Physics-informed framework for predictive modeling of product distributions in steam cracking of heterogeneous polymeric waste

Fuel 429 (2027) 140963

Contents lists available at ScienceDirect

Fuel

journal homepage: www.elsevier.com/locate/fuel

Full Length Article

Physics-informed framework for predictive modeling of product distributions in steam cracking of heterogeneous polymeric waste

Renesteban Forero-Franco [a,*], Teresa Berdugo-Vilches [a], Chahat Mandviwala [a], Nidia Díaz Perez [a], Isabel Cañete-Vela [b], Henrik Thunman [a], Martin Seemann [a]

[a] Department of Space, Earth and Environment (SEE), Division of Energy Technology, Chalmers University of Technology, 41296 Gothenburg, Sweden
[b] Borealis AB, 444 86 Stenungsund, Sweden

ARTICLE INFO

Keywords:
Thermochemical recycling
Steam cracking in Dual Fluidized Beds (DFB)
Plastic waste recycling
Physics-Informed Machine Learning
Carbon Bond Group (CBG) model

ABSTRACT

Thermochemical recycling of heterogeneous plastic waste via steam cracking offers a scalable pathway for carbon circularity, yet predictive control of product distributions remains limited by feedstock variability and complex radical chemistry. This work introduces a physics-informed data-driven model, referred to as the Carbon Bond Group (CBG) model, that describes feedstock–product relationships at a reduced but chemically meaningful level of abstraction. The approach reformulates thermochemical conversion as a constrained transformation between compact structure–based carbon spaces, where feedstocks and products are represented by key carbon bonding environments and linked through a column–stochastic operator matrix enforcing conservation laws. Operating conditions are incorporated through a temperature–dependent parameterization, enabling the mapping to evolve systematically while preserving physical admissibility. The model is trained on a multi–year experimental dataset from a semi–industrial Dual Fluidized Bed steam cracker spanning diverse waste streams and temperatures (720–835 °C). It accurately reproduces product group distributions (R^2 up to 0.98 for COx and aliphatics, and 0.90 for aromatics) and demonstrates robust performance under cross–validation. Analysis of the temperature-dependent conversion matrix reveals chemically consistent carbon redistribution trends, such as increasing routing of oxygen-bound carbon toward COx and enhanced gasification of aromatic carbon at high temperatures, demonstrating the model's ability to provide insights into the system's carbon conversion behavior. By embedding structural descriptors, physical constraints, and operating variability, the CBG model provides an interpretable, scalable surrogate for thermochemical conversion behavior of complex feedstocks. The approach supports predictive feedstock assessment, scenario exploration, and development of reduced–order digital twins for heterogeneous waste steam-cracking systems.

1. Introduction

Plastics have become essential to modern life due to their low cost, versatility, and durability, yet their rapidly increasing production has generated a global waste challenge. Globally, plastics and fibers reached approximately 437 Mt in 2022 and are projected to double up by 2050, while effective global recycling rates remain below 10 %, leaving the majority of waste to landfilling, incineration, or uncontrolled dumping [1–3]. Current recycling systems rely mostly on mechanical recycling, which remelts and reprocesses plastic materials into new products. Although mechanical recycling is a comparatively energy-efficient method, it faces intrinsic limitations: it requires clean, homogeneous feedstocks, demands extensive sorting, and suffers from progressive quality degradation with each cycle [4,5]. As a result, heterogeneous

Abbreviations: Aliph, aliphatics group in the CBG model; Arom, aromatics groups in the CBG model; BFB, bubbling fluidized bed; BTXS, benzene toluene xylenes styrene; C-xO, carbons bonded to oxygen; C-AL, carbons bonded in aliphatics; C-AR, carbons bonded in aromatics; Cell, Cellulose; CBG, carbon bond group; CFB, circulating fluidized bed; CRR, cardboard recycling reject; CV, cross-validation; LOO, leave-one-out; DFB, dual fluidized bed; HTR, high temperature reactor; MAE, mean absolute error; ML, machine learning; MPW, mixed packaging waste; MRP, mechanically recycled PO; MSE, mean squared error; PA, polyamide; PAN, polyacrylonitrile; PE, polyethylene; PET, polyethylene terephthalate; PO, polyolefins; PP, polypropylene; PS, polystyrene; PU, polyurethane; PVC, polyvinyl chloride; Rel-MAE, relative mean absolute error; RMSE, root mean squared error; R^2, coefficient of determination; SPA, solid phase adsorption.

* Corresponding author.
E-mail address: rforero@chalmers.se (R. Forero-Franco).

https://doi.org/10.1016/j.fuel.2026.140963
Received 16 March 2026; Received in revised form 5 August 2026; Accepted 8 August 2026
Available online 18 August 2026

post-consumer waste streams often fall outside the recycling operational window, requiring extensive sorting which leads to large rejected fractions that need to be recycled by other means or incinerated. Moreover, mechanical recycling inherently preserves polymer structures and configurations of today's plastics, which risks locking legacy formulations into future products. These shortcomings underscore the need for complementary and robust strategies capable of processing diverse materials and recovering carbon at the molecular level, enabling production pathways aligned with long-term material transitions.

Thermochemical recycling has emerged as a robust alternative in this context. This approach aims to recover carbon by breaking down polymer chains at high temperatures to produce valuable chemicals [6–9]. Among such technologies, steam cracking in Dual Fluidized Bed (DFB) systems has gained increasing attention as a scalable route for converting mixed polymeric waste into valuable chemical building blocks. In conventional petrochemical applications, steam cracking consists of exposing hydrocarbons (e.g. naphtha) to high temperatures (~800 °C) in the presence of steam, initiating complex free-radical reaction networks that fragment carbon chains into light olefins and other hydrocarbons. When applied to plastic waste, the process generates a spectrum of products including hydrogen, carbon oxides, light olefinic monomers, mono-aromatics, and additional hydrocarbon fractions compatible with existing petrochemical value chains [8–10]. DFB systems enhance the suitability of steam cracking for heterogeneous waste. By supplying heat allothermally through circulating bed material, DFB systems enable stable temperature control, high heat-transfer efficiency, and uniform mixing conditions. These characteristics support rapid heating and robust handling of variable feedstocks, making the technology well suited for maximizing monomer recovery and processing mixed or contaminated plastic streams [9,10].

DFB systems offer an additional advantage of increasing relevance under emerging EU regulatory frameworks for chemical recycling. Since the provision of energy and cracking can occur in two physically separated reactors, any carbon released from auxiliary fuels combustion and the carbon originating from the plastic waste conversion can remain distinct from each other within the process. This enables transparent carbon–mass–balance accounting and facilitates the verification of recycled carbon content, a requirement increasingly emphasized in recent EU initiatives, such as Packaging and Packaging Waste Regulation (PPWR) and Single-Use Plastics Directive (SUPD) [11,12], targeting circularity, advanced recycling certification, and traceable material flows. These requirements strengthen the need for methods capable of physically tracing carbon flows, which is a possibility naturally aligned with the DFB configuration and the carbon bond-based framework presented in this study.

Feed-Products Correlation & Modelling Challenges.

While such reactor designs provide the required operational robustness to process complex feeds, predicting how differences in feed composition translate into specific product distributions remains a central scientific and engineering challenge. Reactions governing thermochemical conversion of heterogeneous plastics are chemically complex. Product distributions depend not only on reactor's operating conditions, such as temperature, steam dilution, and residence time, but also on the intrinsic molecular structure of the feedstock. Under steam cracking conditions, polymer chains undergo free-radical fragmentation, β-scission, hydrogen abstraction, cyclization, aromatization, and oxidation reactions, among others [6,13]. The set of reaction pathways accessible to a carbon atom depend strongly on its local bonding environment, whether it belongs to an aliphatic chain, an aromatic ring, or a heteroatom-containing functional group.

Since waste streams consist of diverse polymers with varying structural motifs, their thermochemical conversion leads to significant variability in product slates, even under identical operating conditions [6,8,14]. Structural differences in feedstocks therefore propagate through reaction networks and manifest as systematic shifts in product families. As a result, establishing predictive feed-to-product relationships is inherently complex, particularly when addressing heterogeneous waste streams under varying reactor conditions.

Modeling thermochemical conversion of heterogeneous polymeric waste presents significant challenges due to the interplay of such complex chemistry in highly variable feedstocks with condition-dependent reactor behavior. Traditional deterministic kinetic models based on detailed free radical mechanisms are a common approach in pyrolysis and steam cracking research. While such models provide valuable mechanistic insight, their dimensionality increases rapidly with feedstock complexity. The number of reactions, intermediates, and kinetic parameters can grow dramatically large when multiple polymers and structural motifs are considered simultaneously. Parameter estimation, mechanism reduction, and identifiability therefore become increasingly difficult tasks, often rendering such models impractical for real waste-derived feedstocks or dynamic operating environments.

To address this complexity, lumped or pseudo-component models aggregate species into representative groups based on carbon number, boiling range, or chemical family, enabling tractable reactor-scale descriptions widely used in refining and catalytic cracking [15–17]. However, such lumping schemes are typically empirical and may obscure the structural origin of product distributions, limiting extrapolation to new feed compositions or operating conditions. Alternatively, machine learning (ML) approaches such as neural networks can capture nonlinear relationships between feed descriptors and product yields from large datasets [18]. Yet, purely data-driven models often lack embedded physical constraints, may violate conservation laws, and provide limited system interpretability. Consequently, a gap remains between detailed mechanistic modeling and flexible data-driven prediction for complex thermochemical conversion systems.

Modelling Approach & Aim.

To bridge this gap, recent work has explored physics-informed data analysis methods [19,20]. These approaches integrate conservation laws, chemical feasibility constraints, and structural reductions into data-driven inference to produce compact yet physically consistent representations of complex reacting systems. By combining predictive flexibility with mechanistic interpretability, such frameworks provide a promising direction for modeling heterogeneous thermochemical processes.

Within this context, the Carbon Bond Group (CBG) framework was previously introduced as a structural abstraction to analyze feedstock–product correlations in heterogeneous polymer cracking [14]. By classifying carbon atoms according to their local bonding environments, the CBG representation embeds heterogeneous feedstocks and their resulting product distributions within a shared structure-aware compositional space. The present work advances this framework from a qualitative correlation tool to a predictive, condition-dependent mapping formulation that explicitly enforces conservation laws and physical admissibility.

In this study, steam cracking is formulated as a physics-informed constrained transformation between chemically defined compositional spaces, where feed and product distributions are represented as vectors over CBG-defined structural groups. The proposed model can therefore be viewed as a constrained carbon-conversion operator, parameterized through a SoftMax function to obtain a column-stochastic mapping. Using ML-based parameter estimation, the model is trained on a comprehensive dataset derived from multiyear experimental campaigns in a semi-industrial DFB steam cracker. This dataset spans a broad temperature window and includes diverse feedstocks originating from real waste streams. Through this formulation, the model integrates structural domain knowledge with data-driven inference under physics-enforced constraints, complementing recent physics-informed ML developments in thermochemical conversion while addressing the specific need for a carbon-conserving, interpretable feedstock–product mapping for waste steam cracking.

The resulting condition-dependent CBG model provides a compact, interpretable, and physically admissible surrogate for complex

thermochemical conversion at the evaluated level of abstraction. Beyond predictive accuracy across diverse heterogeneous feedstocks and operating conditions, the model enables chemically consistent interpretation of carbon-routing trends and offers a scalable foundation for reduced-order digital-twin development in thermochemical recycling systems.

2. Method description

2.1. Carbon bond group (CBG) classification for heterogeneous systems

Heterogeneous polymer feedstocks, such as mixed plastics and textile-derived wastes, exhibit a wide diversity of molecular structures, functional groups, and polymer identities. When such materials are subjected to a high-temperature thermochemical conversion, such as steam cracking, their polymer chains undergo fragmentation through free-radical reactions governed by reaction kinetics and by the local chemical environments of the carbon atoms. At the end of the conversion, these chain-breaking events lead to stable molecules which ultimately provide the observed product species distribution. Consequently, a close connection exists between the polymeric chemical structure introduced into the reactor and the product distribution.

To study such correlations, a previous study introduced and validated an approach called carbon bond group (CBG) framework developed as a chemical structure-based representation for analyzing feedstock–product relationships in thermochemical conversion of complex polymer systems. In the present work, this framework is adopted as a foundation for model development. Only a brief summary of the framework is provided here to establish context, while full description, validation, and discussion can be found in the earlier work [14].

For its construction, the CBG framework classifies carbon atoms according to their immediate bonding environment, rather than polymer identity or molecular weight. This abstraction is particularly suited for highly heterogeneous feedstocks, where chemically distinct polymers may share common local bonding environments that govern their conversion behavior at high temperatures. In the CBG representation, all carbon atoms of the polymeric repeating unit structure are assigned to one of three basis groups:

- **C-xE**: corresponds to carbon atoms with x bonds to heteroatoms (E: O, N, S, etc.).
- **C-AL**: corresponds to carbon atoms in aliphatic bonds.
- **C-AR**: corresponds to carbon atoms contained in aromatic rings.

Fig. 1 summarizes the resulting CBG decomposition for representative polymer types commonly found in heterogeneous plastic and textile waste streams illustrating how distinct polymer structures are mapped into the reduced structural basis employed in the present model. The values correspond to carbon ratio with respect to the respective repeating units.

For polymers with relatively low heteroatom content besides oxygen, the C-xE group can be reduced to C-xO representing carbon bonded to oxygen by single or double bonds. A special case is defined for ester-containing structures where both oxygens from the ester are assigned to one carbon (e.g. in PET) [21].

A similar classification is applied to the products, where the species are grouped according to their dominant carbon bonding environments: The COx group, which includes carbon oxides (CO and CO_2). The aliphatic group (Aliph), comprising non-aromatic hydrocarbons including methane (C_1), light olefins and paraffins (C_2–C_3), and heavier aliphatic hydrocarbons (C_{4+}). The aromatic group (Arom), which consists of mono-aromatics (e.g., BTXS), polyaromatic hydrocarbons, and char. Fig. 2 summarizes the input–output transformations into the CBG framework.

Using this classification, and the polymeric composition, chemically complex feedstocks can be translated into the low-dimensional structural space of the framework as normalized carbon fractions, independent of polymer identity.

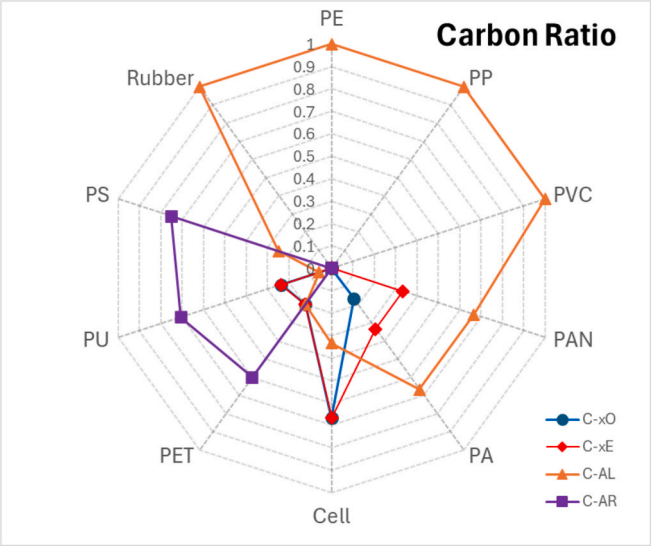

Fig. 1. Transformation of representative polymer types into the CBG framework concept (modified from [14]). The values correspond to carbon ratio from the repeating unit.

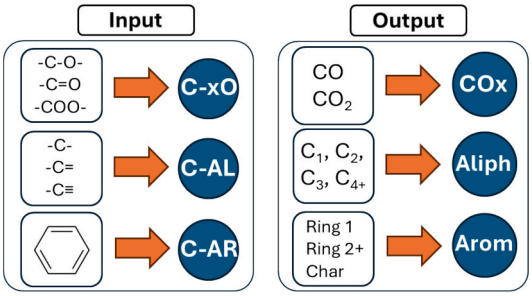

Fig. 2. Schematic of mapping feedstock's basis structures and product species into CBG framework classification.

By expressing both inputs and outputs within the CBG basis, the framework also enables a chemically meaningful mapping that can be expressed in mathematical form between feedstock structure and product distribution at a generalized level. This characteristic lays the foundation for the feedstock–product correlation model that will be presented in this study, as described in the following sections.

2.2. CBG model: feedstock–product mapping formulation

By construction, the group classification performed in the CBG framework enables both feedstock and product to be represented as ternary composition vectors, each lying within a 2-simplex (triangle) defined by carbon conservation. The feedstock vector $\boldsymbol{f}$ and the product yields vector $\boldsymbol{y}$ can be formalized according to Eq. (1).

$$\boldsymbol{f} = \begin{bmatrix} f_{C-xO} \\ f_{C-AL} \\ f_{C-AR} \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_{COx} \\ y_{aliph} \\ y_{arom} \end{bmatrix} \tag{1}$$

Each vector satisfies a unit-sum constraint. In geometrical terms, the elements $(f_{C-xO}, f_{C-AL}, f_{C-AR})$ and $(y_{COx}, y_{aliph}, y_{arom})$ represent barycentric coordinates within the feedstock and product ternary simplex spaces, respectively. Barycentric coordinates are a way to represent a point within a triangle (a simplex) as a weighted average of the vertices.

Considering a random point q in a triangle with vertices a, b, c, then, q can be expressed as a barycentric combination by $q = ka + mb + nc$, where (k, m, n) constitute the barycentric coordinates constrained by $k + m + n = 1$.

The conversion occurring in the reactor can be seen therefore as a mathematical transformation **M** that maps the feedstock simplex onto the products simplex (See Fig. 3).

Within this representation, the feedstock–product relationship is formulated here as a linear mapping between the two CBG spaces and expressed in the form defined by Eq. (2).

$$\boldsymbol{y} = \boldsymbol{K}\boldsymbol{f} \tag{2}$$

Where $\boldsymbol{K}$ is a (3×3) conversion matrix. The elements of $\boldsymbol{K}$, referred to here as conversion coefficients, represent the fraction of carbon originating in feedstock bond group $b \in \{C-xO, C-AL, C-AR\}$ that is redistributed to product family $g \in \{COx, aliph, arom\}$, see Eq. (3).

$$\boldsymbol{K} = \begin{bmatrix} K_{COx,C-xO} & K_{COx,C-AL} & K_{COx,C-AR} \\ K_{aliph,\,C-xO} & K_{aliph,\,C-AL} & K_{aliph,\,C-AR} \\ K_{arom,\,C-xO} & K_{arom,\,C-AL} & K_{arom,\,C-AR} \end{bmatrix} \tag{3}$$

Each column $K_{:,b}$ describes how the carbon in bond group b is partitioned among the groups $\{COx, aliph, arom\}$, To ensure physical consistency, the mapping satisfies two fundamental structural constraints. First, non-negativity: all coefficients are constrained to be non-negative to preserve physical interpretability. Second, column-stochasticity: carbon conservation requires that all carbon originating from a given feedstock bond group is fully redistributed among the products, meaning that each of the columns must sum to unity. Together, these constraints ensure that the mapping remains compatible with the compositional nature of the system and with elemental conservation laws.

The structure of $\boldsymbol{K}$ can provide a compact and chemically interpretable description of the aggregated reaction pathways acting on the feedstock during cracking. Within this structure, the diagonal elements: $K_{COx,C-xO}, K_{aliph,\,C-AL}, K_{arom,\,C-AR}$ represent the dominant retention of carbon, encoding the likelihood that carbon remains in a broadly similar bond-type environment. Conversely, the off-diagonal elements capture cross-group redistribution arising from alternative radical–driven reaction pathways such as: C-AL → arom (cyclization and aromatization), C-AR → COx (steam reforming and ring oxidation), C-xO → COx (decarbonylation/ decarboxylation), and other eventual aliphatic pathways (C-AR → aliph and C-xO → aliph). Since $\boldsymbol{K}$ represents the thermochemical conversion occurring in the reactor, these elements quantify the extent to which heterogeneous feedstocks undergo reactions that affect their structural carbon environments in ways that persist to the final observed product slate.

It is important to emphasize that Eq. (2) formulation constitutes a joint mapping between two chemically defined compositional spaces, in which all product groups must be determined simultaneously as functions of the full set of feedstock bond environments. Each element y_g, with $g \in \{COx, aliph, arom\}$, is defined as the linear predictor of the form given by Eq. (4).

$$y_g = K_{g,C-xO} f_{C-xO} + K_{g,C-AL} f_{C-AL} + K_{g,C-AR} f_{C-AR} \tag{4}$$

Such predictor formulation expresses each product family as the joint contribution of all feedstock bond groups, highlighting that product distributions emerge from the aggregate interaction of multiple structural environments rather than isolated correlations.

Although the CBG model does not prescribe mechanistic details at the species level, the matrix $\boldsymbol{K}$ is intended to capture the net effect of thousands of radical reactions occurring within the reactor. Within this reduced abstraction:

- diagonal dominance may often reflect structural stability or favored termination pathways under the studied conversion conditions,
- non–zero cross–terms reveal active redistribution routes such as aromatization or gasification,
- vanishing entries indicate chemically improbable transitions under the evaluated reactor conditions.

Thus, $\boldsymbol{K}$ serves as a coarse–grained operator linking heterogeneous chemical structure to aggregated chemical outcomes. It embeds essential chemical tendencies, such as aromatic persistence, oxidation favorability, aliphatic fragmentation, while remaining compact enough for statistical estimation and generalization.

In this work, the conversion matrix $\boldsymbol{K}$ is determined from experimental data using constrained optimization techniques, commonly used in machine learning applications, while ensuring consistency with the physical structure of the model defined before. In its simplest form, the mapping set in Eq. (2) describes the conversion behavior at a given

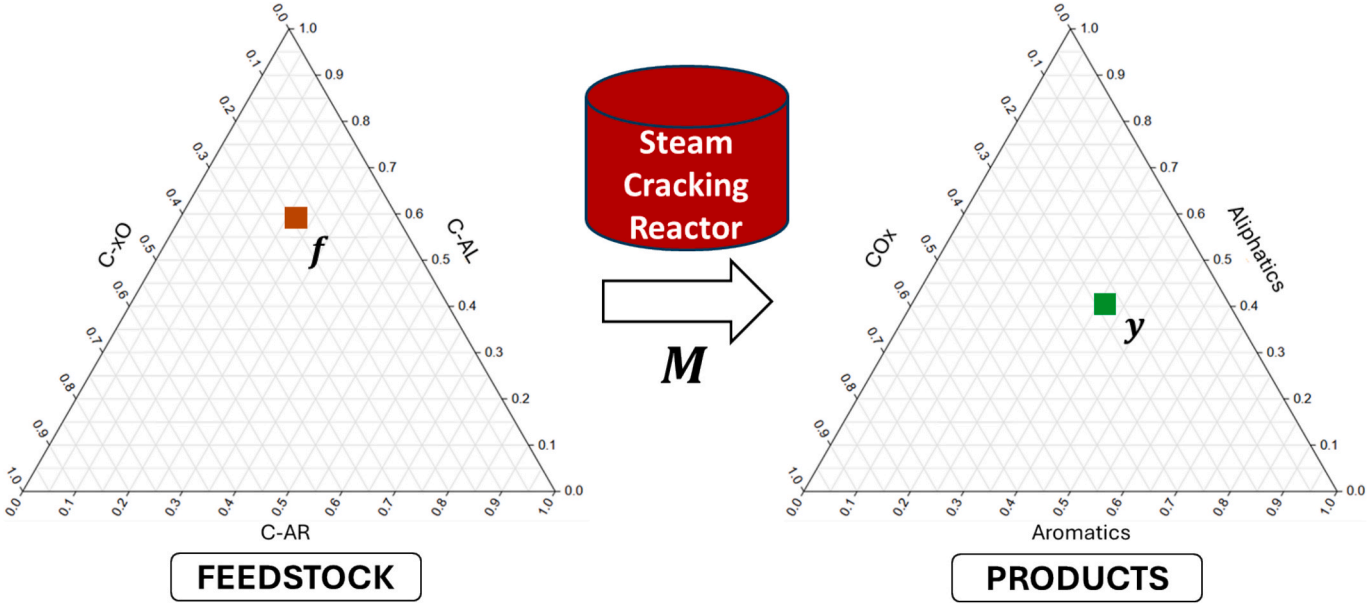

Fig. 3. Schematic representation of the CBG simplex spaces for feedstock and products, and the feedstock–product transformation through the steam cracking reactor.

reactor's operative conditions. In the following section, this formulation is generalized to explicitly incorporate operating temperature as a governing variable, allowing the conversion matrix to evolve systematically with environment-induced severity while preserving the chemical interpretation and compositional constraints of the CBG framework.

2.3. Condition–Dependent CBG model description

In its simplest form, the conversion matrix $\boldsymbol{K}$ of Eq. (2) represents an averaged transformation valid for a single operating regime. However, thermochemical conversion processes exhibit strong dependence on operational severity, particularly temperature, which modulates the balance between radical propagation, β-scission, cyclization, aromatization, and gasification pathways. To account for these effects, the CBG model is generalized to a condition-dependent mapping, allowing the conversion matrix $\boldsymbol{K}$ to evolve systematically with operating conditions. The condition-dependent mapping is written as Eq. (5).

$$\boldsymbol{y}(\boldsymbol{c}) = \boldsymbol{K}(\boldsymbol{c})\boldsymbol{f} \tag{5}$$

where $\boldsymbol{c}$ denotes a vector of operating conditions, and $\boldsymbol{K}(\boldsymbol{c})$ is a condition-dependent conversion matrix.

Given a dataset n of feedstock compositions $\boldsymbol{f}^{(n)}$, product yields $\boldsymbol{y}^{(n)}$, and corresponding conditions $c^{(n)}$, the goal is to estimate a conversion matrix $\mathbf{K}(c)$ that satisfy:

i. Column–stochasticity (carbon conservation):

$$\sum_{g} K_{g,b} = 1; \quad \forall b \in \{C - xO, C - AL, C - AR\} \tag{6}$$

ii. Non–negativity:

$$K_{g,b} \geq 0; \quad \forall g, b \tag{7}$$

iii. Chemical tendency inequalities:

$$K_{arom,C-AR} \geq K_{arom,C-AL}, \quad K_{COx,C-xO} \geq K_{COx,C-AL}, \quad K_{aliph,C-AR} \leq \in \tag{8}$$

It is important to notice that the constraints over $\boldsymbol{K}$ do not force the model to follow any specific chemical mechanism; instead, they restrict the feasible region of the solution space to remain consistent with the respective conservation laws and radical–chemistry tendencies in the evaluated system. So, for instance, the constraint in Eq. (8) leaves certain diagonal terms free to be large if the data supports it, while sets a margin for the unlikeliness of getting certain cross-bond to product transfers (e.g. C-AR to aliph) under non–catalytic conditions.

To ensure that the estimated mapping remains structurally admissible, the elements of $\mathbf{K}(\boldsymbol{c})$ are parameterized using a SoftMax transformation, which guarantees non–negativity and column–stochasticity for all values of $\mathbf{c}$. Each coefficient is defined by Eq. (9).

$$K_{g,b}(\boldsymbol{c}) = \frac{\exp\left(\alpha_{g,b} + \beta_{g,b}\,\phi(\boldsymbol{c})\right)}{\sum_{g'} \exp\left(\alpha_{g',b} + \beta_{g',b}\,\phi(\boldsymbol{c})\right)} = \frac{\exp\left(z_{g,b}(\boldsymbol{c})\right)}{\sum_{g'} \exp\left(z_{g,b}(\boldsymbol{c})\right)} \tag{9}$$

Where $z_{gb}(\boldsymbol{c})$ represents an unconstrained latent score (logit) matrix, and it can be seen as a measure of the relative tendency of a carbon in group b to transition into group g under conditions $\boldsymbol{c}$. The softmax function then transforms these scores into probabilities, ensuring all entries are strictly positive and each column sums to one. $\phi(\boldsymbol{c})$ corresponds to an assembly of operative conditions. Depending on the level of detail and process sensitivity, $\phi(\boldsymbol{c})$ may include: direct variables (e.g. T, Steam/Feed ratio, residence time (τ)), polynomial or nonlinear terms (e.g. T^2, log$T$), interactions (e.g. $T \times \tau$), or spline–based expansions for flexible condition dependence. As a proof of concept, in this study the model is demonstrated using temperature as the dominant operating variable, i.e., $\phi(\boldsymbol{c}) = \phi(T) = T$, and Eq. (5) becomes a temperature-dependent mapping defined by Eq. (10).

$$\boldsymbol{y}(T) = \boldsymbol{K}(T)\boldsymbol{f} \tag{10}$$

Notice that the temperature-dependent formulation can be reduced to the fixed-condition mapping when evaluated at a single temperature. In that case, $\mathbf{K}(T)$ becomes constant, and the model recovers the original linear feedstock–product operator of Eq. (2).

It is important to highlight that the feedstock vector $\boldsymbol{f} = [C - xO, C - AL, C - AR]^T$ is a compositional vector in the feedstock simplex and therefore has two independent degrees of freedom by construction. The three-component representation is retained because it provides barycentric coordinates in the simplex and preserves the physical interpretation of the columns of $\mathbf{K}$. In the temperature-dependent model of Eq. (10), each experimental case is described by the pair $(\boldsymbol{f}, T)$; thus, similar feedstock CBG vectors measured at different temperatures can still contribute to estimating the carbon routing operator $\boldsymbol{K}(T)$.

Finally, given a set of N experiments $\left\{\boldsymbol{f}^{(n)}, T^{(n)}, \mathbf{y}^{(n)}\right\}_{n=1}^{N}$, the model's parameters are jointly determined by optimizing the loss function defined by Eq. (11).

$$L = L_{data} + L_{reg} + L_{phys} \tag{11}$$

- Data-fit term: L_{data} penalizes deviations between measured and predicted carbon allocations (squared Euclidean norm).
- Regularization term: L_{reg} applies ridge penalties to avoid overfitting and to enforce smooth condition-dependence.
- Physics-constraint term: L_{phys} penalizes violations of the inequalities in Eq. (8), ensuring that the estimated operator conforms to the system-related priors.

The model parameters are thus obtained by minimizing the loss function. Mathematical details of Eq. (11) terms can be found in the supplementary material. Optimization was performed in Python/PyTorch using the Adam optimizer with a learning rate of 0.05 and default momentum parameters. Training was done for 2000 epochs, until the loss reached a stable plateau. The predictive performance of the trained model is later evaluated using standard machine–learning metrics, including the coefficient of determination (R^2), mean absolute error (MAE), and relative MAE, which provide complementary measures of predictive accuracy.

In summary, the condition–dependent CBG model extends the fixed–regime mapping to incorporate arbitrary operating variables, ensuring physical plausibility at all conditions through structured constraints and SoftMax parameterization. It also provides a unified mathematical formulation to evaluate how reaction conditions can affect carbon redistribution pathways at an abstracted but physically meaningful level. This generalized formulation forms the basis for the predictive and interpretive analyses presented in Section 4, enabling the CBG framework to describe heterogeneous feedstocks and varying operating conditions within a unified, interpretable modeling approach.

3. Experimental setup

The experiments analyzed in this study were carried out at the Chalmers Power Central facility, where a DFB Gasifier/Cracker is integrated with a Circulating Fluidized Bed (CFB) combustor operating on biomass wood chips. Silica sand served as the bed material, and steam was used as the fluidizing agent. The cracker operated in bubbling fluidized bed (BFB) regime and processed 40–160 kg/h of plastic feedstock, while the CFB combustor was supplied with 1,500–3,000 kg/h of wood

chips. A small flow of high-purity helium (35 Ln/min) was injected with the steam (~150 kg/h) to act as a tracer gas for species quantification.

The DFB system enables continuous withdrawal of the produced char together with the circulating bed material, returning it to the combustor. After conversion of the feedstock, a continuous gas sample (~10 Ln/min) is extracted at the reactor outlet, while the remaining product gas (~3,500 Ln/min) is routed back to the combustor (see Fig. 4).

The sampled gas passed through a ~ 350 °C filter and was split into two lines (Fig. 4 at the right). One line was quenched with isopropanol and chilled before analysis in a Varian CP4900 micro-GC equipped with MS5Å and PoraPLOT Q columns and a Thermal Conductivity Detector (TCD). Permanent gases and C1–C3 hydrocarbons were measured online every 3 min, with weekly calibration using standard gas mixtures. C4 + hydrocarbons were analyzed using a GC-VUV. Gas was taken upstream of the quench loop, dried using a saturated amine filter, and collected in a gas bag. Hydrogen was used as the carrier gas and nitrogen in the detector. More details about this sampling system can be found elsewhere [22].

From the same sampling port, a Solid Phase Adsorption (SPA) method is used to collect the aromatic fraction. The gas is drawn at a controlled rate through an adsorbent cartridge (Supelclean ENVI–Carb/NH2 SPE), which contains a 500 mg amine layer followed by 500 mg of activated carbon. The trapped aromatics are then eluted into a solvent mixture containing hexylbenzene and 4–ethoxyphenol as internal standards (~12,000 mg/L and ~ 250 mg/L). The extract is analyzed using a Bruker GC430 equipped with a Flame Ionization Detector (FID) and a BR–17 MS column with hydrogen as the carrier gas. This method quantifies 28 aromatic compounds, ranging from monoaromatics (BTXS) to heavier polyaromatics such as naphthalene, anthracene, and triphenylene. Further details are provided elsewhere [22,23].

The second hot gas stream is directed through a High–Temperature Reformer (HTR) electrically heated to ~ 1,700 °C, where all hydrocarbons are fully converted by steam reforming. After reforming, the gas passes through a soot filter and a condenser to remove water before entering a second micro–GC (GC2). The resulting composition—primarily H_2, CO, CO_2, and helium—enables determination of the total carbon content in the producer gas and an indirect estimation of char yield. The efficiency of the reformer is confirmed by near–zero methane levels. The HTR operates continuously alongside the main sampling system. After each measurement campaign, the reactor is flushed with air to burn residual soot, and the resulting CO_2 is measured to account for its carbon content.

This procedure, developed by Israelsson et al. [24], provides a robust framework for validating the carbon balance from the cracker gases. The solution of the carbon balance from the reaction system based on the characterization setup is depicted in Fig. 5.

Feedstocks and Operative Conditions.

Table 1 summarizes the materials used in this study, including their origin, key characteristics, and dominant polymer composition. The two polyolefin–rich materials (PE and MRP) are virgin or mechanically recycled polymers with high polyolefin content, negligible oxygenated or aromatic species, and very low ash. All other feedstocks originate from real post–consumer waste streams and therefore contain heterogeneous mixtures of polymers, elevated oxygen content, and a certain ash fractions. These materials were shredded and pelletized prior to feeding into the reactor.

The elemental compositions and lower heating values (LHV) of all eight feedstock batches are presented in Table 2. PE and MRP show the

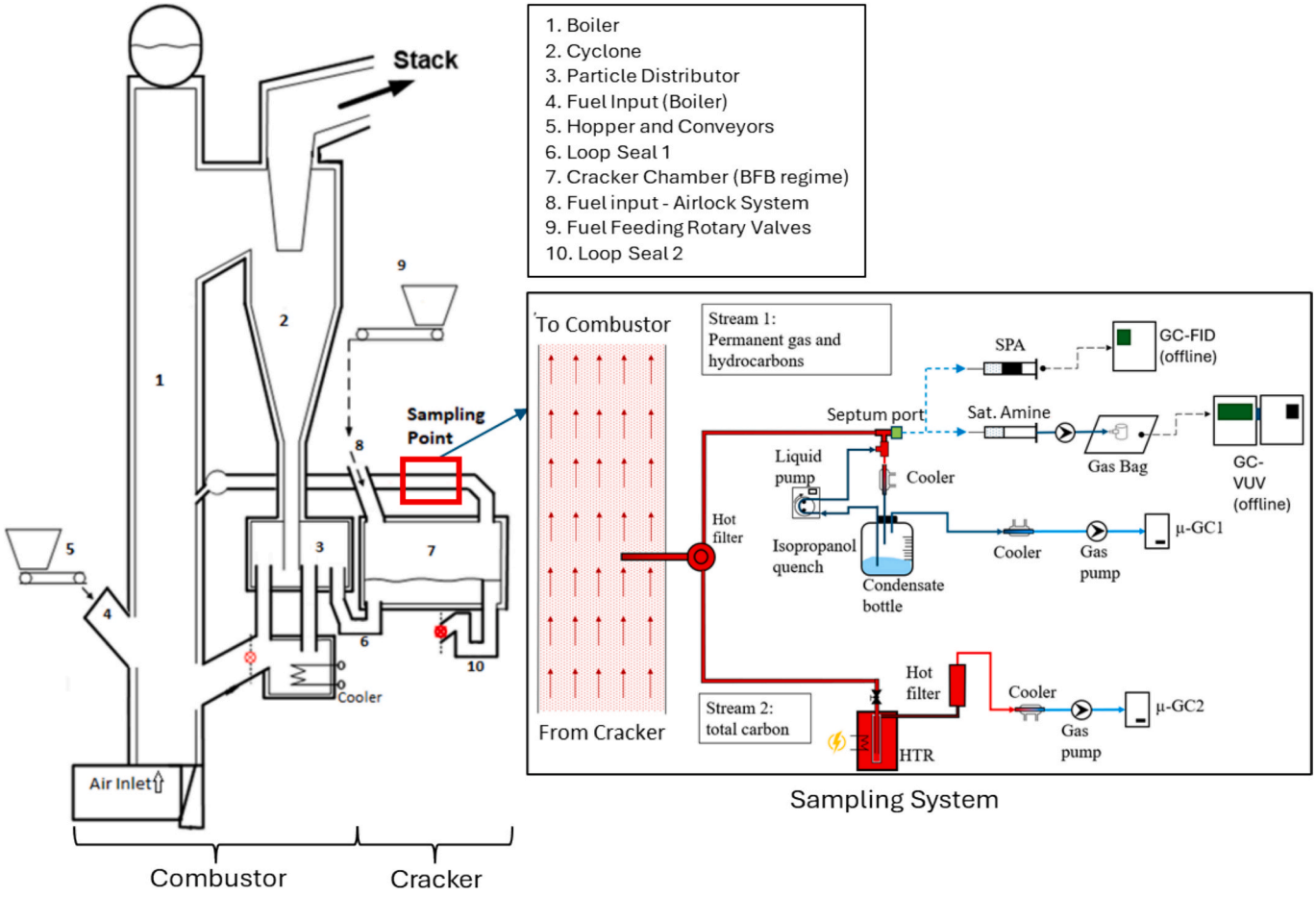

Fig. 4. Schematic of the Chalmers power plant and gasifier as well as the sampling system for species characterization. (adapted from [22]).

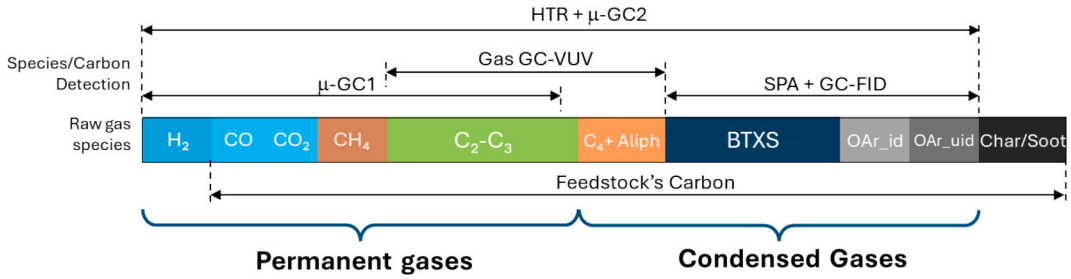

Fig. 5. Schematic of carbon Balance estimation from the implemented characterization setup. AR_kn represents known aromatics such as polyaromatics, while AR_ukn refers to additional unidentified aromatics detected in the SPA–FID analysis. (modified from [14]).

Table 1
Overview of heterogeneous materials used in this work, including their general characteristics and origin. (PE: polyethylene; PP: polypropylene; PET: Polyethylene Terephthalate; PS: polystyrene; PA: polyamide; PAN: polyacrylonitrile; PU: polyurethane; PVC: polyvinyl chloride; Cell: cellulose.).

Material	Polymer Types	Description
Polyethylene (PE)	PE.	Virgin Polymer Pellets. High polyolefin content. None/very low ash content
Mechanical Recycled Polyolefins(MRP)	PE/PP.	Polymer Pellets. High polyolefin content. None/very low ash content
PlasticPackaging Sorting Reject. 2D Packaging (PSR2D)	PE/PP, PET, Cell, PA, PVC.	Reject from post-consumer sorted plastic packaging waste. Mixture of 2D packaging, e.g., multilayer films. Medium high polyolefin content, Medium-low ash content
PlasticPackaging Sorting Reject. 3D Packaging (PSR3D)	PE/PP, PET, PS, PA, PVC.	Reject from post-consumer sorted plastic packaging waste. Mixture of 3D packaging, e.g., bottles, food, cleaning and household recipients. Medium polyolefin content, Medium-low ash content
Mixed Packaging Waste (MPW)	PE/PP, PET, Cell, PVC. (others: PU, PA, PS).	Post-consumer unsorted packaging plastic waste. Medium polyolefin content, Medium ash content
CardBoard Recycling Residue (CRR)	PE/PP, Cell, PET, PVC. (others: PU, PA, PS)	Post-consumer shredded stream of multilayer cardboard/plastic for food packaging. High aliphatic carbon content, Medium ash content
Polyester-based Textiles (TXTp)	PET, PA, PAN, Cell. (others: Wool, PVC, PU).	Household textile waste after sorting the useful pieces of cloth. Complex polymer blends, High oxygen content
Cotton-based Textiles (TXTc)	Cell, Polyester (PET), PA, PAN. (others: Wool, PVC, PU).	Textile waste after sorting the useful pieces of cloth. Complex polymer blends, High oxygen content

highest carbon content and LHVs, consistent with their polyolefinic nature. In contrast, heterogeneous post–consumer streams, including PSR2D, PSR3D, MPW, CRR, and textiles, exhibit higher oxygen and ash contents, reflecting the presence of multilayer structures, natural fibers, and additives.

The polymeric compositions of the different feedstock batches were first estimated through manual sorting and Near–Infrared (NIR) spectrometry prior to pelletization. To further reduce uncertainty, a convex–optimization–based estimation method was applied using a convex optimization model built on elemental balances and low heating values (details can be found in [14]). The resulting polymeric compositions for all evaluated feedstocks are presented in Fig. 6.

The main operating conditions applied in the DFB cracking reactor are summarized in Table 3. The dataset used in this work was collected from experimental campaigns performed over four consecutive operational seasons, spanning four years of reactor operation. The total number of evaluated datasets is N = 32. These campaigns were originally designed to evaluate the technological viability and robustness of the DFB steam cracking process for handling a wide variety of heterogeneous polymeric mixtures under industrially relevant conditions. The investigated waste feedstocks correspond to sorting rejects obtained from different waste management facilities, representing heterogeneous polymeric materials generated within real recycling infrastructures. The results span temperatures between approximately 720 and 835 °C and include feedstocks with widely varying polymeric compositions, from clean polyolefins to complex oxygen-containing waste streams. The cracker was fluidized with steam under bubbling conditions. The "Feeding position" column refers to the locations indicated in Fig. 4, where two feeding configurations are available. Position 6: delivers a molten polymer flow produced by an extruder, which compresses and heats the pelletized feedstock before releasing it onto the cracker bed. Position 8: feeds the material in pellet or granulate form directly into the bed by gravity through an air–lock rotary–valve system. More details of this system can be found elsewhere [25].

Since the model inputs correspond to averaged steady-operation cases rather than time-resolved signals, spectral appraisal and signal-to-noise-ratio analysis were not applicable to the data used for model training. Details on the sampling, product characterization, and stability of operation achieved for this experimental setup can be found elsewhere [22]. The influence of variability on model performance was therefore evaluated at the level of product-group prediction errors, using residual analysis and validation metrics.

The dataset used for model training, validation, and reproduction of the CBG model's calculations implemented here is provided in Table S1 in the supplementary material.

4. Results and analysis

4.1. Carbon yields distribution Overview

Here it is presented the experimental product yields obtained from the thermochemical conversion of heterogeneous polymer feedstocks, with emphasis on how temperature and feedstock composition influence carbon allocation among major product families. The analysis goes from individual species behavior to aggregated carbon groups, establishing the phenomenological basis for the reduced modeling framework introduced in subsequent sections. In order to evaluate the product distribution yields for the different evaluated feedstocks and conditions, Fig. 7 presents the carbon yield distribution of key product species

Table 2
Elemental composition (%w dry basis) and LHV (MJ/kg_{dry}) of the investigated feedstocks. Oxygen content estimated by difference.

Item	PE	MRP	PSR2D	PSR3D	MPW	CRR	TXTp	TXTc
C	85.71	85.71	74.90	75.43	65.03	60.60	60.93	51.95
H	14.29	14.29	11.73	10.76	10.33	9.00	5.54	5.95
O*	0.00	0.00	8.13	8.64	14.64	21.00	28.96	37.96
N	0.00	0.00	0.67	0.54	0.52	0.35	3.15	1.53
S	0.00	0.00	0.04	0.03	0.09	0.07	0.07	0.11
Cl	0.00	0.00	0.20	0.31	0.70	0.20	0.12	0.37
Ash	0.00	0.00	4.34	4.28	8.69	8.75	1.22	2.11
LHV	42.80	44.61	36.03	37.02	30.98	30.14	28.13	17.87

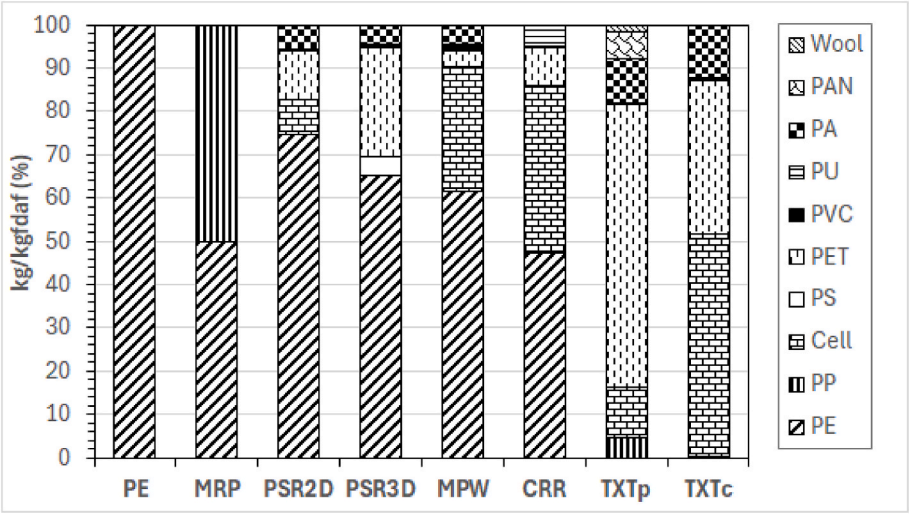

Fig. 6. Estimated polymeric composition of the feedstocks used in this study. (kgfdaf: kilograms of feedstock in dry and ash-free basis).

Table 3
Operating conditions for the DFB cracking reactor used in this work.

Fuel	*Temperature Cracker (°C)*	*Bed material*	*Feed Flow (kg_{daf}/h)*	*Feeding Mode*	*Feeding Position*	*N. samples*
PE	725–831	Silica Sand	63–100	Molten flow via extrusion	6	4
MRP	720–807	Silica Sand	60–75	Molten flow via extrusion	6	4
PSR2D	767–829	Silica Sand	49–58	Molten flow via extrusion	6	4
PSR3D	760–826	Silica Sand	50–60	Molten flow via extrusion	6	4
MPW	721–826	Silica Sand	50	Molten flow via extrusion	6	4
CRR	742–804	Silica Sand	37–50	Molten flow via extrusion	6	3
TXTp	735–815	Silica Sand	37–150	Extrusion / Rotary valves	6,8	7
TXTc	~756	Silica Sand	37–50	Molten flow via extrusion	6	2

groups as a function of cracking temperature for the range of heterogeneous feedstocks evaluated in this work. Carbon yields are expressed as a percentage fraction of the input carbon (%Carbon) to enable direct comparison across conditions and materials. Note that the statistical uncertainty of the results presented is between 3 % and 7 % for all the evaluated cases. According to the product groups defined for the CBG framework (see Fig. 2), the aggregated aliphatics and aromatics groups in panels (h) and (i), contain other aliphatic species and char respectively, which are not depicted in Fig. 7.

At the level of individual species, panels (a)-(g) in Fig. 7 reveal systematic trends with temperature across all investigated feedstocks, which shows the expected dominant role of temperature upon the species distributions of each feedstock at the evaluated conditions. Carbon oxides, shown in panels a and d, increase monotonically with temperature, with CO_2 exhibiting a steeper rise than CO. Due to the immanent steam reforming reactions participating in the process, and some bed oxygen transport [26,27], this positive trend is consistent across feedstocks, but the absolute yield levels are modulated by feedstock type as observed in the general COx yield of panel (g). Oxygen-rich materials, such as textile-derived streams (TXT) and CRR, set the upper range of COx yields, whereas polyolefin-rich feedstocks such as PE and MRP consistently appear at the lower end. For instance, TXTp reaches COx yields of approximately 22 %C and 41 %C at the lowest (~735 °C) and highest (~815 °C) temperature. The cellulose-rich feedstock TXTc shows the highest COx yields at intermediate temperatures (~757 °C), while the pure polyolefin materials remain characterized by comparatively low carbon oxide formation (between 2 %C to 10 %C), throughout the temperature range.

The hydrocarbon species display a more differentiated response to temperature and feedstock composition, as illustrated in panels (b), and (e). Methane yields present a positive monotonic trend with temperature and remain confined to a relative narrow range for most of the feedstocks between 7 %C to 15 %C, showing weak scattering from feedstock composition. This behavior shows methane as a comparatively stable carbon sink under the investigated condition. In contrast, the light olefinic monomers group (ethylene and propylene), shown in panel (e), exhibit overall an inverse temperature dependence, with higher yields at low to intermediate temperatures and a decline at the highest severity

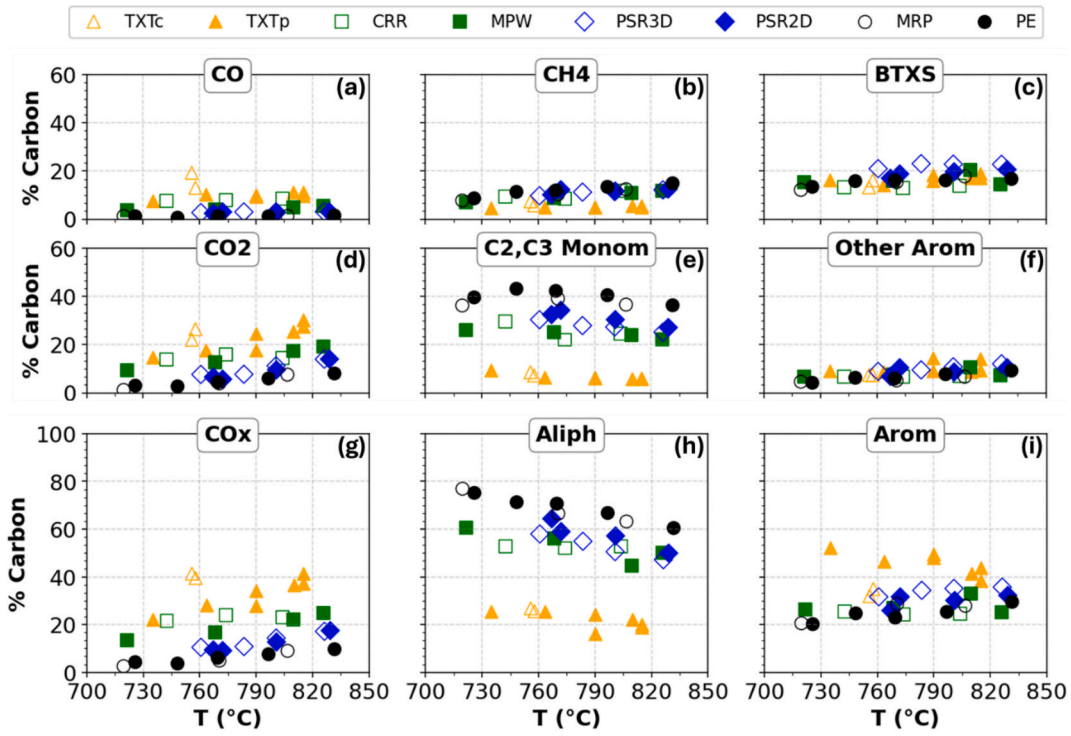

Fig. 7. Carbon Conversion Yields of key chemical species and groups for the evaluated heterogenous feedstocks.

levels. In addition to this temperature-driven trend, a significant stratification across feedstocks is observed. Pure polyolefinic materials (PE and MRP) yield the highest monomer fractions, averaging around 40 %C and reaching peak values at lower temperatures (43 %C for PE at ~740 °C), whereas textile-derived feedstocks yield substantially lower monomer fractions, on the order of 6–9 %C. The decreasing trend behavior is also replicated to the global aliphatics group as shown in panel (h). This contrast highlights the combined influence of feedstock structure and operating severity on the distribution of hydrocarbon products.

For aromatic species, represented by BTXS and other measured aromatics in panels (c) and (f) of Fig. 7, a moderate but systematic increase with temperature is observed, with less pronounced feedstock-to-feedstock variability compared to aliphatic monomeric species. BTXS yields span approximately from 12 to 23 %C across the investigated conditions. Higher yield levels are generally associated with mixed plastic wastes such as PSR3D, which contain a high fraction of polyolefins as well as polymers with aromatic structures (e.g., PET and PS), whereas cellulose-rich feedstocks such as CRR and TXTc occupy the lower range of BTXS yields at comparable temperatures. Despite the wide diversity in polymeric composition among the tested feedstocks, the overall scatter of BTXS and other aromatic carbon conversion yields share similar levels, suggesting that different formation routes converge toward similar carbon allocation levels for aromatic products [9,14,21].

When considering the aggregated aromatics group in panel (i), a consistent upward trend with temperature is observed for most feedstocks. An exception is observed for polyester-rich textile feedstocks (TXTp), for which the total aromatic yield decreases at higher temperatures. This behavior is consistent with enhanced gasification of the substantial char fraction generated from such materials under severe conditions, as reported previously [21].

In general, the results presented in panels (g)-(i) in Fig. 7 show that, despite the chemical heterogeneity of the investigated feedstocks, the main product groups follows structured and interpretable trends with temperature, and also modulated by feedstock composition. To have a closer look at such dependance, Fig. 8 presents the graph of the main

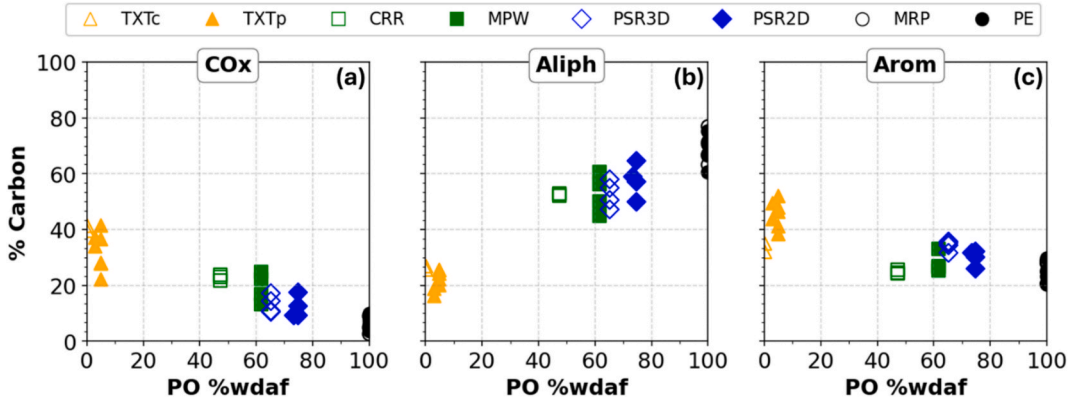

Fig. 8. Main species group yields as a function of the polyolefins share in the evaluated feedstocks. (PO %wdaf: Polyolefin content in the feed in % of mass on a dry and ash-free basis).

species groups yields against the feedstock's polyolefins content for the whole dataset.

See Fig. 8 illustrates the dependence of the main product groups against the polyolefins content of the feedstocks across the entire evaluated dataset. Clear monotonic trends are observed: aliphatic yields increase with increasing polyolefin content, while both COx and aromatics decrease. For feedstocks dominated by polyolefins, aliphatic products account for the majority of converted carbon, reaching approximately 60–80 %C for pure polyolefin streams such as PE and MRP. In contrast, the other feedstocks with low polyolefin content exhibit higher carbon oxide and aromatic yields.

These trends reveal a relevant role that polyolefin content as a feed descriptor has on the product distribution, reflecting the propensity of long aliphatic chains to generate stable aliphatic fragments under cracking conditions. Conversely, feedstocks characterized by a more diverse chemical composition, arising from the presence of oxygenated and aromatic-containing polymers, can promote carbon oxides and aromatic species formation. However, while polyolefin content provides a useful first-order indicator of feedstock–product relationships, a single scalar descriptor cannot capture the influence of such chemical diversity on the resulting product distribution. This limitation motivates the adoption of a more chemically intrinsic representation capable of resolving multiple and more detailed structural contributions simultaneously, which will be detailed in the following sections through the carbon bond group (CBG) framework.

4.2. Feedstocks evaluation into the CBG framework

Using the polymeric shares of the evaluated feedstocks (see Fig. 6), Fig. 9 shows the transformation of the polymeric composition into the CBG chemical framework (see Section 2.1). PO refers to polyolefins content.

In Fig. 9, each feedstock is represented in terms of its fractional contributions of carbon bounded to oxygen (C-xO), carbon in aliphatic bonds (C-AL), and carbon in aromatic bonds (C-AR), revealing a structured but heterogeneous input space that reflects the underlying diversity of polymer chemistries.

Pure polyolefin feedstocks (PE and MRP) are characterized exclusively by C-AL. Mixed plastic wastes and polymer blends exhibit intermediate compositions, with varying shares between C-AL and C-AR carbon fractions. In contrast, textile-derived and cellulose-rich feedstocks (see Fig. 6) display substantial C-xO contributions, together with significant C-AR fractions.

It is important to notice that feedstocks with similar polyolefin contents can differ markedly in their CBG composition, highlighting the additional chemical resolution introduced by the CBG classification. So, this decomposition illustrates how the CBG framework retains

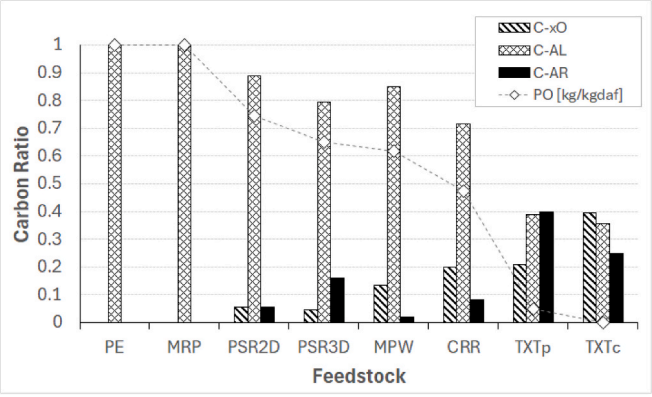

Fig. 9. Transformation of tested feedstocks into the structural classification of the CBG framework.

chemically meaningful distinctions that are not accessible through single-parameter descriptors. By resolving multiple carbon environments simultaneously, the CBG representation can provide a structured and physically interpretable input that enables systematic examination of how different structural features in the feed may contribute to the product distribution. Based on this transformation and the theoretical framework of the CBG model seen in Section 2, the following analysis presents how these carbon bond group fractions map to the main product group yields under fixed operating conditions.

4.3. Fixed-Regime CBG model implementation

To examine the CBG model implementation, the analysis was conducted first for the constant conditions case choosing a relatively narrow operating regime corresponding to 760–785 °C, with an average temperature of approximately 768 °C. The specific temperature window was selected because it provides a representative range for stable cracking conditions and broad experimental coverage across all investigated feedstocks. This fixed-regime implementation is used as a baseline demonstration for a narrow temperature window, whereas the full dataset is evaluated in Section 4.4 using the temperature-dependent formulation presented in Section 2.3. Fig. 10 shows the carbon yields of the three aggregated product groups: COx, aliphatics, and aromatics, plotted as a function of the carbon bond group fractions in the feedstock (C-xO, C-AL, and C-AR) at the chosen temperature regime, as well as the fitting linear projection from the CBG model implementation over the dataset.

From the data points in Fig. 10 it is possible to observe that despite the wide diversity in polymeric composition among the tested materials, clear monotonic trends emerge in all three cases. Carbon oxides increase with the fraction of oxygen-bound carbon in the feedstock, aliphatic products scale with aliphatic carbon content, and aromatic products increase with the aromatic carbon content. The consistency of these trends across heterogeneous materials illustrates how the CBG framework reveals generalized structure–product relationships linking feedstock chemistry to product distribution, in agreement with previous correlation-based observations [14].

Regarding the results of the model's implementation at the evaluated operating regime, the fitted joint CBG mapping reproduces well the observed trends across all feedstocks. Quantitatively, the fixed-regime mapping achieves in-sample R^2 values of [0.994, 0.981, 0.829] for [COx, Aliph, Arom], respectively. The corresponding MAE values are [0.0080, 0.0194, 0.0228] (carbon-fraction basis), while the RMSE values are [0.0097, 0.0221, 0.0256]. The relative MAE values are 6.4 %, 3.9 %, and 8.0 %, respectively. The high R^2 and low RMSE values for COx and aliphatics indicate that these groups are well explained by the feedstock CBG composition within the fixed-regime linear transformation. In contrast, although the aromatic prediction is still acceptable, the lower R^2 values indicate that part of the aromatic variance is not fully resolved by the fixed-regime abstraction. This residual dispersion suggests that additional chemical features may influence aromatic formation beyond the variables considered in the present abstraction.

In the CBG formulation, feedstocks are represented through generalized carbon bonding environments without explicitly resolving distinctions such as linear versus branched aliphatic structures (e.g., from PP) or heteroatoms other than oxygen, such as chlorine, which may affect secondary cyclization and aromatization pathways [28,29]. The remaining variance therefore delineates the boundaries of the current abstraction and indicates a direction for future model refinement through expanded structural descriptors.

In relation to the linear fits shown in Fig. 10, it is important to emphasize that such fits do not represent independent pairwise correlations from standard regression analysis, but rather projections of the underlying joint multivariate transformation between the feedstock and

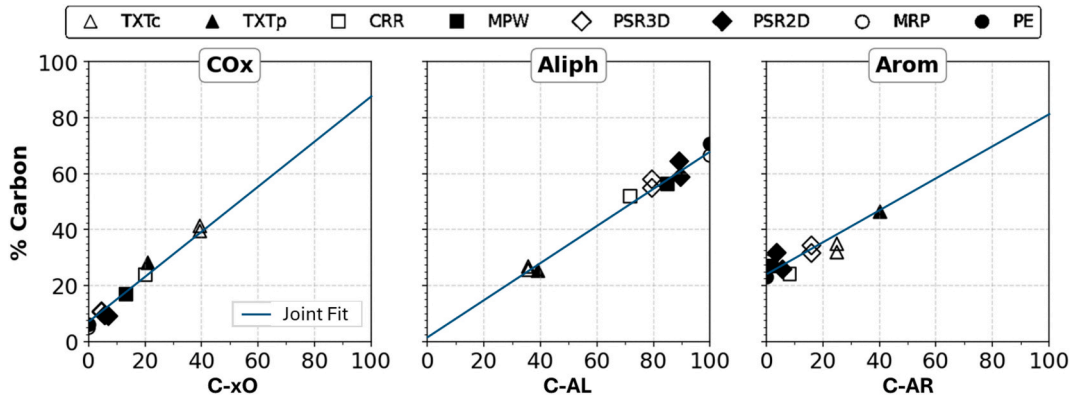

Fig. 10. Results of the dominant projections and the joint fit from the CBG model's implementation for the fixed temperature regime case (average: 768 °C).

product simplex spaces defined in the CBG model. As detailed in Section 2, the CBG model is formulated as a linear mapping between the inputs and outputs carbon ternary vectors. Therefore, the product yields are expressed simultaneously as functions of all feedstock carbon groups (See Eq. (4). This formulation accounts not only for the dominant diagonal contributions evident in Fig. 10, but also for cross-terms that capture redistribution of carbon among the different functional groups. In that regard, the fitted coefficients of the $\boldsymbol{K}$ matrix for the evaluated regime are presented in Eq. (12).

$$\boldsymbol{K} = \begin{bmatrix} K_{COx,C-xO} & K_{COx,C-AL} & K_{COx,C-AR} \\ K_{aliph,C-xO} & K_{aliph,C-AL} & K_{aliph,C-AR} \\ K_{arom,C-xO} & K_{arom,C-AL} & K_{arom,C-AR} \end{bmatrix} = \begin{bmatrix} 0.8749 & 0.0489 & 0.1818 \\ 0.0292 & 0.6904 & 0.0000 \\ 0.0959 & 0.2607 & 0.8182 \end{bmatrix} \quad (12)$$

As shown in Fig. 10, the diagonal terms dominate, where oxygen-bound carbon (C-xO) is routed primarily to COx ($K_{COx,C-xO} \approx 0.87$), carbon in aliphatic bonds (C-AL) is routed primarily to aliphatic product species ($K_{Aliph,C-AL} \approx 0.69$), and carbon in aromatics (C-AR) is routed primarily to aromatics ($K_{Arom,C-AR} \approx 0.82$). At the same time, the presence of non-zero off-diagonal terms, most notably $K_{Arom,C-AL} \approx 0.26$, and $K_{COx,C-AR} \approx 0.18$, indicates non-negligible cross-group redistribution within the CBG kernels, shedding light on the magnitude of relevant aliphatic cyclization and gasification processes taking place in the feedstock conversion, respectively.

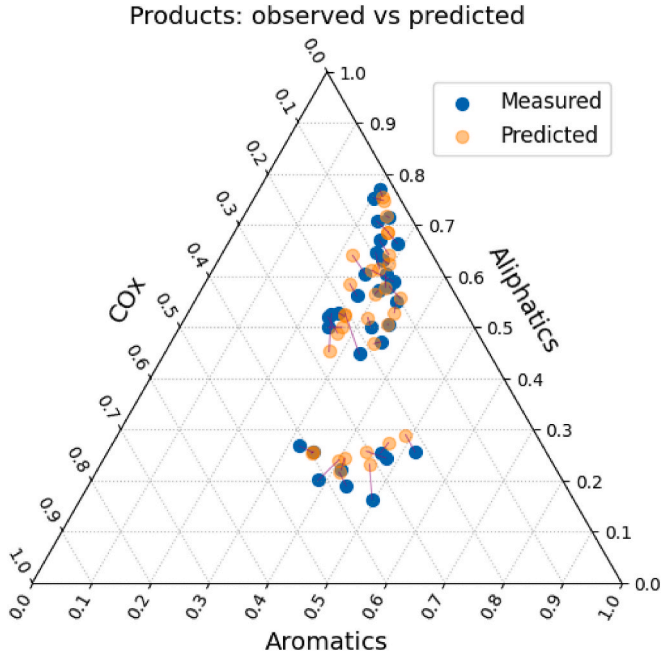

Fig. 11. Measured and predicted yields by the temperature dependent CBG model in the products simplex space.

4.4. CBG model Temperature-Dependance implementation

Using the condition-dependent CBG formulation defined in Section 2.3, the joint mapping was fitted across all investigated feedstocks and temperature regimes presented in Table 3, with a total number of N = 32 datasets. In this generalized model, reactor's temperature is incorporated as an explanatory feature, allowing the conversion matrix $\boldsymbol{K}$ to evolve systematically with it while preserving the structural interpretation of the CBG framework established for fixed conditions. Fig. 11 presents the distribution of experimentally measured and model-predicted product yields in the ternary simplex defined by the three product groups: COx, Aliphatics, and Aromatics. Each point represents an experiment, and the simplex representation highlights the constrained nature of the output space, in which carbon is redistributed among the three product groups.

From Fig. 11 it is possible to observe that the experimental data exhibit structured clustering within the simplex, reflecting systematic shifts in carbon allocation with the varied operating conditions and feedstock composition. The predicted points generated by the generalized CBG model closely overlap the measured distribution, reproducing both the overall distribution within the simplex and the dominant clustering patterns. This indicates that the model captures not only pointwise yield values, but also the global structure of carbon partitioning across the investigated operating window.

The ability of the generalized model to preserve the simplex-spatial distribution of the data demonstrates that the temperature-dependent CBG mapping provides a coherent and internally consistent description of carbon allocation among product families under varying conditions.

Quantitatively, the generalized model achieves in-sample coefficients of determination R^2 of 0.977, 0.972, and 0.903 for COx, aliphatics, and aromatics, respectively. The corresponding MAE values are 0.0126, 0.0230, and 0.0218, while the RMSE values are 0.0177, 0.0302, and 0.0262. The relative MAE values are 8.2 %, 6.4 %, and 7.2 %, respectively. The results obtained for MAE and RMSE indicate that the model captures the products variability with limited average error and the model errors are not dominated by extreme outliers. However, the slightly larger RMSE values may reflect the influence of a small number of relatively higher-deviation cases. Nonetheless, the low Rel-MAE confirms good agreement in general between model's predictions and measurements, with a limited discrepancy considering the whole carbon balance picture.

Although there is still ~10 % of the aromatics variance that remains unexplained, the model's performance increased slightly with respect to the fixed-regime case (from 0.829 to 0.903). This outcome might come in a large part from the incorporation of new data into the model. However, the consideration of temperature variations in the formulation as an explicit variable through the nonlinear SoftMax link can definitely play a role in improving the model's performance, especially since the aromatics formation is sensitive to the system's temperature level.

The predictive accuracy of the generalized CBG model can be further assessed through parity plots comparing observed and predicted yields for each product family, as shown in Fig. 12.

Across the full dataset, Fig. 12 shows that the model exhibits strong agreement with experimental values for COx and aliphatics, with predictions closely aligned along the ideal parity line and with limited dispersion. For all three product groups, the majority of data points lie close to the parity line and within ± 15 % relative deviation, indicating that the generalized CBG model captures the dominant variability over the full yield range.

The largest deviations are observed for a limited number of experiments, primarily affecting aromatic yields and, to a lesser extent, aliphatic and COx yields at high severity. These points correspond predominantly to operating conditions near the upper bounds of the investigated temperature range and to feedstocks characterized by more complex chemical compositions. Under such conditions, competing radical-driven pathways, such as secondary aromatization, and gasification, become increasingly coupled, and their net effect may be only partially resolved within the reduced three-group CBG abstraction. However, the deviations do not exhibit a systematic bias across feedstock classes or temperature levels. In fact, neighboring data points from similar feedstock (e.g. PSR3D) and temperature conditions presented substantially better agreement, which provides an indication that local data sparsity and experimental variability may also contribute to the observed discrepancies, motivating future data acquisition at same conditions to further validate such outcomes.

In this context, the generalized CBG model provides not only predictive capability, but also a way for identifying operating regimes where additional data would be most valuable. Future data acquisition at such boundary conditions can therefore serve a dual purpose: improving model robustness and enabling systematic validation of experimental measurements, which also highlights the potential of the framework as a data-quality diagnostic tool.

4.4.1. *Residuals distribution assessment and model performance tests*

In terms of the error structure of the condition-dependent CBG model, Fig. 13 shows the residuals, calculated as predicted minus measured product-groups, as a function of the measured product-group fractions.

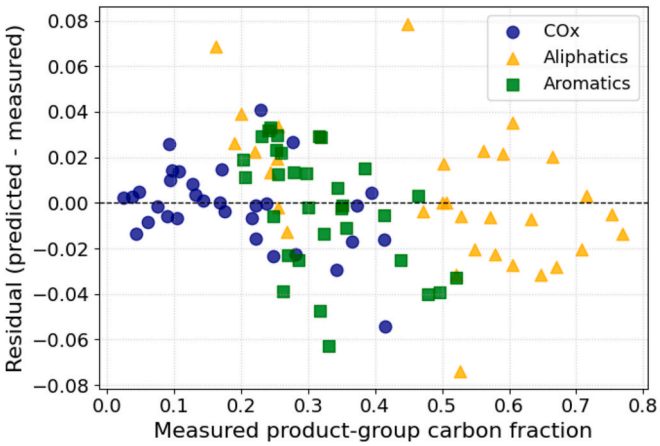

Fig. 13. Residual-errors distribution as a function of measured product group fraction for the condition-dependent CBG model.

Overall, the residuals are distributed around zero and do not indicate a strong systematic overprediction or underprediction across the evaluated range. COx shows the narrowest dispersion, whereas aliphatics and aromatics show broader residual spread, consistent with their larger RMSE values. In particular, aromatics show the broadest dispersion and a mild tendency toward negative residuals at higher measured aromatic fractions, indicating that high-aromatic-yield cases may be slightly underestimated at the current level of model abstraction. However, these deviations remain moderate in absolute carbon-fraction terms. This behavior is consistent with the relatively low R^2 for aromatics (see Table 4) and with the greater sensitivity of aromatic formation to secondary cyclization, aromatization, and gasification pathways which may become more relevant for the high aromatic yield cases. Overall, this analysis supports the reliability of the model over the evaluated range while identifying compositionally complex and high-aromatic-yield cases as priority regions for future data acquisition and model refinement.

To assess the robustness of the generalized CBG model and its ability to generalize beyond the fitted data, case-level cross-validation (CV) analyses were first performed using both k-fold and leave-one-out (LOO) strategies. No duplicated case or replicate of the same operating point was included simultaneously in the training and validation subsets and all scaling transformations were fitted within each training subset and then applied to the corresponding validation subset. For consistency with the grouped validation described below, the main cross-validation metrics were calculated as pooled out-of-fold metrics, obtained by collecting the validation predictions from all folds and then computing the

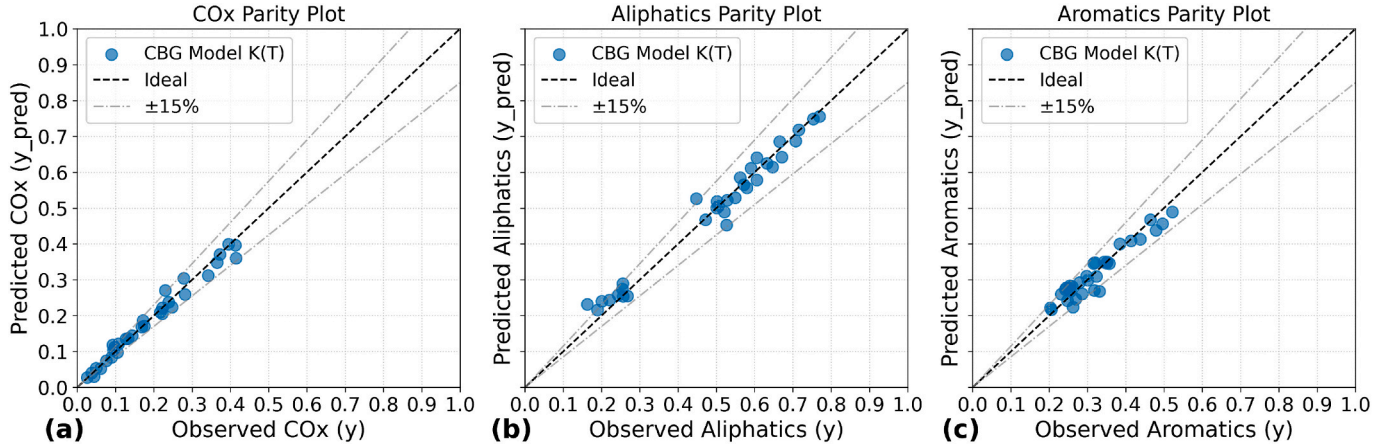

Fig. 12. Measure vs predicted parity plots for COx, Aliph and Arom groups.

Table 4
Summary of models' performance tests. Metrics values correspond to [COx, Aliph, Arom] respectively.

Model	Validation Type	R2	MAE	Rel-MAE	RMSE
Fixed-Regime CBG	In-sample	[0.994, 0.981, 0.829]	[0.008, 0.019, 0.023]	[6.4, 3.9, 8.0]	[0.010, 0.022, 0.026]
Temperature-Dependent CBG	In-sample	[0.977, 0.972, 0.903]	[0.013, 0.023, 0.022]	[8.2, 6.4, 7.2]	[0.018, 0.030, 0.026]
Temperature-Dependent CBG	5-fold CV	[0.962, 0.969, 0.855]	[0.017, 0.025, 0.027]	[10.32, 6.88, 8.78]	[0.023, 0.032, 0.032]
Temperature-Dependent CBG	LOO	[0.968, 0.969, 0.873]	[0.015, 0.025, 0.025]	[9.69, 6.79, 8.26]	[0.021, 0.032, 0.03]
Temperature-Dependent CBG	Grouped CV by CBG vector	[0.961, 0.966, 0.847]	[0.017, 0.026, 0.028]	[11.57, 7.16, 9.22]	[0.023, 0.033, 0.033]

metrics over the combined validation set. The case-level k-fold results, with k = 5, yield R^2 values of 0.962, 0.969, and 0.855 for COx, aliphatics, and aromatics, respectively, with corresponding RMSE values of 0.0228, 0.0321, and 0.0320. The LOO analysis gave comparable predictive performance, with pooled relative errors of 9.7 %, 6.8 %, and 8.3 % respectively, indicating that the model maintains stable performance when individual data points are excluded from training.

To assess the potential influence of repeated feedstock CBG vectors appearing in both training and validation subsets, an additional grouped cross-validation was performed for the condition-dependent model. In this validation, all experimental cases sharing the same rounded feedstock CBG vector $[C-xO, C-Al, C-AR]$ were assigned to the same group and kept within the same validation fold. Therefore, identical feedstock CBG vectors could not appear simultaneously in the training and validation subsets. Because the resultant CBG-vector groups contain different numbers of experimental cases, the groups were distributed across five folds using a size-balanced assignment, in which larger groups were assigned first to the fold with the smallest current number of validation cases. This kept validation fold sizes comparable while preserving the group constraint.

Pooled out-of-fold metrics of the grouped validation gave R^2 of 0.961, 0.966, and 0.847 for COx, aliphatics, and aromatics, respectively, with corresponding RMSE values of 0.0230, 0.0333, and 0.0330. These values are close to the case-level k-fold results, indicating that the predictive performance is not primarily driven by repeated feedstock CBG vectors appearing across training and validation subsets, thereby supporting the robustness and stability of the model.

Table 4 summarizes the model's overall performance. Details of the fold-resolved results are provided in Tables S2a and S2b of the Supplementary Material. Table S2a reports the case-level 5-fold results, whereas Table S2b summarizes the fold-level results from the grouped cross-validation by feedstock CBG vector. Fold-level R^2 is reported for transparency, but it should be interpreted cautiously because grouped validation folds contain restricted CBG-vector regions and may span narrow target ranges, making R^2 sensitive to small absolute errors.

The obtained validation metrics can be placed in the context of recent ML and physics-informed ML models developed for thermochemical conversion. Direct benchmarks for product-group prediction in plastic-waste steam cracking are limited; therefore, biomass gasification, waste gasification, and plastic-waste pyrolysis studies provide a close contextual comparison. These models commonly predict syngas composition, H_2 yield, liquid/gas/solid yields, tar, char, or LHV from feedstock descriptors, operating conditions, reactor variables, and, in catalytic systems, catalyst properties. Representative gasification ML studies report R^2 values of approximately 0.8–0.95 using datasets of about 222–336 samples [30–32]. Physics-informed studies have reported similarly high performance; Ren et al. [33,34] incorporated physical monotonicity constraints in PINN-based frameworks and reported average test R^2 values of 0.91–0.97, while van Wyk [35] introduced carbon mass conservation into a physics-informed ML framework for indirect gasification of waste feedstocks ranging from biomass to plastics, comparing different models such as Random Forests and XGBoost, and obtaining $R^2 = 0.95$ while improving carbon closure. Recent plastic-waste pyrolysis ML studies further show the relevance of composition-based descriptors, although their target outputs differ from the carbon allocation considered here [36,37]. Further comparison details are provided in Table S3 of the supplementary material.

Compared with these studies, the present CBG model uses a smaller dataset (N = 32) of independent semi-industrial steam-cracking cases, and a more compact input representation consisting of the feedstock CBG vector and temperature. The cross-validated performance obtained here therefore falls within the broad range reported for thermochemical ML models, particularly for COx and aliphatics, despite the reduced data and descriptor space. The lower aromatic R^2 reflects the higher complexity of secondary cyclization, aromatization, and gasification pathways.

The comparison should nevertheless be interpreted cautiously because the CBG model is not a conventional direct input–output regressor. It reformulates steam cracking as a constrained transformation between chemistry-aware feedstock and product simplex spaces. Non-negativity and carbon conservation are enforced by the column-stochastic conversion matrix, and each matrix coefficient represents the net routing of carbon from a feedstock bond group to a product family. Thus, the model is interpretable by construction, unlike post-hoc explainability methods such as feature importance or SHAP, which explain trained model sensitivity but do not impose a carbon-conserving conversion structure.

Training-Set Size Sensitivity Analysis.

To evaluate the influence of training-set size on model performance, a dataset-size sensitivity analysis was performed for the model. Random train/test splits were generated by progressively reducing the number of training cases. For each training-set size, the splitting procedure was repeated five times and the resulting RMSE values were averaged. The results are reported in Fig. 14. RMSE was selected as the main sensitivity metric because it directly quantifies prediction error in carbon-fraction units and is less affected than R^2 by changes in the variance of small validation subsets.

The results from Fig. 14 show a rapid reduction in RMSE as the training set increases from very small subsets, followed by stabilization at intermediate-to-large training sizes. For training sets smaller than $N_{train} = 9$, RMSE increases markedly, indicating that very small subsets

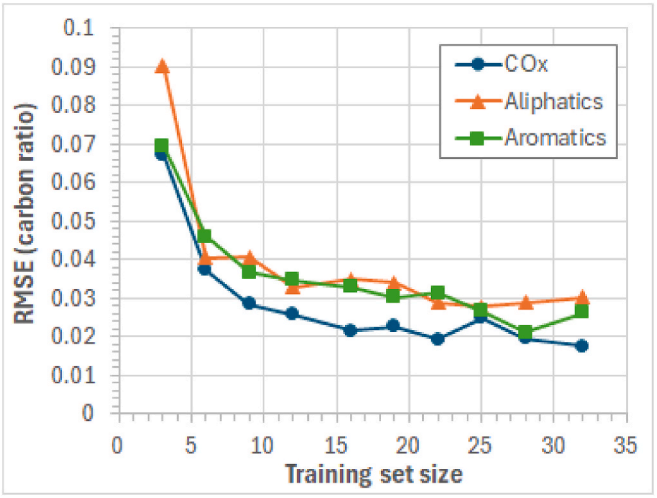

Fig. 14. Dataset-size sensitivity of the condition-dependent CBG model.

cannot provide sufficient coverage of the heterogeneous feedstock composition and temperature space. Between $N_{train} = 9$ and $N_{train} = 16$, the model shows an intermediate transition regime, in which RMSE values decrease toward a more stable range, particularly for COx and aromatics. For training sizes of $N_{train} \geq 16$, RMSE values reach relative stability and remain close to those obtained with the full dataset at N = 32.

These results suggest that the constrained CBG formulation can extract reasonably stable feedstock–product relationships from relatively small amounts of data. However, sufficient data diversity remains essential, especially for representing the aromatic product group and high-temperature conditions. Therefore, dataset size should be interpreted not only in terms of the number of experiments, but also in terms of the coverage of feedstock structural diversity and operating conditions.

Overall, the generalized condition-dependent CBG model provides a unified and predictive description of carbon partitioning among COx, aliphatics, and aromatics across heterogeneous feedstocks and operating severities. By explicitly incorporating temperature, the model improves the description of aromatic outcomes relative to a fixed-regime formulation while retaining the compactness and interpretability of the CBG framework.

Having established the predictive performance and generalization capability of the model, the following section examines how the individual conversion coefficients in the temperature-dependent matrix $\mathbf{K}(T)$ evolve with operating severity, providing insight into the temperature-driven redistribution of carbon among the product groups.

4.4.2. *Temperature dependence of the CBG conversion matrix*

The generalized condition-dependent CBG model enables direct examination of how the feedstock–product conversion process evolves with operating severity through the temperature dependence of the

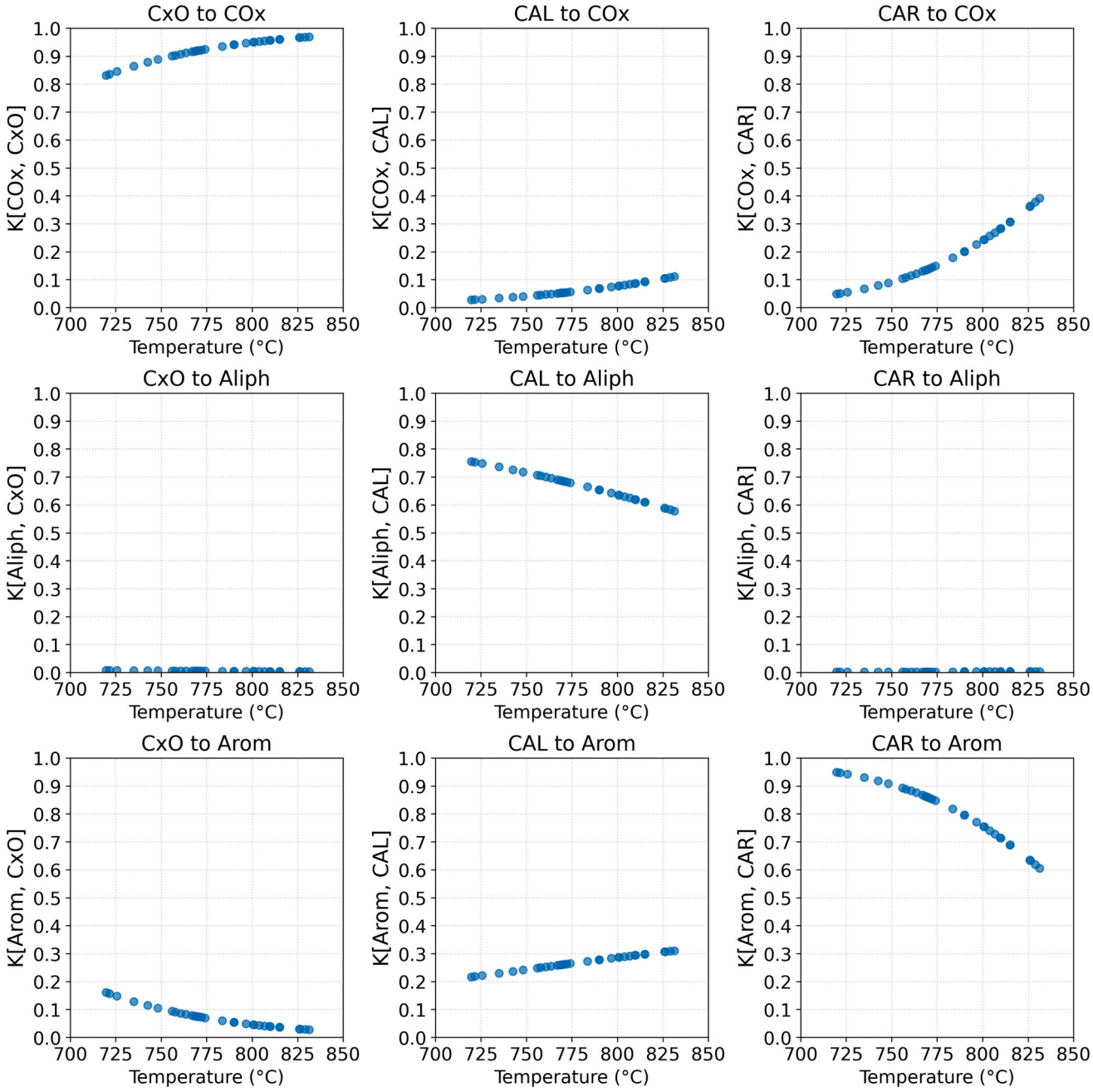

Fig. 15. Results of the conversion coefficients of K matrix as function of temperature.

conversion matrix $\mathbf{K}(T)$. Fig. 15 presents the evolution of the matrix coefficients as a function of temperature for the full set of generalized model solutions. Each panel corresponds to a specific mapping between the feedstock carbon bond (columns) and the product carbon groups (rows).

From Fig. 15, it is possible to observe that the diagonal coefficients of $\mathbf{K}(T)$ describe the dominant routing of carbon within the same bond group category in the evaluated temperature interval. For oxygen-bound carbon, the coefficient C-xO to COX in panel (a) remains high, starting from ~0.83 and takes an asymptotic increasing behavior, approaching unity, as it reaches the highest temperature. This indicates that, as the energy provided increases, the likelihood of oxygen-bound carbon to end up in carbon oxides increases and approaches 100 %. Such trend is consistent with the fact that higher system's energy favors the breaking down of the chains. At the radicals level, oxygenated functionalities form alkoxy and acyl radicals that can undergo rapid β-scission, decarbonylation, and decarboxylation [38–40]. As temperature increases, alternative stabilization pathways become negligible by the newly C-xO released radicals, funneling them almost exclusively toward CO and CO_2 formation. This behavior is also observed by the natural decreasing of the observed C-xO to Aromatics trend in panel (g).

On the other hand, the coefficient C-AL to aliphatics in panel (e), decreases almost linearly with temperature, from ~0.75 to ~0.58 in the evaluated range, reflecting the declining probability that aliphatic radicals terminate as condensed aliphatic products. At lower temperatures, a larger fraction of aliphatic carbon is retained in aliphatic products, with a significant focus towards aromatics through cyclization reactions (see panel (h)), and rather limited towards COx (panel (b)). As temperature increases, net aliphatic retention decreases and contributions towards COx become more relevant in a non-linear way. Conversion to aromatics also increases but following a milder trend closer to a concave-like behavior, which imply that eventually it may achieve a maximum point at some high temperature level while the conversion to oxides will continue to increase in relevance.

In the C-AR case, the conversion coefficient to aromatic products in panel (i) starts very high at ~0.95 and decreases with temperature reaching ~0.6 in the evaluated interval. The high initial coefficient at the low temperature case is an indication of the stability of aromatic ring, which will tend to keep its initial structural configuration. By increasing thermal energy, steam reforming reactions start to kick in dramatically, promoting gasification pathways to produce carbon oxides, as seen in panel (c), until inevitably all aromatic rings will be led to such conversion route.

Negligible conversion index are observed for C-xO to aliphatics and C-AR to aliphatics in panels (d) and (f) respectively, consistent with the mirrored behavior seen in the corresponding $\boldsymbol{K}$-matrix corner plots. Even after relaxing the C-AR to aliphatics in Eq. (8), the associated coefficients remained unchanged. This indicates that the constraint is not binding under the evaluated operating conditions. Instead, the near-zero values emerge from the inferred reaction tendencies encoded in the dataset, which reflect the intrinsic chemistry of the system.

From a more mechanistic perspective, the negligible coefficients in panels (d) and (f) of Fig. 15 are consistent with how oxygen–containing and aromatic carbon groups behave under non–catalytic steam–cracking conditions. As mentioned earlier, oxygenated fragments typically undergo decarbonylation or decarboxylation pathways, forming radicals that are short-lived and may not readily recombine into stable hydrocarbon chains. The carbon is thus mainly funneled towards carbon oxides rather than into C–C chain-forming pathways, and only a minor fraction appear to be routed into generating structural motifs for aromatic precursors (see panel (g)). Likewise, carbon in aromatic rings is largely retained in aromatic structures or being oxidized rather than yield stabilized aliphatic chains. In the absence of catalytic functionalities capable of hydrogen transfer, selective aromatic ring hydrogenation, or ring–opening reactions, C-AR to aliphatics is expected to remain thermodynamically and kinetically disfavored [40,41]. A noticeable increase in these coefficients might be expected in case catalytic effects start to promote alternative reaction pathways that facilitate aliphatic formation (such as ring hydrogenation, ring opening reactions, etc.), which seems to be inactive in the present system.

The results presented in Fig. 11 and Fig. 15 demostrate that the CBG model can capture underlying structure–product relationships at this reduced level of abstraction. The temperature–dependent behavior of the $\boldsymbol{K}(T)$ matrix coefficients reveal consistent and chemically interpretable shifts in carbon redistribution among the evaluated groups. Overall, as temperature increases, the observed trends of $\boldsymbol{K}(T)$ suggest a suppression of pathways leading to stabilized hydrocarbon fragments and a stronger routing of carbon into COx, signaling a drift toward more oxidizing, gasification–like behavior. These trends and carbon redistribution patterns are in general consistent with known reaction tendencies in such high temperature steam environments.

Within the reduced CBG abstraction, the generalized model provides therefore a mechanistically interpretable summary of how the process conversion chemistry drives carbon redistribution among COx, aliphatics, and aromatics as operating severity increases from the reactor's temperature.

While the present validation focuses on datasets with sufficiently high-precision, preliminary tests have also been conducted on earlier experimental campaigns involving highly heterogeneous feedstocks, including materials such as automotive shredder residue (ASR). These legacy datasets were generated prior to the implementation of the current measurement framework and therefore may exhibit larger uncertainty in input composition and output carbon verification. Although initial application of the CBG model to such datasets indicates qualitatively consistent trends, further experiments with improved analytical verification are being prepared before rigorous quantitative validation can be reported. Extending the model assessment to such type of increasingly heterogeneous real-world waste streams remains thus an important direction for future work.

Final remarks.

Beyond the demonstrated predictive and interpretive capabilities, it is important to note that $\boldsymbol{K}$ represents the transformation between the feed and product simplex spaces. Its entries capture the reactor's averaged conversion behavior in a reduced, yet chemically meaningful, form. As the model improves its ability to reproduce trends across diverse feedstocks and temperature conditions, the mapping embodied in $\boldsymbol{K}$ can become more reflective of the reactor's intrinsic conversion behavior. At the same time, the constrained operator formulation does not explicitly encode reactor geometry, residence time distributions, or detailed kinetic parameters. Consequently, the model's mathematical framework is transferable and can be re-parameterized for other steam cracking systems by re-estimating the corresponding conversion matrix from representative data for the new reactor configuration.

With a sufficiently robust estimation of $\boldsymbol{K}$, the generalized CBG model can provide a scalable foundation for the development of a reduced-order digital twin concept for thermochemical conversion systems handling heterogeneous feedstocks. By expressing carbon redistribution through a compact, interpretable, and condition-dependent mapping $\boldsymbol{K}(\boldsymbol{c})$, the framework enables rapid evaluation of feedstock–process–product interactions without resorting to full detailed kinetic mechanisms or complex species-resolved simulations (See Fig. 16).

The modular structure of the CBG formulation allows for systematic extension in several directions. Additional operating variables, such as residence time, steam-to-carbon ratio, heating-rate-related descriptors, bed/catalytic activity indicators, etc., can be incorporated as conditioning features in $\boldsymbol{K}(\boldsymbol{c})$, while finer carbon-bond group resolutions in the inputs or outputs can be introduced to target specific chemical families or functional groups. Moreover, operating time could be incorporated as an additional conditioning variable in future digital-twin implementations. This would help capture slow process drift, including bed-material aging with ash accumulation, changes in oxygen-

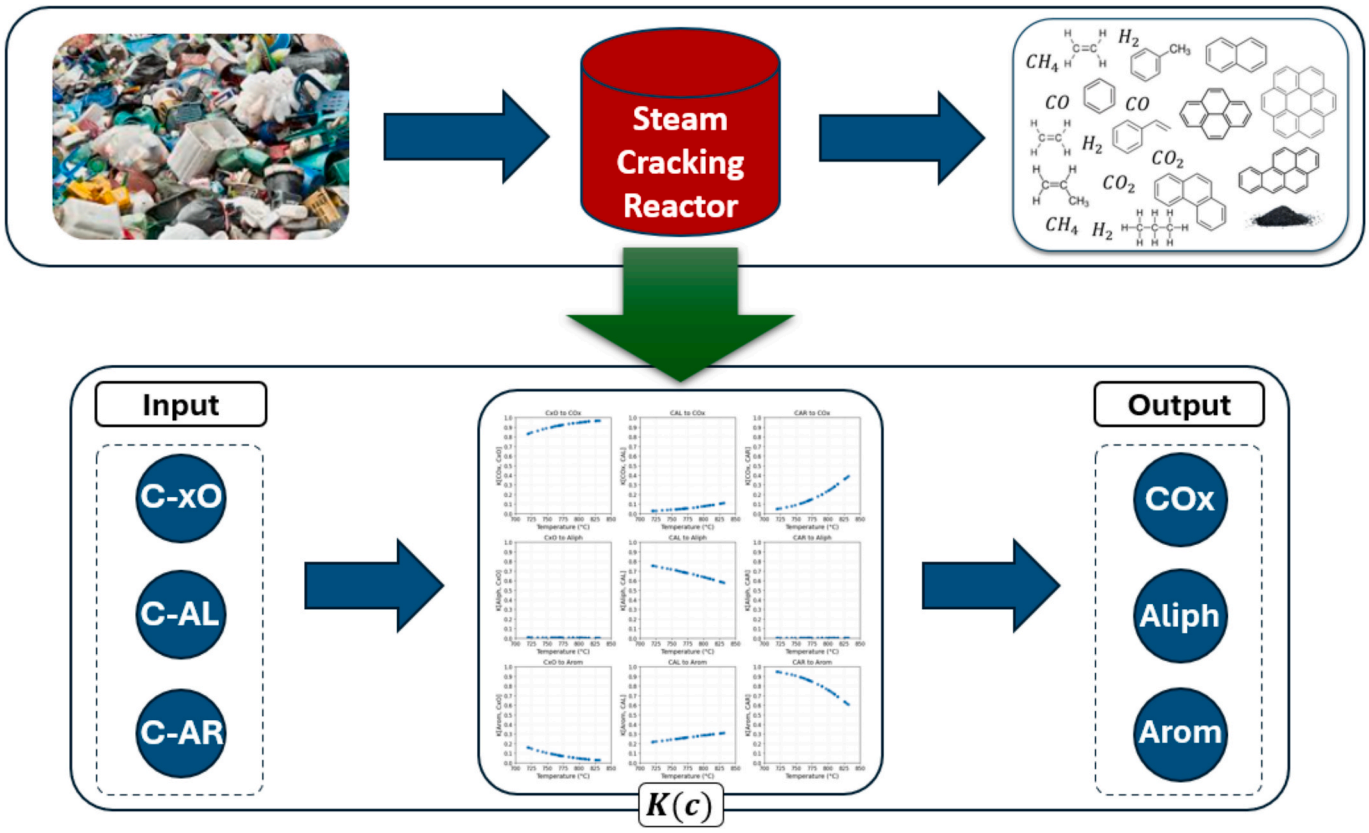

Fig. 16. Schematic of K matrix as a reduced-dimensional digital-twin representation of the steam cracking reactor.

transfer behavior, or other long-term changes during uninterrupted operation.

The explicit representation of conditions-dependent conversion coefficients provides a natural interface for coupling the model with process simulators, optimization routines, and real-time data streams, supporting applications in process design, control, and feedstock valorization. In this context, the CBG-based mapping can be viewed as a physics-informed surrogate model that can bridge experimental data with process-level decision making. Its ability to capture dominant radical-driven redistribution pathways in a reduced yet mechanistically interpretable form, positions it as a promising foundation for future digital twin implementations aimed at managing feedstock variability and operating uncertainty in advanced thermochemical conversion processes.

5. Conclusions

This work introduced a physics-informed Carbon Bond Group (CBG) modeling framework for describing feedstock–product correlations and predicting product-group carbon distributions in steam cracking of heterogeneous polymeric waste. By representing feedstocks and products in chemically meaningful compositional spaces and formulating thermochemical conversion as a physics-aware mathematical transformation between them, the model provides a reduced yet physically interpretable representation of the conversion process.

The condition-dependent implementation, using temperature as the explicit operating variable, was trained and evaluated with 32 independent experimental cases from a semi-industrial DFB steam cracker covering diverse polymeric waste streams and a broad temperature range. The model reproduced the evaluated product-group distributions with high accuracy, reaching in-sample R^2 values of 0.977, 0.972, and 0.903 for COx, aliphatics, and aromatics, respectively, and maintained robust performance under cross-validation. The residual analysis showed no strong systematic bias across the evaluated range, although high-aromatic-yield and compositionally complex cases remain priority regions for future data acquisition.

Beyond prediction, the temperature-dependent conversion matrix $\boldsymbol{K}(T)$ provided chemically consistent insights into carbon redistribution with increasing reactor temperature, including enhanced routing of oxygen-bound carbon toward COx, reduced aliphatic retention, and increased gasification of aromatic structures at higher temperatures. These behaviors emerged directly from the constrained estimation procedure, demonstrating that the reduced-order mapping captures dominant carbon-redistribution tendencies consistent with the process chemistry, without requiring explicit species-level kinetics, while remaining computationally compact.

Overall, these results show that the CBG framework provides a scalable foundation for reduced-order representations of thermochemical conversion processes involving heterogeneous feedstocks. Future extensions should focus on expanding the dataset, incorporating additional structural and operating-condition descriptors, and coupling the CBG mapping with process simulation, optimization, and digital-twin applications.

CRediT authorship contribution statement

Renesteban Forero-Franco: Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Teresa Berdugo-Vilches:** Writing – review & editing, Visualization, Investigation. **Chahat Mandviwala:** Writing – review & editing, Validation, Resources, Investigation. **Nidia Díaz Perez:** Resources, Investigation. **Isabel Cañete-Vela:** Resources, Investigation. **Henrik Thunman:** Writing – review & editing, Project administration, Funding acquisition. **Martin Seemann:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by Energymyndighet project P2024-01719, the project framework Klimatledande Processindustri etapp2 financed by Vinnova under agreement 2022–01445, and Borealis AB, Sweden (Project number: 49514–1). The authors thank Jessica Bohwalli, Johannes Öhlin and Rustan Hvitt for technical support during the experiments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fuel.2026.140963.

Data availability

Data will be made available on request.

References

[1] Houssini K, Li J, Tan Q. Complexities of the global plastics supply chain revealed in a trade-linked material flow analysis. Commun Earth Environ 2025;6:257. https://doi.org/10.1038/s43247-025-02169-5.
[2] Dokl M, Copot A, Krajnc D, Van Fan Y, Vujanović A, Aviso KB, et al. Global projections of plastic use, end-of-life fate and potential changes in consumption, reduction, recycling and replacement with bioplastics to 2050. Sustain Prod Consum 2024;51:498–518. https://doi.org/10.1016/j.spc.2024.09.025.
[3] OECD Global Plastics Outlook. Policy Scenarios to 2060, OECD Publishing. Paris 2022. https://doi.org/10.1787/aa1edf33-en.
[4] Ragaert K, Delva L, Van Geem K. Mechanical and chemical recycling of solid plastic waste. Waste Manag 2017;69:24–58. https://doi.org/10.1016/j.wasman.2017.07.044.
[5] Anglou E, Boukouvala F, Stathatou PM. A cradle-to-gate life cycle assessment of mechanochemical recycling of poly(ethylene terephthalate) waste. Chem Eng J 2025;524:168719. https://doi.org/10.1016/j.cej.2025.168719.
[6] Yin Y, Wang S, Zhao L, Zhang W, Gao W, Gu B, et al. A review of high-grade pyrolysis oils from plastic waste: Mechanisms, process strategies, and sustainability perspectives. Renew Sustain Energy Rev 2026;238:117035. https://doi.org/10.1016/j.rser.2026.117035.
[7] Lv S, Wu Z, Yang J, Zhao W, Gao A, Gao W, et al. Pyrolytic oil from biomass, plastics, and tires: production, properties, and tailored upgrading strategies for sustainable application. Renew Sustain Energy Rev 2026;226:116480. https://doi.org/10.1016/j.rser.2025.116480.
[8] W. Kaminsky, The Hamburg Fluidized-bed Pyrolysis Process to Recycle Polymer Wastes and Tires, in: Feedstock Recycling and Pyrolysis of Waste Plastics, Wiley, 2006: pp. 475–491. Doi: 10.1002/0470021543.ch17.
[9] Mandviwala C, Forero Franco R, Berdugo Vilches T, Gogolev I, González-Arias J, Cañete Vela I, et al. Steam cracking in a semi-industrial dual fluidized bed reactor: Tackling the challenges in thermochemical recycling of plastic waste. Chemical Engineering Journal 500 2024:156892. https://doi.org/10.1016/j.cej.2024.156892.
[10] Thunman H, Berdugo Vilches T, Seemann M, Maric J, Vela IC, Pissot S, et al. Circular use of plastics-transformation of existing petrochemical clusters into thermochemical recycling plants with 100% plastics recovery. Sustain Mater Technol 2019;22:e00124. https://doi.org/10.1016/j.susmat.2019.e00124.
[11] Regulation (EU) 2025/40 on packaging and packaging waste, amending Regulation (EU) 2019/1020 and Directive (EU) 2019/904, and repealing Directive 94/62/EC, 2024. http://data.europa.eu/eli/reg/2024/1781/oj.
[12] Directive (EU) 2019/904 on the reduction of the impact of certain plastic products on the environment, (2019). http://data.europa.eu/eli/dir/2019/904/oj (accessed February 19, 2026).
[13] Levine SE, Broadbelt LJ. Detailed mechanistic modeling of high-density polyethylene pyrolysis: Low molecular weight product evolution. Polym Degrad Stab 2009;94:810–22. https://doi.org/10.1016/J.POLYMDEGRADSTAB.2009.01.031.
[14] Forero-Franco R, Cañete-Vela I, Berdugo-Vilches T, González-Arias J, Maric J, Thunman H, et al. Correlations between product distribution and feedstock composition in thermal cracking processes for mixed plastic waste. Fuel 2023;341:127660. https://doi.org/10.1016/j.fuel.2023.127660.
[15] Van Geem KM, Reyniers MF, Marin GB. Challenges of Modeling Steam Cracking of Heavy Feedstocks. Oil & Gas Science and Technology - Revue de l'IFP 2008;63:79–94. https://doi.org/10.2516/ogst:2007084.
[16] Dai F, Wang H, Gong M, Li C, Li Z, Zhang S. Carbon-Number-based Kinetics, Reactor Modeling, and Process simulation for Coal Tar Hydrogenation. Energy Fuel 2015;29:7532–41. https://doi.org/10.1021/acs.energyfuels.5b01476.
[17] Zhang X, Sarathy SM. A functional-group-based approach to modeling real-fuel combustion chemistry – II: Kinetic model construction and validation. Combust Flame 2021;227:510–25. https://doi.org/10.1016/j.combustflame.2020.10.039.
[18] Plehiers PP, Symoens SH, Amghizar I, Marin GB, Stevens CV, Van Geem KM. Artificial Intelligence in Steam Cracking Modeling: a Deep Learning Algorithm for Detailed Effluent Prediction. Engineering 2019;5:1027–40. https://doi.org/10.1016/j.eng.2019.02.013.
[19] Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. Nat Rev Phys 2021;3:422–40. https://doi.org/10.1038/s42254-021-00314-5.
[20] Maria Tronci E, Downey ARJ, Mehrjoo A, Chowdhury P, Coble D. Physics-Informed Machine Learning Part I: different strategies to Incorporate Physics into Engineering Problems. In: Data Science in Engineering, Vol. 10. New York: River Publishers; 2025. p. 1–6. https://doi.org/10.1007/978-3-031-68142-4_1.
[21] Forero-Franco R, Cañete-Vela I, Berdugo-Vilches T, Mandviwala C, Perez ND, Gogolev I, et al. Towards sustainable textile waste management: Exploring valuable chemicals production through steam cracking in a dual fluidized bed. Fuel 2025;397:135731. https://doi.org/10.1016/j.fuel.2025.135731.
[22] C. Mandviwala, R. Forero Franco, I. Gogolev, J. González-Arias, T. Berdugo Vilches, I.C. Cañete Vela, H. Thunman, M. Seemann, Method development and evaluation of product gas mixture from a semi-industrial scale fluidized bed steam cracker with GC-VUV, Fuel Processing Technology 253 (2024) 108030. Doi: 10.1016/j.fuproc.2023.108030.
[23] Israelsson M, Seemann M, Thunman H. Assessment of the Solid-phase Adsorption Method for Sampling Biomass-Derived Tar in Industrial Environments. Energy Fuel 2013;27:7569–78. https://doi.org/10.1021/ef401893j.
[24] Israelsson M, Larsson A, Thunman H. Online Measurement of Elemental yields, Oxygen Transport, Condensable Compounds, and heating Values in Gasification Systems. Energy Fuel 2014;28:5892–901. https://doi.org/10.1021/ef501433n.
[25] Larsson A, Seemann M, Neves D, Thunman H. Evaluation of Performance of Industrial-Scale dual Fluidized Bed Gasifiers using the Chalmers 2–4-MW $_{th}$ Gasifier. Energy Fuel 2013;27:6665–80. https://doi.org/10.1021/ef400981j.
[26] Berdugo Vilches T, Thunman H. Experimental Investigation of Volatiles–Bed Contact in a 2–4 MWth Bubbling Bed Reactor of a dual Fluidized Bed Gasifier. Energy Fuel 2015;29:6456–64. https://doi.org/10.1021/acs.energyfuels.5b01303.
[27] Forero-Franco R, Berdugo-Vilches T, Mandviwala C, Seemann M, Thunman H. Developing a parametric system model to describe the product distribution of steam pyrolysis in a dual Fluidized bed. Fuel 2023;348:128518. https://doi.org/10.1016/j.fuel.2023.128518.
[28] Sun J, Ding T, Zhao X, Wang P, Chen W, Yu J. Dehydrochlorination of PVC at low-temperatures and the pyrolysis behavior of dechlorinated PVC. Waste Dispos Sustain Energy 2025. https://doi.org/10.1007/s42768-025-00252-3.
[29] Zhou H, Wu C, Onwudili JA, Meng A, Zhang Y, Williams PT. Influence of process conditions on the formation of 2–4 ring polycyclic aromatic hydrocarbons from the pyrolysis of polyvinyl chloride. Fuel Process Technol 2016;144:299–304. https://doi.org/10.1016/J.FUPROC.2016.01.013.
[30] Xue P, Chen T, Huang X, Hu Q, Hu J, Zhang H, et al. Prediction of syngas properties of biomass steam gasification in fluidized bed based on machine learning method. Int J Hydrogen Energy 2024;49:356–70. https://doi.org/10.1016/j.ijhydene.2023.08.259.
[31] Kim JY, Kim D, Li ZJ, Dariva C, Cao Y, Ellis N. Predicting and optimizing syngas production from fluidized bed biomass gasifiers: a machine learning approach. Energy 2023;263:125900. https://doi.org/10.1016/j.energy.2022.125900.
[32] Ascher S, Wang X, Watson I, Sloan W, You S. Interpretable machine learning to model biomass and waste gasification. Bioresour Technol 2022;364:128062. https://doi.org/10.1016/j.biortech.2022.128062.
[33] Ren S, Wu S, Weng Q. Physics-informed machine learning methods for biomass gasification modeling by considering monotonic relationships. Bioresour Technol 2023;369:128472. https://doi.org/10.1016/j.biortech.2022.128472.
[34] Ren S, Wu S, Weng Q, Zhu B, Deng Z. Disentangled Representation aided Physics-Informed Neural Network for predicting Syngas Compositions of Biomass Gasification. Energy Fuel 2024;38:2033–45. https://doi.org/10.1021/acs.energyfuels.3c03496.
[35] van Wyk S. Development of a novel physics-informed machine learning model for advanced thermochemical waste conversion. Chemical Engineering Journal Advances 2025;21:100699. https://doi.org/10.1016/j.ceja.2024.100699.
[36] Cheng Y, Ekici E, Yildiz G, Yang Y, Coward B, Wang J. Applied machine learning for prediction of waste plastic pyrolysis towards valuable fuel and chemicals production. J Anal Appl Pyrol 2023;169:105857. https://doi.org/10.1016/j.jaap.2023.105857.

[37] Xu D, Zhang Y, Wang S, Jiang J. Harnessing the optimization strategy via interpretable machine learning framework for high-efficiency pyrolysis-catalysis engineering of plastic wastes. Energy Convers Manag 2026;348:120770. https://doi.org/10.1016/j.enconman.2025.120770.
[38] Çit İ, Sınağ A, Yumak T, Uçar S, Mısırlıoğlu Z, Canel M. Comparative pyrolysis of polyolefins (PP and LDPE) and PET. Polym Bull 2010;64:817–34. https://doi.org/10.1007/s00289-009-0225-x.
[39] Liu C, Huang J, Huang X, Li H, Zhang Z. Theoretical studies on formation mechanisms of CO and CO2 in cellulose pyrolysis, Comput. Theor Chem 2011;964:207–12. https://doi.org/10.1016/j.comptc.2010.12.027.
[40] Gebre SH, Sendeku MG, Bahri M. Recent Trends in the Pyrolysis of Non-Degradable Waste Plastics. ChemistryOpen 2021;10:1202–26. https://doi.org/10.1002/open.202100184.
[41] Carrasco Díaz A, Abdelouahed L, Brodu N, Montes-Jiménez V, Taouk B. Upgrading of Pyrolysis Bio-Oil by Catalytic Hydrodeoxygenation, a Review Focused on Catalysts, Model Molecules, Deactivation, and Reaction Routes. Molecules 2024;29:4325. https://doi.org/10.3390/molecules29184325.