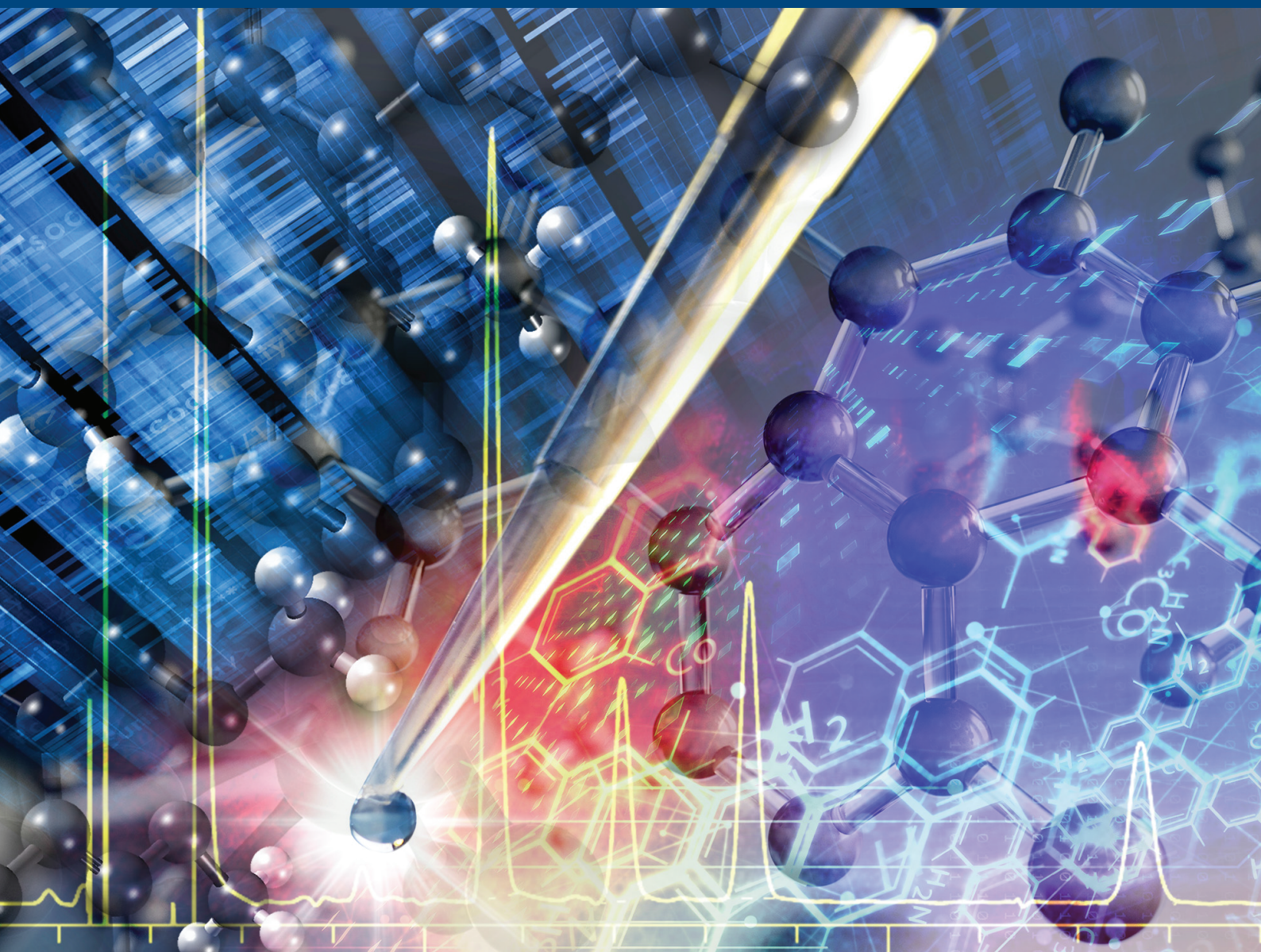




# JOURNAL OF SEPARATION SCIENCE

1 | 2021



**Methods**  
Chromatography · Electroseparation

**Applications**  
Biomedicine · Foods · Environment

[www.jss-journal.com](http://www.jss-journal.com)

**WILEY-VCH**

## REVIEW ARTICLE

# Advanced mono- and multi-dimensional gas chromatography–mass spectrometry techniques for oxygen-containing compound characterization in biomass and biofuel samples

Marco Beccaria<sup>1</sup>  | Anna Luiza Mendes Siqueira<sup>2,4</sup> | Adrien Maniquet<sup>2,4</sup>  |  
Pierre Giusti<sup>3,4</sup>  | Marco Piparo<sup>3,4</sup>  | Pierre-Hugues Stefanuto<sup>1</sup>  |  
Jean-François Focant<sup>1</sup> 

<sup>1</sup> Organic and Biological Analytical Chemistry Group, MolSys Research Unit, University of Liège, Liège, Belgium

<sup>2</sup> TOTAL Marketing Services, Research Center, Solaize, France

<sup>3</sup> TOTAL Refining and Chemicals, Total Research and Technologies Gonfreville, Harfleur, France

<sup>4</sup> International Joint Laboratory - iC2MC: Complex Matrices Molecular Characterization, TRTG, Harfleur, France

## Correspondence

Marco Beccaria, Jean-François Focant, Organic and Biological Analytical Chemistry Group, MolSys Research Unit, University of Liège, Liège, Belgium.  
Email: [marcobecos@gmail.com](mailto:marcobecos@gmail.com);  
[jf.focant@uliege.be](mailto:jf.focant@uliege.be)

A wide variety of biomass, from triglycerides to lignocellulosic-based feedstock, are among promising candidates to possibly fulfill requirements as a substitute for crude oils as primary sources of chemical energy feedstock. During the feedstock processing carried out to increase the H:C ratio of the products, heteroatom-containing compounds can promote corrosion, thus limiting and/or deactivating catalytic processes needed to transform the biomass into fuel. The use of advanced gas chromatography techniques, in particular multi-dimensional gas chromatography, both heart-cutting and comprehensive coupled to mass spectrometry, has been widely exploited in the field of petroleomics over the past 30 years and has also been successfully applied to the characterization of volatile and semi-volatile compounds during the processing of biomass feedstock. This review intends to describe advanced gas chromatography–mass spectrometry-based techniques, mainly focusing in the period 2011–early 2020. Particular emphasis has been devoted to the multi-dimensional gas chromatography–mass spectrometry techniques, for the isolation and characterization of the oxygen-containing compounds in biomass feedstock. Within this context, the most recent advances to sample preparation, derivatization, as well as gas chromatography instrumentation, mass spectrometry ionization, identification, and data handling in the biomass industry, are described.

## KEYWORDS

analytical pyrolysis, biofuels, biomass, GC-MS, oxygen compounds

**Article Related Abbreviations:** 1D, mono-dimensional; 2D, two-dimensional; <sup>1</sup>D, first dimension; <sup>2</sup>D, second dimension; ASTM, American society for testing and materials; CI, chemical ionization; CPO, catalytic pyrolysis oil; DA, discriminatory analysis; EI, electron-impact ionization; FAME, fatty acid methyl ester; FI, field ionization; FID, flame ionization detector; HCA, hierarchical clustering

analysis; HD, hydro-deoxygenation; HDO, hydro-deoxygenated oil; HTL, hydrothermal liquefaction; HTO, hydro-treated oil; IL, ionic liquid; LR, low-resolution; LRI, linear retention index; MD, multi-dimensional; NEFA, not esterified fatty acid; NIST, national institute of standards and technology; PCA, principal component analysis; PLS, partial least squares; PW, pinewood; PWT, pinewood bio-oil after thermal

## 1 | INTRODUCTION

The production of chemicals from biomass is highly desirable in the medium/long term to replace petrochemical products. Dependency on fossil-based energy sources, decreasing accessibility to crude oil, and environment protection from carbon dioxide or greenhouse gases emissions has led to increased interest in becoming less fossil-energy-dependent than in the past, although, together with the increasing of the world population, the global energy demand is rising [1]. Biomass is one of the few resources showing high potential to meet sustainable and green energy challenges. Hence, its use is expected to grow in the foreseeable future [2]. The best approach to compete with fossil-based refineries is biomass processing in integrated biorefineries (facilities integrating biomass conversion processes and equipment to produce fuels, power, and chemicals from biomass) [3]. A biorefinery produces high-value low-volume chemical products and low-value high-volume biofuels using different unit operations, maximizing the mass and energy efficiency and minimizing the waste streams. The biomass transformation in biorefineries is still a process in the early stage, intending to obtain products with a high H:C ratio, transforming and drastically reducing the amount of heteroatoms containing compounds for fuel applications [1,4]. Heteroatoms are not always efficiently removed, and undesired products are often formed during the process. Heteroatoms, in particular oxygen and sulfur, can promote corruptions, thus limiting and/or deactivating the catalytic processes needed for the transformation of the biomass into fuel, compromising the quality of the final chemical products [5,6].

In light of these considerations, the characterization of these heteroatom-containing compounds is necessary to isolate and ride them out to preserve the quality of the final chemical products.

The most suitable and applied technique for the analytical characterization of biomass and bio-oil volatile and semi-volatile heteroatom compounds is GC. GC techniques are often employed, after fat hydrolysis, for the analysis of (i) esterified fatty acids (FAs), usually, methyl esterified (FAME); (ii) analysis of FAMEs oil after thermochemical processing such as pyrolysis or hydrothermal liquefaction deriving from the oleaginous feedstock. In bio-fuels derived from carbohydrates (pyrolysis- or bio-oil), GC usually covers the analysis of the medium-polar monomers (around the 40 wt%) and, after derivatization, can also cover the polar monomers (10–15 wt% in bio-oils) [7]. The hyphenation between GC and MS has become

very important in biomass analysis due to its high throughput, selectivity, and sensitivity, adding an analytical dimension for the identification and separation of heteroatom-containing compounds. Although the great potentiality of MS as a detector, quantification of volatile and semi-volatile biomass organic compounds has been preferentially done over the years using a flame ionization detector (FID), limiting the coupling of the MS to qualitative purposes.

Although conventional GC has been extensively and successfully employed for the analysis of biofuels, many samples are not completely resolved and thus characterized [8]. The use of multi-dimensional (MD)GC proved to be a powerful approach to improve analyte resolution in derived biomass characterization, increasing the peak capacity, and enhancing the identification accuracy [9,10]. The scope of this review is to describe, over the past decade, the most recent advances in GC–MS-based techniques for the analysis of oxygen-containing compounds in biomass feedstock, with particular emphasis on the MDGC-MS techniques. Except for some contributions worthy to note, the attention has been focused on the most innovative applications in the period 2011–early 2020. Within this context, dedicated sample preparation, derivatization, injection, as well as the GC instrumentation, MS ionization and identification, and data handling in the biomass industry are reported too. Moreover, despite the prevalent qualitative use of MS, different quantitative GC–MS solutions have been proposed during the past decade and the most interesting applications are reported. A complete list of all scientific contribution present in this review is reported in Supporting Information Table S1. Specific information, in terms of feedstock investigated, treatment and processing, bio-samples analyzed, sample preparation step, GC mode applied, GC column(s), MS, and statistical approaches, are reported for all contributions.

## 2 | BIOMASS FEEDSTOCK AND PROCESSING

Due to the high availability and the fact that it is renewable, biomass is one of the most promising candidates for the substitution of crude oil as the principal source for chemical energy. The C, H, and O composition of the most common biomass-derived feedstocks are reported in Figure 1. The oil obtained after pyrolysis (pyrolysis- or bio-oil) and crude oil, as “gold standard,” are also reported [11].

As reported in Figure 1, the high percentage of oxygen is in cellulose, bio-oil, and lignin. Due to the very low energy density of these raw materials, their conversion for removing oxygen efficiently from the biomass feedstock is necessary. Biomass feedstock can be grouped into two

decomposition; Py, analytical pyrolysis; Q, quadrupole; QQQ, triple quadrupole; RP, resolving power; SIM, single ion monitoring; SVM, support vector machine; TG, triglyceride; TIC, total ion current

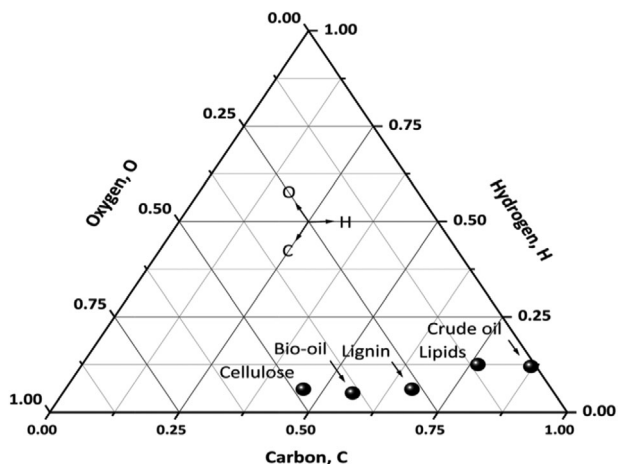


FIGURE 1 Ternary diagram showing the mass composition of biomass-derived raw materials. Figure reprinted with the permission of Royal Society of Chemistry, in Melero et al. [11]

broad categories: oleaginous feedstock and carbohydrates. Before a description of the products that can be obtained, a brief overview of the different processes used for their conversion is necessary. The challenge is to produce processed biomass with high energy density and excellent combustion properties. To make this real, the amount of heteroatoms containing compounds as oxygenated and sulfurated has to be minimized as much as possible [12,13].

A schematic presentation of the most common source of biomass resources, conversion technologies, bio-fuels obtained, and end of use products is reported in Figure 2 [14].

Biomasses can be transformed using biochemical, physical-chemical, and thermo-chemical processing. Among these, thermo-chemical represents the most common processing technique to convert biomass in bio-energy. Within this macro-technique, pyrolysis, gasification, and combustion represent the most common approaches. Pyrolysis is a thermal process in which the biomass is heated at very high temperatures (350–550°C), in the absence of oxygen and air. Fast (few seconds) or flash (<1 s) pyrolysis is usually applied in order to maximize the production of liquid and/or gas phase products (optimizing the content of H) over the formation of charcoal products (higher O content). Gasification is the conversion of non-gaseous feedstock into gaseous fuel (syngas) by reacting the material at high temperatures (>700°C), without combustion, with a controlled amount of oxygen. Thermochemical technologies also include hydrothermal approaches, which utilize an aqueous environment at moderate temperatures (200–600°C) and high pressures to decompose biomass into solid, liquid, and gaseous intermediates [15].

## 2.1 | Oleaginous biomass feedstock

An oleaginous feedstock is generally used for the production of biodiesel and jet fuel. Low-quality lipids are preferred, such as (i) crude vegetable oils (soybean, palm, rapeseed, and sunflower); (ii) used vegetable oils; (iii) wasted animal fats; (iv) non-edible oils [16–19]. Currently, algae and marine organisms have been more intensively investigated as a source of TGs [20–23]. The advantages of oleaginous feedstocks are their high availability and similar physical properties (e.g., density, viscosity) and H/C ratio to crude oil-based materials.

Oleaginous biomass is mainly composed of triglycerides (TGs) containing FA carbon chains, usually in the range of C8–C22, but C16, C18, and C20 are the most common [24,25]. TGs have a high chemical energy conversion due to their very high H:C ratio and low oxygen content. Another important fraction present in low-grade of oleaginous material is formed by not esterified (NE)FAs.

Catalytic and thermal cracking, hydro-treating, and catalytic deoxygenation are the most applied processes to convert oleaginous biomass in fuels [26]. Even if lipids present a rather low amount of oxygen and metals, making them good potential candidates to replace petrochemical products, the high cost of processing makes the use of lignocellulosic feedstock still necessary.

## 2.2 | Lignocellulosic based feedstock

Lignocellulose is the most common form of polysaccharide in feedstock, and it is composed of three components (cellulose (40–50 wt%); hemicellulose (25–40 wt%); and lignin (10–25 wt%)). A high percentage of cellulosic biomass can be produced via dedicated crops. Wastes and residues are other important sources of lignocellulose biomass [27].

### 2.2.1 | Pyrolysis (bio)-oil from lignocellulosic feedstocks

(Pyrolysis) bio-oil from lignocellulosic feedstocks is composed by (i) aqueous phase that represents around 15–30 wt%, where low molecular weight oxygenated organic compounds are present and (ii) a non-aqueous phase, representing from 35 to 50 wt% mainly by several oxygenated molecules, as aliphatic alcohols, carbonyls, acids, phenols, sugar hydroxyl-aldehydes, hydroxyl-ketones, sugars, and aromatic hydrocarbons [28]. Gaseous products (e.g., CO/CO<sub>2</sub>; CH<sub>4</sub>, higher hydrocarbons) and a solid carbonate residue are also formed during the pyrolysis treatment. The chemical composition of bio-oils produced from process conditions that maximize liquid yields is very

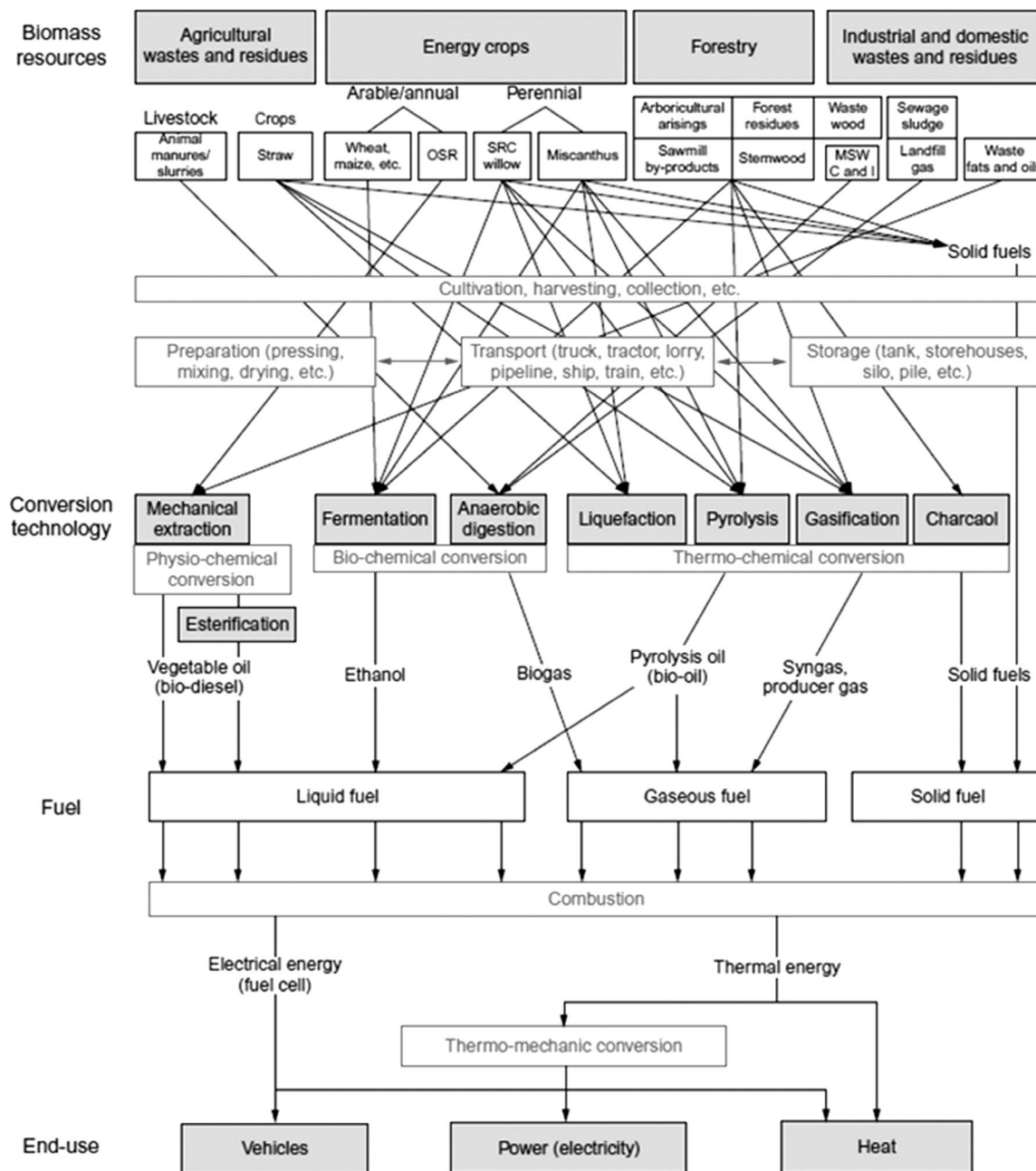


FIGURE 2 Schematic presentation of biomass conversion pathways. Reprinted with the permission of Academic Press (Elsevier), in Adams et al. [14]

complex, and a precise description of bio-oil composition has never been reported yet. Even with considerable analytical efforts, about 20% of the composition still remains unknown [29]. High-lignin containing biomass gives lower bio-oil yield, while a higher percentage of bio-oil is obtained in-high cellulosic biomass. The properties and composition of bio-oils depend both on the specific starting feedstock and the treatment conditions. Bio-oil processing requires a considerable effort in commercial development since commercial production is still in the early stage. Several physical-chemical properties of raw bio-oil from lignocellulosic biomass make it unsuitable for

use in petroleum refineries, as high content in water, high viscosity, low volatility, and high reactivity (which may lead in particular to corrosion issues) that can damage reactors and instrumentations, and a significant content of heteroatoms containing compounds that can poison catalytic processes. With the approval of the standard specification for pyrolysis liquid biofuel by the American Society for Testing and Materials (ASTM) D7544, pyrolysis oil is being strongly considered as a fuel oil replacement. The physical properties of bio-oils obtained from different biomass feedstocks and pyrolysis process steps are well-described by Abdelnur and co-workers [30]

### 3 | OXYGEN-CONTAINING COMPOUNDS IN BIOMASSES

The raw bio-oil material, in particular from lignocellulosic feedstocks, is typically heavily oxygenated, corrosive, unstable, viscous, and immiscible with petroleum feedstocks. These undesirable fuel properties, together with the catalyst poison, are also a result of the several hundreds of reactive oxygenates in the bio-oil. Typically, the concentration of any particular oxygenated containing compound is less than 10 wt% although there is a large variability in the measurements of individual molecular species in the products [29]. A round-robin study, done among different laboratories analyzing the same bio-oil, highlighted the difficulty in measuring these oxygenated containing compounds, leading to non-consistent results [31].

A complete characterization of oxygenated containing compounds in bio-oils is of paramount importance to assess their potential uses or decide upon the proper upgrading method for improving their properties. The oxygen reduction processing (e.g., hydro-treating), necessary for decreasing the amount of oxygen in a range of 2–3 wt%, can have a dramatic incidence on the costs of production. These costs can even exponentially increase in case the goal is to obtain a product with oxygen content < 2 wt% [29]. In fact, certain oxygen functional groups in pyrolysis oils are unlikely to be acceptable in fuel products, even after upgradation of the bio-oil. At a low amount, carboxylic acids can be used in fuels as corrosion inhibitors. At the same time, their corrosion property and poor hydrocarbon solubility at cold temperatures have been observed when present in higher amounts. The possible condensation of aldehydes and ketones can lead to the formation of gums, although no published studies are showing this occurring [29]. Considering that a complete elimination of oxygenated compounds during biomass processing is not possible, some of them, such as esters, ethers, and alcohols, can be accepted in a certain amount in the final product [32].

### 4 | SAMPLE PREPARATION

Biomass feedstocks (in particular lignocellulosic biomass) subjected to thermochemical processing represent a complex assemblage of chemicals containing varied functional groups, although the thermal cracking of these biomolecules generates simpler organic compounds. Lignins, upon pyrolysis, produce methoxyphenols of which monolignols (guaiacol and syringol) are prominent. Cellulose upon pyrolysis gives levoglucosan, levoglucosenone, fur-

fural, substituted furans, and several small molecules comprising aldehydes and ketones. In this context, the role of sample preparation is mainly meant to make the analytes of interest more suitable for separation and detection [33,34]. Here it has been reported a brief description of the most common approach of sample preparation techniques hyphenated to GC–MS in the context of biomass analysis.

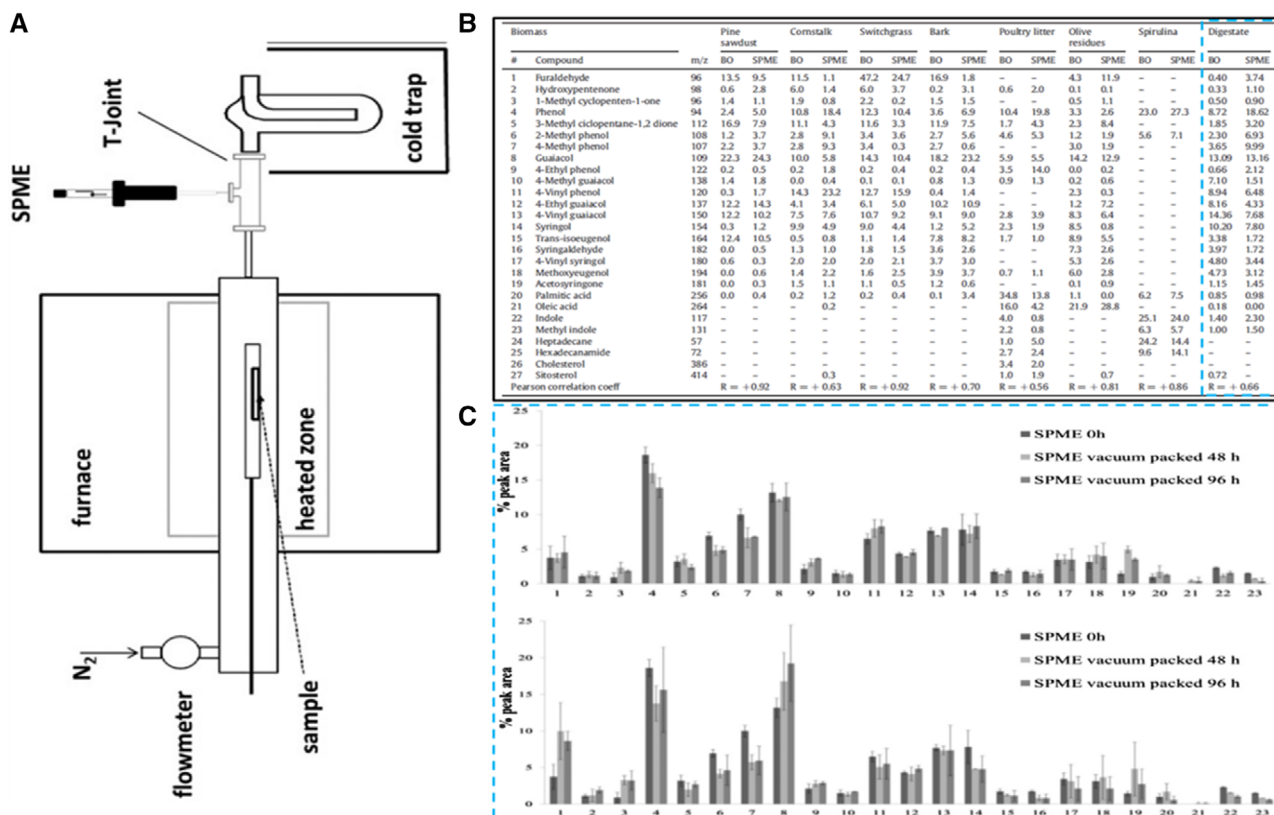
#### 4.1 | Solid-phase micro-extraction

Solid-phase micro-extraction (SPME), although a widely applied technique for volatile and semi-volatile analytes, has found limited application in this field compared to other techniques, probably due to the high number of chemicals present in bio-oil [35,36].

Tessini and co-workers used HS-SPME to extract oxygenated compounds, in particular, low molecular weight aldehydes present in bio-oils after derivatization in solution as well as on-fiber derivatization. After thermal desorption into the injector, these compound derivatives were analyzed by GC–MS [37]. Recently, Conti and co-workers used at-line sampling by SPME to trap and GC–MS to characterize the compounds evolved during intermediate pyrolysis of biomass. The molecular composition resulting from SPME–GC–MS was compared to those obtained by GC–MS analysis of aqueous and organic phase (bio-oil), reporting similar results in terms of oxygenated compounds. Moreover, the authors claimed that SPME sampling to monitor the pyrolysis process could limit laboratory working time, avoiding sample collection, and pretreatment [38]. Figure 3 shows the bench-scale reactor with the addition of a junction for the SPME sampling, a comparison between the GC–MS analysis of bio-oil from different origins and SPME–GC–MS analysis of the compounds evolved during intermediate pyrolysis of biomass, as well as the product distribution from SPME–GC–MS analysis of digestate soon after sampling (0 h) and after storage 48 and 96 h in vacuum-packed bags and under air atmosphere.

#### 4.2 | Liquid–liquid extraction

Liquid–liquid extraction has been used with different combinations of solvent to exploit the selectivity towards specific target compounds. For instance, using *n*-hexane, petroleum ether, or chloroform as the extraction solution, phenols, and guaiacol can be enriched in the solvent phase. At the same time, sugar, acid, and alcohol can be concentrated in the water phase [39,40].



**FIGURE 3** (A) The bench-scale reactor with the addition of a quartz T-junction for the SPME sampling and (B) comparison between the GC-MS analysis of bio-oil (BO) from different origins and SPME-GC-MS analysis of the compounds evolved during intermediate pyrolysis of the same biomass feedstocks, and (C) product distribution from SPME-GC-MS analysis of digestate soon after sampling (0 h) and after storage 48 and 96 h in vacuum-packed bags (above) and under air atmosphere (below). The compounds corresponded to the x-axis number are reported in figure 3b. Reprinted and adapted with the permission of Elsevier, in Conti et al. [38]

### 4.3 | Supercritical-fluid extraction

In 2016, Cheng and co-workers developed a three-step supercritical-CO<sub>2</sub> extraction for the selective fractionation of fast pyrolysis bio-oil, demonstrating the utility of this extraction technique in bio-oil processing in oxygen-containing compound determination. With appropriate optimization of extraction parameters, lipids, hemicellulose, lignin, and condensed aromatics were enriched in three different fractions, and oxygen-containing compounds were characterized using various analytical techniques. Among them, GC-MS was the most useful, allowing the identification of a total of 132 compounds (>95% oxygenated). [41].

### 4.4 | Distillation

Example of distillation used in bio-oil fractionation was reported by Christensen and co-workers in 2011, where hydro-treated bio-oils containing different amount of oxygenated compounds was distilled to produce differ-

ent boiling range fractions (i.e., lights [ $<71^{\circ}\text{C}$ ], naphtha [ $71\text{--}182^{\circ}\text{C}$ ], jet [ $182\text{--}260^{\circ}\text{C}$ ], diesel [ $260\text{--}338^{\circ}\text{C}$ ], and gas oil [ $338\text{--}566^{\circ}\text{C}$ ]) that were characterized for oxygen-containing species by different analytical techniques, including GC-MS [42].

### 4.5 | Column chromatography

Within chromatographic techniques applied to fraction bio-fuels, column chromatography still represents the most conventional method to fractionate bio-oils according to the different adsorption capabilities of bio-fuel components onto the stationary phase (in general, silica gel, or aluminum oxide). Analytes are eluted based on polarity using various combinations of organic solvents. The consumption of a large amount of solvent and the low throughput of the technique make column chromatography only suitable for high value-added compounds [43]. Other chromatographic techniques often used in the fractionation of bio-oils are gel permeation chromatography/size exclusion chromatography, thin-layer

chromatography, ion-exchange chromatography, and adsorption chromatography. Within the latter, SPE has been widely applied in biofuels analysis, especially as a purification step and determination of the impurities in biofuels [44]. The principles of chromatographic techniques are based on the difference of interaction of the components with the stationary phase in the chromatography column, inducing the difference of retention time. Each technique presents some advantages and disadvantages and the selection of a specific one rather than the other often depending on the chemical composition and structure of the analytes to isolate [45].

#### 4.6 | Derivatization

Due to a large amount of non-volatile polar compounds in bio-oil, direct analysis in the GC system is limited, and only a small portion of fractions (from 25 to 40%) can be analyzed. [46]. An example is the high molecular weight carbohydrates obtained by the pyrolysis of lignocellulosic biomass. The low volatility of these high molecular weight molecules limits the use of GC. Moreover, oxygenated compounds present in biofuels can be reactive, causing thermal instability and other problems. The replacement of the active hydrogen (H) in polar groups, such as hydroxyl (OH), thiol (SH), and amine (NH), represents a large percentage of the derivatization process. Among the different derivatization reactions, alkylation and trimethylsilylation are the most common and widely used [47]. Madsen and co-workers used methyl-chloroformate under alkaline conditions in the aqueous phase obtained after hydrothermal liquefaction (HTL) treatment of biomass. A central composing design with surface responses methodology was applied to optimize derivatization conditions, and 32 analytes were quantified by GC-MS [48]. It is worth mentioning that (trans)esterification is also a fundamental step in biofuel production from oleaginous biomass, simplifying the analysis of these products. Different catalysts have been applied in this production step, such as tin organometallic compounds, resins, acids, and enzymes [49,50].

Silylation is generally used in the derivatization of sugars and considerate the thermal stability, even after depolymerization of polysaccharides using pyrolysis. Although several compounds after pyrolysis can be analyzed by GC without any derivatization (as the case of anhydrosugars, that present only three OH groups), derivatization is considered as the best option to avoid the retention of hydroxyl groups of pyrolyzed products that affect GC resolution and peak broadening. Resulting products of pyrolysis and silylation of polysaccharides are, in general, propanones, furanones, anhydro sugars acids, etc. Complete derivatization

of compounds with multiple functional polar groups can be challenging. The use of the derivatization reagent *N*-methyl *N*-trimethylsilyl) trifluoroacetamide was proposed by Joseph and co-workers for the complete derivatization of compounds with multiple -OH and -COOH groups, such as polar sugars and polyhydroxylated benzenes [46]. The complete derivatization of no-reducing sugar gives single peaks in GC analysis since no anomeric form is present. On the contrary, the different anomeric forms in reducing sugars can lead to the formation of multiple peaks due to the interconversion between anomers via the open-chain form. The reduction of saccharides to alditols avoids the problems when mixtures of anomers are present, but a single alditol can be the common final product obtained from the reduction of different saccharides, leading to a loss of structural information. The protection of the carbonyl group before silylation through the formation of oximes limits the chromatogram complexity. Still, each reducing sugar could result in two different peaks due to the formation of the *syn*- and *anti*-forms of the oximes [47].

## 5 | GC-BASED TECHNIQUES

For the determination of the chemical composition of biofuels, several analytical methods are usually employed. Among them, GC-MS represents the preferred method for compositional analysis of bio-oils, bio-crudes, and upgraded products [51]. GC has been applied in both mono-dimensional and multi-dimensional mode coupled to MS. In this review, the main focus will be on multidimensional techniques [10], which have found a wide application in high-end applications dealing with challenging samples, such as biomass characterization and online combination with analytical pyrolysis (Py).

### 5.1 | Multidimensional GC

#### 5.1.1 | Heart-cutting 2D-GC-MS

Although heart-cutting GC (GC-GC) has been widely used in the petrochemical field for the analysis of oxygenates compounds in fuels, only a few applications have been found in biomass analysis [52]. To the best of our knowledge, the first application on the analysis of biomass oils was reported in 2005 by Fullana and co-workers using a multiple heart-cutting 2D-GC-TOFMS to analyze pyrolytic products generated from primary and secondary pyrolysis of cellulose, lignin, and sewage sludge samples. The column set employed was a combination of non-polar 30 m × 0.53 mm × 1 μm as first dimension (<sup>1</sup>D) column and a polar 2.5 m × 0.1 mm × 0.1 μm as second



dimension ( $^2D$ ) column, using a cryogenic trap system to focalize the analytes from the first into the head of the second column. Authors reported that >1000 compounds, mainly oxygenated (containing up to seven oxygen atoms), were identified for each pyrolytic oil analyzed. Sugars and phenols were the most representative chemical families in bio-oils from cellulose and lignin, respectively, while aliphatic compounds and cellulose derivatives were the most abundant in bio-oils from sewage sludge. Moreover, a comparison between heart-cutting 2D-GC-MS and mono-dimensional GC-MS, in terms of MS similarity (match quality), was done. Authors claimed that more than 70% of the total chromatogram peaks could be identified with 2D-GC-MS while only 47% (in the best case) using conventional GC-MS [53]. In 2016, Le Breuch and co-workers used a heart-cutting 2D-GC-FID/MS for the quantitative analysis of condensed vapors from slow pyrolysis of different biomass feedstock (miscanthus, douglas fir, and oak). The compounds from pyrolysis were condensed using cold traps, and tetradecene was used as an internal standard. A combination of  $^1D$  non-polar and  $^2D$  polar columns was employed for the multi-dimensional separation, using a Dean's switch valve [54] to transfer eluates from  $^1D$  to  $^2D$  column. More than 100 oxygenated compounds have been detected, and 30 compounds have been identified based on the literature, and the NIST database. Between 25 and 30 compounds have been quantified for each biomass sample. The authors grouped oxygenated compounds in three categories, depicting the different chemical pathways of pyrolysis as functions of temperature and biomass composition: (i) fragmentation/dehydration pathways  $\rightarrow$  light oxygenated compounds; (ii) dehydration/open-ring/cyclization mechanism  $\rightarrow$  furanic compounds; and (iii) transglycosylation pathway  $\rightarrow$  anhydro-sugars [55].

### 5.1.2 | Comprehensive 2D-GC-MS

Comprehensive 2D-GC (GC $\times$ GC) has been successfully applied for the characterization of samples from lignocellulosic biomass and methyl ester derivative of oleaginous (bio) oils, in particular applying a normal column-set (i.e., non-polar  $\times$  polar column); while the reversed column-set (i.e., polar  $\times$  non-polar) is generally less used. The reversed-set can be useful in the analysis of biodiesel blend in diesel fuels, especially when a separation among FAMES, long branched aromatics, naphthenic compounds, and paraffins is needed [56]. To resolve complex samples as biofuels, normal- and reversed-sets can also be used alternatively in the same GC system. Djokic used a GC $\times$ GC-TOFMS system equipped alternatively with both column-sets for the analysis of crude- and

hydrotreated bio-oil from pine wood, improving the group-type separation of the region of aldehydes, ketones, and furans [57]. The higher peak capacity, selectivity, sensitivity (in particular when cryogenic modulators are employed), and formation of highly organized chromatography structures represent the principal advantages of GC $\times$ GC compared to 1D-GC. A group-type identification approach of petrochemical oil samples by using GC $\times$ GC-TOFMS was introduced by van Deursen and co-workers in 2000. The authors showed that selecting appropriate ion fragments, sulfur- and oxygen-containing compounds (hardly detected in the total ion current (TIC) plot) were easily detected and identified into the 2D-GC space [58]. Since then, this approach has been used in the GC $\times$ GC-MS analysis of petrochemical and biofuel samples.

A group-type identification approach using GC $\times$ GC-TOFMS in TIC and extracted ion chromatogram mode was also carried out by Tessarolo and co-workers, analyzing the pyrolysis products of empty palm fruit bunch, pine wood chips biomass, and sugarcane. Cyclopentenones, furanones, furans, phenols, benzenediols, methoxy- and dimethoxy-phenols and sugars were the more expressed chemical classes (analytes > 0.5% relative area), indicating the suitability of these oils for the production of value-added chemicals [59,60]. In 2014, Joffres and co-workers characterized the products of the wheat straw soda lignin residue after catalytic hydroconversion combining different analytical techniques. A cryogenic GC $\times$ GC-qMS system, using a "normal set," composed of  $^1D$  non-polar (30 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m) column and a  $^2D$  polar (2 m  $\times$  0.1 mm  $\times$  0.1  $\mu$ m) column, was employed for qualitative characterization of oxygenated compounds present into the liquid phase. Identification was carried out considering the 2D structured chromatogram and MS spectra similarity, while the use of linear retention index (LRI) is not reported. The liquid phase represents about 65% of the initial lignin, and it is considered as the most challenging fraction to characterize, not only due to the complexity of the sample but also for the presence of the tetralin solvent. Comparing the GC $\times$ GC-MS chromatograms obtained analyzing the liquid phase with and without catalyst, the authors noticed a notable increased amount of phenol, aromatic, and naphthene compounds in the liquid phase obtained with the presence of a catalyst. Moreover, the authors linked the nature of the compounds detected with the transformation observed in lignin residue [61].

In 2016, Silva and co-workers used GC $\times$ GC-MS technique for quantitative characterization of pinewood (PW) bio-oils derived by different processes, namely real thermal decomposition, catalytic pyrolysis (CP), and hydro-deoxygenation (HD). Bio-oil from thermal decomposition was upgraded by HD to hydro-deoxygenated oil (HDO). Sugars, furans, and alcohols appear as the major

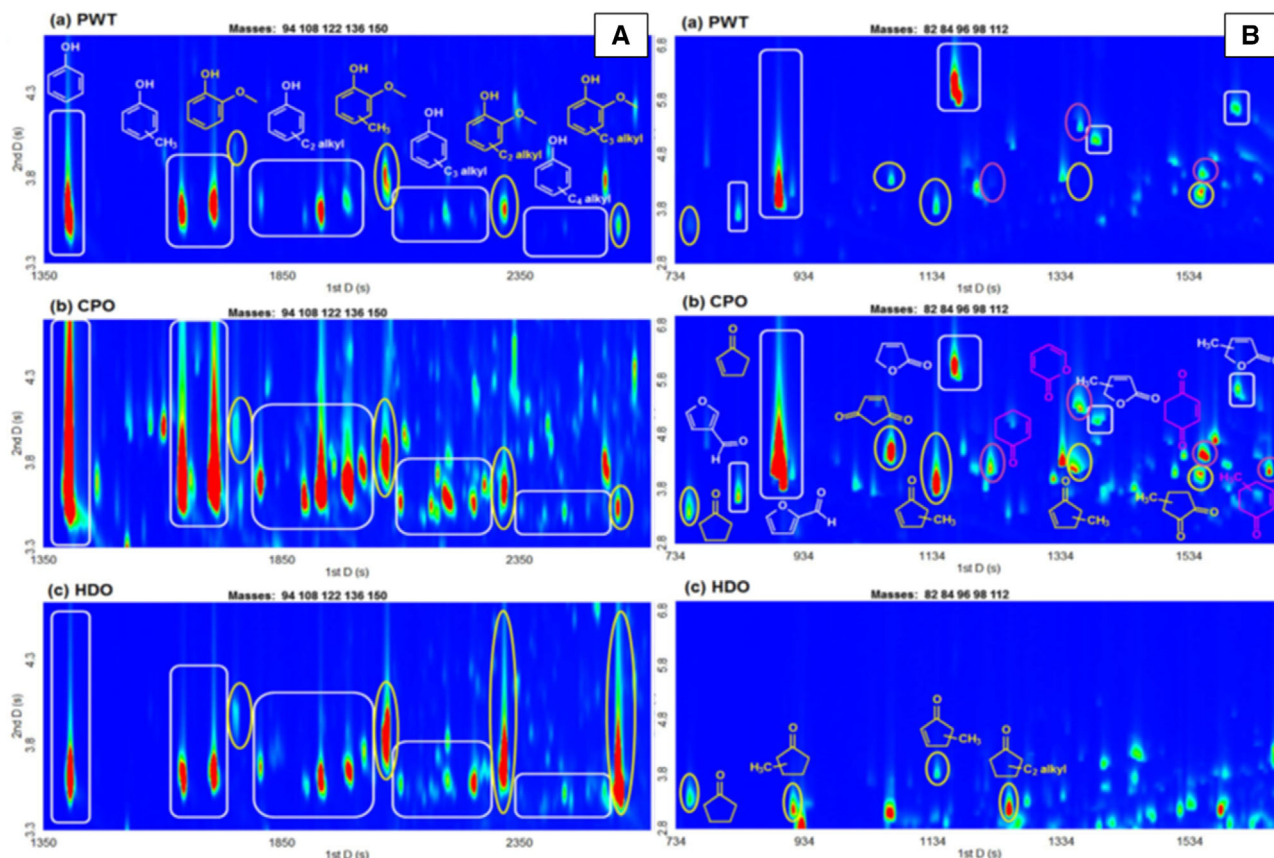


FIGURE 4 GCxGC expansion chromatogram of three samples (a) PWT, (b) CPO, and (c) HDO. (4A) Extracted ion chromatograms (EIC)  $m/z$  94,  $m/z$  108,  $m/z$  122,  $m/z$  136, and  $m/z$  150, illustrating the identification of phenolics and methoxy-phenols. (4B) EICs  $m/z$  82, +  $m/z$  84, +  $m/z$  96, +  $m/z$  98, and +  $m/z$  112 show cyclic ketones and furans. Reprinted and adapted with the permission of Elsevier, in Silva et al. [62]

constituents of the pinewood bio-oil after thermal decomposition (PWT), catalytic pyrolysis oil (CPO), and HDO samples, respectively. Figure 4 shows the expansion 2D chromatogram of the three bio-oils obtained from the different processes. In Figure 4A, using extracted ion chromatograms  $m/z$  94, 108, 122, 136, and 150, a series of phenolics and methoxy-phenols were observed in the three samples (a) PWT, (b) CPO, and (c) HDO, while Figure 4B shows cyclic ketones and furans assigned in the samples by the extracted ion chromatograms  $m/z$  82, +  $m/z$  84, +  $m/z$  96, +  $m/z$  98, and +  $m/z$  112. An increased amount of phenol was observed in CPO, while methoxy-phenols were mainly observed in the HDO. The same behavior was observed for cyclic ketones and furans, more expressed in CPO and less in HDO. Quantification was performed using an internal standard and external calibrations. The limit of quantification was set at 1 ng/ $\mu$ L for major standards, except for hexanoic acid, which was set at 5 ng/ $\mu$ L. The precision (< 10%) and accuracy (recovery range of 70–130%) were also reported for the quantification of oxygenated compounds in bio-oil samples [62].

In 2017, Sajdak and co-workers characterized oxygen compounds present in bio-tars derived from the thermal

conversion of wood biomass by 1D and comprehensive 2D-GC-MS. As a preliminary step, three different stationary phases, namely polar, non-polar, and an ionic liquid (IL-60) were compared to optimized 1D-GC-MS methods. Derivatization through silylation was also applied for allowing the analysis of high boiling substances as polyphenols and polyalcohols. Finally, a comprehensive 2D-GC-TOFMS with a cryogenic modulator, using a  $^1$ D (27 m  $\times$  0.25 mm  $\times$  0.25  $\mu$ m) non-polar column  $\times$   $^2$ D (2 m  $\times$  0.1 mm  $\times$  0.1  $\mu$ m) IL-60 column was employed for the analysis of tars, with and without derivatization step. Analytes were identified, combining MS spectra information with available LRI. Although there was a clear improvement in chromatogram resolution and separation of polyalcohols and polyphenols by using GCxGC, especially when a pre-derivatization step was applied, the authors highlighted the limits to analyze compounds with high mass and high polarity (originated from the decomposition of natural polymers, like lignin, cellulose, and hemicellulose), even when multi-dimensional GC techniques are employed [63]. In the same year, Madsen and co-workers used a highly efficient in situ gas-phase derivatization (silylation by *N*-methyl-*N*-(trimethylsilyl)-trifluoroacetamide)

coupled to GC×GC–TOFMS to characterize semi-volatile organic compounds (in particular oxygenated and nitro compounds) in bio-crude oil obtained from HTL of both algae and lignocellulosic biomass. The GC system was equipped with a <sup>1</sup>D (60 m × 0.25 mm × 0.25 μm) non-polar column and a <sup>2</sup>D (1 m × 0.25 mm × 0.25 μm) polar column with a dual-stage loop modulator. Analytes were identified according to MS similarity and LRI, although no minimum cutoff was reported. The authors identified numerous C<sub>16</sub>, C<sub>18</sub>, and C<sub>20</sub> FAs and several monoacylglycerides. Moreover, FA-amides, indoles, and quinolones, derived from the presence of protein-rich feedstocks, were detected too. A semi-quantification of the identified compounds present in the different samples was reported after TIC normalization [64].

One year later, GC×GC–MS technique was employed for the first time for the analysis of biogas and bio-methane. Volatile compounds in bio-gasses from four different industrial sites were trapped into two-beds sorbent tubes through a vacuum pump, and compounds collected on the first sorbent bed were extracted by pressurized liquid extraction, evaporated until a final volume of 150 μL, and then injected into the GC system. The breakthrough volume was checked by analyzing the second bed of the sorbent tube. Separation was performed on <sup>1</sup>D non-polar column (30 m × 0.25 mm × 0.25 μm) and a <sup>2</sup>D semi-polar column (1.35 m × 0.25 μm × 0.25 μm), using a cryogenic modulator. A mixture of 89 standards (including also oxygen and sulfur organic compounds), representative of ten chemical families, was used for the method optimization. Sensitivity was fivefold more as compared with 1D–GC–MS. A series of oxygenated compounds were found in the bio-samples, although the major part remains unknown. The authors claimed that comprehensive 2D–GC–MS could allow enhanced diagnostics for process performances or risk management in the current contest of the new gas and bioenergy industry [65]. In the past 5 years, GC×GC–MS methodology has been widely employed for the characterization FAMES, phenols and ketones and other oxygenated compounds in bio-oil from the pyrolysis of different biomass feedstock, such as coffee waste, [66,67], coconut fibers [68], rice straw [69,70], tropical seeds [71], residue of vegetable oil [72], as well as in the characterization of creosote from wood coal tars [73], and the hydrogenation products of bio-crude obtained from hydrothermally liquefied algal biomass [74].

Due to the complexity of the samples, the GC×GC–MS approach is mainly employed in the analysis of biofuels from lignocellulosic biomass, especially when an improvement in selectivity and sensitivity is needed (e.g., trace analysis). GC×GC–MS has been applied in the analysis of oleaginous biomass by many researchers, mainly after derivatization and transformation in FAMES [75], but con-

sidering that simpler nature of oleaginous biomass, it is the opinion of the authors that a well-optimized 1D–GC–MS method with the use of LRI as an additional filter for compound identification could be often enough to get the desired information. GC×GC–MS can better fit in the analysis of oleaginous biomass subjected to upgradation (e.g., thermochemical process) [76]. For example, the thermo-conversion of FAMES can lead to the generation of a vast number of compounds that can be difficult to analyze by only using 1D–GC–MS [64].

## 5.2 | Analytical pyrolysis–GC–MS

Py–GC–MS is an online technique that involves the heating of the sample at high temperatures (usually > 600°C) in the absence of oxygen, causing decomposition of the macromolecules (pyrolysis). Py–GC–MS analysis can be performed directly on the raw biomass or the biofuel obtained from conversion technologies. Py–GC–MS has proved to be a reliable technique for fingerprinting lignin, generating diagnostic products without altering the groups attached to the phenyl rings, and preserving the isotopic signature of the original material [77]. In 2013, Brebu and co-workers monitored the thermal degradation products of various lignins of different raw materials and manufacturing processes by Py–GC–MS using a non-polar GC column. Similarities and differences in thermal behavior and composition of degradation products among studied lignins were observed and discussed. In addition to oxygen compounds, representing almost the entire composition of the compounds detected, sulfur-containing compounds were also found in traces in Lignoboost® oil (originating from the Kraft pulping) and in a more significant amount in the pyrolysis oil of Klason wheat straw, probably derived during the hydrolysis process with sulfuric acid (Klason lignin is considered the insoluble residue portion after removing the ash by concentrated acid hydrolysis of the plant tissues). Only qualitative analysis was performed, and compounds were identified based on the NIST database, and those already reported in the literature [78]. In 2015, Lin and co-workers analyzed the bio-oil obtained by fast pyrolysis processing of four types of lignin. Acids, alcohols, aldehydes, ketones, furans, phenols, esters, and hydrocarbons were the main molecules detected. Peak identification was carried out using both the NIST mass spectral library and relevant literature. S-containing compounds were found in wheat straw lignin (one of four lignins analyzed), confirming the kraft pulping origin. The wider variation among the four lignins, in terms of concentration, was in the composition of phenolic products that represented the most prominent class compounds [79]. In 2011, Pyl and co-workers used a cryogenic GC×GC–TOFMS

system, first to analyze rapeseed oil methyl ester then, by a dedicated online sampling system (Py-GC×GC-TOFMS), to analyze the methyl ester oil subjected to a series of pyrolysis experiments. A 50 m × 0.25 mm × 0.5 μm non-polar as <sup>1</sup>D and a 2 m × 0.15 mm × 0.15 μm polar column as <sup>2</sup>D was employed in the GC system. Detailed online monitoring of thermal conversion processes was carried out, and approximately 200 different compounds derived from the FAMES conversion were identified [80]. Py-GC-MS is often used to analyze the non-volatile compounds with boiling point > 350°C present in the residue of biofuels. Pedersen and co-workers analyzed the distillation residue of a bio-crude by Py-GC×GC-qMS. Pyrolysis temperature was set to 600°C, and the chromatographic system was equipped with a cryogenic modulator and a non-polar × semi-polar columns set. Bio-crude was obtained by HTL of biomass, then a fractional distillation of the bio-crude was carried out in agreement with the ASTM D2892, obtaining six distillation fractions, excluding the residue. The fractions were catalytically hydro-treated and then analyzed by GC-qMS, while the residue fraction was subjected to Py-GC×GC-qMS analysis. The identified components of the residue fraction were similar to the compounds from distillation/hydro-treatment of bio-crude detected by GC-MS, suggesting that the cracking of the residue can contribute to increase the yield of the volatile fractions on a chemically similar basis. The analytes detected in the residue were mainly oxygenated aromatics with a small part of non-oxygenated. The authors suggested that the presence of non-oxygenated aromatics could be partially explained by the formation of oxygen-containing pyrolysis products such as CO<sub>2</sub> and H<sub>2</sub>O [81]. Recently, Py-GC×GC-MS methodology was used to investigate the pyrolysis behavior of guaiacol lignin (G-lignin) and its major decomposition mechanisms [82], and the catalytic performances of a selective conversion of lignocellulosic biomass into furan compounds using a bimetal-modified bio-based activated carbon [83]. A Py-GC-MS drawback is the limitation for obtaining quantitative information on pyrolysis products. Although it is possible to perform quantitative analysis by using an internal standard, this rather complex procedure deletes the most attractive aspects of Py-GC-MS, namely the minimal sample preparation. Van Erven and co-workers used for the first time a polymeric <sup>13</sup>C lignin as internal standard for lignin quantification by Py-GC-qMS in SIM mode using a semipolar column. Relative response factors (RRFs) for the various pyrolysis products obtained were determined and applied. The two most abundant fragments per compound were single ion monitored, with a maximum of eight fragments (4 <sup>12</sup>C + 4 <sup>13</sup>C) per segment and with a minimum of 25 data-points per peak. The determination of lignin by <sup>13</sup>C lignin internal standard and Py-GC-SIM-qMS was in good agreement with the total gravi-

metrically determined lignin contents of four common biomass sources (wheat straw, barley straw, corn stover, and sugar cane bagasse), showing high accuracy (>99.9%,  $R^2 > 0.999$ ) and precision (RSD < 1.5%) [84]. Figure 5 shows the workflow of the <sup>13</sup>C-lignin-internal standard/Py-GC-MS approach for the <sup>12</sup>C lignin quantification.

An exhaustive review of the analytical pyrolysis of biomass by GC-MS was reported by Alkalin and Karagoz. Various kind of biomass and biomass components (including carbohydrates, lignin, lignocellulosic biomass, and algal biomass), decomposition pathways, the formation of possible decomposition products, and the effects of catalysts in the degradation of biomass were discussed, together with dedicated sample preparation and instrumentation for analytical pyrolysis [85].

## 6 | MS DETECTION

### 6.1 | Ionization

Since electron impact ionization (EI) can be too energetic to preserve molecular ions and/or diagnostic fragments at the standard electro-kinetic energies of 70 eV, thus soft ionization techniques can be applied in GC-MS ionization to generate abundant molecular ion signals. Although not applied to the analysis of biofuels, worthy of mention are two works that highlight the potential of softer ionization using lower than 70 eV electro-kinetic energies and using conventional EI-QQQ-MS in single Q mode. Both discuss the benefit in identification capability on compounds that can be found in biofuel as well [86,87]. Beccaria and co-workers investigate the effect of using different electro-kinetic energies (20, 30, 50, and 70 eV) to analyze bacteria's FAMES. No loss in sensitivity and high repeatability were reported at lower electro-kinetic energies. The information obtained combining 20 and 70 eV, together with the use of LRI, was fundamental for the mass spectra interpretation and tentative identification of several unknowns FAMES and other oxygenated compounds. An example of the tentative identification of an unknown oxygenated compound, combining the information obtained at 20 and 70 eV in EI-MS together with the LRI, is shown in Figure 6. Considering the mass spectrum at 70 eV, no match with the available libraries was present considering a cut-off of 700/1000, while the ion-ratio present in the lighter part of the MS spectrum at 70 eV indicated the typical fragmentation of the saturated aldehydes. Combining this information with those obtained at 20 eV present in the higher  $m/z$  part of the spectrum, namely the molecular ion  $[M]^+$  at  $m/z$  380 (not present at 70 eV), and the related molecular fragment ions  $[M-H_2O]^+$  corresponding to  $[M-18]^+$  and  $[M-H_2O+CH_2=CH_2]^+$  corresponding to  $[M-46]^+$ ,

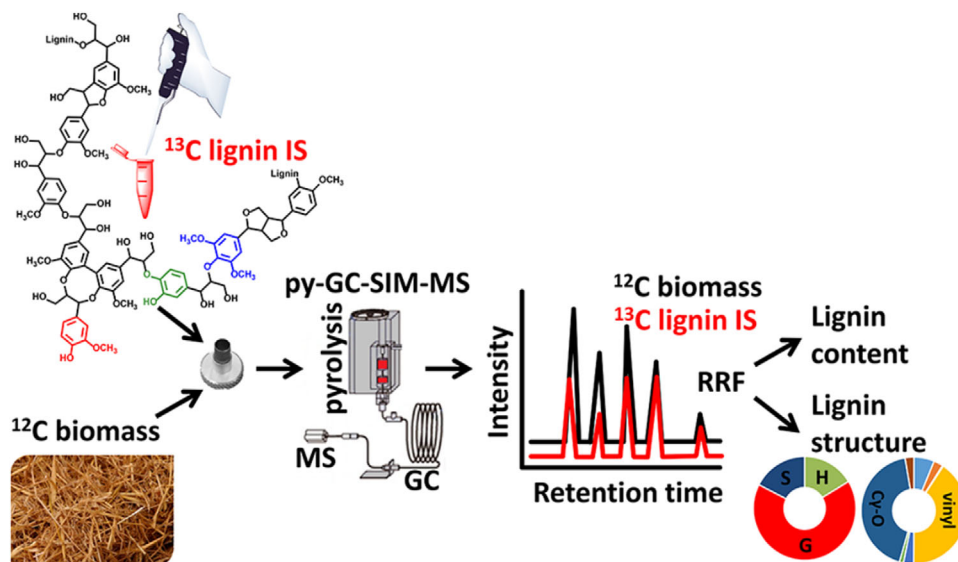


FIGURE 5 Workflow of the quantification of lignin-derived biomass using  $^{13}\text{C}$  lignin as internal standard (IS) and Py-GC-MS analysis. Reprinted with the permission of American Chemical Society, in van Erven et al. [84]

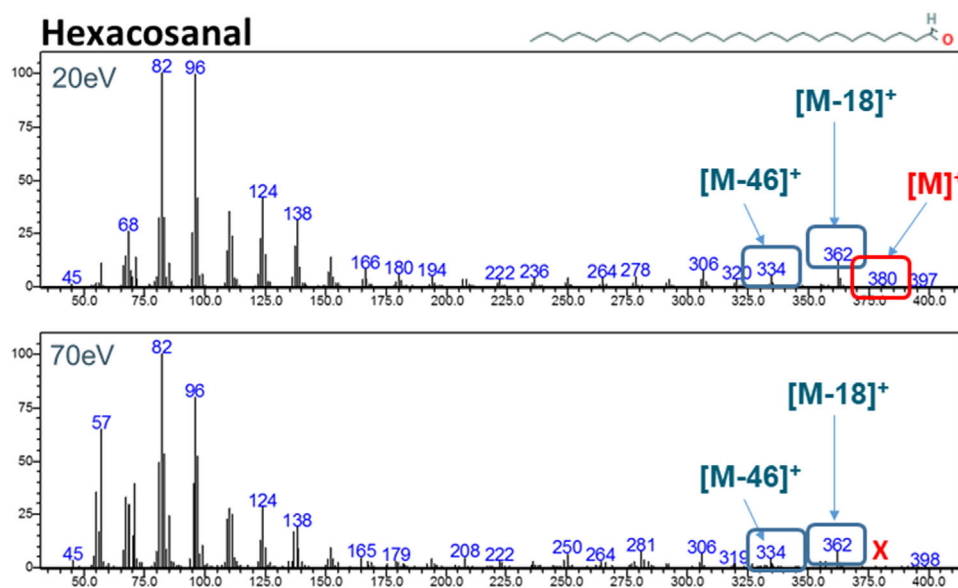


FIGURE 6 Tentative identification of an unknown oxygenated compound, finally identified as hexacosanal, combining the GC-EI-MS information obtained at 20 and 70 eV. Reprinted and adapted with the permission of Springer Nature, in Beccaria et al. [86]

the compound was identified as hexacosanal ( $\text{C}_{26}\text{H}_{52}\text{O}$ ). Furthermore, the experimental LRI corresponded with the data reported in the on-line NIST database [86]. Tranchida and co-workers studied the mass spectral behavior of a series of organic compounds belonging to different chemical classes (among which FAMES and alcohols), applying lower electro-kinetic energies and temperature interface by using a GC×GC-qMS [87].

Furuhashi and co-workers studied the fragmentation patterns of methyloxime-silyl derivatives of constitutive mono- and disaccharide standard isomers analyzed by GC coupled to field ionization (FI)-MS, also comparing the

GC-FI-MS spectra with GC-EI-MS at standard 70 and 18 eV, and GC-chemical ionization (CI)-MS. FI and CI modes generated abundant molecular ions and protonated ions, respectively. At low-voltage EI (18 eV), fewer fragments were observed compared to 70 eV EI, but no molecular ions were detected. Moreover, the authors claimed that the position of disaccharide linkages could be established based on the characteristic fragment ions visible in FI mode, distinguishing co-eluting isomeric disaccharides basing on their fragmentation patterns [88]. Although the coupling of GC with soft ionizations-MS has been reported in the 1970s, this interesting hyphenation was expanded

in the last decade, probably due to the progress in instrument technology. Different soft ionization techniques have been coupled to GC systems in the analysis of petrochemical samples; however, no record seems to be reported by the use of soft ionizations-MS coupled with GC for the analysis of biofuels and related products [89]. Advantages and drawbacks of the hyphenation between the most common atmospheric pressure ionization-MS interfaces and GC systems and new developments in source geometry, ion source miniaturization, multipurpose ion source constructions, and comparison with GC-FID and GC-EI-MS detections have been reviewed by Li and co-workers [90].

## 6.2 | Mass analyzer

The mass analyzer technologies mainly used in biofuels analysis are the Q and low-resolution time-of-flight (LR-TOF). The main limitations of Q compared to TOF technology are the lower data acquisition rate (dependent on the scan speed and interscan delay) and the skewing. These limitations have been for long the main limiting factors in the employment of QMS systems coupled with fast GC separation as GC×GC (where usually the peak width is about 200 ms) in quantitative applications until the beginning of the past decade, when Purcaro and co-workers reported for the first time an accurate full-scan quantification in GC×GC-QMS using a standard comprehensive 2D-GC setup. Due to the continuous improvement in QMS technology, Purcaro's work was the spark for the employment of QMS coupled with GC×GC for quantitative purposes [91]. Even if QMS technology has become a valid option for quantification in fast GC separation, TOFMS remains still necessary for an accurate peak reconstruction. In fact, a characteristic of LR-TOFMS is the high spectral production frequency. Recently, Schena and co-workers studied the effect of the MS acquisition range in qualitative biomass analysis by GC×GC-TOFMS, reporting that the use of a high acquisition rate (200-300 Hz) can increase up to 50% the number of peaks identified in qualitative analysis comparing with 100 Hz traditionally used in GC×GC-TOFMS. Coconut fiber bio-oil obtained by pyrolysis was analyzed by fast GC×GC-TOFMS at five different acquisition rates (30, 50, 100, 200, and 300 Hz). Figure 7 shows expansions of 3D (GC×GC-TOFMS) plots of coconut fiber bio-oil sample chromatograms, considering the effects of acquisition rates applied. In the 2D plots obtained at 300 and 200 Hz (Figures 7A and B), it was possible to differentiate each blob by its shape. On the other hand, at 30 Hz (Figure 7E) some co-elution was present [92].

Kloekhorst and co-workers used a GC×GC-HR-TOFMS system to characterize the compounds present in the lignin oil after catalytic hydro-deoxygenation and hydrocracking

processing [93]. Recently, Hung and co-workers employed an Orbitrap Q-Exactive HR-EI-MS operating at 70 eV coupled with a GC×GC system, equipped with the reverse-set columns, for the analysis of volatile and semi-volatile compounds in biofuel samples. Compounds were identified and quantified using an MS resolving power (RP) of 7.500 with a scan rate of 43 Hz, reaching >10 data points per peak for quantification purpose, with an average mass accuracy of about 1 ppm. Before the real sample analysis, a calibration curve for  $\alpha$ -pinene was successfully established to evaluate the sensitivity and the linear dynamic range. The data acquisition rate represents the main critical point of Orbitrap technology compared with TOF analyzer in GC×GC applications. At a mass RP of 15.000, the frequency average is about 25 Hz, which may not be enough for a proper reconstruction of the narrowest chromatographic peaks. Another limitation is the need to use Orbitrap libraries for the identification since the NIST libraries can lead to some misannotations [94].

## 7 | DATA HANDLING

The complexity of GC-MS data, whether it is second-order data from a 1D-GC-MS or third-order data from a 2D-GC-MS, may require advanced chemometric tools to provide reliable information for class-type classification, pattern recognition, clustering, as well as multivariate calibration, curve resolution, and more. Considering the aims, many statistical approaches can be applied, and each one can include different processing techniques. Statistical techniques can be grouped in: (i) unsupervised and (ii) supervised. (i) Unsupervised approaches usually explore the entire matrix, finding trends and grouping within the dataset (principal component analysis (PCA), and hierarchical clustering analysis (HCA) are the most applied), while (ii) supervised methods are generally used to predict new data, basing on prior knowledge on the data structure. They can be sub-classified as linear (e.g., partial least squares discriminant analysis (PLS-DA)) and non-linear (e.g., random forest) supervised method. Moreover, before applying unsupervised and supervised processing techniques, raw data must be preprocessed to generate a clean data matrix eliminating as much as possible the variance and bias, reducing the complexity, and enhancing significant signals. Different preprocessing techniques, such as noise filtering and baseline correction, peak detection and deconvolution, alignment, normalization, and scaling, are the most commonly applied to handle MS data. A detailed review of advance chemometric methods applied in MS data processing was published by Yi and co-workers in 2016 [95]. An example of the application of unsupervised chemometric techniques for the GC-MS analysis of biodiesel is reported by Flood and co-workers. The

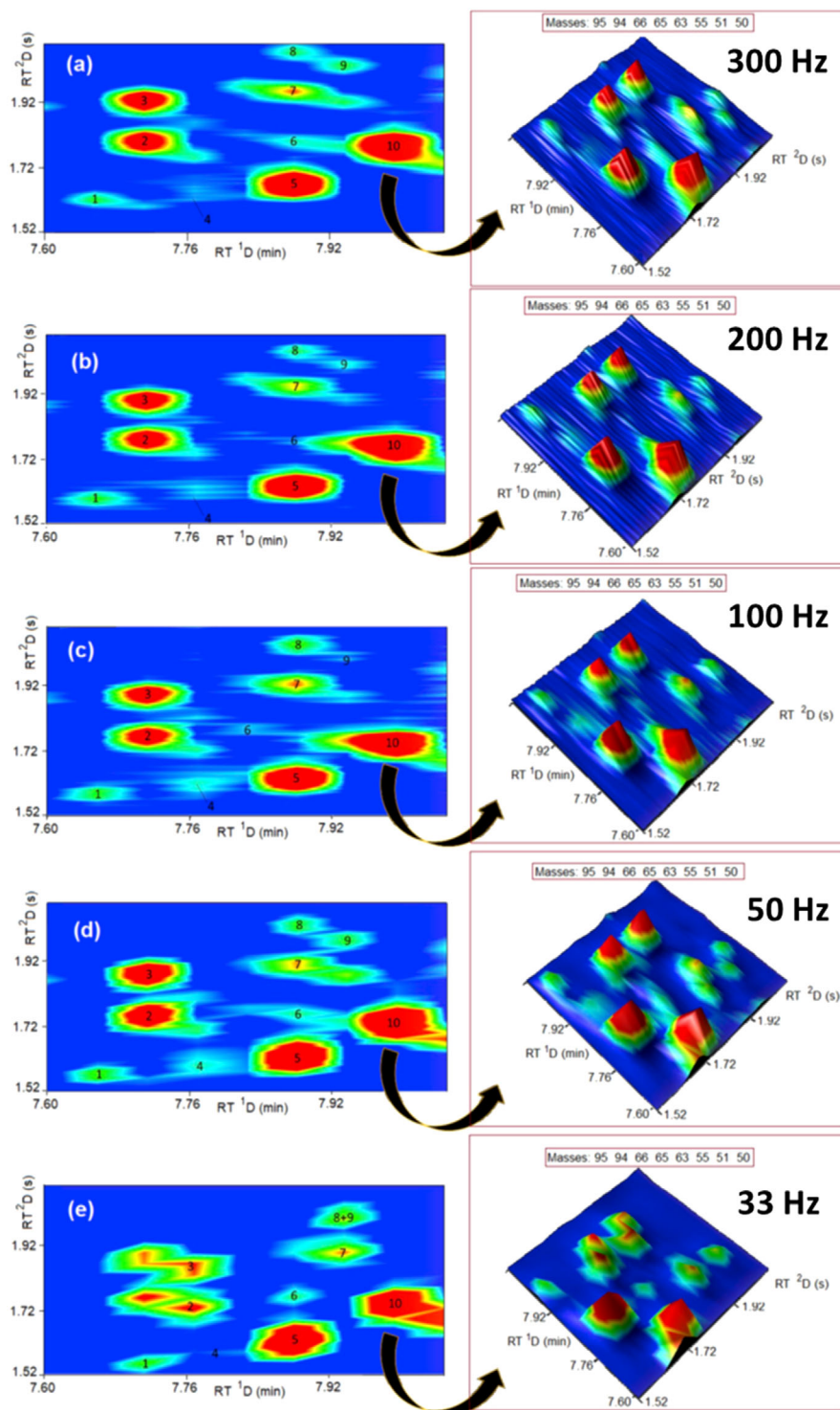


FIGURE 7 Zoomed-in 2D GCxGC-TOFMS plots obtained using acquisition rates of (A) 300 Hz, (B) 200 Hz, (C) 100 Hz, (D) 50 Hz, and (E) 30 Hz. Peak 1: unknown; peak 2: methyl benzenediol; peak 3: methoxy benzenediol; peak 4: dimethyl indene; peak 5: methoxy-ethyl phenol; peak 6- dihydroxyacetophenone; peak 7: benzenediol; peaks 8 and 9: dihydro-indenone isomers; peak 10: methyl benzenediol. Reprinted and adapted with the permission of Elsevier, in Schena et al [92]

authors determined the FAMEs profile present in different biodiesel feedstocks by GC–MS and then investigated the importance of several chromatographic parameters, such as temperature program and column polarity, using PCA and HCA clustering. FAMEs present in biodiesel samples were clustered or linked based on feedstock type regardless of temperature program or column type [96]. Madson and co-workers applied chemometrics tools after an extensive characterization of the bio-crude and aqueous phase from HTL of thermally and chemically pretreated lignocellulosic biomass (*Miscanthus x giganteus*) by GC–MS using pre-derivatization with silylating reagent and methyl chloroformate. A cutoff of 700/1000 was selected as MS similarity. Pre-processing techniques, such as baseline correction, peak alignment, and normalization, as well as mean-centering and auto-scaling, were applied to the dataset. PCA was used to visualize the two datasets from the bio-crude and aqueous phase [97]. However, unsupervised techniques may not be enough to highlight small compositional differences in complex datasets, requiring the use of more advanced chemometric tools, like the supervised techniques, extensively used for examples in pattern recognition, data prediction, and feature selection. An example is reported by Mustafa and co-workers. After the identification of biodiesel FAMEs profile from different plant sources by GC–MS, the authors used the dataset in a pattern-recognition approach applying PLS-DA for the prediction of the class membership of several samples of unknown origin. The statistical model was developed in two different steps: (i) creation of a model class similarity by using the FAMEs dataset of known samples, and (ii) creation of a model classification to classify samples of unknown origin based on the model created in the first step. The results showed that all samples with a previously unknown origin were successfully classified. Authors claimed that, for the first time, a pattern recognition approach was successfully applied in the field of biodiesel classification and modeling tasks by GC–MS [98]. Supervised chemometric tools have also been used to predict the feedstock origin and percent composition of blends of biodiesel and conventional diesel basing on the FAMEs profile by 1D-GC-MS and 2D-GC-MS, as well as in Py-GC-MS analysis of bio-engineered field-grown biomass for the production of high value-added phenolic compounds [99–101]. An interesting approach to highlight differences among samples was proposed by Barcaru and Vivó-Truyols in 2015. The authors developed a statistical method able to extract meaningful differences between two GC×GC–MS analyses, in order to highlight the differences between two samples, to flag differences in composition, or to spot compounds only present in one of the samples. The method is based on the application of the Jensen–Shannon divergence analysis combined with Bayesian hypothesis testing.

The proposed method was applied to spot contamination diesel samples analyzed by GC×GC-MS. It is the opinion of the authors that the proposed method can be successfully applied in different fields. The only requirement is a reasonable resolution [102].

Within advanced chemometrics, machine learning techniques are powerful tools to handle high-dimensional data [103]. However, particular attention has to be devoted to the validation of the statistical method employed to avoid overfitted results, in particular in the prediction models and feature selection by using supervised techniques [104,105]. An optimal validated model usually comes from three sub-datasets, namely, training (to build the model), validation, and test set (to access the robustness). The use of a single data set usually generates an overfitted statistical model. The first one feeds the candidate algorithms during the learning process. The validation set compares the performance of the trained algorithms in terms of accuracy, specificity, and sensitivity, selecting the best ones. The test set is an independent dataset but having in the same probability distribution. The test set determines how the build model fits in terms of accuracy, specificity, and sensitivity on an independent dataset. In general, there is minimal overfitting if a model fits well in the same way in validation and test sets. On the contrary, a better fitting in the validation than in the test set indicates overfitting. Extensive reviews have been published on validation for computational methods with MS data [106,107]. Figure 8 reports a general schematic workflow of the chemometrics methodology for biofuels analysis from the acquired GC–MS data to the application of preprocessing and processing chemometric techniques to the dataset(s).

It is worth mentioning that this generic data handling approach is more common where a considerable amount of samples is analyzed, like in genomics, proteomics, and metabolomics field. In biofuels analysis, only part of this workflow is usually applied.

## 8 | CONCLUSION AND FUTURE PERSPECTIVE

This review describes the main critical aspects and the most advanced applications of GC–MS-based techniques in the analysis of oxygen-containing compounds in biomasses and biofuels, mainly focusing on the period 2011–early 2020. Readers can have an overview of what has been done over the last decade with a look at what it is possible to do using GC–MS based techniques in biomass and biofuels analysis. In this context, the problem of the presence of oxygen compounds in processing biomasses, the most common sample preparation techniques applied, the use of analytical pyrolysis coupled to 1D-GC–MS and 2D-GC–MS, the advantage of 2D-GC separation (both hurt-



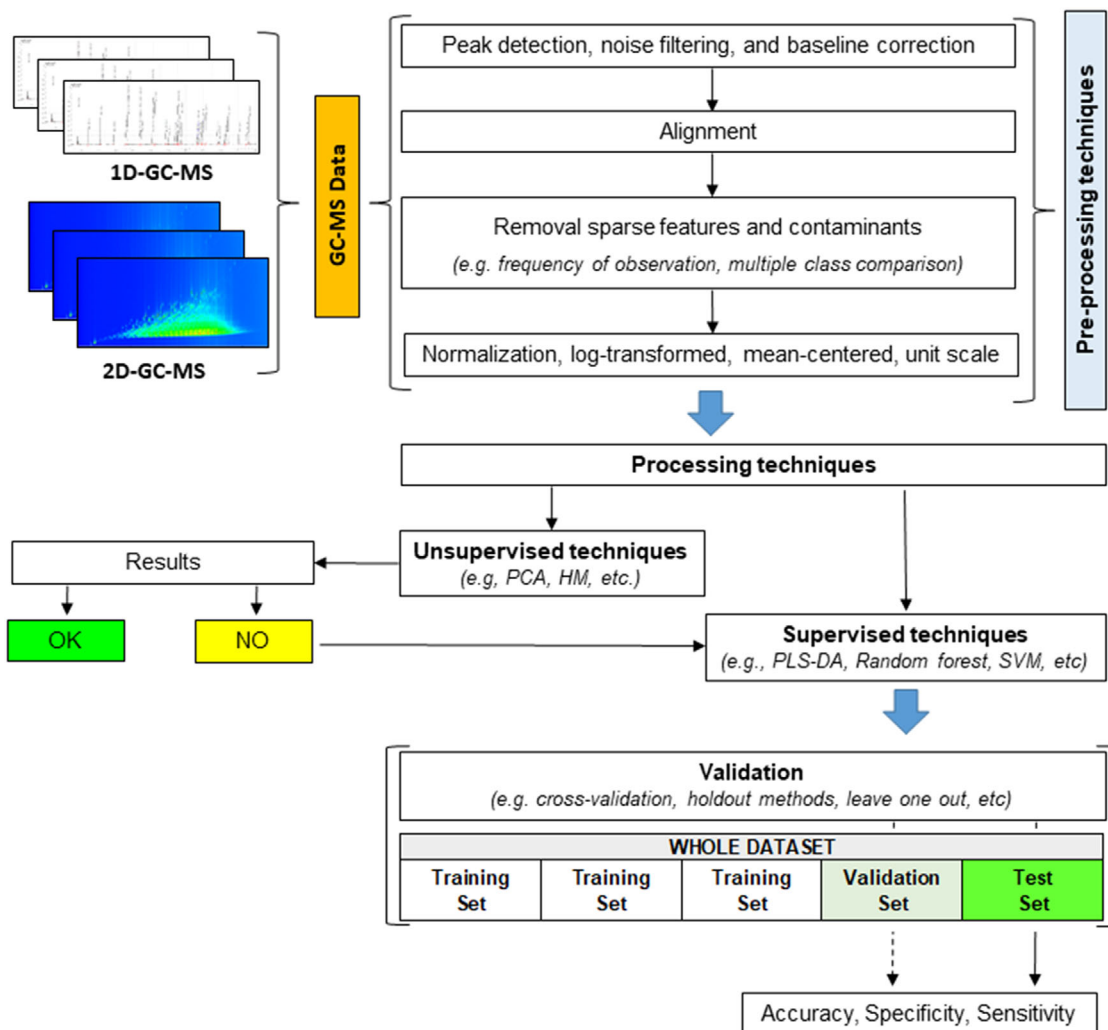


FIGURE 8 A general schematic workflow of the chemometrics methodology from the acquired GC-MS data to the application of pre-processing and processing chemometric techniques to the dataset

cutting and comprehensive) coupled to MS, as well as the most common MS ionization and technology applied, the advanced chemometric tools in data handling, and the future trends in biomass processing are reported.

As already mentioned previously, since the EI-MS spectra information at standard 70 eV could not be enough to distinguish between isomers and/or homolog series, the use of softer and/or soft MS ionization interfaces coupled with GC is taking more interest to the researchers in the petrochemical field. Thanks to the continuous improvement in MS instrument and technology, lower electro-kinetic energies in EI-MS, as well as the use of soft ionization techniques, can be crucial for a successful determination of oxygenated compounds obtained in thermochemical biomass processing. In addition, a complete retention index database of phenols, carbohydrates, and lipids can be considered as an additional filter for the identification of unknown compounds. Moreover, it has been recently shown that the use of Orbitrap technology

coupled to both 1D- and 2D-GC systems are becoming a promising tool for the characterization of complex samples as bio-fuels and related samples.

Considering the complexity of samples obtained from thermochemical biomass processing, an extra-dimension separation coupling an LC system to 1D- or 2D-GC-MS can generate simplified sub-samples, focalizing the region/compounds of interest and, considering the amount of non-volatile material may be present, also reducing problems related to matrix components. This LC-GC hyphenation is possible due to specific evaporative interfaces that, over the years, have been modified and implemented. This LC-GC-MS methodology has been applied in the past for the analysis of oxygenated compounds in edible and essential oils, process waste stream, urban air particulate extract, and petrochemical samples, but no applications seem to be reported in bio-oils/fuels analysis [108]. It is the opinion of the authors that LC-GC could be a valid option in the field, although such systems


are still rather complex and need specialized and trained people

## CONFLICT OF INTEREST

The authors have declared no conflict of interest.

## ORCID

Marco Beccaria  <https://orcid.org/0000-0001-6539-5531>

Adrien Maniquet  <https://orcid.org/0000-0001-8002-3568>

Pierre Giusti  <https://orcid.org/0000-0002-9569-3158>

Marco Piparo  <https://orcid.org/0000-0003-4404-442X>

Pierre-Hugues Stefanuto  <https://orcid.org/0000-0002-1224-8869>

Jean-François Focant  <https://orcid.org/0000-0001-8075-2920>

## REFERENCES

- Kohli K, Prajapati R, Sharma B. Bio-based chemicals from renewable biomass for integrated biorefineries. *Energies*. 2019;12:233.
- Roddy DJ. Biomass in a petrochemical world. *Interface Focus*. 2013;3:20120038.
- Takkellapati S, Li T, Gonzalez MA. An overview of biorefinery-derived platform chemicals from a cellulose and hemicellulose biorefinery. *Clean Technol Environ Policy*. 2018;20:1615–30.
- Kanaujia PK, Sharma YK, Garg MO, Tripathi D, Singh R. Review of analytical strategies in the production and upgrading of bio-oils derived from lignocellulosic biomass. *J Anal Appl Pyrolysis*. 2013;105:55–74.
- Castello D, Haider MS, Rosendahl LA. Catalytic upgrading of hydrothermal liquefaction biocrudes: different challenges for different feedstocks. *Renew Energy*. 2019;141:420–30.
- Elkasabi Y, Chagas BME, Mullen CA, Boateng AA. Hydrocarbons from spirulina pyrolysis bio-oil using one-step hydrotreating and aqueous extraction of heteroatom compounds. *Energy Fuels*. 2016;30:4925–32.
- Veses A, Martínez J, Callén M, Murillo R, García T. Application of upgraded drop-in fuel obtained from biomass pyrolysis in a spark ignition engine. *Energies*. 2020;13:1–15.
- Pereira G, Macedo J, Bispo MD, Krause LC, Jacques RA, Zini CA, Caramao BE. Pyrolysis - Comprehensive Two-Dimensional Gas Chromatography and Its Application to the Investigation of Pyrolytic Liquids. InTech, London, UK 2017.
- Michailof CM, Kalogiannis KG, Sfetsas T, Patiaka DT, Lappas AA. Advanced analytical techniques for bio-oil characterization. *Wiley Interdiscip Rev Energy Environ*. 2016;5: 614–39.
- Seeley JV, Seeley SK. Multidimensional gas chromatography: Fundamental advances and new applications. *Anal Chem*. 2013;85:557–78.
- Melero JA, Iglesias J, Garcia A. Biomass as renewable feedstock in standard refinery units. Feasibility, opportunities and challenges. *Energy Environ Sci*. 2012;5:7393–420.
- Treichel H, Fongaro G, Scapini T, Frumi Camargo A, Spitza Stefanski F, Venturin B. Utilising Biomass in Biotechnology. Springer Nature, Cham, Switzerland 2020.
- Callegari A, Bolognesi S, Cecconet D, Capodaglio AG. Production technologies, current role, and future prospects of biofuels feedstocks: A state-of-the-art review. *Crit Rev Environ Sci Technol*. 2020;50:384–436.
- Adams P, Bridgwater T, Lea-Langton A, Ross A, Watson I, Chapter 8: Biomass Conversion Technologies. In: Thornley P, Adams P, editors. *Greenhouse Gas Balances of Bioenergy Systems*. 1st Edition. Academic Press, Elsevier Inc. 2018, pp. 107–39.
- Nagappan S, Bhosale RR, Nguyen DD, Chi NTL, Ponnusamy VK, Woong CS, Kumar G. Catalytic hydrothermal liquefaction of biomass into bio-oils and other value-added products – a review. *Fuel*. 2021;285:119053.
- Beccaria M, Oteri M, Micalizzi G, Bonaccorsi IL, Purcaro G, Dugo P, Mondello L. Reuse of dairy product: evaluation of the lipid profile evolution during and after their shelf-life. *Food Anal Methods*. 2016;9:3143–54.
- Costa R, Beccaria M, Grasso E, Albergamo A, Oteri M, Dugo P, Fasulo S, Mondello L. Sample preparation techniques coupled to advanced chromatographic methods for marine organisms investigation. *Anal Chim Acta*. 2015;875:41–53.
- Wang X, Liu Si-F, Qin Zi-H, Balamurugan S, Li H-Ye, Lin CSKi. Sustainable and stepwise waste-based utilisation strategy for the production of biomass and biofuels by engineered microalgae. *Environ Pollut*. 2020;265:114854.
- Maza DD, Viñarta SC, Su Y, Guillamón JM, Aybar MJ. Growth and lipid production of *Rhodotorula glutinis* R4, in comparison to other oleaginous yeasts. *J Biotechnol*. 2020;310:21–31.
- Wang Y, Han Y, Hu W, Fu D, Wang G. Analytical strategies for chemical characterization of bio-oil. *J Sep Sci*. 2020;43:360–71.
- Arif M, Bai Y, Usman M, Jalalah M, Harraz FA, Al-Assiri MS, Li X, Salama El-S, Zhang C. Highest accumulated microalgal lipids (polar and non-polar) for biodiesel production with advanced wastewater treatment: role of lipidomics. *Bioresour Technol*. 2020;298:122299.
- Zhu Y, Albrecht KO, Elliott DC, Hallen RT, Jones SB. Development of hydrothermal liquefaction and upgrading technologies for lipid-extracted algae conversion to liquid fuels. *Algal Res*. 2013;2:455–64.
- Chaudry S, Bahri PA, Moheimani NR. Pathways of processing of wet microalgae for liquid fuel production: a critical review. *Renew Sustain Energy Rev*. 2015;52:1240–50.
- Li-beisson Y, Nakamura Y, Harwood J. Lipids in plant and algae development. Chapter 1. Lipids: From Chemical Structures, Biosynthesis, and Analyses to Industrial Applications. In: Nakamura Y, Li-Beisson, Y, editors. *Lipids in Plant and Algae Development*. Subcellular Biochemistry 86. Springer International Publishing Switzerland 2016, pp. 1–18.
- Viswanathan MB, Park K, Cheng M-H, Cahoon EB, Dweikat I, Clemente T, Singh V. Variability in structural carbohydrates, lipid composition, and cellulosic sugar production from industrial hemp varieties. *Ind Crops Prod*. 2020;157:112906.
- Kim SKI, Han JY, Lee H-S, Yum T, Kim Y, Kim J. Production of renewable diesel via catalytic deoxygenation of natural triglycerides: Comprehensive understanding of reaction intermediates and hydrocarbons. *Appl Energy*. 2014;116:199–205.
- Negi S, Jaswal G, Dass K, Mazumder K, Elumalai S, Roy JK. Torrefaction: a sustainable method for transforming of agri-wastes to high energy density solids (biocoal). *Rev Environ Sci Biotechnol*. 2020;19:463–88.

28. Lyu G, Wu S, Zhang H. Estimation and comparison of bio-oil components from different pyrolysis conditions. *Front Energy Res.* 2015;3:1–11.
29. Talmadge MS, Baldwin RM, Bidy MJ, McCormick RL, Beckham GT, Ferguson GA, Czernik S, Magrini-Bair KA, Foust TD, Metelski PD, Hetrick C, Nimlos MR. A perspective on oxygenated species in the refinery integration of pyrolysis oil. *Green Chem.* 2014;16:407–53.
30. Abdelnur PV, Vaz BG, Rocha JD, De Almeida MBB, Teixeira MAG, Pereira RCL. Characterization of bio-oils from different pyrolysis process steps and biomass using high-resolution mass spectrometry. *Energy and Fuels.* 2013;27:6646–54.
31. Oasmaa A, Meier D. Norms and standards for fast pyrolysis liquids: 1. Round robin test. *J Anal Appl Pyrolysis.* 2005;73:323–34.
32. Sikarwar VS, Zhao M, Fennell PS, Shah N, Anthony EJ. Progress in biofuel production from gasification. *Prog Energy Combust Sci.* 2017;61:189–248.
33. Chan YiH, Loh SK, Chin BLF, Yiin CL, How BS, Cheah KW, Wong MK, Loy ACM, Gwee YL, Lo SLY, Yusup S, Lam SuS. Fractionation and extraction of bio-oil for production of greener fuel and value-added chemicals: recent advances and future prospects. *Chem Eng J.* 2020;397:125406.
34. Maqbool W, Dunn K, Doherty W, Mckenzie N, Cronin D, Hobson P. Extraction and purification of renewable chemicals from hydrothermal liquefaction bio-oil using supercritical carbon dioxide: a techno-economic evaluation. *Ind Eng Chem Res.* 2019;58:5202–14.
35. Dugheri S, Mucci N, Bonari A, Marrubini G, Cappelli G, Ubiali D, Campagna M, Montalti M, Arcangeli G. Solid phase microextraction techniques used for gas chromatography: a review. *Acta Chromatogr.* 2019;32:1–9.
36. Jalili V, Barkhordari A, Ghiasvand A. A comprehensive look at solid-phase microextraction technique: A review of reviews. *Microchem J.* 2020;152:104319.
37. Tessini C, Müller N, Mardones C, Meier D, Berg A, Von Baer D. Chromatographic approaches for determination of low-molecular mass aldehydes in bio-oil. *J Chromatogr A.* 2012;1219:154–60.
38. Conti R, Fabbri D, Torri C, Hornung A. At-line characterisation of compounds evolved during biomass pyrolysis by solid-phase microextraction SPME-GC-MS. *Microchem J.* 2016;124:36–44.
39. Wei Yi, Lei H, Wang Lu, Zhu L, Zhang X, Liu Y, Chen S, Ahring B. Liquid-liquid extraction of biomass pyrolysis bio-oil. *Energy Fuels.* 2014;28:1207–12.
40. Ren S, Ye XP, Borole AP. Separation of chemical groups from bio-oil water-extract via sequential organic solvent extraction. *J Anal Appl Pyrolysis.* 2017;123:30–9.
41. Cheng T, Han Y, Zhang Y, Xu C. Molecular composition of oxygenated compounds in fast pyrolysis bio-oil and its supercritical fluid extracts. *Fuel.* 2016;172:49–57.
42. Christensen ED, Chupka GM, Luecke J, Smurthwaite T, Alleman TL, Iisa K, Franz JA, Elliott DC, McCormick RL. Analysis of oxygenated compounds in hydrotreated biomass fast pyrolysis oil distillate fractions. *Energy Fuels.* 2011;25:5462–71.
43. Zeng F, Liu W, Jiang H, Yu H-Q, Zeng RJ, Guo Q. Separation of phthalate esters from bio-oil derived from rice husk by a basification-acidification process and column chromatography. *Bioresour Technol.* 2011;102:1982–7.
44. Charon N, Ponthus J, Espinat D, Broust F, Volle G, Valette J, Meier D. Multi-technique characterization of fast pyrolysis oils. *J Anal Appl Pyrolysis.* 2015;116:18–26.
45. Lu Y, Li G-S, Lu Y-C, Fan X, Wei X-Y. Analytical strategies involved in the detailed componential characterization of biooil produced from lignocellulosic biomass. *Int J Anal Chem.* 2017;2017:1.
46. Joseph J, Rasmussen MJ, Fecteau JP, Kim S, Lee H, Tracy KA, Jensen BL, Frederick BG, Stemmler EA. Compositional changes to low water content bio-oils during aging: an NMR, GC/MS, and LC/MS study. *Energy Fuels.* 2016;30:4825–40.
47. Ruiz-Matute AI, Hernández-Hernández O, Rodríguez-Sánchez S, Sanz ML, Martínez-Castro I. Derivatization of carbohydrates for GC and GC-MS analyses. *J Chromatogr B Anal Technol Biomed Life Sci.* 2011;879:1226–40.
48. Madsen RB, Jensen MM, Mørup AJ, Houlberg K, Christensen PS, Klemmer M, Becker J, Iversen BoB, Glasius M. Using design of experiments to optimize derivatization with methyl chloroformate for quantitative analysis of the aqueous phase from hydrothermal liquefaction of biomass. *Anal Bioanal Chem.* 2016;408:2171–83.
49. Schulzke T, Conrad S, Shumeiko B, Auersvald M, Kubička D, Raymakers LFJM. Fuels from reliable bio-based refinery intermediates: BioMates. *Waste Biomass Valoriz.* 2020;11:579–98.
50. Hu X, Gunawan R, Mourant D, Hasan MDM, Wu L, Song Y, Lievens C, Li C-Z. Upgrading of bio-oil via acid-catalyzed reactions in alcohols — A mini review. *Fuel Process Technol.* 2017;155:2–19.
51. Dayton DC, Foust TD. Analytical Methods in Thermochemical Conversion. *Anal Methods Biomass Charact Convers.* 2020;75–88.
52. David F, in: Chapter 2.1: Classical two-dimensional GC combined with mass spectrometry; Tranchida, P., Mondello, L., (Eds.), *Hyphenations of Capillary Chromatography with Mass Spectrometry.* Elsevier, Amsterdam, the Netherlands 2020. pp. 135–82.
53. Fullana A, Contreras JA, Striebich RC, Sidhu SS. Multidimensional GC/MS analysis of pyrolytic oils. *J Anal Appl Pyrolysis.* 2005;74:315–26.
54. Boeker P, Leppert J, Mysliwicz B, Lammers PS. Comprehensive theory of the deans' switch as a variable flow splitter: Fluid mechanics, mass balance, and system behavior. *Anal Chem.* 2013;85:9021–30.
55. Le Brech Y, Jia L, Cissé S, Mauviel G, Brosse N, Dufour A. Mechanisms of biomass pyrolysis studied by combining a fixed bed reactor with advanced gas analysis. *J Anal Appl Pyrolysis.* 2016;117:334–46.
56. Waktola HD, Zeng AXu, Chin S-T, Marriott PJ. Advanced gas chromatography and mass spectrometry technologies for fatty acids and triacylglycerols analysis. *TrAC - Trends Anal Chem.* 2020;129:115957.
57. Djokic MR, Dijkmans T, Yildiz G, Prins W, Van Geem KM. Quantitative analysis of crude and stabilized bio-oils by comprehensive two-dimensional gas-chromatography. *J Chromatogr A.* 2012;1257:131–40.
58. Van Deursen M, Beens J, Reijnga J, Lipman P, Cramers C, Blomberg J. Group-type identification of oil samples using comprehensive two-dimensional gas chromatography coupled

- to a time-of-flight mass spectrometer. *J High Resol Chromatogr.* 2000;23:507–10.
59. Tessarolo NS, Dos Santos LRM, Silva RSF, Azevedo DA. Chemical characterization of bio-oils using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry. *J Chromatogr A.* 2013;1279:68–75.
  60. Tessarolo NS, Silva RVS, Vanini G, Casilli A, Ximenes VL, Mendes FL, De Rezende Pinho A, Romão W, De Castro EVR, Kaiser CR, Azevedo DA. Characterization of thermal and catalytic pyrolysis bio-oils by high-resolution techniques: <sup>1</sup>H NMR, GC × GC-TOFMS and FT-ICR MS. *J Anal Appl Pyrolysis.* 2016;117:257–67.
  61. Joffres B, Lorentz C, Vidalie M, Laurenti D, Quoineaud A-A, Charon N, Daudin A, Quignard A, Geantet C. Catalytic hydroconversion of a wheat straw soda lignin: Characterization of the products and the lignin residue. *Appl Catal B Environ.* 2014;145:167–76.
  62. Silva RVS, Tessarolo NS, Pereira VB, Ximenes VL, Mendes FL, De Almeida MBB, Azevedo DA. Quantification of real thermal, catalytic, and hydrodeoxygenated bio-oils via comprehensive two-dimensional gas chromatography with mass spectrometry. *Talanta.* 2017;164:626–35.
  63. Sajdak M, Chrubasik M, Muzyka R. Chemical characterisation of tars from the thermal conversion of biomass by 1D and 2D gas chromatography combined with silylation. *J Anal Appl Pyrolysis.* 2017;124:426–38.
  64. Madsen RB, Zhang H, Biller P, Goldstein AH, Glasius M. Characterizing Semivolatile Organic Compounds of Biocrude from Hydrothermal Liquefaction of Biomass Energy and Fuels. 2017;31:4122–34.
  65. Hilaire F, Basset E, Bayard R, Gallardo M, Thiebaut D, Vial J. Comprehensive two-dimensional gas chromatography for biogas and biomethane analysis. *J Chromatogr A.* 2017;1524:222–32.
  66. Primaz CT, Schena T, Lazzari E, Caramão EB, Jacques RA. Influence of the temperature in the yield and composition of the bio-oil from the pyrolysis of spent coffee grounds: Characterization by comprehensive two dimensional gas chromatography. *Fuel.* 2018;232:572–80.
  67. Polidoro ADS, Scapin E, Lazzari E, Silva AN, Dos Santos AL, Caramão EB, Jacques RA. Valorization of coffee silverskin industrial waste by pyrolysis: From optimization of bio-oil production to chemical characterization by GC × GC/qMS. *J Anal Appl Pyrolysis.* 2018;129:43–52.
  68. Schena T, Farrapeira R, Bjerck TR, Krause LC, Mühlen C, Caramão EB. Fast two-dimensional gas chromatography applied in the characterization of bio-oil from the pyrolysis of coconut fibers. *Sep Sci Plus.* 2019;2:89–99.
  69. Wang Y, Yin R, Chai M, Nishu, Li C, Sarker M, Ronghou Liu. Comparative study by GC × GC-TOFMS on the composition of crude and composite-additives bio-oil before and after accelerated aging treatment. *J Energy Inst.* 2020;93:2163–75.
  70. Lazzari E, Polidoro ADS, Onorevoli B, Schena T, Silva AN, Scapin E, Jacques RA, Caramão EB. Production of rice husk bio-oil and comprehensive characterization (qualitative and quantitative) by HPLC/PDA and GC × GC/qMS. *Renew Energy.* 2019;135:554–65.
  71. Nunes VO, Silva RVS, Romeiro GA, Azevedo DA. The speciation of the organic compounds of slow pyrolysis bio-oils from Brazilian tropical seed cake fruits using high-resolution techniques: GC × GC-TOFMS and ESI(±)-Orbitrap HRMS. *Microchem J.* 2020;153:104514.
  72. Kalogiannis KG, Stefanidis SD, Michailof CM, Lappas AA. Castor bean cake residues upgrading towards high added value products via fast catalytic pyrolysis. *Biomass and Bioenergy.* 2016;95:405–15.
  73. Gallacher C, Thomas R, Taylor C, Lord R, Kalin RM. Comprehensive composition of Creosote using comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS). *Chemosphere.* 2017;178:34–41.
  74. Rathsack P, Wollmerstaedt H, Kuchling T, Kureti S. Analysis of hydrogenation products of biocrude obtained from hydrothermally liquefied algal biomass by comprehensive gas chromatography mass spectrometry (GC × GC-MS). *Fuel.* 2019;248:178–88.
  75. Purcaro G, Tranchida PQ, Mondello L. Chapter 11 - Comprehensive gas chromatography methodologies for the analysis of lipids. In Tranchida PQ, Mondello L, editors. *Handbook of Advanced Chromatography/Mass Spectrometry Techniques.* Elsevier Inc., Amsterdam, the Netherlands 2017, pp. 407–44.
  76. Onorevoli B, Machado ME, Dariva C, Franceschi E, Krause LC, Jacques RA, Caramão EB. A one-dimensional and comprehensive two-dimensional gas chromatography study of the oil and the bio-oil of the residual cakes from the seeds of *Crambe abyssinica*. *Ind Crops Prod.* 2014;52:8–16.
  77. Zhang T, Wang Z, Wang Y, Wei Z, Li X, Hou X, Sun Z, Wang G, Qian Yu. The characteristics of free/bound biomarkers released from source rock shown by stepwise Py-GC-MS and thermogravimetric analysis (TGA/DTG). *J Pet Sci Eng.* 2019;179:526–38.
  78. Brebu M, Tamminen T, Spiridon I. Thermal degradation of various lignins by TG-MS/FTIR and Py-GC-MS. *J Anal Appl Pyrolysis.* 2013;104:531–9.
  79. Lin X, Sui S, Tan S, Pittman C, Sun J, Zhang Z. Fast pyrolysis of four lignins from different isolation processes using Py-GC/MS. *Energies.* 2015;8:5107–21.
  80. Pyl SP, Schietekat CM, Van Geem KM, Reyniers M-F, Vercammen J, Beens J, Marin GB. Rapeseed oil methyl ester pyrolysis: on-line product analysis using comprehensive two-dimensional gas chromatography. *J Chromatogr A.* 2011;1218:3217–23.
  81. Pedersen TH, Jensen CU, Sandström L, Rosendahl LA. Full characterization of compounds obtained from fractional distillation and upgrading of a HTL biocrude. *Appl Energy.* 2017;202:408–19.
  82. Wang L, Li J, Chen Y, Yang H, Shao J, Zhang X, Yu H, Chen H. Investigation of the pyrolysis characteristics of guaiacol lignin using combined Py-GC × GC/TOF-MS and in-situ FTIR. *Fuel.* 2019;251:496–505.
  83. Lu Yi, Zheng Y, He R, Li W, Zheng Z. Selective conversion of lignocellulosic biomass into furan compounds using bimetal-modified bio-based activated carbon: Analytical Py-GC × GC/MS. *J Energy Inst.* 2020;93:2371-80 in Press.
  84. Van Erven G, De Visser R, Merckx DWH, Strolenberg W, De Gijssel P, Gruppen H, Kabel MA. Quantification of lignin and its structural features in plant biomass using <sup>13</sup>C lignin as internal standard for pyrolysis-GC-SIM-MS. *Anal Chem.* 2017;89:10907–16.

85. Akalın MK, Karagöz S. Analytical pyrolysis of biomass using gas chromatography coupled to mass spectrometry. *TrAC - Trends Anal Chem.* 2014;61:11–6.
86. Beccaria M, Franchina FA, Nasir M, Mellors T, Hill JE, Purcaro G. Investigation of mycobacteria fatty acid profile using different ionization energies in GC–MS. *Anal Bioanal Chem.* 2018;410:7987–96.
87. Tranchida PQ, Aloisi I, Giocastro B, Zoccali M, Mondello L. Comprehensive two-dimensional gas chromatography-mass spectrometry using milder electron ionization conditions: A preliminary evaluation. *J Chromatogr A.* 2019;1589:134–40.
88. Furuhashi T, Ishii K, Tanaka K, Weckwerth W, Nakamura T. Fragmentation patterns of methyloxime-trimethylsilyl derivatives of constitutive mono- and disaccharide isomers analyzed by gas chromatography/field ionization mass spectrometry. *Rapid Commun Mass Spectrom.* 2015;29:238–46.
89. Borisov RS, Kulikova LN, Zaikin VG. Mass Spectrometry in Petroleum Chemistry (Petroleomics) (Review). *Pet Chem.* 2019;59:1055–76.
90. Li Du-X, Gan L, Bronja A, Schmitz OJ. Gas chromatography coupled to atmospheric pressure ionization mass spectrometry (GC-API-MS): Review. *Anal Chim Acta.* 2015;891:43–61.
91. Tranchida PQ, Franchina FA, Dugo P, Mondello L. Comprehensive two-dimensional gas chromatography-mass spectrometry: recent evolution and current trends. *Mass Spectrom Rev.* 2016;35:524–34.
92. Schena T, Bjerk TR, Von Mühlen C, Caramão EB. Influence of acquisition rate on performance of fast comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry for coconut fiber bio-oil characterization. *Talanta.* 2020;219:121186.
93. Kloekhorst A, Shen Yu, Yie Y, Fang Ma, Heeres HJ. Catalytic hydrodeoxygenation and hydrocracking of Alcell® lignin in alcohol/formic acid mixtures using a Ru/C catalyst. *Biomass and Bioenergy.* 2015;80:147–61.
94. Hung NV, Mohabeer C, Vaccaro M, Marcotte S, Agasse-Peulon V, Abdelouahed L, Taouk B, Cardinael P. Development of two-dimensional gas chromatography (GC×GC) coupled with Orbitrap-technology-based mass spectrometry: Interest in the identification of biofuel composition. *J Mass Spectrom.* 2020;55:1–11.
95. Yi L, Dong N, Yun Y, Deng B, Ren D, Liu S, Liang Y. Chemometric methods in data processing of mass spectrometry-based metabolomics: A review. *Anal Chim Acta.* 2016;914:17–34.
96. Flood ME, Goding JC, O'connor JB, Ragon DY, Hupp AM. Analysis of biodiesel feedstock using GCMS and unsupervised chemometric methods. *JAOCs J Am Oil Chem Soc.* 2014;91:1443–52.
97. Madsen RB, Lappa E, Christensen PS, Jensen MM, Klemmer M, Becker J, Iversen BoB, Glasius M. Chemometric analysis of composition of bio-crude and aqueous phase from hydrothermal liquefaction of thermally and chemically pretreated *Miscanthus x giganteus*. *Biomass and Bioenergy.* 2016;95:137–45.
98. Mustafa Z, Milina R, Simeonova PA, Tsakovski SL, Simeonov VD. Prediction of class membership of biodiesels using chemometrics. *J Environ Sci Heal - Part A Toxic/Hazardous Subst Environ Eng.* 2015;50:72–80.
99. Pierce KM, Schale SP. Predicting percent composition of blends of biodiesel and conventional diesel using gas chromatography-mass spectrometry, comprehensive two-dimensional gas chromatography-mass spectrometry, and partial least squares analysis. *Talanta.* 2011;83:1254–9.
100. Schale SP, Le TM, Pierce KM. Predicting feedstock and percent composition for blends of biodiesel with conventional diesel using chemometrics and gas chromatography-mass spectrometry. *Talanta.* 2012;94:320–7.
101. Toraman HE, Abrahamsson V, Vanholme R, Van Acker R, Ronse F, Pilate G, Boerjan W, Van Geem KM, Marin GB. Application of Py-GC/MS coupled with PARAFAC2 and PLS-DA to study fast pyrolysis of genetically engineered poplars. *J Anal Appl Pyrolysis.* 2018;129:101–11.
102. Barcaru A, Vivó-Truyols G. Use of Bayesian statistics for pairwise comparison of megavariable data sets: extracting meaningful differences between GC×GC-MS chromatograms using Jensen-Shannon divergence. *Anal Chem.* 2016;88:2096–104.
103. Izadmanesh Y, Garreta-Lara E, Ghasemi JB, Lacorte S, Matorros V, Tauler R. Chemometric analysis of comprehensive two dimensional gas chromatography–mass spectrometry metabolomics data. *J Chromatogr A.* 2017;1488:113–25.
104. Xu Y, Goodacre R. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *J Anal Test.* 2018;2:249–62.
105. Gu GHo, Noh J, Kim I, Jung Y. Machine learning for renewable energy materials. *J Mater Chem A.* 2019;7:17096–117.
106. Gatto L, Hansen KD, Hoopmann MR, Hermjakob H, Kohlbacher O, Beyer A. Testing and Validation of Computational Methods for Mass Spectrometry. *J Proteome Res.* 2016;15:809–14.
107. Liland KH. Multivariate methods in metabolomics - from preprocessing to dimension reduction and statistical analysis. *TrAC - Trends Anal Chem.* 2011;30:827–41.
108. Purcaro G. Chapter 2.3: Classical and comprehensive 2D LC-GC. In: Tranchida P, Mondello L, editors. *Hyphenations of Capillary Chromatography with Mass Spectrometry.* Elsevier, Amsterdam, the Netherlands 2020, pp. 227–75.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Beccaria M, Siqueira ALM, Maniquet A, et al. Advanced mono- and multi-dimensional gas chromatography–mass spectrometry techniques for oxygen-containing compound characterization in biomass and biofuel samples. *J Sep Sci.* 2021;44:115–134.  
<https://doi.org/10.1002/jssc.202000907>