Electronic Supplementary Material (ESI) for Analytical Methods.


# Geographical traceability of cultivated *Paris polyphylla* var. *yunnanensis* using ATR-FTMIR spectroscopy with three mathematical algorithms

Yi-Fei Pei [a, b], Li-Hua Wu [a], Qing-Zhi Zhang [b] and Yuan-Zhong Wang [a†]

[a] Institute of Medicinal Plants, Yunnan Academy of Agricultural Sciences, Kunming 650200, PR China.

[b] College of Traditional Chinese Medicine, Yunnan University of Traditional Chinese Medicine, Kunming 650500, PR China.

[†]Corresponding author:

**Mr Yuan-Zhong Wang**, Institute of Medicine Plants, Yunnan Academy of Agricultural Science, 2238, Beijing Road, Panlong District, Kunming 650200, China.

Tel: +86 871-65033575, Fax: +86 871-65033441, E-mail: boletus@126.com

**Support material table captions**

**Table S1.** The major parameters of raw and preprocessing PLS-DA models based on 3th to 8th year cultivated *P. yunnanensis* samples combined with ATR-FTMIR spectra.

**Table S2.** The major parameters of VIP PLS-DA models based on 3th to 8th year cultivated *P. yunnanensis* samples combined with ATR-FTMIR spectra.

**Table S3.** The major parameters of each class of raw and the best preprocessing ATR-FTMIR spectral PLS-DA models based on 3th to 8th year cultivated *P. yunnanensis* samples.

[a] 1: Central Yunnan, 2: Western Yunnan, 3: Northwest Yunnan, 4: Southeast Yunnan, 5: Southwest Yunnan.

**Table S4.** The major parameters of each class of raw and the best preprocessing ATR-FTMIR spectral RF models based on 3th to 8th year cultivated *P. yunnanensis* samples.

[a] 1: Central Yunnan, 2: Western Yunnan, 3: Northwest Yunnan, 4: Southeast Yunnan, 5: Southwest Yunnan.

**Table S5.** The major parameters of raw and preprocessing PLS-DA models based on cultivated *P. yunnanensis* samples combined with ATR-FTMIR spectra.

**Table S6.** The major parameters of each class of raw and the best preprocessing ATR-FTMIR spectral PLS-DA models based on cultivated *P. yunnanensis* samples.

[a] 1: Central Yunnan, 2: Western Yunnan, 3: Northwest Yunnan, 4: Southeast Yunnan, 5: Southwest Yunnan.

**Table S7.** The 27 classes Person correlation coefficient of HCA based on the best preprocessing ATR-FTMIR spectral data.

[a] 1: Central Yunnan (3th year), 2: Western Yunnan (3th year), 3: Northwest Yunnan (3th year), 4: Southeast Yunnan (3th year), 5: Southwest Yunnan (3th year), 6: Central Yunnan (4th year), 7: Western Yunnan (4th year), 8: Northwest Yunnan (4th year), 9: Southeast Yunnan (4th year), 10: Southwest Yunnan (4th year), 11: Central Yunnan (5th year), 12:

Western Yunnan (5th year), 13: Northwest Yunnan (5th year), 14: Southeast Yunnan (5th year), 15: Southwest Yunnan (5th year), 16: Central Yunnan (6th year), 17: Western Yunnan (6th year), 18: Northwest Yunnan (6th year), 19: Southeast Yunnan (6th year), 20: Central Yunnan (7th year), 21: Western Yunnan (7th year), 22: Northwest Yunnan (7th year), 23: Southeast Yunnan (7th year), 24: Central Yunnan (8th year), 25: Western Yunnan (8th year), 26: Northwest Yunnan (8th year), 27: Southeast Yunnan (8th year).

**Support material figure captions**

**Fig. S1.** $T^2$ Hotelling values of PCA plots of 3th to 8th year cultivated *P. yunnanensis* samples.

**Fig. S2.** Raw ATR-FTMIR spectra of cultivated *P. yunnanensis* samples after eliminating outliers.

**Fig. S3.** The 10-fold cross validation error rates of random forest model (sequentially reduce each five variables) based on 4th year cultivated *P. yunnanensis* samples.

**Fig. S4.** The 10-fold cross validation error rates of random forest model (sequentially reduce each five variables) based on 5th year cultivated *P. yunnanensis* samples.

**Fig. S5.** The 10-fold cross validation error rates of random forest model (sequentially reduce each five variables) based on 6th year cultivated *P. yunnanensis* samples.

**Fig. S6.** The 10-fold cross validation error rates of random forest model (sequentially reduce each five variables) based on 7th year cultivated *P. yunnanensis* samples.

**Fig. S7.** The 10-fold cross validation error rates of random forest model (sequentially reduce each five variables) based on 8th year cultivated *P. yunnanensis* samples.

**Fig. S8.** The $n_{tree}$ screening in random forest model of 4th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S9.** The $n_{tree}$ screening in random forest model of 5th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S10.** The $n_{tree}$ screening in random forest model of 6th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S11.** The $n_{tree}$ screening in random forest model of 7th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S12.** The $n_{tree}$ screening in random forest model of 8th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S13.** The $m_{try}$ screening in random forest model of 4th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S14.** The $m_{try}$ screening in random forest model of 5th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S15.** The $m_{try}$ screening in random forest model of 6th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S16.** The $m_{try}$ screening in random forest model of 7th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S17.** The $m_{try}$ screening in random forest model of 8th year cultivated *P. yunnanensis* samples for variables ranked by permutation accuracy importance.

**Fig. S18.** The importance variables in the random forest models and the VIP values in PLS-DA models of 4th year cultivated *P. yunnanensis* samples: (a) random forest model, (b) PLS-DA.

**Fig. S19.** The importance variables in the random forest models and the VIP values in PLS-DA models of 5th year cultivated *P. yunnanensis* samples: (a) random forest model, (b) PLS-DA.

**Fig. S20.** The importance variables in the random forest models and the VIP values in PLS-DA models of 6th year cultivated *P. yunnanensis* samples: (a) random forest model, (b) PLS-DA.

**Fig. S21.** The importance variables in the random forest models and the VIP values in PLS-DA models of 7th year cultivated *P. yunnanensis* samples: (a) random forest model, (b) PLS-DA.

**Fig. S22.** The importance variables in the random forest models and the VIP values in PLS-DA models of 8th year cultivated *P. yunnanensis* samples: (a) random forest model, (b) PLS-DA.

**Fig. S23.** $T^2$ Hotelling values of PCA plots of all cultivated *P. yunnanensis* samples.

97

**Table S1**

| Cultivation year | Preprocessing | $R^2$ | $Q^2$ | Calibration set | | | Validation set | |
|---|---|---|---|---|---|---|---|---|
| | | | | RMSEE | RMSECV | Accuracy | RMSEP | Accuracy |
| 3th | **Raw** | **0.130** | **0.095** | **0.3720760** | **0.3741626** | **42.99%** | **0.3686538** | **34.55%** |
| | **SNV-SG5-SD** | **0.924** | **0.721** | **0.1146762** | **0.2261066** | **100%** | **0.1450830** | **100%** |
| | SNV-SG7-SD | 0.901 | 0.704 | 0.1303506 | 0.2312564 | 100% | 0.1502088 | 100% |
| | SNV-SG9-SD | 0.911 | 0.692 | 0.1251111 | 0.2401040 | 100% | 0.1483418 | 100% |
| | SNV-SG11-SD | 0.908 | 0.680 | 0.1273949 | 0.2496370 | 100% | 0.1594702 | 100% |
| | SNV-SG13-SD | 0.853 | 0.502 | 0.1593638 | 0.2897692 | 100% | 0.1657708 | 98.18% |
| | SNV-SG15-SD | 0.808 | 0.503 | 0.1802360 | 0.2821198 | 99.07% | 0.1927210 | 96.36% |
| 4th | **Raw** | **0.872** | **0.578** | **0.1551412** | **0.2651592** | **100%** | **0.2039218** | **96.30%** |
| | **SNV-SG5-SD** | **0.891** | **0.710** | **0.1364512** | **0.2266002** | **100%** | **0.1783110** | **100%** |
| | SNV-SG7-SD | 0.866 | 0.697 | 0.1396532 | 0.2323352 | 100% | 0.1782930 | 100% |
| | SNV-SG9-SD | 0.883 | 0.710 | 0.1414784 | 0.2243538 | 100% | 0.1780812 | 100% |
| | SNV-SG11-SD | 0.882 | 0.681 | 0.1421946 | 0.2359512 | 100% | 0.1783928 | 100% |
| | SNV-SG13-SD | 0.883 | 0.680 | 0.1410890 | 0.2438798 | 100% | 0.1789962 | 100% |
| | SNV-SG15-SD | 0.891 | 0.655 | 0.1373834 | 0.2377426 | 100% | 0.1791938 | 100% |
| 5th | **Raw** | **0.657** | **0.478** | **0.2382252** | **0.2795564** | **92.86%** | **0.2553344** | **86.27%** |
| | SNV-SG5-SD | 0.875 | 0.763 | 0.1421154 | 0.2161942 | 99.00% | 0.1356692 | 100% |
| | **SNV-SG7-SD** | **0.906** | **0.768** | **0.1251872** | **0.2091728** | **100%** | **0.1318970** | **100%** |
| | SNV-SG9-SD | 0.899 | 0.761 | 0.1290902 | 0.2101728 | 100% | 0.1306526 | 100% |
| | SNV-SG11-SD | 0.873 | 0.751 | 0.1432070 | 0.2230904 | 98.98% | 0.1375952 | 100% |
| | SNV-SG13-SD | 0.920 | 0.762 | 0.1157530 | 0.2173356 | 100% | 0.1440970 | 100% |
| | SNV-SG15-SD | 0.875 | 0.769 | 0.1428344 | 0.2106780 | 98.98% | 0.1431756 | 100% |
| 6th | **Raw** | **0.342** | **0.221** | **0.3585155** | **0.3694278** | **71.19%** | **0.3300400** | **68.97%** |
| | **SNV-SG5-SD** | **0.958** | **0.888** | **0.0929050** | **0.1954237** | **100%** | **0.1322281** | **100%** |
| | SNV-SG7-SD | 0.966 | 0.901 | 0.0845265 | 0.1922495 | 100% | 0.1292994 | 100% |
| | SNV-SG9-SD | 0.964 | 0.898 | 0.0860050 | 0.1892953 | 100% | 0.1299654 | 100% |
| | SNV-SG11-SD | 0.960 | 0.884 | 0.0911349 | 0.1915370 | 100% | 0.1305449 | 100% |
| | SNV-SG13-SD | 0.960 | 0.892 | 0.0908156 | 0.1864963 | 100% | 0.1269629 | 100% |
| | SNV-SG15-SD | 0.918 | 0.860 | 0.1271845 | 0.1771825 | 100% | 0.1710573 | 100% |
| 7th | **Raw** | **0.432** | **0.329** | **0.3387485** | **0.3526643** | **78.46%** | **0.3081238** | **77.00%** |
| | **SNV-SG5-SD** | **0.910** | **0.740** | **0.1364935** | **0.2453135** | **100%** | **0.1553158** | **100%** |
| | SNV-SG7-SD | 0.908 | 0.734 | 0.1377408 | 0.2485798 | 100% | 0.1537353 | 100% |
| | SNV-SG9-SD | 0.906 | 0.735 | 0.1398590 | 0.2453420 | 100% | 0.1564418 | 100% |
| | SNV-SG11-SD | 0.906 | 0.748 | 0.1390683 | 0.2440255 | 100% | 0.1544920 | 100% |
| | SNV-SG13-SD | 0.900 | 0.742 | 0.1435710 | 0.2424915 | 100% | 0.1537558 | 100% |
| | SNV-SG15-SD | 0.946 | 0.750 | 0.1080286 | 0.2498123 | 100% | 0.1821420 | 97.14% |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 8th | **Raw** | **0.515** | **0.377** | **0.3127565** | **0.3397203** | **82.14%** | **0.3113498** | **74.00%** |
| | SNV-SG5-SD | 0.925 | 0.784 | 0.1264000 | 0.2226270 | 100% | 0.1834178 | 93.55% |
| | SNV-SG7-SD | 0.782 | 0.653 | 0.2045613 | 0.2659275 | 92.86% | 0.2088145 | 96.77% |
| | SNV-SG9-SD | 0.732 | 0.620 | 0.2298555 | 0.2720580 | 92.86% | 0.2271950 | 87.10% |
| | SNV-SG11-SD | 0.925 | 0.774 | 0.1271910 | 0.2248368 | 100% | 0.1828263 | 93.55% |
| | **SNV-SG13-SD** | **0.936** | **0.770** | **0.1181445** | **0.2246970** | **100%** | **0.1826703** | **93.55%** |
| | SNV-SG15-SD | 0.692 | 0.576 | 0.2485943 | 0.2862348 | 89.29% | 0.2583608 | 87.10% |

98
99

## Table S2

| Cultivation year | $R^2$ | $Q^2$ | PCs | Calibration set | | | Validation set | |
|---|---|---|---|---|---|---|---|---|
| | | | | RMSEE | RMSECV | Accuracy | RMSEP | Accuracy |
| 3th | 0.855 | 0.675 | 10 | 0.130181 | 0.271757 | 100% | 0.131931 | 100% |
| 4th | 0.869 | 0.705 | 10 | 0.149641 | 0.230972 | 100% | 0.151579 | 100% |
| 5th | 0.868 | 0.759 | 8 | 0.144489 | 0.208389 | 100% | 0.137904 | 100% |
| 6th | 0.912 | 0.846 | 5 | 0.127312 | 0.222636 | 100% | 0.121431 | 100% |
| 7th | 0.931 | 0.811 | 8 | 0.120194 | 0.207619 | 100% | 0.130088 | 100% |
| 8th | 0.787 | 0.656 | 4 | 0.205517 | 0.288590 | 96.43% | 0.191470 | 97.70% |

## Table S3

| Cultivation year | Set | Classes [a] | Raw | | | | The best preprocessing | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | SE | SP | Accuracy | MCC | SE | SP | Accuracy | MCC |
| 3th | Calibration set | 1 | 0 | 1.000 | 0.785 | 0 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.719 | 0.547 | 0.598 | 0.244 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0 | 0.957 | 0.841 | -0.073 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.263 | 0.920 | 0.804 | 0.222 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 0.900 | 0.816 | 0.832 | 0.600 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0 | 1.000 | 0.782 | 0 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.563 | 0.436 | 0.473 | -0.001 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0 | 0.896 | 0.782 | -0.121 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.200 | 0.911 | 0.782 | 0.137 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 0.800 | 0.911 | 0.891 | 0.664 | 1.000 | 1.000 | 1.000 | 1.000 |
| 4th | Calibration set | 1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0.941 | 0.973 | 0.963 | 0.914 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 1.000 | 0.977 | 0.981 | 0.943 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 0.900 | 1.000 | 0.981 | 0.938 | 1.000 | 1.000 | 1.000 | 1.000 |
| 5th | Calibration set | 1 | 0.944 | 0.919 | 0.929 | 0.851 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 1.000 | 0.987 | 0.990 | 0.970 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.900 | 1.000 | 0.990 | 0.943 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.75 | 0.988 | 0.949 | 0.804 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0.833 | 0.879 | 0.863 | 0.704 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 1.000 | 0.925 | 0.941 | 0.853 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 4 | 0.556 | 1.000 | 0.922 | 0.712 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 6th | Calibration set | 1 | 0.765 | 0.976 | 0.915 | 0.789 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.600 | 0.959 | 0.898 | 0.613 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.500 | 0.939 | 0.864 | 0.481 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.818 | 0.703 | 0.746 | 0.504 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0.556 | 0.900 | 0.793 | 0.493 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.600 | 1.000 | 0.931 | 0.744 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.600 | 0.875 | 0.828 | 0.443 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.900 | 0.789 | 0.828 | 0.659 | 1.000 | 1.000 | 1.000 | 1.000 |
| 7th | Calibration set | 1 | 0.500 | 0.959 | 0.846 | 0.548 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.941 | 0.917 | 0.923 | 0.817 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.769 | 0.942 | 0.908 | 0.712 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.895 | 0.891 | 0.892 | 0.756 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0.222 | 0.962 | 0.771 | 0.287 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.889 | 0.923 | 0.914 | 0.786 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 0.929 | 0.943 | 0.850 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 1.000 | 0.88 | 0.914 | 0.823 | 1.000 | 1.000 | 1.000 | 1.000 |
| 8th | Calibration set | 1 | 0.722 | 0.947 | 0.875 | 0.706 | 0.8 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.667 | 0.977 | 0.911 | 0.719 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 0.978 | 0.982 | 0.943 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.938 | 0.85 | 0.875 | 0.735 | 1.000 | 1.000 | 1.000 | 1.000 |
| | Validation set | 1 | 0.600 | 1.000 | 0.871 | 0.710 | 0.8 | 1.000 | 0.935 | 0.855 |
| | | 2 | 0.429 | 0.958 | 0.839 | 0.483 | 1.000 | 0.958 | 0.968 | 0.916 |
| | | 3 | 1.000 | 0.962 | 0.968 | 0.895 | 0.935 | 1.000 | 1.000 | 1.000 |
| | | 4 | 1.000 | 0.727 | 0.806 | 0.661 | 0.855 | 0.955 | 0.968 | 0.927 |

## Table S4

| Cultivation year | Set | Classes [a] | Raw | | | | The best preprocessing | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | SE | SP | Accuracy | MCC | SE | SP | Accuracy | MCC |
| 3th | Calibration set | 1 | 0.304 | 0.869 | 0.748 | 0.190 | 0.870 | 0.964 | 0.944 | 0.834 |
| | | 2 | 0.688 | 0.760 | 0.738 | 0.423 | 0.969 | 0.960 | 0.963 | 0.913 |
| | | 3 | 0.154 | 0.957 | 0.860 | 0.158 | 0.846 | 1.000 | 0.981 | 0.91 |
| | | 4 | 0.474 | 0.886 | 0.813 | 0.360 | 0.947 | 1.000 | 0.911 | 0.968 |
| | | 5 | 0.550 | 0.851 | 0.794 | 0.374 | 1.000 | 0.989 | 0.991 | 0.970 |
| | Validation set | 1 | 0.417 | 0.977 | 0.855 | 0.521 | 0.917 | 0.977 | 0.964 | 0.893 |
| | | 2 | 0.813 | 0.821 | 0.818 | 0.598 | 1.000 | 0.949 | 0.964 | 0.918 |
| | | 3 | 0.571 | 0.958 | 0.909 | 0.566 | 0.571 | 1.000 | 0.945 | 0.733 |
| | | 4 | 0.900 | 0.933 | 0.927 | 0.778 | 1.000 | 0.978 | 0.982 | 0.943 |
| | | 5 | 0.900 | 0.956 | 0.945 | 0.825 | 1.000 | 1.000 | 1.000 | 1.000 |
| 4th | Calibration set | 1 | 0.618 | 0.783 | 0.728 | 0.395 | 0.912 | 0.899 | 0.903 | 0.790 |
| | | 2 | 0.800 | 0.880 | 0.864 | 0.619 | 0.950 | 0.976 | 0.971 | 0.909 |
| | | 3 | 0.182 | 0.935 | 0.854 | 0.135 | 0.727 | 1.000 | 0.971 | 0.839 |
| | | 4 | 0.353 | 0.837 | 0.757 | 0.178 | 0.700 | 0.940 | 0.893 | 0.652 |
| | | 5 | 0.722 | 0.965 | 0.922 | 0.720 | 0.889 | 0.988 | 0.971 | 0.897 |
| | Validation set | 1 | 0.786 | 0.825 | 0.815 | 0.568 | 0.941 | 1.000 | 0.981 | 0.957 |
| | | 2 | 1.000 | 0.907 | 0.926 | 0.816 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.667 | 0.979 | 0.944 | 0.700 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.200 | 0.955 | 0.815 | 0.229 | 1.000 | 0.977 | 0.981 | 0.943 |
| | | 5 | 0.900 | 1.000 | 0.981 | 0.938 | 1.000 | 1.000 | 1.000 | 1.000 |
| 5th | Calibration set | 1 | 0.833 | 0.871 | 0.857 | 0.697 | 0.917 | 0.935 | 0.929 | 0.847 |
| | | 2 | 0.800 | 0.910 | 0.888 | 0.675 | 0.950 | 0.987 | 0.980 | 0.937 |
| | | 3 | 0.200 | 0.943 | 0.867 | 0.168 | 0.900 | 0.989 | 0.980 | 0.889 |
| | | 4 | 0.563 | 0.976 | 0.908 | 0.630 | 0.813 | 0.976 | 0.949 | 0.809 |
| | | 5 | 0.875 | 0.939 | 0.929 | 0.761 | 0.938 | 0.988 | 0.980 | 0.925 |
| | Validation set | 1 | 0.833 | 0.727 | 0.765 | 0.537 | 1.000 | 0.788 | 0.863 | 0.753 |
| | | 2 | 1.000 | 0.925 | 0.941 | 0.853 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.400 | 1.000 | 0.944 | 0.614 | 1.000 | 1.000 | 1.000 | 1.000 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 4 | 0.333 | 0.976 | 0.863 | 0.439 | 0.222 | 1.000 | 0.863 | 0.436 |
| | | 5 | 0.875 | 1.000 | 0.980 | 0.925 | 1.000 | 1.000 | 1.000 | 1.000 |
| 6th | Calibration set | 1 | 0.588 | 0.786 | 0.729 | 0.362 | 0.941 | 0.977 | 0.967 | 0.918 |
| | | 2 | 0.500 | 0.959 | 0.881 | 0.533 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.600 | 0.878 | 0.831 | 0.445 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.591 | 0.784 | 0.712 | 0.378 | 0.955 | 0.973 | 0.966 | 0.928 |
| | Validation set | 1 | 0.667 | 1.000 | 0.897 | 0.761 | 0.667 | 1.000 | 0.897 | 0.761 |
| | | 2 | 0.800 | 0.958 | 0.931 | 0.758 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.800 | 0.875 | 0.862 | 0.596 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.900 | 0.895 | 0.897 | 0.779 | 1.000 | 0.833 | 0.897 | 0.809 |
| 7th | Calibration set | 1 | 0.500 | 0.898 | 0.800 | 0.429 | 0.875 | 0.980 | 0.954 | 0.874 |
| | | 2 | 0.412 | 0.771 | 0.677 | 0.179 | 0.941 | 1.000 | 0.985 | 0.960 |
| | | 3 | 0.462 | 0.865 | 0.785 | 0.327 | 0.923 | 0.981 | 0.969 | 0.904 |
| | | 4 | 0.526 | 0.761 | 0.692 | 0.279 | 0.950 | 0.956 | 0.954 | 0.894 |
| | Validation set | 1 | 0.778 | 0.962 | 0.914 | 0.770 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 2 | 0.778 | 0.962 | 0.914 | 0.770 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 0.857 | 0.929 | 0.914 | 0.748 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 4 | 0.900 | 0.920 | 0.914 | 0.798 | 1.000 | 1.000 | 1.000 | 1.000 |
| 8th | Calibration set | 1 | 0.833 | 0.789 | 0.804 | 0.591 | 0.889 | 0.921 | 0.911 | 0.799 |
| | | 2 | 0.167 | 0.886 | 0.732 | 0.066 | 0.667 | 1.000 | 0.929 | 0.782 |
| | | 3 | 0.600 | 0.978 | 0.911 | 0.670 | 1.000 | 0.978 | 0.982 | 0.943 |
| | | 4 | 0.625 | 0.775 | 0.732 | 0.382 | 1.000 | 0.950 | 0.964 | 0.919 |
| | Validation set | 1 | 0.600 | 1.000 | 0.871 | 0.710 | 1.000 | 0.950 | 0.967 | 0.929 |
| | | 2 | 0.429 | 0.958 | 0.839 | 0.483 | 1.000 | 1.000 | 1.000 | 1.000 |
| | | 3 | 1.000 | 0.923 | 0.935 | 0.812 | 0.800 | 1.000 | 0.968 | 0.878 |
| | | 4 | 1.000 | 0.773 | 0.839 | 0.705 | 1.000 | 1.000 | 1.000 | 1.000 |

## Table S5

| Preprocessing | $R^2$ | $Q^2$ | Calibration set | | | Validation set | |
|---|---|---|---|---|---|---|---|
| | | | RMSEE | RMSECV | Accuracy | RMSEP | Accuracy |
| **Raw** | **0.571** | **0.391** | **0.263298** | **0.319616** | **87.16%** | **0.261651** | **87.55%** |
| **SNV-SG5-SD** | **0.735** | **0.574** | **0.205275** | **0.267485** | **98.34%** | **0.216638** | **93.78%** |
| SNV-SG7-SD | 0.739 | 0.558 | 0.203830 | 0.272793 | 98.14% | 0.221316 | 92.95% |
| SNV-SG9-SD | 0.745 | 0.548 | 0.201681 | 0.280961 | 98.14% | 0.224041 | 92.95% |
| SNV-SG11-SD | 0.712 | 0.561 | 0.213827 | 0.271916 | 96.48% | 0.231361 | 92.12% |
| SNV-SG13-SD | 0.724 | 0.525 | 0.210035 | 0.280905 | 97.10% | 0.226049 | 95.02% |

## Table S6

| Set | Classes [a] | Raw | | | | The best preprocessing | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | SE | SP | Accuracy | MCC | SE | SP | Accuracy | MCC |
| Calibration set | 1 | 0.883 | 0.931 | 0.917 | 0.800 | 0.971 | 0.988 | 0.983 | 0.959 |
| | 2 | 0.755 | 0.960 | 0.913 | 0.745 | 0.982 | 0.997 | 0.994 | 0.982 |
| | 3 | 0.899 | 0.986 | 0.973 | 0.889 | 0.986 | 0.995 | 0.994 | 0.975 |
| | 4 | 0.912 | 0.970 | 0.957 | 0.879 | 0.991 | 0.997 | 0.996 | 0.988 |
| | 5 | 0.963 | 0.986 | 0.983 | 0.920 | 1.000 | 1.000 | 1.000 | 1.000 |
| Validation set | 1 | 0.812 | 0.977 | 0.929 | 0.824 | 0.870 | 0.983 | 0.950 | 0.877 |
| | 2 | 0.873 | 0.930 | 0.917 | 0.775 | 0.982 | 0.984 | 0.983 | 0.954 |
| | 3 | 0.853 | 0.990 | 0.971 | 0.877 | 0.941 | 1.000 | 0.992 | 0.965 |
| | 4 | 0.912 | 0.946 | 0.938 | 0.834 | 0.965 | 0.962 | 0.963 | 0.901 |
| | 5 | 1.000 | 0.995 | 0.996 | 0.979 | 0.962 | 0.991 | 0.988 | 0.937 |

## Table S7

| Classes [a] | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1.000 | 0.665 | 0.390 | 0.452 | 0.418 | 0.314 | 0.476 | 0.327 | 0.369 | 0.239 | 0.387 | 0.447 | 0.150 | 0.246 | 0.283 | -0.05 | 0.175 | 0.148 | 0.341 | 0.113 | 0.098 | 0.002 | 0.351 | -0.08 | -0.09 | -0.44 | 0.051 |
| 2 | 0.665 | 1.000 | 0.190 | 0.420 | 0.232 | 0.106 | 0.643 | 0.120 | 0.260 | 0.062 | 0.315 | 0.615 | -0.05 | 0.147 | 0.141 | -0.15 | -0.02 | -0.05 | 0.312 | 0.107 | 0.219 | 0 | 0.345 | -0.21 | -0.18 | -0.49 | -0.01 |
| 3 | 0.390 | 0.190 | 1.000 | 0.367 | 0.362 | 0.300 | 0.123 | 0.425 | 0.342 | 0.105 | 0.173 | 0.036 | 0.432 | 0.241 | 0.078 | 0.063 | 0.130 | 0.487 | 0.190 | 0.092 | -0.07 | 0.446 | 0.220 | 0.149 | 0.149 | -0.18 | 0.141 |
| 4 | 0.452 | 0.420 | 0.367 | 1.000 | 0.435 | 0.117 | 0.291 | 0.270 | 0.454 | 0.387 | 0.408 | 0.250 | 0.342 | 0.209 | 0.487 | -0.17 | 0.055 | 0.319 | 0.488 | 0.136 | 0.372 | 0.156 | 0.513 | -0.23 | 0.057 | -0.45 | 0.515 |
| 5 | 0.418 | 0.232 | 0.362 | 0.435 | 1.000 | 0.244 | 0.240 | 0.583 | 0.499 | 0.637 | 0.057 | 0.179 | 0.437 | 0.270 | 0.777 | -0.05 | 0.097 | 0.361 | 0.323 | -0.09 | 0.117 | -0.01 | 0.352 | -0.02 | 0.027 | -0.36 | 0.380 |
| 6 | 0.314 | 0.106 | 0.300 | 0.117 | 0.244 | 1.000 | 0.219 | 0.556 | 0.723 | 0.368 | 0.374 | -0.23 | 0.232 | 0.464 | 0.135 | 0.696 | 0.504 | 0.450 | 0.551 | 0.276 | -0.31 | 0.171 | 0.087 | 0.544 | 0.352 | 0.081 | 0.135 |
| 7 | 0.476 | 0.643 | 0.123 | 0.291 | 0.240 | 0.219 | 1.000 | 0.340 | 0.452 | 0.206 | 0.139 | 0.569 | -0.04 | 0.100 | 0.154 | -0.07 | 0.047 | 0.083 | 0.441 | -0.01 | 0.192 | -0.07 | 0.185 | -0.22 | -0.12 | -0.41 | 0.103 |
| 8 | 0.327 | 0.120 | 0.425 | 0.270 | 0.583 | 0.556 | 0.340 | 1.000 | 0.668 | 0.378 | 0.036 | -0.02 | 0.524 | 0.236 | 0.405 | 0.204 | 0.344 | 0.639 | 0.425 | -0.13 | -0.13 | 0.024 | 0.121 | 0.170 | 0.336 | -0.14 | 0.296 |
| 9 | 0.369 | 0.260 | 0.342 | 0.454 | 0.499 | 0.723 | 0.452 | 0.668 | 1.000 | 0.563 | 0.283 | -0.01 | 0.358 | 0.434 | 0.450 | 0.407 | 0.414 | 0.542 | 0.727 | 0.104 | -0.06 | 0.071 | 0.281 | 0.209 | 0.269 | -0.25 | 0.329 |
| 10 | 0.239 | 0.062 | 0.105 | 0.387 | 0.637 | 0.368 | 0.206 | 0.378 | 0.563 | 1.000 | 0.183 | 0.050 | 0.222 | 0.191 | 0.701 | 0.103 | 0.288 | 0.308 | 0.520 | -0.16 | 0.184 | -0.08 | 0.270 | -0.09 | 0.111 | -0.32 | 0.450 |
| 11 | 0.387 | 0.315 | 0.173 | 0.408 | 0.057 | 0.374 | 0.139 | 0.036 | 0.283 | 0.183 | 1.000 | 0.213 | 0.189 | 0.378 | 0.205 | 0.183 | 0.232 | 0.071 | 0.429 | 0.511 | 0.316 | 0.376 | 0.467 | 0.012 | 0.190 | -0.09 | 0.223 |
| 12 | 0.447 | 0.615 | 0.036 | 0.250 | 0.179 | -0.23 | 0.569 | -0.02 | -0.01 | 0.050 | 0.213 | 1.000 | -0.10 | 0.022 | 0.154 | -0.36 | -0.15 | -0.18 | 0.107 | -0.05 | 0.571 | -0.06 | 0.349 | -0.54 | -0.22 | -0.52 | 0.064 |
| 13 | 0.150 | -0.05 | 0.432 | 0.342 | 0.437 | 0.232 | -0.04 | 0.524 | 0.358 | 0.222 | 0.189 | -0.10 | 1.000 | 0.162 | 0.367 | 0.040 | 0.300 | 0.623 | 0.245 | 0.064 | 0.091 | 0.284 | 0.275 | 0.150 | 0.439 | -0.08 | 0.366 |
| 14 | 0.246 | 0.147 | 0.241 | 0.209 | 0.270 | 0.464 | 0.100 | 0.236 | 0.434 | 0.191 | 0.378 | 0.022 | 0.162 | 1.000 | 0.336 | 0.452 | -0.01 | 0.200 | 0.388 | 0.437 | -0.07 | 0.185 | 0.312 | 0.350 | -0.03 | -0.08 | 0.077 |
| 15 | 0.283 | 0.141 | 0.078 | 0.487 | 0.777 | 0.135 | 0.154 | 0.405 | 0.450 | 0.701 | 0.205 | 0.154 | 0.367 | 0.336 | 1.000 | -0.08 | 0.034 | 0.246 | 0.388 | -0.03 | 0.253 | -0.03 | 0.388 | -0.14 | 0.009 | -0.35 | 0.462 |
| 16 | -0.05 | -0.15 | 0.063 | -0.17 | -0.05 | 0.696 | -0.07 | 0.204 | 0.407 | 0.103 | 0.183 | -0.36 | 0.040 | 0.452 | -0.08 | 1.000 | 0.270 | 0.201 | 0.385 | 0.480 | -0.35 | 0.217 | 0.073 | 0.668 | 0.254 | 0.323 | -0.01 |
| 17 | 0.175 | -0.02 | 0.130 | 0.055 | 0.097 | 0.504 | 0.047 | 0.344 | 0.414 | 0.288 | 0.232 | -0.15 | 0.300 | -0.01 | 0.034 | 0.270 | 1.000 | 0.387 | 0.280 | -0.05 | 0.019 | -0.04 | 0.027 | 0.174 | 0.420 | 0.055 | 0.213 |
| 18 | 0.148 | -0.04 | 0.487 | 0.319 | 0.361 | 0.450 | 0.083 | 0.639 | 0.542 | 0.308 | 0.071 | -0.18 | 0.623 | 0.200 | 0.246 | 0.201 | 0.387 | 1.000 | 0.352 | 0.011 | -0.08 | 0.222 | 0.087 | 0.265 | 0.429 | -0.07 | 0.318 |
| 19 | 0.341 | 0.312 | 0.190 | 0.488 | 0.323 | 0.551 | 0.441 | 0.425 | 0.727 | 0.520 | 0.429 | 0.107 | 0.245 | 0.388 | 0.388 | 0.385 | 0.280 | 0.352 | 1.000 | 0.256 | 0.172 | 0.031 | 0.560 | 0.056 | 0.184 | -0.28 | 0.502 |
| 20 | 0.113 | 0.107 | 0.092 | 0.136 | -0.09 | 0.276 | -0.09 | -0.13 | 0.104 | -0.16 | 0.511 | -0.05 | 0.064 | 0.437 | -0.03 | 0.480 | -0.05 | 0.011 | 0.256 | 1.000 | -0.04 | 0.403 | 0.351 | 0.409 | 0.133 | 0.258 | 0.030 |
| 21 | 0.098 | 0.219 | -0.07 | 0.372 | 0.117 | -0.31 | 0.192 | -0.13 | -0.06 | 0.184 | 0.316 | 0.571 | 0.091 | -0.07 | 0.253 | -0.30 | 0.019 | -0.08 | 0.172 | -0.04 | 1.000 | 0.058 | 0.530 | -0.61 | -0.02 | -0.37 | 0.440 |
| 22 | 0.002 | 0 | 0.446 | 0.156 | -0.01 | 0.171 | -0.07 | 0.024 | 0.071 | -0.08 | 0.376 | -0.06 | 0.284 | 0.185 | -0.03 | 0.217 | -0.04 | 0.222 | 0.031 | 0.403 | 0.058 | 1.000 | 0.222 | 0.215 | 0.183 | 0.198 | 0.050 |
| 23 | 0.351 | 0.345 | 0.220 | 0.513 | 0.352 | 0.087 | 0.185 | 0.121 | 0.281 | 0.270 | 0.467 | 0.349 | 0.275 | 0.312 | 0.388 | 0.073 | 0.027 | 0.087 | 0.560 | 0.351 | 0.530 | 0.222 | 1.000 | -0.17 | -0.01 | -0.39 | 0.471 |
| 24 | -0.08 | -0.21 | 0.149 | -0.23 | -0.02 | 0.544 | -0.22 | 0.170 | 0.209 | -0.09 | 0.012 | -0.54 | 0.150 | 0.350 | -0.14 | 0.668 | 0.174 | 0.265 | 0.056 | 0.409 | -0.61 | 0.215 | -0.17 | 1.000 | 0.370 | 0.565 | -0.07 |
| 25 | -0.08 | -.018 | 0.149 | 0.057 | 0.027 | 0.352 | -0.12 | 0.336 | 0.269 | 0.111 | 0.190 | -0.22 | 0.439 | -0.03 | 0.009 | 0.254 | 0.420 | 0.429 | 0.184 | 0.133 | -0.02 | 0.183 | -0.01 | 0.370 | 1.000 | 0.327 | 0.324 |
| 26 | -0.44 | -0.49 | -0.17 | -0.44 | -0.36 | 0.081 | -0.41 | -0.14 | -0.25 | -0.32 | -0.09 | -0.52 | -0.08 | -0.08 | -0.35 | 0.323 | 0.055 | -0.07 | -0.28 | 0.258 | -0.37 | 0.198 | -0.39 | 0.565 | 0.327 | 1.000 | -0.18 |

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 27 | 0.051 | -0.01 | 0.141 | 0.515 | 0.380 | 0.135 | 0.103 | 0.296 | 0.329 | 0.450 | 0.223 | 0.064 | 0.366 | 0.077 | 0.462 | -0.01 | 0.213 | 0.318 | 0.502 | 0.030 | 0.440 | 0.050 | 0.471 | -0.07 | 0.324 | -0.18 | 1.000 |

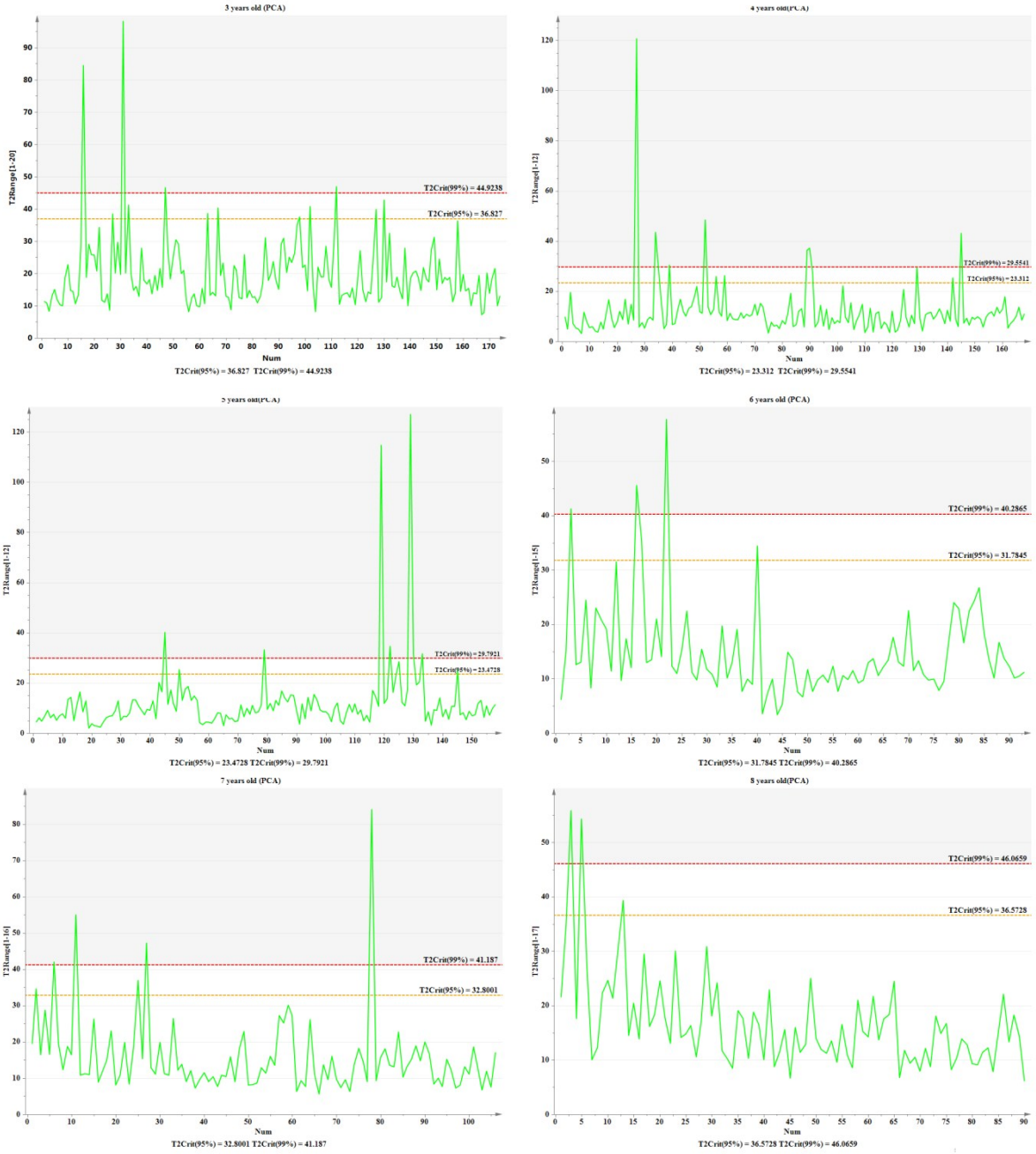
3 years old (PCA)
T2Range[1-20]
T2Crit(99%) = 44.9238
T2Crit(95%) = 36.827
Num
T2Crit(95%) = 36.827 T2Crit(99%) = 44.9238
4 years old(PCA)
T2Range[1-12]
T2Crit(99%) = 29.5541
T2Crit(95%) = 23.312
Num
T2Crit(95%) = 23.312 T2Crit(99%) = 29.5541
5 years old(PCA)
T2Range[1-12]
T2Crit(99%) = 29.7921
T2Crit(95%) = 23.4728
Num
T2Crit(95%) = 23.4728 T2Crit(99%) = 29.7921
6 years old (PCA)
T2Range[1-15]
T2Crit(99%) = 40.2865
T2Crit(95%) = 31.7845
Num
T2Crit(95%) = 31.7845 T2Crit(99%) = 40.2865
7 years old (PCA)
T2Range[1-16]
T2Crit(99%) = 41.187
T2Crit(95%) = 32.8001
Num
T2Crit(95%) = 32.8001 T2Crit(99%) = 41.187
8 years old(PCA)
T2Range[1-17]
T2Crit(99%) = 46.0659
T2Crit(95%) = 36.5728
Num
T2Crit(95%) = 36.5728 T2Crit(99%) = 46.0659

**Fig. S1**

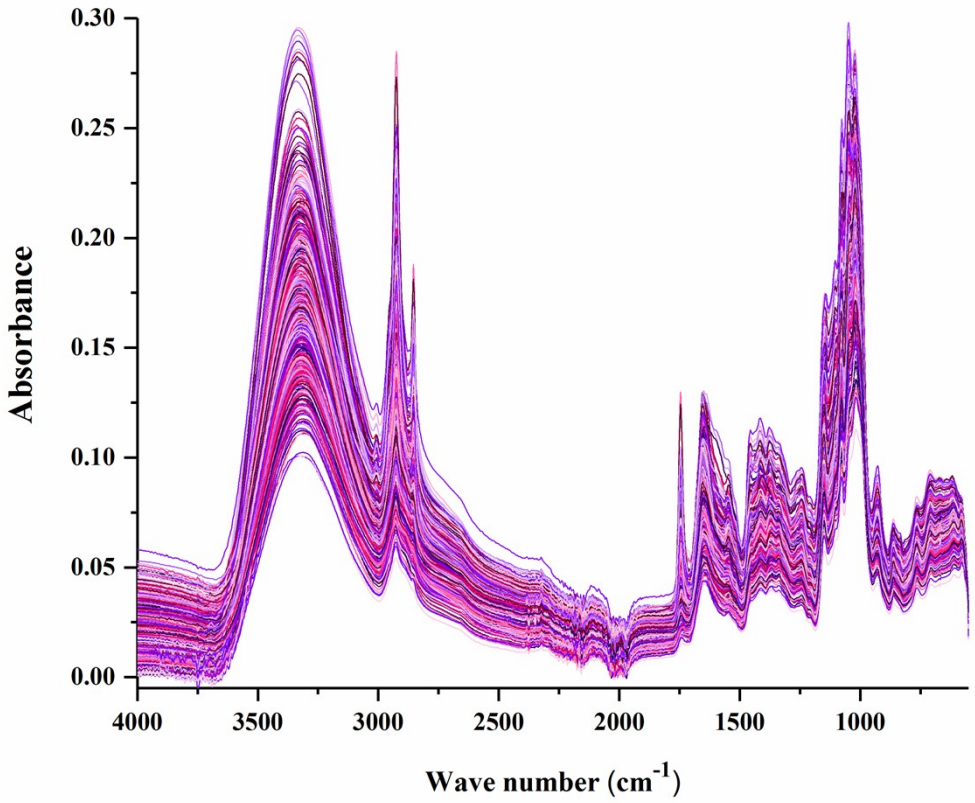
Absorbance
0.30
0.25
0.20
0.15
0.10
0.05
0.00
4000
3500
3000
2500
2000
1500
1000
Wave number (cm-1)

**Fig. S2**

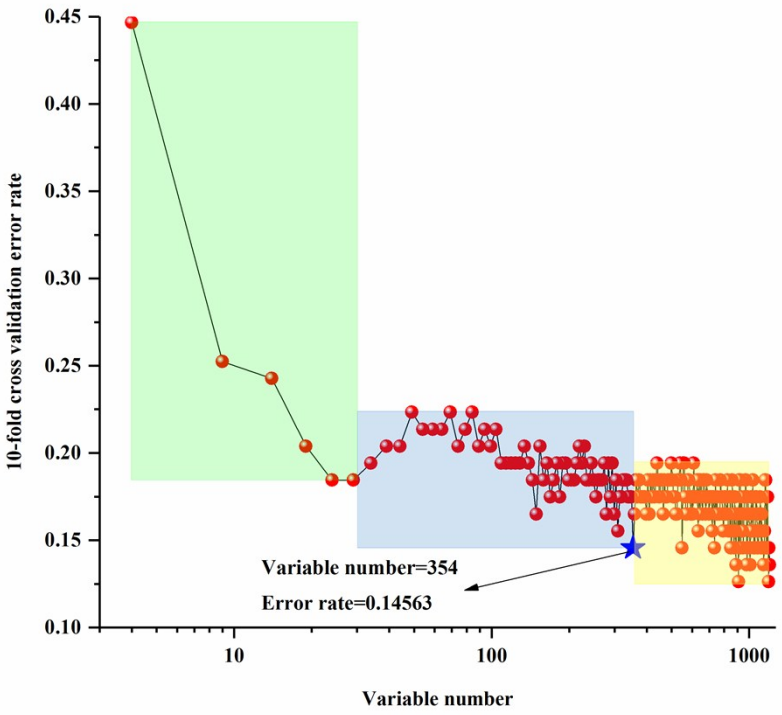
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
10-fold cross validation error rate
Variable number=354
Error rate=0.14563
10
100
1000
Variable number

**Fig. S3**

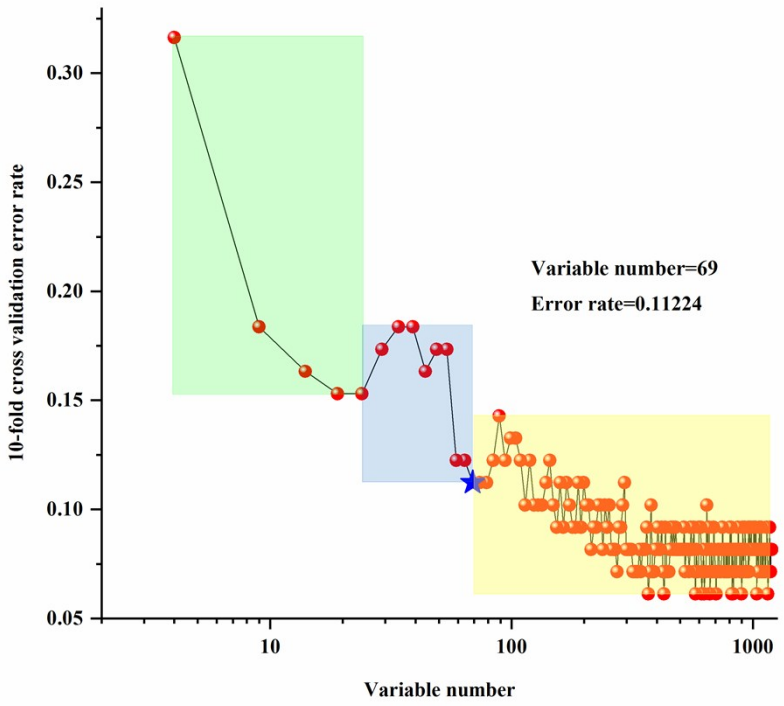
10-fold cross validation error rate
0.30
0.25
0.20
0.15
0.10
0.05
10
100
1000
Variable number
Variable number=69
Error rate=0.11224

**Fig. S4**

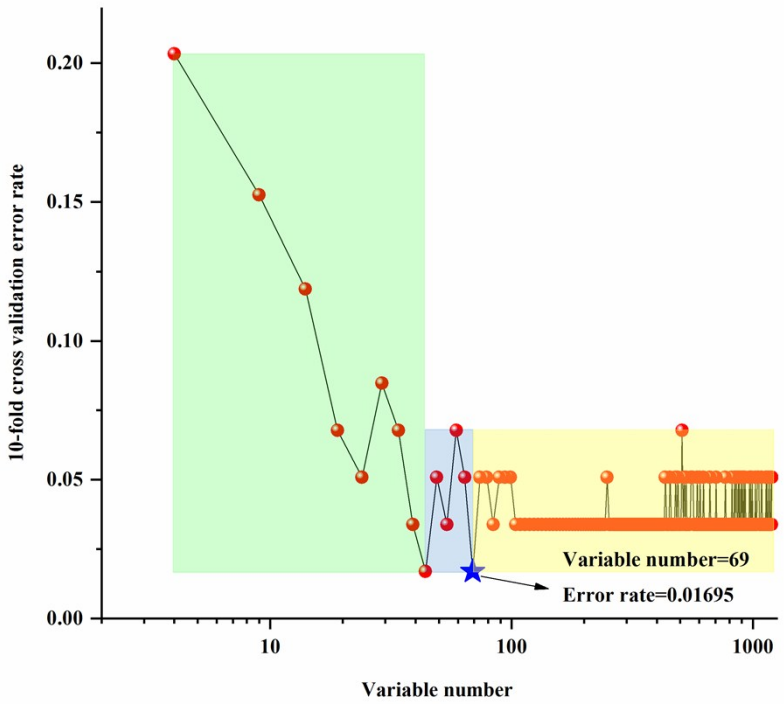
10-fold cross validation error rate
0.20
0.15
0.10
0.05
0.00
10
100
1000
Variable number
Variable number=69
Error rate=0.01695

**Fig. S5**

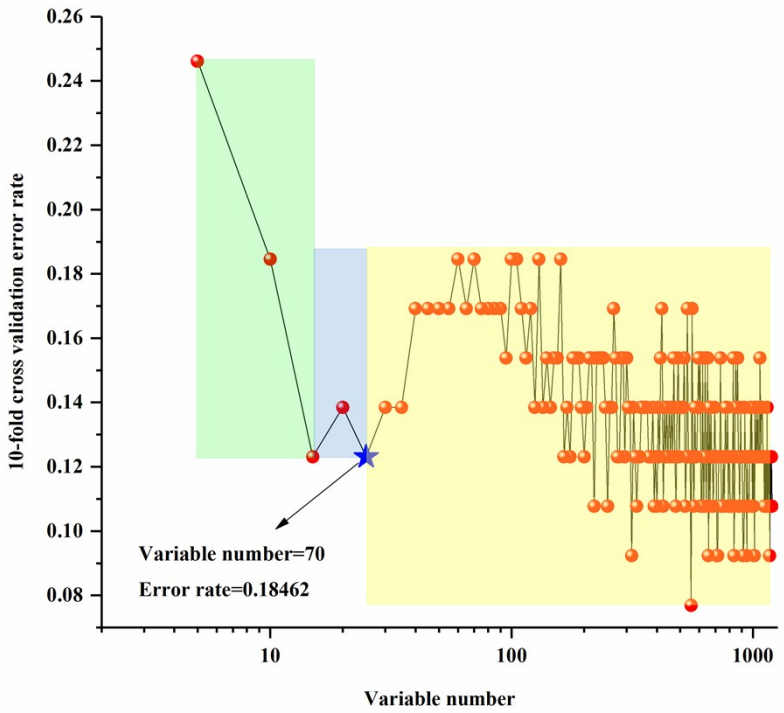
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0.10
0.08
10-fold cross validation error rate
Variable number=70
Error rate=0.18462
10
100
1000
Variable number

**Fig. S6**

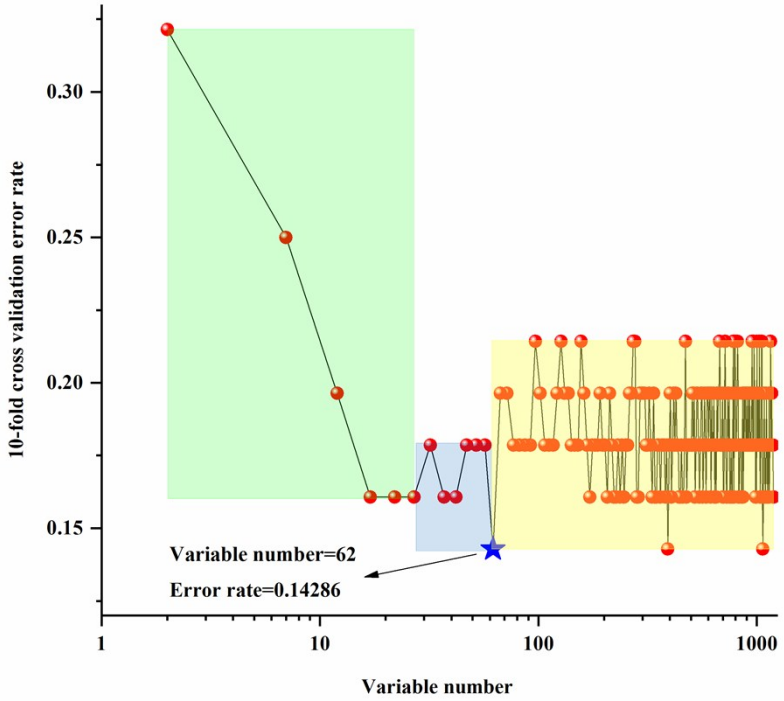
10-fold cross validation error rate
0.30
0.25
0.20
0.15
Variable number=62
Error rate=0.14286
1
10
100
1000
Variable number

**Fig. S7**

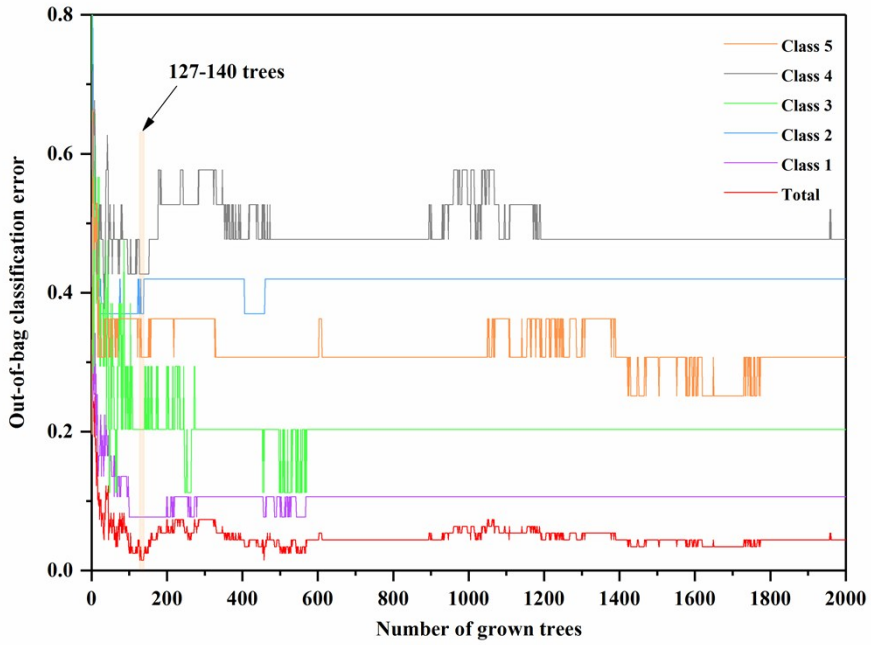
0.8
0.6
0.4
0.2
0.0
Out-of-bag classification error
127-140 trees
Class 5
Class 4
Class 3
Class 2
Class 1
Total
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Number of grown trees

**Fig. S8**

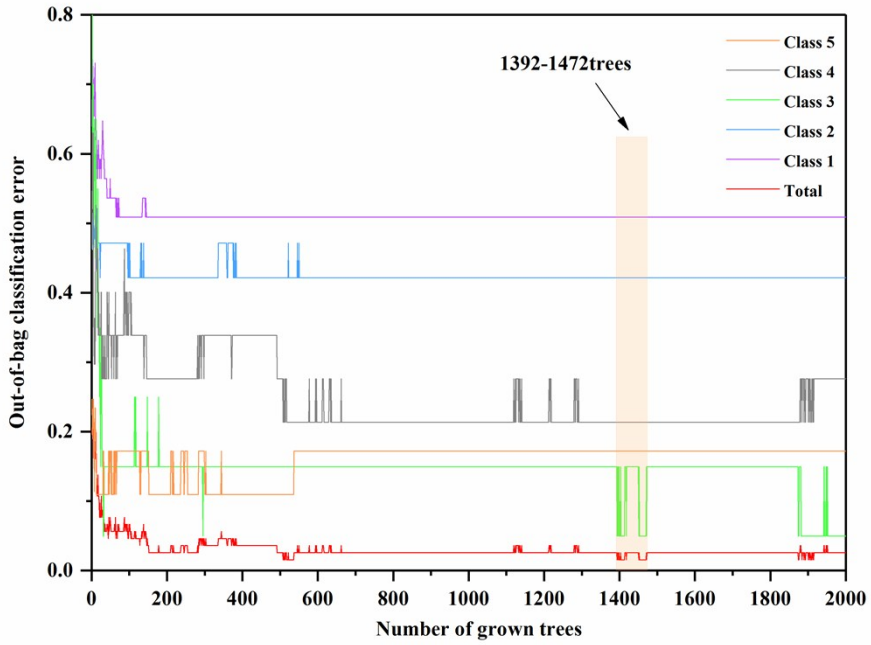
0.8
0.6
0.4
0.2
0.0
Out-of-bag classification error
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Number of grown trees
1392-1472trees
Class 5
Class 4
Class 3
Class 2
Class 1
Total

**Fig. S9**

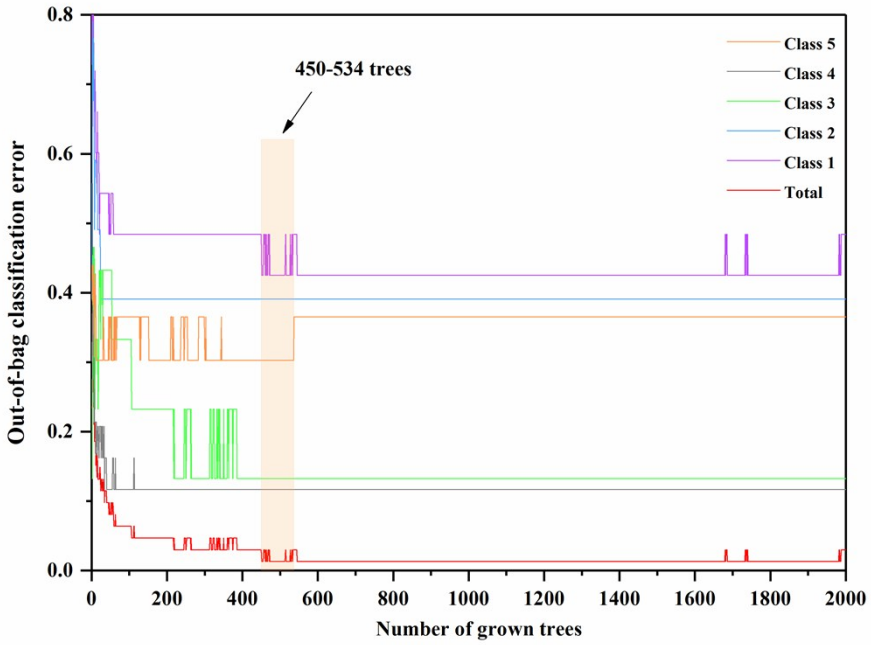
0.8
0.6
0.4
0.2
0.0
Out-of-bag classification error
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Number of grown trees
450-534 trees
Class 5
Class 4
Class 3
Class 2
Class 1
Total

**Fig. S10**

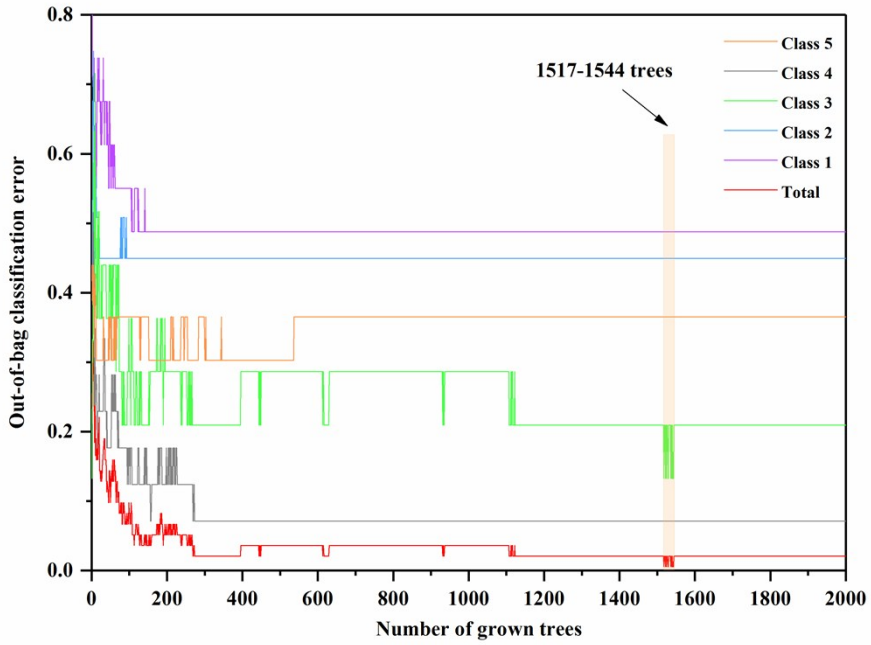
0.8
0.6
0.4
0.2
0.0
Out-of-bag classification error
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Number of grown trees
1517-1544 trees
Class 5
Class 4
Class 3
Class 2
Class 1
Total

**Fig. S11**

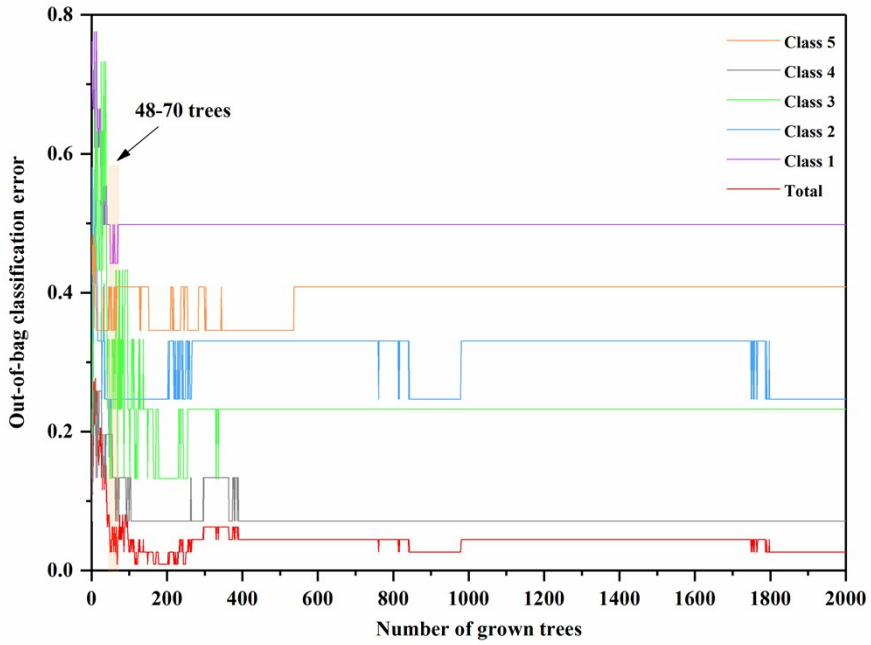
0.8
0.6
0.4
0.2
0.0
Out-of-bag classification error
48-70 trees
Class 5
Class 4
Class 3
Class 2
Class 1
Total
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Number of grown trees

**Fig. S12**

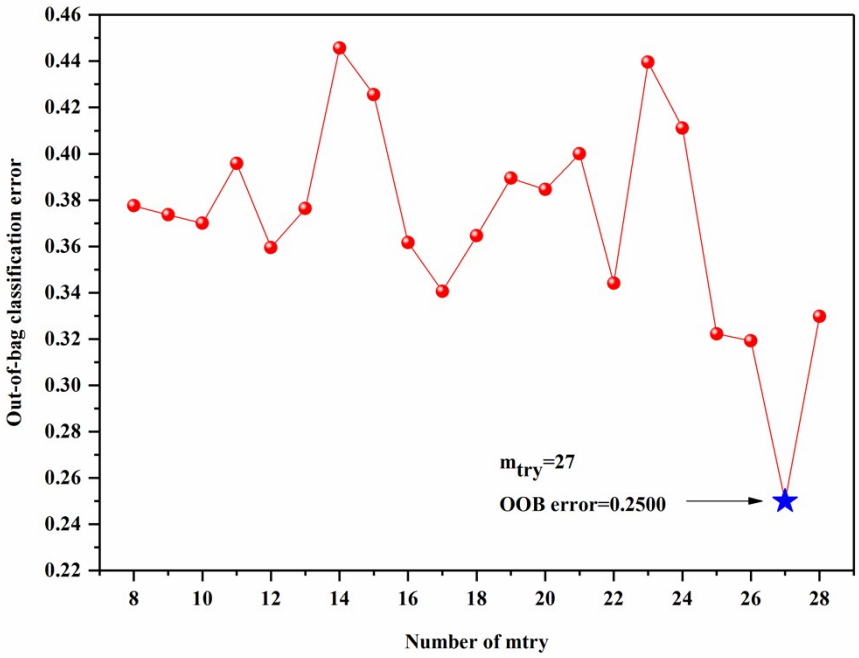
Out-of-bag classification error
0.46
0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
8
10
12
14
16
18
20
22
24
26
28
Number of mtry
m try=27
OOB error=0.2500

**Fig. S13**

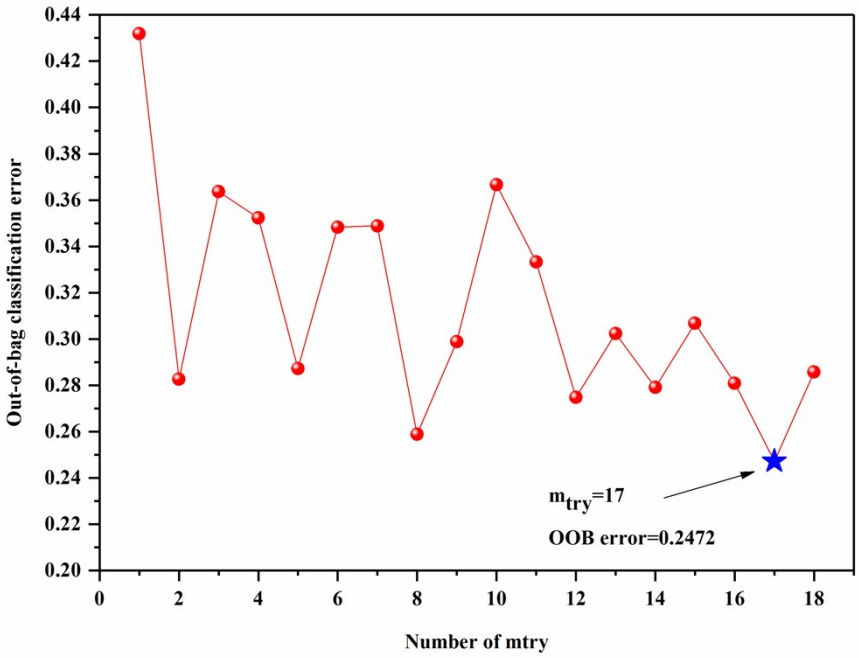
Out-of-bag classification error
0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0
2
4
6
8
10
12
14
16
18
Number of mtry
m try=17
OOB error=0.2472

**Fig. S14**

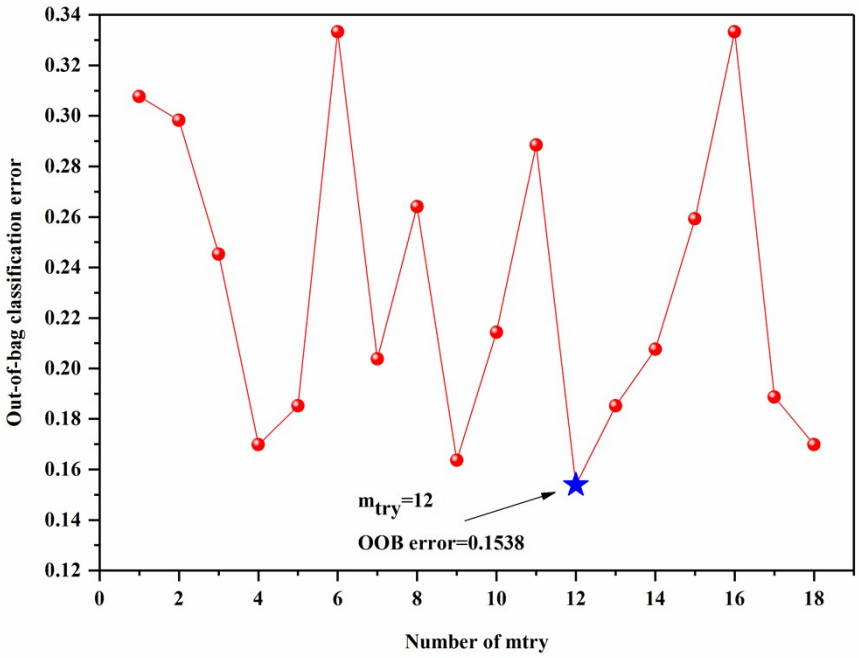
Out-of-bag classification error
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0
2
4
6
8
10
12
14
16
18
Number of mtry
m try=12
OOB error=0.1538

**Fig. S15**

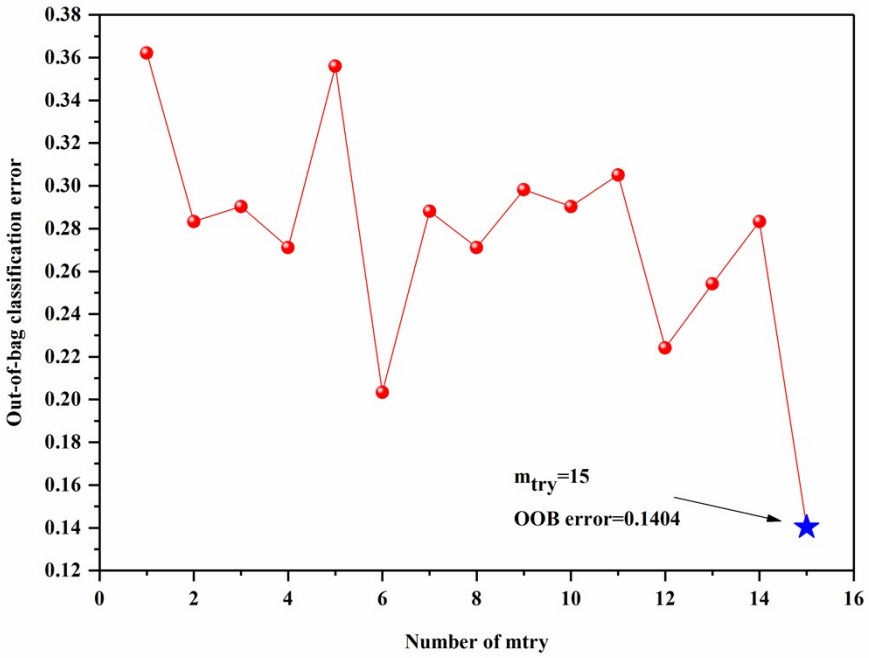
Out-of-bag classification error
Number of mtry
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0
2
4
6
8
10
12
14
16
m_try=15
OOB error=0.1404

**Fig. S16**

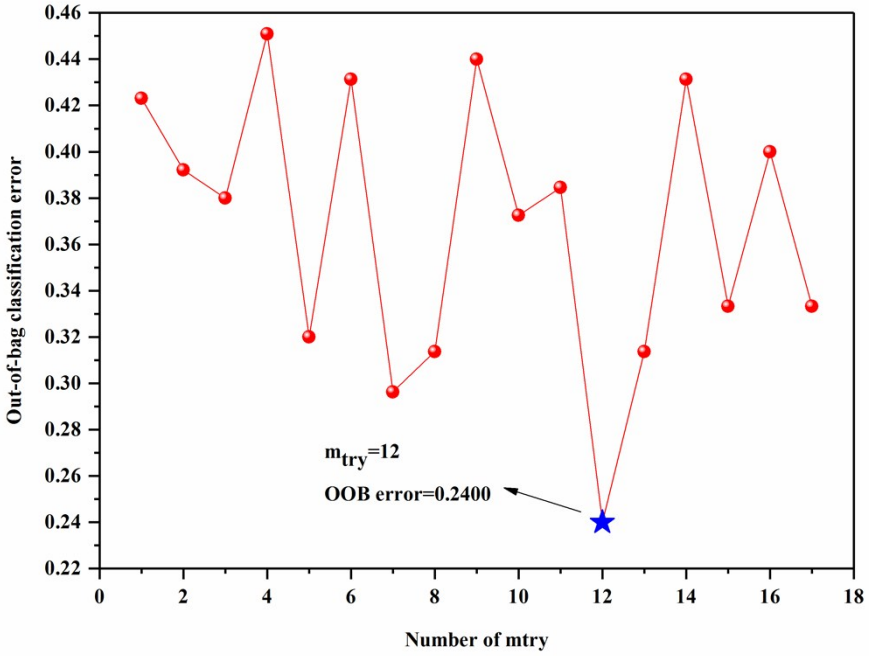
Out-of-bag classification error
0.46
0.44
0.42
0.40
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0
2
4
6
8
10
12
14
16
18
Number of mtry
m_try=12
OOB error=0.2400

**Fig. S17**

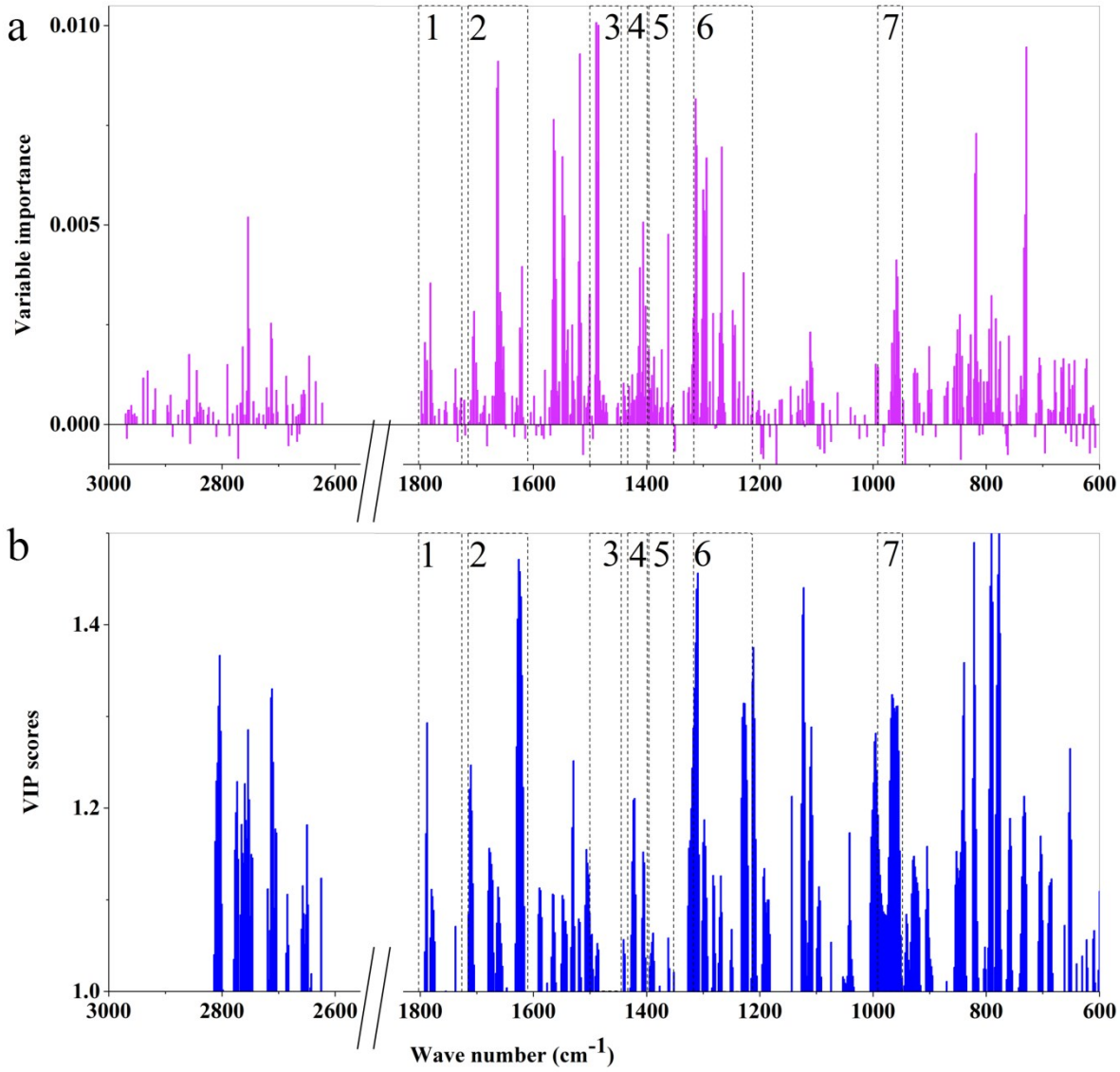
a
b
1
2
3
4
5
6
7
Variable importance
VIP scores
0.010
0.005
0.000
1.4
1.2
1.0
3000
2800
2600
1800
1600
1400
1200
1000
800
600
Wave number (cm-1)

**Fig. S18**

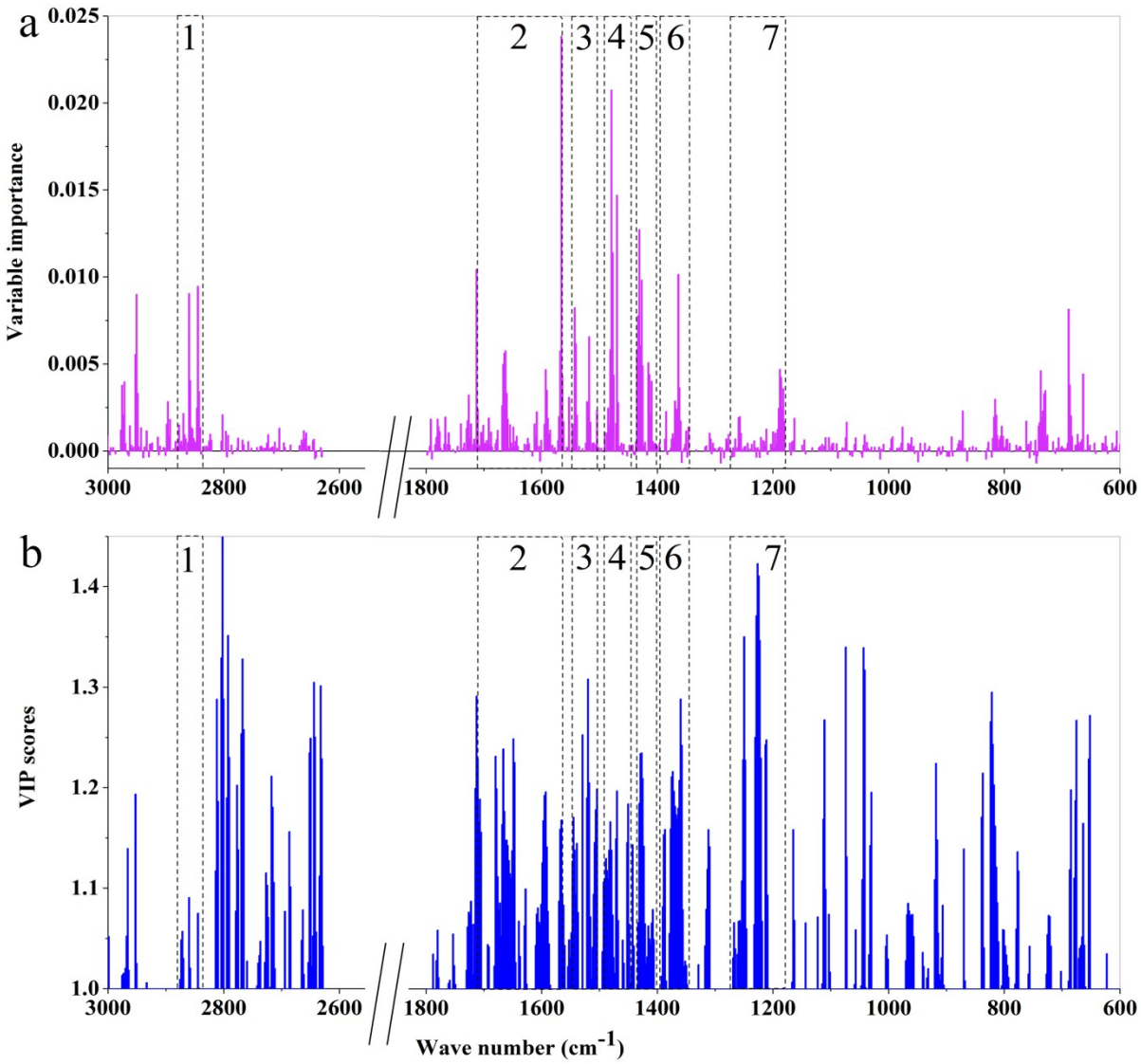
a
b
1
2
3
4
5
6
7
Variable importance
VIP scores
Wave number (cm-1)
0.025
0.020
0.015
0.010
0.005
0.000
1.4
1.3
1.2
1.1
1.0
3000
2800
2600
1800
1600
1400
1200
1000
800
600

**Fig. S19**

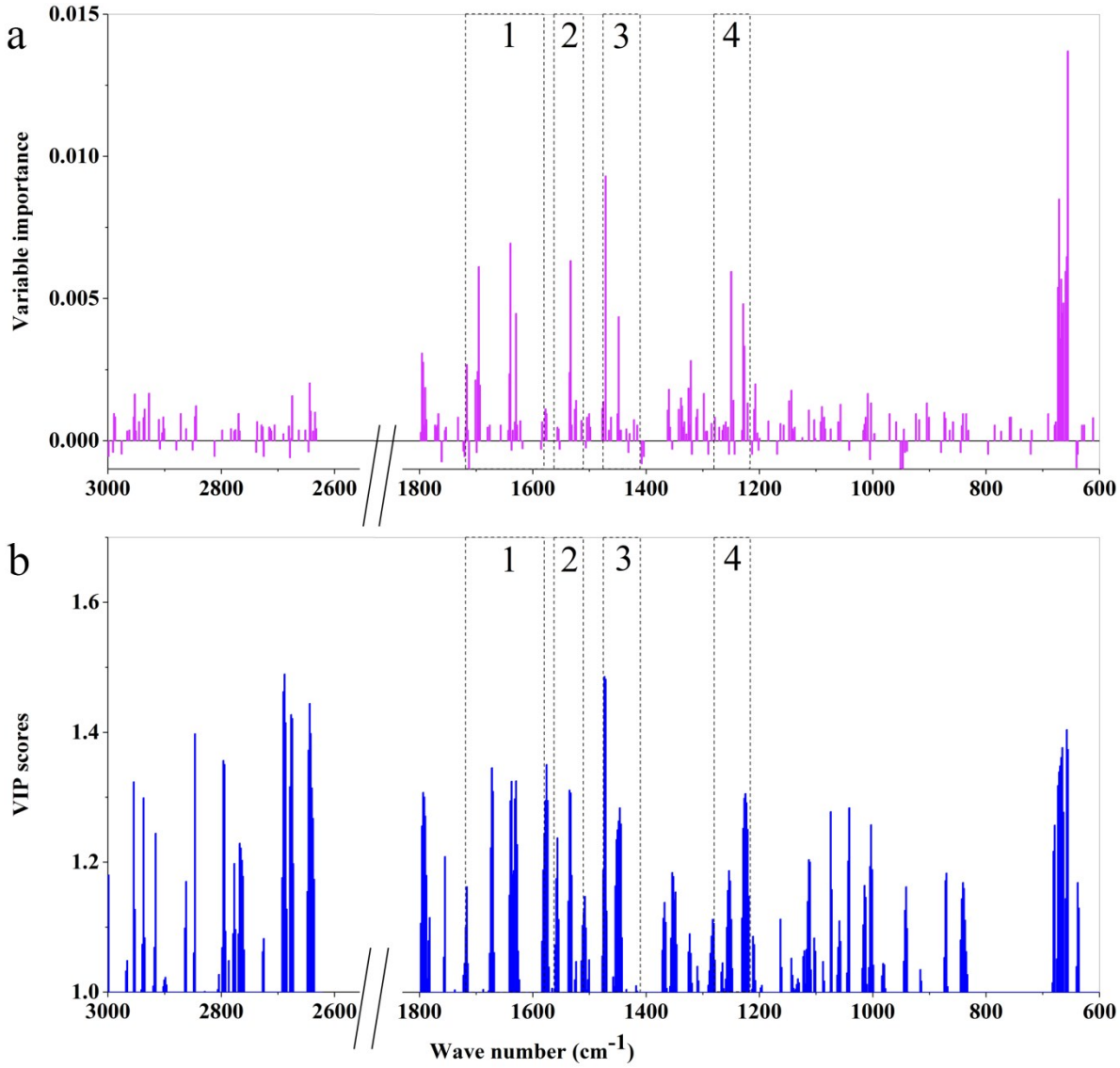
a
b
1
2
3
4
Variable importance
VIP scores
0.015
0.010
0.005
0.000
1.6
1.4
1.2
1.0
3000
2800
2600
1800
1600
1400
1200
1000
800
600
Wave number (cm-1)

**Fig. S20**

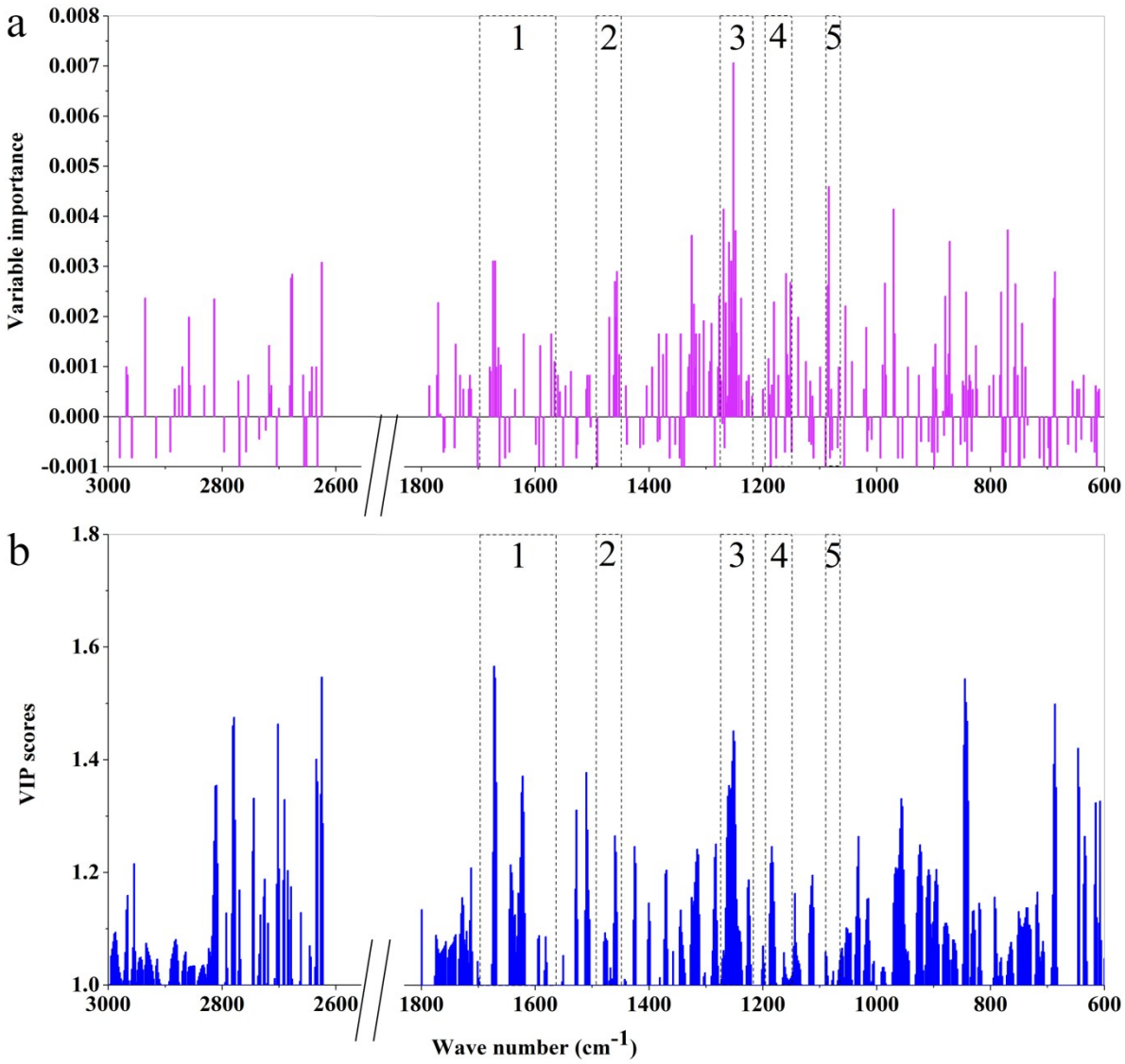
a
Variable importance
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
-0.001
1
2
3
4
5
3000
2800
2600
1800
1600
1400
1200
1000
800
600
b
VIP scores
1.8
1.6
1.4
1.2
1.0
Wave number (cm-1)

**Fig. S21**

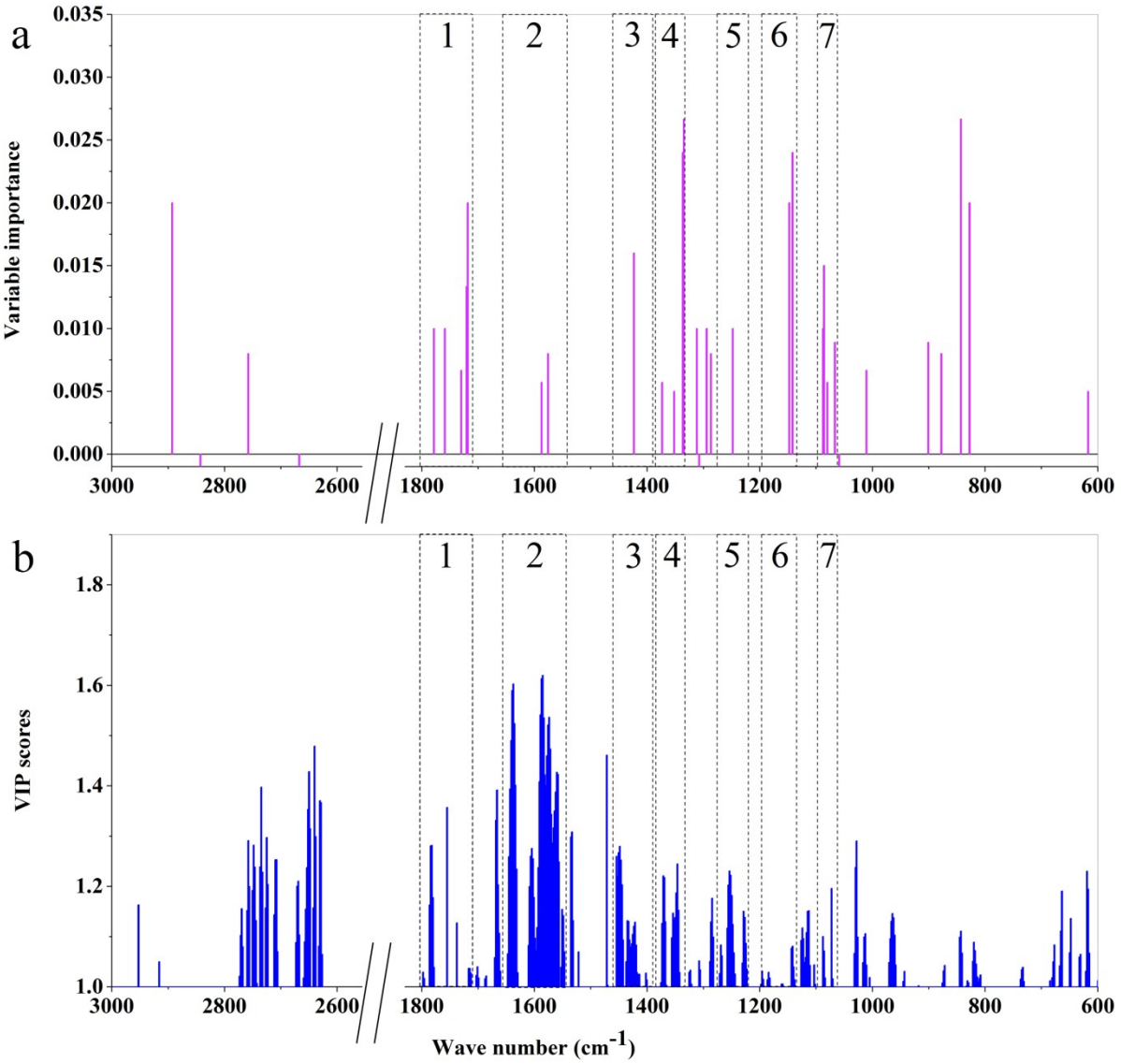
a
Variable importance
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
1
2
3
4
5
6
7
3000
2800
2600
1800
1600
1400
1200
1000
800
600
b
VIP scores
1.8
1.6
1.4
1.2
1.0
Wave number (cm-1)

**Fig. S22**

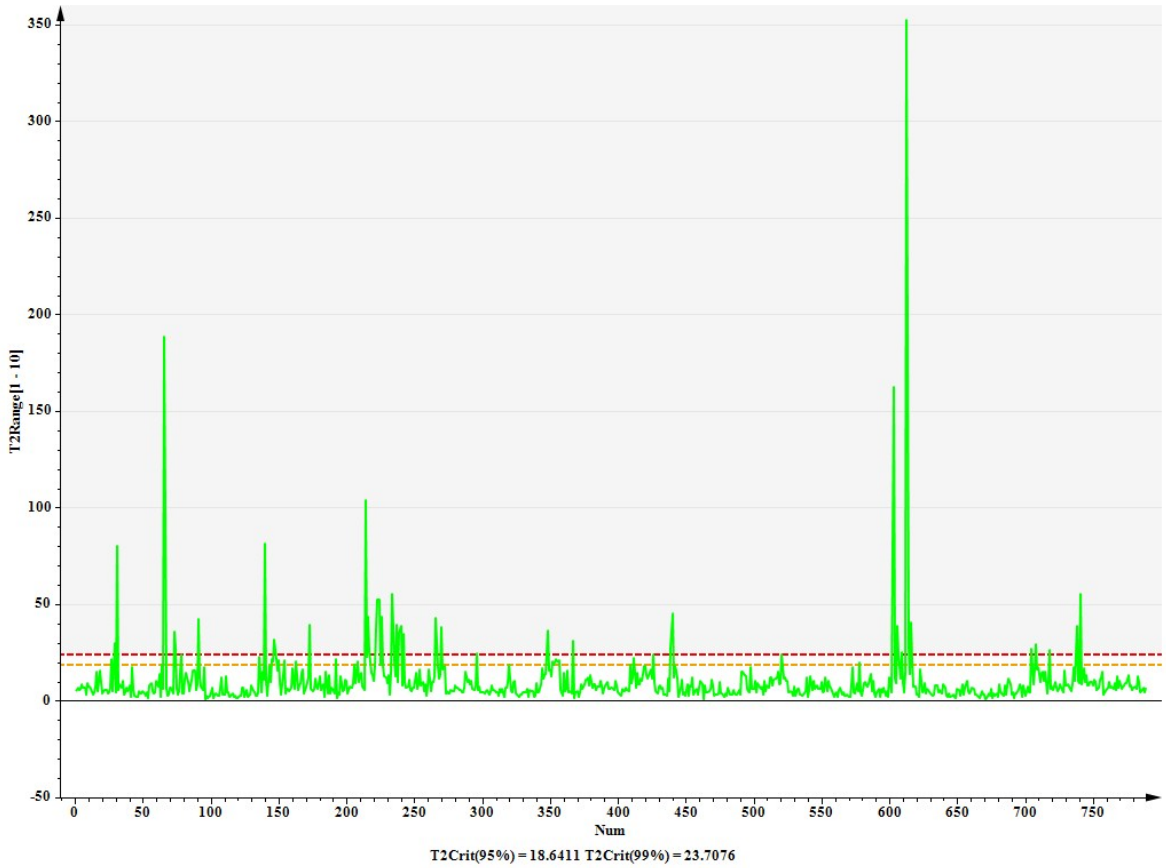
T2Range[1 - 10]
350
300
250
200
150
100
50
0
-50
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
Num
T2Crit(95%) = 18.6411 T2Crit(99%) = 23.7076

**Fig. S23**