

**Travel time and source variation explain the molecular transformation of dissolved organic matter in an Alpine stream network**

H. Peter<sup>1\*</sup>, G. Singer<sup>2</sup>, A. J. Ulseth<sup>1,3</sup>, T. Dittmar<sup>4,5</sup>, Y. T. Prairie<sup>6</sup>, and T.J. Battin<sup>1</sup>

<sup>1</sup> Stream Biofilm and Ecosystem Research Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

<sup>2</sup> Leibniz-Institute of Freshwater Ecology and Inland Fisheries, 12587 Berlin, Germany

<sup>3</sup> Department of Biological Sciences, Sam Houston State University, Huntsville, Texas, 77341 USA

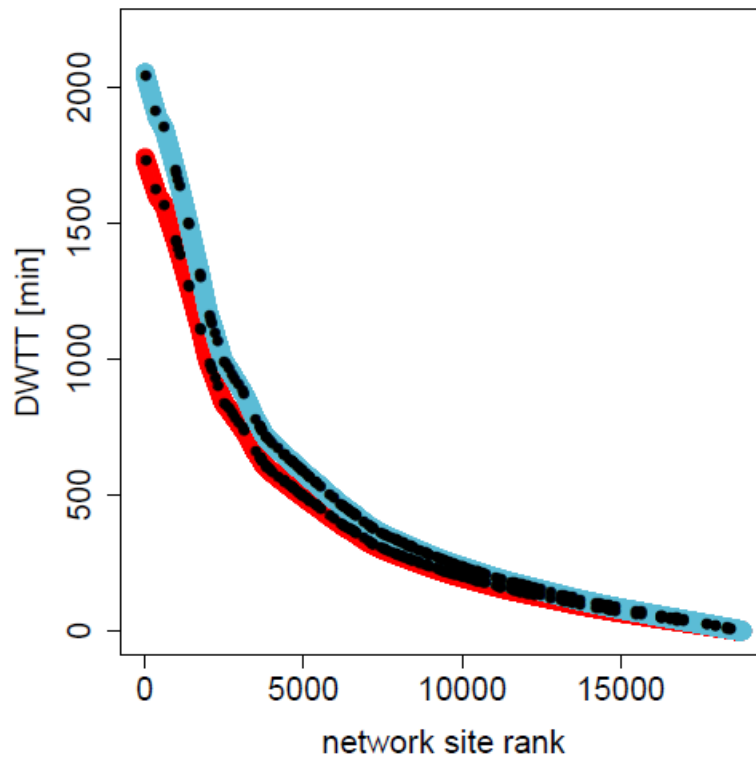
<sup>4</sup> Institute for Chemistry and Biology of Marine Environment (ICBM), ICBM-MPI Bridging, Group for Marine Geochemistry, 26129 Oldenburg, Germany

<sup>5</sup> Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), 26129 Oldenburg, Germany

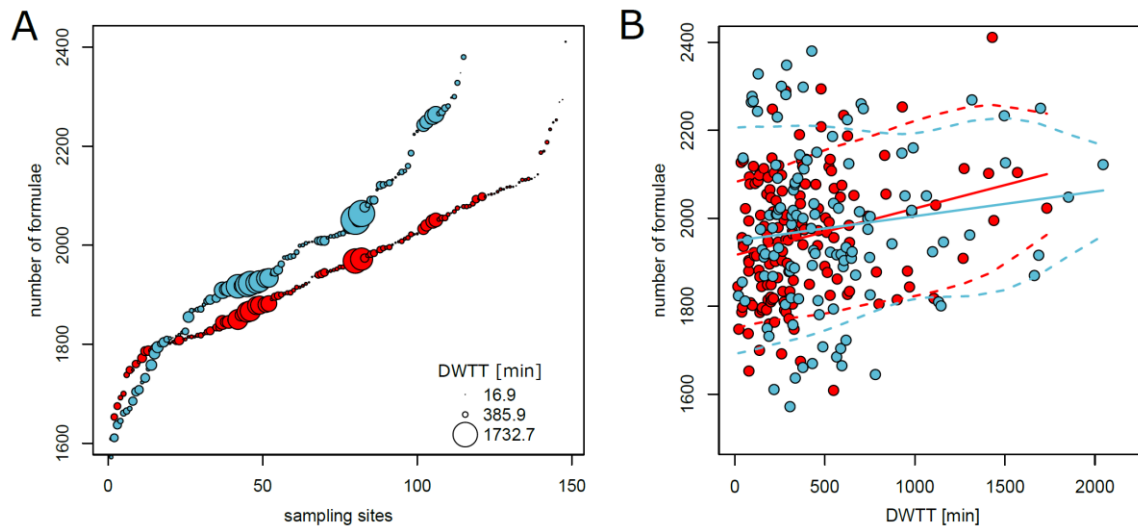
<sup>6</sup> Groupe de Recherche Interuniversitaire en Limnologie, (GRIL), Département des Sciences Biologiques, Université du Québec à Montréal, Montréal, Québec, Canada

**Contents of this file**

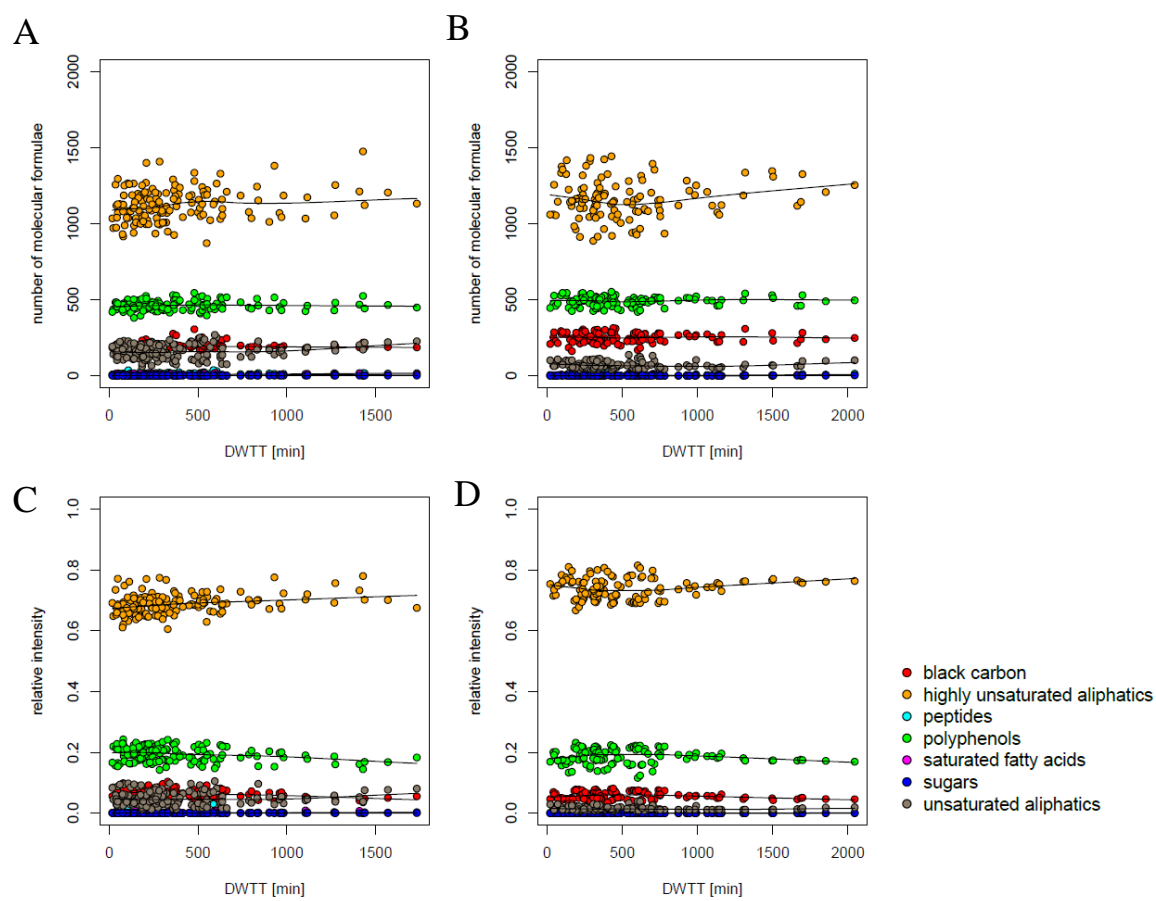
Figures S1 to S10



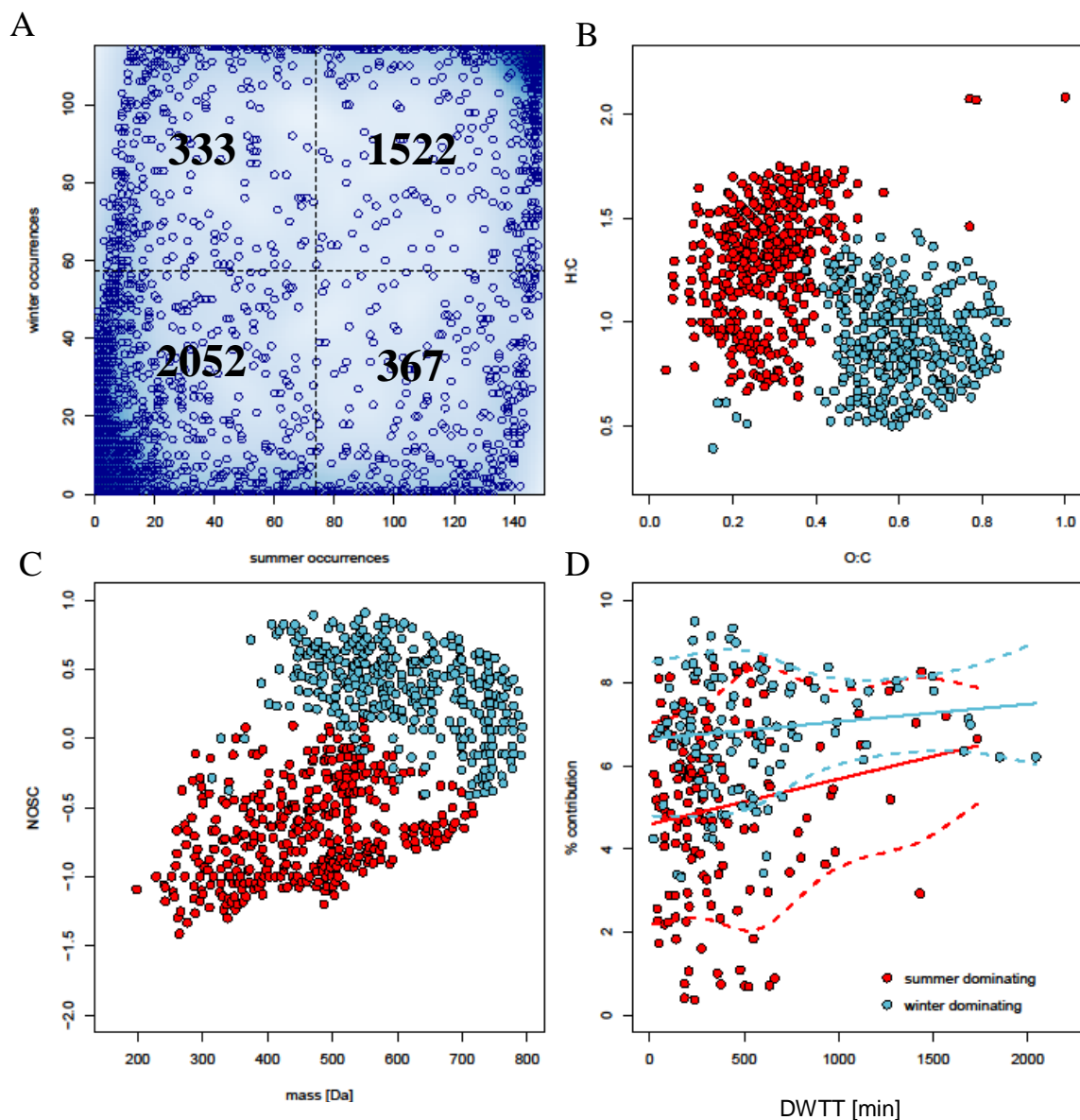
**Figure S1.** Shown is the distribution of DWTT across the entire Ybbs River network in summer (red) and winter (blue). Black dots indicate the distribution of samples along the DWTT gradient. Notably, marked seasonal differences in travel time distribution arose only at large DWTT.



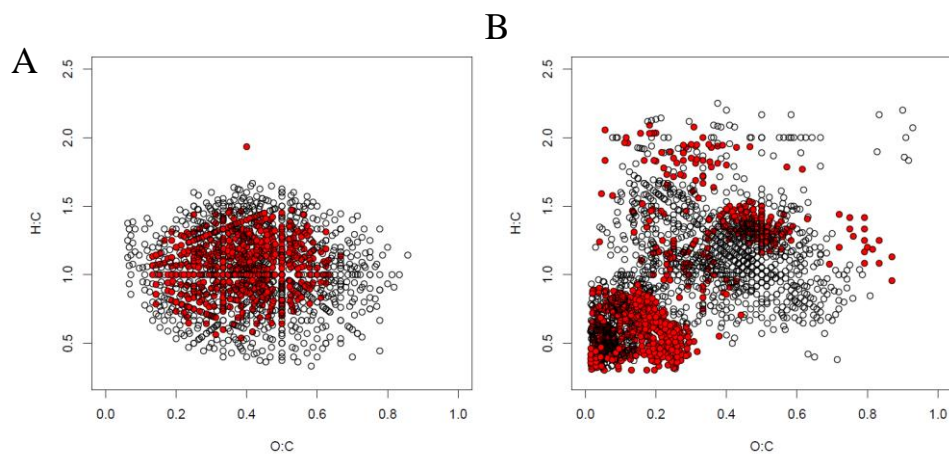
**Figure S2.** No distinct pattern of the number of formulae exists along the DWTT gradient. Shown is the ranked number of formulae for the 115 sites in winter (blue) and the 148 sites in summer (red) (A) with the size of the dots reflecting DWTT. Between 1609 and 2411 (median: 1954) formulae were detected in summer, whereas between 1572 and 2380 (median: 1975) molecular formulae were detected in winter. GAMLSS model fits (solid line for means and dashed lines for 10 and 90% confidence intervals, respectively) indicate a slight increase in the average number of molecules detected (from 1917 to 2100 in summer and from 1949 to 2063 in winter) (B).



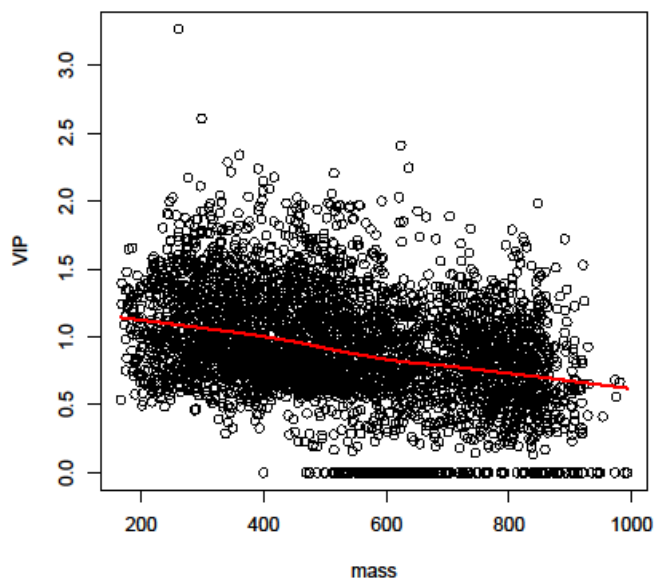
**Figure S3.** Distribution of the total number of molecular formulae (A,B) and summed relative intensities (C,D) among compound categories along the DWTT gradient in summer (A,C) and winter (B,D). Locally weighted scatterplot smoothing lines (LOESS) for each compound category are provided for visual guidance.



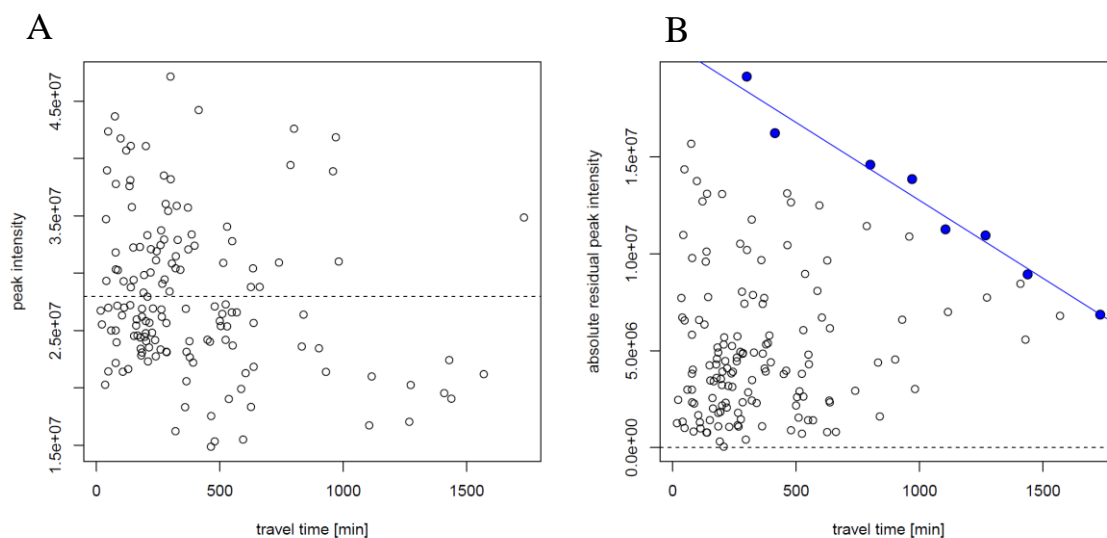
**Figure S4.** Contrasting the number of samples, in which a molecular formula is detected in summer and winter (A), we classified 1522 formula as being common (present in >50% of sites in both seasons), rare (n=2052), summer-specific (n=367) or winter-specific (333). Summer- and winter-specific formulae formed chemically distinct populations in a van Krevelen diagram (B). Summer-specific formulae were dominated by lignin- and protein-like compounds, while winter-specific molecules were dominated by tannin-like compounds (B). Also, summer-specific formulae were smaller and had a lower NOSC (C). Panel D shows the relative contribution of season-specific formulae to total peak intensity at every site in the stream network. The population of tannin-like compounds had a stable average contribution of about 7% throughout the network in winter, the total relative contribution of summer-specific formulae, in contrast, had higher variability and increased from the headwaters to the outlet (GAMLSS,  $p < 0.01$ ).



**Figure S5.** Van Krevelen diagram of common (A) and rare (B) formulae. Red-colored formulae indicate the presence of N in these compounds.

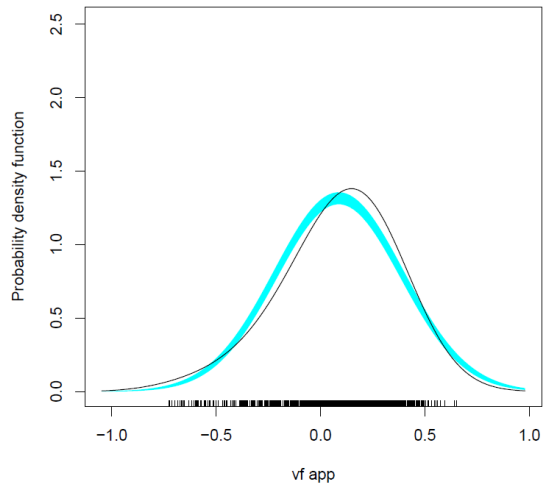


**Figure S6.** Formulae with lower nominal mass had a greater predictive power of DWTT (high Variable Influence on Projection (VIP) scores in the Partial Least Squares (PLS) analysis). The red line shows a Locally Weighted Scatterplot Smoothing (LOESS) fit for visual guidance.

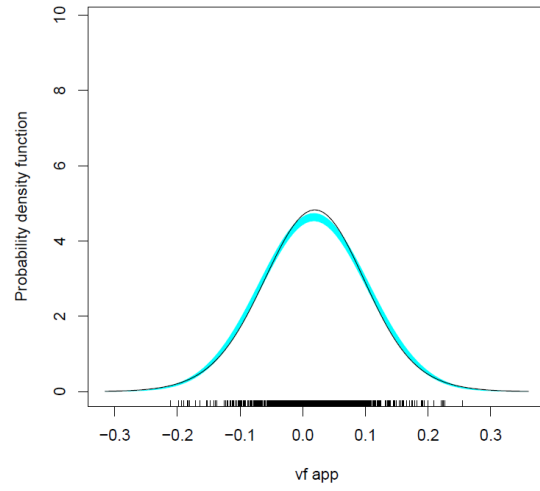


**Figure S7.** Explanation of the calculation of convergence rate. Panel A shows the distribution of peak intensity for a single compound around the catchment average (dotted line). In case of a “funnel-effect” the absolute residual intensity results in a triangular distribution (B) in which at least four pareto-optimal points (blue; i.e. there is no higher value on both axis compared to this point) were identified. The slope of the linear regression (blue line) fitted through these points indicates whether the funnel effect occurs rapidly or slowly.

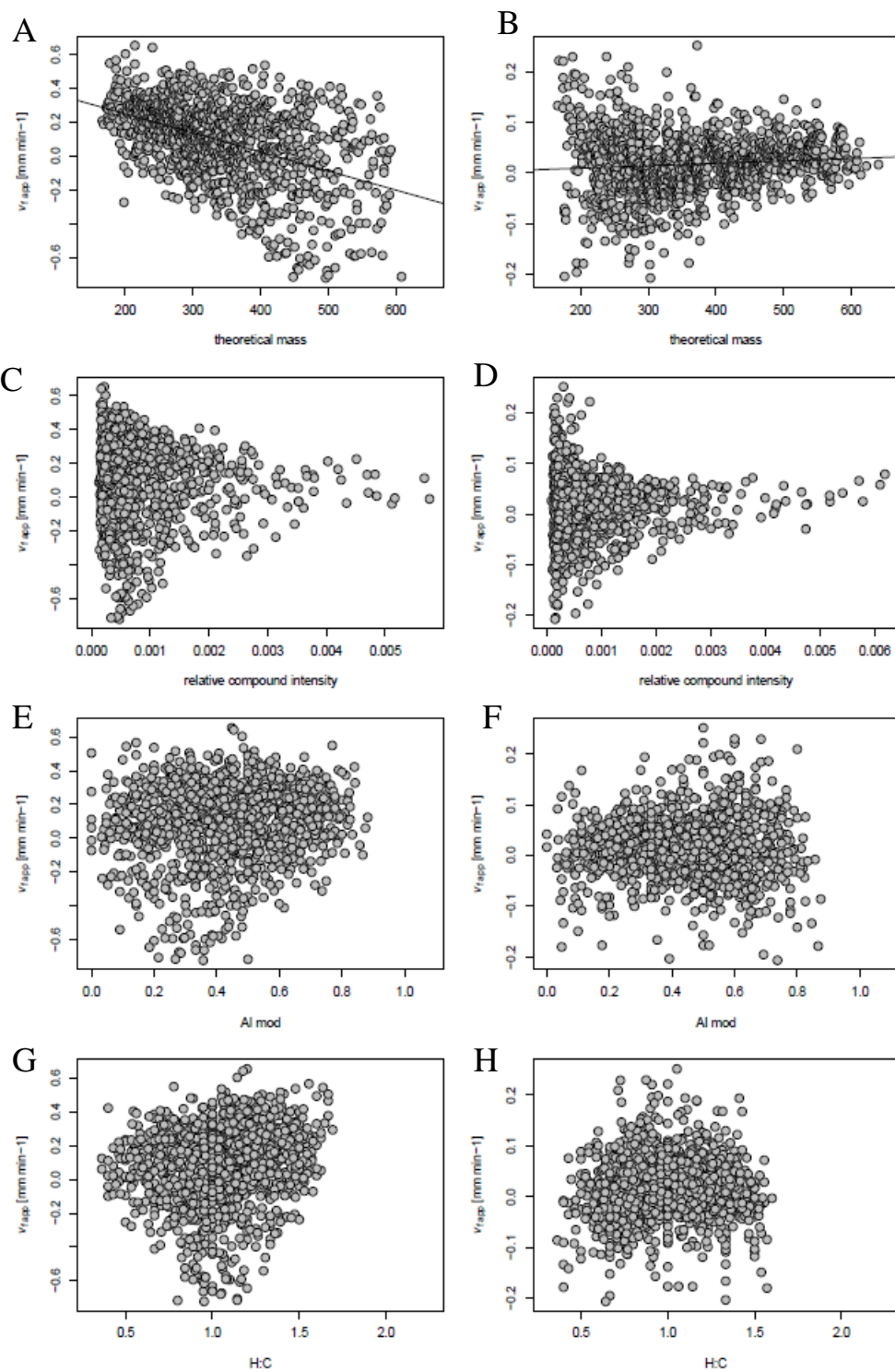
A



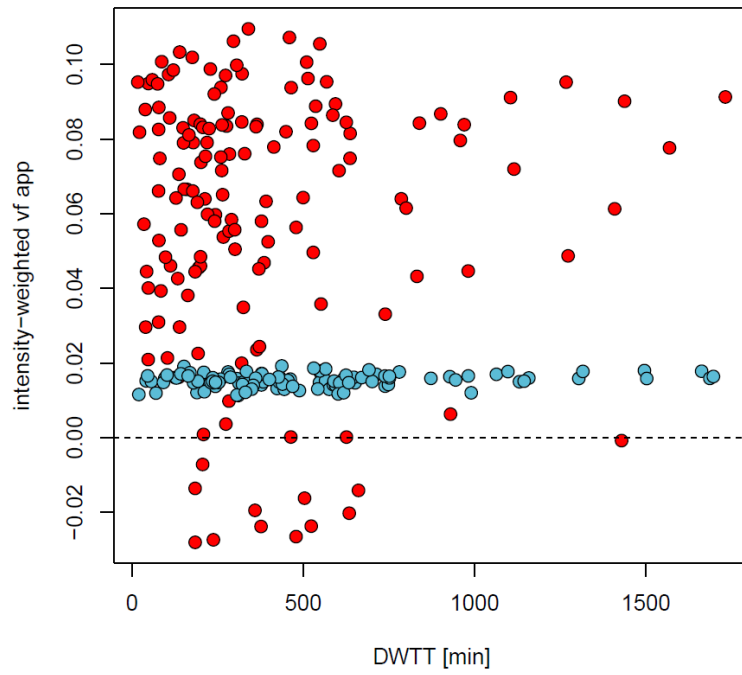
B



**Figure S8.** Distribution of  $v_{f\text{ app}}$  estimates in summer (A) and winter (B). Deviation from the shaded area indicates significant differences from a normal distribution.



**Figure S9.** Relationships between molecular mass (A, B), relative compound intensity (C, D), modified aromaticity index (E, F) and H:C ratio (G, H) and  $v_{fapp}$  for summer and winter respectively.



**Figure S10.** Distribution of intensity-weighted  $v_{f_{app}}$  across the DWTT gradient in summer (red) and winter (blue). Note the points that lie below the zero line (dotted) in summer.