TabNet Interpretability Report

SHAP (KernelExplainer) + PDP / ICE

Classes: low, medium, high

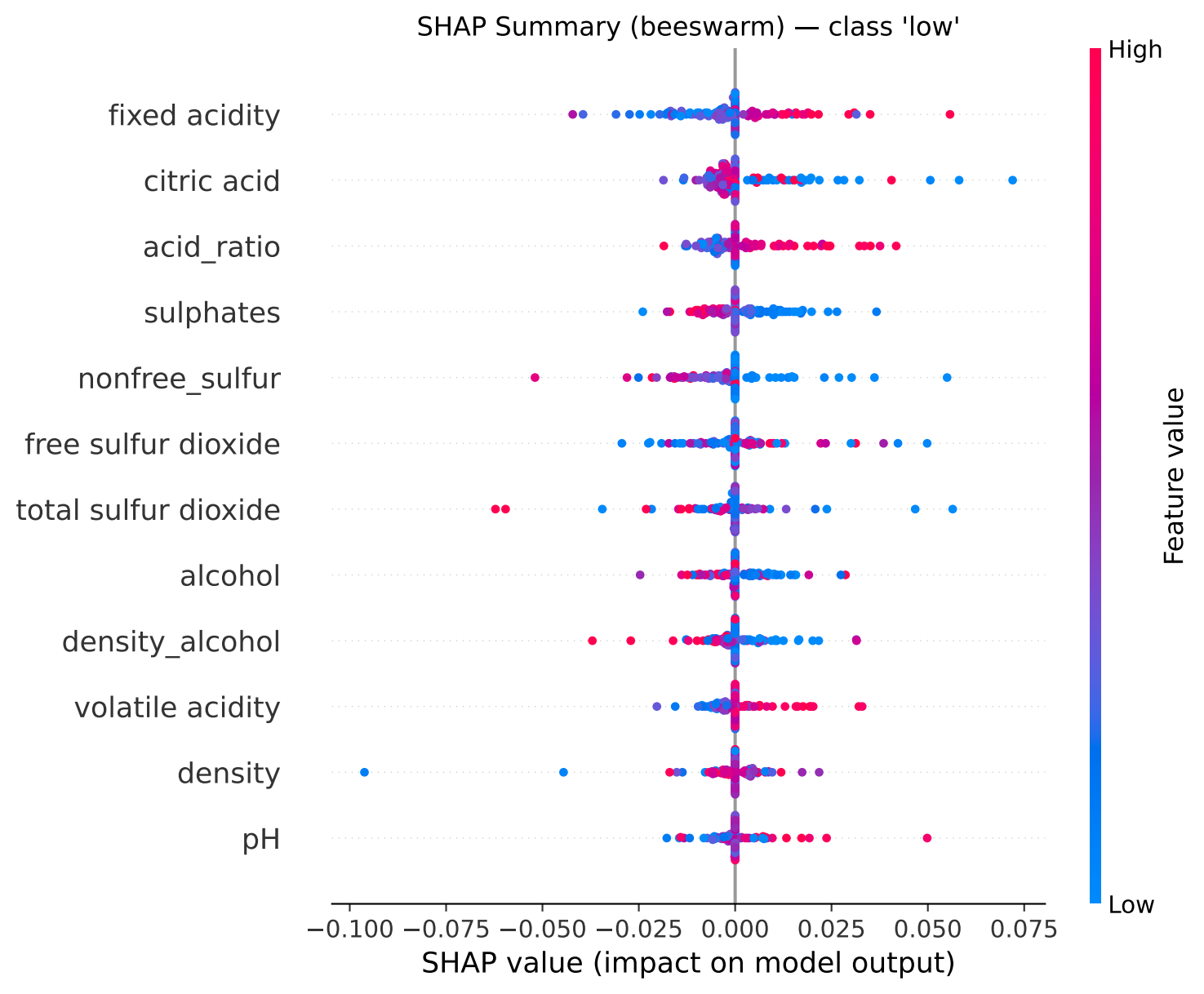
Top plotted features (by overall mean|SHAP|): sulphates, alcohol, citric acid, sulphur_ratio, density_alcohol, fixed acidity, residual sugar, free sulfur dioxide

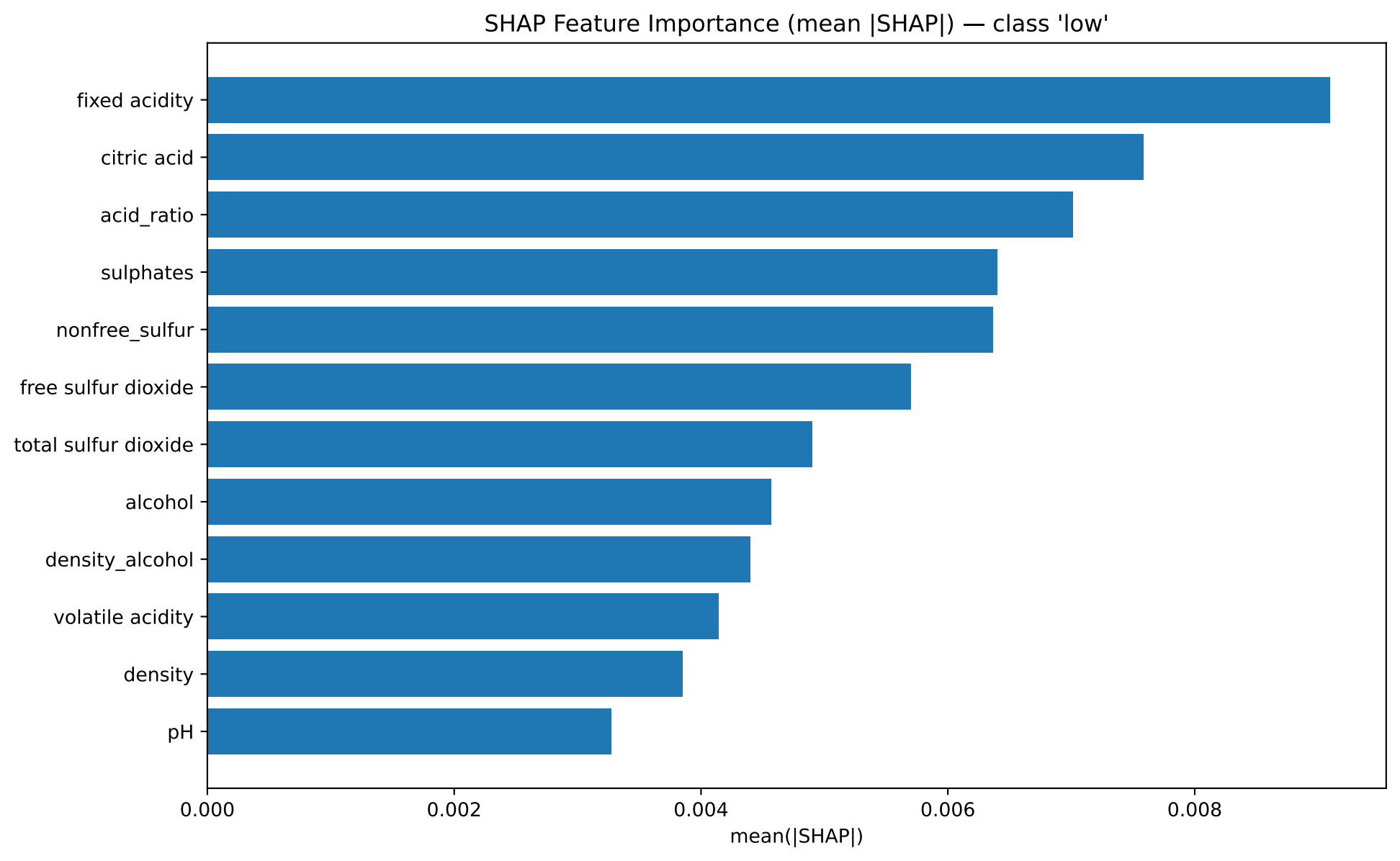
Performance Summary

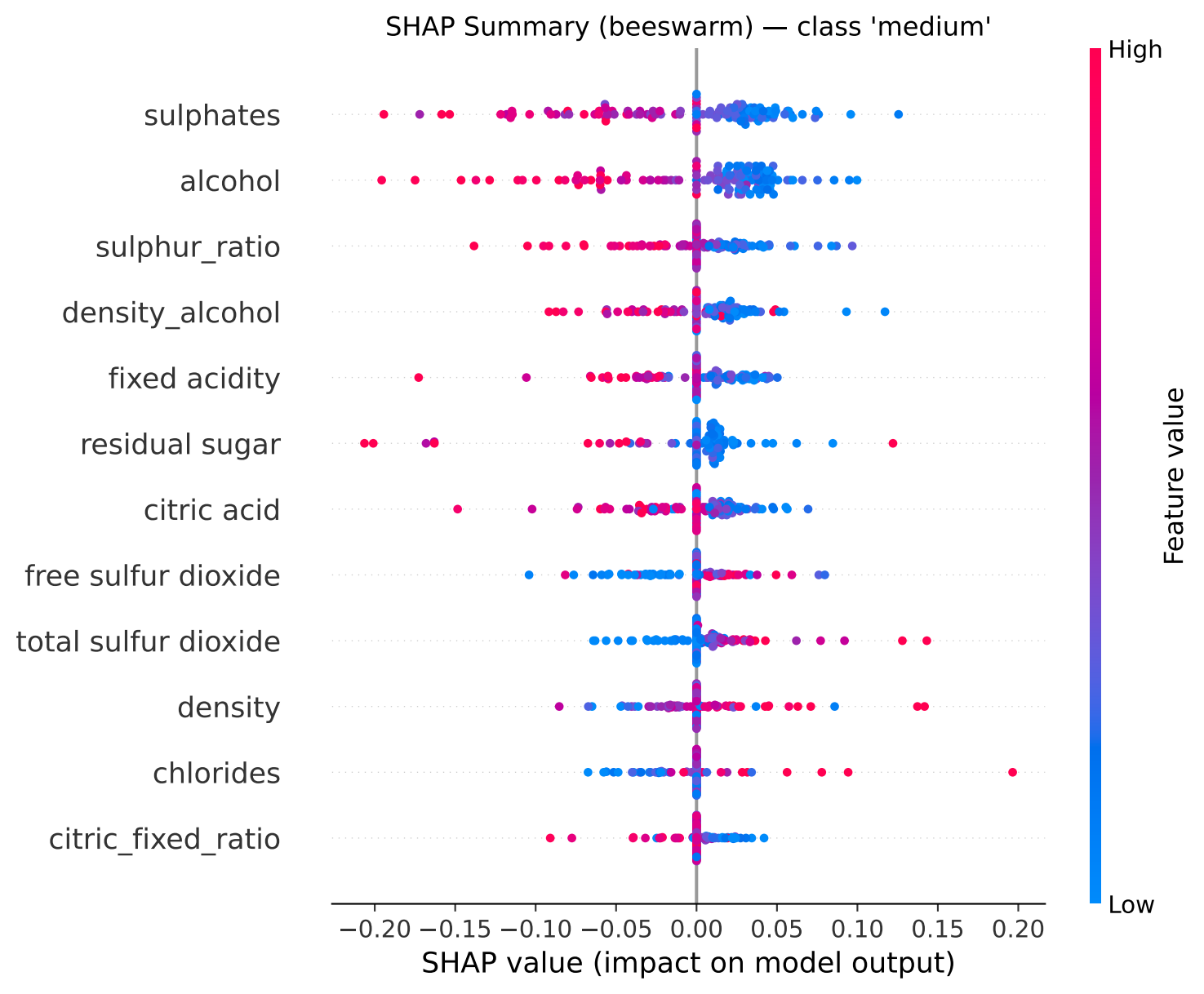
```
Accuracy:          0.8656
Balanced accuracy: 0.5428
Macro F1:          0.5628

Confusion matrix:
             pred_low  pred_medium  pred_high
true_low            1           12          0
true_medium         1          250         13
true_high           0           17         26
```

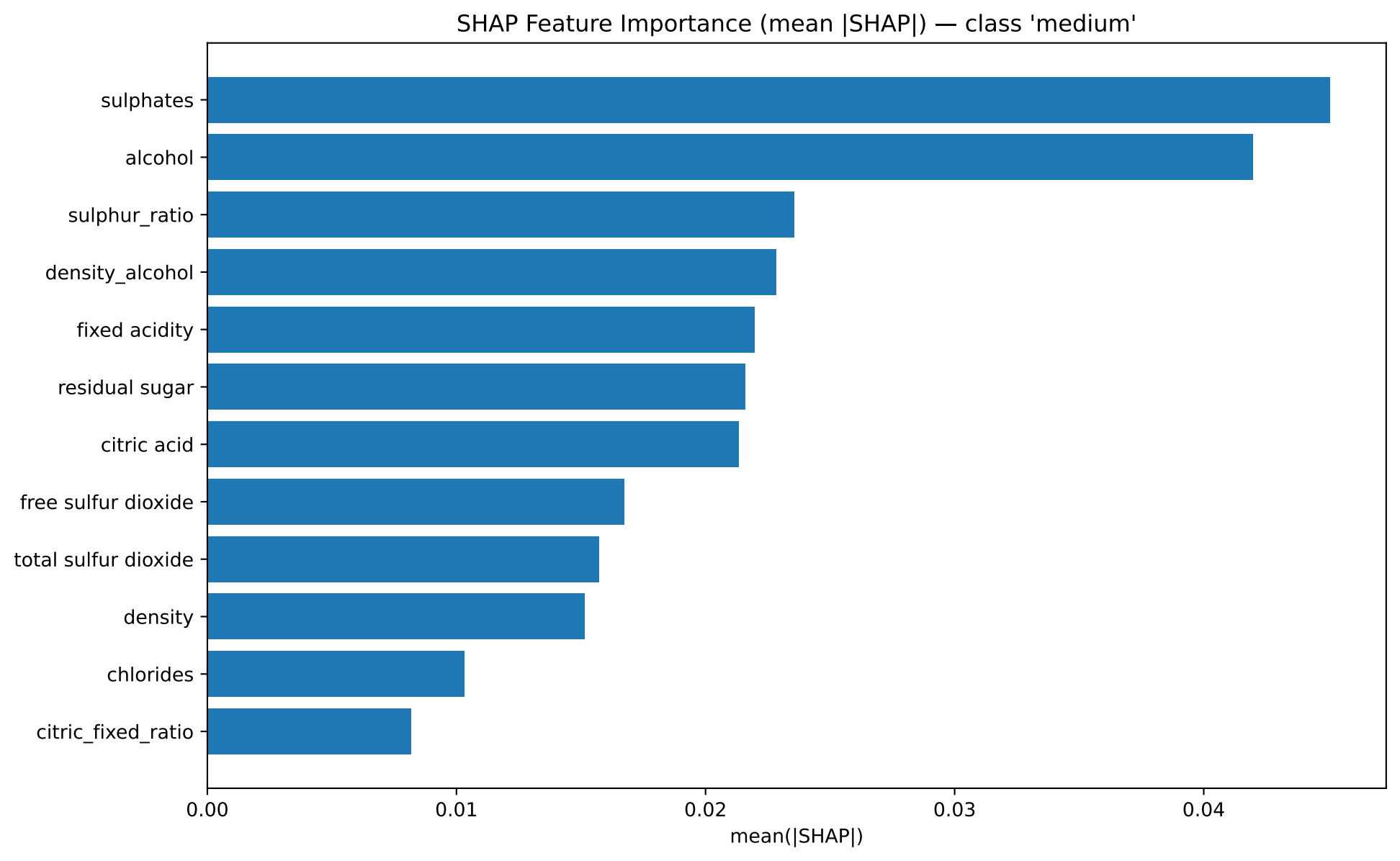
Top overall SHAP features (mean|SHAP| averaged across classes)

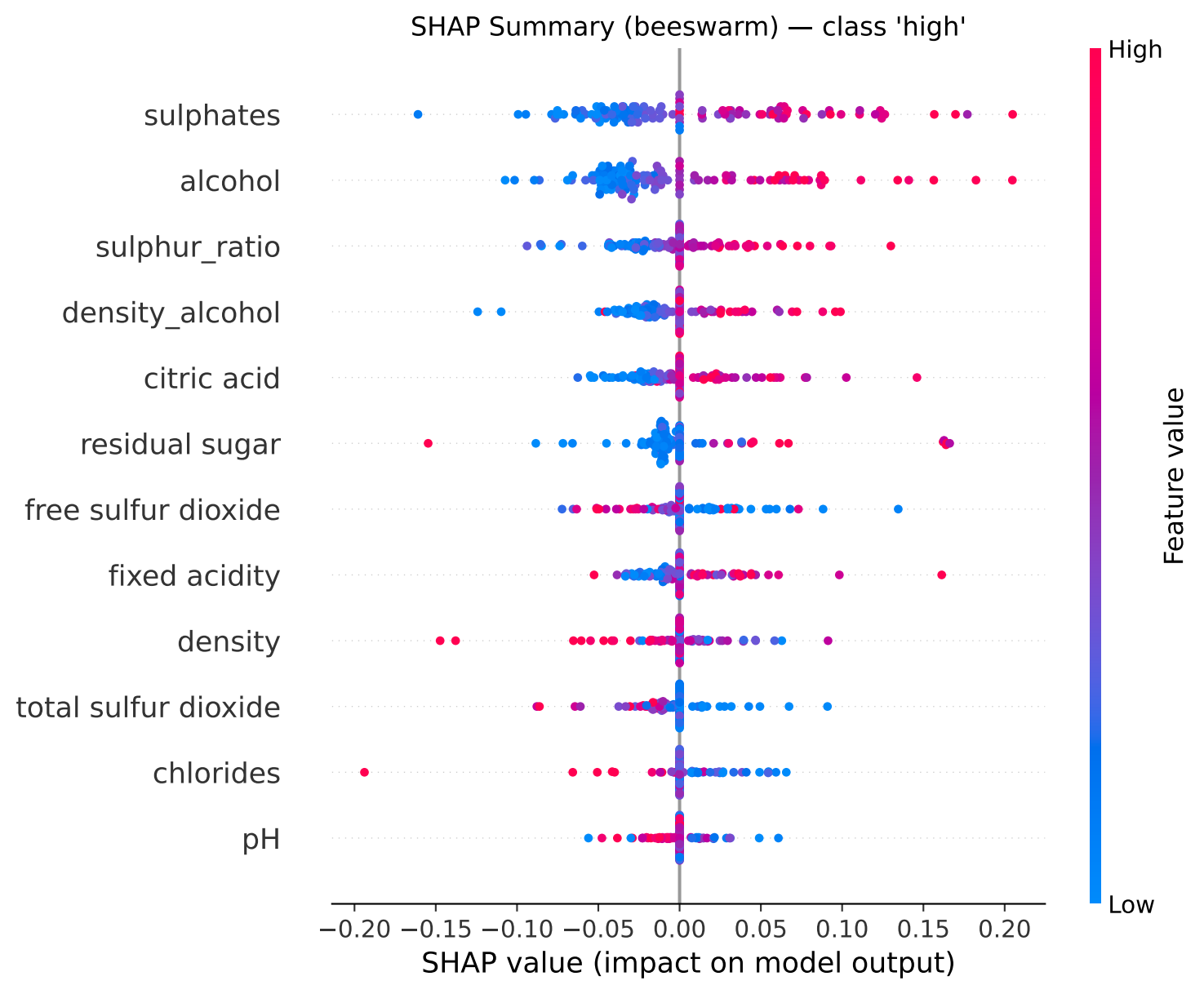
feature	mean_abs_shap_overall
sulphates	0.034285
alcohol	0.030360
citric acid	0.016981
sulphur_ratio	0.016801
density_alcohol	0.016791
fixed acidity	0.015577
residual sugar	0.015127
free sulfur dioxide	0.012912
total sulfur dioxide	0.011034
density	0.010575
chlorides	0.007521
citric_fixed_ratio	0.006002
acid_ratio	0.005814
pH	0.005647
volatile acidity	0.005468
nonfree_sulfur	0.005042
alc_sulphates	0.004197

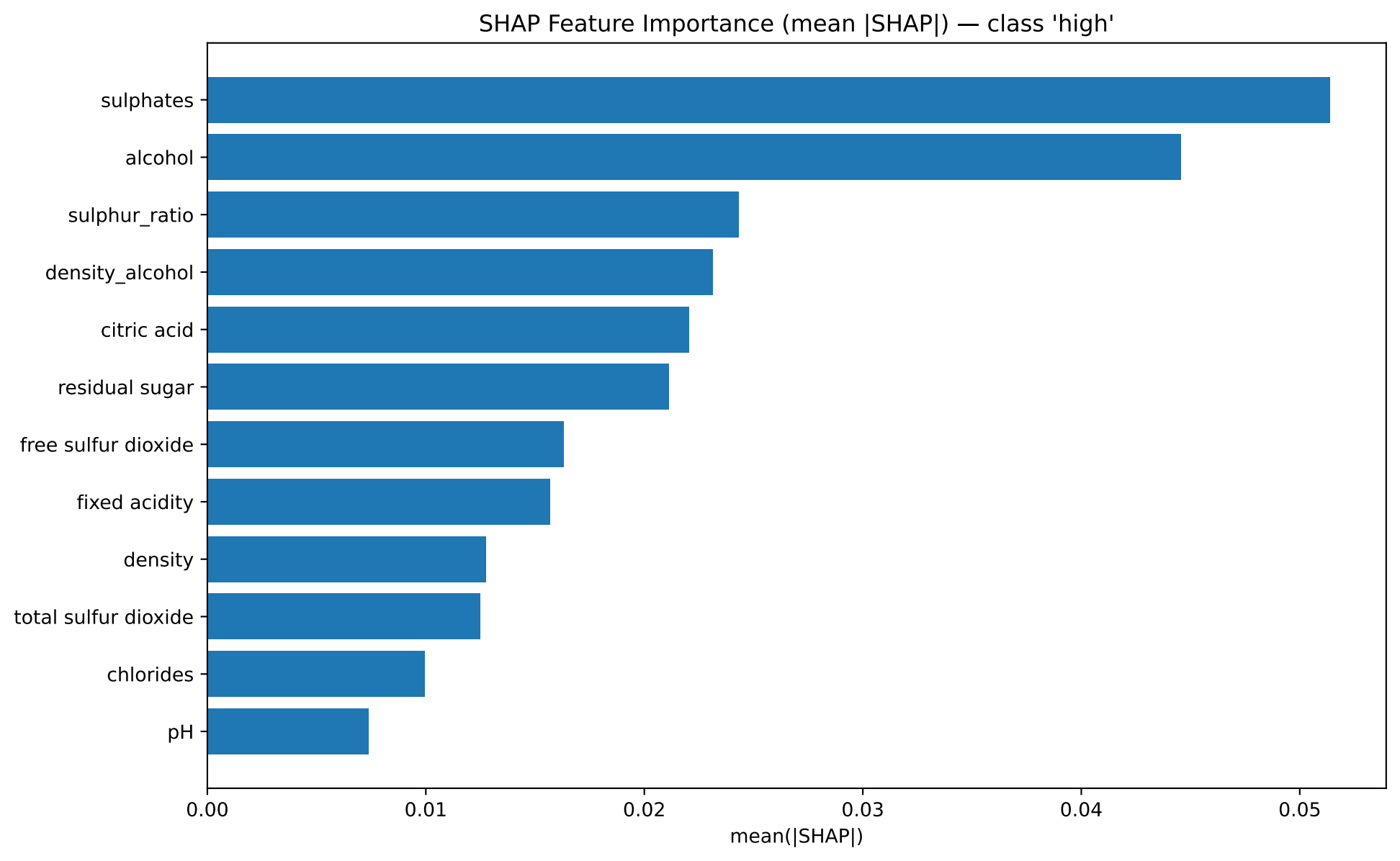

SHAP Summary (beeswarm) — class 'low'
fixed acidity
citric acid
acid_ratio
sulphates
nonfree_sulfur
free sulfur dioxide
total sulfur dioxide
alcohol
density_alcohol
volatile acidity
density
pH
−0.100 −0.075 −0.050 −0.025 0.000 0.025 0.050 0.075
SHAP value (impact on model output)
High
Low
Feature value

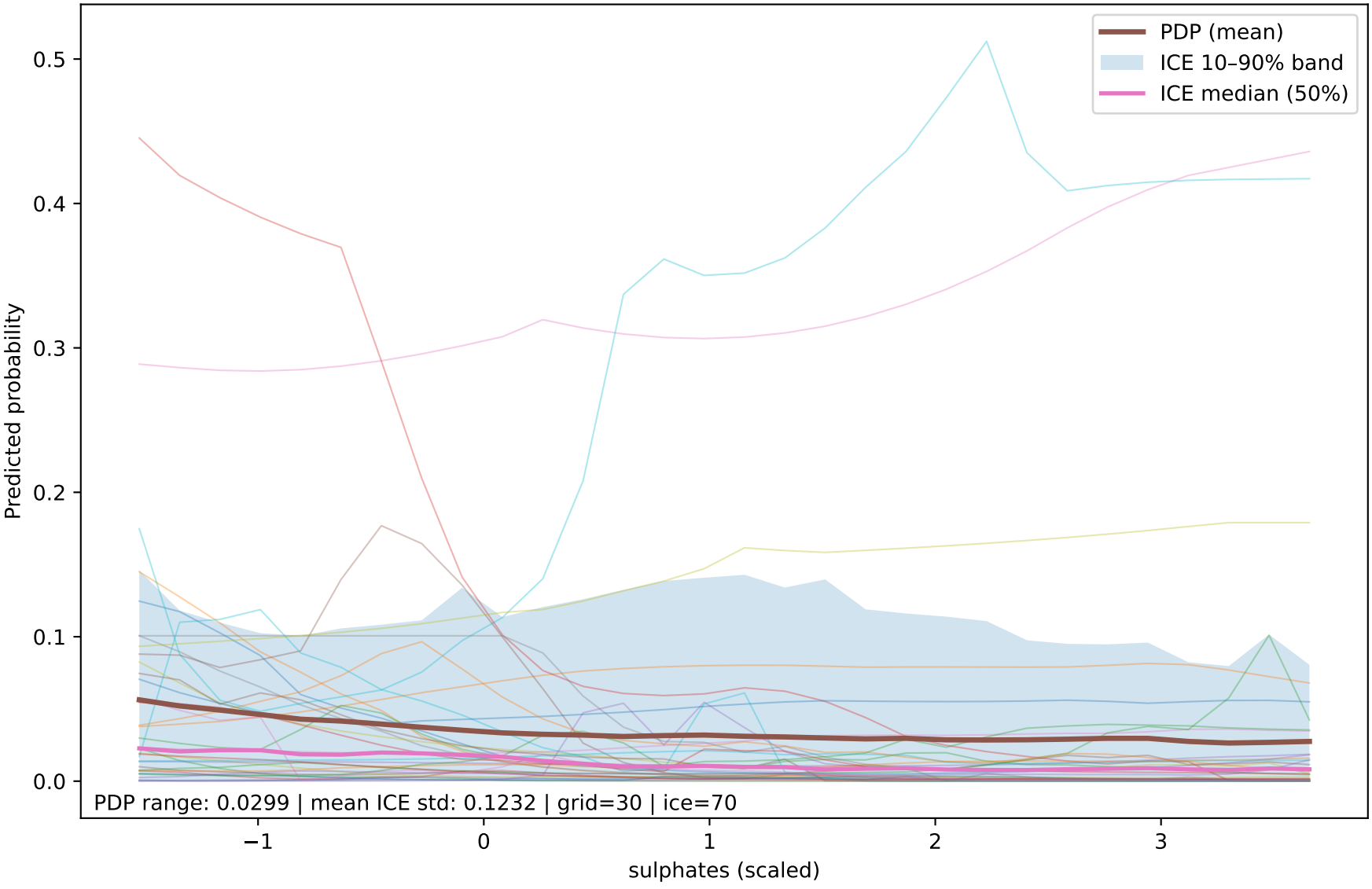

SHAP Feature Importance (mean |SHAP|) — class 'low'
fixed acidity
citric acid
acid_ratio
sulphates
nonfree_sulfur
free sulfur dioxide
total sulfur dioxide
alcohol
density_alcohol
volatile acidity
density
pH
0.000
0.002
0.004
0.006
0.008
mean(|SHAP|)

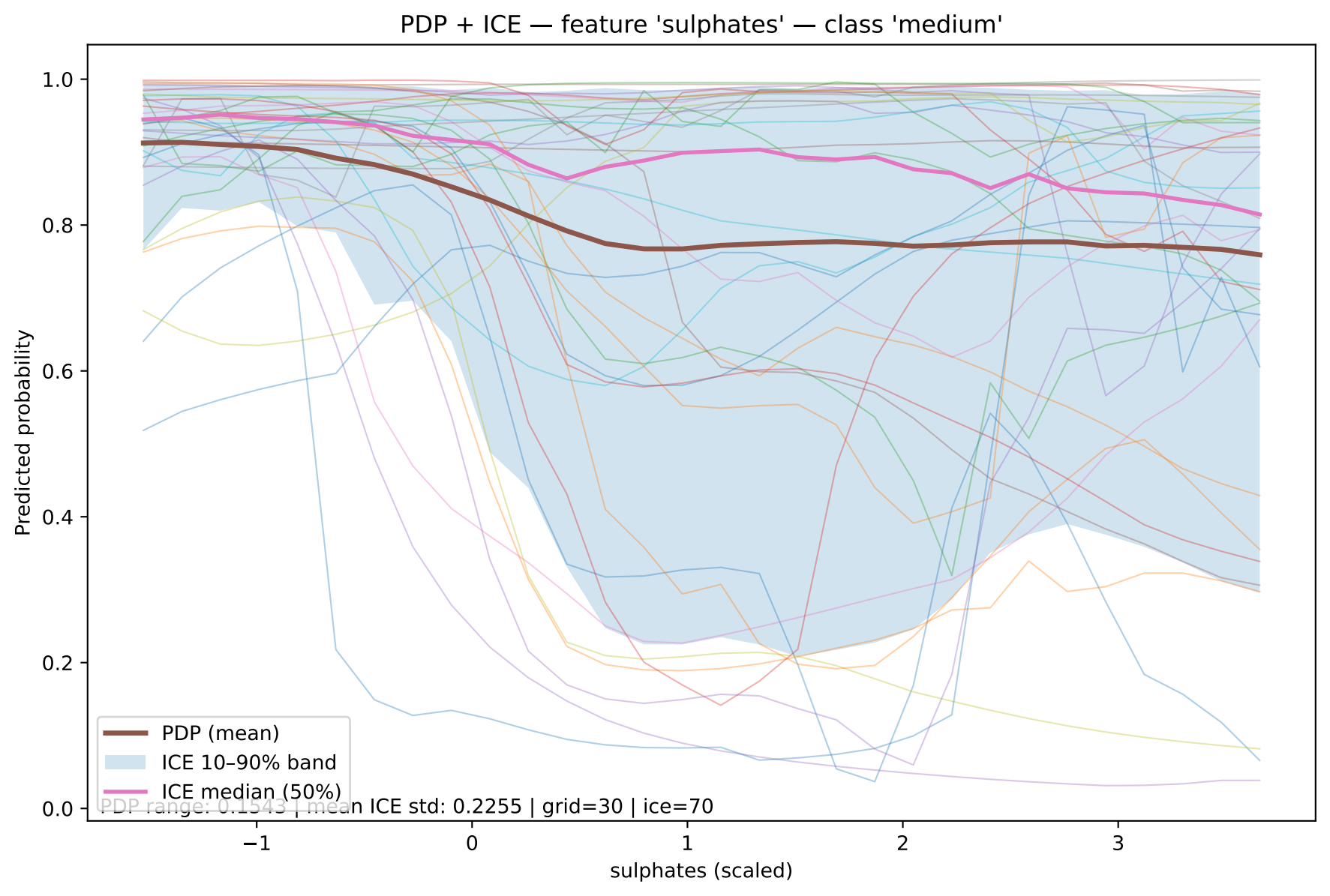

SHAP Summary (beeswarm) — class 'medium'
sulphates
alcohol
sulphur_ratio
density_alcohol
fixed acidity
residual sugar
citric acid
free sulfur dioxide
total sulfur dioxide
density
chlorides
citric_fixed_ratio
−0.20 −0.15 −0.10 −0.05 0.00 0.05 0.10 0.15 0.20
SHAP value (impact on model output)
High
Low
Feature value

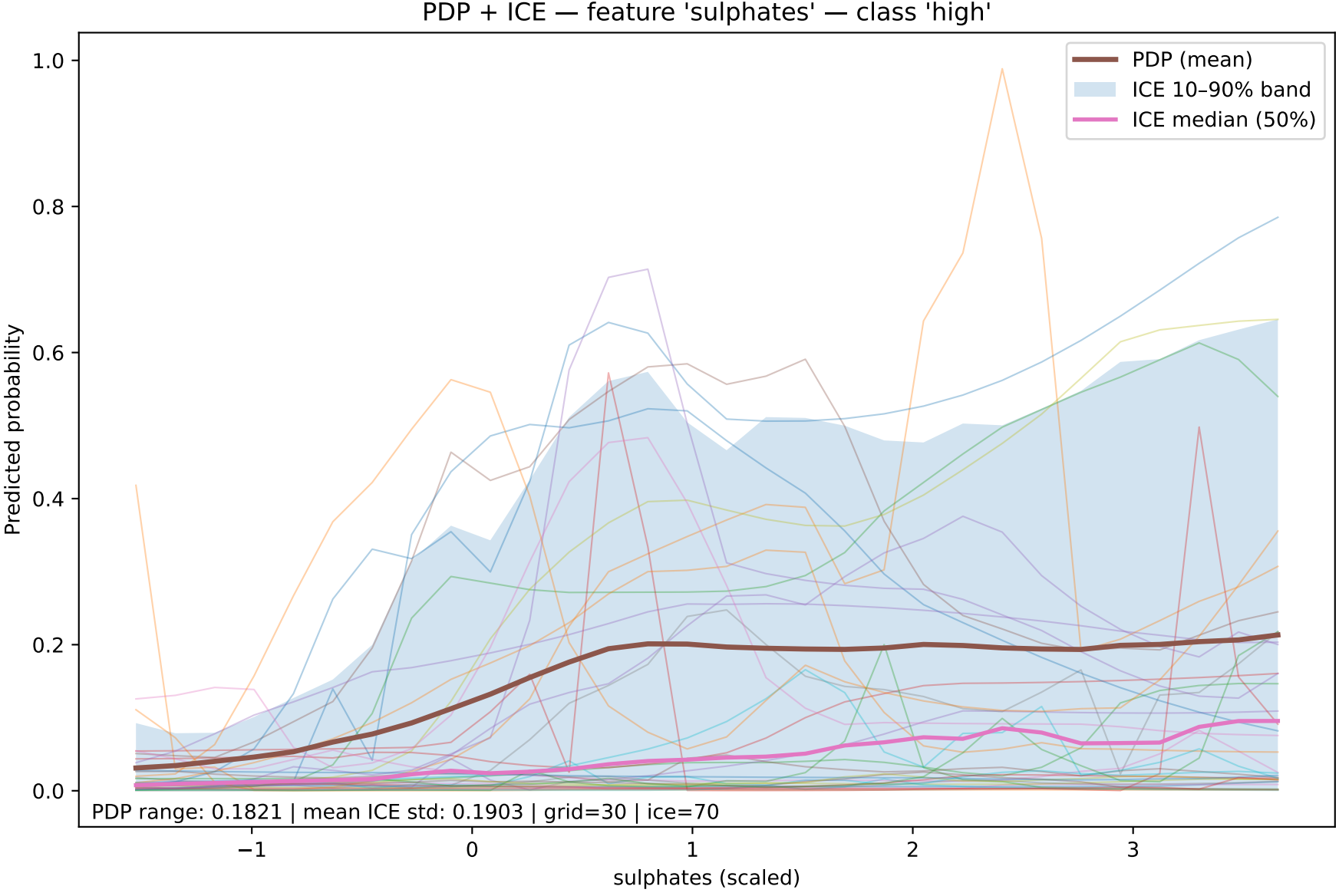

SHAP Feature Importance (mean |SHAP|) — class 'medium'
sulphates
alcohol
sulphur_ratio
density_alcohol
fixed acidity
residual sugar
citric acid
free sulfur dioxide
total sulfur dioxide
density
chlorides
citric_fixed_ratio
0.00
0.01
0.02
0.03
0.04
mean(|SHAP|)

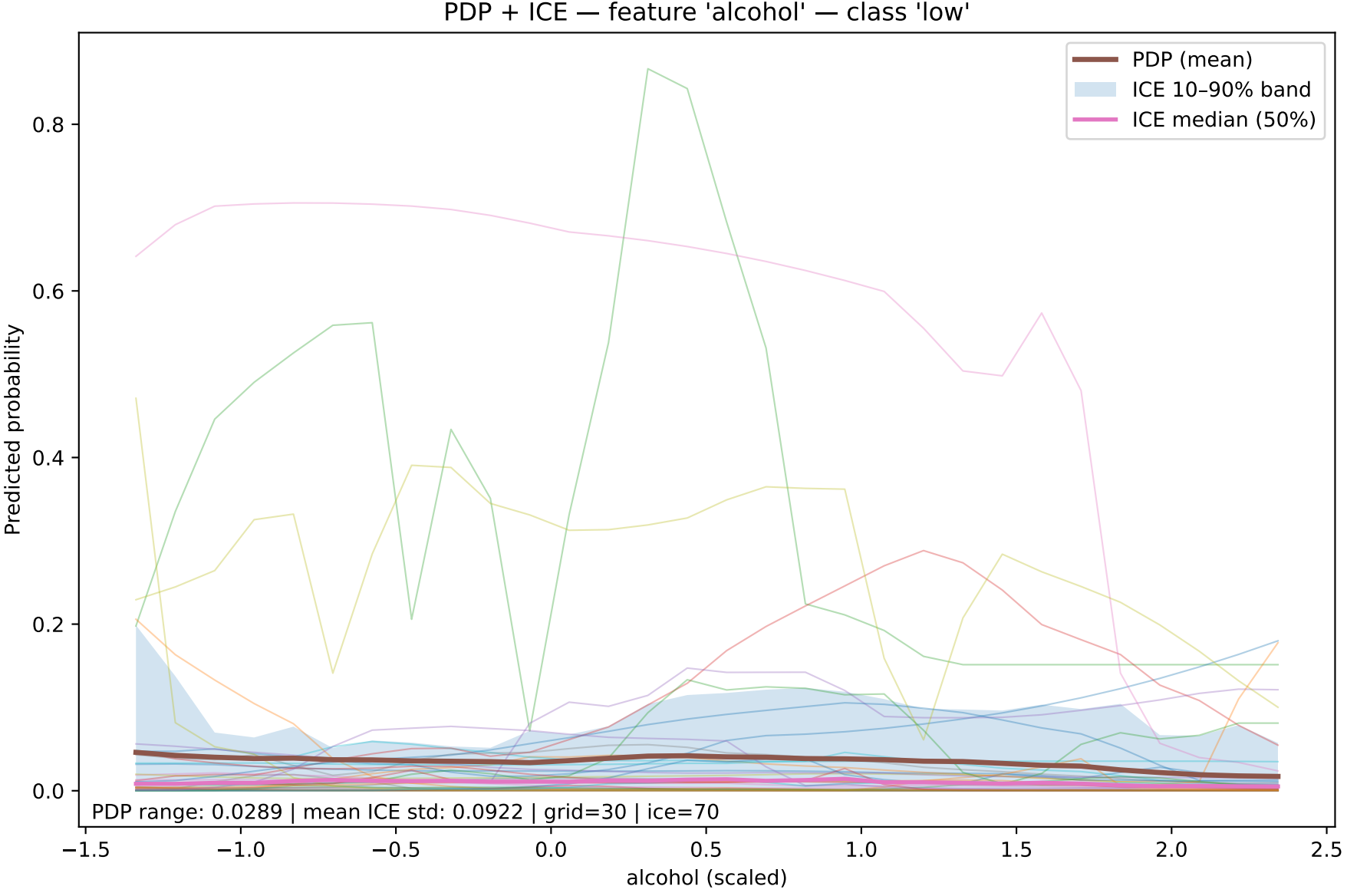

SHAP Summary (beeswarm) — class 'high'
sulphates
alcohol
sulphur_ratio
density_alcohol
citric acid
residual sugar
free sulfur dioxide
fixed acidity
density
total sulfur dioxide
chlorides
pH
−0.20 −0.15 −0.10 −0.05 0.00 0.05 0.10 0.15 0.20
SHAP value (impact on model output)
High
Feature value
Low

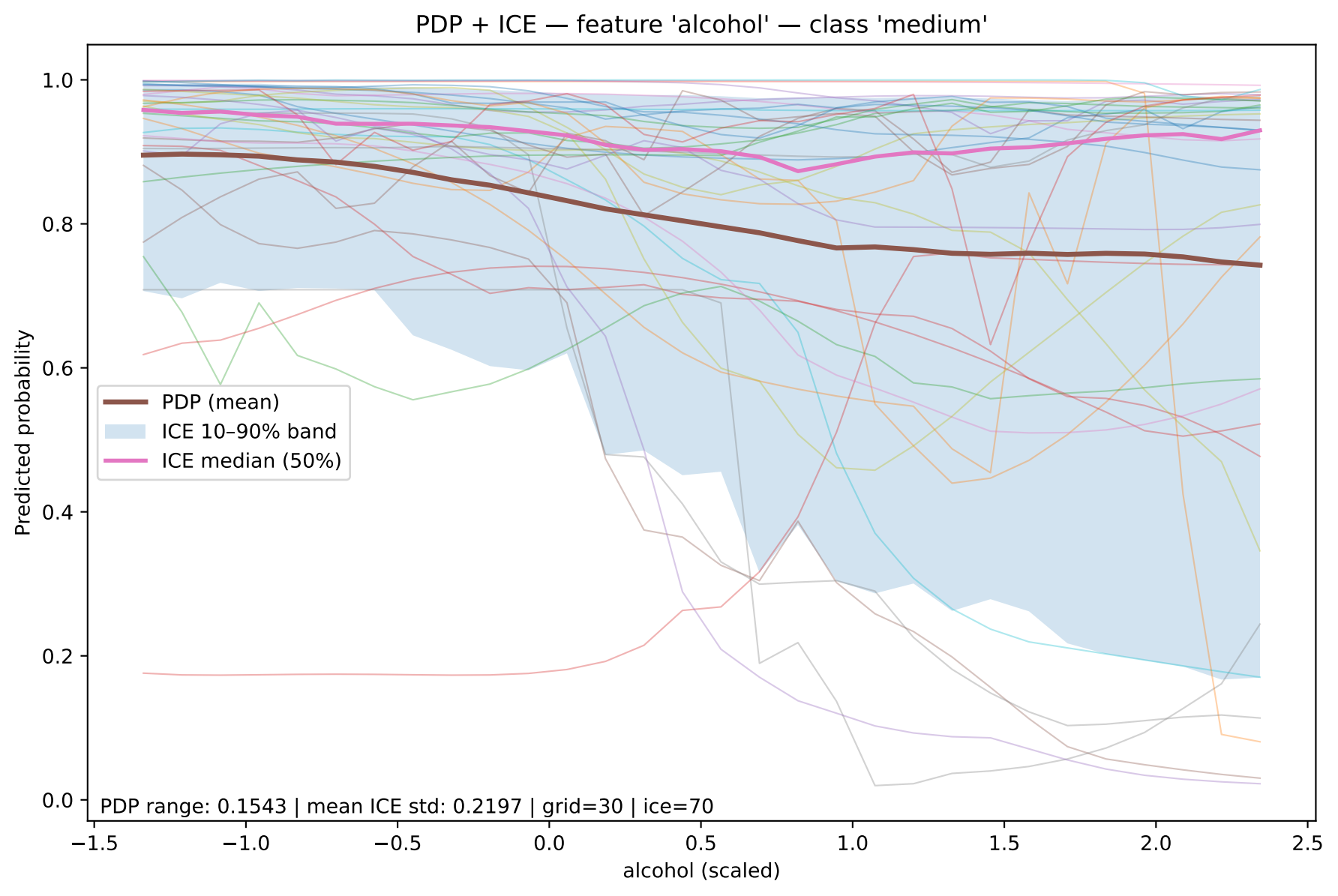

SHAP Feature Importance (mean |SHAP|) — class 'high'
sulphates
alcohol
sulphur_ratio
density_alcohol
citric acid
residual sugar
free sulfur dioxide
fixed acidity
density
total sulfur dioxide
chlorides
pH
0.00
0.01
0.02
0.03
0.04
0.05
mean(|SHAP|)

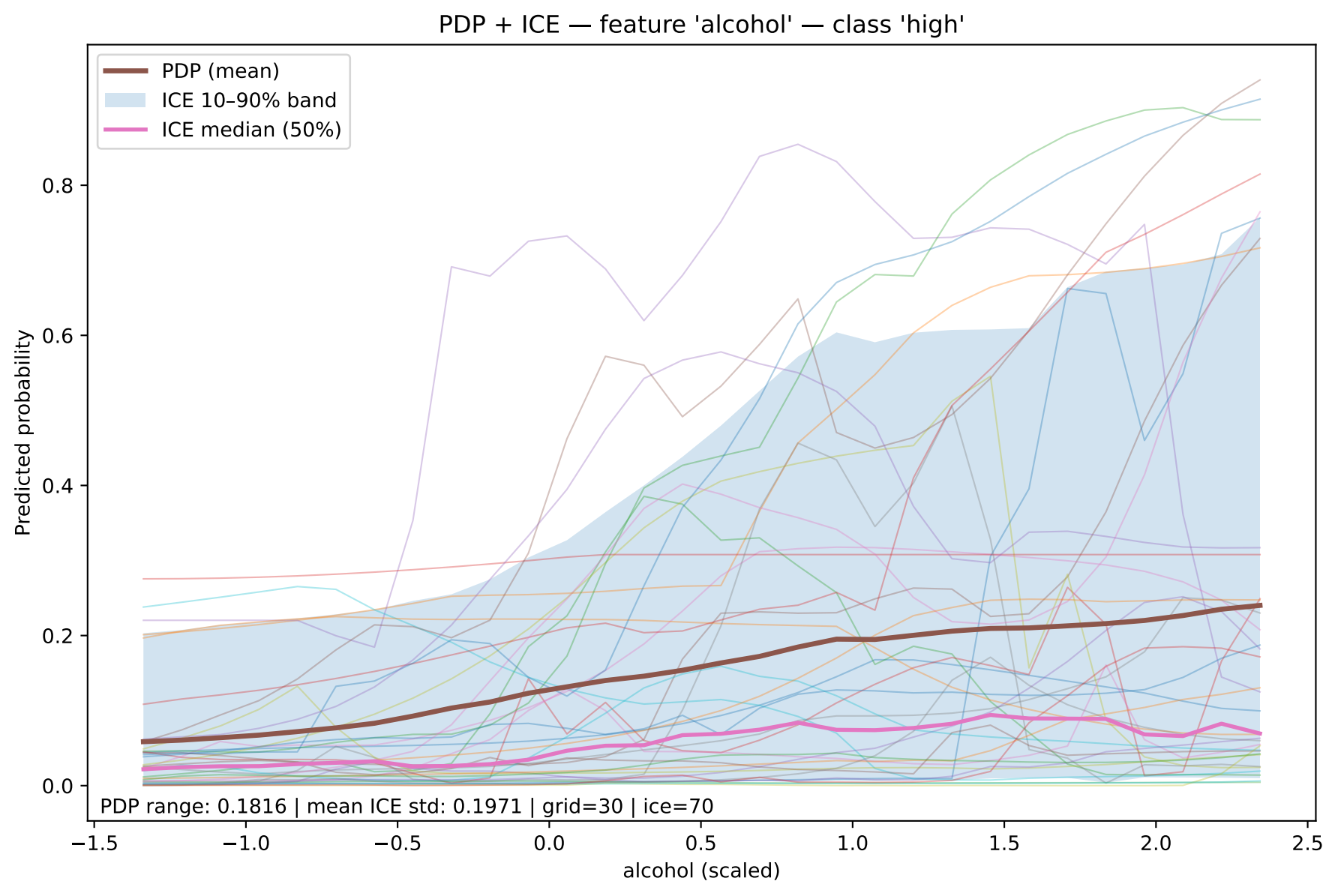

PDP + ICE — feature 'sulphates' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.1
0.2
0.3
0.4
0.5
−1
0
1
2
3
sulphates (scaled)
PDP range: 0.0299 | mean ICE std: 0.1232 | grid=30 | ice=70

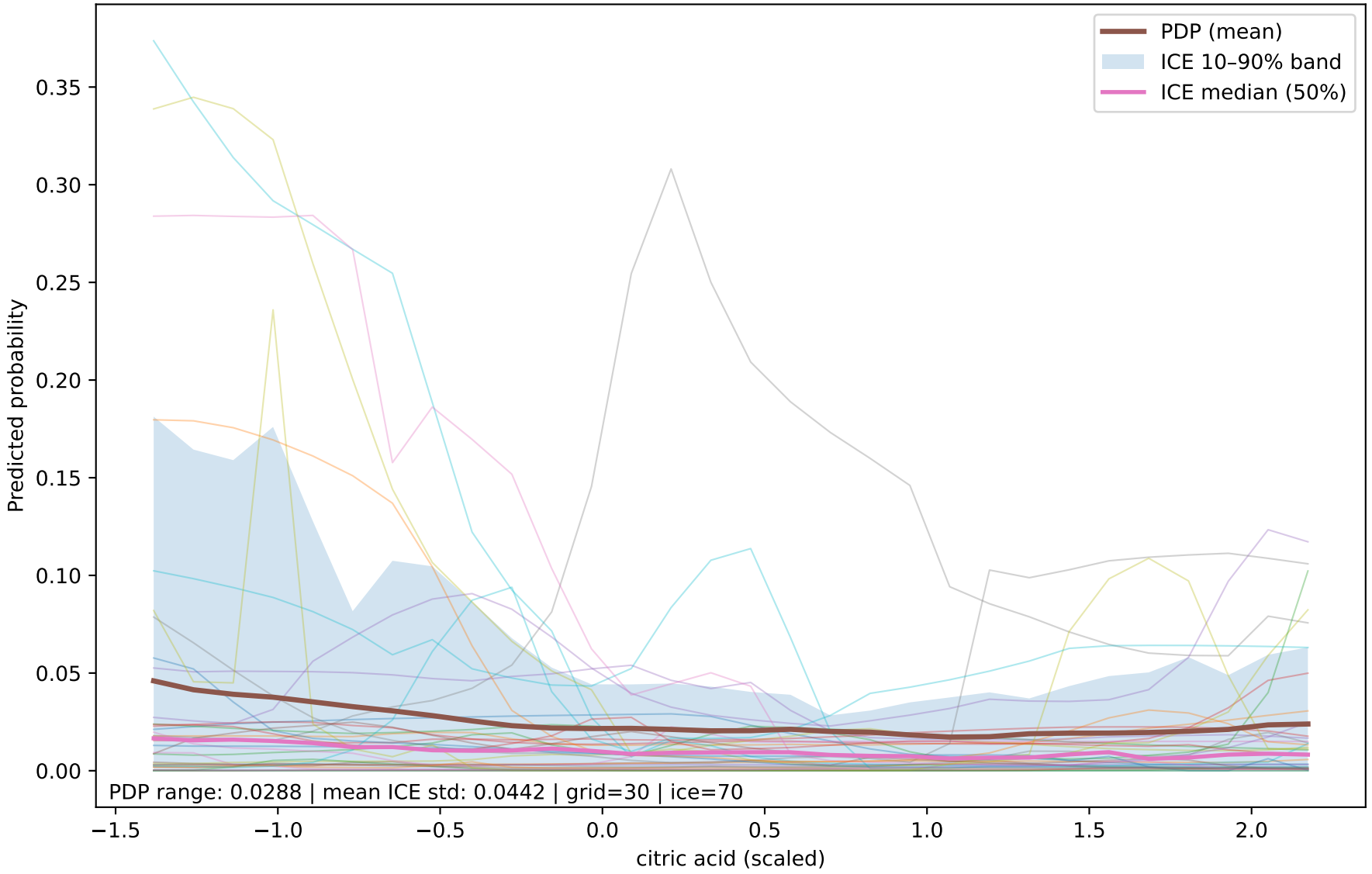

PDP + ICE — feature 'sulphates' — class 'medium'
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
−1
0
1
2
3
sulphates (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.1543 | mean ICE std: 0.2255 | grid=30 | ice=70

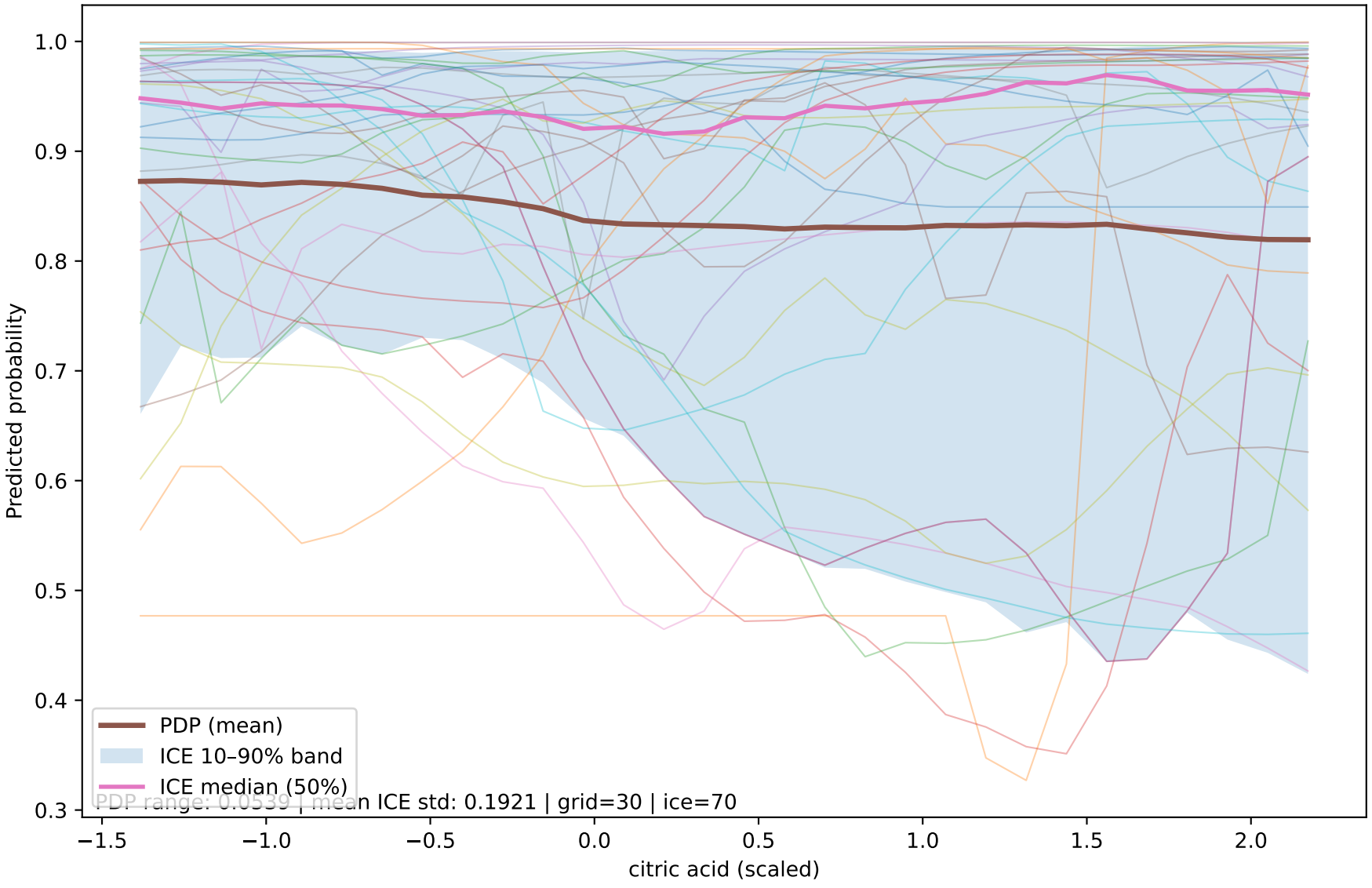

PDP + ICE — feature 'sulphates' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
PDP range: 0.1821 | mean ICE std: 0.1903 | grid=30 | ice=70
−1
0
1
2
3
sulphates (scaled)

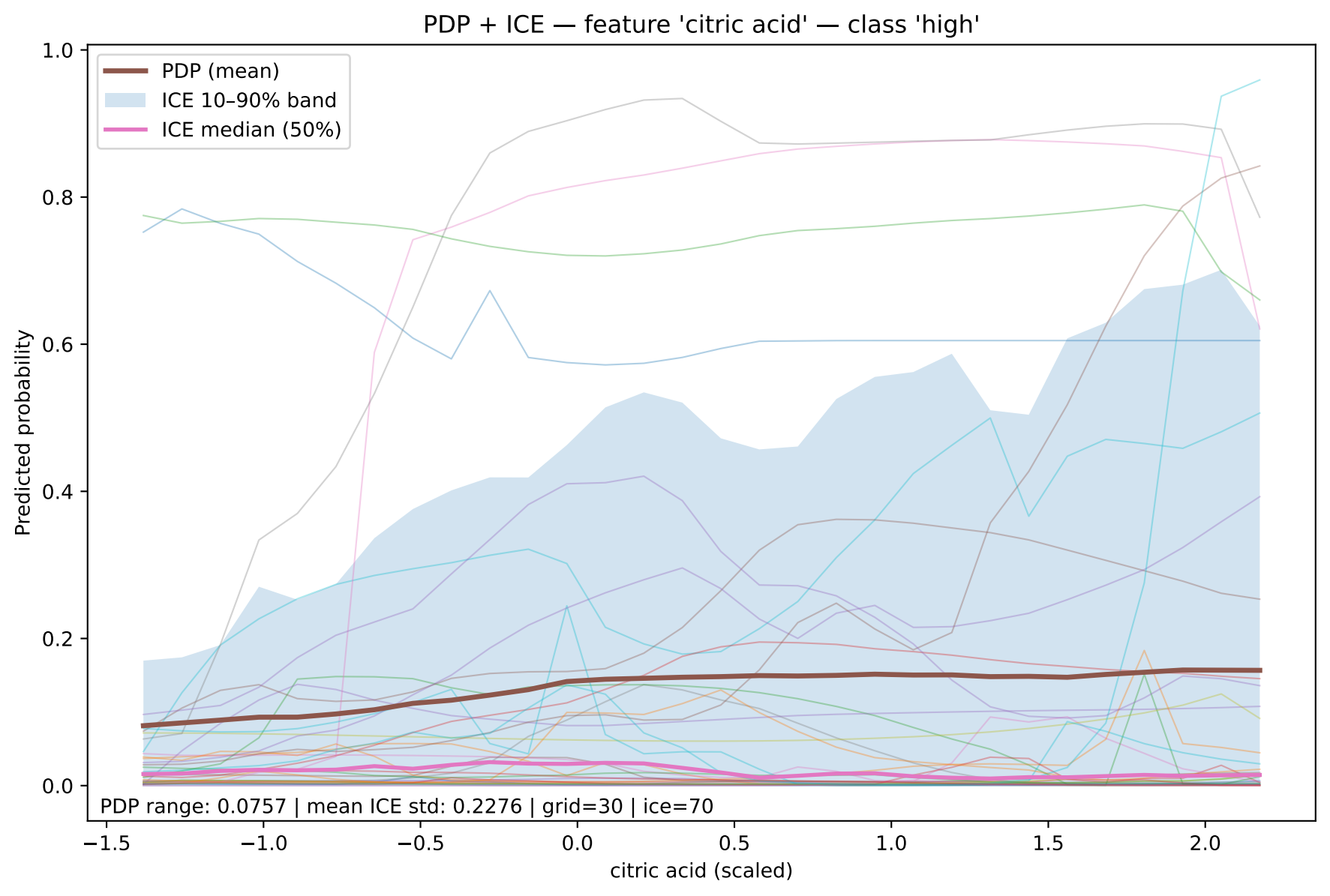

PDP + ICE — feature 'alcohol' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
PDP range: 0.0289 | mean ICE std: 0.0922 | grid=30 | ice=70
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
alcohol (scaled)

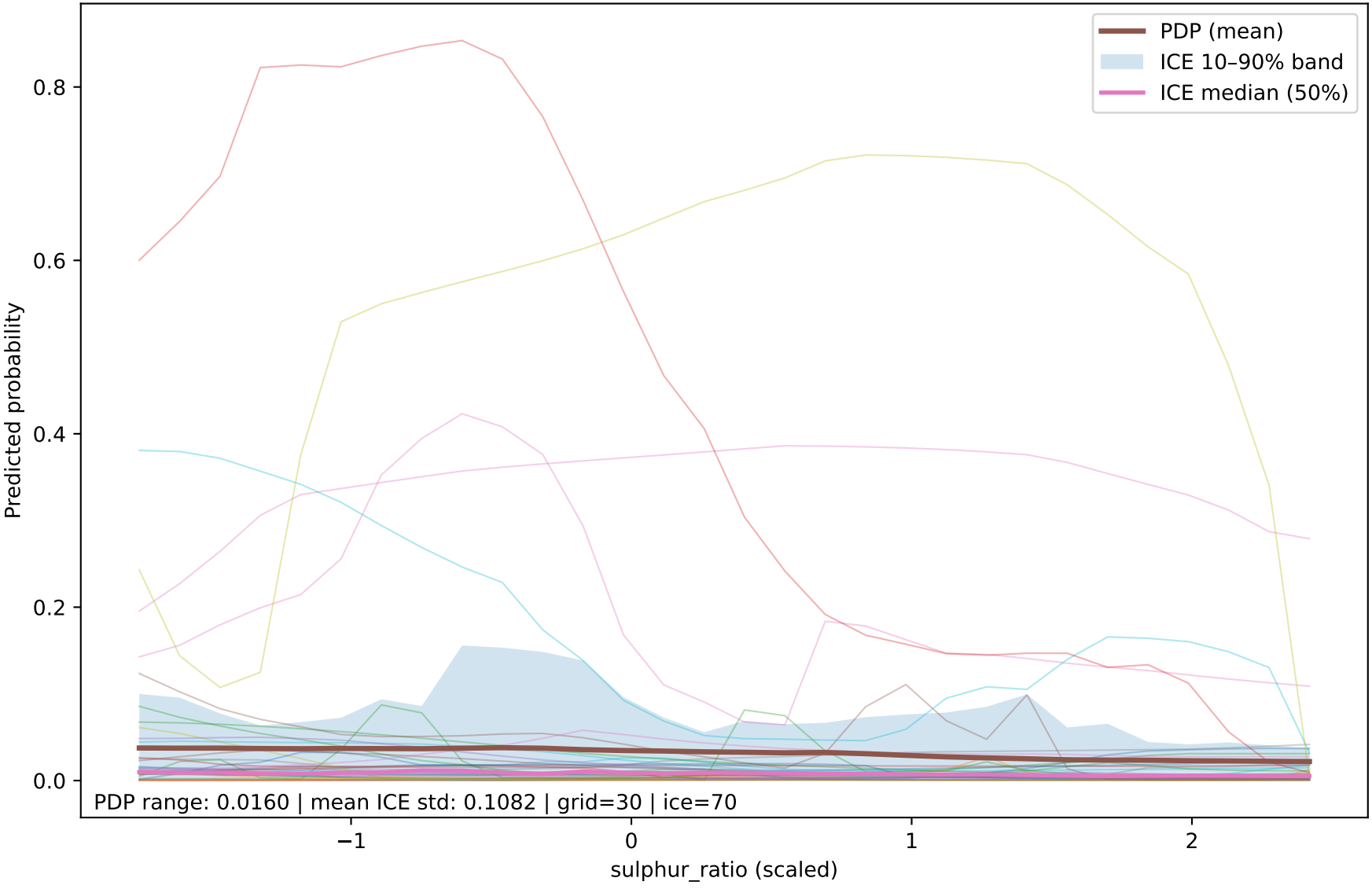

PDP + ICE — feature 'alcohol' — class 'medium'
Predicted probability
alcohol (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.1543 | mean ICE std: 0.2197 | grid=30 | ice=70
1.0
0.8
0.6
0.4
0.2
0.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5

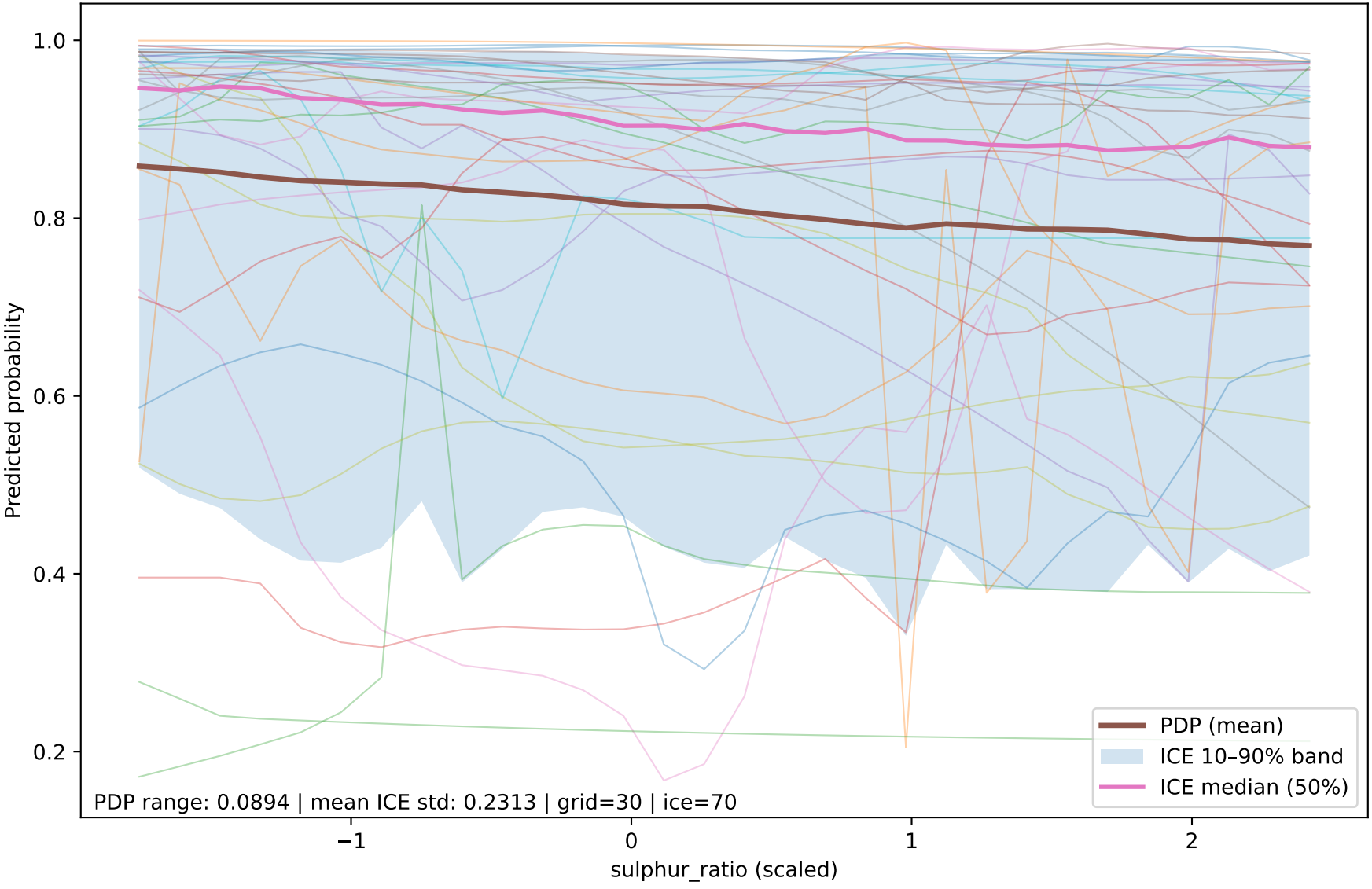

PDP + ICE — feature 'alcohol' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
PDP range: 0.1816 | mean ICE std: 0.1971 | grid=30 | ice=70
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
alcohol (scaled)

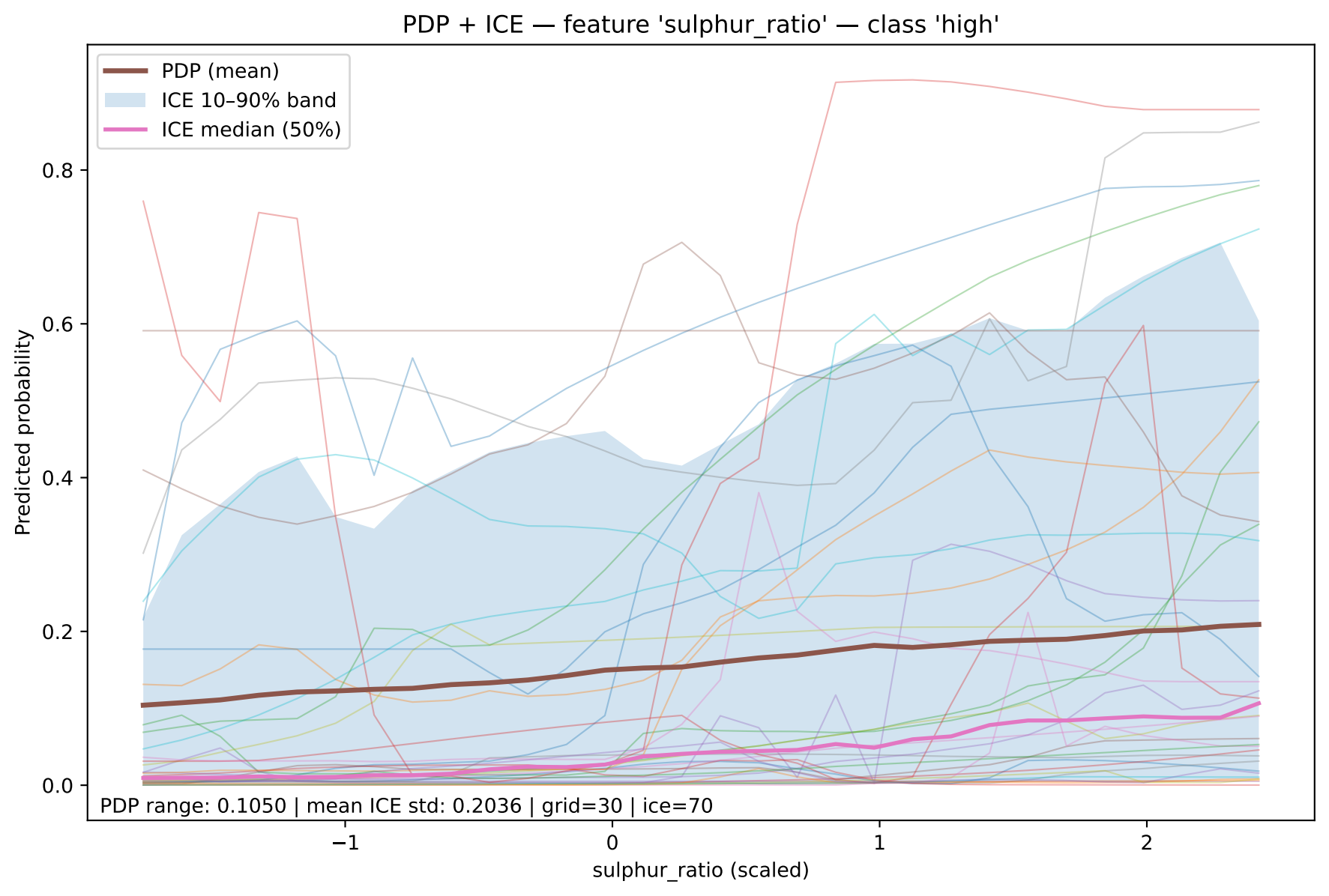

PDP + ICE — feature 'citric acid' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
PDP range: 0.0288 | mean ICE std: 0.0442 | grid=30 | ice=70
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
citric acid (scaled)

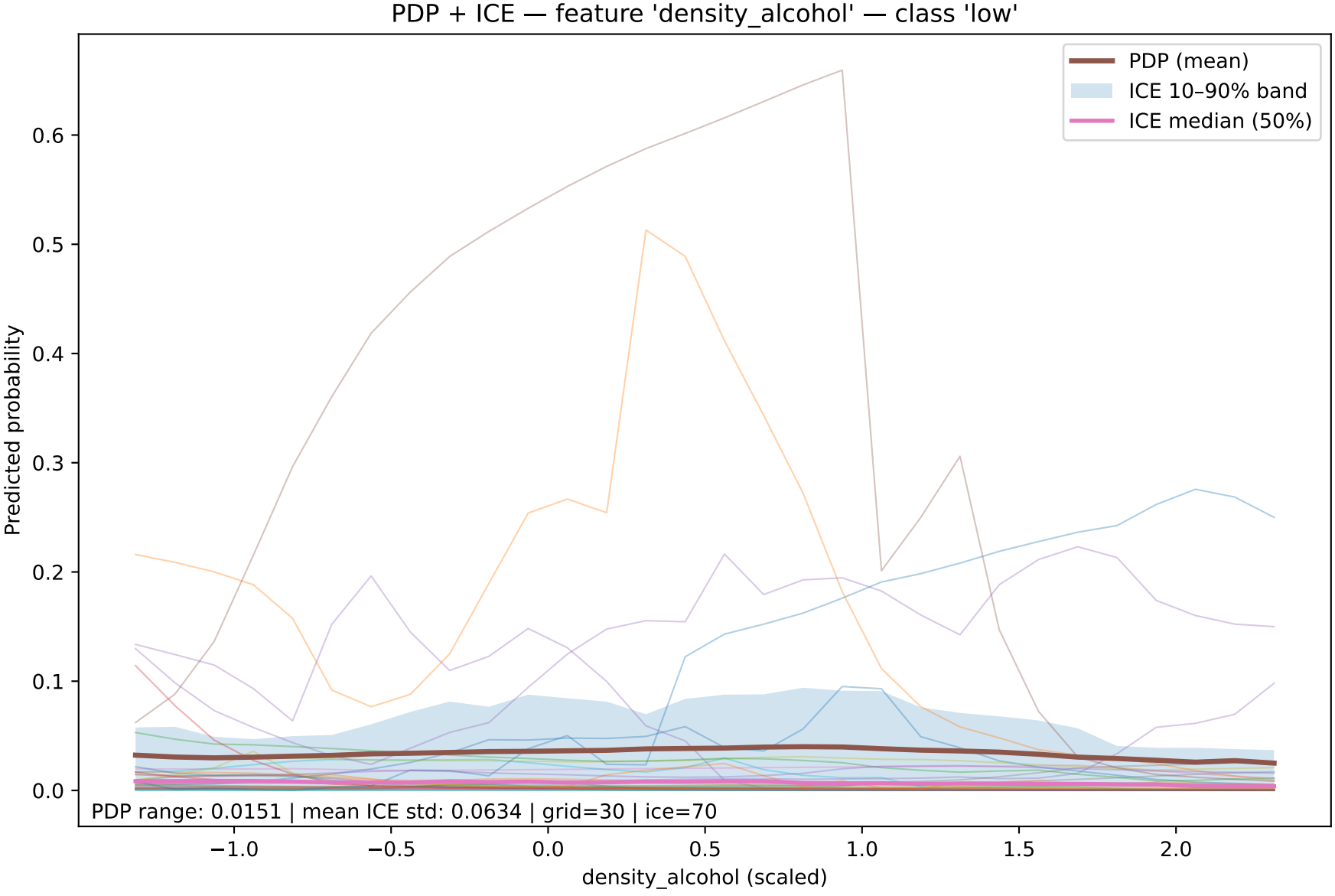

PDP + ICE — feature 'citric acid' — class 'medium'
Predicted probability
citric acid (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.0539 | mean ICE std: 0.1921 | grid=30 | ice=70

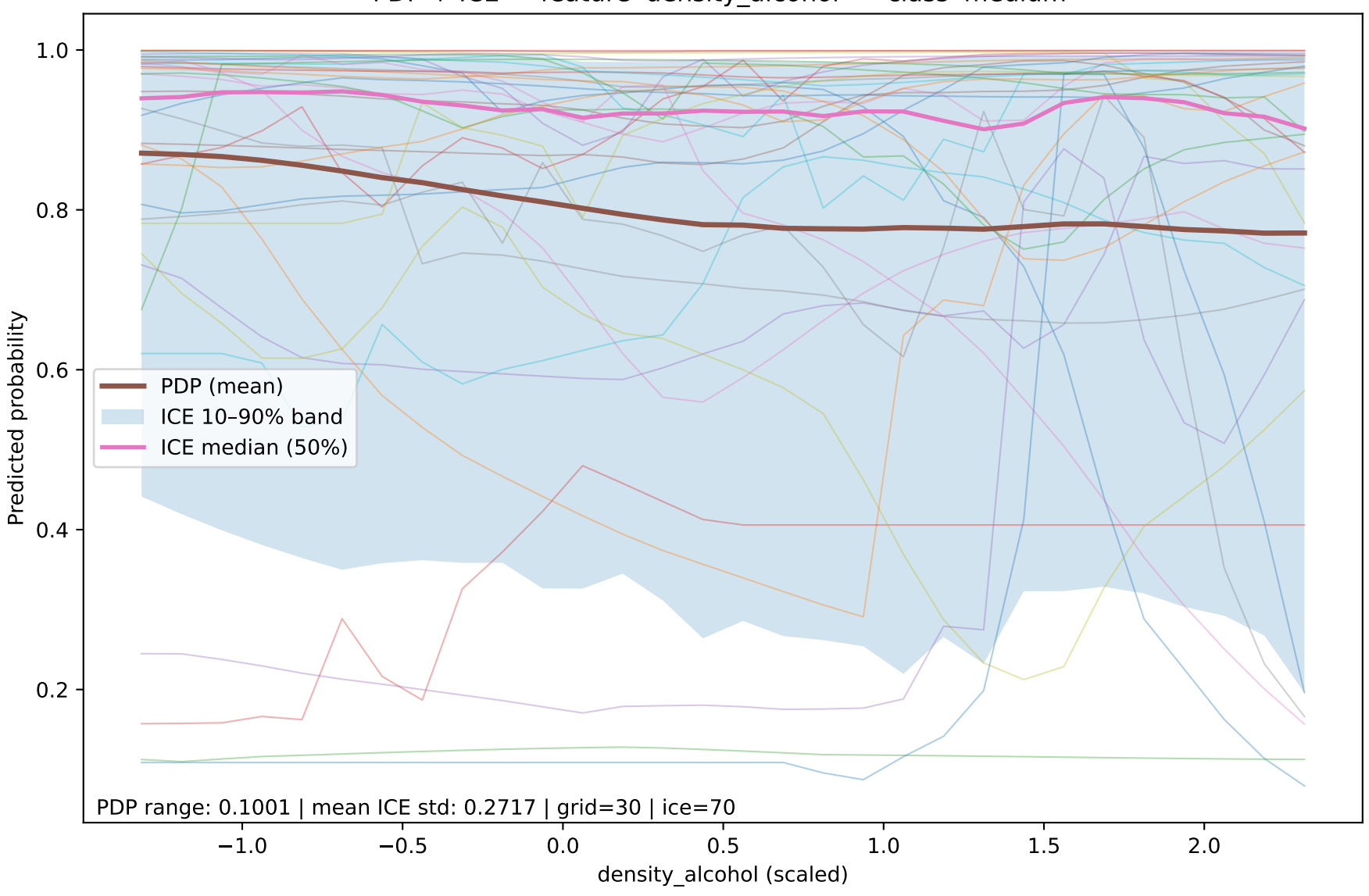

PDP + ICE — feature 'citric acid' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
PDP range: 0.0757 | mean ICE std: 0.2276 | grid=30 | ice=70
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
citric acid (scaled)

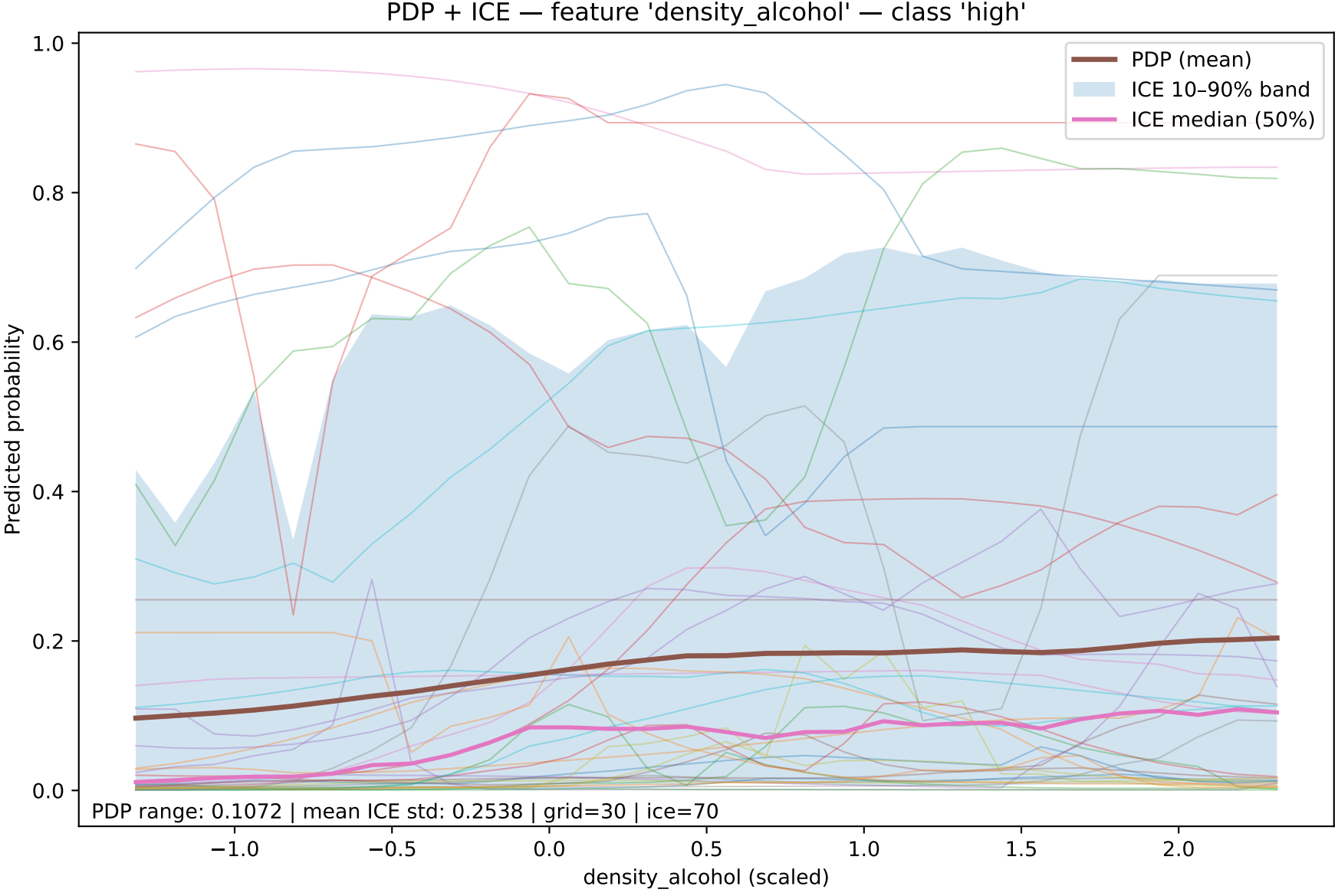

PDP + ICE — feature 'sulphur_ratio' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
−1
0
1
2
sulphur_ratio (scaled)
PDP range: 0.0160 | mean ICE std: 0.1082 | grid=30 | ice=70

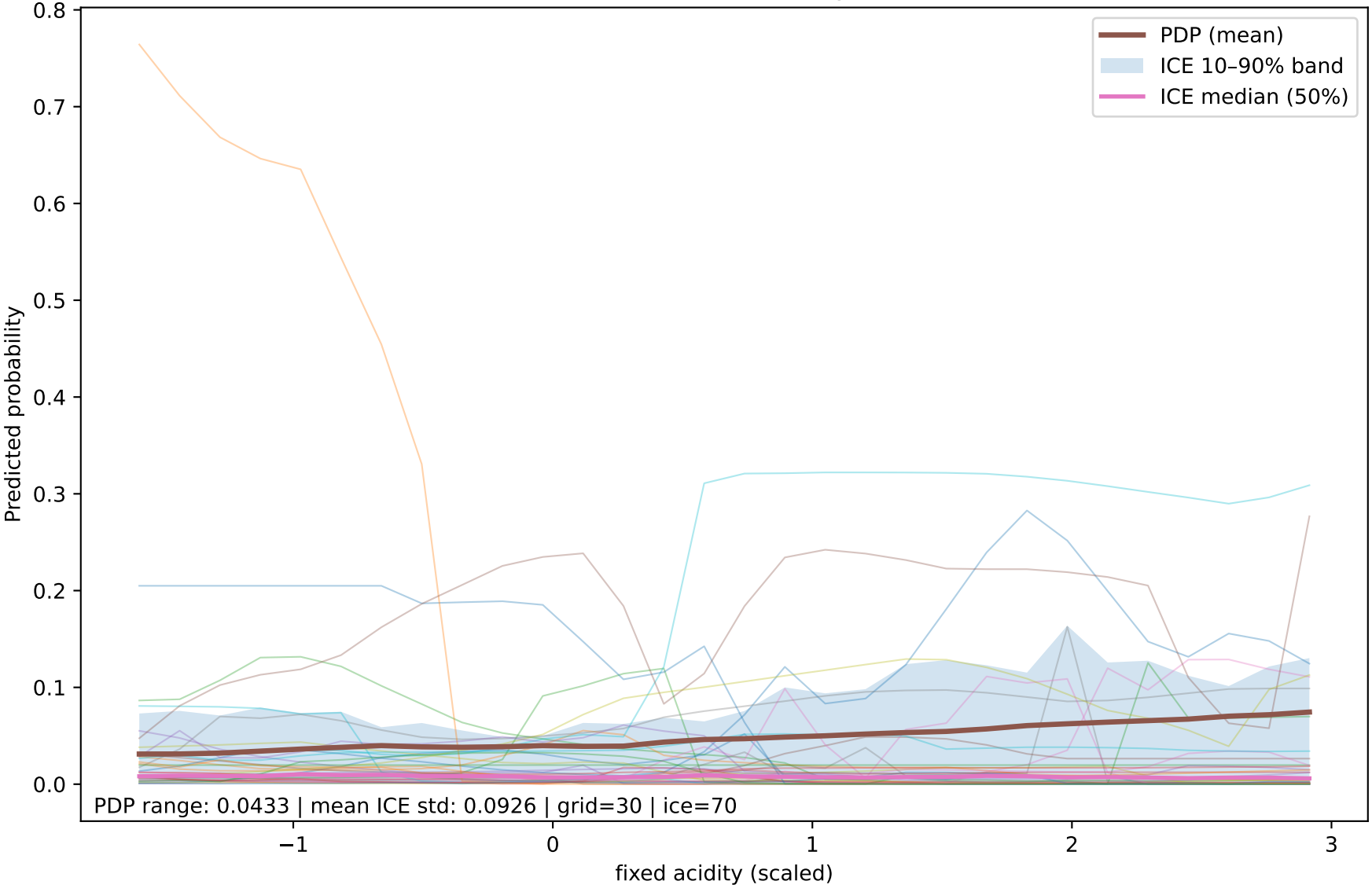

PDP + ICE — feature 'sulphur_ratio' — class 'medium'
Predicted probability
1.0
0.8
0.6
0.4
0.2
−1
0
1
2
sulphur_ratio (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.0894 | mean ICE std: 0.2313 | grid=30 | ice=70

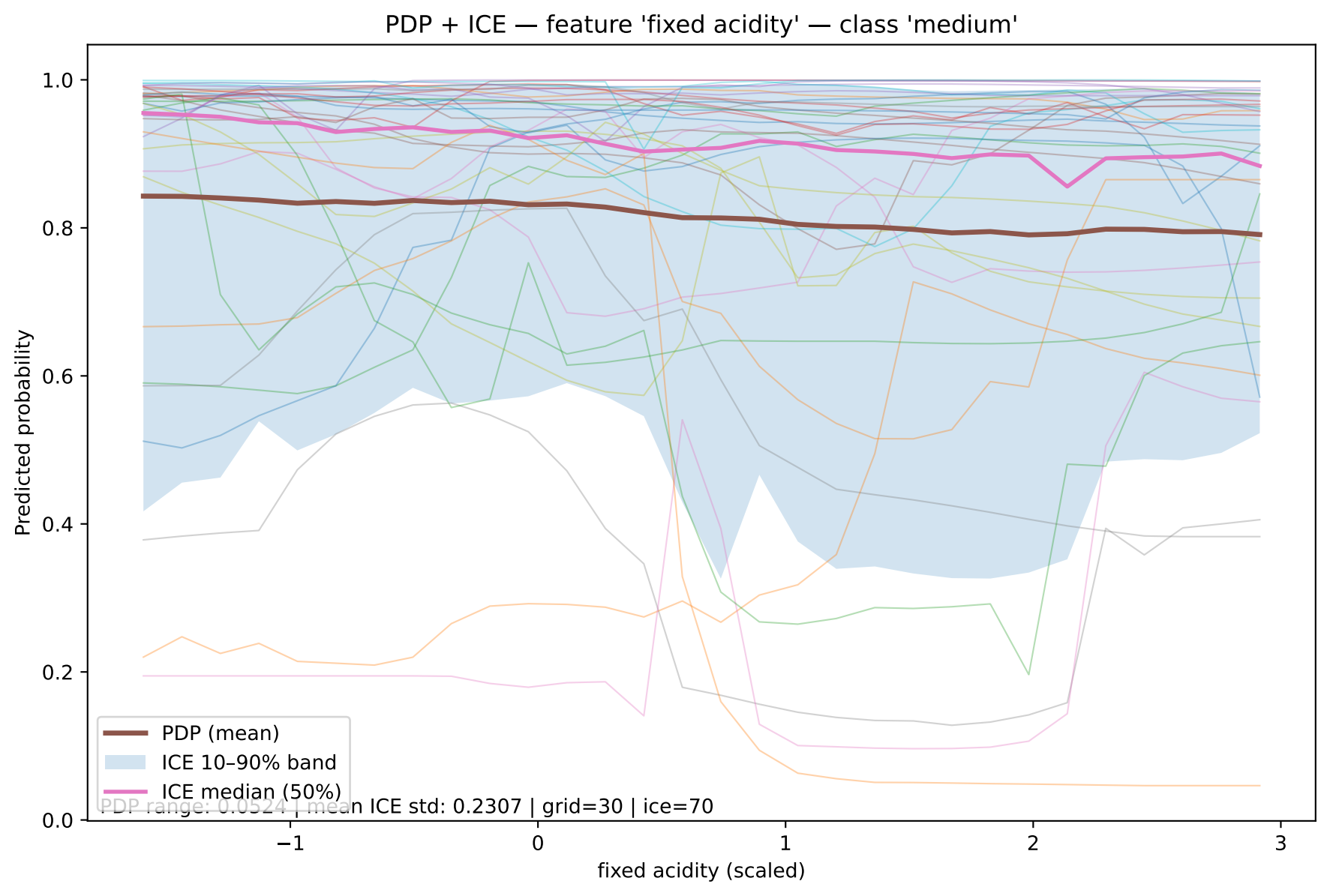

PDP + ICE — feature 'sulphur_ratio' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
−1
0
1
2
sulphur_ratio (scaled)
PDP range: 0.1050 | mean ICE std: 0.2036 | grid=30 | ice=70

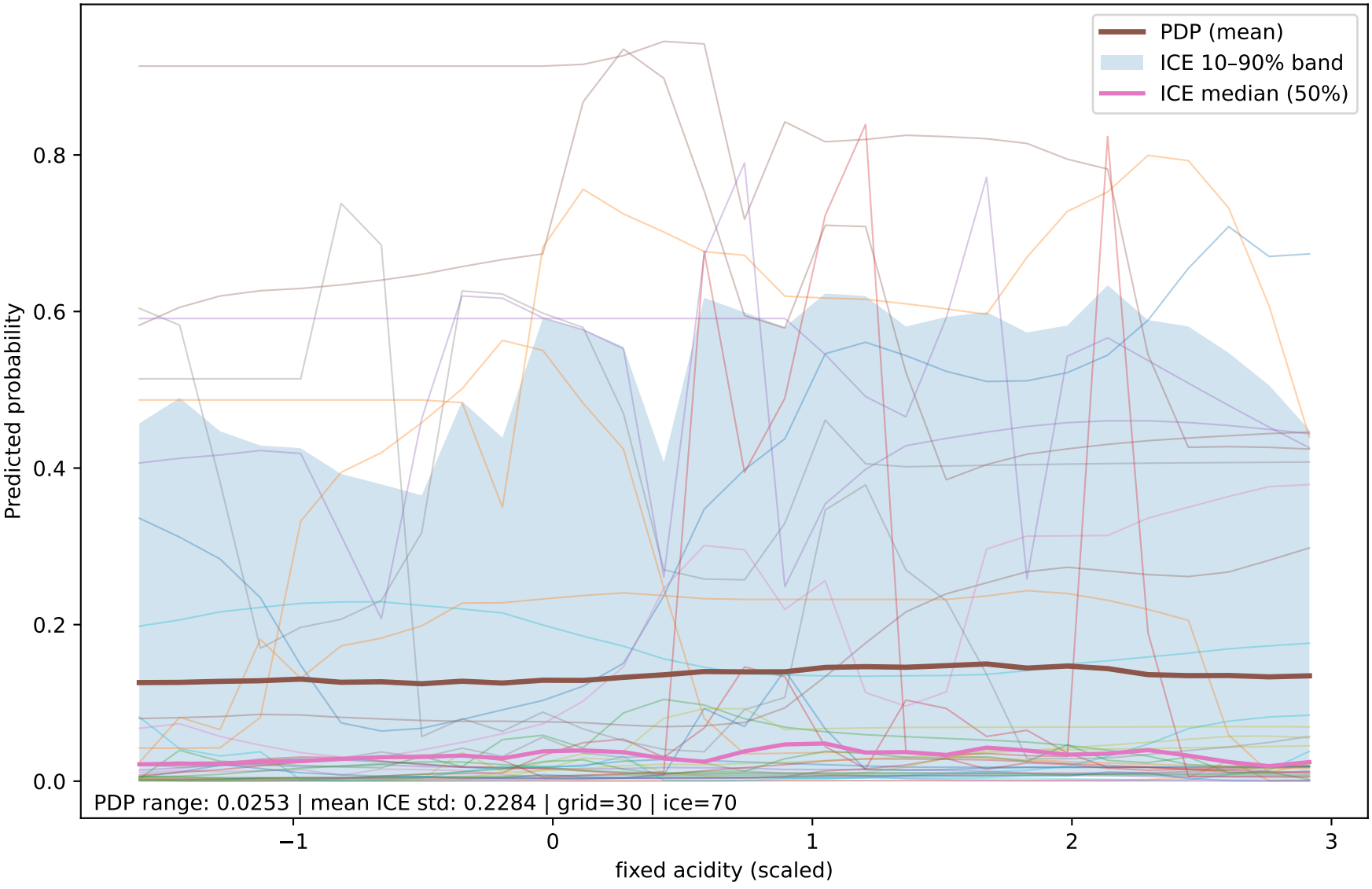

PDP + ICE — feature 'density_alcohol' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
PDP range: 0.0151 | mean ICE std: 0.0634 | grid=30 | ice=70
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
density_alcohol (scaled)

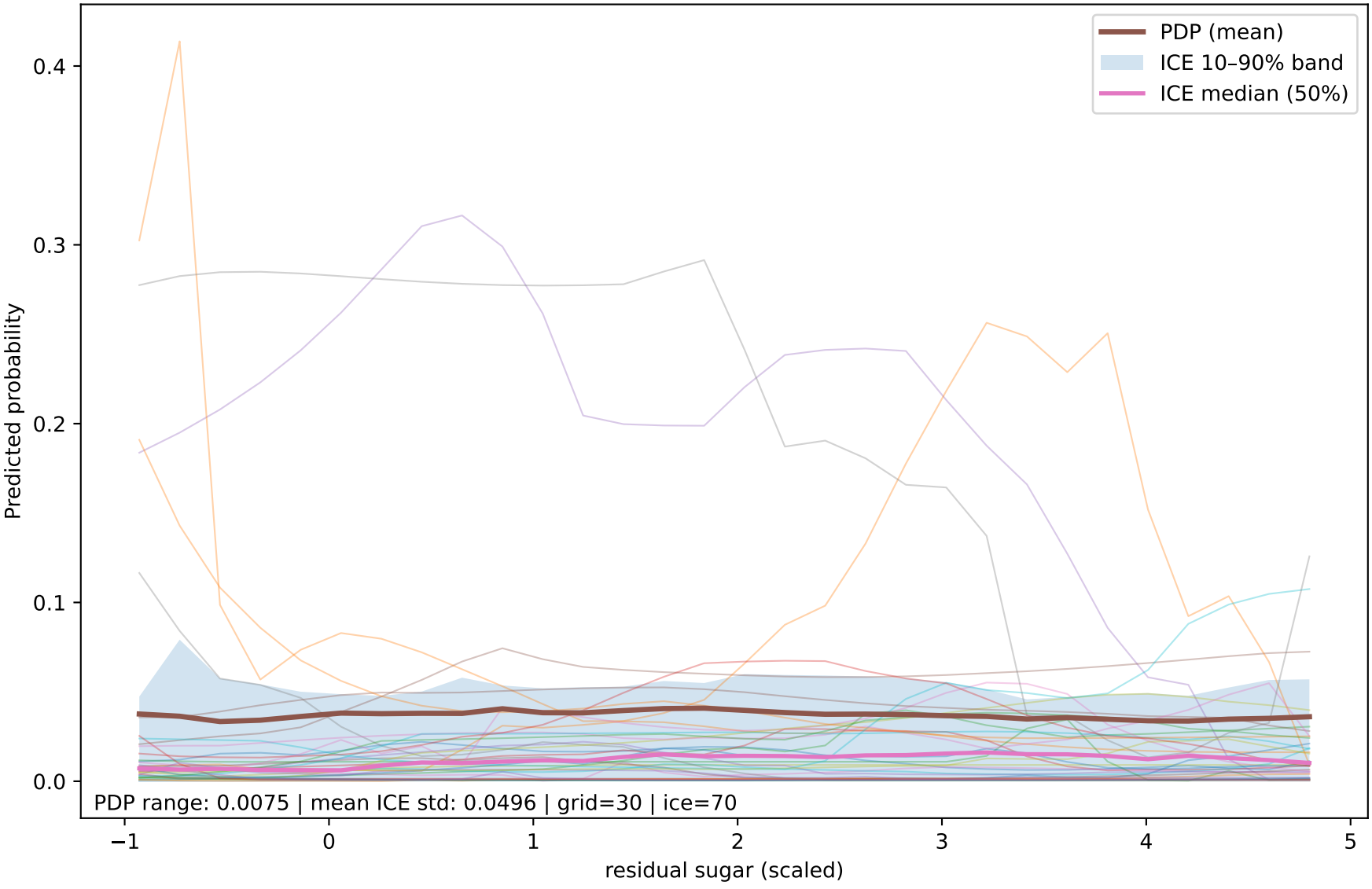

PDP + ICE — feature 'density_alcohol' — class 'medium'
Predicted probability
1.0
0.8
0.6
0.4
0.2
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.1001 | mean ICE std: 0.2717 | grid=30 | ice=70
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
density_alcohol (scaled)

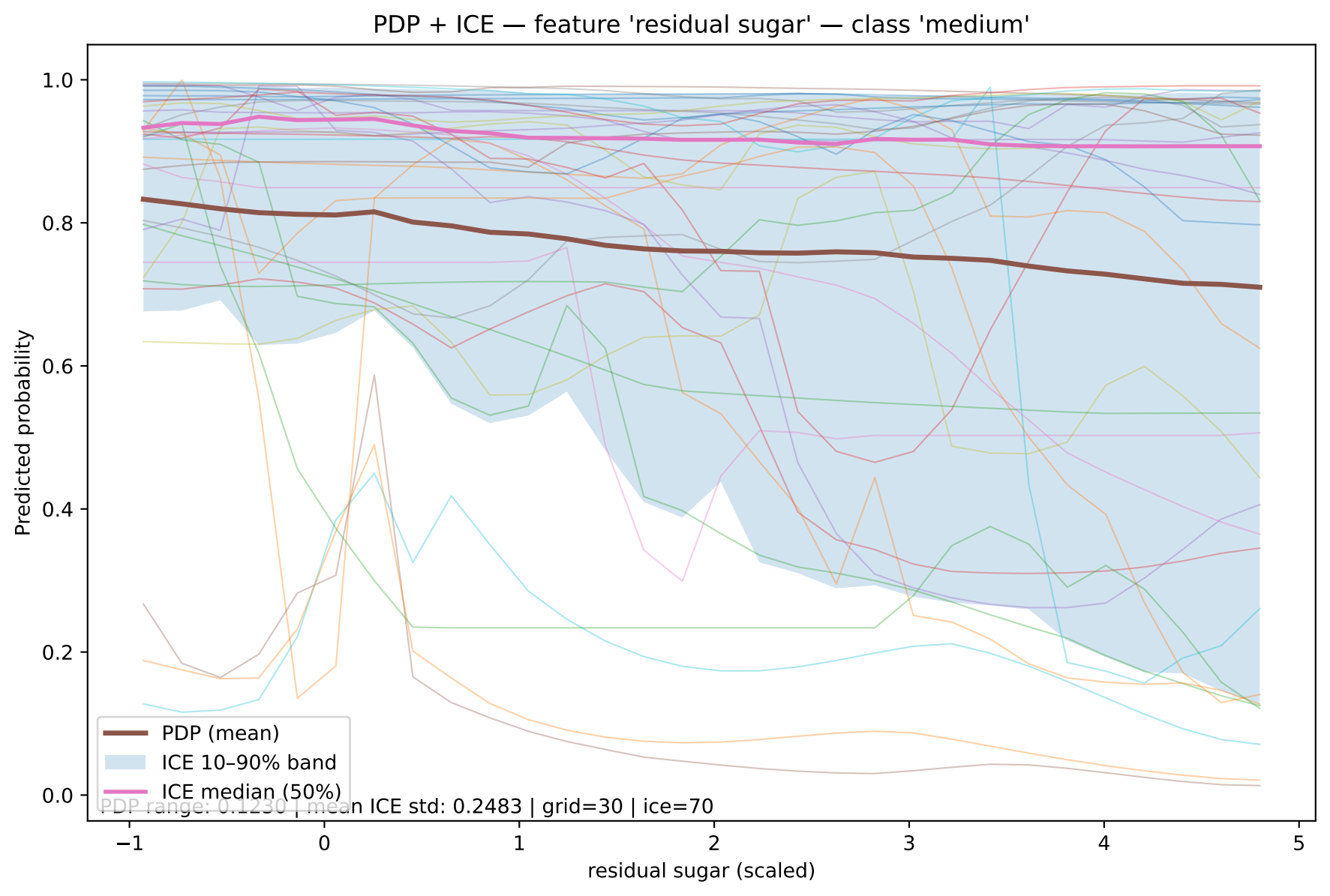

PDP + ICE — feature 'density_alcohol' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
PDP range: 0.1072 | mean ICE std: 0.2538 | grid=30 | ice=70
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
density_alcohol (scaled)

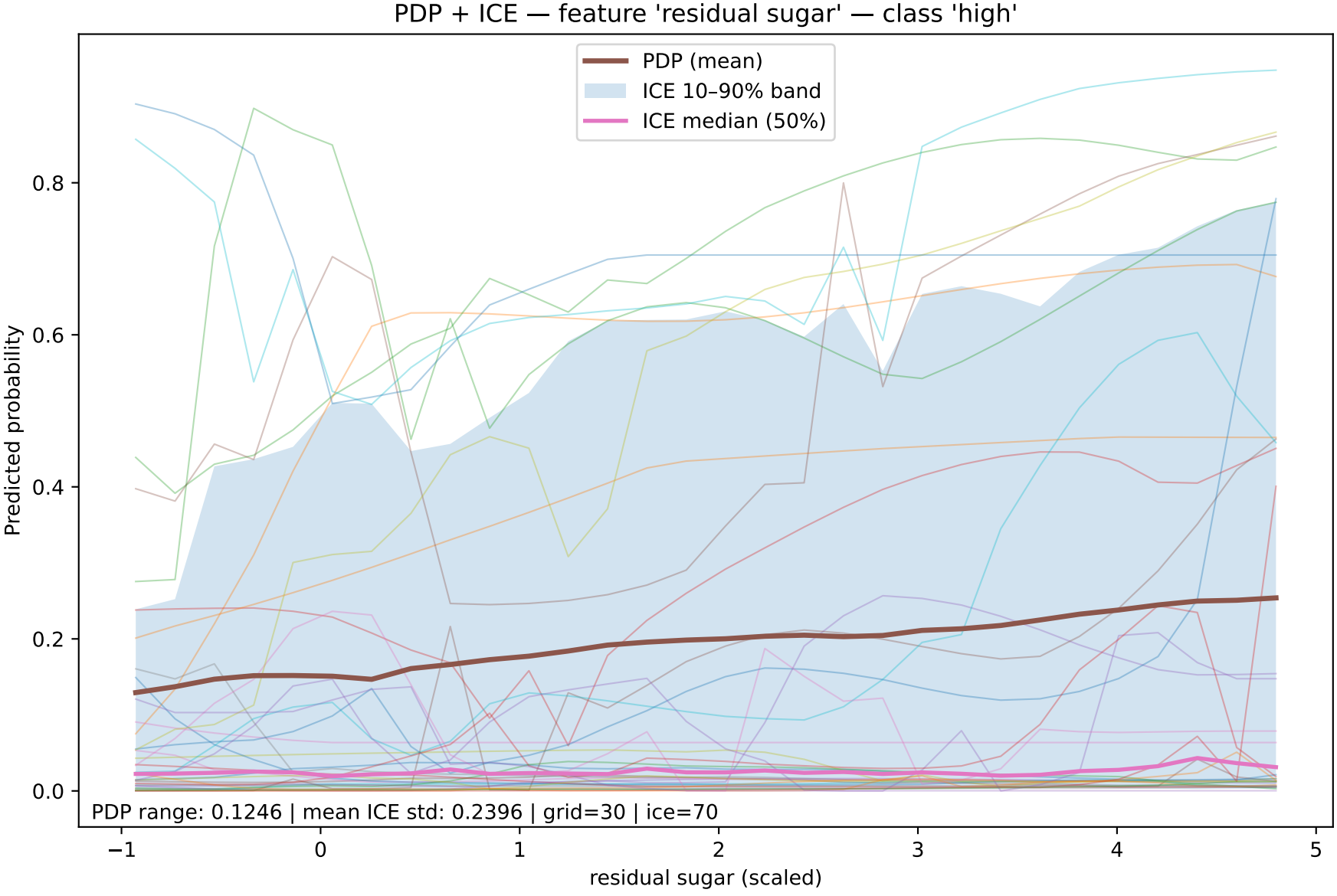

PDP + ICE — feature 'fixed acidity' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
fixed acidity (scaled)
PDP range: 0.0433 | mean ICE std: 0.0926 | grid=30 | ice=70
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
−1
0
1
2
3

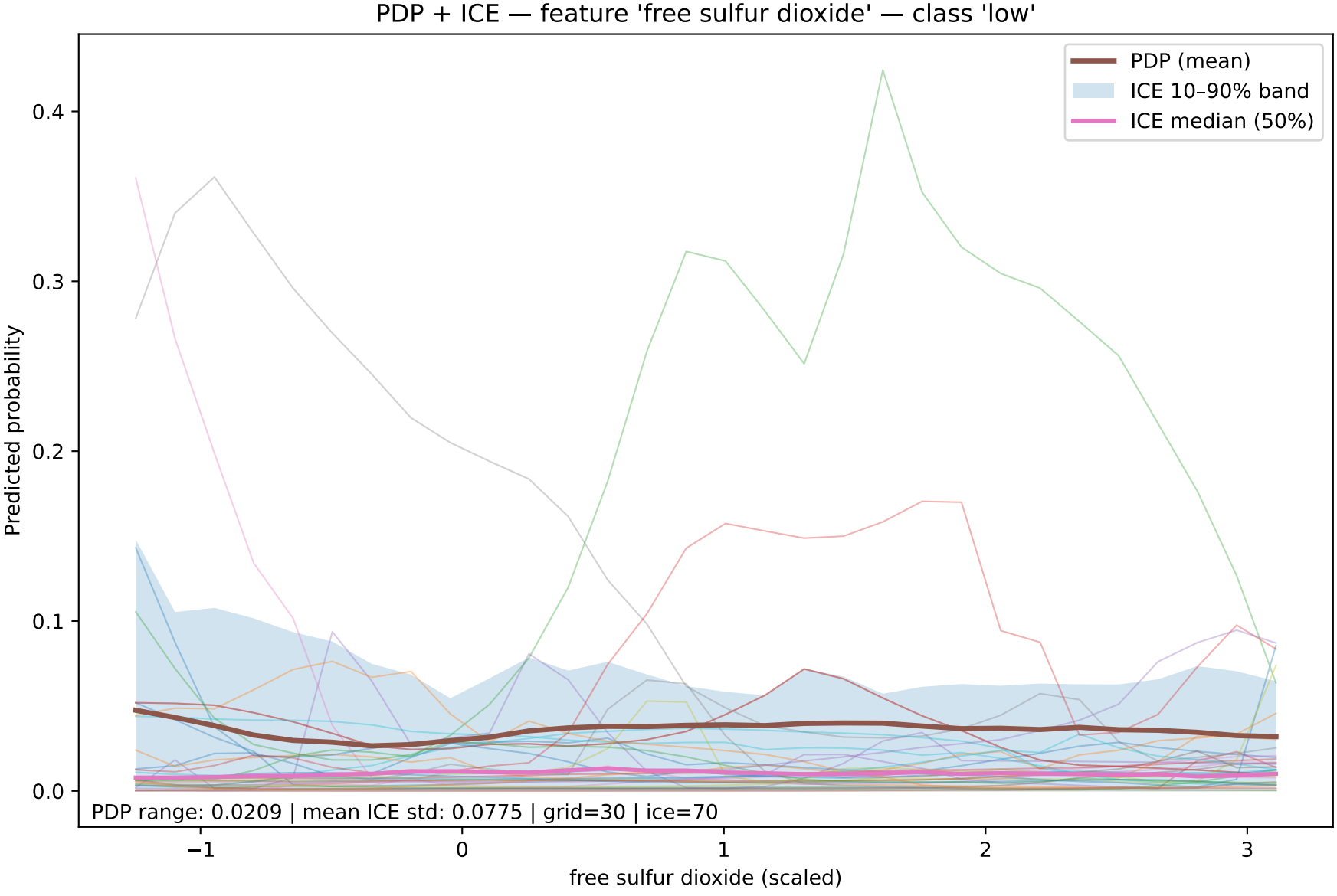

PDP + ICE — feature 'fixed acidity' — class 'medium'
Predicted probability
fixed acidity (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.0524 | mean ICE std: 0.2307 | grid=30 | ice=70

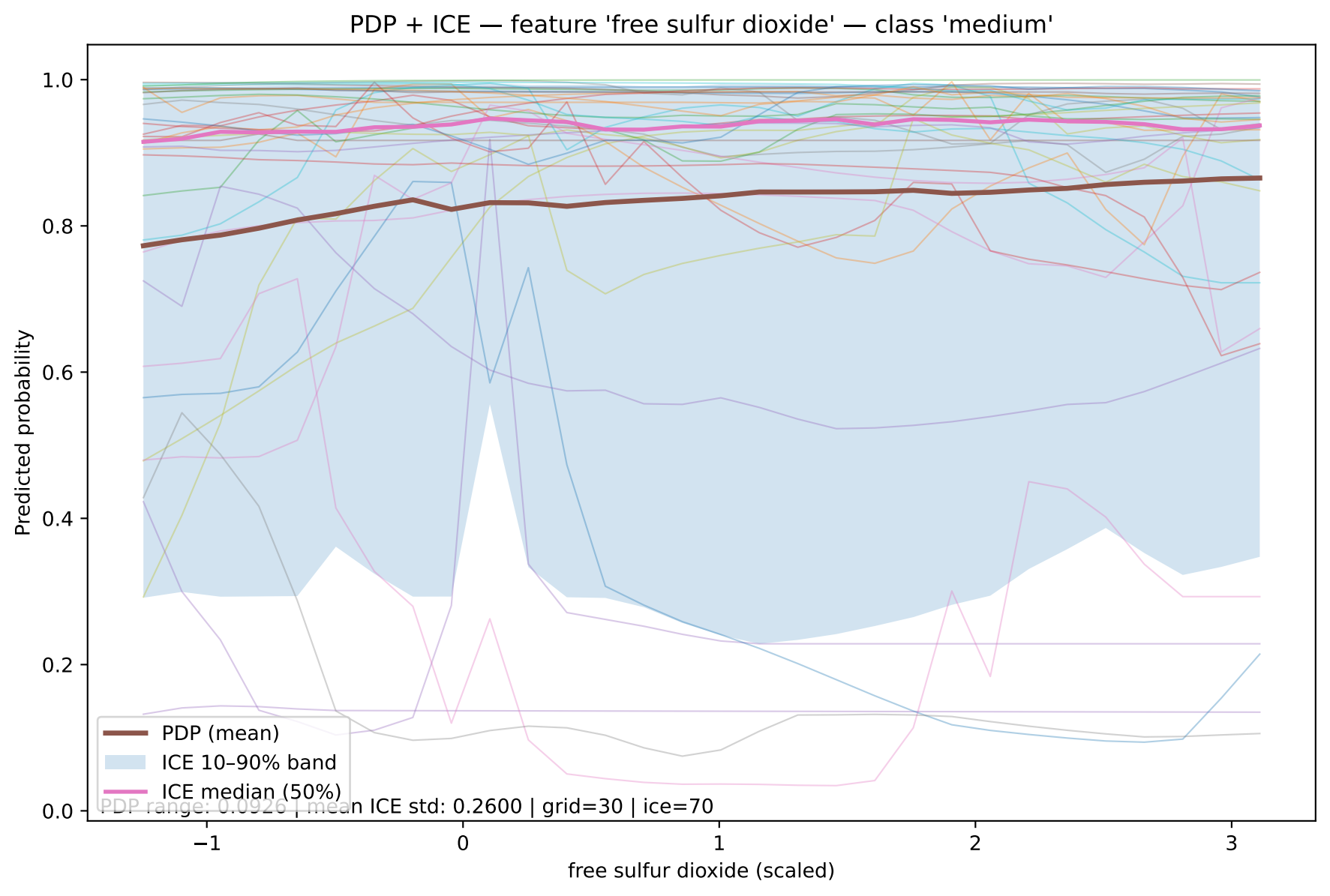

PDP + ICE — feature 'fixed acidity' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
−1
0
1
2
3
fixed acidity (scaled)
PDP range: 0.0253 | mean ICE std: 0.2284 | grid=30 | ice=70

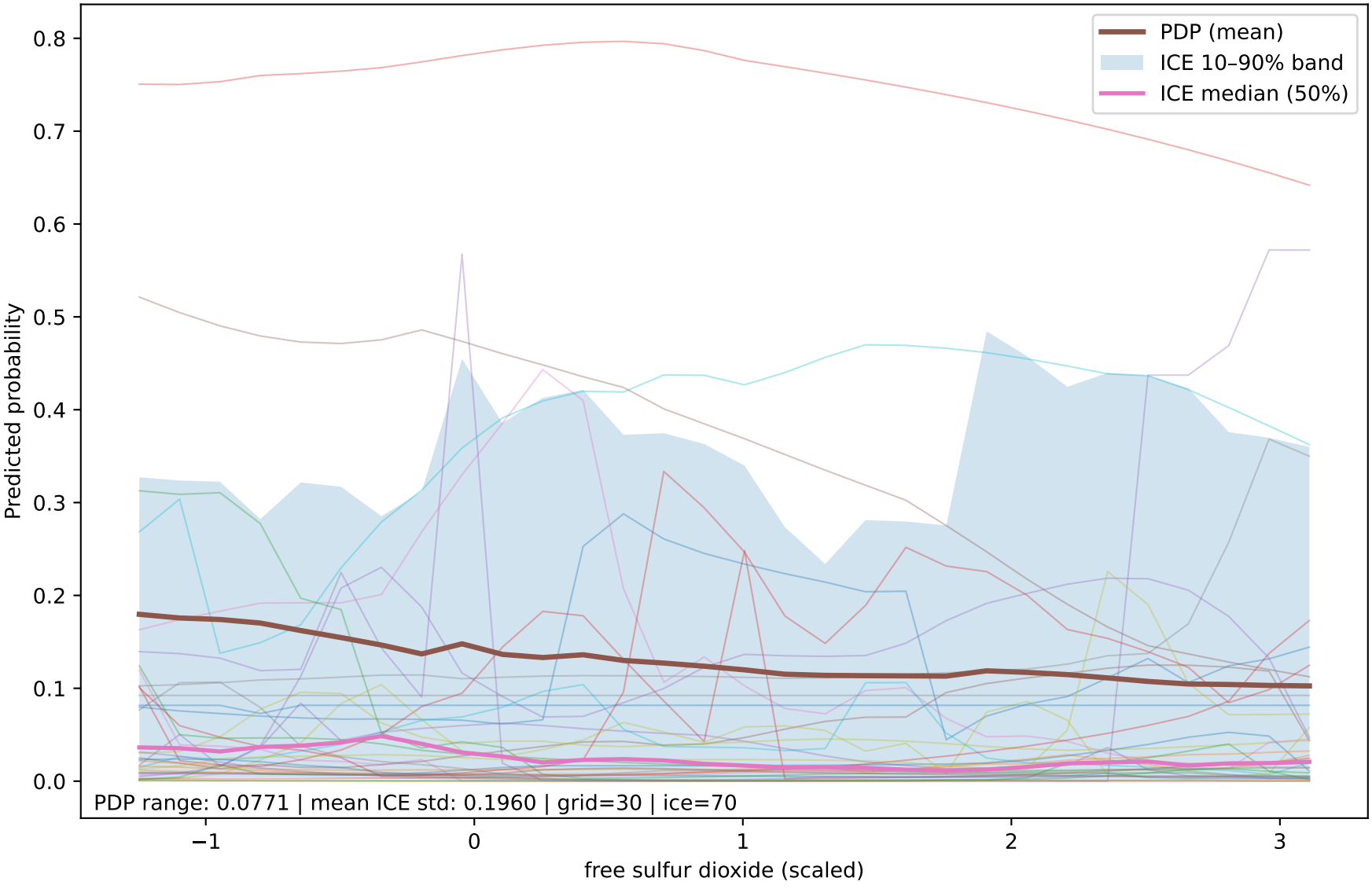

PDP + ICE — feature 'residual sugar' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.1
0.2
0.3
0.4
PDP range: 0.0075 | mean ICE std: 0.0496 | grid=30 | ice=70
−1
0
1
2
3
4
5
residual sugar (scaled)


PDP + ICE — feature 'residual sugar' — class 'medium'
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
−1
0
1
2
3
4
5
residual sugar (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.1230 | mean ICE std: 0.2483 | grid=30 | ice=70


PDP + ICE — feature 'residual sugar' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.2
0.4
0.6
0.8
PDP range: 0.1246 | mean ICE std: 0.2396 | grid=30 | ice=70
−1
0
1
2
3
4
5
residual sugar (scaled)


PDP + ICE — feature 'free sulfur dioxide' — class 'low'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.1
0.2
0.3
0.4
−1
0
1
2
3
free sulfur dioxide (scaled)
PDP range: 0.0209 | mean ICE std: 0.0775 | grid=30 | ice=70


PDP + ICE — feature 'free sulfur dioxide' — class 'medium'
Predicted probability
1.0
0.8
0.6
0.4
0.2
0.0
−1
0
1
2
3
free sulfur dioxide (scaled)
PDP (mean)
ICE 10–90% band
ICE median (50%)
PDP range: 0.0926 | mean ICE std: 0.2600 | grid=30 | ice=70


PDP + ICE — feature 'free sulfur dioxide' — class 'high'
PDP (mean)
ICE 10–90% band
ICE median (50%)
Predicted probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
PDP range: 0.0771 | mean ICE std: 0.1960 | grid=30 | ice=70
−1
0
1
2
3
free sulfur dioxide (scaled)