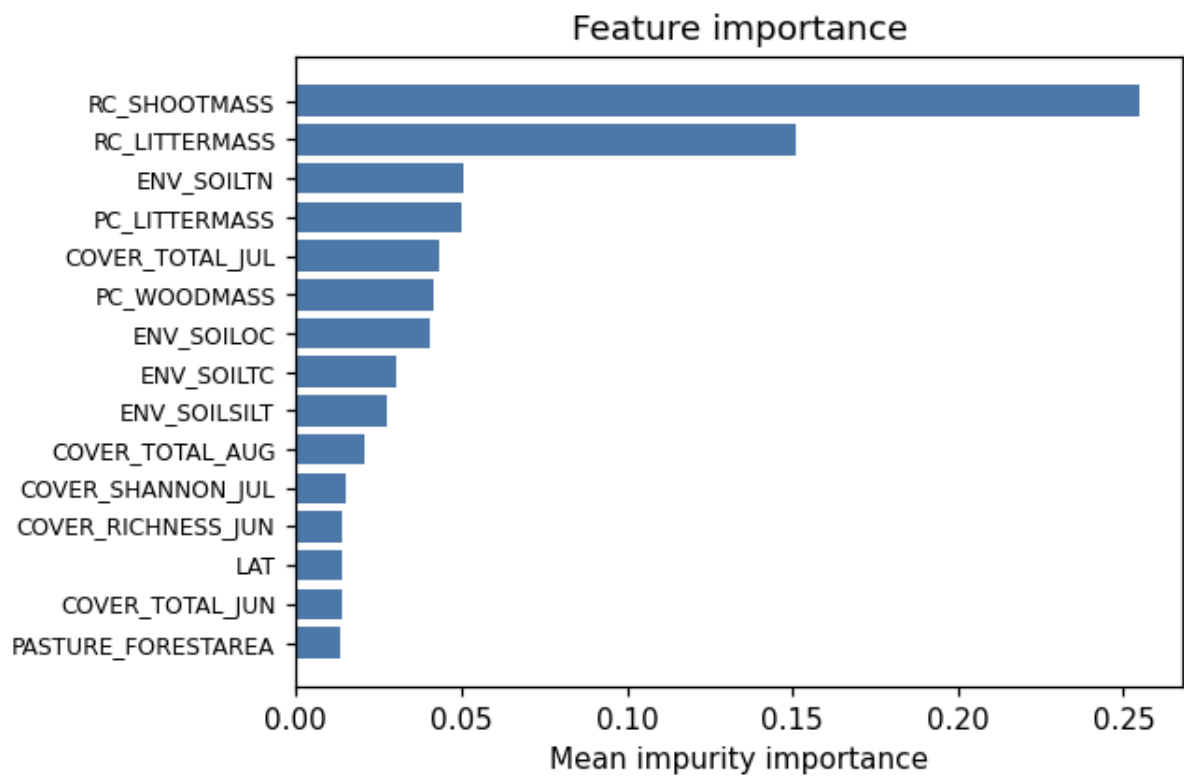
name	mean	std
PLOT	0.001104	0.001748
LAT	0.00181	0.000957
LONG	0.002725	0.002277
ENV_SOILNH4	0.002134	0.00171
ENV_SOILNO3	0.002311	0.003533
ENV_SOILP	0.001566	0.003499
ENV_SOILOC	0.016012	0.006552
ENV_SOILTN	0.011141	0.012211
ENV_SOILTC	-0.000762	0.006243
ENV_PH	-0.002819	0.005171

shap importance

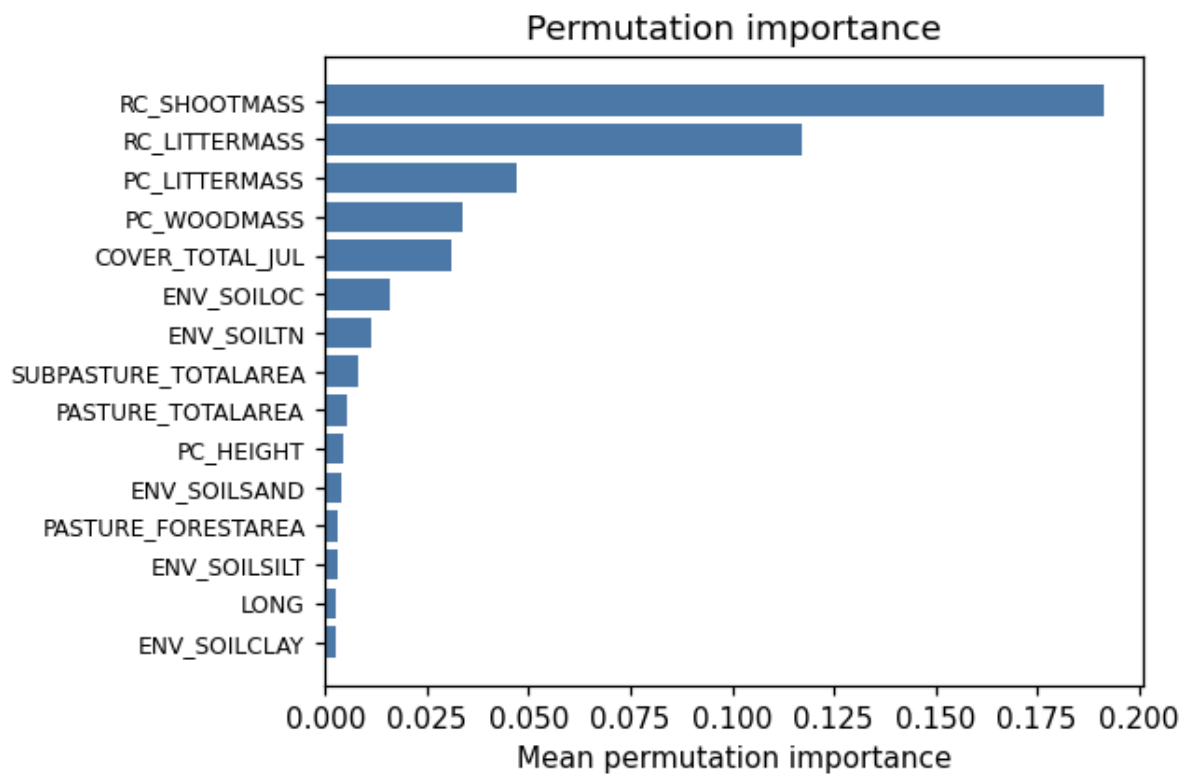
name	mean_abs
PLOT	0.092551
LAT	0.199099
LONG	0.047196
ENV_SOILNH4	0.111703
ENV_SOILNO3	0.050621
ENV_SOILP	0.06099
ENV_SOILOC	0.221387
ENV_SOILTN	0.734706
ENV_SOILTC	0.240267
ENV_PH	0.192217

split metrics

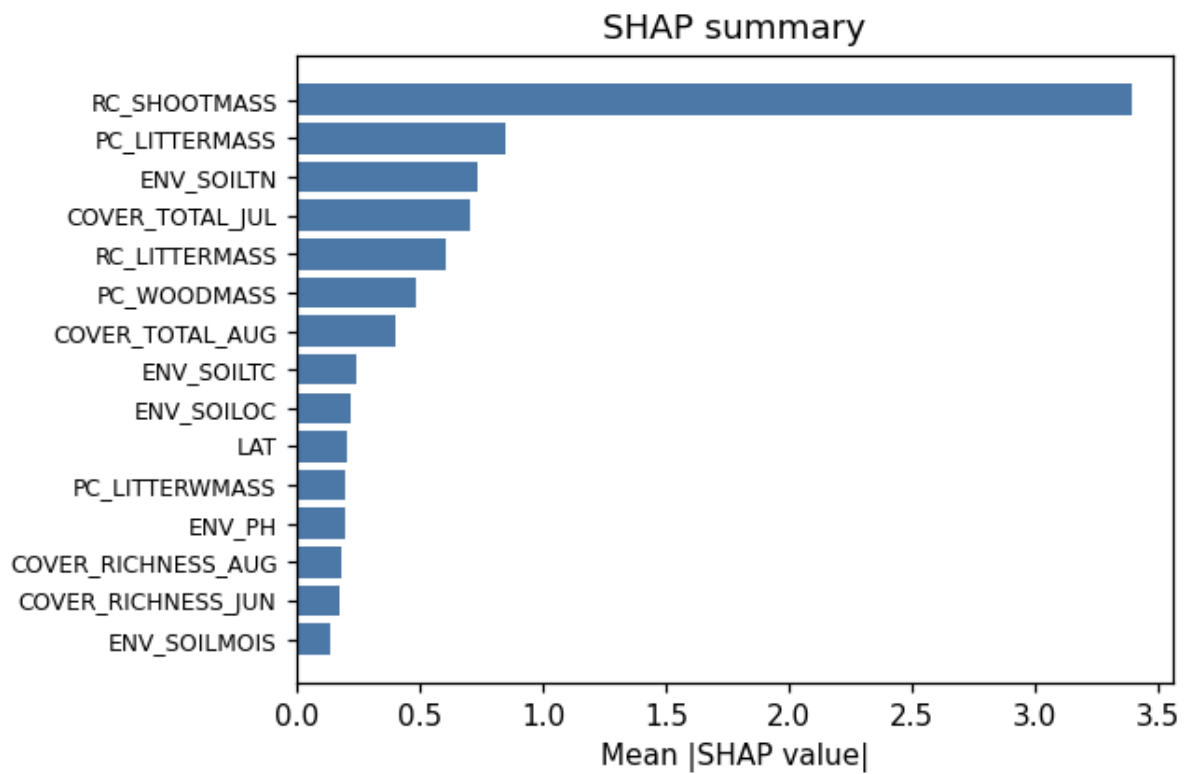
r2	mse
0.562705	44.270396
0.516368	40.052147
0.52441	36.052445



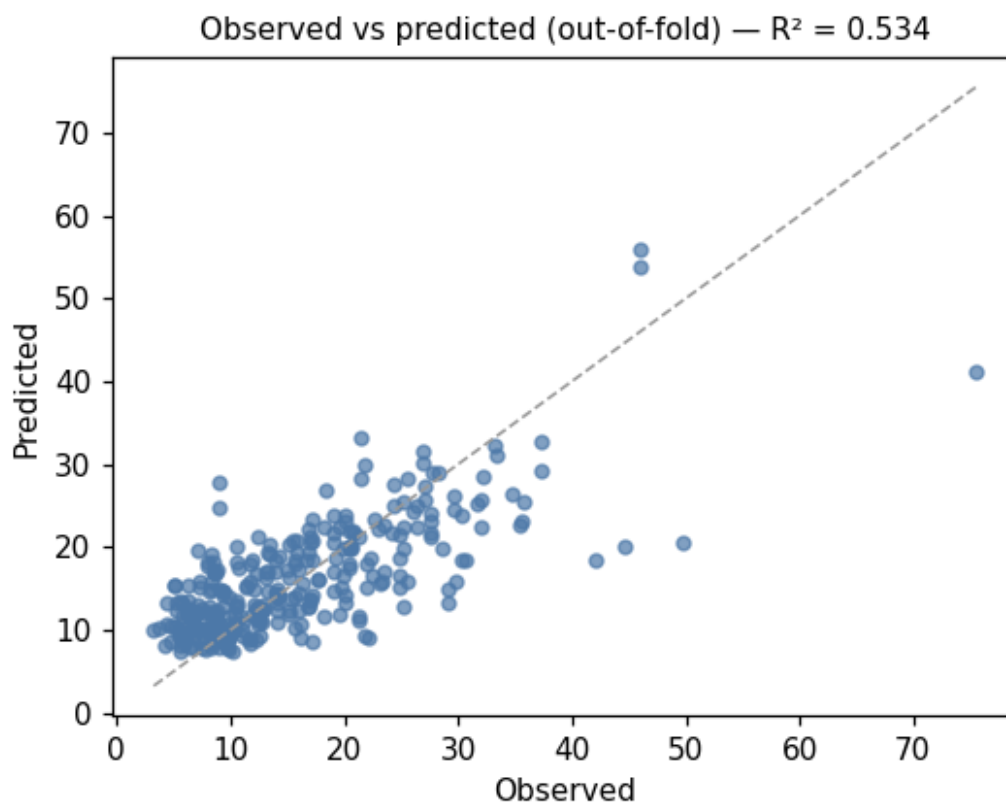
importance bar



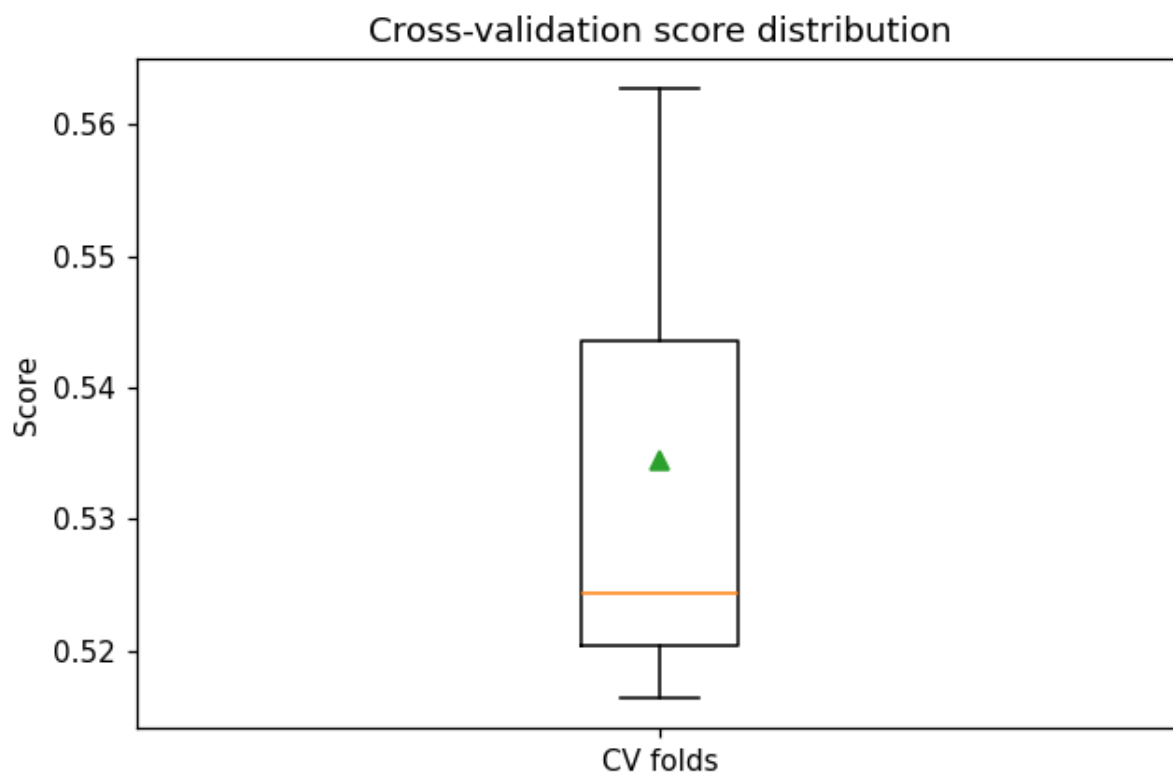
permutation bar



shap summary



pred vs obs



cv boxplot

Method & Provenance

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

engine	Python · scikit-learn 1.8.0
datasets	gallery — cat_g_field_data_plots.csv (preprocessed)
started	2026-09-21 02:11:43.291308
permanent link	https://cat-g-omxpert.wishartlab.com/results/by-token/4ae9de70c16e41858d39d893714

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