Principal component analysis and partial least square regression models to understand sorption-enhanced biomass gasification

María Soledad Callén[1] · Isabel Martínez[1] · Gemma Grasa[1] · José Manuel López[1] · Ramón Murillo[1]

Abstract

Gasification represents a potential technology for the conversion of biomass into usable energy. The influence of the main gasification parameters, i.e. the type of biomass used and its composition, as well as the composition of the outlet gas, were studied by a multivariate statistical analysis based on principal component analysis (PCA) and partial least square (PLS) regression models in order to identify the main correlations between them and to the contents of methane, ethylene and tar in the outlet gas. In this work, the experimental data used as input for the multivariate statistical analysis came from a TRL-4 gasification plant running under sorption enhanced conditions, i.e. using steam as the gasifying agent and CaO as the bed material. The composition of the biomass feed played an important role in the quality of the outlet gas composition. In fact, biomasses with high ash and sulphur contents (municipal solid waste) increased ethylene content, while those with high-volatile matter content and fixed C content (wood pellets, straw pellets and grape seeds) mainly increased CO and CO_2 formation. By increasing the gasification bed temperature and the CaO/C ratio, it was possible to reduce the methane and the collected tar contents in the outlet gas. Other light hydrocarbons could also be reduced by controlling the $T_{reactor}$ and T_{FB}. Methane, ethylene and tar contents were modelled, cross-validated and tested with a new set of samples by PLS obtaining results with an average overall error between 8 and 26%. The statistically significant variables to predict methane and ethylene content were positively associated to the thermal input and negatively to the CaO/C ratio. The biomass composition was also remarkable for both variables, as mentioned in the PCA analysis. As far as the tar content, which is undesirable in all gasification processes, the decrease in the tar content was favoured by high bed temperature, low thermal input and biomass with high-volatile matter content. In order to produce an outlet gas with adequate quality (e.g. low tar content), a compromise should be found to balance average bed temperature, sorbent-to-mass ratio, and ultimate and proximate analyses of the biomass feed.

Keywords Biomass gasification · Statistical tools · Principal component analysis · Partial least square regression · Modelling

Abbreviations and nomenclature

Ash	Content of ashes in the biomass [wt%]
BFB	Bubbling fluidised bed
C	Carbon content in the biomass [wt%]
CaO/C	Sorbent to carbon molar ratio
Fixed C	Content of fixed carbon in the biomass [wt%]
CH_4	Methane content in the outlet gas [vol%]
C_2H_4	Ethylene content in the outlet gas [vol%]
C_2H_6	Ethane content in the outlet gas [vol%]
C_3H_8	Propane content in the outlet gas [vol%]
CO	Content of carbon monoxide in the outlet gas [vol%]
CO_2	Content of carbon dioxide in the outlet gas [vol%]
Cl	Chlorine content in the biomass [wt%]
GC-MS	Gas chromatography with mass spectrometry detection
GS	Grape seeds
G1-G5	Gasification experiments performed with GS
ΔH	Enthalpy [kJ/mol]
H	Hydrogen content in the biomass [wt%]
H_2	Hydrogen content in the outlet gas [vol%]
HHV	Higher heating value [MJ/kg]

✉ María Soledad Callén
marisol@icb.csic.es

1 Instituto de Carboquímica (ICB-CSIC), C/Miguel Luesma Castán, 4, 50018 Zaragoza, Spain

ICP-OES	Inductively coupled plasma optical emission spectroscopy
LHV	Lower heating value [MJ/kg]
LV	Latent variables
IC	Ion chromatography
IQR	Interquartile range = Q3-Q1
Min	Minimum value of a variable
Max	Maximum value of a variable
Moisture	Content of water in the biomass [wt%]
MSW	Municipal solid waste
M1-M11	Gasification experiments performed with MSW
N	Nitrogen content in the biomass [wt%]
n	Number of samples
O	Oxygen content in the biomass [wt%]
OE	Overall error for the calibration samples Eq. (3) [%]
OE_{CV}	Overall error for the cross-validation samples Eq. (3) [%]
OE_{ext}	Overall error for the external validation samples Eq. (3) [%]
PCA	Principal component analysis
PLS	Partial least square
Q^2	Coefficient of determination for the cross-validation samples
Q^2_{ext}	Coefficient of determination for the external validation samples
R^2	Coefficient of determination for the calibration samples Eq. (1)
RSD	Relative standard deviation [%]
RMSEC	Root mean square error of calibration Eq. (2)
$RMSE_{CV}$	Root mean square error of cross-validation
$RMSEP_{ext}$	Root mean square error of prediction for the external validation samples
S	Sulphur content in the biomass [wt%]
S/C	Steam to carbon ratio
SEG	Sorption enhanced gasification
SP	Straw pellets
S1-S2	Gasification experiments performed with SP
Std Deviation	Standard deviation, units of the variable
T_{bed}	Temperature of the bed [°C]
T_{FB}	Temperature of the freeboard [°C]
$T_{reactor}$	Temperature of the reactor [°C]
TRL-4	Technology readiness level 4
Volatile	Content of volatile matter in the biomass [wt%]
WP	Wood pellets
W1-W2	Gasification experiments performed with WP
WChips	Wood chips

WC1	Gasification experiment performed with WChips
$\widehat{y}_i$	Predicted Y values according to the PLS model Eq. (1)
$\bar{y}_i$	Means values Eq. (1)

1 Introduction

Effective climate change mitigation needs a substantial reduction of greenhouse gas emissions, and to do so, an important shift from fossil to renewable energy is needed [1]. Biomass is a renewable energy source that could play a crucial role in climate change mitigation due to its consideration as carbon–neutral energy compared to fossil fuels, since the CO_2 released when biomass is burned is the CO_2 that plants capture through photosynthesis as they grow [2]. In this way, biofuels produced from biomass could be decisive in the decarbonisation of heat, transport, electricity and high-value chemical production sectors [3].

Gasification is one of the main thermochemical routes for converting biomass and waste materials into syngas that can be directly used as a fuel or be converted through a catalytic process [4]. In a biomass gasification process, a solid or / liquid organic compound is converted to a gas/vapour phase and a solid phase [5]. In steam gasification, the high-calorific value syngas produced is mainly composed of hydrogen (H_2), carbon monoxide (CO), methane (CH_4), and carbon dioxide (CO_2) and also includes some light hydrocarbons and heavier hydrocarbons which condense between 250 and 300 °C, known as tars [6]. The resulting solid phase, referred to as char, is composed of the inert and the unconverted organic fraction, mainly carbon and ash.

Multiple reactions take place at the same time during the gasification process, which hinders understanding of the process and the effect of the main variables involved. This complexity can be overcome by conducting more research on how different variables affect the process and how these variables interact with each other [7]. In the particular case of gasification, these variables are related, on the one hand, to the gasification parameters (temperature, gasifying agent, reactor type, thermal power, etc.) and to the material used as the feedstock and on the other hand, to the reactor technology used. In addition, it is also necessary to take into consideration the nature and quality of the gas obtained in the process, which should meet different specifications depending on its final use.

Different studies into biomass gasification with fluidised bed gasifiers have been reported in the literature [8–16] that assess the effect of the operating parameters [8–10] on the gas produced, feedstock characteristics, gasifier design [11, 12], gasification agent [10, 13–16] and the use of catalysts [8]. However, modelling and simulation are helpful tools

for better understanding these complex phenomena and so facilitating the experimental work at large-scale pilot plants [7]. Nevertheless, each one of these models has limitations, unique features and capabilities that limit the scale of applicability.

Identification of the most influential parameters on the gas quality is one of the main outcomes of developing a general model for biomass gasification. Different models can be applied: theoretical model (equilibrium model), semi-empirical model (modified equilibrium model, kinetic combined with hydrodynamic model) and empirical model (statistical model) [17]. In this particular work, attention will be mainly focussed on statistical models.

Multivariate statistical tools have been applied to assist the study of data composed of large sets of experimental data and different variables for easier interpretation drawing conclusions in several areas [18]. It merits mention that these studies have not been frequently applied to energy research.

Principal component analysis (PCA) and partial least square (PLS) regression models are two powerful analytical tools that have been successfully applied in order to determine which parameters meaningfully affect a process and how to model dependent variables. Statistical tools [19] have been applied to gasification research in recent years, in particular, the application of PCA [20, 21] to study the influence of the biomass characteristics and PLS to the gasification process [22–24]. However, the complexity of the process, the different types of gasifiers, biomass used and experimental conditions complicate comparison with and extrapolation to dissimilar experimental installations. Studies have also been conducted on circulating fluidised bed gasifiers to correlate input and output parameters in order to determine the most important parameters affecting product gas quality [25].

 PCA was also applied to examine which variables had a more determinant influence on the differentiation found in the syngas resulting from the steam gasification of particular feedstocks [22] based on (co)-gasification. Linear regression models have also been applied to biomass in fluidised bed for chemical looping gasification for future simulations of devolatilisations [26]. Artificial neural network (ANN) models have also been developed to simulate the gasification process in a fluidised bed gasifier with different biomasses at different operating parameters based on data from experimental investigation [27, 28]. Some of these PLS modelling studies have been performed by compiling bibliographic data on gasification [22, 23], whereas others have taken into account the properties of the biomass and gasification parameters [20, 29].

The purpose of this research was to study the application of statistical analysis to the behaviour of sorption-enhanced gasification (SEG). The main characteristic of this indirect gasification process is that a CO_2 sorbent is fed into the gasifier together with the biomass, usually at high temperature, from a second reactor or calciner [30, 31]. Dolomite and limestone are typically used as the CO_2 sorbent in this process. In this way, the CO_2 produced by gasification reacts with the CaO present in these sorbents through an exothermic reaction, supplying part of the energy required for endothermic gasification reactions. For this work, SEG experiments were performed in an atmospheric fluidised bed gasifier using steam as the gasifying agent for five different types of biomass feedstocks. The gasification parameters, nature of the biomass (proximate and ultimate analysis) and the composition of the outlet gas were studied statistically as a first approximation in order to gain a better understanding of the parameters with the greatest influence over the process. Moreover, certain variables associated with gas quality had to be modelled owing to the knowledge gap that hinders the ability to predict gasification processes.

To our knowledge, this is the first statistical study (PCA and PLS) to be applied to an SEG process. One of the main novelties of this work is that all the data used to apply statistical tools in this research were obtained experimentally from the same installation by modifying the operating variables associated with the SEG process (i.e. sorbent-to-biomass proportion, steam-to-carbon ratio and gasification temperature, as the most significant) for different biomass feedstocks. This involves that the same reactor has been utilised for all experiments, and in this particular case, no literature data have been considered. Furthermore, given that the gasification facility has a significant technology readiness level (TRL4), with a fluidised bed reactor of 3-m height and 0.20-m internal diameter [32–34], the application of statistical tools and outlet gas modelling tools could help to reduce the number of experiments to be performed, simplifying the process and reducing time and expense in biomass gasification research. In fact, other of the expectations of this research relies on the modelling of the tar content that is a variable affecting negatively to the gasification and involves, not only the sampling, but also the further analysis with analytical techniques such as gas chromatography with mass spectrometer detection.

In the literature, there are few models applied to tar prediction [19, 28, 35, 36] compared to gasification modelling [37–41]. The lack of comparability of experimental data and the chance of obtaining previous information on this variable would be advantageous for reducing experimental work and costs. Definitively, this first approach for improving knowledge on biomass gasification process will also be of great interest for scaling up the facility to an industrial scale. All this justifies the main aim of this research that has been focused on the application of PCA to better understand the correlations among gasification parameters, biomass and outlet gas composition, besides the development of three PLS models validated for the prediction of the contents of

collected tar, CH_4 and C_2H_4 with experimental data of the same SEG installation.

2 Materials and methods

2.1 Gasification plant

SEG tests were performed in the 30 kW_{th} (in reference to the nominal thermal input of biomass on LHV basis) bubbling fluidised-bed (BFB) reactor shown in Fig. 1 (3 m of height, 0.20-m internal diameter). Details regarding the configuration of this BFB installation can be found elsewhere [32–34]. As indicated in this figure, CO_2 sorbent (in its calcined form) and biomass were fed separately into the reactor from two independent closed hoppers by different screw feeders. Partially converted solids left the BFB reactor through a lateral overflow, which is shown schematically on the left-hand side of the reactor

in Fig. 1, and were collected in a hopper that was periodically discharged during operation. Solid bed inventory in the bottom dense zone of the reactor was regulated to around 5–7 kg (depending on the fluidisation conditions) thanks to this lateral overflow. A constant flow rate of 0.009 Nm^3/min of N_2 was fed with the H_2O to facilitate its flow through the evaporation system. Moreover, a second N_2 flow rate of 0.003 Nm^3/min was introduced into the reactor through the CO_2 sorbent screw feeder, the overflow solids hopper and the biomass hopper (not all indicated in Fig. 1 for the sake of simplicity) to prevent the back-flow of H_2O-rich syngas into the hoppers, which would have caused operational problems.

2.2 Materials characterisation

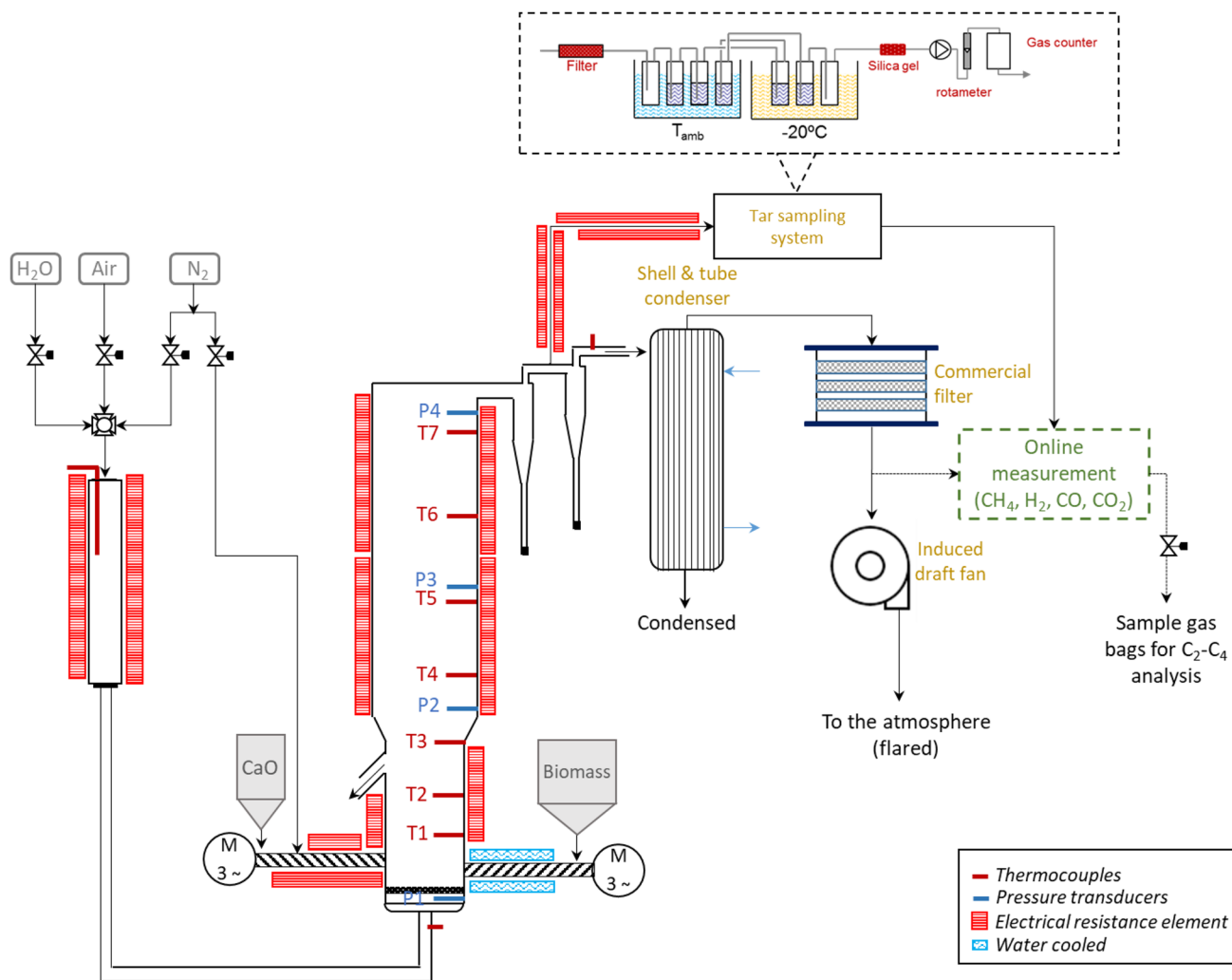Five biomass feedstocks were used for the SEG experiments. Four of them were lignocellulosic materials (pine wood

Fig. 1 Schematic diagram of the 30 kW_{th} BFB gasification plant at ICB-CSIC

pellets and chips (WP, WChips), grape seeds (GS) and straw pellets (SP)), while the fifth was a feedstock derived from municipal solid waste (MSW) delivered by Econward [33].

The ultimate analysis, proximate analysis and calorific values of the different fuels are shown in Table 1. The ultimate analysis was determined in a Thermo Flash 111 (UNE-EN-5104) whereas moisture, volatile matter and ash contents were calculated according to standards UNE-EN 18,134–3, UNE-EN 14,775 and UNE-EN 15,148, respectively.

The content of chlorine was performed by ion chromatography with thermal conductivity detector. Analyses were carried out by a Metrohm serie 820 ion chromatography system with Chromeleon version 6.60SP2 software.13 The anion method used the Metrosep A Supp 5–250/4.0 analytical column (2 mm × 250 mm) using standards with known concentrations of chlorine to establish calibration curves. Previously to the analysis, solid samples were combusted in a calorimetric bomb Parr 6400 and the washing water was concentrated, filtered and rinsed to a final volume. Samples were determined by duplicate with RSD lower than 5%.

The content of sulphur was performed by inductively coupled plasma optical emission spectroscopy (ICP-OES). An Xpectroblue-EOP-TI FMT26 (Spectro) allowed determining the atomic content of S. Previously, the samples were fused with Na_2O_2, at 460℃ in a muffle for 1.5 h, and further, they were attacked with 5 ml of HCl at 37%, rinsing with de-ionised water to a final volume of 100 ml. Measurements were carried out using external standard calibration. Samples were determined by duplicate and RSD < 5% according to a reference material was obtained.

The calcined material used as the CO_2 sorbent and bed material was obtained from parent limestone calcination in the BFB reactor. Calcination was performed at c.a. 910 ℃ by means of air combustion using the same wood pellets that appear in Table 1 as a fuel.

The calcined material was also analysed by inductively coupled plasma-optical emission spectrometry (ICP-OES) to determine its chemical composition, which consisted of 92.2 wt% CaO, 0.9 wt% Fe_2O_3, 0.9 wt% MgO and 0.4 wt% Al_2O_3 (the remaining oxides being based on Na, Si and K).

A total of 21 gasification experiments (named correlative to the initial letter of biomass, e.g. for MSW: M1 to M11; for SP: S1 to S5; for GS, G1, G2, for WP, W1, W2 and for WChips: WC1) were carried out with the five different materials at the conditions shown in Table 2 using steam as gasifying agent. The temperatures included in this table correspond to the average of the thermocouples placed within the dense solid bed (i.e. T1 and T2 in Fig. 1, T_{bed}), to the average of the thermocouples along the freeboard (i.e. from T3 to T7, T_{FB}), and the average of all the thermocouples placed along the gasifier (i.e. from T1 until T7, $T_{reactor}$). As a consequence of the specific design of this pilot plant (mainly electrical resistance element power input, pilot plant heat

Table 1 Proximate analysis, ultimate analysis and calorific values of the different biomass feedstocks tested in the 30 kW_{th} BFB gasifier (HHV = higher heating value; LHV = lower heating value)

wt%	WP	GS	MSW	SP	WChips
Proximate analysis (mass basis)					
% Moisture	5.55	6.30	5.90	6.51	8.09
% Ash	0.36	4.30	32.2	4.87	1.30
% Volatile matter	78.8	65.1	55.4	70.3	72.9
% Fixed Carbon	15.3	24.3	6.60	18.3	17.7
Ultimate analysis (mass basis)					
% C	49.3	53.9	34.8	45.1	49.2
% H*	6.30	6.58	4.40	5.96	5.83
% N	0.10	2.20	1.70	0.70	0.14
% S**	0.02	0.12	0.66	0.13	0.05
% O	40.8	32.4	30.9	38.6	39.5
%Cl***	0.00	0.06	0.43	0.36	0.01
Calorific value					
HHV [MJ/kg]	19.1	22.1	13.8	17.4	18.1
LHV [MJ/kg]	17.6	20.5	12.8	16.1	16.7

*Includes H in moisture

**The S content was determined by ICP-OES

***The Cl content was determined by ion chromatography since this technique is able to determine the small amounts of Cl present in these samples

losses, calcined sorbent inlet temperature), there is a relation between the biomass thermal input, solid bed temperature and CaO/C ratio. In general, experiments with a higher biomass thermal input or a higher CaO flow to the reactor led to lower T_{bed} stabilisation.

An additional set of three experiments were used to validate model predictions. One of them was carried out with SP, whereas the other two were performed with WP. More details regarding the gasification variables can be found in Table S1. Supplementary Information.

2.3 Gas analysis

Permanent gas composition (CO_2, CO, H_2 and CH_4) was determined by means of an online gas analyser (SICK GMS810) that was placed downstream of the gas filter and tar sampling system, as depicted in Fig. 1. In addition to this continuous gas measurement, Tedlar sampling bags were also taken during steady-state operation in order to determine the composition of permanent gases and light hydrocarbons (up to C4) by gas chromatography.

Permanent gases and light hydrocarbons (C1–C2) were analysed in a Bruker 450 gas chromatograph (GC) equipped with a thermal conductivity detector (TCD). Separation was performed by using two stainless steel packed columns (Molsieve 13X and HayeSep Q). An oven temperature of 60 ℃ was maintained for 10 min. The carrier gas was Ar at

Table 2 Gasification parameters (i.e. biomass thermal input, steam-to-carbon ratio, molar ratio of CaO in sorbent to C in the biomass, solid bed temperature, freeboard temperature and reactor temperature) for each of the experiments considered in this work

Experiment	Biomass thermal input [kW$_{th}$]	S/C ratio	CaO/C ratio	T_{bed} [°C]	T_{FB} [°C]	$T_{reactor}$ [°C]
WP						
W1	16.6	1.4	0.3	717	696	702
W2	14.6	1.6	0.4	665	639	647
GS						
G1	20.8	1.1	0.3	630	617	620
G2	12.7	1.8	0.3	672	641	649
MSW						
M1	16.6	1.4	0.4	680	663	668
M2	11.3	1.4	0.5	683	637	650
M3	11.3	1.4	0.4	696	658	669
M4	18.4	1.4	0.3	635	653	648
M5	8.40	1.4	0.3	664	736	716
M6	15.9	1.0	0.4	617	748	710
M7	11.4	1.4	0.2	680	696	692
M8	11.7	1.3	0.6	629	697	677
M9	9.10	1.3	0.3	735	699	708
M10	9.00	1.3	1.2	704	687	692
M11	8.70	1.0	1.1	708	686	692
SP						
S1	10.5	1.4	0.3	644	663	658
S2	9.60	1.5	0.5	651	631	637
S3	10.5	1.3	0.3	665	644	650
S4	10.8	1.3	0.2	735	711	718
S5	10.5	1.4	0.8	697	652	665
WChips						
WC1	5.30	3.0	1.81	705	681	688

Each biomass experiment was named with the initial of the corresponding biomass and consecutive numeration

a constant column flow rate of 30 Nml/min. The TCD temperature was 200 °C. The permanent gases analysed included H_2, CO_2, O_2, N_2 and CO, while the light hydrocarbons analysed included methane (CH_4), ethane (C_2H_6) and ethylene (C_2H_4). Certified gas mixtures (Air Products) were used for identification and quantification purposes. The standards were injected by triplicate with relative standard deviations lower than 1.5%. The samples were injected by duplicate with relative standard deviations lower than 3% except for O_2 and ethane that reached 10%.

Subsequently, light hydrocarbons (C1–C4) were analysed in a PerkinElmer Clarus 590 gas GC equipped with a flame ionisation detector (FID). Separation was performed by a 30-m-long and 0.32-mm-wide Alumina chloride capillary column. The light hydrocarbons analysed included methane (CH_4), ethane (C_2H_6), ethylene (C_2H_4), propane (C_3H_8), propylene (C_3H_6), isobutene and n-butane (C_4H_{10}), and trans-2-butene, 1-butene, isobutene, cis-2-butene and 1,3-butadiene (C_4H_8). An initial oven temperature of 40 °C was maintained for 2.5 min, after which a heating rate of 5 °C/min was implemented to reach an oven temperature of

90 °C. The heating rate was then increased to 15 °C/min to reach the final oven temperature of 180 °C. This temperature was maintained for 4.5 min. The carrier gas was He with an initial pressure of 9 psi for 16 min, then raised to 18 psi. The injector and FID temperatures were both 250 °C. Certified gas mixtures (Air Products) were used for identification and quantification purposes. The standards were injected by triplicate with relative standard deviations lower than 5%. The samples were injected by duplicate with relative standard deviations lower than 5% in the case of hydrocarbons.

2.4 Tar sampling and analysis

Because one of the undesirable products of gasification is related to tar formation, an offline method for tar sampling was used based on the specific protocol for biomass gasification units described elsewhere [42, 43]. This tar sampling method involved absorbing tars by flowing the gas through a series of 7 impinger bottles filled with isopropanol, which were placed in two different cooling baths as shown in Fig. 1: four impingers (1, 2, 3 and 5) at room temperature

and three (4, 6 and 7) at − 20 ℃. All of them contained 100 ml of isopropanol, with the exception of impingers 1 and 7, which were empty. The gas flow rate sampled through the tar system was regulated by a pump, and the total volume passed was measured by a gas metre before the gas was sent to the online analyser. Once the tar sampling test was completed, the isopropanol from the impingers was collected in an amber glass bottle, and both the impingers and connection tubes were rinsed with isopropanol. The final volume of the isopropanol collected, referred to as 'collected tar', was measured and the tar content determined by gas chromatography–mass spectrometry (GC–MS).

The chemical composition of the collected tar was determined using a Varian CP-3800 gas chromatograph connected to a Saturn 2200 Ion Trap Mass Spectrometer. 1 μl of the sample was injected in the split mode with a ratio of 20:1. A low-bleed capillary column, CP-Sil 8 CB: 5% phenyl, 95% dimethylpolysiloxane (60 m × 0.25 mm i.d. × 0.25-μm film thickness) was used. An initial oven temperature of 60 ℃ was maintained for 3 min keeping a ramp rate of 7 ℃/min until a final temperature of 300 ℃ for 15.57 min. The carrier gas was He (BIP quality) at a constant column flow of 1 ml/min. The injector, detector and transfer line temperatures were 300 ℃, 200 ℃ and 300 ℃, respectively.

The MS was operated in electron ionisation mode within 35–550 m/z range. The identification and quantification of organic compounds were carried out by external standard calibration for a total of 22 compounds according to the quantitation ion (benzene (m/z = 78), toluene (m/z = 91), ethylbenzene (m/z = 91), p + m xylene (m/z = 91), o-xylene (m/z = 91), phenol (m/z = 94), 4-methyl phenol (m/z = 107), naphthalene (m/z = 128), acenaphthylene (m/z = 152), acenaphthene (m/z = 153), fluorene (m/z = 165), phenanthrene (m/z = 178), anthracene (m/z = 178), fluoranthene (m/z = 202), pyrene (m/z = 202), benz(a)anthracene (m/z = 228), chrysene (m/z = 228), benzo(a)pyrene (m/z = 252), benzo(b)fluoranthene (m/z = 252), indeno(1,2,3-cd)pyrene (m/z = 276), dibenz(a,h)anthracene (m/z = 278) and benzo(g,h,i) perylene (m/z = 276)).

The tar content was obtained as the sum of each compound based on these standards, according to the final volume of isopropanol and the gas sampled during each experiment. More details regarding the individual composition of the tar content have been studied in previous papers [33, 34]. RSD values relative to standards were lower than 9% for three injections. Samples were analysed by duplicate with RSD lower than 15%.

2.5 Multivariate statistical tools

Two multivariate statistical tools, PCA and PLS, were used to determine the relationships between the gasification parameters, biomass properties and outlet gas quality, and to model particular parameters of gas quality.

SPSS 15.0 and Unscrambler X 10.3 software were used to obtain the PCA and the PLS regression models.

2.5.1 Principal component analysis

PCA [44, 45] is a statistical procedure that allows the orthogonal transformation of a set of possibly correlated variables into a set of linearly uncorrelated variables called principal components. PCA is one of the most powerful exploratory data analysis tools available for investigations of large data sets. It has been applied widely and successfully in different scientific fields [46–51].

The first component explains the highest portion of the total variance in the data and the successive components explain in decreasing order the rest of the total variance in such a way that only few components explain most of the total variance in the data.

The main aim of PCA was to study correlations among the process parameters, biomass characteristics and gas product quality. In this particular study, all these variables were obtained experimentally in the same installation described in Sect. 2.1 and characterised by the different analytical techniques described in Sects. 2.2, 2.3 and 2.4 so that no bibliographic data were compiled. PCA was applied to the total of 21 experiments performed with the five different biomasses (Table 2) and to a total of 26 variables including the parameters shown in Table 3. All data were mean-centred and divided by the standard deviation so that variable standardisation was carried out because different kinds of variables and scales were taken into account. This ensured that all variables had the same chance to influence the estimation of the components. The cross-fold validation, the uncertainty Martens' test using the optimal number of components and nonlinear iterative partial least square (NIPALS) algorithm were applied in the model input.

A total of 5 PCs in accordance to the Kaiser's rule (and the percent of the cumulative data variance) were retained, and only those values (either positive or negative) with the maximum loading (in absolute values) per each retained PC were considered for the interpretation of the links among the initial variables and outlet gas composition similarities/dissimilarities.

A brief summary with the potential information provided by PCA: score values and loading values has been added. The score values describe the properties of the samples and allow visualising if groupings or separations among samples could be observed, indicating sample differences or similarities. In this research, score values provided information on how different gasification experiments were related to each other. The loading plots describe the relationships between variables and

Table 3 Descriptive statistics of the variables involved in the gasification process ($N = 21$ samples) (IQR = interquartile range)

	Mean	Max	Min	Range	Std Deviation	Variance	Skewness	Kurtosis	Median	IQR	RSD [%]
Gasification conditions											
T_{bed} [ºC]	677	735	617	118	34.5	1190	0.0	-0.9	680	53.1	5.1
T_{FB} [ºC]	673	748	617	131	34.8	1209	0.5	-0.3	664	52.2	5.2
$T_{reactor}$ [ºC]	674	718	620	97.8	28.3	801	0.0	-1.0	669	42.4	4.2
CaO/C	0.5	1.8	0.2	1.6	0.4	0.2	2.1	4.5	0.3	0.2	80
S/C	1.4	3.0	1.0	2.0	0.4	0.2	3.3	12.9	1.4	0.1	28.6
Thermal power [kW]	12.1	20.8	5.3	15.5	3.8	14.2	0.7	0.2	11.3	4.94	31.4
Composition of the outlet gas											
Collected tar (g/m^3 N)	33.6	62.2	6.7	55.5	13.9	195	0.6	0.5	29.6	11.6	41.4
CO [vol%]	6.1	13.8	2.1	11.7	3.4	11.8	0.9	0.3	5.5	5.4	55.7
CO_2 [vol%]	7.7	12.3	4.3	8.1	2.3	5.2	0.2	-0.8	7.7	4.3	29.9
CH_4 [vol%]	9.9	14.9	4.7	10.3	2.5	6.4	0.4	0.3	9.4	2.4	25.3
H_2 [vol%]	37.3	45.6	29.2	16.5	5.6	30.9	0.0	-1.6	36.8	10.5	15.0
C_2H_6 [vol%]	1.1	1.9	0.5	1.4	0.4	0.1	0.5	0.1	1.1	0.5	36.4
C_2H_4 [vol%]	3.3	6.4	0.8	5.6	1.4	1.9	0.5	-0.1	3.0	1.9	42.4
C_3H_8 [vol%]	0.0	0.2	0.0	0.2	0.0	0.0	2.7	9.2	0.0	0.0	
C_3H_6 [vol%]	0.3	1.5	0.0	1.5	0.3	0.1	2.8	10.0	0.3	0.3	100
Composition of the biomass fed											
Moisture [%]	6.2	8.1	5.6	2.5	0.5	0.3	2.4	7.8	5.9	0.6	8.1
Ash [%]	18.5	32.2	0.4	31.8	14.8	218	-0.1	-2.2	32.2	27.3	80
% Volatile matter	62.9	78.8	55.4	23.4	8.7	75.4	0.5	-1.3	55.4	14.9	13.8
% Fixed C	12.4	24.3	6.6	17.7	6.6	43.6	0.5	-1.4	6.6	11.7	53.2
% C	41.1	53.9	34.8	19.1	7.2	52.2	0.5	-1.4	34.8	10.3	17.5
% H	5.2	6.6	4.4	2.2	0.9	0.8	0.2	-1.9	4.4	1.6	17.3
% N	1.3	2.2	0.1	2.1	0.7	0.5	-0.6	-1.1	1.7	1.0	53.8
% S	0.4	0.7	0.0	0.6	0.3	0.1	-0.1	-2.1	0.7	0.6	75.0
% O	34.2	40.8	30.9	9.9	4.1	16.9	0.6	-1.7	30.9	7.7	12.0
% Cl	0.3	0.4	0.0	0.4	0.2	0.0	-1.2	-0.3	0.4	0	66.7
HHV* [MJ/kg]	16.1	22.1	13.8	8.3	2.8	7.9	0.8	-0.3	13.8	3.6	17.4

provide information on variable contribution and correlations. Variables that are close to each other are positively correlated, contributing similar information and those opposite to each other are negatively correlated and positioned in diagonally opposed quadrants. The direction of PC1 in relation to the original variable is given by the cosine of the angles. Geometrically, loading is the cosine of the angle between the variable and the PC, ranging between -1 and $+1$ so that the smaller the angle, the larger the loading. When the angle between variable and PC is close to zero, completely describes the variable. When the angle between variables is 90º, both variables are uncorrelated whereas angles greater than 90º and 180º indicate that variables are negatively correlated [52].

The distance to the origin also conveys important information so that the further away from the plot origin a variable lies, the stronger the impact that variable has on the model. Variables located at the origin of the PC or very close to the origin are not well described by both PC. In order to

better understand the interpretation of loading plots, it will be necessary to search for each PC those variables with high loadings. The higher the loading of a variable for a PC, the higher the contribution to the sample shown by the score plot. Variables lying close are highly correlated. The sign of a loading value (positive or negative) informs the positive or negative correlation of that variable with the PC [53].

2.5.2 Partial least square regression

PLS was developed by Wold in 1960 for application to social science modelling [54], and it has been used widely in different scientific statistical approaches [55, 56].

In this particular work, PLS was applied in order to model particular parameters that determine the quality of the outlet gas. The dependent variables considered, i.e. the variables to be modelled, associated with the outlet gas were CH_4 and C_2H_4 contents and collected tar. On the other hand, the independent variables considered were the gasification

parameters, i.e. biomass thermal input, steam-to-carbon ratio (S/C), CaO-to-C molar ratio (CaO/C), solid bed temperature (T_{bed}), freeboard temperature (T_{FB}) and reactor temperature ($T_{reactor}$) and the composition of the biomass feed, i.e. proximate analysis (moisture, ash, volatile and fixed C contents) and ultimate analysis (C, H, N, S, O and Cl contents), as well as higher heating value (HHV). CaO/C refers to the molar ratio between the moles of CaO in the CO_2 sorbent fed to the gasifier and the moles of C in the biomass introduced. This material was the same for all experiments. PLS was applied to the entire set of experiments (21 samples) and to the 17 independent variables in order to establish a calibration PLS model for each dependent variable obtaining a total of 3 PLS models.

Prior to statistical analysis, all the initial variables were mean-centred and divided by the standard deviation in order to avoid influence from the different scale units. When the main aim of modelling is the prediction, one of the most important considerations is to validate the model. Once the model was elaborated or calibrated, the validation of the model was performed by two methods: internal validation and external validation. The internal validation was carried out by the cross-validation method (CV) with a random number of segments equal to 20 was used to measure the prediction performance of the multivariate model since there was a relatively low number of samples. This involves partitioning a data sample into complementary subsets, performing the analysis on one subset and validating the analysis on the other subset [57]. A set of three external validation experiments were also used to assess the predictive performance of the developed models (Table S1, Supplementary Information) with variables covering the range of the ones used for the calibration model. An uncertainty Martens' test with the optimum number of latent variables (LV) was applied to identify non-significant variables. The NIPALS algorithm was also applied to each PLS model.

The quality of the PLS model was evaluated by calculating different parameters, including the coefficient of determination R^2 for each predicted variable (Eq. 1),

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2} \tag{1}$$

where y_i represents reference y values, while $\widehat{y}_i$ represents predicted Y values according to the PLS model, $\overline{y}_i$ means value and n refers to the number of samples. The higher the R^2, the better the model as a function of the descriptor variables, with the upper limit of 1 for a perfect fit and 0.5 for an acceptable fit [44].

The R^2 explains how well a model performs when replicating the observed outcomes. It is the proportion of the variance in the dependent variable that is predictable from the independent variable. For the cross-validation, this

determination coefficient is denoted by Q^2 (cross-validated correlation coefficient) while Q^2_{ext} is the corresponding determination coefficient of the predicted variable for the external validation.

In order to assess the predictive ability of multivariate models and as a measurement of precision uncertainty, the root mean square error of calibration (RMSEC) (Eq. 2) was used to express differences between the values predicted by the PLS model and the experimentally measured values. This value is expressed in the same unit as the Y or dependent variable. The lower the RMSEC, the better the model prediction. The $RMSE_{CV}$ and the $RMSEP_{ext}$ denoted this value for the cross-validation and external validation, respectively.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2}{n}} \tag{2}$$

Finally, an average overall error was used (Eq. 3) to relate the RMSEC to the experimental values according to the formula:

$$Average\ OE = \frac{100 * RMSEC}{\widehat{y}_i} \tag{3}$$

As mentioned above, the OE_{CV} and OE_{ext} measured this value for the cross-validation and external validation giving information on the relative precision. Residuals, which are of diagnostic value for the quality of a model and inform on the non-explained variance, enabled identifying outliers that did not fit the model. A normal probability plot of the Y residuals of the model showed a fairly straight line.

3 Results and discussion

Usually in statistical studies, a normal distribution of the variables is used because most of the normal phenomena can be modelled with this distribution. The normal distribution, which is defined by two parameters (i.e. the mean and the standard deviation), describes how the values of a variable are distributed. It is symmetrical, being the most important probability distribution in statistics for independent, random variables. Firstly, a study of the descriptive statistics for all the variables involved in the gasification process was performed as shown in Table 3. As measures of central tendency, the mean and median were included for normal and non-normally distributed variables. When variables are normal, these values are the same. This happened for most of the variables considered, whereas the largest differences were found for CaO/C, ash, volatile matter, fixed C and S.

Measures of variability or dispersion included range, standard deviation, variance, minimum and maximum and

relative standard deviations (RSD) for normal variables. It was observed that some variables had a very narrow range, such as those variables affecting the ultimate composition of the biomass feed, e.g. the S and Cl contents due to the low content of these elements in the different biomasses, whereas others like T_{bed}, T_{FB}, $T_{reactor}$, collected tar and the ash content of the feed biomass presented a wider variation in part due to the different conditions at which the gasification process was carried out. This was interesting from the modelling point of view as it would permit verification of the adequacy of the model for a wider range of variables. Out of these ranges of variables, the model should be checked and validated. Therefore, this was one of the model limitations. Only for the range of these variables, the PLS model could be applied to predict the three dependent variables of this research. Out of this range, the model should be validated to confirm its adequacy in the prediction of new samples.

Another important parameter to take into account was related to RSD%, which might be useful for detecting which variables had the largest absolute variation. In this case, the highest values corresponded to C_3H_6, CaO/C, the CO content in the product gas, and the S, Cl, N and fixed C contents in the biomass, whereas the lowest RSD% was associated with variables covering a wider range, such as gasifier temperatures.

Other interesting parameters to study were the skewness and kurtosis, which also provide information on data distribution. Skewness and kurtosis were introduced for those variables with no normal distribution with the interquartile range as dispersion indicator [58]. Kurtosis determines the heaviness of the distribution tails, while skewness determines the distribution symmetry. From Table 3, it is observed that most of the variables had skewness values close to zero (exception CaO/C, S/C, collected tar, CO, C_3H_8, C_3H_6, moisture, %N, %Cl and HHV) indicating symmetry while the kurtosis remarked the tailing of some variables relative to a normal distribution. In general, it is acceptable the values for asymmetry and kurtosis between −2 and +2 to prove a normal distribution [59]. In this way, the S/C, C_3H_8 and C_3H_6 and moisture were the exception.

Regarding the interquartile range as a measure of spread, it indicates where most of the values lie and it has been shown in Table 3 for those variables which could follow a non-normal distribution. The main advantage of this parameter is that it is not sensible to atypical values, while the standard deviation and the range would change drastically with outliers.

To avoid the lack of normality in some variables, standardisation of variables is common in order to compare different types of observations based on the own distribution of each observation. To each observation of each variable, subtract the mean and divide by the standard deviation obtaining the standard normal distribution which, is a special case of the normal distribution (the mean is zero and the standard deviation is 1). This practice is very common in statistical studies like PCA and PLS, and it has been applied to this particular research.

3.1 Principal component analysis

PCA was applied to all experiments (21 samples) and all variables, including the independent variables affecting the gasification process, the biomass used (ultimate and proximate analyses) and the dependent variables associated with syngas quality, totalling 26 variables.

Results obtained for PCA indicated that a total of 5 factors (Fig. 2) with eigenvalues higher than 1, Kaiser's rule, were able to explain 89% of the total variance where the two first factors were able to explain most of the variance, 70%. The first factor explained 44% of the total variance whereas the second, third, fourth and fifth factors explained 24%, 12%, 6% and 4% of the variance, respectively. The extraction communalities found for almost all the variables were higher than 0.8 (with the lowest values for $T_{FB} = 0.71$ and CaO/C = 0.77), and no variable was removed from the analysis.

In terms of the different variables studied, the loading plots between the two first components (Fig. 2b) explaining most of the variance shows that ash, S, Cl and C_2H_4 contents, and the volatile matter, C, H, fixed C, O, HHV and CO contents correlated significantly with PC1, although inversely. Ash, S, Cl and C_2H_4 correlated positively, whereas the other parameters related to the biomass used were negatively correlated. With the exception of C_2H_4 and CO contents, all these parameters were related to the composition of the biomass feed. Therefore, the nature of the biomass feedstock seemed to play a remarkable role in the gasification process, as shown by their link to the first factor, which is referred to as 'biomass feed'.

By analysing the second factor (Fig. 2b), PC2, it was observed that thermal power, collected tar, C_2H_6, C_3H_6, C_3H_8 and CH_4 were positively correlated with PC2 while $T_{reactor}$, T_{bed} and CaO/C correlated negatively. This second factor was related to the gasification conditions and hydrocarbons composition and is referred to as 'gasification conditions and hydrocarbons composition'. Into this group of different light hydrocarbons (Fig. 2b), it was possible to distinguish on the one hand, C_2H_6, C_3H_6 and C_3H_8 that were negatively affected by the T_{FB} and $T_{Reactor}$. The important role of the gasification temperature in order to reduce the presence of the light hydrocarbons has also been shown in previous research by these authors [32–34]. On the other hand, the closeness of both variables, CH_4 and the collected tar, in the graph (Fig. 2b) remarked a similar behaviour, being negatively affected by the T_{bed} and CaO/C ratio.
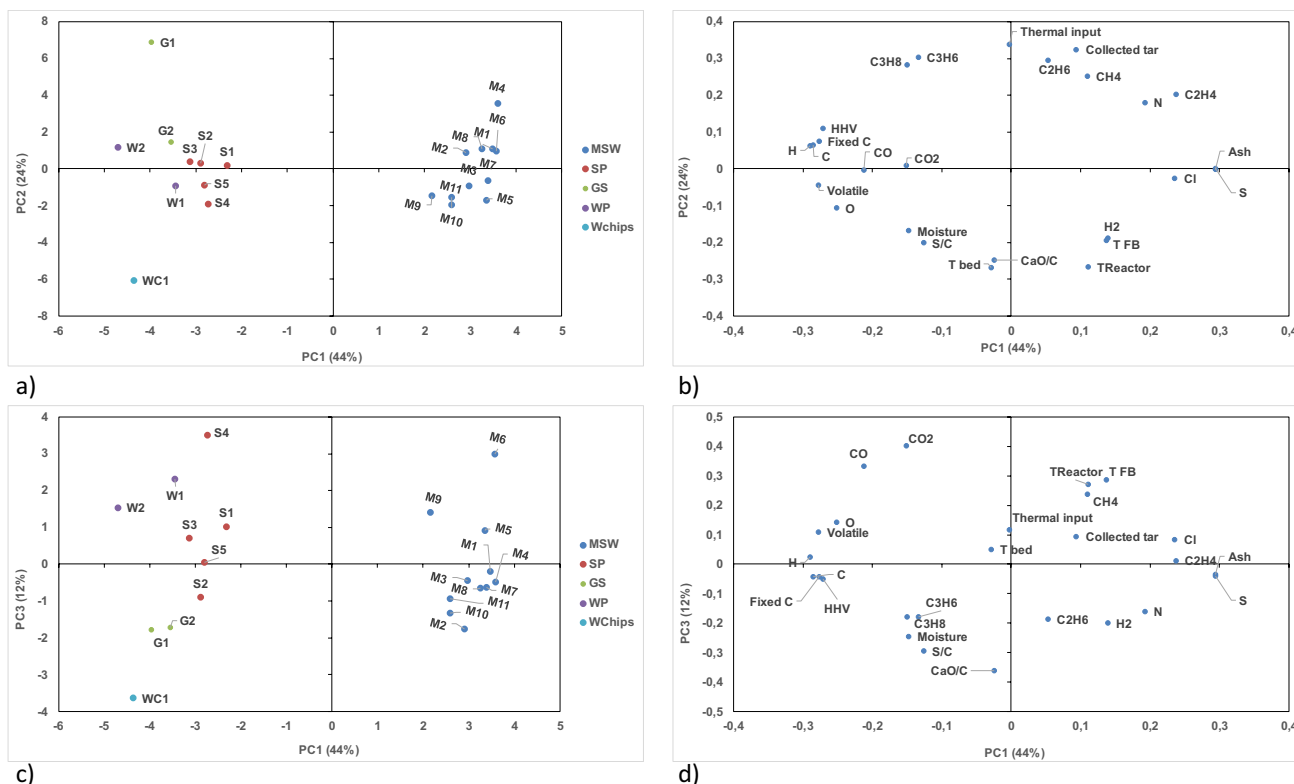
Fig. 2 Equation: PCA results for the whole set of samples, where **a**, **c**, **e**, and **g** represent the scores plot PC1 versus PC2, PC1 versus PC3, PC1 versus PC4 and PC1 versus PC5 and **b**, **d**, **f**, and **h** represent the loading plot PC1 versus PC2, PC1 versus PC3, PC1 versus PC4 and PC1 versus PC5, respectively. Samples are named with the first letter of the biomass feedstock with consecutive numbers. MSW samples: M1-M11; SP samples: S1-S5; GS samples: G1, G2; WP samples: W1, W2 and WChips sample: WC1

This corroborated previous results found for this gasification installation with MSW and straw by this research group [32–34]. What this statistical analysis reinforces is the possibility of acting on operating variables (T_{bed}, CaO/C, $T_{Reactor}$) to improve gas quality by reducing the collected tar, CH_4 and C_2H_6 content correlation between collected tar and light hydrocarbons, indicated their important role in polycyclic aromatic hydrocarbon formation by Diels–Alder pathways and/or hydrogen abstraction-acetylene addition [60, 61]. Rakesh and Dasappa [62], among others, have pointed out that tar formation in biomass gasification is related to two different pathways: (i) the direct release of aromatic structures in lignin devolatilisation and (ii) a mechanism of hydrogen abstraction and acetylene addition.

The loadings plot of the third and first factor (Fig. 2d) provided information regarding which variables were affecting these two factors. In fact, the CO_2 content was positively correlated to the third factor and negatively to CaO/C. This third factor was associated with the carbonation reaction that the sorbent undergoes, which is referred to as 'carbonation reaction' and whose main role is to capture the CO_2 released during the gasification process. This would explain these statistical results which prove that the higher the CaO/C ratio,

the lower the CO_2 content released in the gasification process because this CO_2 is mainly captured by the sorbent used in the SEG process (CaO to produce $CaCO_3$). This process is also explained through the carbonation reaction (Eq. 11) which is produced in this SEG process [34]. The proximity between CO_2 and CO would also indicate a similar behaviour with a positive correlation.

With regard to the fourth factor (Fig. 2f), a moderate negative correlation between T_{bed} and the CH_4 content was observed. A moderate positive correlation between T_{bed} and H_2 content in the product gas was obtained in the fifth factor (Fig. 2h), whereas the moisture of the biomass feedstock was negatively correlated to this factor. However, because the variance explained by the third, fourth and fifth factors was quite low, some variables such as H_2 and S/C were correlated not only to one factor, but also to two or even more factors, highlighting the complexity of the gasification process.

Taking into consideration the Fig. 2a related to scores plot, it was possible to distinguish two different groups of samples classified as a function of the biomass. Firstly, all experiments carried out with MSW as biomass feed (right side of the graph), which were characterised by a high ash and sulphur contents compared to the other biomasses (left
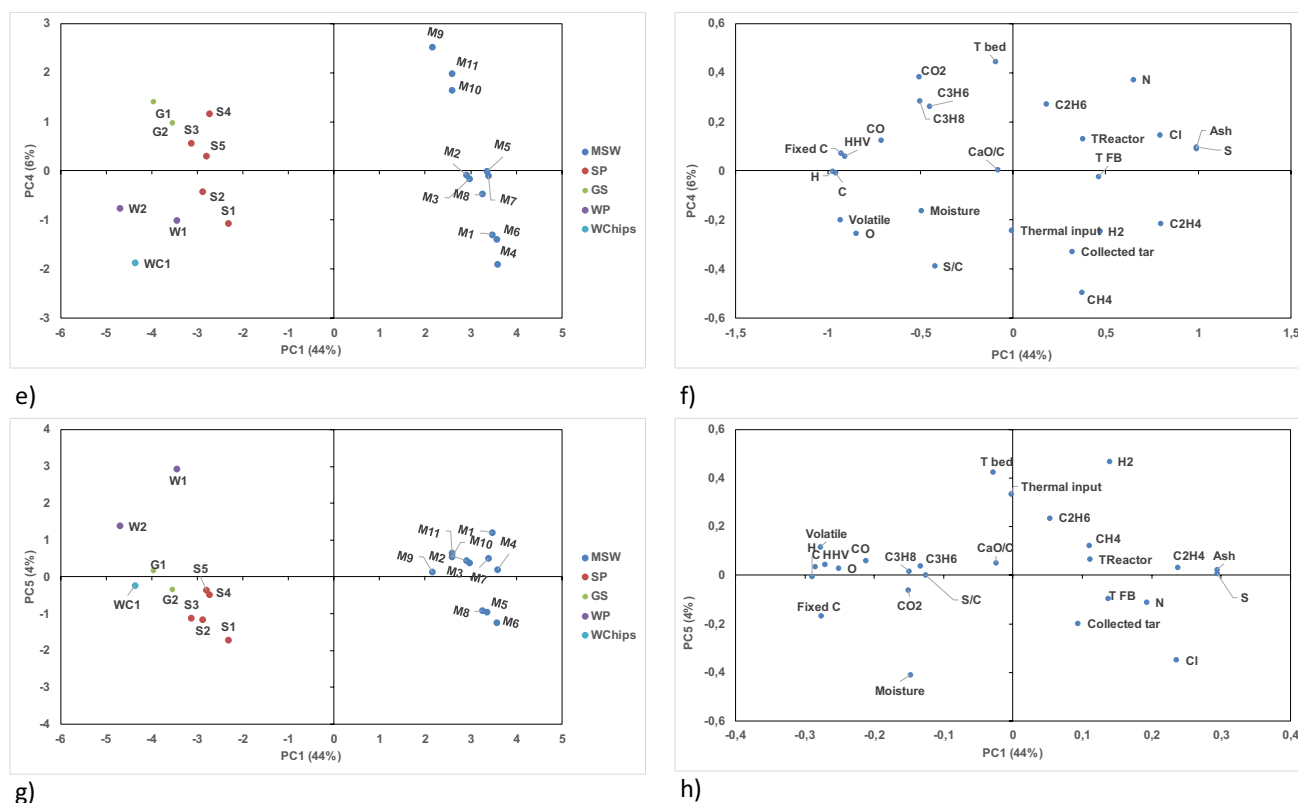
e)

f)

g)

h)

Fig. 2 (continued)

side of the graph). This distinction between groups of samples was observed for the four scores plots of the PCA analysis (Fig. 2a, c, e and g). A possible explanation could be found in the different nature of the biomass. While MSW cannot be considered as 'a natural biomass' because also contains other components at low proportion such as plastic, cardboard, etc., samples represented on the left side of the graph are considered as natural biomass (agricultural and forestry residues of lignocellulosic biomass).

The quality of the outlet gas product for these samples is mainly depending on the gasification conditions. According to the abovementioned, this kind of biomass, MSW, would promote the formation of light hydrocarbons, especially C_2H_4 for those experiments that would be close to C_2H_4. In fact, sample M4 showed the highest concentration of C_2H_4 (6.4%) and CH_4 (Fig. 2a, b) (see Table 4). However, this sample also favoured the formation of tar with the highest content of collected tar for all MSW samples (62.3 g/m^3) due to its proximity to this variable.

These samples could be also distinguished according to the composition of the outlet gas due to their proximity to some variables associated to PC1 and PC4 (Fig. 2b, f). The group at the bottom comprised samples with low CH_4 contents (i.e. less than 8.0 vol%) (M9, M10, M1) (Fig. 2e, 2f) and low tar content, whereas the group at the top (M1,

M2, M3, M4, M5, M6, M7, M8) were characterised by a high content in CH_4 (8.9 to 14.9 vol%) and C_2H_4 (4.0 to 5.3 vol%) and collected tar (Table 4). As mentioned previously in the first factor, those experiments performed at low T_{bed} (M1–M8), in particular, lower than 680 ℃ favoured the CH_4 formation; on the contrary, those carried out at T_{bed} higher than 680 ℃ (M9, M10, M11) promoted the CO formation (Table 4).

In fact, current research [33] carried out by these authors have shown that for MSW feedstock samples, devolatilisation of biomass seems to be the main process occurring at T_{bed} lower than 680 ℃ (including CH_4 formation), whereas at T_{bed} higher than 680 ℃, solid char gasification reactions become more significant, which would favour the increase in CO (Eq. 10), CO_2 (Eq. 9) and H_2 contents (Eq. 7) in the product gas. Taking into account Eq. 7, the steam cracking reaction favoured the formation of CH_4 at the expense of CO at high temperatures. These results have also been remarked by other authors [21] where it has also shown that gasification from feedstock with higher content of S and lower content of O produced higher CH_4 content and lower CO_2.

On the other hand, the other group of samples (GS, SP, WP and WChips) placed on the left side of the graph (Fig. 2a) were mainly distinguished by their high volatile matter, H, C and fixed C contents, and their HHV

Table 4 Mean gasification parameters and mean content of the outlet gas composition as a function of biomass feed. For the total of experiments, the minimum and maximum of each variable are also included in brackets. Specific samples have been remarked for each biomass; n = number of samples

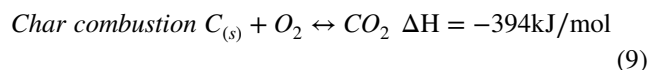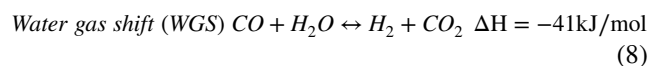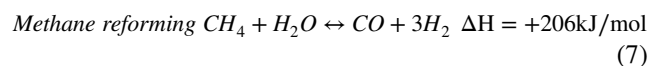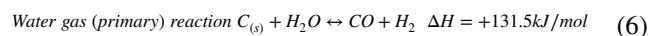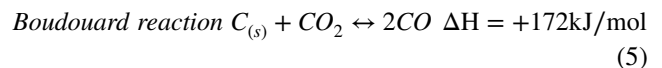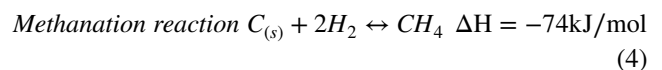|	MSW				SP			GS		WP	WChips
	Mean $n=11$	M4 $n=1$	M1-M8 $n=8$	M9-M11 $n=3$	Mean $n=5$	S1 $n=1$	S4 $n=1$	Mean $n=2$	G1 $n=1$	Mean $n=2$	WC1 $n=1$
T_{bed} [°C]	675 (617–735)	635	660 (617–696)	716 (704–735)	678 (645–735)	645	735	651 (630–672)	630	691 (665–717)	705
CaO/C	0.508 (0.19–1.22)	0.28	0.372 (0.19–0.55)	0.870 (0.28–1.22)	0.422 (0.20–0.81)	0.30	0.20	0.291 (0.25–0.33)	0.25	0.327 (0.29–0.36)	1.81
S/C	1.27 (0.98–1.39)	1.39	1.31 (0.98–1.38)	1.18 (1.04–1.26)	1.35 (1.26–1.47)	1.35	1.31	1.47 (1.11–1.82)	1.11	1.47 (1.37–1.56)	2.99
Collected tar [g/m³N]	36.9 (14.8–62.2)	62.2	42.9 (27.5–39.8)	20.9 (14.8–24.0)	30.4 (23.2–38.2)	38.2	23.2	42.7 (28.5–57.0)	57.0	28.2 (26.7–29.6)	6.70
CO [vol%]	3.83 (2.15–8.52)	2.97	3.32 (2.15–6.33)	5.20 (3.52–8.52)	9.11 (6.63–13.8)	7.44	13.8	6.32 (5.35–7.30)	7.29	11.1 (8.66–13.5)	5.47
CO_2 [vol%]	6.59 (4.27–10.2)	5.28	5.98 (4.27–9.67)	8.21 (6.88–10.2)	9.27 (7.0–12.3)	8.19	12.3	8.52 (7.70–9.34)	9.34	9.83 (9.75–9.91)	5.43
CH_4 [vol%]	10.6 (7.26–14.9)	14.9	11.7 (8.89–14.9)	7.57 (7.26–7.92)	9.18 (8.4–10.7)	10.7	9.45	8.98 (7.22–10.74)	10.7	11.1 (10.84–11.37)	4.67
H_2 [vol%]	40.0 (33.7–44.9)	36.9	40.2 (33.7–44.9)	39.51 (34.3–42.6)	32.0 (29.2–35.1)	29.7	29.2	31.2 (30.63–31.81)	30.6	38.0 (31.85–44.07)	45.6
C_2H_6 [vol%]	1.13 (0.56–1.70)	1.69	1.10 (0.56–1.69)	1.25 (1.17–1.30)	0.882 (0.81–0.95)	0.81	0.81	1.53 (1.18–1.87)	1.87	0.958 (0.82–1.09)	0.492
C_2H_4 [vol%]	4.23 (2.93–6.40)	6.40	4.69 (3.94–6.40)	2.98 (2.93–3.03)	2.14 (1.74–2.65)	2.65	2.48	3.10 (2.55–3.65)	3.65	2.05 (1.95–2.15)	0.818

(Fig. 2b). In this group of samples, the formation of CO and CO_2 seemed to be favoured due to their proximity to these variables in the PC1. This is mainly fulfilled for SP samples, being the S4 the one with the highest concentration for these compounds, as can be observed from the scores plot 2a and c (Fig. 2a, c). The influence of the biomass properties on the gas composition has been remarked by other authors [20], finding that the C and H contents and the HHV of the biomass promote the CO and CH_4 production.
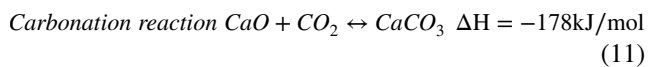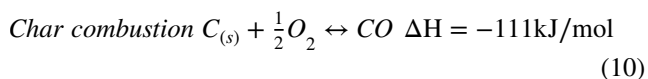
Other samples which could be differentiated and compared were the samples S1 and S4 corresponding to SP (Fig. 2a). In fact, the distinction between these two samples was more remarked in Fig. 2c, d. For both samples, the rest of the gasification parameters were kept constant (S/C, thermal input, CaO/C ratio) with the exception of the T_{FB} and $T_{reactor}$. The higher the T_{FB} and the $T_{reactor}$, the higher the CO and CO_2 contents as observed for samples S4 versus S1 (Table 4) due to its proximity to these variables (Fig. 2c, d). This could also be explained through the thermodynamic equilibrium of the carbonation reaction as the CO_2 equilibrium pressure increased with the increasing of the reaction temperature (Eq. 11).

The sample of GS at the top left side (sample G1) (Fig. 2a) also presented the highest concentrations of C_3H_8 and C_3H_6 (Fig. 2b) according to its proximity to these

variables compared to sample G2. Both samples, differed in the gasification temperature, being the lowest for sample G1. This reflected again what was previously explained by the correlation found between T_{bed} and the formation of light hydrocarbons.

The reactions involved in the gasification process would also help to understand the complexity of predicting variables such as CO, CO_2, H_2 and CH_4 contents, which were correlated with different factors according to the PCA analysis. When a biomass particle enters inside the gasification reactor, different reactions and processes (Eqs. 4–11) occur.

$$\text{Methanation reaction } C_{(s)} + 2H_2 \leftrightarrow CH_4 \quad \Delta H = -74 \text{kJ/mol} \tag{4}$$

$$\text{Boudouard reaction } C_{(s)} + CO_2 \leftrightarrow 2CO \quad \Delta H = +172 \text{kJ/mol} \tag{5}$$

$$\text{Water gas (primary) reaction } C_{(s)} + H_2O \leftrightarrow CO + H_2 \quad \Delta H = +131.5 \text{kJ/mol} \tag{6}$$

$$\text{Methane reforming } CH_4 + H_2O \leftrightarrow CO + 3H_2 \quad \Delta H = +206 \text{kJ/mol} \tag{7}$$

$$\text{Water gas shift (WGS) } CO + H_2O \leftrightarrow H_2 + CO_2 \quad \Delta H = -41 \text{kJ/mol} \tag{8}$$

$$\text{Char combustion } C_{(s)} + O_2 \leftrightarrow CO_2 \quad \Delta H = -394 \text{kJ/mol} \tag{9}$$

Char combustion $C_{(s)} + \frac{1}{2}O_2 \leftrightarrow CO$ $\Delta H = -111 kJ/mol$

(10)

Carbonation reaction $CaO + CO_2 \leftrightarrow CaCO_3$ $\Delta H = -178 kJ/mol$

(11)

Firstly, the fuel drying and a primary devolatilisation [62] take place, releasing volatile compounds from the biomass. Rakesh and Dasappa [62], among others, have pointed out that when this process is carried out in the absence of air, pyrolysis products have a composition that mainly depends on biomass composition and not on the gasifying agent used. The released volatile compounds will react with the different gases present in the reactor (H_2O when the gasifying agent is steam, CO_2 and H_2) (Eq. 6). At this point, methane reforming and water gas shift reactions (Eqs. 7 and 8) take place with their different enthalpy changes, with CH_4 being one of the volatile compounds released.

Afterwards, the solid carbon in the form of char undergoes gasification reactions with different gases (H_2, CO_2, CH_4, O_2) (Eqs. 4, 5 and 6) so that char combustion (exothermic reaction) (Eqs. 9 and 10) occurs. In addition, a reaction takes place between the CaO and the CO_2 formed in the gasification reactions to form $CaCO_3$ (Eq. 11). When gasification is led to produce an H_2-rich gas, endothermic reactions are also involved. This competition between different reactions was also reflected in the PCA analysis, where H_2 was affected not only by one specific factor, but also by different factors, as shown in Fig. 2b–d, indicating the complexity of predicting this gas in the gasification process. On the contrary, the exothermic reactions would promote the formation of CH_4, which would be mainly encouraged by thermal power and discouraged by the increase in T_{bed} and $T_{reactor}$.

The positive correlation between T_{bed} and CaO/C and negative correlation to the thermal power found in the PCA analysis could also be explained by the energy balance in this particular gasifier. For the same thermal power (maintaining the biomass feed constant), an increase in CaO/C involves a decrease in T_{bed} because the sensible heat required to heat the CaO stream is higher. On the other hand, by keeping CaO/C constant and if thermal power decreases, the gasification reaction becomes less endothermic, achieving a higher T_{bed} (because there is less biomass feed and because CaO decreases).

Understanding the relationship between variables can be a great advantage when working with experimental plants of considerable dimensions in which different variables are involved and their control is relatively difficult. Once the interrelations between the different parameters involved in the gasification process were studied, the statistical analysis focused on modelling the gas quality, searching for particular final product compositions. In some cases, the aim was to optimise the composition of the outlet gas trying to maximise (for example) CH_4 content when the production of synthetic natural gas was the objective downstream, while minimising the amount of tars produced in the gas product. In attempting to meet these requirements, PLS regression was used to model each individual parameter of the outlet gas, in particular collected tar, CH_4 and C_2H_4 contents, without taking into account the composition of the other outlet gases.

3.2 Partial least square regression models

A total of three PLS models were obtained for each one of the three independent variables: collected tar, CH_4 and C_2H_4 with the main aim of predicting these variables avoiding a high number of gasification experiments that are really costly in time and expenses. It is important to point out that the main implications of the model are related to the availability of compiling enough experimental data and variables that can be reduced to a new set of explanatory variables just maintaining the same information of the original data. The simplicity of the model, as it is not necessary to include kinetic reactions and the ability of handling more descriptor variables than compounds, constitute one of the main advantages of applying PLS. As disadvantages, it is worth mentioning that only a limited number of available experiments have been used to predict the model.

More experimental data would be required for this present model in order not only, to improve the prediction in all the range of the model but also the accuracy, especially with those biomasses that have only been represented by a low number of experiments such as GS, WP and WChips. The wider the range of gasification conditions and the feedstock composition, the more robust the model for predicting the outlet gas composition.

3.2.1 Partial least square regression for collected tar

It is well known that tars are undesirable subproduct of the gasification process that requires minimisation owing to potential problems of fouling and soot formation on the cold surfaces downstream of the gasifier. In this sense, a PLS regression model was produced in order to determine which parameters mainly affected collected tar as a function of the independent variables (17) associated with the gasification plant and biomass used (Fig. 3).

A total of six factors explained 91% of the variance, with the first factor explaining (61%), more than half of total variance (Table 5). As found in the PCA analysis, and as a function of the PLS scores (Fig. 3a), it was possible to distinguish between different groups of samples according to the different biomass feedstocks based on their nature and their proximate and ultimate composition. Separate groups were obtained between the MSW feedstock and SP, GS and the WP. But not only the nature of the biomass fed affected

these samples, but also the gasification conditions. In fact, the lowest average tar content was obtained for the WChips sample, the result of the favourable conditions in which the gasification process took place (high S/C and CaO/C and moderately high T_{bed}) in comparison with the GS and MSW feedstocks (see Tables 2 and 4), which showed the highest tar content.

For collected tar, all variables included in the two outer circumferences of the correlation loadings plot (Fig. 3b) were found to influence tar content. However, it was observed that the statistically significant variables were thermal power, T_{bed} and volatile matter content, related to gasification conditions and biomass feedstocks. The Beta coefficients, which provide information on the most important variables affecting the predicted variable are also shown in Fig. 3c, where in red colour appear those coefficients statistically significant at 95% level according to Martens' test. While thermal power was positively correlated to collected tar, T_{bed} and volatile matter content were negatively correlated (Fig. 3b, c). In this way, these model results confirmed previous results shown in this research for the PCA analysis where the increase of the T_{bed} decreased the tar content, despite the PCA analysis considers all variables in the same way as input and PLS model includes the variable to be modelled as dependent variable.

In fact, it has been proved in the literature that high temperatures (bed temperatures) produce a reduction in tar content [63, 64] mainly owing to the fact that thermal cracking and reforming reactions become more favoured. Wolfesberger et al. [65], among others, found decreasing amounts of tar at higher temperatures during the gasification of biomass in a 100 kW steam-blown dual fluidised-bed gasifier pilot plant. Previous studies [33, 34] also pointed out the important role that T_{bed} had on the tar content obtained for the MSW samples, resulting in lower concentrations of tars (gravimetric and collected tar) as T_{bed} increased, indicating that this gasification parameter was quite relevant for decreasing tar content.

These results were also corroborated by the results previously mentioned in the PCA section where collected tar was positively correlated to thermal power and negatively correlated to T_{bed} and $T_{reactor}$. The adequacy of the PLS model was evaluated using the coefficient of determination $R^2 = 0.91$, which reached a value of 0.80 for the validation between the experimental and the modelled results (Table 5).

In addition, the RMSEC was obtained in the validation samples as well as the average overall error (OE). For the collected tar, $RMSE_{CV} = 6.62$ (g/m³N) and average $OE_{CV} = 19.8\%$. By applying this model to a new set of samples not used in the modelling, it was found that the average $OE_{ext} = 26.2\%$ with $RMSEP_{ext} = 5.74$ (g/m³N) (Table 5 and Table S2,
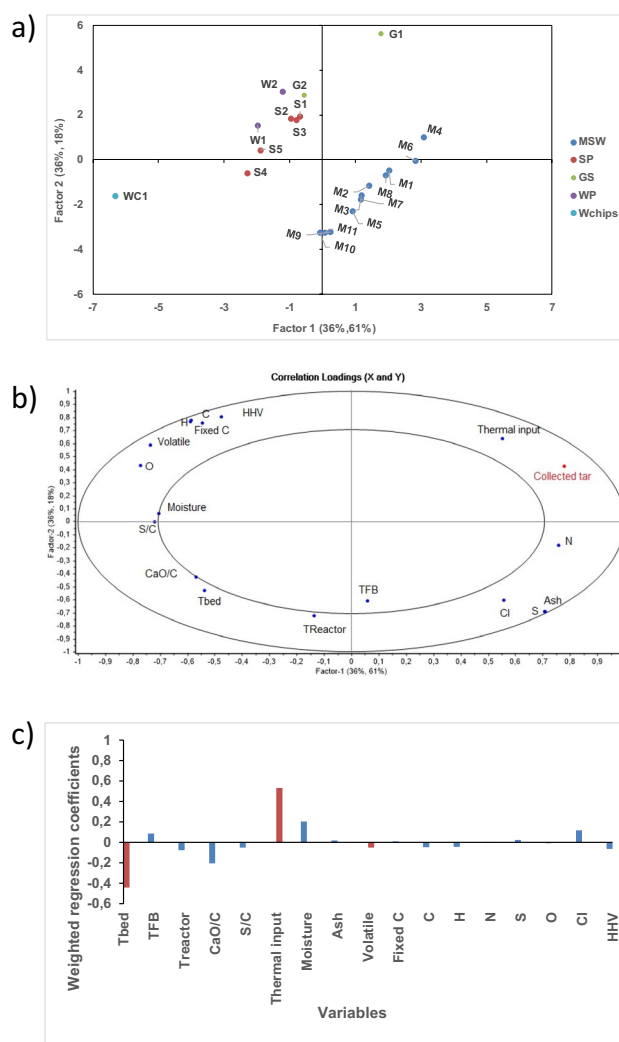

Fig. 3 PLS results for collected tar. **a** Factor 1 versus factor 2 scores plot, **b** factor 1 versus factor 2 correlation loadings plot and **c** weighted regression coefficients plot (bars in red are statistically significant variables)

Supplementary Information). According to Dellavedova et al. [22], a good agreement between the experimental and the prediction results can be achieved when the average OE is within ± 30%. Therefore, this variable, which is time consuming, could be modelled through this PLS model by knowing the gasification parameters and the biomass composition.

Compared to the literature, high coefficients of determination with R^2 higher than 0.90 and 0.98 were found by different authors [28, 66] by using artificial networks ANN to predict tar from biomass gasification. Despite the essence of this model is the same that PLS, the ANN has also got limitations related to an oversimplified approach, where some parameters are assumed to be constant. It requires a very large number of samples and it has got

Table 5 Parameters obtained in the PLS regression model of collected tar, CH_4 and C_2H_4 as a function of the biomass properties and the gasification conditions (17 variables) for the calibration, cross-validation and external validation samples ($RMSEC$, $RMSE_{CV}$ and $RMSEP_{ext}$ in g/m^3N for the collected tar, in vol% for CH_4 and C_2H_4 contents; LV = latent variables)

Dependent variable		Variance		Calibration		Cross validation			External validation	
	LV	X	Y	R^2	RMSEC	Q^2	$RMSE_{CV}$	OE_{CV}	$RMSEP_{ext}$	OE_{ext}
Collected tar [g/m^3N]	6	89	92	0.91	4.08	0.80	6.62	19.8	5.74	26.2
CH_4 [vol%]	5	89	96	0.96	0.51	0.85	0.98	9.90	0.92	7.9
C_2H_4 [vol%]	2	72	86	0.86	0.51	0.80	0.61	18.5	0.34	16.3

'black box' nature in comparison with other models. Other authors have mainly proposed models based on kinetic models finding overestimation of the tar content or even the models are not able to foresee the tar content in the product gas [35, 66–69]. The lack of kinetic data and the theoretical comprehension of the tar reaction during gasification point out the difficult task of modelling this variable. For this reason, these modelling results could be considered as promising for this SEG plant.

3.2.2 Partial least square regression for methane

A total of five factors were used to model the CH_4 content in the product gas by PLS in order to explain 96% of total variance, where the dependent variable was CH_4 content and a total of 17 variables associated with the gasification parameters and the biomass feed were introduced as independent variables. The first factor explained 51% of total variance and the following factors explained 23%, 16%, 4% and 2% of total variance, respectively (Fig. 4a, 4b).

As mentioned in the PCA analysis, the scores plot between the factors 1 and 2 allowed two different groups of samples to be distinguished based on the type of biomass (Fig. 4a), as pointed out by tar content. On the lower right-hand side of the graph, all samples from MSW biomass with a high ash content constituted one group. On the left side of the graph, the other biomasses formed another group comprising SP, GS and WP. These samples were characterised by their high content in volatile matter, fixed C, C and H, and their HHV. An exception was the sample corresponding to WChips with the highest moisture content and tested under conditions of high S/C and CaO/C, which produced the lowest CH_4 content (see Tables 1, 2, and 4).

According to the CH_4 content for the different biomass feed, a decreasing order was obtained for WP, MSW, SP and GS under similar conditions of S/C, CaO/C and T_{bed} (Table 4). MSW samples could be distinguished according to gasification parameters (Fig. 4a, b). In fact, thermal power had an opposite effect on CH_4 content to that of T_{bed} and CaO/C (Fig. 4b), and the samples were located at different

points of the graph, with those close to these variables being the ones with the highest values for these gasification parameters, as also remarked in the previous PCA analysis.

When the weighted regression coefficients that provide information on the most important variables affecting the predicted variable were plotted (Fig. 4c), it was found that thermal input was the variable exerting the most influence on CH_4 production as it was positively correlated to CH_4 and negatively correlated to the CaO/C, C, H, N and HHV of the biomass feed. The Martens' test allowed assessing which variables were statistically significant ($p < 0.05$) for beta coefficients (Fig. 4c, variables in red colour).

It was observed that thermal input, CaO/C ratio, C, H, N and HHV were variables statistically significant at 95% level. Although the most influential variable, thermal power, is a variable related to this particular gasification process, the CaO/C ratio and other variables associated with the nature of the biomass feed exerted influence on CH_4 production (Fig. 4b). This contrasts with the findings of authors such as Dellavedova et al. [22] who studied the influence of the gasification conditions and biomass variables on the dependent variables such as CO, CO_2, CH_4 and H_2 contents and found that although the nature of the syngas produced was mainly dependent on the gasification process (equivalent ratio (ER), steam to biomass ratio (SB), T and type of reactor), the biomass variables presented weak correlations with the syngas composition: HHV, C, particle size and O.

However, based on the findings in Fig. 4, it has been found that CH_4 production is affected not only by the gasification variable but also by the nature of the biomass used. In fact, Gil et al. [20], among others, also concluded that gasification parameters and biomass type affected the gasification reaction by applying PCA and hierarchical cluster analysis (HCA). Mirmoshtaghi et al. [25] also indicated that the carbon content was a parameter influencing the biomass gasification process. For the CH_4, it was obtained the best $Q^2 = 0.85$ with an average overall error of 9.9% (Table 5) indicating a good agreement between the experimental and the predicted values. An overall error of 7.9% with a $RMSEP_{ext} = 0.92$ vol % was obtained for the new set of gasification experiments not used in the model (Table 5; Table S2 Supplementary Information).
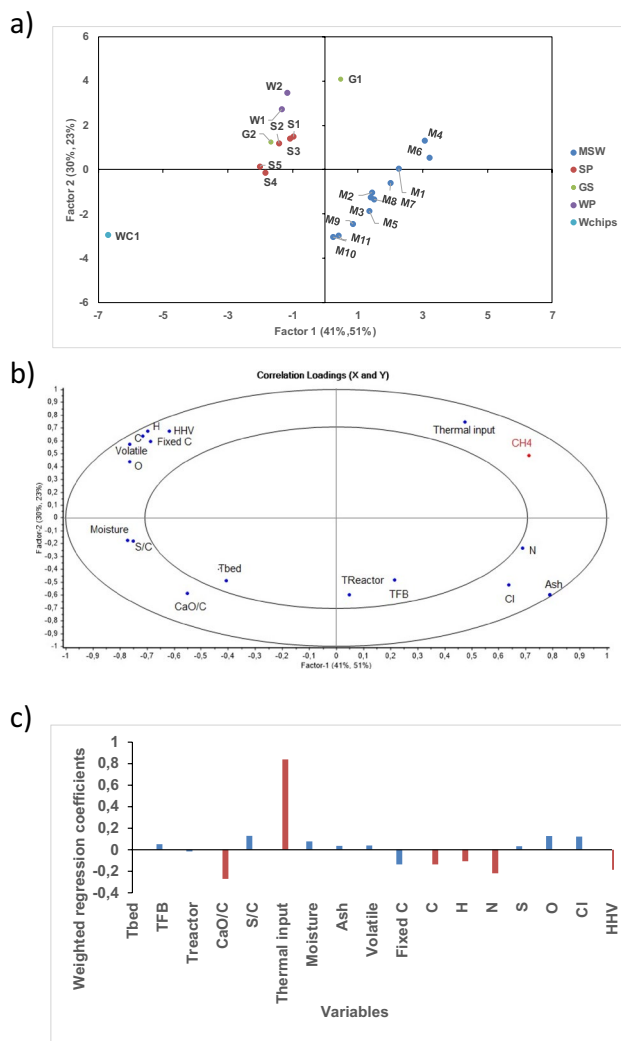
Fig. 4 PLS results for methane. **a** Factor 1 versus factor 2 scores plot, **b** factor 1 versus factor 2 correlation loadings plot and **c** weighted regression coefficients plot (bars in red are statistically significant variables)

According to previous studies [33], when the main aim of biomass gasification is the production of synthetic biofuels such as CH_4, an M-module of 3 (defined as the molar ratio $(H_2\text{-}CO_2)/(CO+CO_2)$) is preferred. For the MSW samples, it was found that a decrease in CH_4 content with increasing gasification temperature was due to steam cracking and reforming reactions into CO, CO_2 and H_2. However, temperature was not directly included among the most influential parameters on CH_4 production in the modelling of this parameter; instead, thermal power was positively correlated to CH_4. It is true, nonetheless, that thermal power is indirectly related to the gasification temperature in the gasifier shown in Fig. 1 since an increase in thermal power (i.e. an increase in the biomass input into the gasifier) involves a decrease in the gasification temperature reached. Therefore, these statistical

results on the concentration of CH_4 confirm experimental results previously obtained when the reactions involved in the gasification process were taken into account [33].

3.2.3 Partial least square regression for ethylene

A two-factor solution, explaining 86% of the total variance, was obtained to model C_2H_4 production by PLS regression modelling (Fig. 5a, b). The first factor (72% of total variance) was positively correlated to ash, S and N, and negatively correlated to the volatile matter, fixed C, HHV and C, H and O contents of the biomass (Fig. 5b). The second factor (Fig. 5b) was associated with thermal input and negatively correlated to CaO/C, T_{bed}, $T_{reactor}$ and T_{FB} and confirmed the results also shown in the PCA analysis. Similar results were also found by Mirmoshtaghi et al. [25], who determined that the C_2H_4 content in the product gas obtained in a CFB gasifier had a significant negative correlation with the load, heating value and H content of the biomass and a positive correlation with the moisture and ash content of the biomass.

When the weighted regression coefficients were plotted (Fig. 5c), it was observed that ethylene content was positively correlated to thermal power and the N, ash and S contents of the biomass used, and negatively correlated to the CaO/C and O, volatile matter, H, fixed C and C contents of the biomass used. These were the most influential variables, statistically significant at 95% level, for C_2H_4 content. As mentioned previously for CH_4, these results also corroborated previous studies [33, 34] on the formation of non-condensable hydrocarbons (C2-C4) whose content decreased with increasing gasification temperature for MSW samples, highlighting the indirect relation between gasification temperature and thermal power obtained in this research. In addition, the role of CaO/C noted in these previous studies was also reflected in this statistical study as having a negative correlation on ethylene production.

There was found to be discrimination for the ethylene content between the different biomass samples (Fig. 5a) as a function of the nature of the biomass feed in the gasification process. The MSW samples are represented on the right side of the graph with high-ash content, N and Cl, whereas the remaining samples associated with other biomasses present high volatile matter, fixed C, C, H and O contents in the biomass feed, as previously mentioned in the section on PCA, confirming the important role of particular variables on C_2H_4 production. Based on these effects, the MSW samples produced higher amounts of C_2H_4 compared to the other biomasses (Table 4), such as WChips and WP, which had the lowest concentrations. These results could have been predicted based on the proximate and ultimate analyses in Table 1 for the different biomasses according to results shown in PCA analysis.

Finally, a good correlation was also achieved with $Q^2=0.80$ for validation, with $RMSE_{CV}=0.61$ vol % and with an overall error of
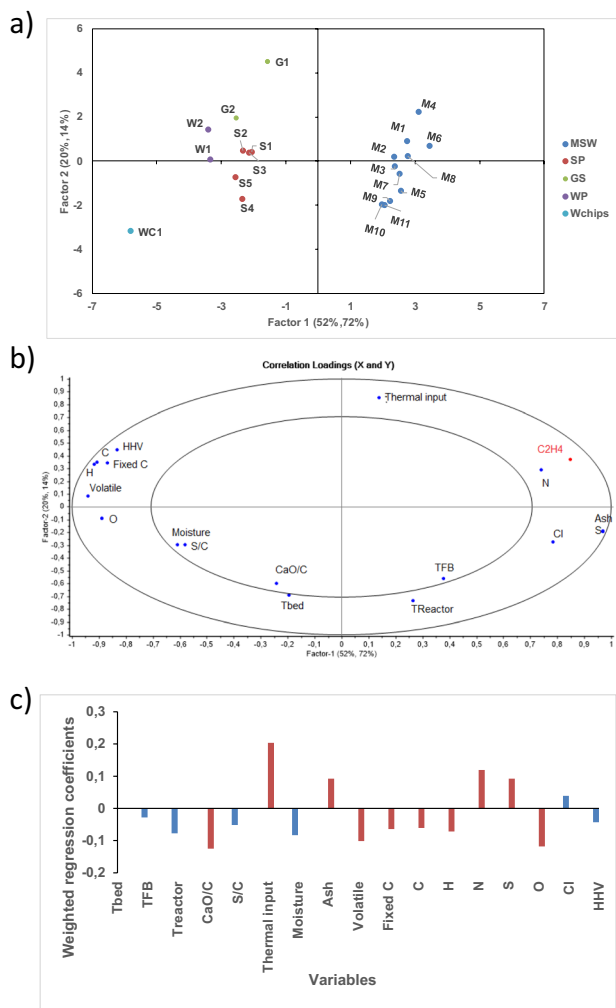
Fig. 5 PLS results for ethylene. **a** Factor 1 versus factor 2 scores plot, **b** factor 1 versus factor 2 correlation loadings plot and **c** weighted regression coefficients plot (bars in red are statistically significant variables)

18.5% (Table 5). For the new set of gasification samples not used in the model, the overall error for C_2H_4 prediction was 16.3% with $RMSEP_{ext} = 0.34$ vol % (Table 5, Table S2, Supplementary Information), allowing modelling of the C_2H_4 content in the outlet gas always working in the same range of studied variables. For different biomasses and different ranges of variables, new models should be checked and validated. However, this first attempt seems to be quite promising in order to better understand the biomass gasification process and to model the composition of the outlet gas.

4 Conclusions and prospects

The PCA study provided valuable information concerning the relationships between the gasification parameters in an SEG process, the biomass feedstock used and the composition of the outlet gas, indicating which variables had the greatest effect on this composition.

It could be concluded as follows:

- Five factors explained most of the variance in the PCA analysis. From these factors, the nature of the biomass used had an important role in the SEG process explaining most of the variance in the PCA analysis (44%). The CO and the C_2H_4 contents in the outlet gas composition were also correlated to this factor although in an opposite way. Samples associated to different biomasses were classified into two main groups so that those biomasses with high volatile matter, C, fixed C, C and H (GS, WP, SP) seemed to promote the formation of CO whereas those with high ash and sulphur contents favoured the C_2H_4 formation as the case of MSW samples.
- In relation to gasification parameters, thermal power, average reactor temperature and solid bed temperature, as well as the CaO/C ratio were the variables affecting most the gas outlet composition, highlighting the important role of the sorbent in the gas obtained for a SEG process. An increase of these variables resulted in a decrease of the hydrocarbon composition of the outlet gas (CH_4, C_2H_6, C_3H_6 and C_3H_8) and of the tar content.
- Related to the outlet gas composition, the CH_4, C_2H_6 and collected tar contents were positively correlated between them and found to be negatively correlated to the bed temperature and the CaO/C ratio. C_3H_6 and the C_3H_8 were also positively correlated and were mainly dependent on thermal power and negatively associated with the reactor temperature. With regard to C_2H_4 content, it was mainly affected by the ultimate and proximate composition of the biomass feedstock. In fact, C_2H_4 content was enhanced for those biomasses with high ash, S and N contents and low volatile matter, fixed C, C, H and O.

With regard to the PLS analysis, it could be concluded as follows:

- It was possible to model the CH_4 and C_2H_4 contents as well as the tar content in the outlet gas product with internal and external validation in the range of variables used for the model construction, obtaining good results with overall errors for the prediction between 8 and 26% for these three variables.
- The main gasification parameters decreasing the formation of CH_4 and C_2H_4 were the temperature and the CaO/C ratio for this SEG installation.
- The abatement of the tar content was promoted by high-bed temperature, low thermal power and biomass with high-volatile matter content.

This study constitutes a first approximation to assess the influence of the biomass used, the gasification parameters

and the outlet gas composition in a SEG process by multivariate statistical analysis. Furthermore, these results seem to corroborate previous results on biomass gasification taking into account the different reactions involved in the process, where one of the most influential variables on gas composition and tar content was gasification temperature in addition to bed material.

This research has been performed and validated for this SEG experimental installation using CaO as a bed material and steam as gasification agent. The modelling is limited to the range of the experimental conditions shown in this paper and to this particular SEG process. However, these PLS models can be considered as an alive model so that they can be improved by including more experimental data obtained from this experimental installation, i.e. other sorbents and/or gasification agents, biomasses, increasing in this way the accuracy of the models.

It should also be noticed that this modelling has also taken into account the tar content prediction, which is a variable that requires more effort in sampling and analysis of different organic compounds. Nevertheless, and because of the considerable dimensions of this installation, experimental data are time-consuming. Thus, any progress with modelling, no matter how small is, will be really helpful to reduce the number of experiments for scaling this SEG process.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. IPCC (2018) Summary for Policymakers. In: Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, Pirani A, Moufouma-Okia W, Péan C, Pidcock R, Connors S, Matthews JBR, Chen Y, Zhou X, Gomis MI, Lonnoy E, Maycock T, Tignor M and Waterfield T (eds). Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. World Meteorological Organization, Geneva, Switzerland, pp 32.

2. Tursi A (2019) A review on biomass: importance, chemistry, classification and conversion. Biofuel Res J 22:962–979. https://doi.org/10.18331/BRJ2019.6.2.3

3. Dhanavath KN, Shah K, Bhargava SK, Bankupalli S, Parthasarathy R (2018) Oxygen-steam gasification of karanja press seed cake: fixed bed experiments, ASPEN Plus process model development and benchmarking with saw dust, rice husk and sunflower husk. J Environ Chem Eng 6(2):3061–3069. https://doi.org/10.1016/j.jece.2018.04.046

4. Bridgwater AV (2003) Renewable fuels and chemicals by thermal processing of biomass. Chemical Eng J 91:87–102. https://doi.org/10.1016/S1385-8947(02)00142-0

5. Molino A, Chianese S, Musmarra D (2016) Biomass gasification technology: The state of the art overview. J Energy Chem 25:10–25. https://doi.org/10.1016/j.jechem.2015.11.005

6. Zhang L, Charles C, Champagne P (2010) Overview of recent advances in thermo-chemical conversion of biomass. Energy Convers Manage 51:969–982. https://doi.org/10.1016/j.enconman.2009.11.038

7. Perera SMHD, Wickramasinghe C, Samarasiri BKT, Narayana M (2021) Modeling of thermochemical conversion of waste biomass-a comprehensive review. Biofuel Res J 32:1481–1528. https://doi.org/10.18331/BRJ2021.8.4.3

8. Ian N, Orio A, Aznar MP, Corella J (1996) Biomass gasification with air in an atmospheric bubbling fluidized bed. Effect of six operational variables on the quality of the produced raw gas. Ind Eng Chem Res 5885:2110–2120. https://doi.org/10.1021/ie9507540

9. Skoulou V, Koufodimos G, Samaras Z, Zabaniotou A (2008) Low temperature gasification of olive kernels in a 5-kW fluidized bed reactor for H2-rich producer gas. Int J Hydrogen Energy 33:6515–6524. https://doi.org/10.1016/j.ijhydene.2008.07.074

10. Pio DT, Tarelho LAC, Matos MAA (2017) Characteristics of the gas produced during biomass direct gasification in an autothermal pilot-scale bubbling fluidized bed reactor. Energy 120:915–928. https://doi.org/10.1016/j.energy.2016.11.145

11. Cordiner S, De Simone G, Mulone V (2012) Experimental-numerical design of a biomass bubbling fluidized bed gasifier for paper sludge energy recovery. Appl Energy 97:532–542. https://doi.org/10.1016/j.apenergy.2011.11.024

12. Das B, Bhattacharya A, Datta A (2020) Kinetic modeling of biomass gasification and tar formation in a fluidized bed gasifier using equivalent reactor network (ERN). Fuel 280:118582. https://doi.org/10.1016/j.fuel.2020.118582

13. Caballero MA, Corella J, Aznar MP, Gil J (2000) Biomass gasification with air in fluidized bed. Hot gas cleanup with selected commercial and full-size nickel-based catalysts. Ind Eng Chem Res 1143–1154. https://doi.org/10.1021/ie990738t

14. Campoy M, Go A, Fuentes-cano D, Ollero P (2010) Tar reduction by primary measures in an autothermal air-blown fluidized bed

biomass gasifier. Ind Eng Chem Res 49:11294–11301. https://doi.org/10.1021/ie101267c

15. Esfahani RM, Wan Ab Karim Ghani WA, MohdSalleh MA, Ali S (2012) Hydrogen-rich gas production from palm kernel Shell by applying air gasification in fluidized bed reactor. Energy Fuels 26:1185–1191. https://doi.org/10.1021/ef2010892

16. Kim YD, Yang CW, Kim BJ, Kim KS, Lee JW, Moon JH, Yang W, Yu TU, Lee UD (2013) Air-blown gasification of woody biomass in a bubbling fluidized bed gasifier. Appl Energy 112:414–420. https://doi.org/10.1016/j.apeergy.2013.03.072

17. Miromshtaghi G (2016) Biomass gasification in fluidized bed gasifiers. Modelling and simulation. Thesis November 2016. Mälardalen University Press Dissertations. No.216, Sweden

18. Gil MV, Calvo LF, Blanco D, Sánchez ME (2008) Assessing the agronomic and environmental effects of the application of cattle manure compost on soil by multivariate methods. Bioresour Technol 99:5763–5772. https://doi.org/10.1016/j.biortech.2007.10.014

19. Benedikt F, Kuba M, Schmid JC, Müller S, Hofbauer H (2019) Assessment of correlations between tar and product gas composition in dual fluidized bed steam gasification for online tar prediction. Appl Energy 238:1138–1149. https://doi.org/10.1016/j.apenergy.2019.01.181

20. Gil MV, González-Vázquez MP, García R, Rubiera F, Pevida C (2019) Assessing the influence of biomass properties on the gasification process using multivariate data analyst. Energ Convers Manage 184:649–660. https://doi.org/10.1016/j.enconman.2019.01.093

21. Škrbić BD, Đurišić-Mladenović N, Cvejanov J (2018) Differentiation of syngases produced by steam gasification of mono- and mixed sources feedstock: a chemometric approach. Energ Convers Manage 171:1193–1201. https://doi.org/10.1016/j.enconman.2018.06.060

22. Dellavedova M, Derudi M, Biesuz R, Lunghi A, Rota R (2012) On the gasification of biomass: data analysis and regressions. Process Saf Environ Prot 90:246–254. https://doi.org/10.1016/j.psep.2011.08.001

23. Dahlquist E, Naqvi M, Thorin E, Yan J, Kyprianidis K, Hartwell P (2016) Modeling of wood gasification in an atmospheric CFB plant. Proceedings of the 9th EUROSIM & the 57th SIMS 877 September 12th-16th, 2016, Oulu, Finland, 872–877. https://doi.org/10.3384/ecp17142872

24. Dahlquist E, Naqvi M, Thorin E, Yan J, Kyprianidis K (2017) Comparison of gas quality from black liquor and wood pellet gasification using Modelica simulation and pilot plant results. Energy Procedia 105:992–998. https://doi.org/10.1016/j.egypro.2017.03.437

25. Mirmoshtaghi G, Skvaril J, Campana PE, Li H, Thorin E, Dhlquist E (2016) The influence of different parameters on biomass gasification in circulating fluidized bed gasifiers. Energ Convers Manage 126:110–123. https://doi.org/10.1016/j.enconman.2016.07.031

26. Di Giuliano A, Gallucci M, Malsegna B, Lucantonio S, Gallucci K (2022) Pretreated residual biomasses in fluidized beds for chemical looping gasification: analysis of devolatilization data by statistical tools. Bioresour Technol Reports 17:100926. https://doi.org/10.1016/j.biteb.2021.100926

27. George J, Arun P, Muraleedharan C (2018) Assessment of producer gas composition in air gasification of biomass using artificial neural network model. Int J Hydrogen Energ 43:9558–9568. https://doi.org/10.1016/j.ijhydene.2018.04.007

28. Serrano D, Castelló D (2020) Tar prediction in bubbling fluidized bed gasification through artificial neural networks. Chem Eng J 402:12622. https://doi.org/10.1016/j.cej.2020.126229

29. González-Vázquez MP, García R, Gil MV, Pevida C, Rubiera F (2018) Comparison of the gasification performance of multiple

biomass types in a bubbling fluidized bed. Energy Convers Manage 176:309–323. https://doi.org/10.1016/j.enconman2018.09.020

30. Florin NH, Harris AT (2008) Enhanced hydrogen production from biomass with in situ carbon dioxide capture using calcium oxide sorbents. Chem Eng Sci 63:287–316. https://doi.org/10.1016/j.ces.2007.09.011

31. Heidenreich S, Foscolo PU (2015) New concepts in biomass gasification. Prog Energy Combust Sci 46:72–95. https://doi.org/10.1016/j.pecs.2014.06.002

32. Martínez I, Kulakova V, Grasa G, Murillo R (2020) Experimental investigation on sorption enhanced gasification (SEG) of biomass in a fluidized bed reactor for producing a tailored syngas. Fuel 259:116252. https://doi.org/10.1016/j.fuel.2019.116252

33. Martínez I, Grasa G, Callén MS, López JM, Murillo R (2020) Optimised production of tailored syngas from municipal solid waste (MSW) by sorption enhanced gasification. Chemical Eng J 401:126067. https://doi.org/10.1016/j.cej.2020.126067

34. Martínez Callén MS, Grasa G, López JM, Murillo R (2022) Sorption-enhanced gasification (SEG) of agroforestry residues: Influence of feedstock and main operating variables on product gas quality. Fuel Proc Technol 226:107074. https://doi.org/10.1016/j.fuproc.2021.107074

35. Stark AK, Bates RB, Zhao Z, Ghoniem AF (2015) Prediction and validation of major gas and tar species from a reactor network model of air-blown fluidized bed biomass gasification. Energy Fuels 29:2437–2452. https://doi.org/10.1021/ef5027955

36. Yu J, Smith JD (2018) Validation and application of a kinetic model for biomass gasification simulation and optimization in updraft gasifiers. Chem Eng Process -Process Intensif 125:214–226. https://doi.org/10.1016/j.cep.2018.02.003

37. Gómez-Barea A, Leckner B (2010) Modelling of biomass gasification in fluidized bed. Prog Energy Combust Sci 36:444–509. https://doi.org/10.1016/j.pecs.2009.12.002

38. Yan L, Jim Lim C, Yue G, He B, Grace JR (2016) One-dimensional modelling of a dual fluidized bed for biomass steam gasification. Energy Convers Manag 127:612–622. https://doi.org/10.1016/j.enconman.2016.09.027

39. Nemtsov DA, Zabaniotou A (2008) Mathematical modelling and simulation approaches of agricultural residues air gasification in a bubbling fluidized bed reactor. Chem Eng J 143:10–31. https://doi.org/10.1016/j.cej.2008.01.023

40. Loha C, Chattopadhyay H, Chatterjee PK (2014) Three dimensional kinetic modeling of fluidized bed biomass gasification. Chem Eng Sci 109:53–64. https://doi.org/10.1016/j.ces.2014.01.017

41. Nikoo MB, Mahinpey N (2008) Simulation of biomass gasification in fluidized bed reactor using ASPEN PLUS. Biomass Bioenergy 32:1245–1254. https://doi.org/10.1016/j.biombioe.2008.02.020

42. Neeft JPA (2005) Rationale for setup of impinger train as used in the technical specification of sampling and analysis of tar and particles in the product gases of biomass gasification. Technical background document

43. Goog J, Ventress L, Knoef H, Zielke U, Hansen PL, van de Kamp W et al. (2005) Sampling and analysis of tar and gases particles in biomass producer gases. Technical Report prepared under CEN BT/TF 143 Organic contaminants (tar) in biomass producer gases

44. Jackson JE (1991) A user's guide to principal components. John Wiley & Sons, New York, pp 43. https://doi.org/10.1002/0471725331

45. Jolliffe IT (1986) Principal component analysis. In: Springer Series in Statistics, 1st edn. Springer-Verlag, New York, pp 487. ISBN 0–387–96269–7

46. Škrbić B, Đurišić-Mladenović N, Cvejanov J (2005) Principal component analysis of trace elements in Serbian wheat. JAgric Food Chem 53(6):2171–2175. https://doi.org/10.1021/jf0402577

47. Škrbić B, Szyrwińska K, Đurišić-Mladenović N, Nowicki P, Lulek J (2010) Principal component analysis of indicator PCB profiles in breast milk from Poland. Environ Int 36(8):862–872. https://doi.org/10.1016/j.envint.2009.04.008

48. Hébergera K, Škrbić B (2012) Ranking and similarity for quantitative structure–retention relationship models in predicting Lee retention indices of polycyclic aromatic hydrocarbons. Anal Chim Acta 716:92–100. https://doi.org/10.1016/j.aca.2011.11.061

49. Škrbić B, Cvejanov J, Đurišić-Mladenović N (2015) Chemometric characterization of vegetable oils based on the fatty acid profiles for selection of potential feedstocks for biodiesel production. J Biobased Mater Bioenergy 9:358–371. https://doi.org/10.1166/jbmb.2015.1527

50. Škrbić B, Marinković V, Spaić S (2020) Assessing the impact of combustion and thermal decomposition properties of locally available biomass on the emissions of BTEX compounds by chemometric approach. Fuel 282:118824. https://doi.org/10.1016/j.fuel.2020.118824

51. Pattiya A, Titiloye JO, Bridgwater AV (2010) Evaluation of catalytic pyrolysis of cassava rhizome by principal component analysis. Fuel 89:244–253. https://doi.org/10.1016/j.fuel.2009.07.003

52. The Unscrambler® X v10.3 User manual version 1.0 CAMO SOFTWARE AS Nedre Vollgate 8, N-0158, Oslo, Norway, www.camo.com. Accessed 25 Nov 2021

53. Đurišić-Mladenović N, Škrbić B, Zabaniotou A (2016) Chemometric interpretation of different biomass gasification processes based on the syngas quality: assessment of crude glycerol co-gasification with lignocellulosic biomass. Renew Sust Energ Rev 59:649–661. https://doi.org/10.1016/j.rser.2016.01.002

54. Wold HOA (1966) Estimation of principal components and related models by iterative least squares. In: Krishnaiaah PR (ed) Multivariate Analysis. Academic Press, New York, pp 391–420

55. Škrbić B, Đurišić-Mladenović N, Cvejanov J (2004) Discrimination between linear and non-linear models for retention indices of polycyclic aromatic hydrocarbons in the so-called Lee's scale. Chemom Intel Lab Syst 72(2):167–171. https://doi.org/10.1016/j.chemolab.2004.01.011

56. Veses A, López JM, García T, Callén MS (2018) Prediction of elemental composition, water content and heating value of upgraded biofuel from the catalytic cracking of pyrolysis bio-oil vapors by infrared spectroscopy and partial least square regression models. J Anal Appl Pyrolysis 132:102–110. https://doi.org/10.1016/j.jaap.2018.03.010

57. Gramatica P, Giani E, Papa E (2007) Statistical external validation and consensus modeling: a QSPR case study for KOC prediction. J Mol Graph Model 25:755–766. https://doi.org/10.1016/j.jmgm.2006.06.005

58. Madadizadeh F, Ezati Asar M, Hosseini M (2015) Common statistical mistakes in descriptive statistics reports of normal and non-normal variables in biomedical sciences research. Iran J Public Health. 44(11):1557–1558. Available at: http://ijph.tums.ac.ir. Accessed 01/02/2022

59. George D, Mallery P (2010) SPSS for windows step by step: a simple guide and reference 17.0 Update. 10th Edition, Pearson, Boston.

60. Moldeveanu SC (2009) Pyrolysis of organic molecules. In: Applications to health and environmental issues. 1st Edition 28, p 217. ISBN 9780080931256

61. Kislov VV, Islamova NI, Kolker AM, Lin SH, Mebel AM (2005) Hydrogen abstraction acetylene addition and diels−alder mechanisms of PAH formation: a detailed study using first principles calculations. J Chem Theory Comput 1(5):908–924. https://doi.org/10.1021/ct500491

62 Rakesh N, Dasappa S (2018) A critical assessment of tar generated during biomass gasification -Formation, evaluation, issues and mitigation strategies. Renew. Sust Energ Rev 91:1045–1064. https://doi.org/10.18331/BRJ2016.3.4.3

63. Trubetskaya A, Souihi N, Umeki K (2019) Categorization of tars from fast pyrolysis of pure lignocellulosic compounds at high temperature. Renew Energ 141:751–759. https://doi.org/10.1016/j.renene.2019.04.033

64. Yang H, Wang C, Xu S, Liu R (2019) Biomass gasification over hematite in a decoupled dual loop gasifier. Fuel Process Technol 192:140–146. https://doi.org/10.1016/j.fuproc.2019.04.025

65. Wolfesberger U, Aigner I, Hofbauer H (2009) Tar content and composition in producer gas of fluidized bed gasification of Wood- Influence of temperature and pressure. Environ Proc Sustain Energy 28(3):372–379. https://doi.org/10.1002/ep.10387

66. Rameshkumar R, Mayilsamy K (2014) Prediction of tar and particulate in biomass gasification using adaptive neuro fuzzy inference system. J Intell Fuzzy Syst 27:361–365. https://doi.org/10.3233/IFS-131004

67. Wojnicka B, Ściążko M, Schmid JC (2019) Modelling of biomass gasification with steam. Biomass Convers Biorefinery 11:1787–1805. https://doi.org/10.1007/s13399-019-00575-2

68. Sridhar R (2016) Chemistry models for major gas species estimation and tar prediction in fluidized bed biomass gasification, Massachusetts Institute of Technology (2016)

69. Palma CF (2013) Model for biomass gasification including tar formation and evolution. Energy Fuels 27:2693–2702. https://doi.org/10.1021/ef4004297