Data-driven Assessment of Desulfurization Fly Ash Catalysis in Cotton-stalk and Model-compound Pyrolysis

Jiuwei Huang*
Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia
*Corresponding email: jiuwei.lab@gmail.com

Abstract

Desulfurization fly ash (DFA) is an alkaline mineral by-product with potential value as a low-cost catalyst for biomass pyrolysis upgrading. The original study was organized as a mathematical-modeling contest report, which fragmented the scientific logic and mixed descriptive analysis with unsupported high-order fitting claims. In the present manuscript, the study is reconstructed as a journal-style article centered on a single mechanistic question: How does DFA loading redistribute tar, water, char and syngas yields, and how do the gas fingerprints differ among cotton stalk (CS), cellulose model compound (CMCpd) and lignin model compound (LMC)? Product-yield values were reconstructed from the reported equations and figures in the source report, whereas gas-yield values were taken from the reported table for CS and digitized from the original gas-yield plots for CMCpd and LMC when editable spreadsheets were unavailable. The reanalysis reveals three distinct response modes. In CS pyrolysis, increasing DFA suppresses tar and markedly enriches H2, indicating a shift from condensable to gaseous products. CMCpd shows the strongest H_2-selective pathway, but tar suppression is not observed; instead, the system moves toward hydrogen-rich yet tar-promoted behavior. LMC remains char-dominated and CO_2- and CH_4-rich, showing only limited hydrogen upgrading. A mechanism-guided and leave-one-out cross-validated interpretation is proposed in place of over-parameterized polynomial or generic artificial-intelligence fitting. The results support the view that DFA can promote cracking, deoxygenation and gas reforming, but that the attainable upgrading route depends strongly on feedstock structure.

Keywords

Biomass pyrolysis, Cotton stalk, Desulfurization fly ash, Cellulose, Lignin, Catalytic upgrading, Hydrogen enrichment, Small-sample modelling

Introduction

The thermochemical conversion of agricultural residues has become an increasingly important route for obtaining renewable fuels, platform chemicals and carbonaceous materials [1]. Cotton stalk is a representative lignocellulosic residue and a promising pyrolysis feedstock because it is abundant, inexpensive and underutilized [2]. However, pyrolysis product distribution is highly sensitive to feedstock composition, thermal conversion pathway and kinetic behavior [3,4].

Desulfurization fly ash (DFA) is an industrial by-product rich in inorganic and alkaline mineral phases. Its possible reuse as a catalytic additive is attractive because it may simultaneously reduce waste-disposal pressure and enhance biomass conversion. Existing reports show that desulfurized fly ash can influence cracking, deoxygenation and gas-reforming reactions, thereby shifting the balance among tar, char and permanent gases [5].

The source report addressed this topic as a Shuwei Cup mathematical-modelling study. Although the original work contained useful observations, its contest-style organization obscured the research hypothesis and overstated the predictive strength of a very small dataset. The present manuscript therefore reconstructs the study into a journal format that emphasizes mechanism -guided interpretation, data provenance and visually clearer figures.

Study objective and contribution

The revised objective is to determine how DFA loading changes product and gas fingerprints of cotton stalk

(CS), cellulose model compound (CMCpd) and lignin model compound (LMC), and identify which response mode is practically most favorable for gas upgrading.

Compared with the original report, the present manuscript makes four improvements: (1) The five contest questions are integrated into a single hypothesis-led narrative. (2) All redrawn figures are standardized and data-audited. (3) The previous visually odd multivariate PCA map is replaced by a directly interpretable operating map. (4) The modelling discussion is constrained by mechanism and small-sample uncertainty rather than by high-order curve fitting alone [6].

Materials and methods

Data source and reconstruction strategy

The analysis is based on the Portable Document Format (PDF) source document provided by the author. Product-yield series for CS, CMCpd and LMC were reconstructed from the polynomial equations reported in the source text and checked against the original bar charts. Gas-yield series for CS were taken from the reported table, while CMCpd and LMC gas values were digitized from the original figure panels because editable data tables were not provided.

The revised manuscript avoids any claim that reconstructed values are equivalent to raw experimental measurements. Instead, the figures are designed to preserve the trend structure visible in the source document while clearly documenting their origin. All reconstructed values used in the figures are included in the accompanying data package.

Analysis workflow

Figure 1 briefly summarizes the restructured workflow. The analysis begins with three distinct feedstock layers (CS, CMCpd and LMC), introduces DFA as the catalytic factor in fixed-bed pyrolysis, and then traces the resulting changes in product and gas outputs. The final interpretation layer integrates data auditing, effect-size normalization, mechanism-consistent fitting, leave-one-out cross-validation (LOOCV) and response-map visualization.

This workflow replaces the old task-by-task setup and makes scientific logic explicit: Statistical and graphical analysis aims to test mechanistic hypotheses rather than maximize curve-fitting complexity.

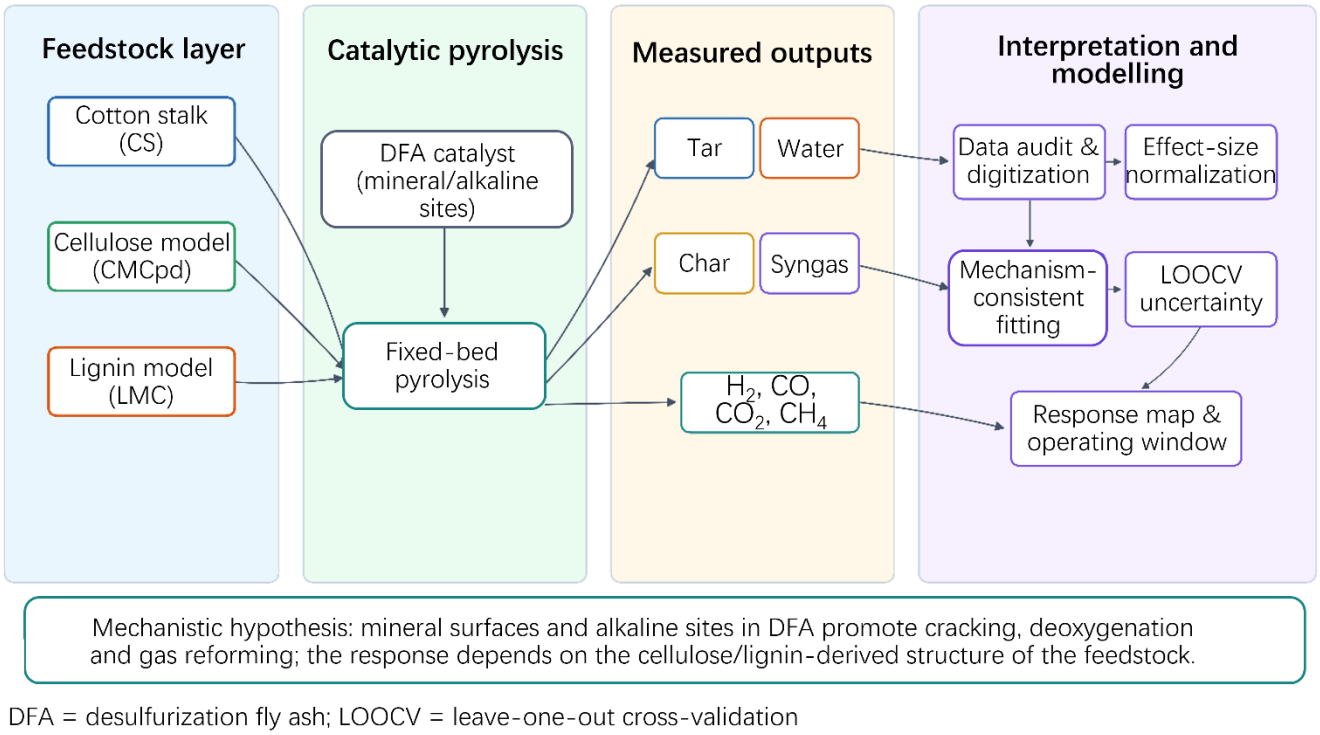

Figure 1. Reframed workflow and mechanism-guided data-analysis structure.

Statistical and graphical principles

Because the dataset is small and partly reconstructed, the revised manuscript prioritizes transparent descriptive analysis, effect-size interpretation and low-order interpolation over formal significance claims that would require replicate structure. Product yields are visualized through stacked distributions and selectivity ratios, gas yields through feedstock-specific response

curves, and integrated trends through normalized effect-size and a direct tar-H_2 operating map. This choice is more appropriate than black-box predictive framing, because machine-learning models for biomass pyrolysis generally rely on substantially larger and richer datasets [7,8].

For predictive stability, a LOOCV procedure was applied to low-order polynomial models [9]. This approach does not claim broad predictive power; rather, it evaluates whether the observed trends remain stable when individual points are omitted (see Table 1 for data provenance and confidence grading).

Table 1. Data provenance and confidence audit.

Dataset	Reconstruction source	Confidence
Product yields, CS	Reported cubic equations and original bar chart	High
Product yields, CMCpd	Reported cubic equations and original bar chart	Medium
Product yields, LMC	Reported cubic equations and original bar chart	Medium
Gas yields, CS	Printed table in source PDF	High
Gas yields, CMCpd	Digitized from embedded gas-yield figure	Medium
Gas yields, LMC	Digitized from embedded gas-yield figure	Medium
Figure 5 operating-map metrics	Calculated from reconstructed products and gas series	High

Results and discussion

DFA redistributes condensed and gaseous products in a feedstock-dependent manner

Figure 2 presents reconstructed product distributions of CS, CMCpd and LMC. Table 2 quantifies yield variations from baseline to endpoint. CS exhibits the clearest upgrading response: with rising DFA loading, tar falls, syngas-to-tar selectivity rises. This indicates DFA shifts part of cotton-stalk conversion pathway away from condensable tar toward gaseous products.

CMCpd behaves distinctly differently. The tar fraction remains consistently high across the entire studied range, and the gas fraction does not increase in parallel with H_2. This clearly confirms that a hydrogen-rich gas response does not necessarily imply global tar suppression. This behavior is consistent with the distinct thermal response of cellulose-type polysaccharides [10]. LMC remains char-dominated throughout the full range, which is consistent with the aromatic and cross-linked intrinsic nature of lignin and its tendency toward solid-residue formation [11].

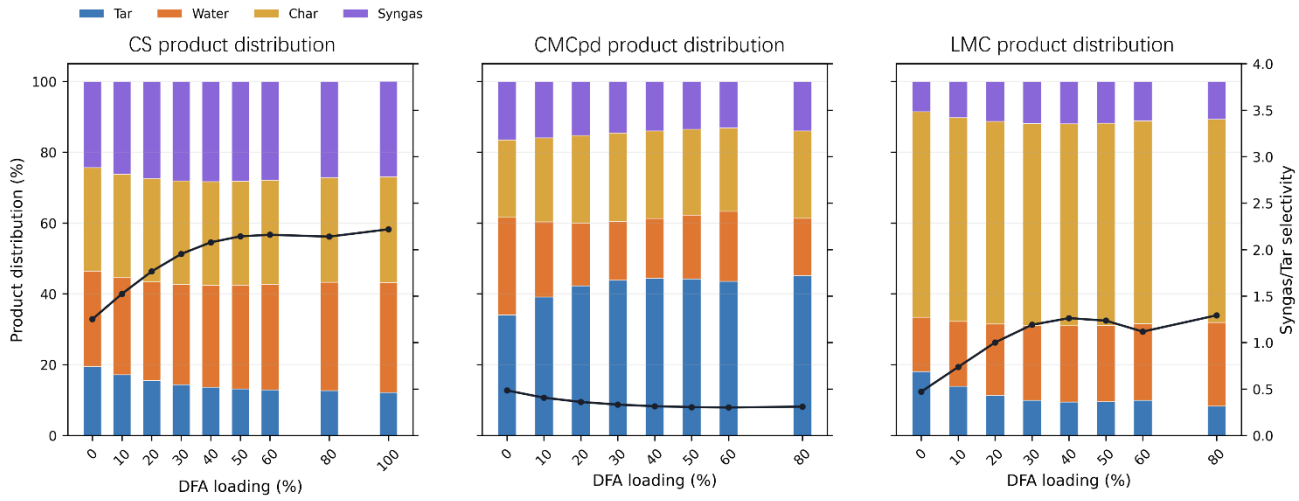

Figure 2. Product redistribution and syngas-to-tar selectivity reconstructed from the source report.

Table 2. Baseline-to-final product-yield summary used for Figures 2 and 4.

Feedstock	DFA range (%)	Tar (%)	Water (%)	Char (%)	Syngas (%)
CS	0-100	19.44 → 12.14	27.04 → 31.08	29.20 → 29.88	24.33 → 26.93
CMCpd	0-80	34.06 → 45.17	27.54 → 16.20	21.90 → 24.65	16.50 → 13.98
LMC	0-80	18.06 → 8.28	15.34 → 23.64	58.10 → 57.38	8.50 → 10.70

Gas-phase fingerprints reveal distinct catalytic pathways

Gas composition provides stronger mechanistic discrimination than total product yield alone, with full gas yield variations quantified in Table 3. In Figure 3, during CS catalytic pyrolysis, H_2 increases substantially across the entire loading range, while CO_2 decreases and CH_4 changes moderately. This pattern aligns with gas upgrading through cracking, deoxygenation and reforming, consistent with ash-catalyzed pyrolysis observations [12].

CMCpd displays the strongest H_2-selective response, whereas CO_2, CO and CH_4 decline. LMC shows the opposite behavior: H_2 remains very low, while CO_2 and CH_4 remain high. These contrasting signatures explain why whole cotton stalk behaves as an intermediate system rather than as a simple linear combination of its model compounds [13].

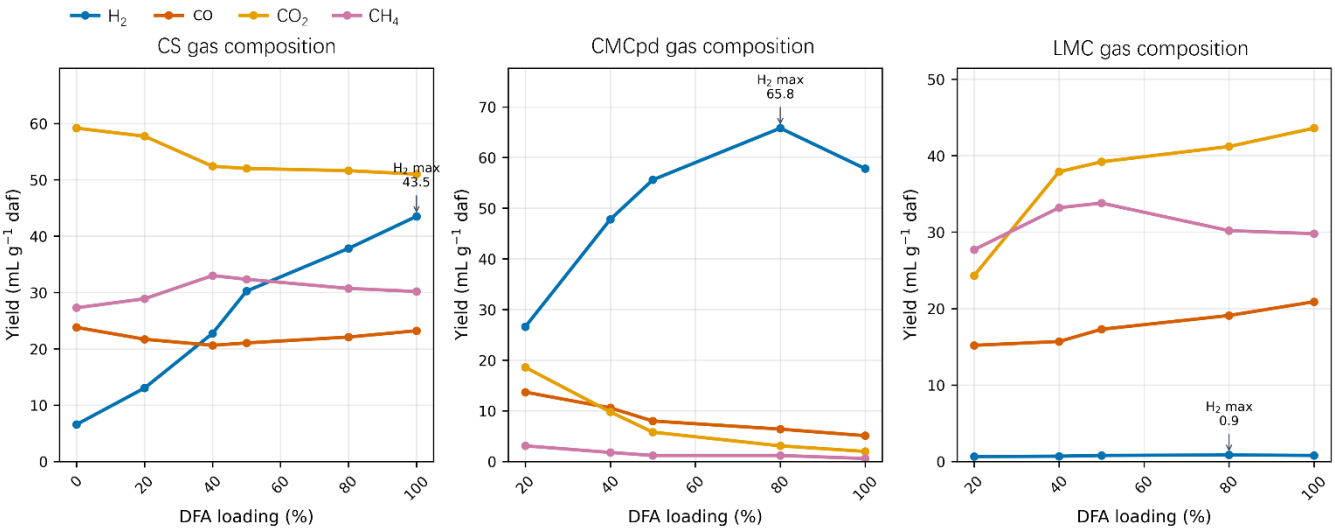

Figure 3. Gas-yield response fingerprints for CS, CMCpd and LMC.

Table 3. Baseline-to-final gas-yield summary used for Figures 3 and 4 (Units: mL g^{-1} daf).

Feedstock	DFA range (%)	H_2	CO	CO_2	CH_4
CS	0-100	6.58 → 43.52	23.81 → 23.22	59.17 → 51.01	27.30 → 30.19
CMCpd	20-100	26.60 → 57.80	13.70 → 5.10	18.60 → 2.00	3.10 → 0.60
LMC	20-100	0.65 → 0.80	15.20 → 20.90	24.30 → 43.60	27.70 → 29.80

Effect-size mapping identifies robust response directions

Figure 4 summarizes final-to-baseline responses. CS shows the most favorable combination: lower tar, higher syngas, strong H_2 enrichment and lower CO_2. CMCpd shows strong H_2 enrichment but no simultaneous tar suppression. LMC exhibits high char retention and minimal hydrogen upgrading.

The heatmap clarifies a key conceptual point: DFA is not a universal yield enhancer. Instead, it redirects feedstock structures toward distinct dominant pathways. Analysis must emphasize feedstock-specific selectivity over generic correlation.

Mechanistic response trajectories reveal the practical upgrading window

Figure 5 provides an integrated upgrading-response view by jointly comparing hydrogen generation, tar suppression, the upgrading index of H_2 to tar and char-retention tendency. This four-panel structure avoids compressing all variables into a single overloaded map and makes the feedstock-specific response to DFA loading easier to interpret.

The four-panel response map shows that CS develops a balanced upgrading pathway: H_2 yield increases, tar yield decreases, and the upgrading index of H_2 to tar rises continuously with DFA loading. CMCpd produces a higher absolute H_2 yield but maintains a high tar fraction, indicating that hydrogen enrichment does not automatically imply tar suppression. LMC remains in a low-H_2, high-char regime, confirming that lignin-derived structures primarily favor solid-residue retention rather than hydrogen-rich gas formation.

Taken together, Figure 5 clearly identifies CS at medium-to-high DFA loading as the most practically useful catalytic upgrading window, because this range combines increasing hydrogen generation with decreasing tar yield and an improved upgrading index of H_2 to tar.

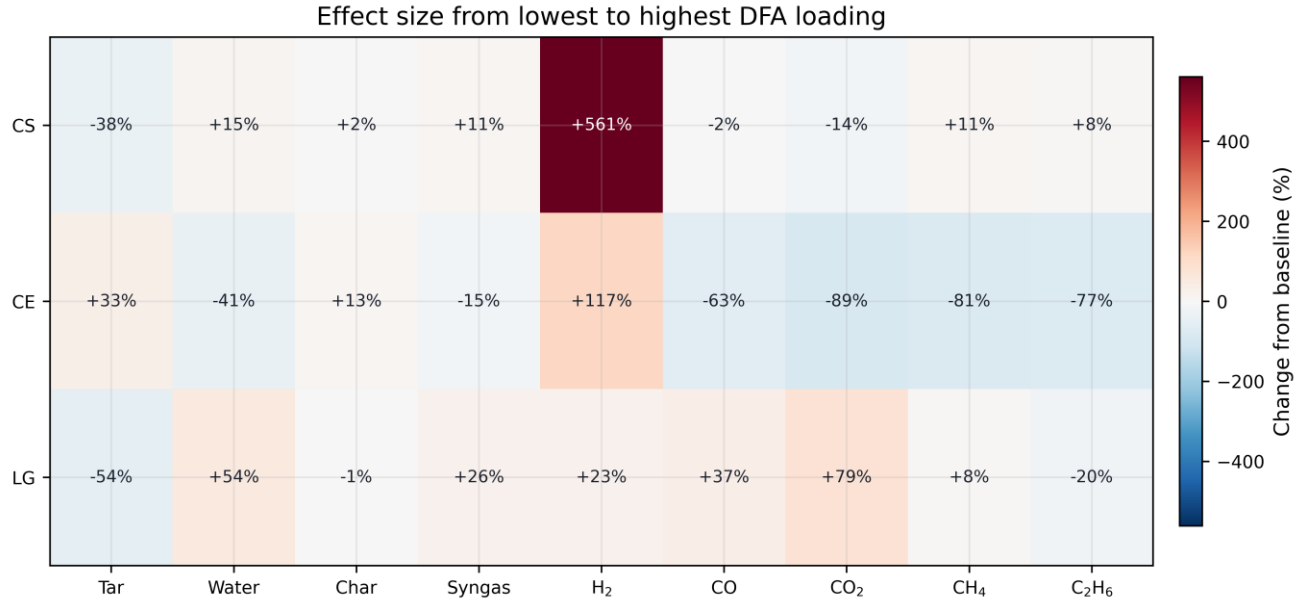

Figure 4. Final-to-baseline effect-size heatmap for products and selected gases.

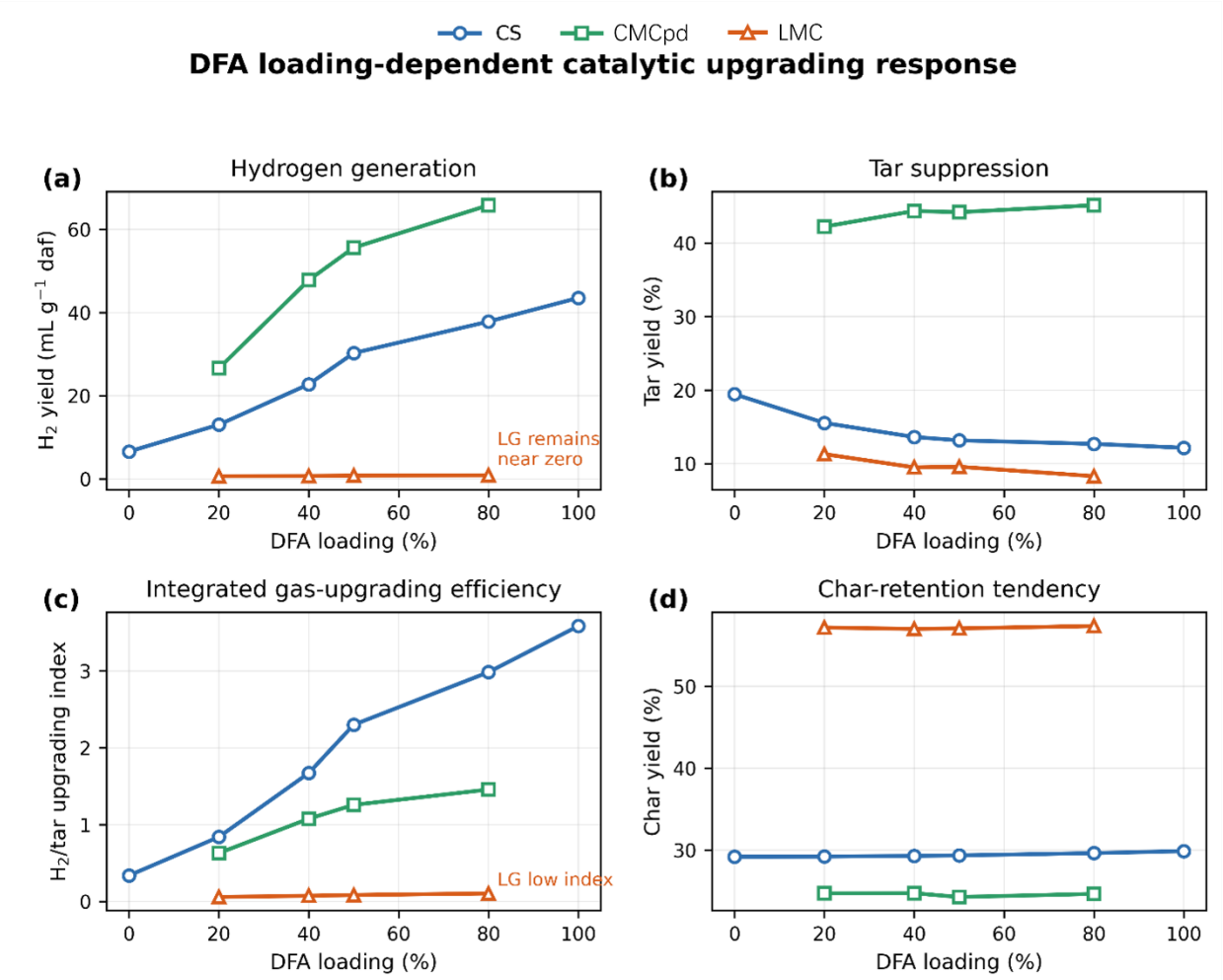

Figure 5. Tar-H_2 operating map integrating absolute tar yield, H_2 yield and char retention.

Note: Only matched loading points with both product-yield and gas-yield data are plotted: CS=0%, 20%, 40%, 50%, 80%, 100%; CMCpd and LMC=20%, 40%, 50%, 80%.

Conservative prediction is preferable to overfitted AI claims

The original report used high-order polynomial and artificial-intelligence language, but the available data does not justify unrestricted predictive claims. Figure 6 therefore presents a conservative LOOCV analysis based on low-order regression. Product variables generally show smoother interpolation behavior than gas variables, while variables with sharper peaks or digitized endpoints exhibit larger normalized error. This cautious stance is consistent with the wider biomass-pyrolysis modelling literature, where predictive AI models are usually strengthened by substantially larger datasets [14].

The appropriate modelling conclusion is not that the dataset can support accurate prediction over untested conditions. Rather, it can support mechanism-guided interpolation within the observed range and identify which variables urgently require additional experiments and raw-data recovery [15].

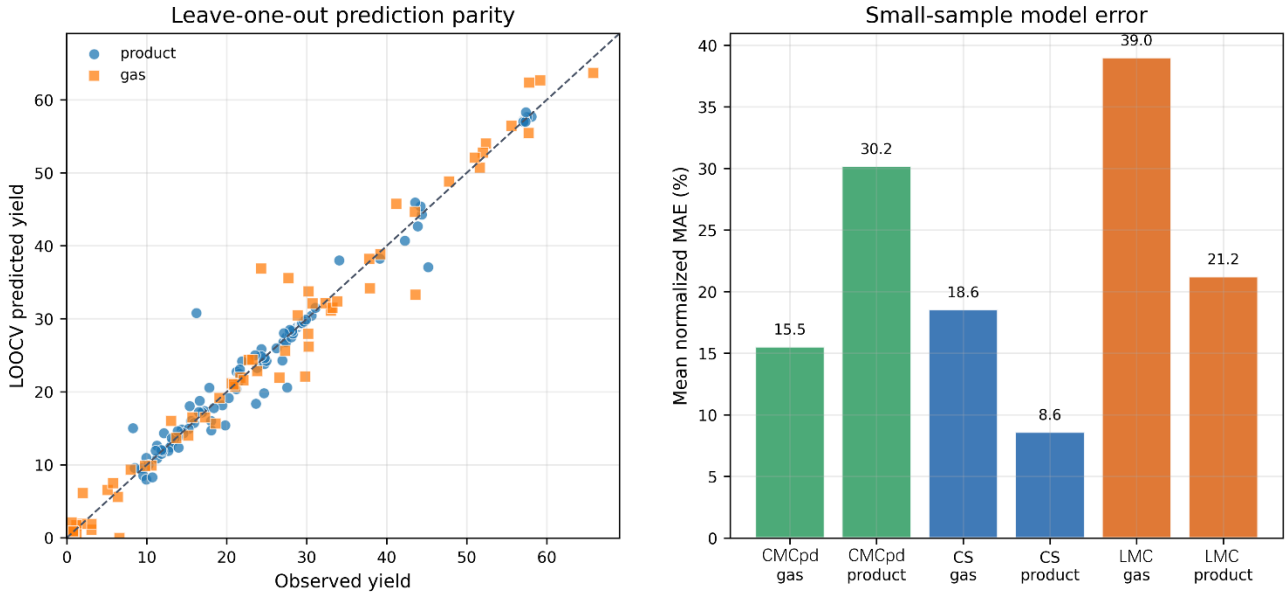

Figure 6. Leave-one-out cross-validation diagnostics for conservative small-sample modelling.

Conclusion

Increasing DFA loading shifts cotton-stalk pyrolysis from tar-rich toward gas-favored behavior and strongly enriches H_2, indicating a practically meaningful catalytic upgrading effect.

CMCpd exhibits a hydrogen-selective pathway, but Figure 5 shows that this does not coincide with tar suppression. LMC remains char-dominated and CO_2- and CH_4-rich, indicating a structurally constrained upgrading route.

Figure 5 clarifies that the most useful practical window occurs in CS at medium-to-high DFA loading, where lower tar yield, higher H_2 yield and an improved the upgrading index of H_2 to tar are achieved simultaneously.

Funding

This work was not supported by any funds.

Data availability

Reconstructed datasets and redrawn figures are documented in the supplementary tables and accompanying data package. The values were reconstructed from the source report using displayed equations, printed tables and embedded figure panels.

Acknowledgements

The author would like to show sincere thanks to those techniques who have contributed to this research.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] Bridgwater, A. V. (2012) Review of fast pyrolysis of biomass and product upgrading. *Biomass and Bioenergy*, 38, 68-94.

[2] Fawzy, S., Osman, A. I., Farrell, C., Al-Muhtaseb, A. A. H., Harrison, J., Al-Fatesh, A. S., Fakeeha, A. H., Doran, J., Yang, H., Rooney, D. W. (2021) Characterization and kinetic modeling for pyrolytic conversion of cotton stalks. *Energy Science & Engineering*, 9(10), 1908-1918.

[3] Várhegyi, G., Antal Jr, M. J., Jakab, E., Szabó, P. (1997) Kinetic modeling of biomass pyrolysis. *Journal of Analytical and Applied Pyrolysis*, 42(1), 73-87.

[4] Makepa, D. C., Chihobo, C. H., Musademba, D. (2023) Advances in sustainable biofuel production from fast pyrolysis of lignocellulosic biomass. *Biofuels*, 14(5), 529-550.

[5] Song, J., Tang, C., An, X., Wang, Y., Zhou, S., Huang, C. (2022) Catalytic pyrolysis of sawdust with desulfurized fly ash for pyrolysis gas upgrading. *International Journal of Environmental Research and Public Health*, 19(23), 15755.

[6] Yang, X., Chen, N., Ge, S., Sheng, Y., Yang, K., Lin, P., Guo, X., Lam, S. S., Ming, H., Zhang, L. (2022) Components interaction of cotton stalk under low-temperature hydrothermal conversion: a bio-oil pyrolysis behavior perspective analysis. *Polymers*, 14(20), 4307.

[7] Akinpelu, D. A., Adekoya, O. A., Oladoye, P. O., Ogbaga, C. C., Okolie, J. A. (2023) Machine learning applications in biomass pyrolysis: from biorefinery to end-of-life product management. *Digital Chemical Engineering*, 8, 100103.

[8] Dong, Z., Bai, X., Xu, D., Li, W. (2023) Machine learning prediction of pyrolytic products of lignocellulosic biomass based on physicochemical characteristics and pyrolysis conditions. *Bioresource Technology*, 367, 128182.

[9] Kale, R. D., Lenka, M., Rao, C. S. (2025) Leveraging explainable AI framework for predictive modeling of products of microwave pyrolysis of lignocellulosic biomass using machine learning. *Journal of Analytical and Applied Pyrolysis*, 192, 107249.

[10] Shen, D. K., Gu, S., Bridgwater, A. V. (2010) The thermal performance of the polysaccharides extracted from hardwood: cellulose and hemicellulose. *Carbohydrate Polymers*, 82(1), 39-45.

[11] Lu, Y., Wei, X., Zong, Z., Lu, Y., Zhao, W., Cao, J. (2013) Structural investigation and application of lignins. *Progress in Chemistry*, 25(05), 838.

[12] An, X., Tang, C., Liu, J., Song, X., Huang, C. (2025) Catalytic pyrolysis of biomass using desulfurization ash: enhanced bio-oil yields and reaction pathway modulation. *Energy*, 330, 136817.

[13] Jia, W., Zhang, H., Liu, B., Xiao, J., Yang, H., Chen, H. (2025) Pyrolytic interactions in biomass: a review across molecular, component, and feedstock scales. *Energy & Fuels*, 39(47), 22463-22488.

[14] Tsekos, C., Tandurella, S., De Jong, W. (2021) Estimation of lignocellulosic biomass pyrolysis product yields using artificial neural networks. *Journal of Analytical and Applied Pyrolysis*, 157, 105180.

[15] Lee, H., Choi, I. H., Hwang, K. R. (2024) Machine learning prediction of bio-oil production from the pyrolysis of lignocellulosic biomass: recent advances and future perspectives. *Journal of Analytical and Applied Pyrolysis*, 179, 106486.