PAPER

Pressurized solvent extraction of secondary char from hydrothermal carbonization: sample preparation and analytical strategy shape GC-MS results

James L. Adair, [ID] †[a] Matteo Pecchi [ID] †[ab] and Jillian L. Goldfarb [ID] *[ab]

Pressurized solvent extraction (PSE) reduces extraction time and solvent use compared to ambient extraction. This study examines the potential to use PSE to extract secondary char, the amorphous tarry phase formed during hydrothermal carbonization (HTC) from hydrochar, and two key sample preparation factors affecting SC characterization. Hydrochar was produced *via* HTC of cellulose, apple pomace, miscanthus, and a representative food waste. Solvent drying and reconstitution had minimal impact on the secondary char composition, except for minor losses of short-chain acids in cellulose-derived samples. Derivatization improved detection of polar compounds, particularly alcohols, though 5-hydroxymethylfurfural became undetectable, likely due to polymerization. Critically, the common practice of using GC-MS peak areas as biofuel concentration proxies may introduce substantial error when comparing relative abundances of compounds and functional groups across samples or studies. Ratios of chromatogram areas should only be used to indicate relative concentration within the same analytical group, not to compare absolute yields across samples or disparate studies. Overall, while sample preparation decisions modestly affect GC-MS analysis of PSE secondary char, data analysis decisions profoundly influence interpretation. We recommend the transition from qualitative peak-area comparisons to quantitative GC-MS methods to ensure accurate characterization of PSE-derived secondary char and thermochemically derived biofuels in general.

1 Introduction

The thermochemical conversion of agro-industrial biomass and other wet carbonaceous wastes produces sustainable biofuels and carbonaceous materials.[1] While wastes' high moisture and oxygen content render combustion unappealing, hydrothermal carbonization (HTC) leverages water content as a reaction medium, solvent, and catalyst.[2,3] At HTC conditions, temperatures of 180–250 °C and subcritical pressures, biomass transforms into a multi-phase hydrochar, comprised of a carbonized primary char (PC) and amorphous secondary char (SC) deposited on the PC surface.[4] While PC often resembles biochars produced from pyrolysis, SC comprises volatile organics that often form an amorphous tar and microspheres on the surface of the PC as a result of the recondensation of organics that hydrolyzed into the aqueous phase and subsequently recondensed onto the solid PC through polymerization and

condensation pathways. In this paper, as in prior work, we define SC as the fraction of hydrochar that can be separated *via* solvent extraction.[4–6] SC causes potential issues with combustion (*e.g.* high reactivity, difficult fluidization)[7] and limits environmental applications (owing to its potential phytotoxicity).[8,9] Yet upon extraction, SC may serve as a precursor to fuels and advanced carbon materials.[10,11] At lab scale, room temperature solvent extraction is often employed to separate the PC and SC.[5] At industrial scale, pressurized solvent extraction (PSE) may be a preferred option to reduce solvent use and extraction time while maximizing yield.[12,13] Thus, PSE is used here to produce a more commercially relevant SC.

PSE, sometimes referred to as pressurized liquid extraction or commercially as accelerated solvent extraction, uses high temperatures (50–200 °C) and pressures to accelerate the desorption, diffusion, and solubilization of organics from solids while preventing evaporation.[14] Due to its high throughput and low solvent usage, PSE is a commonly used alternative to Soxhlet extraction in environmental analysis, particularly for contaminated soils.[15,16] Although PSE requires higher initial capital costs than other extraction methods, its long-term advantages have made it suitable for industrial applications such as organosolv pretreatment of biomass[17] and

[a]Biological and Environmental Engineering, Cornell University, Ithaca, NY 14850, USA. E-mail: goldfarb@cornell.edu; Tel: +1607.255.5789

[b]Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14850, USA

† Co-first authors (equal contribution).

sorbent regeneration.[18] Despite its widely established use for efficient extraction from complex solid matrices in analytical and environmental applications, the application of PSE for separating the solvent-extractable secondary-char fraction from HTC hydrochar appears to be under-reported; we are unable to locate such a study in the literature. Here, we demonstrate the feasibility of PSE for this purpose and evaluate how downstream GC-MS preparation and quantification decisions influence reported SC composition.

SC composition is typically analyzed *via* gas chromatography-mass spectrometry (GC-MS),[5,19,20] though sample preparation methods vary widely. If the extraction solvent is GC-compatible, SC samples can be injected directly. Otherwise, the sample is usually dried under vacuum and reconstituted in a compatible solvent. When compounds of interest are difficult to detect, often due to high polarity or low volatility, samples are derivatized to increase detectability.[21–23] While any of these three sample preparation techniques (direct injection, drying and reconstitution, and derivatization) may be appropriate for specific studies, their impact on GC-MS results in SC analysis is poorly understood.

The literature norm to report GC-MS data for hydrothermal products (SC, biocrude) is to use chromatogram peak area to support mechanistic discussion, often as a proxy for concentration.[24–27] While peak areas (absolute to the sample, relative to a batch of samples analyzed under the same conditions) can be used to compare the relative concentration of the same compound between samples analyzed under identical conditions, area-to-concentration responses differ between compounds. For complex mixtures like SC and biocrude, relying on peak area comparisons or internal standard usage without accounting for differing response rates between compounds can lead to large quantitative errors and poor reproducibility across laboratories. These issues can be mitigated by calibrating the GC-MS for specific compounds, but this is rarely done in the biofuels literature due to time constraints, costs of standards, and the large number of different compounds typically present in biofuels.[28] Therefore, assessing the error that interpreting peak area as a proxy for concentration introduces *versus* relying on calibrated concentrations is paramount to inform researchers about the embedded uncertainty in their assumptions.

We hypothesized that drying SC samples would cause a decrease in small molecule concentration due to increased volatility, and that derivatizing would increase the visibility of alcoholic compounds to our GC-MS system, with large differences in reported composition between the calibrated and area comparison quantification methods. We expect the results of this study to inform SC sample preparation methods both in the laboratory and industrial settings, allowing for better understanding of SC composition. Because secondary char chemistry depends on both feedstock composition and HTC severity, we intentionally include multiple feedstocks to ensure that the extraction and analytical conclusions are robust across chemically distinct SCs. However, we do not attempt to deconvolute feedstock – or reaction condition-driven mechanistic differences; rather, our focus is on how extraction, sample preparation, and GC-MS interpretation shape the reported composition of the GC-amenable SC fraction.

2 Experimental

We prepare extracted SC samples in three ways for GC-MS analysis: direct injection (named NN [non-dried, non-derivatized]), dried and reconstituted with dichloromethane (DCM) (named DN [dried, non-derivatized]), and dried and derivatized with *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) before DCM reconstitution (named DD [dried, derivatized]). We calibrated the GC-MS for both non-derivatized (88 standard compounds) and derivatized compounds (11 compounds) commonly seen in our laboratory's previous SC and other biofuel analyses[29] which were used to quantify concentrations. Through analysis of PSE-extracted SC from four biomass feedstocks, we vary sample preparation techniques and use a quantitative GC-MS analysis to uncover how both effect the reported SC composition.

2.1 Feedstock and hydrothermal conditions

To ensure robustness with SCs of varied compositions, we studied SCs from four feedstocks processed in previous works from our lab and others.[29–31] HTC conditions differed slightly between feedstocks, which does not affect the goals of this work: to test the viability of PSE and compare analysis methods on reported composition.

Microcrystalline cellulose (CLS) was purchased from Alfa Aesar, stored in a plastic container at room temperature and used as received. CLS was hydrothermally carbonized in a 1 L Parr reactor (Moline, USA) at 250 °C for 1 hour with a volumetric loading of 60% and a 15 : 85 biomass/water weight ratio. CLS is chosen as a "model compound" and is found throughout the literature.[31–34]

Apple pomace (AP) was produced by coring and crushing Ruby Frost apples (*Malus domestica*) sourced from the Cornell Orchard; the apple pomace was stored in a freezer at −4 °C and used wet.[29] AP was carbonized in the same 1 L reactor at 50% volumetric loading, again with a 15 wt% biomass to water loading at 250 °C for 2 hours. AP was chosen due to the large size of the apple market, and it being comparatively under-studied with respect to usage as biofuel feedstock.

Miscanthus giganteus (MIS) was harvested in Tompkins County, NY in 2020, dried at 55 °C, ground to <1 mm, and stored in plastic bags at room temperature. Prior to use, it was dried again overnight at 80 °C. MIS was carbonized in a 0.3 L Parr benchtop reactor at 60% loading with a 15 : 85 biomass/water weight ratio at 250 °C for 1 h. MIS was chosen since it is considered a promising lignocellulosic energy crop for HTC.[30,35–37]

The food waste (FW) mixture was produced by blending and mixing the ingredients listed in 38 to recreate the typical US supermarket waste composition. The full procedure and the calculation for yields is described in 5 with the full composition of the mixture given in the SI. The mixture was stored at −18 °C before utilization. FW was carbonized in the same 0.3 L Parr

benchtop reactor at 60% volumetric loading with a 15 wt% biomass loading at 250 °C for 1 h. FW was chosen because of its complex composition and high lipid content that make it a good candidate for SC production.

The common HTC temperature of 250 °C was chosen as it approaches the transition point between HTC and hydrothermal liquefaction and produces a high yield of SC.[5,33] In each case, the resulting slurries were separated via vacuum filtration with Whatman 42 (2.5 μm) cellulose filter paper. Solid HCs were dried in a benchtop oven overnight at 80 °C then stored at room temperature in plastic containers. Results of chemical and physical analysis of the different feedstocks are reported in SI Table S2.

2.2 Pressurized solvent extraction of secondary char

Secondary char extraction was performed on a ThermoFisher Dionex ASE 350 Accelerated Solvent Extractor. Extraction cells (internal volume 10 mL) were loaded with 0.5 g HC onto a cellulose filter (Thermo). HPLC grade dichloromethane (DCM, Thermo) was used at 50% rinse volume and 100 s purge time. Extraction temperature was 100 °C (chosen to avoid significantly exceeding the temperature at which HCs were dried) across two 5 min cycles with 4 min static time in between (based on initial tests with the apparatus). Extraction conditions were kept constant between runs to minimize experimental variables when looking at analysis methods; optimization of PSE conditions falls outside the scope of this work. It is important to note that SC composition and yield is dependent on the solvent choice. Our prior work uncovered that acetone and DCM are able to extract oxygenated and polar compounds more efficiently than methanol, ethanol, isopropanol and hexanes. The alcohols selectively recover hydrocarbon chains but miss some of the aromatics. Given this, DCM was chosen as it allows direct GC-MS injection (more details in Section 2.4).

After extraction, PC-containing cells were opened in an aspirated hood and left uncapped for several hours (to evaporate DCM) and then transferred to a benchtop oven for overnight drying at 80 °C (to ensure dryness). The solvent phase (containing the SC) was transferred immediately after extraction to glass vials and stored at −18 °C.

The PC fraction was obtained gravimetrically by weighing the cell's bottom after the overnight drying; the SC fraction was obtained by difference. The SC concentration was computed using the total SC yield and the total solvent used for the extraction for each sample.

For MIS only, we also performed manual SC extraction as described in previous work[5] using halved quantities, to allow comparison between manual and PSE yields.

2.3 Thermogravimetric analysis of secondary char

Proximate analysis of SC was performed on a TGA 650 thermogravimetric analyzer (TA Instruments) as described in previous works[5] to estimate the fraction of the sample with boiling point up to the temperature reached during GC-MS injection. As-extracted SC was pipetted into 90 μL alumina crucibles and the DCM was allowed to evaporate. This process

was repeated until at least 1 mg of DCM-free SC was in the crucible. The crucibles were loaded into the TGA and the nitrogen flow rate was set to 100 mL min^{-1} (Airgas), at which point the sample was heated to 110 °C at 10 °C min^{-1} and held for 30 minutes to assess moisture content, and then heated at 10 °C min^{-1} to 900 °C to measure volatile matter (VM) fraction. The gas was switched to air after the 900 °C isothermal step and the sample heated at 10 °C min^{-1} to 950 °C to assess fixed carbon fraction FC. Residual mass was attributed to ash. Each sample was analyzed in triplicate; AP was duplicated due to lack of sample. Analysis of TGA data relied on the "tga_data_analysis" package, version 2.2.0.[39] Given its supporting and not primary role in this work, the results of this TGA analysis are provided in SI Fig. S1–S3.

2.4 GC-MS sample preparation

Three different preparation methods for GC-MS analysis are compared: NN (non-dried, non-derivatized), DN (dried, non-derivatized), and DD (dried, derivatized).

For NN samples, the solvent-SC solution was directly transferred from the PSE extract to 2 mL amber vials and diluted with HPLC grade DCM to reach a vial content of 1 mL for analysis, since samples dissolved in DCM could be directly injected into our GC-MS system.

For DN samples, the same amount of solvent-SC solution as for the NN samples was dried in 2 mL amber vials under vacuum at room temperature overnight and then reconstituted with 1 mL DCM before GC-MS analysis. The dilution factor is the same for NN and DN. The DN case is included to investigate the effect of drying on GC-MS analysis, as some extraction solvents may not be GC-MS compatible, and the drying step is necessary for derivatization to prevent solvent interactions but may cause relevant compounds to evaporate, losing them before analysis. There is inherent variability in this process for which our experiment cannot account. For example, if a less volatile solvent is used in place of DCM for extraction, the sample may need to be heated and thus even more high-volatile compounds could be lost to evaporation during the drying step. Likewise, drying could also be done at lower temperatures and under vacuum or flowing gas, which further changes the solvent and solutes' vapor pressures and volatilization rates. As such, this work is offered as a point of comparison for drying versus direct injection, and further research should be done to investigate the impact of solvent and drying system choice.

For DD samples, the solvent-SC solution was dried under the same conditions as for DN and then excess (100 μL) N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) was added before heating the vials at 70 °C for 1 hour. (The lack of underivatized peaks noted suggests complete derivatization was achieved under these conditions). The same vacuum drying step was employed to remove BSTFA before DCM reconstitution to 1 mL and subsequent analysis. Derivatization was investigated as it is reported to improve the quantification of long chain fatty acids (LCFA)[21] usually present in SC.[5]

Samples were run at different dilution levels to maximize compound visibility without risking GC-MS detector damage;

an arbitrary threshold of 1 M TIC (total ion current) for the highest peaks was chosen as the maximum allowed. We determined these GC-MS-safe concentrations prior to the study, resulting in approximately 100–1000 mg L^{-1} in each GC vial. Data on dilution factors are available in Table S4. For example, FW SC required 25-fold dilution in the NN preparation method to avoid exceeding the arbitrary peak height limit of 1 M TIC for LCFAs; smallest peaks were not visible due to the high dilution. Therefore, FW SC was also run at maximum concentration with the GC-MS detector turned off at the retention time noted for the tallest peaks to analyze the compounds that are missed when concentration is lowered for machine safety. This last method is named NNd (non-dried, non-derivatized, detector-off). Each analysis was triplicated on each SC extracted.

2.5 GC-MS calibration

Two GC-MS calibrations were performed using compounds that typically appear in HTC SC and pyrolysis bio-oils. Calibration results (compound names, concentrations, and chromatogram areas obtained) are available in the SI, Table S3.

The non-derivatized calibration was performed by analyzing mixtures with known concentrations of pure standard compounds (purity >98%) dissolved in HPLC grade DCM with up to 6 concentration levels (from 5 mg L^{-1} to 300 mg L^{-1}). A total of 89 chemical standards were included in the set.

The derivatized calibration was performed by derivatizing known amounts of a subset of the abovementioned pure standard compounds (11 compounds, prioritizing identified peaks in derivatized samples) following the same procedure as for DD sample preparation (see Section 2.4).

2.6 GC-MS and data analysis

GC-MS analysis was performed on a Shimadzu GC 2010 Plus Gas Chromatograph-Mass Spectrometer with a Restek Rtx-5Sil MS 30 m fused silica column. Samples were run with a split ratio of 10, the column oven at 40 °C, injection temperature at 280 °C, ion source temperature at 230 °C, and interface temperature at 320 °C. 5 sample rinses were performed pre- and post-run with DCM. The oven heating method started at 40 °C for 6 min, ramped at 5 °C min^{-1} to 300 °C where it held for 5 min, and then cooled to 250 °C at a rate of 20 °C min^{-1} before quenching to room temperature.

Peak integration and identification were performed automatically using the manufacturer's software and the embedded NIST libraries; a minimum slope of 250 mV min^{-1}, minimum width of 2.4 s, and 0 drift were employed. Smoothing was not performed. A minimum similarity of 70% was used for identification.

For each analyzed sample, a qualitative table was produced; all qualitative tables (38 files, including replicates of samples) were analyzed using the open-source Python module "gcms_data_analysis" package, version 1.2.0.[40] The main steps performed by the software are summarized here. First, the properties of every identified compound were retrieved by querying the PubChem[41] database and then each molecule was split (on a mass fraction basis) into its functional groups using

a heuristic fragmentation algorithm.[42] For each sample, the concentration of identified compounds was computed applying the available calibrations (for non-derivatized and derivatized samples) to the measured area. For compounds without a calibration, the calibration of the closest compound in terms of Tanimoto similarity[43] was used; a Tanimoto similarity threshold of 0.4 and a maximum molecular weight difference of 100 atomic units were chosen to minimize error.[40,44] When no calibrated compound satisfied the criteria, no concentration was computed. For each sample, aggregated results that describe the mass fraction (or yield) of each functional group in it were computed using the functional group mass fraction for each compound and its concentration (or yield). Each sample was analyzed in triplicate, and results report the average and the standard deviation of triplicates. Only aggregated plots and summarizing tables are reported in the manuscript for brevity, all data necessary to produce them (table of compounds' properties, calibrations, single sample results, aggregated tables, *etc.*) are provided as raw data files accessible as described in Data availability.

3 Results and discussion

Four feedstocks were subjected to HTC followed by PSE to obtain their secondary char, which was subsequently analyzed *via* GC-MS to assess the impact of sample preparation method and data analysis decisions on the reported SC composition. The chosen feedstocks are used to stress-test the methodology across distinct SC chemistries, not to infer general feedstock-process-SC structure and composition relationships. Readers interested in mechanistic models of secondary-char formation and the influence of reactor conditions are referred to recent HTC/SC studies.[6,45]

3.1 SC yield and GC-MS identification

Cumulative results for each sample (defined as each SC type and its preparation method) are provided in Table 1; results report average and standard deviations of triplicates.

SC yields varied greatly, with FW producing far more secondary char than the other samples. Food waste has a high lipid content;[46] triglycerides hydrolyze into LCFAs such as linoleic acid, the most prominent compound detected in FW SC. SC yields from the lignocellulose-based HCs are lower and feature levulinic acid (CLS, AP) and 5-HMF (MIS). Notably, both of these chemicals stem from dehydration of the cellulosic component, where sugar dehydrates to 5-HMF then further dehydrates to levulinic acid.[11] PSE FW SC yield was roughly equal to that from manual extraction in previous work,[5] while manual MIS SC yield was on average 0.25 g_{SC} g_{HC}^{-1}, with a standard deviation of 0.1, which was lower than PSE MIS SC yield.

In GC-MS analysis of these SCs, the identified area is always above ~40% and mostly above ~80%, suggesting that most compounds in the samples that can be volatilized and ionized are successfully identified (with a similarity matching to NIST library or calibrated compounds of >70%). However, smaller
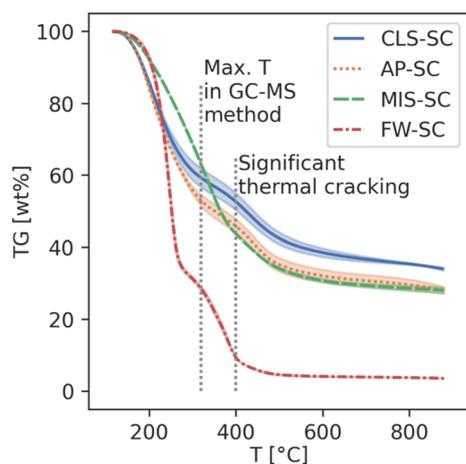
Table 1 SC yields on a HC basis, run once each due to sample limitations. TGA volatilization and GC–MS cumulative results (average and standard deviation of triplicates) for each sample analyzed. Total identified area refers to the chromatogram peaks, concentration at the injection, sample identified fraction and total identified sample yield are obtained applying the corresponding calibration (Table S3) to each qualitative table[a]

Sample	SC yield [g_{SC} g_{HC}^{-1}]	% Mass volatilized at 320 °C	Identified fraction of total area	Identified fraction [g g_{sample}^{-1}]	Identified conc. in SC [mg L^{-1}]	Compound with maximum concentration
CLS-NN	0.14	39.71 ± 3.65	0.90 ± 0.02	0.41 ± 0.01	1138 ± 25	4-Oxopentanoic acid
CLS-DN			0.90 ± 0.01	0.14 ± 0.01	386 ± 27	4-Oxopentanoic acid
CLS-DD			0.42 ± 0.05	0.28 ± 0.04	772 ± 103	(3,4-Dihydroxyphenyl)-phenylmethanone
AP-NN	0.20	47.15 ± 2.41	0.85 ± 0.02	0.16 ± 0.01	655 ± 56	4-Oxopentanoic acid
AP-DN			0.84 ± 0.03	0.17 ± 0.03	692 ± 126	4-Oxopentanoic acid
AP-DD			0.53 ± 0.02	0.12 ± 0.00	474 ± 13	3-Hydroxybenzoic acid; benzene-1,3-diol
MIS-NN	0.32	37.02 ± 0.54	0.69 ± 0.05	0.16 ± 0.02	1016 ± 99	5-(Hydroxymethyl)furan-2-carbaldehyde
MIS-DN			0.69 ± 0.04	0.15 ± 0.01	940 ± 61	5-(Hydroxymethyl)furan-2-carbaldehyde
MIS-DD			0.48 ± 0.04	0.10 ± 0.02	612 ± 157	1-(3-Hydroxyphenyl)ethanone
FW-NN	0.70	73.80 ± 3.92	0.97 ± 0.01	0.71 ± 0.13	9945 ± 1758	(9Z,12Z)-Octadeca-9,12-dienoic acid
FW-DN			0.96 ± 0.01	0.78 ± 0.09	10 860 ± 1194	(9Z,12Z)-Octadeca-9,12-dienoic acid
FW-DD			0.99 ± 0.02	1.00 ± 0.13	14 048 ± 1850	(9Z,12Z)-Octadeca-9,12-dienoic acid
FW[d]-NN			0.85 ± 0.02	0.13 ± 0.01	1773 ± 178	2,3-Dihydroxypropyl (Z)-octadec-9-enoate

[a] d = run with detector off to boost signal of unidentified compounds in FD-NN run.

values were obtained for the sample identified fraction, obtained as the sum of all identified compound fractions in the sample (g g^{-1}), indicating that while compounds could be identified if detected, a significant portion of the sample could not be properly separated by the GC-MS column used or identified by our detector. This result agrees with prior work that performed a careful calibration of compounds in pyrolysis oil analyzed with GC-MS,[44] indicating that there is a range of compounds that GC-MS can detect, but the lightest (*e.g.* C1, C2 compounds) and heaviest compounds (*e.g.* dimers, trimers) may be elusive.

The TGA analysis of SC presented in Fig. 1 is performed under an inert atmosphere to replicate the conditions of the injection in the GC-MS; it suggests that this low identified yield could be due to the presence of compounds with volatilization temperature greater than the maximum temperature that the GC-MS method reaches (320 °C). The full set of TG, DTG, and second DTG curves for HCs, PCs, and SCs are shown in Fig. S1–S3. At 400 °C a bio-oil would start pyrolyzing or polymerizing, which alters compound identification.[47–49] Under inert conditions at $T \leq 320$ °C, only about 37–47 wt% SC for CLS, AP, and MIS and ~74 wt% SC for FW was volatilized. The identified fractions and the volatilized fraction at 320 °C in Table 1 follow the same trends: while the DN SC samples for CLS, AP, MIS showed 14–17 wt% identified fraction, FW showed 78 wt%. The differences between the TGA and GC-MS identified fractions suggests that thermal accessibility, rather than extraction efficiency, governs which SC components appear in chromatographic analysis. Compounds that fail to volatilize below the GC-MS maximum temperature (particularly oligomeric phenolics and partially polymerized sugars) remain undetected despite being present in the SC matrix. This underscores a methodological limitation: GC-MS provides insight into the thermally volatile subset of SC (which of course is itself dependent on column and injection/oven conditions used), not necessarily its full molecular inventory.[20,49] Therefore, relying solely on chromatogram areas without considering the identified fraction can lead to large overestimations of the coverage of GC-MS analysis. When GC-MS chromatograms are assumed to represent the entire sample, all compound quantifications are underestimated by at least the inverse of the quantified fraction, or 2–5× for identification rates of 50 to 20%.

3.2 SC composition

As with yield, the identified SC concentration suggests that the effect of drying and derivatization are dependent on the feedstock composition. Drying did not impact the SC concentration

Fig. 1 TG curves for pyrolysis at 10 °C min⁻¹ of SC from all feedstocks; shading indicates one standard deviation of triplicate runs.

© 2026 The Author(s). Published by the Royal Society of Chemistry

to a significant level; CLS is the only exception where almost 75% of the initial concentration was lost during drying. Derivatization decreased the total identified concentration for AP and MIS, while it increased the identified concentration for CLS and FW.

For CLS, AP, and MIS, the identified compound with the highest concentration for NN samples (and for DN) is different than the case of DD samples, as derivatization detects more polar compounds in the SC that are present at high concentration. For FW, since long chain fatty acids can be effectively observed with and without derivatization, there was no difference between the DN and the DD sample.

A detailed view of the main identified compounds (those with concentration >0.05 in at least one sample) in all samples is provided in Fig. 2. Most compounds are excluded based on the 0.05 threshold; for example, 5-HMF (IUPAC: 5-(hydroxymethyl)furan-2-carbaldehyde) is present in all NN samples except for FW, and is the compound with the highest concentration in MIS (see Table 1), yet its fraction does not exceed 0.03 in any sample so it is not shown in Fig. 2.

For CLS SC, the presence of ~25% of 4-oxopentanoic acid suggests PSE with DCM extracts the embedded 4-oxopentanoic

acid in the hydrochar matrix.[50] CLS SC is dominated by 4-oxopentanoic acid because cellulose-derived intermediates preferentially degrade through dehydration pathways that produce short-chain acids rather than undergo lipid-derived reactions; this finding is supported by several studies that found –OH and –COOH groups on CLS HC surface for HTC temperatures around 250 °C.[34,51] The contrast with FW highlights how HTC reaction mechanisms diverge between carbohydrate-rich and lipid-rich feedstocks: lipids generate thermally stable LCFAs that survive HTC largely intact, whereas carbohydrates fragment and repolymerize into oxygenated species.[5,52] In contrast, the SC obtained from AP and MIS reflected their heterogenous, lignocellulosic composition,[53,54] with no single class of compounds dominant. Compounds found in FW also appear in these samples, albeit in much smaller amounts. These mechanistic distinctions explain why derivatization enhances detectability for lignocellulosic SC (rich in phenolic OH groups) but has minimal effect on FW SC.

The small amounts of LCFA in CLS, AP, and MIS are more likely due to contamination inside the reactor than as a result of feedstock processing. During FW HTC, LCFA are likely adsorbed onto the small coke layer on the reactor walls that is
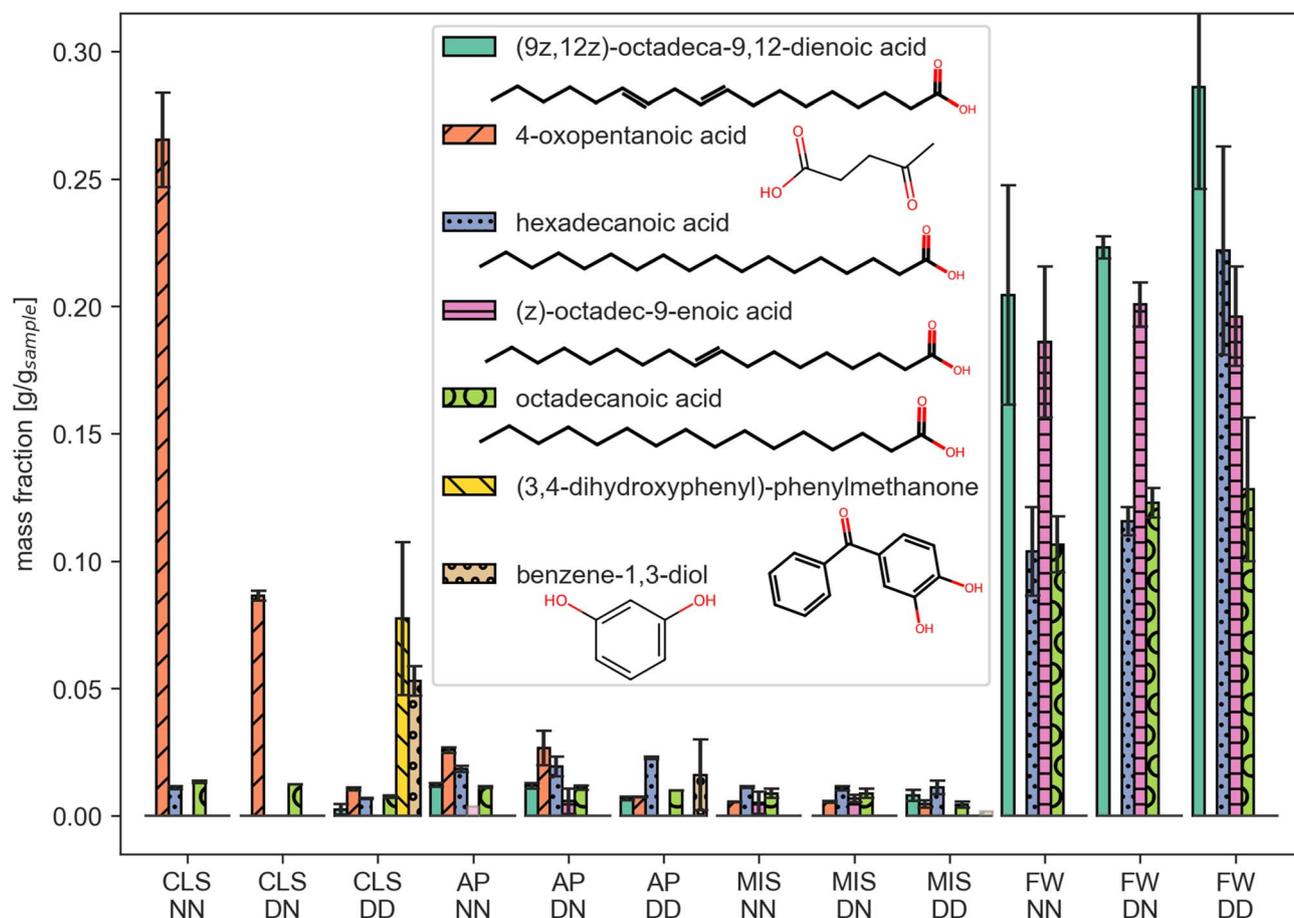

Fig. 2 Sample fraction for major identified compounds in all samples (CLS cellulose, AP apple pomace, MIS miscanthus, and FW food waste extracted SC) for all employed extraction techniques (NN non-dried-non-derivatized, DN dried-non-derivatized, and DD dried-derivatized). The complete list of identified compounds in tabular form is available in the SI. FWd (FW analysis where LCFA are avoided by turning off the GCMS detector) is not included since no compounds exceed 0.02 fraction.

difficult to fully remove with manual cleaning. During HTC runs with the non-FW feedstocks, LCFA can desorb from the coke layer and adsorb onto the forming HC. As this contamination occurs before the experimental step and affects all samples of from a feedstock equally, it does not impact our analyses of sample preparation or GC-MS calibration,‡ though researchers should be mindful of this potential in studies that compare samples across different feedstocks.

Fig. 2 shows that drying had no significant effect on the total sample yield for AP, MIS, and FW (2-tailed t-tests, $p = 0.88, 0.61, 0.50$, respectively). The exception was CLS SC, which significantly decreased in yield post-drying (2-tailed t-test, $p = 0.003$). CLS SC was primarily comprised of 4-oxopentanoic acid, which was mostly lost during the drying due to its low molecular weight and absence of long hydrophobic chain, increasing volatility. Although vapor pressure data at ambient temperatures are scarce, extrapolated constants[55,56] indicate the vapor pressure of levulinic acid is much higher than that of LCFAs (on the order of 10^{-3} mmHg for levulinic, compared to 10^{-20} mmHg for stearic acid, for example). The effect of derivatization (DN $vs.$ DD samples) was more pronounced. For CLS, AP, and MIS; DN and DD samples showed markedly different compositions: levulinic acid (IUPAC: 4-oxopentanoic acid) dominated DN samples, for DD samples were rich in 3-hydroxybenzoic acid, (3,4-dihydroxyphenyl)-phenylmethanone, and benzene-1,3,diol; compounds with two polar hydroxy functionalities that require derivatization for detection under the current GC-MS configuration. For FW, since SC is composed of long chain fatty acids like (9Z,12Z)-octadeca-9,12-dienoic acid (~30 wt%), (Z)-octadec-9-enoic acid (~20 wt%), and hexadecenoic and octadecanoic acid (~10–20 wt%), derivatization only increased the fraction of linoleic and octadecanoic acids detected.

To understand changes in detected SC composition due to differences in sample preparation, the feedstock-dependent SC functional group composition must first be analyzed. The aggregated sample concentrations by functional group, obtained by computing the product of the concentration of each compound by the mass fraction of its functional groups and summing over all the compounds, is shown in Fig. 3. The axis limits were chosen to facilitate a visual comparison among the majority of compounds, the few compounds with higher concentration are shown as "outliers" and their values with standard deviation is shown on the bars (the "±" sign centered on the bar is the value described).

CLS was the only SC where ketones were the most abundant functional group in the NN sample and aromatic C in the DD sample (Fig. 3), due to its high 4-oxopentanoic acid and phenolic concentrations (Fig. 2), respectively. Since in SC from CLS at 220 °C the only identifiable compound was 5-HMF,[5] the larger content of aromatic C (phenols) over aromatic O (furan) at 250 °C suggests furans tend to form more DCM-insoluble (stable) PC compared to phenols and short chain acids (4-oxopentanoic).[50] The aromatic C trend for CLS SC samples is due to

phenolic ketones that were lost to evaporation for the DN sample, while derivatization allowed identification of large amounts of phenolic alcohols in the DD sample. CLS was also the only feedstock where short-chain fatty acids were more abundant than LCFA, as indicated by the ratio between carboxyl concentration and aliphatic C concentration (higher for short-chain fatty acids).

AP and MIS showed similar compositions, and the main differences were ketone (AP > MIS) and ester (MIS > AP) contents; AP showed a higher fraction of 4-oxopentanoic acid and benzene-diols, while MIS yielded a higher 5-HMF fraction. In general, the ratio between aromatic-C and aromatic-O in AP and MIS were similar to CLS, suggesting that furans tend to form more stable structures in the HC compared to phenols.

FW SC is dominated by LCFA, as indicated by the ratio between carboxyl and aliphatic carbon. Analysis of the NNd sample (undiluted sample with detector turned off for known major peaks) shows that only about 10% of functional groups detected were unrelated to the primary LCFA identified in Fig. 2. Most of these were additional LCFA or their derivatives, such esters and amides, though small concentrations of other compounds were detected that do not necessarily correspond to small yields on feedstock basis.

It is important to interpret the GC-MS compositional profiles (Fig. 2 and 3) as representing the volatile and GC-amenable subset of SC rather than the entirety of extractable organic matter. Secondary char contains oligomeric and partially polymerized material that may not volatilize, elute, or ionize under the GC-MS temperature program and detector settings, and therefore can be under-represented even when chromatographic identification rates are high. As discussed earlier, our identified fraction (Table 1) and fraction of each SC sample does not volatilize by the GC-MS inlet temperature (Fig. 1) suggest varying degrees of sample that cannot be identified. Together, these results emphasize that GC-MS provides a detailed view of the volatile portion of SC rather than a complete accounting of all extractable organics. This distinction is particularly important when evaluating the effects of drying and derivatization in Section 3.3, as these preparation steps act solely on (and could therefore selectively bias) the GC-visible portion of the sample. Complementary bulk techniques (e.g., FTIR or elemental analysis) could, in future work, help contextualize overall functional group distributions, HTC reaction mechanisms, and close HTC mass balances beyond the GC-MS window.

3.3 Effect of drying and derivatization

A complete picture of the effect of drying and derivatizing on SC composition is given by Fig. 4 that shows, for each functional group, the absolute and relative aggregated concentration difference between NN and DN samples and between NN and DD samples. Relative concentration differences are computed by dividing the absolute differences by the NN aggregated concentration. When bars equal 1, that indicates that the functional group was only present in the NN samples and not in the DN ones. FWd is not included since only the NN sample was analyzed. In Fig. 4a, all differences are either insignificant or

---

‡ After consulting with technicians at Parr, we developed a protocol for deep cleaning of the reactor by adding oxygen gas (<40 bar) and enough deionized water to cover the bottom of the thermocouple, then running at 200 °C for 1 hour.
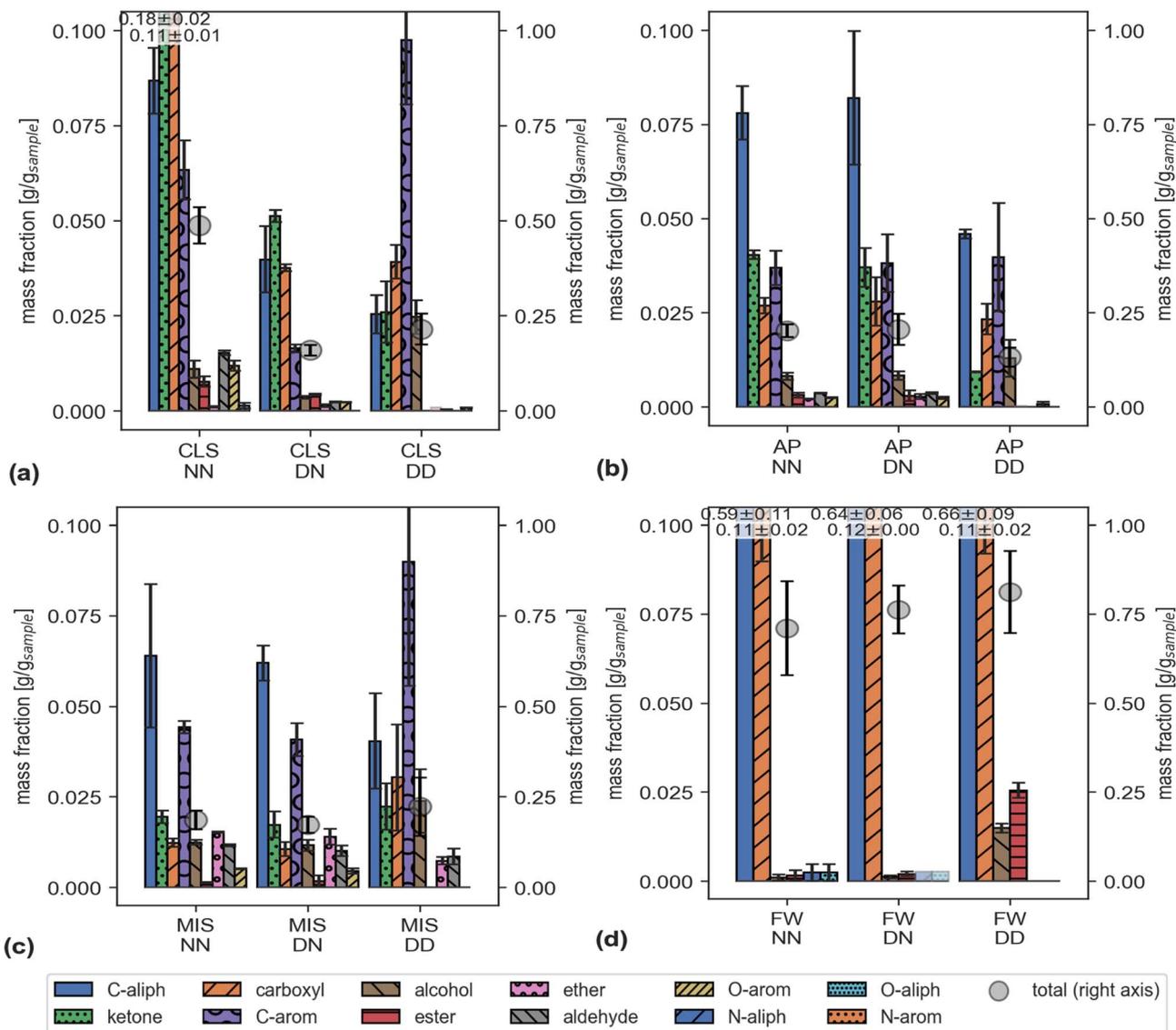
Fig. 3   GC-MS aggregated concentration of functional groups in SC samples ((a) CLS cellulose, (b) AP apple pomace, (c) MIS miscanthus, and (d) FW food waste extracted SC) on a sample fraction basis. Results for all preparation techniques are shown (NN non-dried-non-derivatized, DN dried-non-derivatized, and DD dried-derivatized). Bars represent the average of triplicates, error bars the standard deviation. Only values above 0.1% are shown. Outliers are annotated (average and standard deviation). The total (grey dot) refers to the right axis.

positive, thus indicating a loss upon drying of samples, which is expected. Differences in Fig. 4a and c are mostly insignificant outside of CLS, which also suggests relatively minor modification of the samples upon drying. The apparent increase in carboxyls in MIS DD is attributed to the presence of terephthalic acid (a benzenedicarboxylic acid) in only one of the three replicates.

CLS samples lost half or more of their functional groups during drying, except for esters. Comparing yields between CLS NN and CLS DN shows most of the 4-oxopentanoic acid (the dominant compound in both) is lost in the drying process, in addition to all of the furfural and the majority of the short, aliphatic ketones. The effect of drying on other samples was largely insignificant, except for an unexpected increase in esters in MIS samples post-dry. As there was minimal ester content to

begin with, slight variations in magnitude caused a large relative increase, due to the slight presence of vinyl hexanoate in MIS DN and not MIS NN samples. Derivatization necessitated dilution of DN samples to concentrations low enough for many previously seen compounds to go missing in DD analysis, and underivatized peaks were removed to isolate derivatized compounds. These factors explain why no aromatic oxygen was recorded in the CLS, AP, and MIS DD samples (FW contained none initially). As derivatization replaces alcohols with TMS esters,[23] increased alcohol concentration detected *via* phenolics (CLS & AP) and phenolic acids (AP) derivatizing are associated with increased aromatic carbon. In FW DD, there was an increase in (di)enoates detected, associating the ester increase with the alcohols. LCFAs made up the increase in carboxyls, and by virtue of the long chains, both LCFAs and the enoates
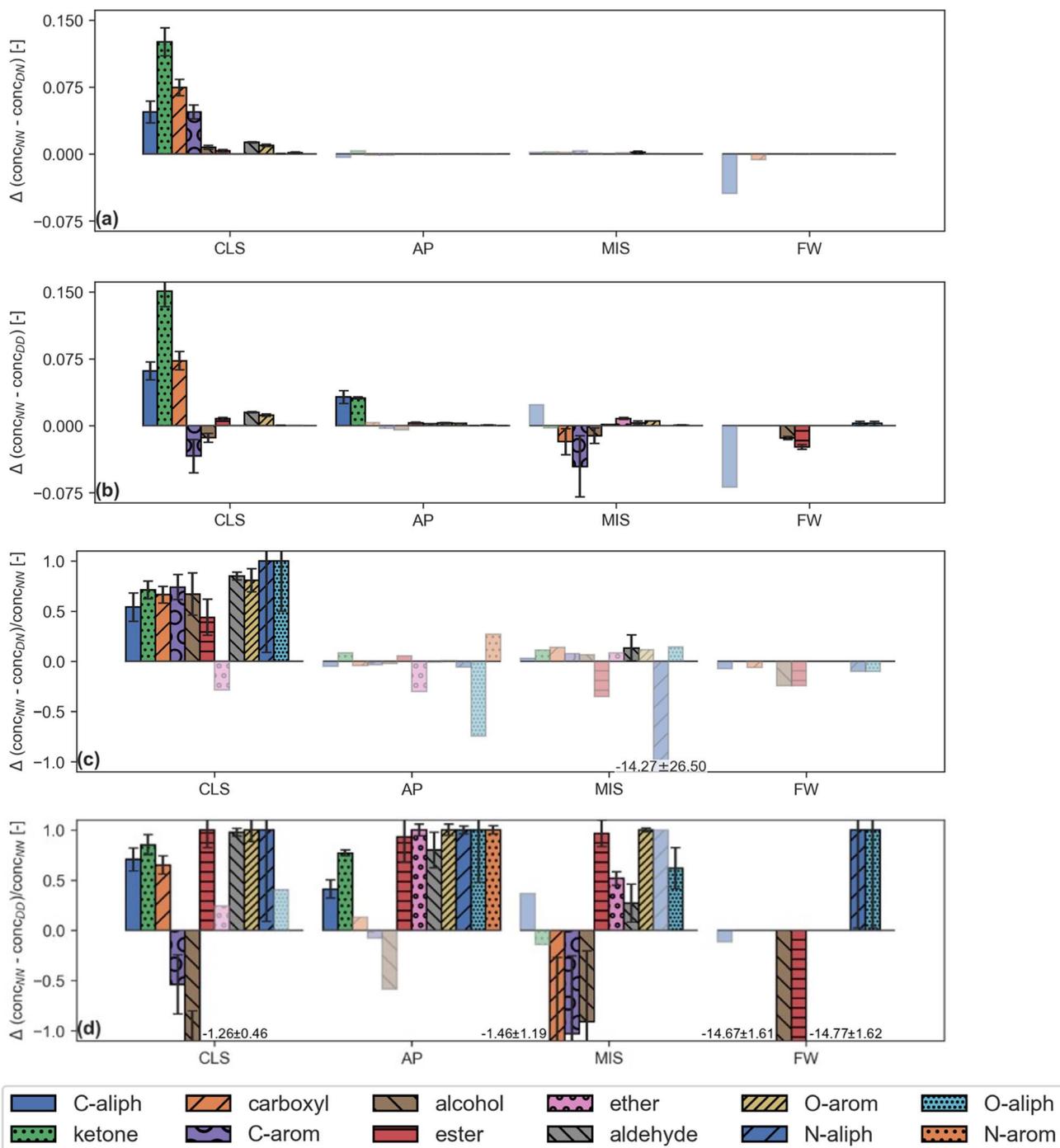
Fig. 4 Sample fraction difference between NN and DN samples: (a) absolute, (c) relative; and NN and DD samples: (b) absolute, (d) relative. Outliers are annotated, transparent bars represent non-significant datapoints (average < standard deviation).

contributed to the rise in aliphatic carbon. In general, derivatization worked as expected and increased detection of alcoholic compounds. One notable exception is the inability of 5-HMF to be detected post-derivatization. While derivatized 5-HMF was not in our calibration, at no point was any peak assigned to a TMS derivative of 5-HMF; we believe we lose it as it polymerizes during the 70 °C heating. As 5-HMF was the most

present compound in MIS NN and DN, its absence caused a net decrease in alcohols in the DD sample.

The same approach is applied to NN and DD samples to investigate the differences in aggregated results with and without derivatization for SC. The results are shown in Fig. 4b. Derivatization (which includes drying) increases or decreases the concentration of different classes of compounds depending on their chemical nature, and in some cases even offsets the

losses due to the drying phase. The alcohol functionality concentration is better detected after derivatization, since derivatized alcohols lose their polarity and are better separated by the non-polar column here adopted. MIS is the only exception, but this seems to be due to the smaller derivatized calibration set compared to the non-derivatized one. For MIS-DD, excluded peaks are more likely than for MIS-NN and DN (due to the Tanimoto similarity threshold of 0.4 and the smaller calibration dataset) and few compounds with relatively high concentration and alcohol functionalities were not included. Indeed, if the Tanimoto similarity is lowered to 0.3, MIS behaves like the other feedstocks (Fig. S4). Looking at aromatic C in CLS, while drying alone decreased its concentration significantly, derivatization increased its content beyond the NN value. A similar behavior is found for AP, which is reasonable since the aromatic carbon fraction in SC for CLS and AP derives from the same glucose decomposition. The aromatic C in MIS has a less obvious trend (deviation is relatively high) and it is hard to draw meaningful conclusions from this sample.

Overall, the impact of drying and derivatization on apparent SC composition reflects the underlying chemistry of each feedstock rather than simple procedural artifacts. The magnitude of these changes depends on the volatility and polarity of the dominant compound classes produced during HTC. For lignocellulosic feedstocks (CLS, AP, MIS), drying disproportionately removed short-chain oxygenates such as 4-oxopentanoic acid and furfural derivatives, which possess comparatively high vapor pressures and weak intermolecular interactions. Their loss indicates that a portion of the "light ends" observed in NN samples are not robust to these sample preparation steps, meaning that reported yields of volatile acids could vary significantly depending on solvent compatibility and drying protocol. In contrast, FW SC, dominated by long-chain fatty acids, showed minimal compositional change upon drying. This outcome highlights that lipid-derived SC components are chemically resilient, and thus far less sensitive to evaporation-driven biases.

Derivatization introduced a second, more mechanistic layer of differentiation among feedstocks. BSTFA modification selectively enhanced detection of alcohol- and phenol-rich compounds by reducing polarity and improving chromatographic separation (on the non-polar column used here). This explains the large increases in phenolic and benzylic species in derivatized CLS, AP, and MIS samples, and supports the interpretation that a substantial portion of lignocellulosic SC is masked in non-derivatized sample preparations due to poor volatility or co-elution. At the same time, derivatization suppressed or eliminated certain compounds, most notably 5-HMF, likely due to polymerization during the heating step. As such, while derivatization reveals hidden functionality, it can also chemically transform reactive intermediates, biasing the apparent chemical space toward more thermally stable or successfully derivatized species. For FW, the limited effect of derivatization further reflects the dominance of already GC-amenable LCFAs, whose detectability is not significantly improved by silylation. These mechanistic insights support the broader methodological conclusion that sample preparation

steps can selectively bias GC-MS datasets, and that interpretations of SC composition must account for these preparation biases, especially when comparing across chemically diverse feedstocks.

3.4 Error associated with peak area comparison

In the biofuels literature, GC-MS chromatogram areas are often used to draw mechanistic conclusions when calibration for absolute concentrations is not performed. However, relative areas are sometimes conflated with yields on feedstock basis, which can mislead readers into overestimation of the level of quantification presented. The typical approach of aggregating compounds by their functional groups by summing their area can introduce substantial errors because compounds within the same group often exhibit different area-to-concentration response factors.[40]

To estimate the potential error from the area-summing approach, the difference between actual concentrations (obtained from applying the calibrations to the obtained areas) and normalized area (normalized based on the total identified peak area for each sample, mimicking the typical approach adopted when a calibration is not available) is shown in Fig. 5a. To compute this difference, the aggregated concentration of each functional group is subtracted by its normalized aggregated area (ignoring the calibration) and divided by its concentration; the result is the fractional error in the measured concentration that using areas would entail. Since normalized areas sum to an unknown total concentration (sometimes assumed to be 1), while sample fractions reflect the actual concentration of compounds in the sample (so the unidentified fraction is accounted for, and total fractions rarely exceed 0.4 outside of certain samples like FW), this method largely overestimates the importance of relative areas when compared to the actual concentrations (see Fig. 5a). FW (the diluted version, not FWd) is more identifiable due to its composition: since the only identified compounds in FW SC are LCFA, they produce very similar concentration to area response in GC-MS analysis due to similar chemistry similarity; in this case, discussing areas rather than concentrations leads to relatively smaller error compared to the other feedstocks. In contrast, phenolics, furans, and short-chain acids in lignocellulosic SC vary widely in ionization cross-section, resulting in orders-of-magnitude discrepancies when areas are treated as concentration surrogates.

To evaluate how well relative areas describe the identified fraction, they must be compared with normalized concentrations, obtained by setting the highest aggregated functional group concentration to 1 and scaling the others accordingly, excluding the unidentified fraction. When employing this approach, it is critical to clarify that areas only represent the identified fraction (which can be far from unity). Fig. 5 illustrates the difference between normalized areas and normalized concentrations. The same considerations as for Fig. 5a apply to FW and FWd. Owing to small initial concentrations in functional groups like heteroatoms, minor absolute errors can result in large relative errors, sometimes over hundred-fold. These
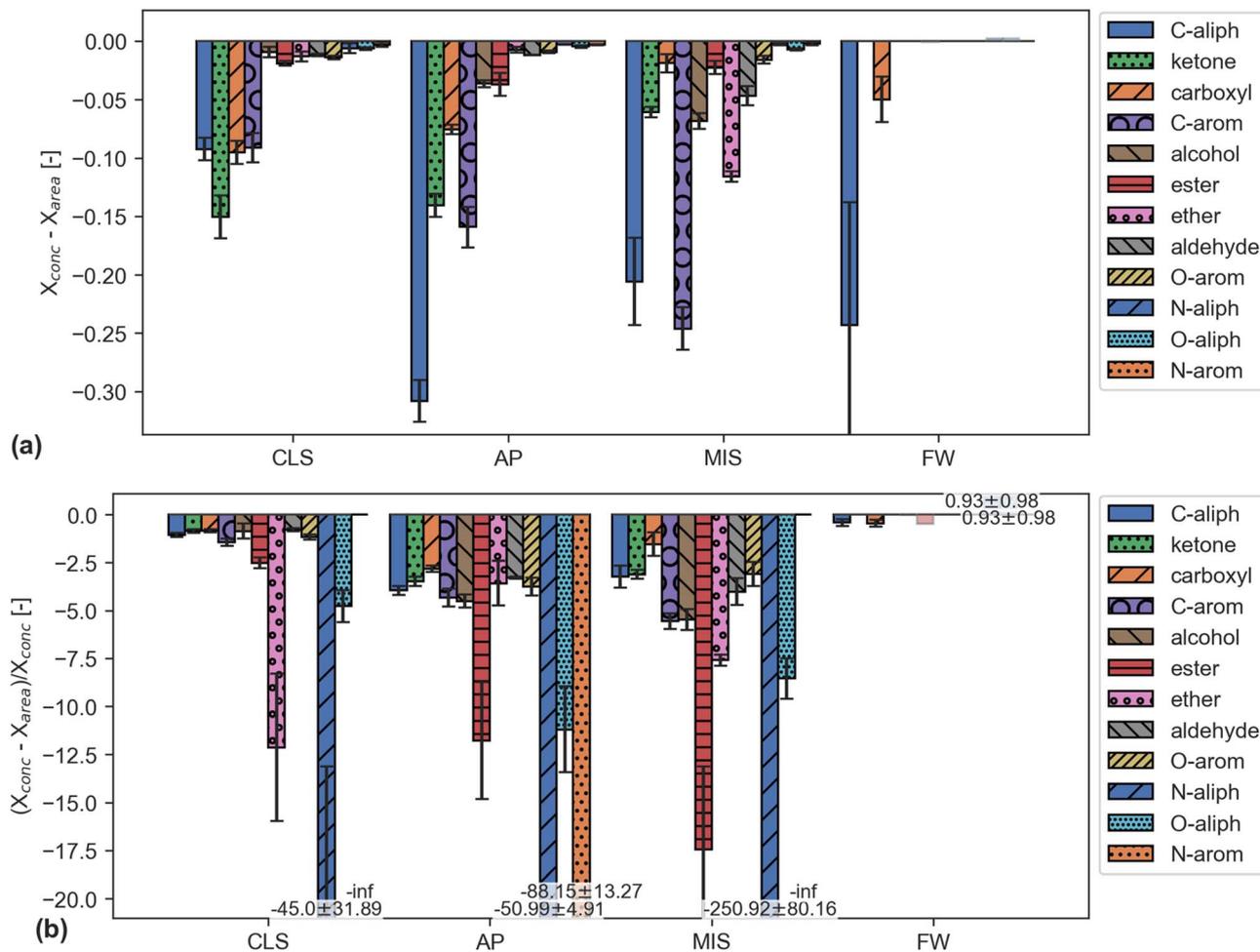
Fig. 5 Error associated with using area instead of concentration. X refers to the reported concentration of a given compound calculated one of two different ways. X_{area} represents the area of a compound peak divided by total area, used when calibration is unavailable. X_{conc} represents the concentration of a compound as calculated using calibration, divided by the total concentration of the sample. The discrepancies between these two methods are shown in (a) absolute and (b) relative to calibrated concentrations. Outliers are annotated. Transparent bars represent non-significant datapoints (average < standard deviation).

limitations do not render the relative area method obsolete; it remains useful for comparing specific compounds across samples analyzed using identical conditions. However, this work suggests that while peak-area summation is modestly reliable within chemically homogeneous families, calibrated or semi-quantitative approaches are essential for cross-compound comparisons.

3.5 Summary of implications

PSE enabled faster, automated extraction of SC from HC while using less solvent than manual methods, achieving equal or greater yields. These results confirm PSE is a viable extraction technique for SC.

FW HC contained more SC than the lignocellulosic HC feedstocks, and its SC was rich in LCFAs due to triglyceride hydrolysis during HTC. FW SC also volatilized more readily at 320 °C in TGA, leading to increased identification by GC-MS. As a homogenous feedstock, CLS SC was (expectedly) also the most homogenous, comprised mostly of levulinic acid formed *via*

dehydration of glucose monomers. CLS SC was the only SC where ketones were the most abundant class, whereas the others were largely aliphatic carbons. AP and MIS SCs were more heterogenous, with major components being levulinic acid and the less-dehydrated 5-HMF. In general, SC composition varied between feedstocks as expected, with the three lignocellulosic feedstocks yielding SCs more similar to each other than the FW SC.

Drying disproportionately affected CLS because its dominant constituents (short-chain oxygenates) possess significantly higher vapor pressures than the LCFAs enriched in FW SC. Levulinic acid, as well as 5-HMF, disappear during derivatization with BSTFA, with the latter likely due to polymerization upon heating. Thus, the drying step selectively strips away volatile species from carbohydrate-derived SC, while leaving lipid-derived SC essentially unchanged. Derivatization improved detection of alcoholic compounds by increasing molecular weight and decreasing polarity (given the nonpolar GC column), but reduced visibility of other compounds due to

chemical reactions, evaporation, or dilution. While drying samples and reconstituting them appears to be viable for most SCs, at least one extraction with a GC-compatible solvent is prudent to avoid missing compounds. BSTFA derivatization performed as expected, and can be used to identify large, polar compounds, with the understanding that most other compounds will not appear.

This work underscores the importance of quantitative GC-MS calibration, due to the widely varying response between concentration and TIC across compounds. While the gold standard is calibrating for every visible compound of interest, this is generally impractical for complex samples like SC. In this context, a semi-quantitative calibration based on chemical similarity reduces error compared to peak area comparison. Peak-area comparison produced concentration estimates that were orders of magnitude off of calibrated concentrations in AP, CLS, and MIS SCs, while FW SC showed reasonable agreement due to similar LCFAs response factors. Thus, a semi-quantitative calibration should serve as a minimum requirement for presenting quantitative results regarding GC-MS output of compound-diverse SC samples.

These methodological insights provide the context needed to interpret the relevance of the extracted fraction itself, motivating a shift toward considering its potential roles in fuel and chemical value chains. While the focus of this work was methodological in nature, we offer several insights on how this work can inform the implementation of HTC for waste upcycling. Solvent-extracted secondary char could serve as a separable, upgradable organic stream rather than a byproduct, enhancing HTC as a waste valorization pathway in which the extracted organics and the cleaned primary char are directed to different value chains. First, when the extract is enriched in long-chain carboxylic acids and lipid-derived oxygenates, it aligns with established upgrading routes to transportation fuels: acids can be esterified to produce biodiesel-range molecules[57] or processed *via* catalytic deoxygenation/hydrotreating toward hydrocarbon fuels,[58] suggesting a fuel-precursor role for SC extracts from lipid-rich substrates. Second, extracts containing higher fractions of platform oxygenates (notably levulinic-acid-type products and furanics such as 5-HMF) are perhaps better chemical feedstocks than as direct fuels, as they are widely used as intermediates to solvents, fuel additives, and polymer building blocks (*e.g.*, levulinic-acid derivative pathways; HMF upgrading to furan-based monomers[59]). Third, phenolic and multifunctional oxygenated aromatics observed in some SC extracts may hold potential as specialty chemicals (*e.g.*, resin precursors or performance additives), where selective recovery and fractionation can be more valuable than bulk fuel upgrading, particularly if the extract is chemically diverse.[60] In summary, these use-cases imply a feedstock- and severity-dependent valorization strategy: lipid-rich feedstocks and conditions that promote lipid hydrolysis favor targeting SC extraction as a liquid fuel-precursor stream, whereas carbohydrate-rich feedstocks motivate strategies that recover and upgrade oxygenated platforms (potentially integrating SC extraction with aqueous-phase upgrading). In all cases, because GC-MS interrogates only the volatile/GC-amenable portion of

SC, the applications discussion should be interpreted as stream-targeting guidance within the GC-visible chemical space, not as a full accounting of total extractable organics.

4 Conclusions

We used pressurized solvent extraction to produce four different secondary chars carbonized at 250 °C from relevant biomass feedstocks: cellulose, apple pomace, miscanthus, and retail-level food waste. The extensive dataset encompasses 38 samples, including triplicates, and hundreds of uniquely identified compounds. Data analysis and statistics were performed using a previously published automated algorithm.

Minimal losses were recorded when samples were pre-dried. One exception was the notable reduction in functional groups within cellulose, excluding esters, where more than half were lost during the drying process. We infer that similar behavior could apply to other short-chain acids. Derivatization with BSTFA enhanced the detectability of most alcohol-containing compounds, thereby increasing the measured quantities of their respective functional groups. However, an exception was noted with 5-HMF, which became undetectable after undergoing derivatization, possibly due to its polymerization during the reaction step at 70 °C.

Using GC-MS peak areas as a proxy for concentrations can result in substantial errors, sometimes over a hundred-fold, in assessing the relative importance of compounds and functional groups, which could significantly alter the conclusions of such analyses. Since calibration is not always practical, we emphasize the need for transparency in discussing the uncertainty of area results, avoiding confusion with concentrations/yields.

We determined that PSE is an effective method for separating out SC from HC, and can be used to reduce solvent volumes required. Drying and reconstituting SC has minimal effect, while derivatization can cause some compounds to go undetected. We recommend using absolute areas for comparing only relative concentrations of the same compounds across samples rather than different compounds within the same sample, due to the large potential for error across compounds. Taken together, the findings of this work can be used to increase the efficiency and accuracy of SC production and characterization, allowing it to meet the challenges associated with upscaling.

Author contributions

JLA: conceptualization, methodology, investigation, writing – original draft, data curation, validation, formal analysis. MP: conceptualization, methodology, investigation, writing – original draft, software, data curation, visualization. JLG: conceptualization, resources, writing – review & editing, supervision, funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Abbreviations

5-HMF	5-(Hydroxymethyl)furan-2-carbaldehyde
AP	Apple pomace
CLS	Cellulose
DCM	Dichloromethane
DD	Dried, derivatized
DN	Dried, non-derivatized
FW	Food waste
GC-MS	Gas-chromatography mass-spectrometry
HC	Hydrochar
HTC	Hydrothermal carbonization
LCFA	Long-chain fatty acids
MIS	Miscanthus
NN	Non-dried, non-derivatized
PC	Primary char
PSE	Pressurized solvent extraction
SC	Secondary char

Data availability

The data used to make all figures and tables in this manuscript is available at Jillian L. Goldfarb's Dataverse, archived at https://doi.org/10.7910/DVN/6E2ORZ.

Supplementary information (SI): additional feedstock and char data, thermogravimetric analysis, and details for GC-MS analysis. See DOI: https://doi.org/10.1039/d5ra09293k.

Acknowledgements

References

1 S. Y. Lee, R. Sankaran, K. W. Chew, C. H. Tan, R. Krishnamoorthy, D.-T. Chu, *et al.*, Waste to bioenergy: a review on the recent conversion technologies, *BMC Energy*, 2019, **1**, 4, DOI: 10.1186/s42500-019-0004-7.

2 A. A. Peterson, F. Vogel, R. P. Lachance, M. Fröling, J. Michael, J. Antal and J. W. Tester, Thermochemical biofuel production in hydrothermal media: A review of sub- and supercritical water technologies, *Energy Environ. Sci.*, 2008, **1**, 32–65, DOI: 10.1039/B810100K.

3 S. S. Toor, L. Rosendahl and A. Rudolf, Hydrothermal liquefaction of biomass: A review of subcritical water technologies, *Energy*, 2011, **36**, 2328–2342, DOI: 10.1016/j.energy.2011.03.013.

4 M. Lucian, M. Volpe, L. Gao, G. Piro, J. L. Goldfarb and L. Fiori, Impact of hydrothermal carbonization conditions on the formation of hydrochars and secondary chars from the organic fraction of municipal solid waste, *Fuel*, 2018, **233**, 257–268, DOI: 10.1016/j.fuel.2018.06.060.

5 M. Pecchi, M. Baratieri, A. R. Maag and J. L. Goldfarb, Uncovering the transition between hydrothermal carbonization and liquefaction via secondary char extraction: A case study using food waste, *Waste Manage.*, 2023, **168**, 281–289, DOI: 10.1016/j.wasman.2023.06.009.

6 M. Pecchi, M. Baratieri, J. L. Goldfarb and A. R. Maag, Effect of solvent and feedstock selection on primary and secondary chars produced via hydrothermal carbonization of food wastes, *Bioresour. Technol.*, 2022, **348**, 126799, DOI: 10.1016/j.biortech.2022.126799.

7 P. Motiei, M. Pecchi, J. L. Adair, J. L. Goldfarb and J. O'Connor, Role of volatile secondary char on the combustion behavior of cellulose-based hydrochars, *Combust. Flame*, 2024, **269**, 113703, DOI: 10.1016/j.combustflame.2024.113703.

8 D. Bona, D. Scrinzi, G. Tonon, M. Ventura, T. Nardin, F. Zottele, *et al.*, Hydrochar and hydrochar co-compost from OFMSW digestate for soil application: 2. agro-environmental properties, *J. Environ. Manage.*, 2022, **312**, 114894, DOI: 10.1016/j.jenvman.2022.114894.

9 D. Kalderis, G. Papameletiou and B. Kayan, Assessment of Orange Peel Hydrochar as a Soil Amendment: Impact on Clay Soil Physical Properties and Potential Phytotoxicity, *Waste Biomass Valorization*, 2019, **10**, 3471–3484, DOI: 10.1007/s12649-018-0364-0.

10 Y. Gong, L. Xie, C. Chen, J. Liu, M. Antonietti and Y. Wang, Bottom-up hydrothermal carbonization for the precise engineering of carbon materials, *Prog. Mater. Sci.*, 2023, **132**, 101048, DOI: 10.1016/j.pmatsci.2022.101048.

11 P. Modugno and M.-M. Titirici, Influence of Reaction Conditions on Hydrothermal Carbonization of Fructose, *ChemSusChem*, 2021, **14**, 5271–5282, DOI: 10.1002/cssc.202101348.

12 W. Chen, Y. Liu, L. Song, M. Sommerfeld and Q. Hu, Automated accelerated solvent extraction method for total lipid analysis of microalgae, *Algal Res.*, 2020, **51**, 102080, DOI: 10.1016/j.algal.2020.102080.

13 Y. Tang, Y. Zhang, J. N. Rosenberg, N. Sharif, M. J. Betenbaugh and F. Wang, Efficient lipid extraction and quantification of fatty acids from algal biomass using accelerated solvent extraction (ASE), *RSC Adv.*, 2016, **6**, 29127–29134, DOI: 10.1039/C5RA23519G.

14 B. E. Richter, B. A. Jones, J. L. Ezzell, N. L. Porter, N. Avdalovic and C. Pohl, Accelerated Solvent Extraction: A Technique for Sample Preparation, *Anal. Chem.*, 1996, **68**, 1033–1039, DOI: 10.1021/ac9508199.

15 N. Saim, J. R. Dean, Md P. Abdullah and Z. Zakaria, Extraction of polycyclic aromatic hydrocarbons from contaminated soil using Soxhlet extraction, pressurised and atmospheric microwave-assisted extraction, supercritical fluid extraction and accelerated solvent extraction, *J. Chromatogr., A*, 1997, **791**, 361–366, DOI: 10.1016/S0021-9673(97)00768-1.

16 M. Klees, C. Bogatzki and E. Hiester, Selective pressurized liquid extraction for the analysis of polychlorinated

biphenyls, polychlorinated dibenzo-*p*-dioxins and dibenzofurans in soil, *J. Chromatogr., A*, 2016, **1468**, 10–16, DOI: 10.1016/j.chroma.2016.09.029.

17 X. Zhao, K. Cheng and D. Liu, Organosolv pretreatment of lignocellulosic biomass for enzymatic hydrolysis, *Appl. Microbiol. Biotechnol.*, 2009, **82**, 815–827, DOI: 10.1007/s00253-009-1883-1.

18 O. N. Fedyaeva, A. V. Shishkin and A. A. Vostrikov, Dynamic Adsorption and Desorption of Phenol on Activated Carbon in Media of Sub- and Supercritical Water, *Russ. J. Phys. Chem. B*, 2022, **16**, 1285–1293, DOI: 10.1134/S1990793122070053.

19 C. Klüpfel, B. Herklotz and P. Biller, Influence of processing conditions and biochemical composition on the hydrothermal liquefaction of digested urban and agricultural wastes, *Fuel*, 2023, **352**, 129016, DOI: 10.1016/j.fuel.2023.129016.

20 H. Sudibyo, M. Pecchi and J. W. Tester, Experimental-based mechanistic study and optimization of hydrothermal liquefaction of anaerobic digestates, *Sustain. Energy Fuels*, 2022, **6**, 2314–2329, DOI: 10.1039/D2SE00206J.

21 E. M. Antolín, D. M. Delange and V. G. Canavaciolo, Evaluation of five methods for derivatization and GC determination of a mixture of very long chain fatty acids (C24:0–C36:0), *J. Pharm. Biomed. Anal.*, 2008, **46**, 194–199, DOI: 10.1016/j.jpba.2007.09.015.

22 X. Gao, E. Pujos-Guillot and J.-L. Sébédio, Development of a Quantitative Metabolomic Approach to Study Clinical Human Fecal Water Metabolome Based on Trimethylsilylation Derivatization and GC/MS Analysis, *Anal. Chem.*, 2010, **82**, 6447–6456, DOI: 10.1021/ac1006552.

23 P. Gries, A. S. Rathore, X. Lu, J. Chiou, Y. B. Huynh, A. Lodi, *et al.*, Automated Trimethyl Sulfonium Hydroxide Derivatization Method for High-Throughput Fatty Acid Profiling by Gas Chromatography–Mass Spectrometry, *Molecules*, 2021, **26**, 6246, DOI: 10.3390/molecules26206246.

24 V. Benavente, C. Pérez and S. Jansson, Co-hydrothermal carbonization of microalgae and digested sewage sludge: Assessing the impact of mixing ratios on the composition of primary and secondary char, *Waste Manage.*, 2024, **174**, 429–438, DOI: 10.1016/j.wasman.2023.11.039.

25 M. S. Jensen, R. B. Madsen, D. Salionov and M. Glasius, Predicting the chemical composition of biocrude from hydrothermal liquefaction of biomasses using a multivariate statistical approach, *Sustain. Energy Fuels*, 2024, **8**, 3877–3891, DOI: 10.1039/D4SE00860J.

26 S. I. Hussain Shah, T. H. Seehar, M. Raashid, R. Nawaz, Z. Masood, S. Mukhtar, *et al.*, Biocrude from hydrothermal liquefaction of indigenous municipal solid waste for green energy generation and contribution towards circular economy: A case study of urban Pakistan, *Heliyon*, 2024, **10**, e36758, DOI: 10.1016/j.heliyon.2024.e36758.

27 J. Yang, C. Hong, Y. Xing, Z. Zheng, Z. Li, X. Zhao, *et al.*, Study on the hydrothermal liquefaction of antibiotic residues with molecular sieve catalysts in the ethanol–water system: focus on product distribution and

characterization, *RSC Adv.*, 2021, **11**, 26763–26772, DOI: 10.1039/D1RA03860E.

28 J.-W. Ahn, S. K. Pandey and K.-H. Kim, Comparison of GC-MS Calibration Properties of Volatile Organic Compounds and Relative Quantification Without Calibration Standards, *J. Chromatogr. Sci.*, 2011, **49**, 19–28, DOI: 10.1093/chrsci/49.1.19.

29 J. L. Adair, M. Karod and J. L. Goldfarb, Addition of in situ clay catalysts at different process points in a cascaded hydrothermal carbonization-pyrolysis process for agro-industrial waste valorization, *Bioresour. Technol.*, 2023, 128649, DOI: 10.1016/j.biortech.2023.128649.

30 M. Mihajlović, J. Petrović, S. Maletić, M. K. Isakovski, M. Stojanović, Z. Lopičić, *et al.*, Hydrothermal carbonization of Miscanthus × giganteus: Structural and fuel properties of hydrochars and organic profile with the ecotoxicological assessment of the liquid phase, *Energy Convers. Manage.*, 2018, **159**, 254–263, DOI: 10.1016/j.enconman.2018.01.003.

31 M. Diakité, A. Paul, C. Jäger, J. Pielert and J. Mumme, Chemical and morphological changes in hydrochars derived from microcrystalline cellulose and investigated by chromatographic, spectroscopic and adsorption techniques, *Bioresour. Technol.*, 2013, **150**, 98–105, DOI: 10.1016/j.biortech.2013.09.129.

32 N. Paksung, J. Pfersich, P. J. Arauzo, D. Jung and A. Kruse, Structural Effects of Cellulose on Hydrolysis and Carbonization Behavior during Hydrothermal Treatment, *ACS Omega*, 2020, **5**, 12210–12223, DOI: 10.1021/acsomega.0c00737.

33 M. Pecchi, A. Cascioli, A. R. Maag, J. L. Goldfarb and M. Baratieri, Uncovering the transition between hydrothermal carbonization and liquefaction using differential scanning calorimetry, *Fuel Process. Technol.*, 2022, **235**, 107349, DOI: 10.1016/j.fuproc.2022.107349.

34 M. Sevilla and A. B. Fuertes, The production of carbon materials by hydrothermal carbonization of cellulose, *Carbon*, 2009, **47**, 2281–2289, DOI: 10.1016/j.carbon.2009.04.026.

35 X. Jin, X. Chen, C. Shi, M. Li, Y. Guan, C. Y. Yu, *et al.*, Determination of hemicellulose, cellulose and lignin content using visible and near infrared spectroscopy in Miscanthus sinensis, *Bioresour. Technol.*, 2017, **241**, 603–609, DOI: 10.1016/j.biortech.2017.05.047.

36 Y.-H. Li, H.-T. Lin, K.-L. Xiao and J. Lasek, Combustion behavior of coal pellets blended with Miscanthus biochar, *Energy*, 2018, **163**, 180–190, DOI: 10.1016/j.energy.2018.08.117.

37 Y. Zhang, I. Zahid, A. Danial, J. Minaret, Y. Cao and A. Dutta, Hydrothermal carbonization of miscanthus: Processing, properties, and synergistic Co-combustion with lignite, *Energy*, 2021, **225**, 120200, DOI: 10.1016/j.energy.2021.120200.

38 J. C. Buzby, H. Farah-Wells and J. Hyman, *The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States*, USDA-ERS

Economic Information Bulletin Number 121, 2014, DOI: 10.2139/ssrn.2501659.

39 M. Pecchi, tga-data-analysis: Tool for automatic analysis of multiple TGA results n.d.

40 M. Pecchi and J. L. Goldfarb, Open-source Python module to automate GC-MS data analysis developed in the context of bio-oil analyses, *RSC Sustain.*, 2024, **2**, 1444–1455, DOI: 10.1039/D3SU00345K.

41 S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, *et al.*, PubChem 2023 update, *Nucleic Acids Res.*, 2023, **51**, D1373–D1380, DOI: 10.1093/nar/gkac956.

42 S. Müller, Flexible heuristic algorithm for automatic molecule fragmentation: application to the UNIFAC group contribution model, *J. Cheminf.*, 2019, **11**, 57, DOI: 10.1186/s13321-019-0382-3.

43 D. Bajusz, A. Rácz and K. Héberger, Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?, *J. Cheminf.*, 2015, **7**, 20, DOI: 10.1186/s13321-015-0069-3.

44 G. Lyu, S. Wu and H. Zhang, Estimation and Comparison of Bio-Oil Components from Different Pyrolysis Conditions, *Front. Energy Res.*, 2015, **3**, DOI: 10.3389/fenrg.2015.00028.

45 O. M. Abdeldayem, C. Dupont, D. Ferras and M. Kennedy, An experimental and numerical investigation of secondary char formation in hydrothermal carbonization: revealing morphological changes via hydrodynamics, *RSC Adv.*, 2025, **15**, 12723–12738, DOI: 10.1039/D4RA08995B.

46 J. C. Buzby, H. Farah-Wells and J. Hyman, *The Estimated Amount, Value, and Calories of Postharvest Food Losses at the Retail and Consumer Levels in the United States*, 2014, DOI: 10.2139/ssrn.2501659.

47 F. Jin, P. Liu, L. Chen, D. Hua and X. Yi, Study on the thermal stability of the bio-oil components by Py-GC/MS, *Energy Rep.*, 2023, **9**, 280–288, DOI: 10.1016/j.egyr.2023.04.001.

48 Q. Lu, X. Yang and X. Zhu, Analysis on chemical and physical properties of bio-oil pyrolyzed from rice husk, *J. Anal. Appl. Pyrolysis*, 2008, **82**, 191–198, DOI: 10.1016/j.jaap.2008.03.003.

49 M. Garcia-Perez, A. Chaala, H. Pakdel, D. Kretschmer and C. Roy, Characterization of bio-oils in chemical families, *Biomass Bioenergy*, 2007, **31**, 222–242, DOI: 10.1016/j.biombioe.2006.02.006.

50 N. Baccile, G. Laurent, F. Babonneau, F. Fayon, M.-M. Titirici and M. Antonietti, Structural Characterization of Hydrothermal Carbon Spheres by Advanced Solid-State MAS 13C NMR Investigations, *J. Phys. Chem. C*, 2009, **113**, 9644–9654, DOI: 10.1021/jp901582x.

51 K. Sheng, S. Zhang, J. Liu, S. E, C. Jin, Z. Xu, *et al.*, Hydrothermal carbonization of cellulose and xylan into hydrochars and application on glucose isomerization, *J. Clean. Prod.*, 2019, **237**, 117831, DOI: 10.1016/j.jclepro.2019.117831.

52 M. C. Johnson and J. W. Tester, Lipid Transformation in Hydrothermal Processing of Whole Algal Cells, *Ind. Eng. Chem. Res.*, 2013, **52**, 10988–10995, DOI: 10.1021/ie400876w.

53 Y. A. Gismatulina, V. V. Budaeva, A. N. Kortusov, E. I. Kashcheyeva, E. K. Gladysheva, G. F. Mironova, *et al.*, Evaluation of Chemical Composition of Miscanthus × giganteus Raised in Different Climate Regions in Russia, *Plants*, 2022, **11**, 2791, DOI: 10.3390/plants11202791.

54 A. C. Gowman, M. C. Picard, A. Rodriguez-Uribe, M. Misra, H. Khalil, M. Thimmanagari, *et al.*, Physicochemical analysis of apple and grape pomaces, *BioResources*, 2019, **14**, 3210–3230, DOI: 10.15376/biores.14.2.3210-3230.

55 J. Dykyj, J. Svoboda, R. C. Wilhoit, M. Frenkel, K. R. Hall and K. R. Hall, in *Datasheet from Landolt-Börnstein – Group IV Physical Chemistry Volume 20B: "Vapor Pressure and Antoine Constants for Oxygen Containing Organic Compounds"*, Organic Compounds, C1 to C57 Part 1, SpringerMaterials, DOI: 10.1007/10688583_3.

56 J. Dykyj, J. Svoboda, R. C. Wilhoit, M. Frenkel, K. R. Hall and K. R. Hall, in *Datasheet from Landolt-Börnstein – Group IV Physical Chemistry Volume 20B: "Vapor Pressure and Antoine Constants for Oxygen Containing Organic Compounds"*, 2 Organic Compounds, C1 to C57 Part 2, SpringerMaterials, pp. , pp. 111–205, DOI: 10.1007/10688583_4.

57 J. Liu, E. Robles Molina and N. S. Mosier, Sustainable biokerosene from lipids using efficient ozone cracking, *Biofuel Res. J.*, 2025, **12**, 2412–2431, DOI: 10.18331/BRJ2025.12.2.4.

58 J. Wang, C. Ye, Z. Ding, M. Fu, Y. Chen, D. Hou, *et al.*, Catalytic Deoxygenation of Oleic Acid and Oils into a Gasoline-Range Hydrocarbon over a Highly Efficient La-Based Al-SAPO-11 Catalyst: A Combined Experimental and DFT Calculation, *ACS Sustainable Chem. Eng.*, 2025, **13**, 15107–15125, DOI: 10.1021/acssuschemeng.5c06220.

59 A. A. Rosatella, S. P. Simeonov, R. F. M. Frade and C. A. M. Afonso, 5-Hydroxymethylfurfural (HMF) as a building block platform: Biological properties, synthesis and synthetic applications, *Green Chem.*, 2011, **13**, 754–793, DOI: 10.1039/C0GC00401D.

60 D. Subramani, S. Mohan, H. Srinivasan and A. Muthukaruppan, Eco-Inspired Multifunctional Cyanate Esters Derived from Bio-Phenols and Fossil-Based Phenols for High Thermal and Super Hydrophobic Applications, *J. Polym. Environ.*, 2025, **33**, 2902–2919, DOI: 10.1007/s10924-025-03572-x.