SUPPLEMENTARY DATA

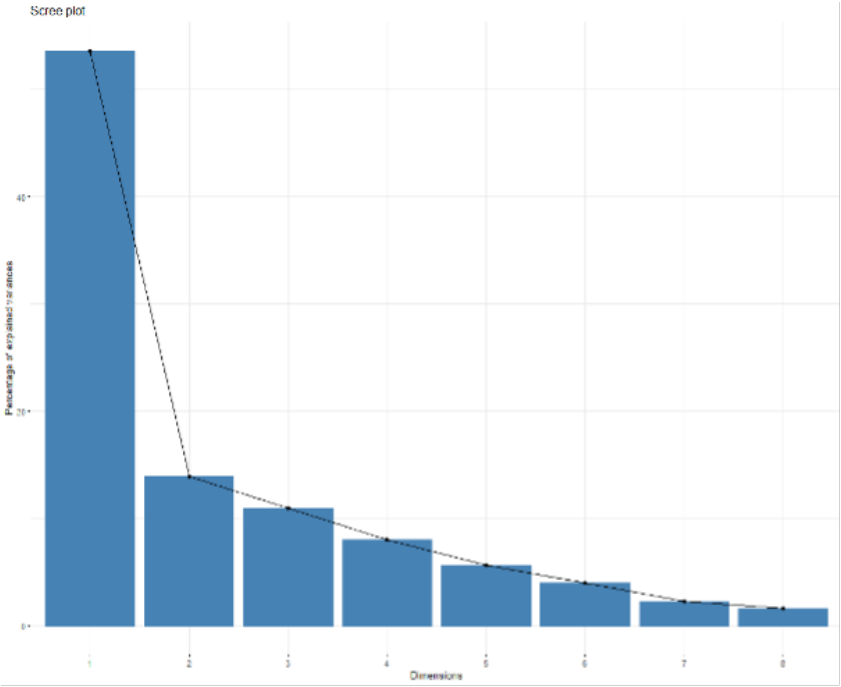
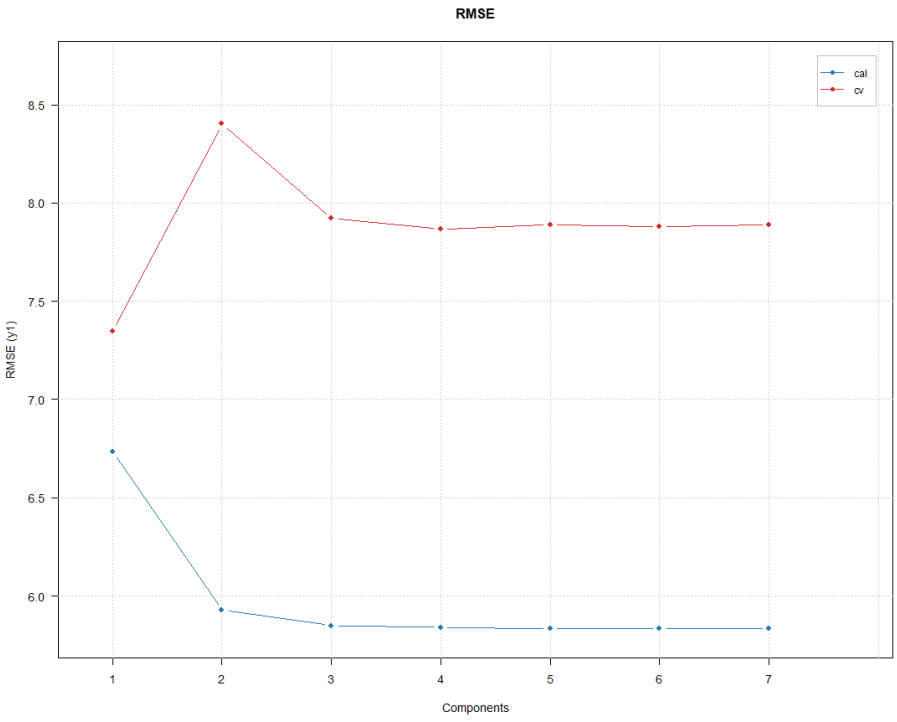
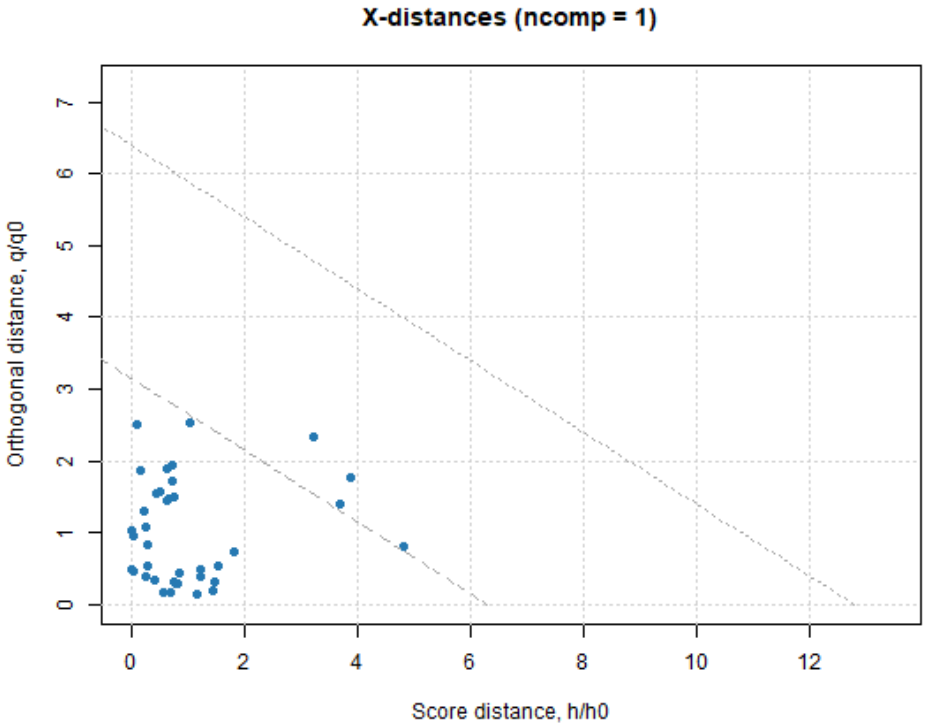


Figure S1. Principal component analysis (PCA) screeplot, showing relative contributions of principal components.

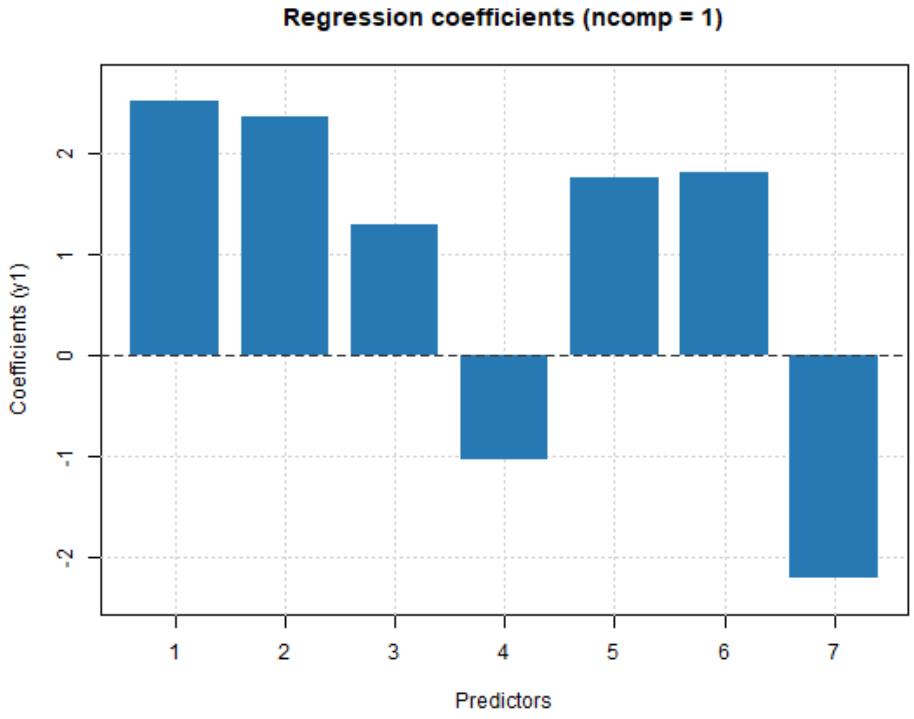
A.



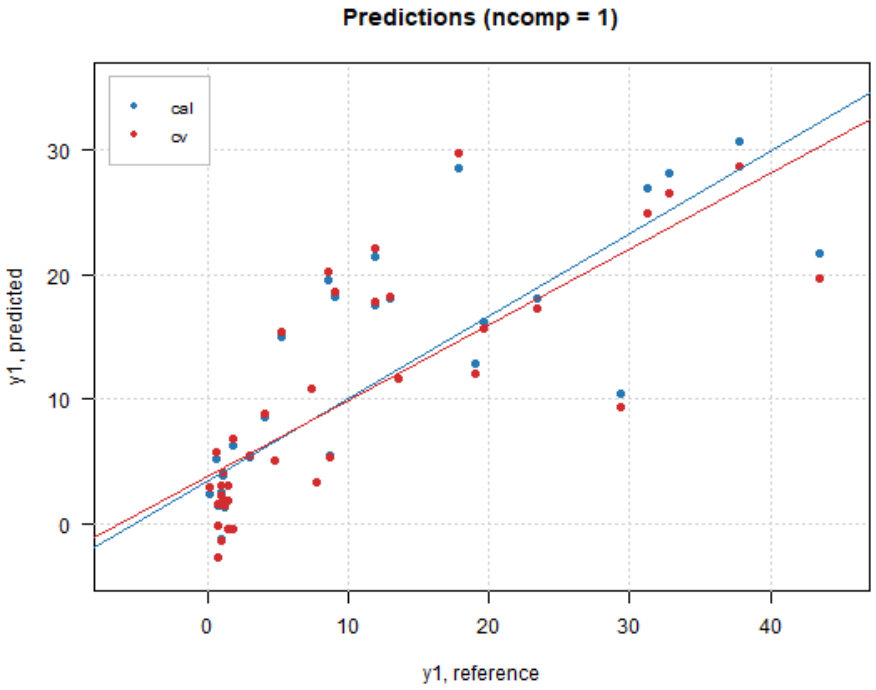
B.


X-distances (ncomp = 1)
Orthogonal distance, q/q0
Score distance, h/h0
0
1
2
3
4
5
6
7
0
2
4
6
8
10
12

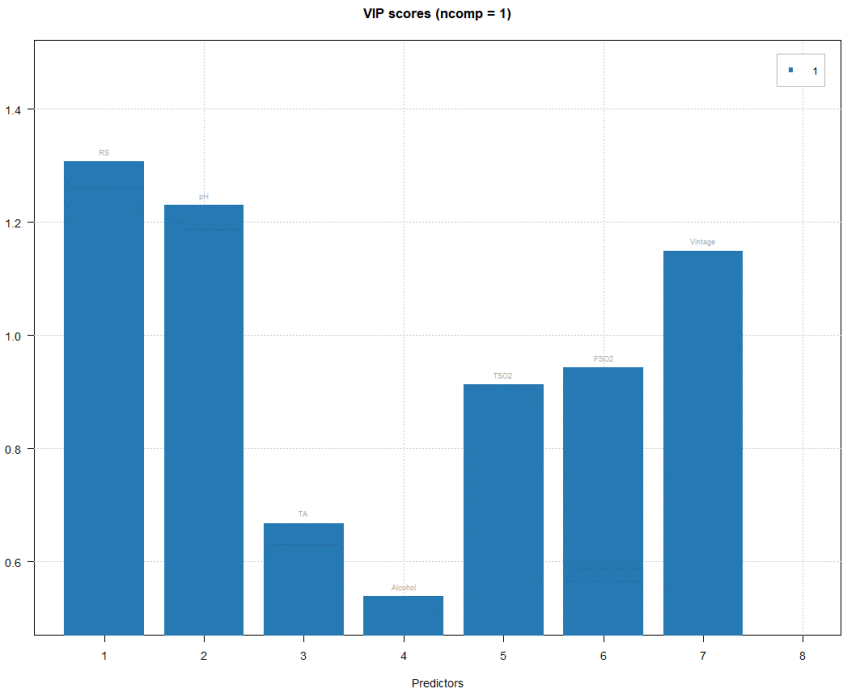
C.


Regression coefficients (ncomp = 1)
Coefficients (y1)
Predictors
-2
-1
0
1
2
1
2
3
4
5
6
7

D.


Predictions (ncomp = 1)
cal
cv
y1, predicted
y1, reference
0
10
20
30
40

E.


VIP scores (ncomp = 1)
1
RS
pH
TA
Alcohol
TSO2
FSO2
Vintage
0.6
0.8
1.0
1.2
1.4
1
2
3
4
5
6
7
8
Predictors

F.

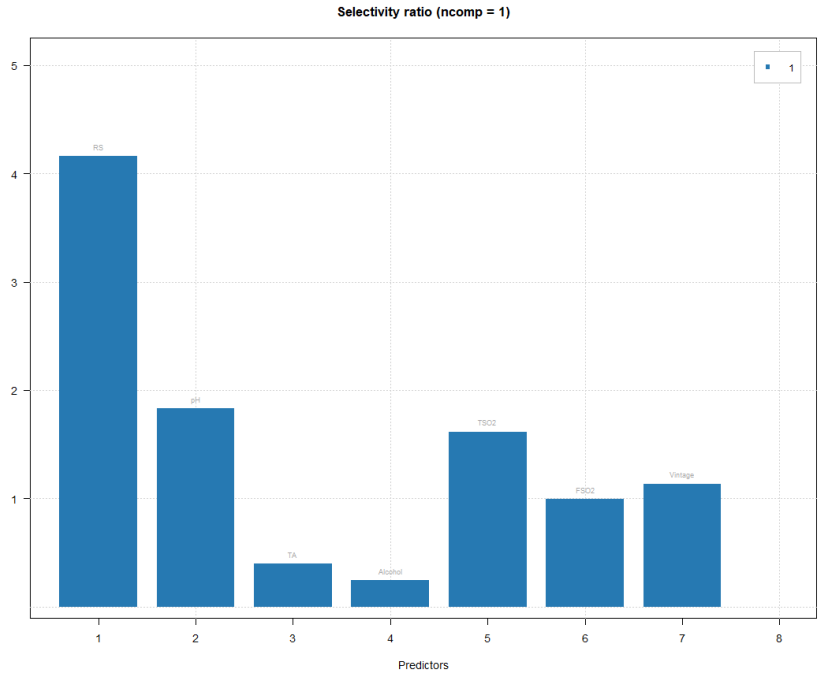


Figure S2. Figures showing the result of fitting a partial-least-squares regression (PLS-R) model (A) root-mean-square error (RMSE) associated with each component; (B) X-Distances; (C) regression coefficients; (D) predictions; (E) variable importance (VIP) scores and (F) selectivity ratios for predictors.

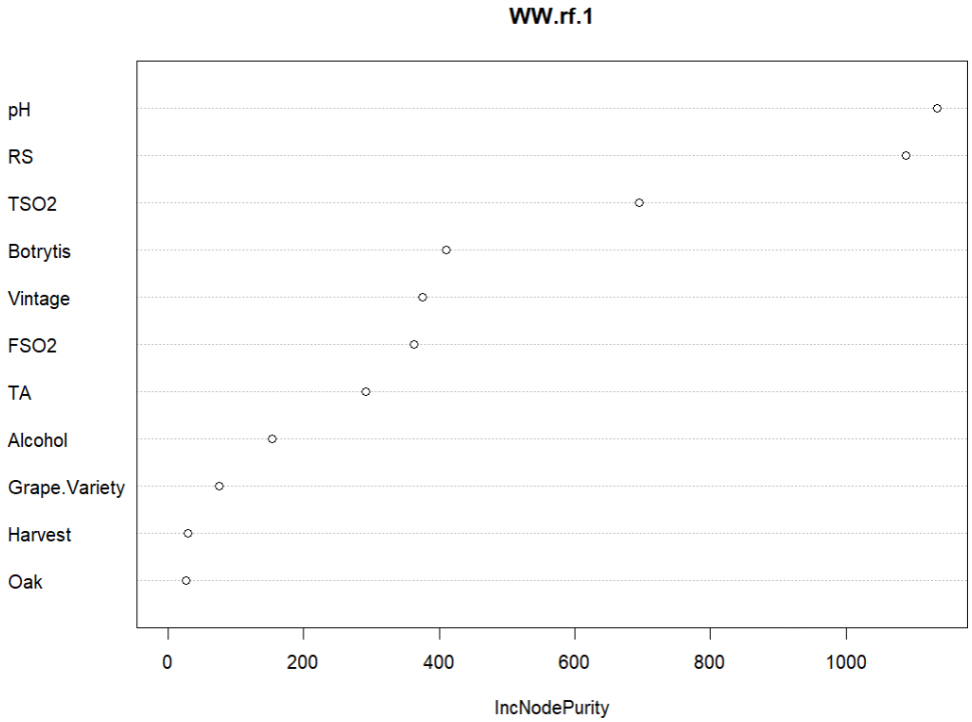


Figure S3. Random Forest output showing the contribution of different variables to predictions in one of the developed models.