

Figure 1. Histograms (top) and scatter plots (bottom) of RT difference, m/z difference, and Pearson correlation of FI between putative isotopologue pairs.

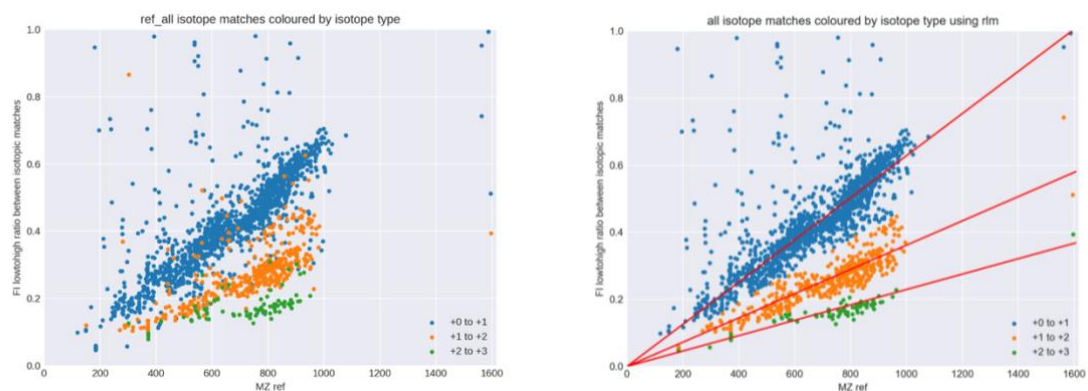


Figure 2. Feature intensity ratio vs m/z scatter plot of (left): all connections within large threshold, (right): reassigned results using RLM. The red lines on the right plot are the regression lines for each class.

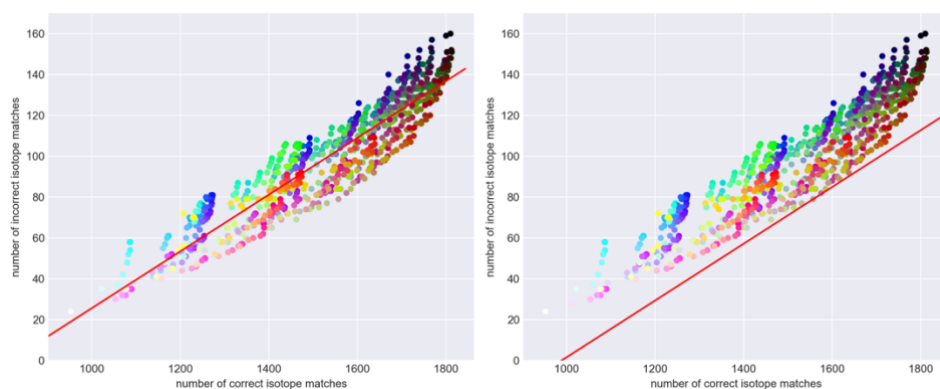


Figure 3. Number of incorrect vs number of correct matches (Left) linear regression results (Right) process of selecting thresholds. The colour used in these plots is RGB code encoded by [RT, m/z, correlation] rescaled from [[0.005, 0.001], [0.005, 0.001], [0.5, 0.9]] to [[0, 1], [0, 1], [0, 1]].

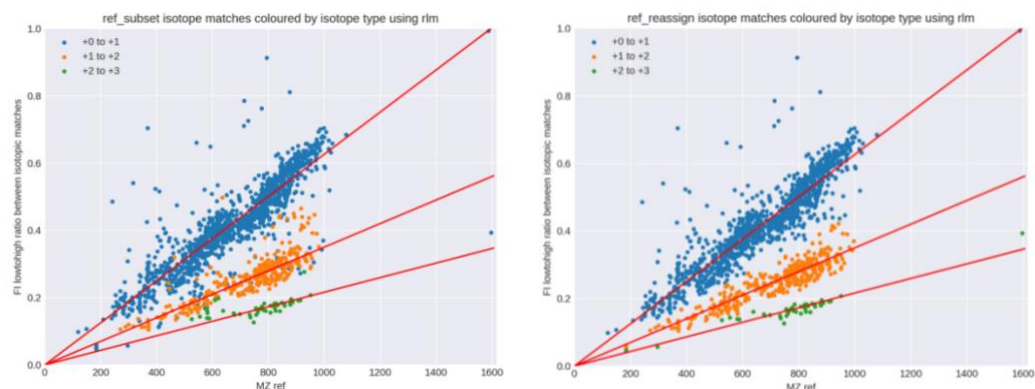


Figure 4. FI ratio vs m/z before (left) and after (right) reclassification using RLM with the optimal set of thresholds.

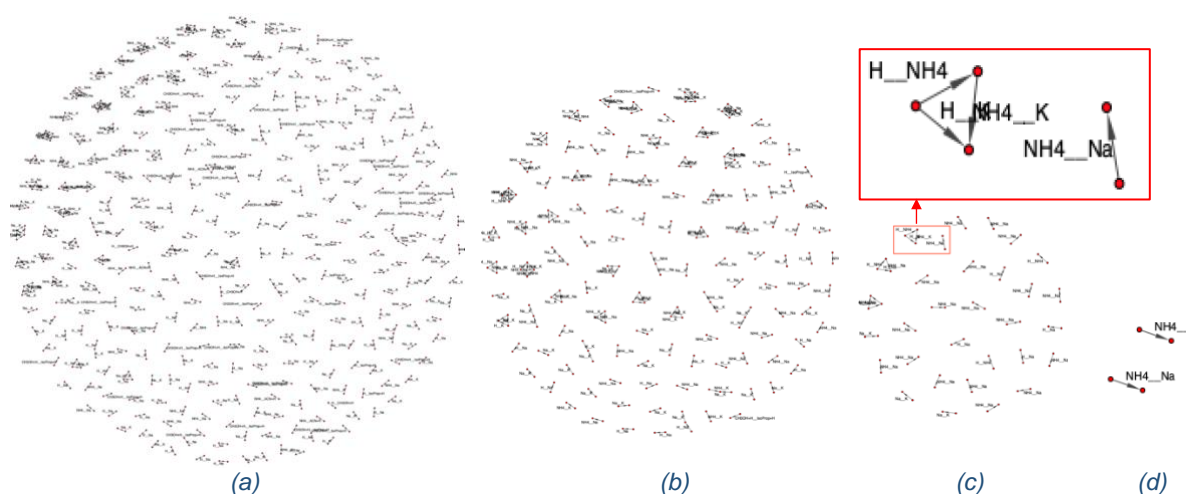


Figure 5. Adduct matching results within isotopologue layers $M+0$ (a), $M+1$ (b), $M+2$ (c), and $M+3$ (d).

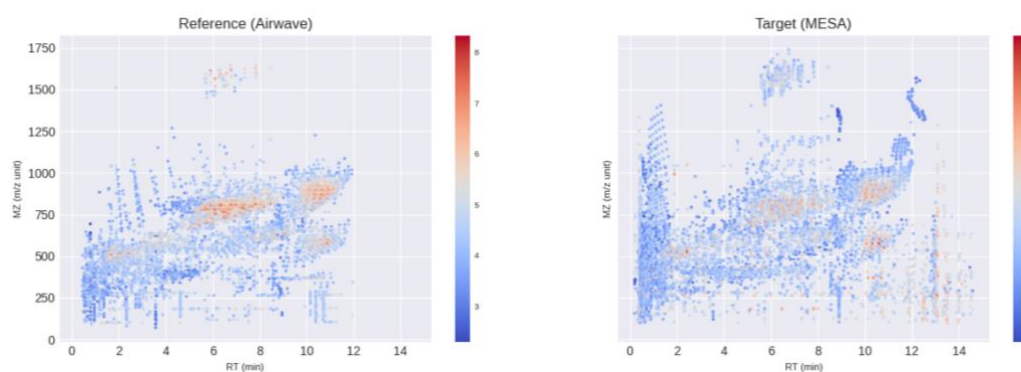


Figure 1. m/z vs RT plots of Airwave reference dataset (left) and MESA target dataset (right) coloured by \log_{10} FI.

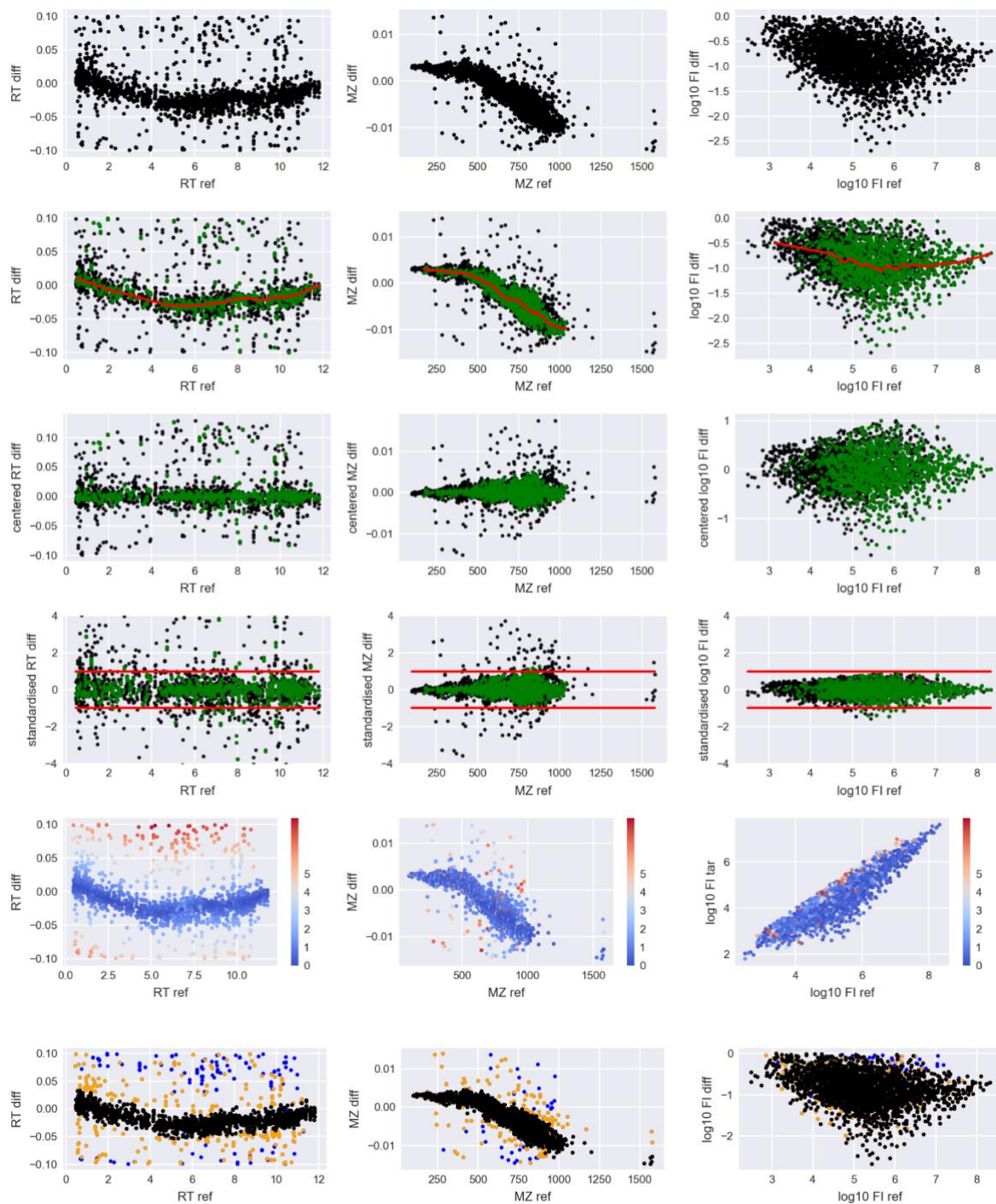


Figure 7. Workflow of matching Airwave and MESA. 1st row: inter-dataset difference for matched feature pairs in RT, m/z and log10 FI. Each black dot is an initial feature match. 2nd row: green dots are matched-subnetwork feature pairs, and the red line is the interpolated LOWESS regression using only green dots. 3rd row: the residuals from each point to the red line. 4th row: standardised residuals by dividing by 5 times MAD respectively. 5th row: the same plot coloured by penalisation score with weight $W = [1, 1, 0]$. 6th row: final matching results with 5 times penalisation score as the threshold where blue points are deleted multiple matches, yellow points are poor matches, and black points are good matches.

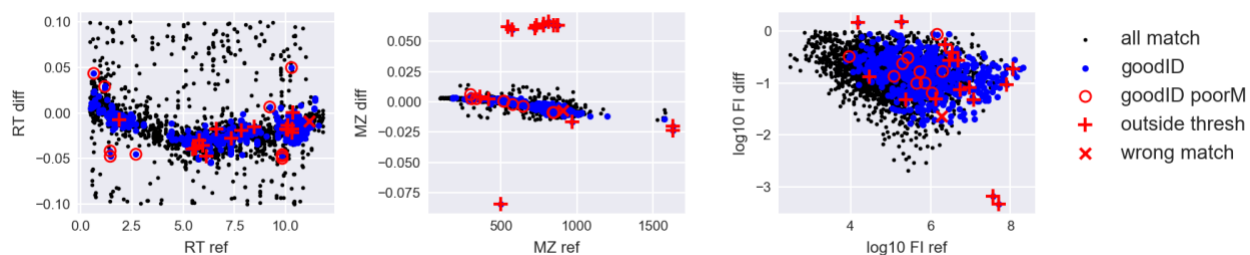


Figure 8. Feature matching results shown in RT, m/z and log10 FI domains. Black dots are all matches within the initial thresholds, blue dots are all the matches with identical annotations, red circles are matches with identical annotations that are incorrectly considered as poor matches, + symbols are matches with identical annotations that are outside of initial thresholds, and x symbol is an incorrect match.

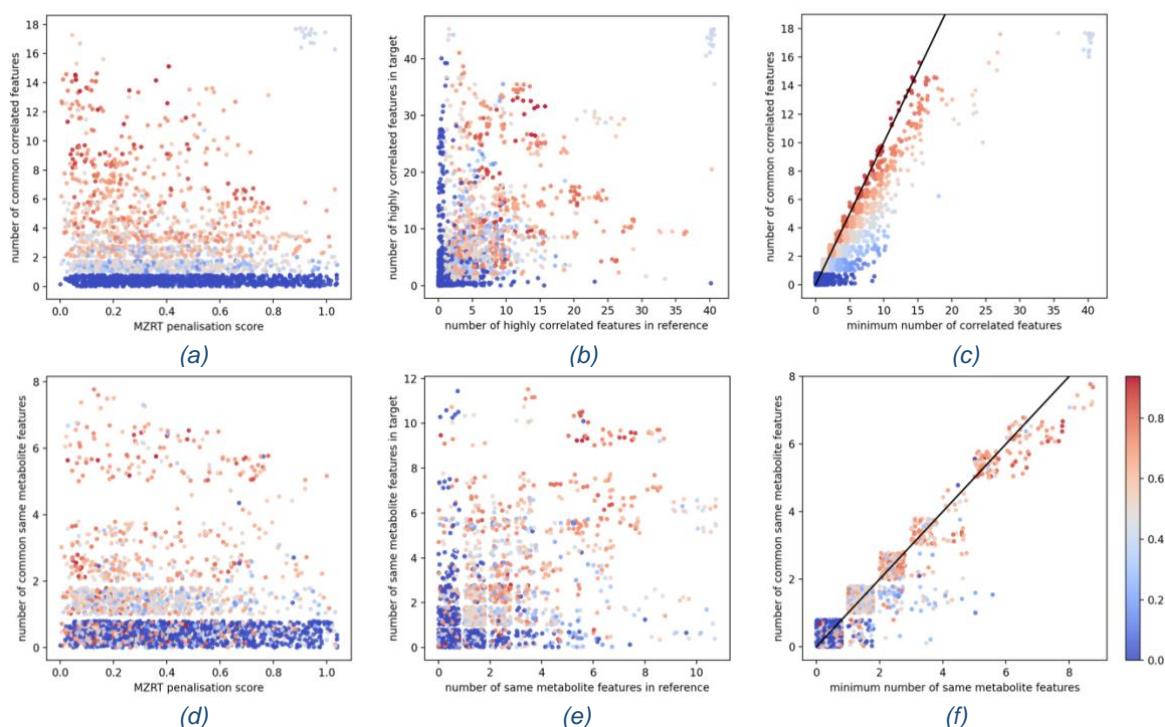


Figure 9. Evaluation plots are coloured by CMR score. The x-axis is jittered in (b), (c), (e) and (f), and the y-axis is jittered in all plots. Top plots consider all features, while the bottom plots consider only the ones in isotopologues-adduct subnetworks.

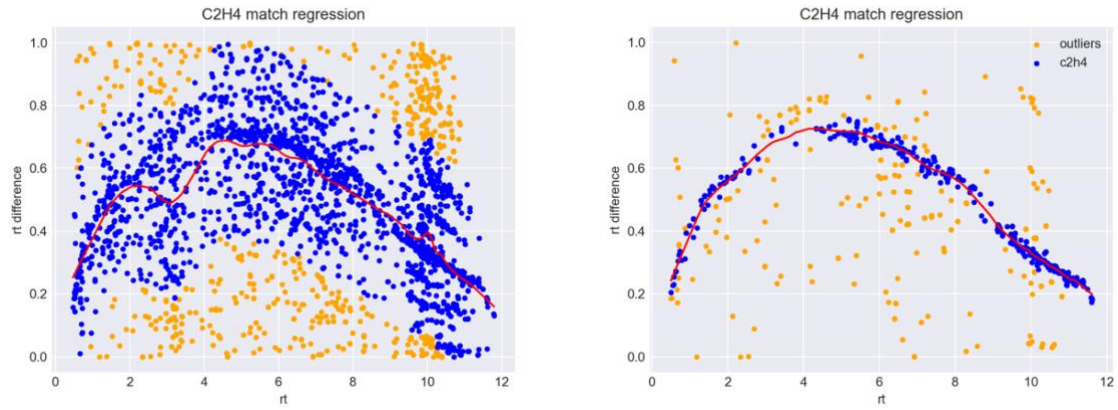


Figure 10. RT difference vs RT for C2H4 matches with the LOWESS regression in the Airwave dataset. The left plot is C2H4 matches using all features in the reference datasets, the right plot is C2H4 matches using only features that are matched across datasets. The red line is the linear interpolator of LOWESS regression. Blue dots are within $3 \times \text{MAD}$ of all points, and yellow dots are outside.