

Modeling the Association Between Operating Conditions and Catalytic Performance in CO₂-to-Methanol Conversion

Emily Cheong
cheong.chaeyoon@gmail.com

ABSTRACT

Converting a greenhouse gas into a useful chemical and energy carrier has been a pressing societal concern, and carbon dioxide hydrogenation to methanol is a promising pathway for it. However, catalytic performance substantially varies across experimental studies. This study focused on examining how four commonly reported operating conditions, namely; temperature, pressure, hydrogen-to-carbon dioxide ratio, and gas hourly space velocity (GHSV), were associated with methanol yield and CO₂ conversion. Data were obtained from an open literature-derived dataset of thermocatalytic CO₂-to-methanol experiments named TheMeCat. For methanol yield ($n = 712$) and CO₂ conversion ($n = 696$), separate ordinary least squares regression models were estimated. Due to the nature of strongly right-skewed distribution of GHSV and multiplicative differences in gas flow, natural logarithm was used for GHSV. Heteroskedasticity-robust standard errors were reported. The methanol-yield model explained 29.7% of observed variation, while the CO₂-conversion model explained 31.8% of observed variation. A positive association was reported between both temperature and H₂/CO₂ ratio with both outcomes. However, a negative association was reported between $\ln(\text{GHSV})$ and both outcomes. Pressure behaved differently in both models. It was positively associated with methanol yield. However, it was not statistically significant in the CO₂-conversion model. With these findings, the hypothesis was partially supported, suggesting that a small set of operating variables may explain a meaningful yet incomplete portion of performance differences across published experiments. The analysis also demonstrates how a simple and reproducible model may provide interpretable trends from a curated catalysis dataset without claiming causal proof.

Keywords: CO₂ hydrogenation; methanol synthesis; catalytic performance; gas hourly space velocity; multiple linear regressions

INTRODUCTION

Converting carbon dioxide into useful chemicals has become one of the important research topics due to its link of carbon utilization with the production of fuels and industrial feedstocks. In this context, methanol is especially valuable due to its nature of being used directly as a chemical raw material,

Sep 2026
Vol 11, No 2.

converted into other products, and considered as a liquid energy carrier within a broader “methanol economy (Olah et al., 2011). In addition, A practical complement to emerging photochemical and electrochemical pathways is CO₂-to-methanol synthesis as thermochemical conversion is also compatible with existing industrial knowledge and infrastructure (Hu et al., 2013; Tountas et al., 2019). In this perspective, industrial-scale studies have examined how CO₂ and low-carbon hydrogen were captured and integrated into methanol and dimethyl ether production (Pontzen et al., 2011).

The central reaction from CO₂ + 3H₂ to CH₃OH + H₂O is chemically and operationally challenging because CO₂ is stable, but methanol formation competes with other pathways, such as the reverse water-gas shift reaction. Therefore, high methanol yield or selectivity is not automatically produced by high conversion. For example, performance is influenced by catalyst identity, active-site structure, support materials, and reaction environment. Copper-zinc oxide systems have long been used as a benchmark catalyst. However, promising behavior has also been shown by other materials based on palladium-zinc and indium oxide (Arena et al., 2007; Bahruji et al., 2016; Chou and Lobo, 2019). According to the research conducted on zinc-containing model catalysts and indium-oxide systems, the chemical state of active components and oxygen-vacancy-related mechanisms have been shown to substantially alter methanol formation (Amann et al., 2022; Wang et al., 2021).

Operating conditions are also important due to their nature of affecting reaction rate, equilibrium, contact time, and product distribution. According to high-pressure experiments, pressure and promoter selection has been demonstrated to influence the conversion of CO₂ and the formation of methanol over copper-based catalysts (Bansode et al., 2013). Increasing temperature may accelerate reaction kinetics. However, methanol selectivity may also be reduced by higher temperatures. In addition, higher temperatures may also make methanol formation increasingly constrained by thermodynamic equilibrium. Therefore, a positive temperature coefficient in the analysis should not be interpreted as a universal chemical rule but as an empirical association in the range of conditions specified in TheMeCat. The feed ratio of H₂/CO₂ determines reactant availability. At the same time, gas hourly space velocity controls the amount of gas passing over a given catalyst mass. Therefore, it is related to resident time. According to experimental work conducted on copper, palladium, indium-zirconium, and supported catalyst systems, performance has been illustrated to depend on interactions among materials and operating conditions over a single isolated factor (Chen et al., 2019; Jiang et al., 2020; Samson et al., 2014). Furthermore, the need to evaluate both catalyst design and process variable together has been emphasized similarly by the reviews of copper-based and indium-oxide catalysts (Niu et al., 2022; Wang et al., 2021).

Prior studies have provided detailed chemical insights without direct comparisons across studies. Different catalyst compositions, temperature and pressure ranges, feed ratios, flow definitions, and reporting conventions have been used in individual experiments. Many studies examined the best-performing catalyst or a narrow mechanistic question. Therefore, it was difficult to identify broad, interpretable trends across the larger evidence base. To address this data-fragmentation problem, TheMeCat was developed by standardizing catalyst information, process parameters, and performance measure from 44 thermocatalytic CO₂-to-methanol studies, and 822 experimental records were produced

as a result (Toldy et al., 2026). The structure is based on the investigation to use real experimental data through an application of a transparent statistical model over a complex microkinetic simulation.

Therefore, the present study focuses on evaluating whether a limited set of commonly reported operating conditions may explain variation in two measures of catalytic performance: CO₂ conversion and methanol yield. The contribution is a reproducible cross-study analysis that quantifies associations within a curated dataset, not a new catalyst or causal mechanism. This study addressed the following research question: To what extent are temperature, pressure, the H₂/CO₂ ratio, and GHSV associated with CO₂ conversion and methanol yield in literature-reported thermocatalytic experiments? It was hypothesized that higher temperature, pressure, and H₂/CO₂ ratio would be associated with higher CO₂ conversion and methanol yield, whereas higher GHSV would be associated with lower conversion and yield as the extent of reaction may be limited by shorter residence time.

METHODS AND MATERIALS

In this study, TheMeCat version 1.0, an open dataset created from published experiments on the thermocatalytic conversion of CO₂ to methanol, was used. The dataset was downloaded from Zenodo that was described by Toldy et al. (2026). In the source database, a total of 822 experimental records were contained as drawn from 44 publications. The database was limited by the original curators to single-pass, packed-bed thermocatalytic experiments that CO₂ and H₂ were the only reactants. Therefore, catalyst composition could be quantified. In the database, catalyst composition and physical properties, process conditions, and performance measures were included. As this study focused on a simple and interpretable operating-condition model, the regression equations did not include catalyst-composition variables.

Two dependent variables, CO₂ conversion and methanol yield, were defined in the source dataset and analyzed separately. First, CO₂ conversion was defined as the percentage of inlet CO₂ consumed during the experiment. Methanol yield was defined as the percentage of inlet CO₂ converted into methanol. This could be calculated through conversion multiplied by methanol selectivity. In addition, reaction temperature in Kelvin, pressure in bar, the H₂/CO₂ feed ratio, and gas hourly space velocity were four independent variables. In TheMeCat, GHSV is standardized as the normal-liter flow rate of reactants per hour divided by catalyst mass in grams. Therefore, more gas passes over each gram of catalyst per unit time and shorter effective contact or resident time correspond to a larger GHSV.

In the analysis, a complete-case approach was used for each outcome. A record was included only if all four predictors and the outcome were available. There were 712 observations in the methanol-yield model, while there were 696 observations in the CO₂-conversion model. As some records reported one performance measures, separate samples were retained. Missing values were not replaced through imputation but were treated as unavailable information. No outcome values were generated or simulated, or no observations were removed only because they produced high or low catalytic performance.

There was a much wider numerical range in GHSV when compared with the other predictors. Therefore, it represented multiplicative differences in flow conditions. This was the reason why the natural logarithm

of GHSV was used in both models, but only positive reported GHSV values were log-transformed. The remaining variables still kept the original units to directly interpret the unstandardized coefficients. The two ordinary least squares models were generated as follows:

$$\textbf{Methanol Yield: } \beta_0 + \beta_1(\text{Temperature}_i) + \beta_2(\text{Pressure}_i) + \beta_3(\text{H}_2/\text{CO}_2 \text{ Ratio}_i) + \beta_4 \ln(\text{GHSV}_i) + \varepsilon_i$$

$$\textbf{CO}_2 \text{ Conversion: } \beta_0 + \beta_1(\text{Temperature}_i) + \beta_2(\text{Pressure}_i) + \beta_3(\text{H}_2/\text{CO}_2 \text{ Ratio}_i) + \beta_4 \ln(\text{GHSV}_i) + \varepsilon_i$$

In both equations, β_0 is the intercept, β_1 through β_4 are slope coefficients showing the adjusted association between each predictor and the outcome, and ε_i is the residual error for observation i . For the log-transformed GHSV predictor, a multiplicative change in GHSV by a factor of k corresponds to an estimated change of $\beta_4 \ln(k)$ percentage points in the outcome. Therefore, the estimated effect of doubling GHSV is calculated as $\beta_4 \ln(2)$. Percentage-point units were used for coefficients for yield and conversion. For example, a coefficient of 0.025 for temperature in the yield model shows an estimated 0.025-percentage-point difference in methanol yield for a 1 K difference in temperature, given that the remaining variables were held constant.

Python 3 was used for the data analysis. Pandas was used for data preparation, and statsmodels was used for regression estimation. Ordinary least squares regression was chosen since both outcomes were continuous percentage measures. At the same time, the main objective was to estimate transparent and directly interpretable adjusted associations instead of maximizing predictive accuracy using a more complex model. With the application of the same predictor set in both models, it was possible to directly compare the associations with methanol yield and CO₂ conversion. Since the pooled dataset contained experiments from multiple publications and catalyst systems, constant residual variance could not be assumed. Therefore, heteroskedasticity-robust standard errors were used. Before the models were finalized, predictor distributions and bivariate scatterplots were inspected for skewness, outliers, and possible nonlinear patterns. Residual-versus-fitted and Q-Q plots were reviewed, while examining standardized residuals, leverage values, and Cook's distance to identify potential influential observations. There were no problems revealed by these diagnostics to justify excluding otherwise eligible observations. Therefore, all complete-case records were retained. $\alpha = 0.05$ was used for the evaluation of statistical significance, and 95% confidence intervals were calculated from the robust standard errors. Using R^2 , model fit was summarized. Since the analysis was observational and cross-sectional across experimental records, coefficients were not interpreted as causal effects but associations. Furthermore, no interaction terms, polynomial terms, automated variable selection, or machine-learning algorithms were used to keep the model simple. The same predictor set was applied to both outcomes. Therefore, the two models may be directly compared.

RESULTS

To examine the relationships between operating conditions and catalytic performance, two multiple linear regression models were estimated. In the first model, methanol yield was used as the dependent variable where a total of 712 complete observations were included. The model explained 29.7% of the observed variation in methanol yield ($R^2 = 0.297$). As shown in this level of fit, a meaningful portion of cross-study differences was captured by the four operating variables. However, around 70.3% of the variation remained associated with catalyst composition, experimental design, measurement differences, nonlinear behavior, interactions, and other factors not included in the model.

The fitted methanol-yield equation was as follows:

$$\text{Predicted Methanol Yield} = -9.569 + 0.025(\text{Temperature}) + 0.078(\text{Pressure}) + 1.808(\text{H}_2/\text{CO}_2 \text{ Ratio}) - 3.453\ln(\text{GHSV}).$$

There was a positive and statistically significant association between temperature and methanol yield. When holding pressure, feed ratio, and $\ln(\text{GHSV})$ constant, a 1 K increase in temperature was associated with a 0.025-percentage-point increase in yield (robust SE = 0.004, 95% CI [0.017, 0.033], $p < 0.001$). Pressure was also positively associated with yield. Each additional bar of pressure was associated with a 0.078-percentage-point increase in predicted yield (robust SE = 0.013, 95% CI [0.053, 0.103], $p < 0.001$). A larger positive coefficient was reported with the H₂/CO₂ ratio. Specifically, a one-unit increase in the ratio was associated with a 1.808-percentage-point increase in yield (robust SE = 0.271, 95% CI [1.277, 2.339], $p < 0.001$). In contrast, $\ln(\text{GHSV})$ was negatively associated with methanol yield. A one-unit increase in $\ln(\text{GHSV})$, equivalent to multiplying GHSV by e , or approximately 2.72, was associated with a 3.453-percentage-point decrease in predicted methanol yield (robust SE = 0.368, 95% CI [-4.174, -2.732], $p < 0.001$). On the original GHSV scale, doubling GHSV was associated with an estimated 2.39-percentage-point decrease in methanol yield (95% CI [-2.89, -1.89]), when holding the other predictors constant (Table 1; Figure 1).

Table 1. Multiple Linear Regression Predicting Methanol Yield

Predictor	B	Robust SE	95% CI	p
Intercept	-9.569	2.104	[-13.693, -5.445]	< .001
Temperature (K)	0.025	0.004	[0.017, 0.033]	< .001
Pressure (bar)	0.078	0.013	[0.053, 0.103]	< .001
H ₂ /CO ₂ ratio	1.808	0.271	[1.277, 2.339]	< .001
$\ln(\text{GHSV})$	-3.453	0.368	[-4.174, -2.732]	< .001

Note: B represents the unstandardized regression coefficients. SE = robust standard error; CI = confidence interval; $n = 712$; $R^2 = 0.297$

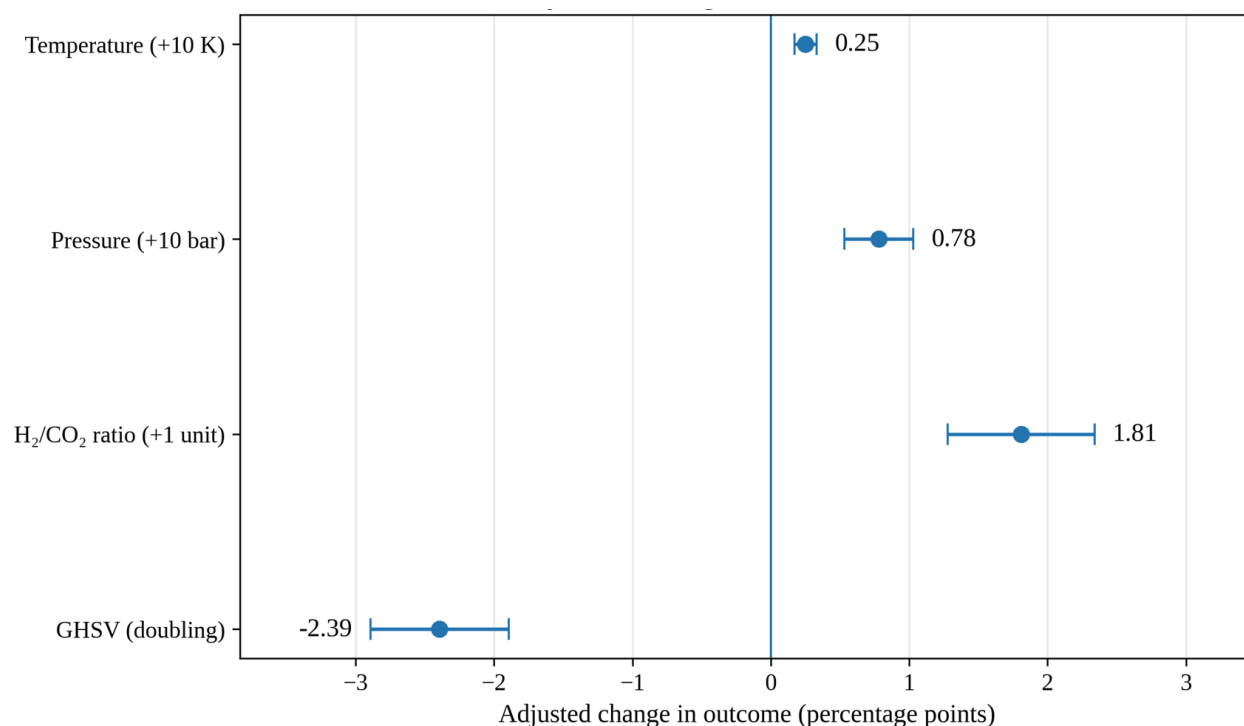


Figure 1. Adjusted changes in methanol yield associated with meaningful changes in the four operating conditions. Points show regression estimates and horizontal bars show robust 95% confidence intervals; the GHSV estimate represents the predicted effect of doubling GHSV.

In the second model, CO₂ conversion was used as the dependent variable where a total of 696 complete observations were included. This model explained 31.8% of the observed variation in conversion ($R^2 = 0.318$). This was slightly higher than the methanol-yield model. The fitted equation was as follows.

$$\text{Predicted CO}_2 \text{ Conversion} = -20.896 + 0.061(\text{Temperature}) + 0.037(\text{Pressure}) + 2.353(\text{H}_2/\text{CO}_2 \text{ Ratio}) - 5.647\ln(\text{GHSV})$$

A positive association was reported between temperature and conversion. A 1 K increase was associated with a 0.061-percentage-point increase in predicted CO₂ conversion when holding the other variables constant (robust SE = 0.006, 95% CI [0.049, 0.073], $p < 0.001$). A positive and statistically significant association was reported between the H₂/CO₂ conversion and conversion. A one-unit increase in the ratio was associated with a 2.353-percentage-point increase in conversion (robust SE = 0.335, 95% CI [1.696, 3.010], $p < 0.001$). $\ln(\text{GHSV})$ was also negatively associated with CO₂ conversion. A one-unit increase in $\ln(\text{GHSV})$, equivalent to multiplying GHSV by approximately 2.72, was associated with a 5.647-percentage-point decrease in predicted conversion (robust SE = 0.577, 95% CI [-6.778, -4.516], $p < 0.001$). On the original GHSV scale, doubling GHSV was associated with an estimated

3.91-percentage-point decrease in CO₂ conversion (95% CI [-4.70, -3.13]), when holding the other predictors constant (Table 2; Figure 2).

In the conversion model, pressure behaved differently. Its coefficient was positive but small that each additional bar of pressure was associated with a 0.037-percentage-point increase in predicted conversion. At the 0.05 significance level, the estimated was not statistically significant (robust SE = 0.022, 95% CI [-0.006, 0.080], $p = 0.082$). As zero was included in the confidence interval, no independent pressure-conversion association was established after controlling for the other variables (Table 2) (Figure 2).

Table 2. Multiple Linear Regression Predicting CO₂ Conversion

Predictor	B	Robust SE	95% CI	p
Intercept	-20.896	2.945	[-26.668, -15.124]	< .001
Temperature (K)	0.061	0.006	[0.049, 0.073]	< .001
Pressure (bar)	0.037	0.022	[-0.006, 0.080]	.082
H ₂ /CO ₂ ratio	2.353	0.335	[1.696, 3.010]	< .001
ln(GHSV)	-5.647	0.577	[-6.778, -4.516]	< .001

Note: B represents the unstandardized regression coefficient. SE = robust standard error; CI = confidence interval; $n = 696$; $R^2 = 0.318$

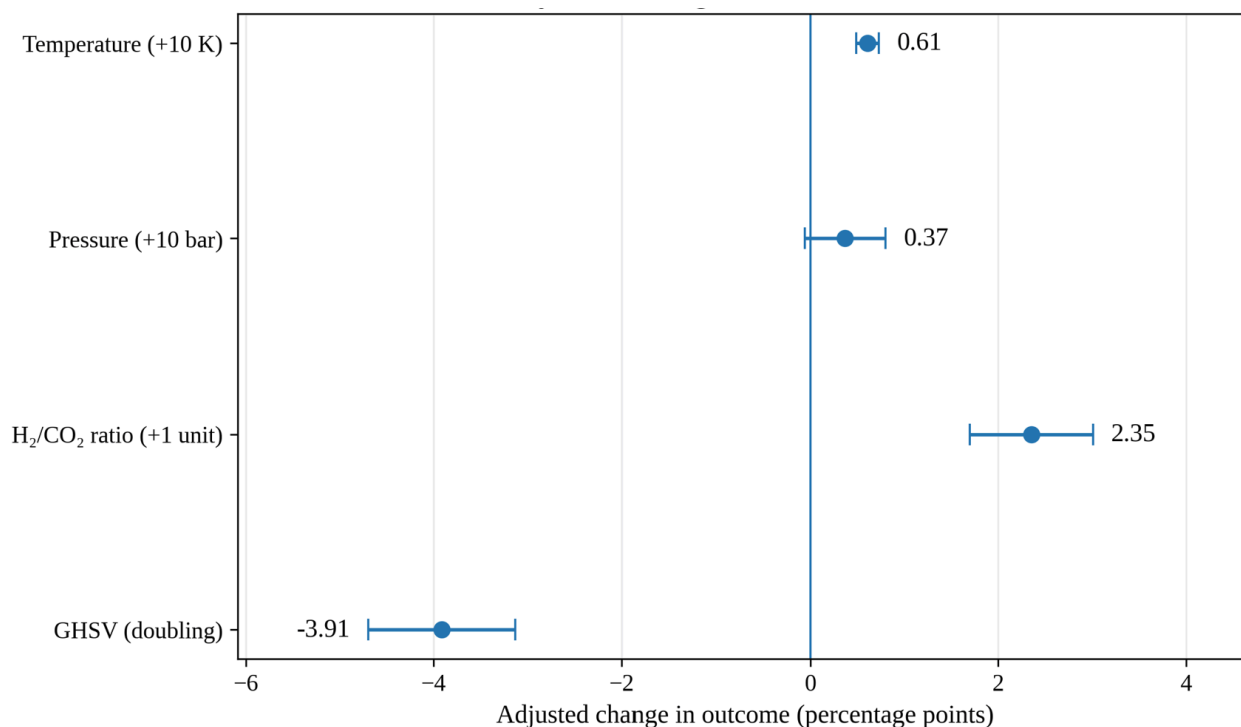


Figure 2. Adjusted changes in CO₂ conversion associated with meaningful changes in the four operating conditions. Points show regression estimates, and horizontal bars show robust 95% confidence intervals; the pressure interval crosses zero, while the GHSV estimate shows the effect of doubling GHSV.

Across the two models, a consistently positive association was reported between both temperature and H₂/CO₂ ratio and the outcomes. However, ln(GHSV) was consistently negatively associated with the outcomes. The predicted direction for methanol yield was supported by the pressure, but not conclusively for CO₂ conversion. Therefore, the hypothesis established in this study was partially supported rather than fully confirmed. Broad associations across published experiments were identified by the results. However, the moderate R^2 values and observational design do not support causal claims or precise prediction for an individual catalyst. In both models, the intercepts were statistically different from zero. However, they were not chemically interpreted due to a condition with zero temperature, pressure, feed ratio, and ln(GHSV) lying outside the meaningful experimental range. The directions, sizes, uncertainty intervals, and consistency of the predictor coefficients were more useful findings in the data analysis. In particular, the repeated negative GHSV association was observed for both performance measures. However, the role of pressure was different between yield and conversion models.

DISCUSSION

The findings in this study show that consistent patterns in literature-reported CO₂-to-methanol performance have been captured by simple operating condition variables. In particular, there was an association reported between higher temperature and H₂/CO₂ ratio and higher conversion and yield. On the other hand, there was a negative association between ln(GHSV) and the values of both outcomes. Due to faster gas throughput generally reducing effective contact time between reactants and the catalyst bed, it is chemically plausible to have negative GHSV relationship. The positive feed-ratio result may indicate greater hydrogen availability for methanol-forming pathways. These interpretations support prior experimental literature indicating that performance is jointly determined by operating conditions and catalyst structure rather than independently acting (Arena et al., 2007; Bansode et al., 2013; Niu et al., 2022).

More cautious interpretation is required for pressure as it behaved differently in two models. Pressure was always positively associated with methanol yield. However, it did not provide statistical significance for CO₂ conversion after controlling for other predictors. This difference may be from correlations among process conditions, restricted pressure ranges within particular studies, or catalyst-dependent responses that cannot be presented by a single pooled linear coefficient. Therefore, the hypothesis established in this study was partially supported.

Although this study provides an important insight between temperature, pressure, H₂/CO₂ ratio, and GHSV, collectively, and CO₂ conversion and methanol yield, there are several limitations that reduce the strength of inference. First, TheMeCat combines records from multiple publications, catalyst families, and measurement procedures. Therefore, the simple models tend to omit catalyst composition, surface area,

pore structure, and study-specific effects. Second, observations from the same publication may be similar with the ones from different publications. Third, nonlinear temperature effects, equilibrium limits, or interactions may not be fully captured by a linear regression. The models established in this study explained around 30% of outcome variation. This is useful for broad comparison. However, it is insufficient for precise catalyst-level prediction at the same time. It is recommended for future work to focus on adding catalyst descriptions, publication-level controls, and carefully selecting interaction terms while preserving interpretability.

CONCLUSION

In this study, a curated experimental dataset and two simple regression models were used to examine the association between operating conditions and catalytic performance in thermocatalytic CO₂-to-methanol conversion. A positive association was reported between temperature and H₂/CO₂ ratio and methanol yield and CO₂ conversion. However, a negative association was reported between ln(GHSV) and both outcomes. Pressure was reported to be a statistically significant predictor of methanol yield but was not independently significant for conversion. Therefore, the hypothesis in this study was partially supported. The models established in this study explained 29.7% and 31.8% of the variation in yield and conversion, respectively. This demonstrates that performance may be explained by common operating variables usefully but incompletely at the same time. The findings in this study should not be interpreted as causal as heterogeneous experiments were combined in the datasets, while omitting important catalyst characteristics. However, the analysis still shows that statistical modeling may produce meaningful insight from open catalysis data. It is recommended for future study to improve explanatory power by incorporating catalyst composition, surface properties, and variation while preserving interpretability and reproducible framework.

REFERENCES

- Amann, P., et al. (2022). The state of zinc in methanol synthesis over a Zn/ZnO/Cu(211) model catalyst. *Science*, 376, 603–608. <https://doi.org/10.1126/science.abj7747>
- Arena, F., et al. (2007). Synthesis, characterization and activity pattern of Cu-ZnO/ZrO₂ catalysts in the hydrogenation of carbon dioxide to methanol. *Journal of Catalysis*, 249, 185–194. <https://doi.org/10.1016/j.jcat.2007.04.003>
- Bahruji, H., et al. (2016). Pd/ZnO catalysts for direct CO₂ hydrogenation to methanol. *Journal of Catalysis*, 343, 133–146.
- Bansode, A., Tidona, B., von Rohr, P. R., & Urakawa, A. (2013). Impact of K and Ba promoters on CO₂ hydrogenation over Cu/Al₂O₃ catalysts at high pressure. *Catalysis Science & Technology*, 3, 767–778. <https://doi.org/10.1039/C2CY20604H>

- Chen, T.-Y., et al. (2019). Unraveling highly tunable selectivity in CO₂ hydrogenation over bimetallic In-Zr oxide catalysts. *ACS Catalysis*, 9, 8785–8797. <https://doi.org/10.1021/acscatal.9b01869>
- Chou, C.-Y., & Lobo, R. F. (2019). Direct conversion of CO₂ into methanol over promoted indium oxide-based catalysts. *Applied Catalysis A: General*, 583, Article 117144.
- Hu, B., Guild, C., & Suib, S. L. (2013). Thermal, electrochemical, and photochemical conversion of CO₂ to fuels and value-added products. *Journal of CO₂ Utilization*, 1, 18–27.
- Jiang, X., et al. (2020). A combined experimental and DFT study of H₂O effect on In₂O₃/ZrO₂ catalyst for CO₂ hydrogenation to methanol. *Journal of Catalysis*, 383, 283–296.
- Niu, J., et al. (2022). Comprehensive review of Cu-based CO₂ hydrogenation to CH₃OH: Insights from experimental work and theoretical analysis. *International Journal of Hydrogen Energy*, 47, 9183–9200.
- Olah, G. A., Goeppert, A., & Prakash, G. K. S. (2011). *Beyond oil and gas: The methanol economy*. Wiley.
- Pontzen, F., Liebner, W., Gronemann, V., Rothaemel, M., & Ahlers, B. (2011). CO₂-based methanol and DME—Efficient technologies for industrial scale production. *Catalysis Today*, 171, 242–250.
- Samson, K., et al. (2014). Influence of ZrO₂ structure and copper electronic state on activity of Cu/ZrO₂ catalysts in methanol synthesis from CO₂. *ACS Catalysis*, 4, 3730–3741.
- Toldy, Á. I., Pisal, P., Krejčí, O., Rinke, P., & Santasalo-Aarnio, A. (2026). TheMeCat: Dataset of thermocatalytic conversion of CO₂ to methanol. *Scientific Data*, 13, Article 165. <https://doi.org/10.1038/s41597-025-06482-8>
- Tountas, A. A., et al. (2019). Towards solar methanol: Past, present, and future. *Advanced Science*, 6, Article 1801903.
- Wang, J., et al. (2021). CO₂ hydrogenation to methanol over In₂O₃-based catalysts: From mechanism to catalyst development. *ACS Catalysis*, 11, 1406–1423. <https://doi.org/10.1021/acscatal.0c03665>