

Evaluation of steady-state kinetic models for the Cu/ZnO/Al₂O₃ catalytic synthesis of methanol

By: *Karthik Rajeev Annigeri and Nicholas Diego Daponte*

Introduction:

Improving accuracy of methanol production models has gained increased interest in research due to the chemicals application in storing energy and to produce important precursors such as ethylene and propylene. Methanol is additionally capable of being produced by renewable carbon sources such as CO₂ and CO which allow a sustainable approach to produce this valuable feedstock. Kinetic models derived for commodity chemicals such as methanol should be able to accurately predict reaction rates under a wide range of reactor conditions and hence, play an important role in the design of industrial processes for large-scale production. Over recent years there has been a debate about which reaction pathway is correct for methanol production under the Cu/ZnO/Al₂O₃ catalyst leading to varying kinetic models described in literature based on experts' assumptions of the correct underlying mechanisms.

In this project, we will first attempt to replicate the literature paper's goal of fitting four previously established models as well as the paper's proposed model containing fewer parameters to the given supplementary experimental data set. The paper finally concludes that the model representing the steady state catalytic synthesis of methanol the best is the model that contains the fewest number of parameters while simultaneously being capable of predicting experimental data within one standard deviation, for which the papers new fewer-parameter model is selected as the best one after evaluation of the final metrics described in the literature's results.

The second part of this project aims to implement a machine learning solution that eliminates the need for an expert in kinetics to derive a new model by using the Automated Learning of Algebraic Models (ALAMO) machine learning software to predict a simpler and more accurate model that satisfies the statistical constraint previously discussed (the best model being the one with fewer parameters that can predict a new set of data within 1 standard deviation). ALAMO is a powerful Black-Box modeling software that allows the construction of accurate models from small data sets while ensuring that the models are as simple as possible. ALAMO's methodology is said to be superior to other competing software's such as LASSO. A 2015 academic study showed that ALAMO models are simpler while maintaining higher accuracy than LASSO output surrogate models. ALAMO has been successfully used to model industrial processes such as optimal syngas production as well as producing accurate models for ethylene and propylene plants. Thus, the overarching idea of this project is to determine if ALAMO's capabilities can replicate and output a simpler proposed model, like the novel proposed model outlined in the literature paper, by training the software to the older higher-parameter models along with the supplementary experimental data. ALAMO should theoretically be able to determine which parameters can be eliminated or even changed (e.g., rate dependence power terms) to arrive at a model that predicts the data within one standard deviation and has the least number of parameters.

Part I: Parameter Estimation by Regression (Least Squares Error Minimization):

The goal of the first part of the project is to determine the optimal values of the parameters (the Arrhenius-type coefficients of the lumped kinetic constants) that would best fit the data to different kinetic models. Here, we compare the results obtained for different kinetic models described in the paper and attempt to assess which among them would be the best fit. This is done by determining the minimum value of the mean sum of squares of errors (between the predicted and experimental weight fractions of components) and identifying the kinetic model which would give the lowest such value.

The kinetic models that will be assessed are listed below:

- 1) Graaf et al.
- 2) Vanden Bussche and Froment
- 3) Novel proposed model (6-parameter)

The new model proposed in this paper is supposed to be the best model for the experimental range considered and we will attempt to validate that by solving a series of optimization problems.

Data Set:

The experiments for this project were divided into two: The first set of experiments were done in a CSTR. Then, another set of experiments were conducted under carefully controlled conditions in a set-up consisting of a series of four parallel isothermal fixed bed reactors. Since the reactor is maintained under constant temperature (isothermal conditions), this means that we do not need to account for the energy balance equations which will greatly simplify our objective function formulation. Furthermore, the reactor diameter was kept small enough so that we can neglect any radial variations. As can be seen in the formulation, only axial variations in the gas concentrations will be considered. This assumption simplifies the governing equations and reduces the complexity of the ODE system whose solution is embedded in the optimization formulation.

The data set consists of 234 points, each of which describe the gas mixture composition in weight percentages at the inlet and outlet of the reactor setup. The experiments are conducted over a wide range of reaction conditions to ensure that the rate equations being tested are not biased by narrow sampling ranges. The temperature is varied from 483 K to 533 K, the pressure is varied from 20 to 50 bara and the gas flow rates are varied from 2000 to 10,000 GHSV. Additionally, the amount of catalyst is also varied slightly and is accounted for in the data set by differing values of the apparent catalyst density.

The dataset used in the optimization problem was eventually truncated to include only the 94 samples sourced from the PFR. The other points were sourced from CSTR experiments which would require a different mass balance equation to make use of. We believe this to be an oversight in the original paper.

Pre-processing of the data set prior to starting the optimization problem will not be required as there is not any significant noise expected in the data. Any deviations would be minor and attributed to the variance in measurements while conducting the experiments. A snapshot of the data set is provided below:

E. Experimental data

Table E.10: Experimental values. The density is given as the apparent density. The compositions are in mole percentage.

N	T (°C)	P (bar)	ϕ_V° (mL/min)	T_{in}^0 (°C)	CO_2 <i>in</i>	CO <i>in</i>	H_2 <i>in</i>	N_2 <i>in</i>	CO_2 <i>out</i>	CO <i>out</i>	H_2 <i>out</i>	MeOH <i>out</i>	H_2O <i>out</i>	N_2 <i>out</i>	γ	ρ_{app} (kg m ⁻³)
141	210	20	66.79	20	22.58	0.005	67.72	9.69	21.70	0.38	65.82	0.96	1.32	9.82	0.98	870
142	220	20	66.79	20	22.59	0.005	67.71	9.69	21.38	0.63	65.30	1.12	1.74	9.84	0.98	870
143	230	20	66.79	20	22.58	0.005	67.79	9.61	20.85	1.10	64.79	1.16	2.25	9.84	0.98	870
144	240	20	66.79	20	22.56	0.005	67.87	9.56	20.07	1.74	63.86	1.36	3.09	9.87	0.97	870
145	250	20	66.79	20	22.55	0.005	67.80	9.64	19.33	2.69	63.51	0.98	3.66	9.83	0.98	870
146	260	20	66.79	20	22.55	0.005	67.80	9.64	18.85	3.32	63.32	0.71	4.01	9.79	0.99	870
147	210	30	65.77	20	22.24	0.005	66.51	11.24	21.16	0.44	64.12	1.20	1.62	11.47	0.98	860
148	220	30	65.77	20	22.21	0.005	66.56	11.22	20.73	0.76	63.60	1.35	2.10	11.46	0.97	860
149	230	30	65.77	20	22.22	0.005	66.57	11.20	20.15	1.30	62.92	1.43	2.72	11.48	0.97	860
150	240	30	65.77	20	22.25	0.005	66.49	11.25	19.36	2.14	62.06	1.37	3.50	11.57	0.97	860
151	250	30	65.77	20	22.26	0.005	66.42	11.31	18.61	2.96	61.37	1.27	4.22	11.58	0.97	860
152	260	30	65.77	20	22.26	0.005	66.41	11.32	18.19	3.50	61.16	1.07	4.56	11.52	0.98	860
153	210	40	65.77	20	22.66	0.005	67.64	9.69	21.37	0.49	64.73	1.50	1.98	9.93	0.97	860
154	220	40	65.77	20	22.79	0.005	67.48	9.72	21.01	0.88	63.85	1.71	2.57	9.98	0.97	860
155	230	40	65.77	20	22.82	0.005	67.47	9.70	20.39	1.45	63.02	1.84	3.28	10.02	0.96	860
156	240	40	65.77	20	22.37	0.005	66.76	10.86	19.40	2.09	62.01	1.53	3.62	11.35	0.97	860
157	250	40	65.77	20	22.26	0.005	66.51	11.23	18.58	2.66	60.74	1.85	4.50	11.68	0.96	860
158	260	40	65.77	20	22.37	0.005	66.37	11.25	18.29	3.09	60.25	1.80	4.88	11.69	0.96	860
159	210	50	65.77	20	22.87	0.005	67.34	9.78	21.29	0.61	63.76	1.83	2.43	10.08	0.96	860
160	220	50	65.77	20	22.97	0.005	67.23	9.79	20.63	1.19	62.51	2.16	3.34	10.17	0.96	860

Figure. 1: Snapshot of experimental data set used in the referenced literature.

Final Optimization Formulation:

The basic framework of the optimization problem will remain the same for all models, however due to differing assumptions and complexity, the models will differ in the number of parameters to be estimated. For instance, the model proposed by Graaf et al. has 12 parameters (Arrhenius coefficients for various rate and adsorption equilibrium constants) whereas the novel model proposed in the paper has only 6 parameters and hence, is a considerably simpler optimization problem. For the formulation discussed here, we will consider only the 6-parameter model for sake of simplicity. However, the same method can be extended to any model in a similar manner.

The problem should be set up so that we find the optimal values of the parameters that, when plugged into the rate equations, should predict values that match the experimental values as closely as possible. The values of the objective functions can be compared between models to get a rough idea as to which model does a better job of predicting the rates.

Variables:

$$K_i = A_i \exp\left(\frac{B_i}{RT}\right) = \exp\left(\frac{\Delta S_{ads}^\circ}{R}\right) \exp\left(\frac{-\Delta H_{ads}^\circ}{RT}\right)$$

$$k_i = A_i \exp\left(\frac{B_i}{RT}\right) = A \exp\left(\frac{-E_A}{RT}\right)$$

Figure 2. Arrhenius equations containing the constants A_i and B_i fitted to the models in the literature.

The above equations in Figure 2 describe the Arrhenius-type dependence of the adsorption equilibrium constants and rate constants on temperature. We need to evaluate the optimal values of the coefficients in these equations.

For the 6-parameter model, some of the adsorption equilibrium and rate constants have been combined to give lumped kinetic constants by making certain simplifying assumptions. This reduces the number of parameters in the optimization problem. For this problem, we have the following parameters:

- 1) A_{CO_2}
- 2) B_{CO_2}
- 3) A_{RWGS}
- 4) B_{RWGS}
- 5) k_{H_2}
- 6) k_{H_2O}

Where the first four parameters are the Arrhenius coefficients of the respective rate constants, and the last two parameters are the temperature independent rate constants.

Constraints:

$k > 0$ and $K > 0$, imposed by logarithmic Arrhenius form

$E_A > 0$ or $b_i > 0$, imposed by bounds

$-\Delta H_{ads}^\circ > 0$ or $b_i < 0$, imposed by bounds

$0 < -\Delta S_{ads}^\circ < S_{gas}^\circ$, checked afterwards

Figure 3. Parameter constraints used for model fitting in the referenced literature.

The above constraints in Figure 3 are general and are utilized to improve numerical accuracy. The last two constraints would not be required for the 6-parameter model since the adsorption equilibrium constants have been combined into simpler lumped fitting parameters based on the derivation.

Objective function:

The governing mass balance function in Figure 4 below is used in the literature for model fitting and is highly non-convex while also being non-linear due to the nature of the rate equations which appear indirectly in the function. This equation gives us a system of coupled ODEs, the solution to which describes the weight fractions of each component in the reaction mixture as a function of axial position down the reactor bed. Solving these ODEs by integrating over the total length of the reactor bed gives us the expected values of the weight fractions at the exit of the reactor. These values can be subtracted from the corresponding experimental values of the gas composition at the reactor outlet to generate

the error terms. The mean error for each data point is hence calculated, squared, and summed up for all data points to generate the objective function.

Our goal is to minimize this objective function by adjusting the values of the parameters described previously.

$$\frac{\phi_m}{A_r} \frac{dw_i}{dx} = M_{w_i} \sum_{j=1}^3 v_{i,j} (1 - \epsilon_b) \rho_{cat} r_j$$

Figure 4. Governing mass balance equation used for parameter fitting in the referenced literature.

Where, $i = 1, 2, 3, 4, 5$ represent the various components involved in the reaction pathway

The above set of equations represent the system of coupled ODEs which need to be integrated over the length of the reactor to obtain the outlet weight fractions. If we integrate the above equation with respect to x from $x = 0$ to $x = L$, where L the length of the reactor bed is, we will get the weight fractions at the reactor exit for each of the i components. Let the solution to the above ODE system be given by

$$w_{i,pred} = w_{i,out} \quad i = 1, 2, 3, 4, 5$$

The sum of mean squared errors can be determined as follows:

$$\min. MSE = \sum_{k=1}^{94} \frac{1}{5} \sum_{i=1}^5 (w_{i,pred} - w_{i,expt})^2$$

Figure 5. The mean squared error objective function utilized in the referenced literature through use of outlet weight fractions derived from the governing mass balance equation and experimental data. The outer summation implies that the MSE is being evaluated and summed up for all 234 data points.

Once parameters are fitted to each model. We will perform a 5-fold Cross Validation (CV) analysis as well as the statistical metric “wp” presented in the literature that determines the percentage of correctly predicted data points within one standard deviation. The MSE, CV, and wp values for each of the models were final metrics utilized in the literature to determine which model was the most appropriate for use in industrial process modeling of methanol synthesis.

Optimization Method:

The detailed steps involved in the optimization formulation and the use of ODE solvers will be discussed here. The problem is inherently complex since the coupled ODE system is embedded within the optimization problem.

Initially, we attempted to use Pyomo to solve the problem. However, the need to use ‘**pyomo.DAE**’ to solve the ODEs within Pyomo proved to be too complex, so we decided to use a standard gradient-free black-box optimizer to solve the problem.

First, we convert all the volumetric flow rates given in the data set to mass flow rates. Then, we set up the system of differential equations after expressing the fugacity terms in terms of respective weight fractions. The basic functions/equations that are required to define the ODE system are mentioned below. These functions are also defined in separate Python code cells prior to starting the optimization formulation.

$$M_{avg.} = \frac{\frac{W_{CO2}}{M_{CO2}} + \frac{W_{H2}}{M_{H2}}}{\frac{W_{CO2}}{M_{CO2}} + \frac{W_{H2}}{M_{H2}}}$$

$$\rho = \frac{P \times 10^5 \times M_{avg.} \times 10^{-3}}{R \times (T + 273.15)}$$

$$\phi_m = \frac{\phi_v^o \times 10^{-6}}{60} \times \rho$$

$$D_r = 6.35 \times 10^{-5}$$

The above function converts all fugacity terms which appear in the ODE into weight fractions.

$$A_r = \frac{\pi}{4} \times D_r^2$$

$$1 - \epsilon_b \cong \frac{\rho_{cat}}{1300}$$

$$f_i = \frac{W_i / M_i}{\sum (W_i / M_i)} P$$

The above function converts all fugacity terms which appear in the ODE into weight fractions.

$$k_{CO_2} = A_{CO_2} \exp \left(\frac{-B_{CO_2}}{RT} \right)$$

$$k_{RWGS} = A_{RWGS} \exp \left(\frac{-B_{RWGS}}{RT} \right)$$

$$k_{H_2O/9} = k_{H_2} = \text{constant}$$

The above equations show the dependence of the lumped kinetic constants on temperature. These are what we will be estimating in the optimization problem.

$$K_{P_{CO_2}}^o = \exp \left[\frac{\left\{ 7.4114 \times 10^4 + (1.8926 \times 10^2) T + (3.2443 \times 10^{-2}) T^2 + (7.0432 \times 10^{-6}) T^3 \right\} - (5.6053 \times 10^{-9}) T^4 + (5.6053 \times 10^{-12}) T^5 - (6.4364 \times 10^1) T \ln T}{RT} \right]$$

$$K_{P_{RWGS}}^o = \exp \left[\frac{\left\{ -3.94121 \times 10^4 - (5.41516 \times 10^1) T - (5.5642 \times 10^{-2}) T^2 + (2.5760 \times 10^{-5}) T^3 \right\} - (7.6594 \times 10^{-9}) T^4 + (1.0161 \times 10^{-12}) T^5 - (1.8429 \times 10^1) T \ln T}{RT} \right]$$

These functions indicate the temperature-dependent equilibrium constants which appear in the rate equations.

$$\frac{\phi_m}{A_r} \frac{d W_{CO}}{dx} = (M_{CO}) [(1 - \epsilon_b) \rho_{cat} (r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{CO_2}}{dx} = (M_{CO_2}) [(1 - \epsilon_b) \rho_{cat} (-1 r_{CO_2} - r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{H_2}}{dx} = (M_{H_2}) [(1 - \epsilon_b) \rho_{cat} (-3 r_{CO_2} - r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{H_2O}}{dx} = (M_{H_2O}) [(1 - \epsilon_b) \rho_{cat} (+ r_{CO_2} + r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{CH_3OH}}{dx} = (M_{CH_3OH}) [(1 - \epsilon_b) \rho_{cat} (r_{CO_2})]$$

Finally, we have the system of coupled ODEs which we need to solve. Note that the above equations are valid for the 6-parameter model and the VB model. For the Graaf et. al model, we have additional non-zero stoichiometric coefficients and an additional rate equation for the consumption of CO.

$$\frac{\phi_m}{A_r} \frac{d W_{CO}}{dx} = (M_{CO}) [(1 - \epsilon_b) \rho_{cat} (-r_{CO} + r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{CO_2}}{dx} = (M_{CO_2}) [(1 - \epsilon_b) \rho_{cat} (-r_{CO_2} - r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{H2}}{dx} = (M_{H2}) [(1 - \epsilon_b) \rho_{cat} (-3 r_{CO2} - 2 r_{CO} - r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{H2O}}{dx} = (M_{H2O}) [(1 - \epsilon_b) \rho_{cat} (r_{CO2} + r_{RWGS})]$$

$$\frac{\phi_m}{A_r} \frac{d W_{CH3OH}}{dx} = (M_{CH3OH}) [(1 - \epsilon_b) \rho_{cat} (r_{CO2} + r_{CO})]$$

After defining all the required functions and making the necessary unit conversions, we defined the system of ODEs in Python and ran it through the ‘odeint’ function in the ‘SciPy.integrate’ package. The initial conditions were referenced from the initial weight fractions as given in the dataset. The step-size specified in the solver was restricted due to memory limitations. The end point of the solution array represents the final weight fractions and the MSE was calculated for all data points by looping the ODE solver through all 94 PFR data points and computing the MSE at each point. The total MSE was then obtained by averaging all 94 MSE values.

This piece of code was then input to a black-box optimizer. We chose to use the ‘BayesianOptimizer’ function which was one of many gradient-free solvers that we could have used to solve this problem. The parameters to be estimated were defined and the MSE loop was run through the function. The search space was restricted based on literature standards and the number of iterations were limited. We had hoped to continue to run the solver through as many iterations as possible till convergence but could not due to memory allocation issues.

After evaluating the parameters and the minimum MSE value, we evaluated the cross-validation coefficients for each model. This CV(5) metric is calculated by averaging the MSEs for five different 80/20 (training/testing) sets of data. Finally, we plotted the variances in the estimated MSE values for different shuffled training sets in the form of a Gaussian distribution. The results of our code are discussed below:

Results:

The different metrics were evaluated and used to compare the three models taken under consideration.

<pre> Results: 'TOTAL_MSE' Best score: 0.09687748354146937 Best parameter: 'Arwgs' : 6.11e+19 'Brwgs' : -180000.0 'Aco2' : 5815000000000000.0 'Bco2' : -230000.0 'kH2' : 4.6000000000000005 'kH2O9' : 292 Random seed: 2036750233 Evaluation time : 118.25650787353516 sec [99.99 %] Optimization time : 0.009905815124511719 sec [0.01 %] Iteration time : 118.26641368865967 sec [11.83 sec/iter] </pre>	<pre> Results: 'TOTAL_MSE' Best score: 0.08889517589755178 Best parameter: 'aCO' : 1110000000000.0 'bCO' : -80000 'aCO2' : 590 'bCO2' : -220000 'aRWGS' : 105100000000.0 'bRWGS' : -80000 'ACO' : 9.1029999999999968e-08 'BCO' : 2400 'ACO2' : 4.4100000000000004e-08 'BCO2' : 27000 'AH2O_H2' : 9.7e-08 'BH2O_H2' : 140000 Random seed: 675017982 Evaluation time : 185.520277261734 sec [99.99 %] Optimization time : 0.012880563735961914 sec [0.01 %] Iteration time : 185.53315782546997 sec [18.55 sec/iter] </pre>	<pre> Results: 'TOTAL_MSE' Best score: 0.004113656374721943 Best parameter: 'aCO2' : 0.5 'bCO2' : 72100 'aRWGS' : 33160000000000.0 'bRWGS' : -20000 'AH2' : 3.9000000000000004 'BH2' : 630 'AH2O' : 4.0069999999999987e-08 'BH2O' : 56000 'KH2O_H2' : 1230 Random seed: 1257484326 Evaluation time : 34.55939292907715 sec [99.97 %] Optimization time : 0.009934425354003906 sec [0.03 %] Iteration time : 34.56932735443115 sec [3.46 sec/iter] </pre>
--	---	--

Fig 6: MSE and parameter estimation in Python for a) 6-parameter model, b) Graaf et. al model and c) VB model

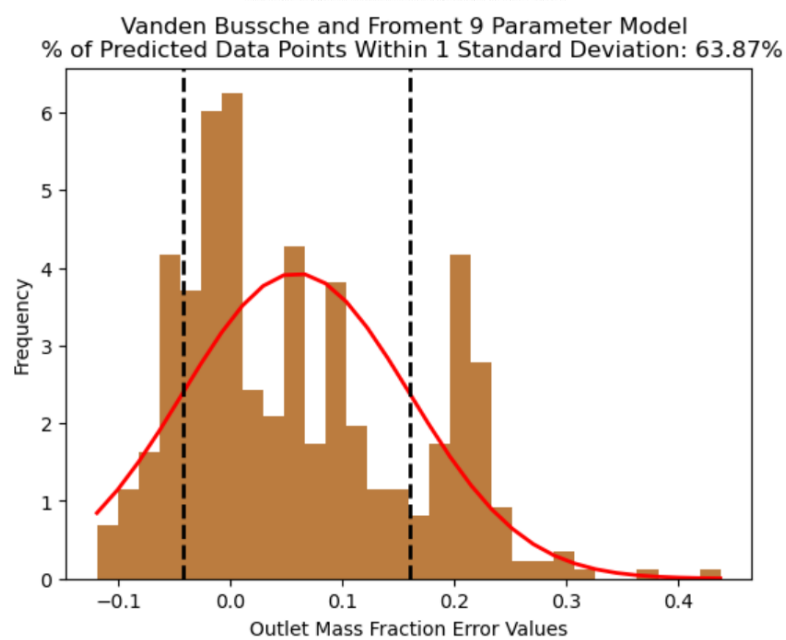
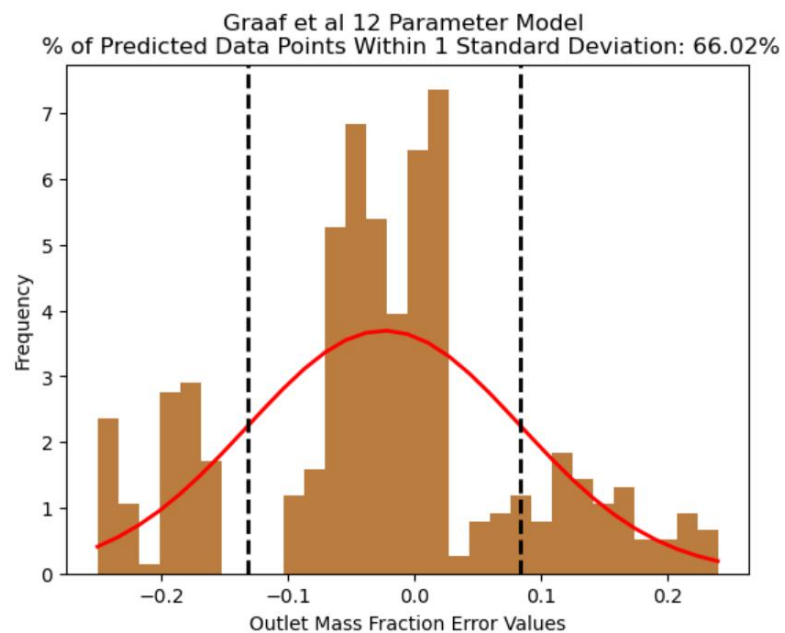
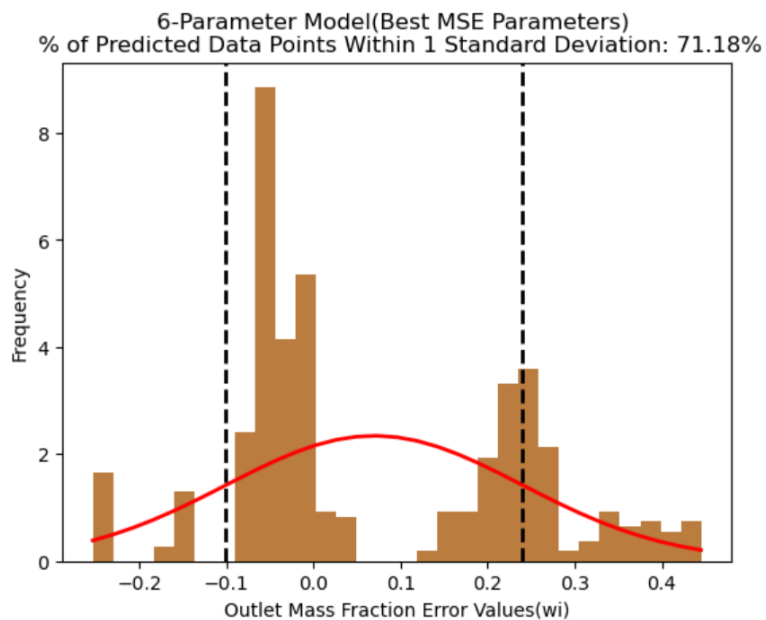


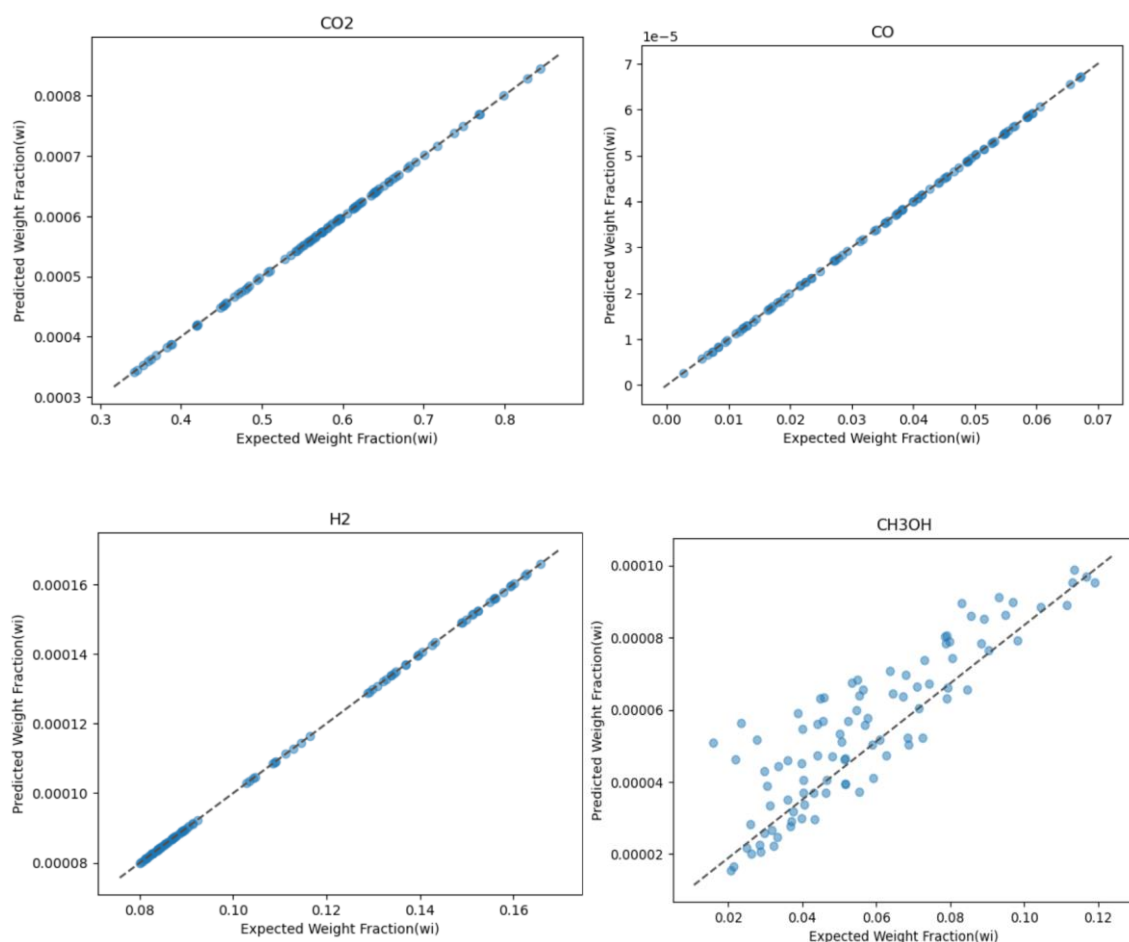
Fig 7: Distribution plots for the MSE deviation/variance from the mean (normalized to zero). Dashed lines represent $-\sigma$ and σ

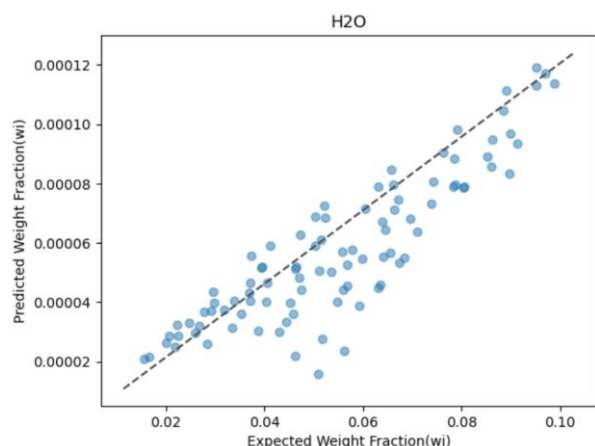
	Total MSE(Full Data Set)	Best MSE(CV5)	CV5	% Correctly Predicted Data Points
6-Parameter	0.0968	0.06673	0.068	71.18
Graaf-12 Parameter	0.0889	0.08224	0.082	66.02
V&B - 9 Parameter	0.0041	0.08222	0.082	63.87

Fig 8: Calculated values of MSE, CV5 and % of correctly predicted points

Based on the MSE values only, we would assume that the VB model is the best since it gives the lowest value among all models. However, the 6-parameter model has a lower CV(5) value implying that it is doing a better job at predicting the data when tested with different sets, i.e., it has more consistency and less variance in the results. This is confirmed from the distribution plots which show the 6-parameter model correctly predicting a larger fraction of the data points within one standard deviation ($-\sigma < x < \sigma$) than the other models. Also, it has the lowest number of parameters (simplest) which makes it desirable to use when modelling process equipment such as reactors etc. Our conclusions align with the paper's conclusions that the 6-parameter model does the best job of predicting the rate compared to the other models. Differences in our estimated values from those in the paper can be attributed to the fact that we chose only 94 data points owing to the difference in the method of sampling as explained previously.

Parity plots which show the discrepancy between the predicted and experimental outlet weight fraction values for each component are shown below for the 6-parameter model to give a better idea of the accuracy.





As can be seen from the parity plots, there is appreciable deviation for methanol and water. This can be attributed to the limited number of iterations run in the optimization solver due to memory limitations.

Part II – Creating a new model using ALAMO:

The goal of the second part of our project is to determine if ALAMO's computational methodology can provide ease of model selection for steady state methanol synthesis under the Cu/ZnO/Al₂O₃ catalyst without the need for manual statistical analysis explained in part 1. ALAMO is a regression and classification model learning methodology that constructs the simplest model with the highest accuracy from minimal data sets or Blackbox simulations. ALAMO uses an integer-programming-based best subset technique that compares many explicit transformations of the original input variables. Once a model is formulated within ALAMO, it is tested and enhanced through derivative-free optimization solvers that sample attractive points. ALAMO uses its own sampling method called error maximization sampling (EMS) that allows greater fidelity in constructing a model than traditional DOE methods such as Latin hypercubes and orthogonal arrays. Due to ALAMO's capability of performing its own regression, sampling method, and final statistical model metrics by simply input and output data, we conclude two formulations to test within ALAMO to replicate or improve results determined in the literature and part 1 of the project.

The first approach we will test within ALAMO is to see if simply providing the 3 derived rate equations in the literature can produce the same statistical conclusions reached within the paper and part 1 of the project. An example of the formulation for this first approach will be discussed along a tutorial of ALAMO's GUI. ALAMO's GUI is surprisingly simple and allows the user to specify multiple inputs and output variables. The initial screen when opening ALAMO is shown below in Figure 6 depicting an example of our 6 input variables and the predicted output rate for the 6-parameter

ALAMO									
File Edit Help									
Enter data		Plot data		Run ALAMO		View results			
	A	B	C	D	E	F	G	H	I
var	T	fco2	fco	fmeth	fh2	fh2O	rate	rate 2	
xmin	451.4762...	2.188	0.286	13.688	0.198	0.504			
xmax	533.15	7.658	1.001	47.908	0.693	1.764			
1	503.15	3.282	0.429	20.532	0.297	0.756	0.000214...	START NDATA	START NDATA
2	513.15	3.282	0.429	20.532	0.297	0.756	0.000695...		
3	523.15	3.282	0.429	20.532	0.297	0.756	0.002079...		
4	533.15	3.282	0.429	20.532	0.297	0.756	0.005832...		
5	503.15	4.376	0.572	27.376	0.396	1.008	0.000250...		
6	513.15	4.376	0.572	27.376	0.396	1.008	0.000814...		
7	523.15	4.376	0.572	27.376	0.396	1.008	0.002432...		
8	533.15	4.376	0.572	27.376	0.396	1.008	0.006824...		
9	503.15	5.47	0.715	34.22	0.495	1.26	0.000282...		
10	513.15	5.47	0.715	34.22	0.495	1.26	0.000918...		
11	523.15	5.47	0.715	34.22	0.495	1.26	0.002745...		
12	523.15	5.47	0.715	34.22	0.495	1.26	0.002745...		
13	533.15	5.47	0.715	34.22	0.495	1.26	0.007699...		
14	503.15	6.564	0.858	41.064	0.594	1.512	0.000311...		
15	513.15	6.564	0.858	41.064	0.594	1.512	0.001012...		
16	523.15	6.564	0.858	41.064	0.594	1.512	0.003027...		
17	533.15	6.564	0.858	41.064	0.594	1.512	0.008491...		
18	503.15	3.282	0.429	20.532	0.297	0.756	0.000214...		
19	513.15	3.282	0.429	20.532	0.297	0.756	0.000695...		
20	523.15	3.282	0.429	20.532	0.297	0.756	0.002079...		
21	533.15	3.282	0.429	20.532	0.297	0.756	0.005832...		
22	503.15	4.376	0.572	27.376	0.396	1.008	0.000250...		

Figure 6. Initial screen of the ALAMO GUI with the 6 input experimental data values with associated bounds(xmin,xmax) and predicted rate values of the 6-paramter model. Once specifying the inputs and outputs you can then move to the “Plot data” tab to get a preliminary idea of the relationships between all variables. Additionally, you can view a histogram of frequency versus specified variable to visualize any missing data gaps. Figure 7 and 8 below depict an example of the “Plot data” tab for the stated example.

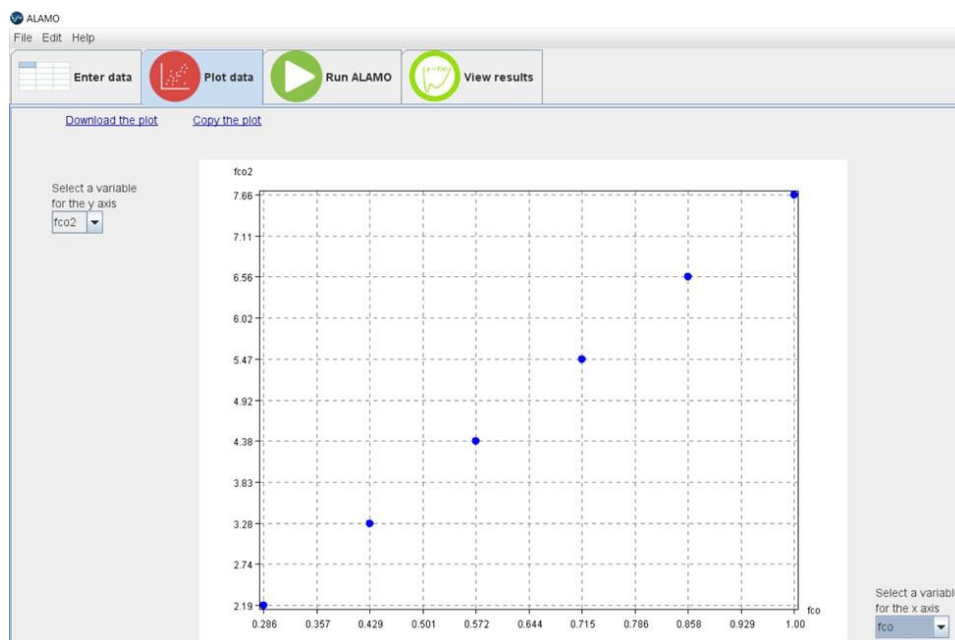


Figure 7. “Plot data” screen on ALAMO GUI depicting the relationship between the fugacity of CO and CO₂.

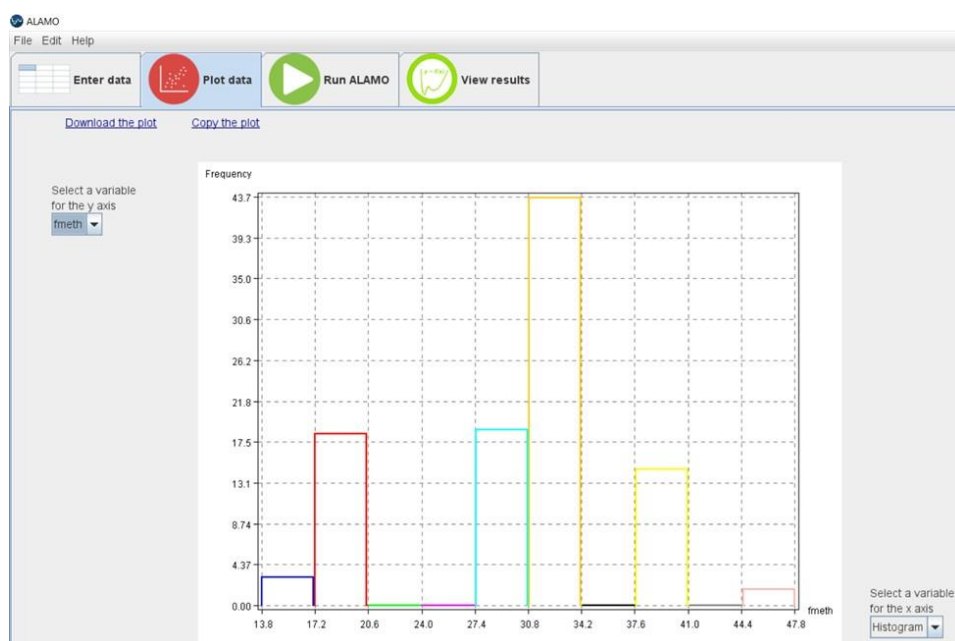


Figure 8. Histogram of the frequency of outlet methanol fugacity's within the experimental data.

Next we turn to the “Run ALAMO” tab which displays many options to be initialized before running the simulation. The “Run tab” is shown in Figure 9 below.

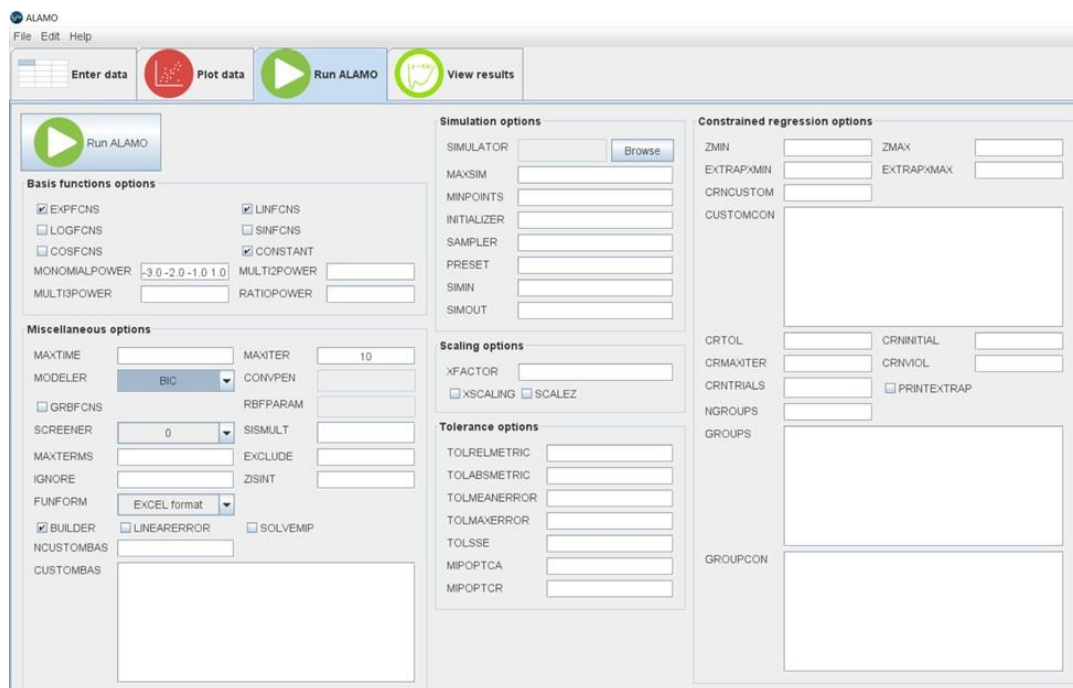


Figure 9. Run ALAMO tab prior to running the simulation where basis functions and model constraints are formulated.

The first set of important options you can provide ALAMO are customized basis functions to consider for model construction. We know the terms in the derived rate determining step equations already, thus we will unselect basis functions such as sin and cos etc. Custom basis functions are also allowable within this ALAMO tab along with other options such as constrained regression. Constrained regression in ALAMO allows formulation of constraints on the response function as well as combinatorial constraints on the basis functions. Customized basis functions will be incorporated for each of the 3 given models within the CUSTOMBAS option and combinatorial constraints will be inputted within the CUSTOMCON tab. Using the 6-paramter model as reference for this application, the general basis set would be $X_1 \in [f^{\pm[0.25,0.5,1,1.5,2,2.5,3]}, \exp(K(T^{-1}))]$, where K is a constant multiplied by the independent temperature variable T as laid out in the Arrhenius formulas for parameter fitting.

The final step is to run ALAMO which will generate two different screens. The first screen depicts the final surrogate model generated by ALAMO aside the final quality metrics and parity plot shown in Figure 10 and 11.

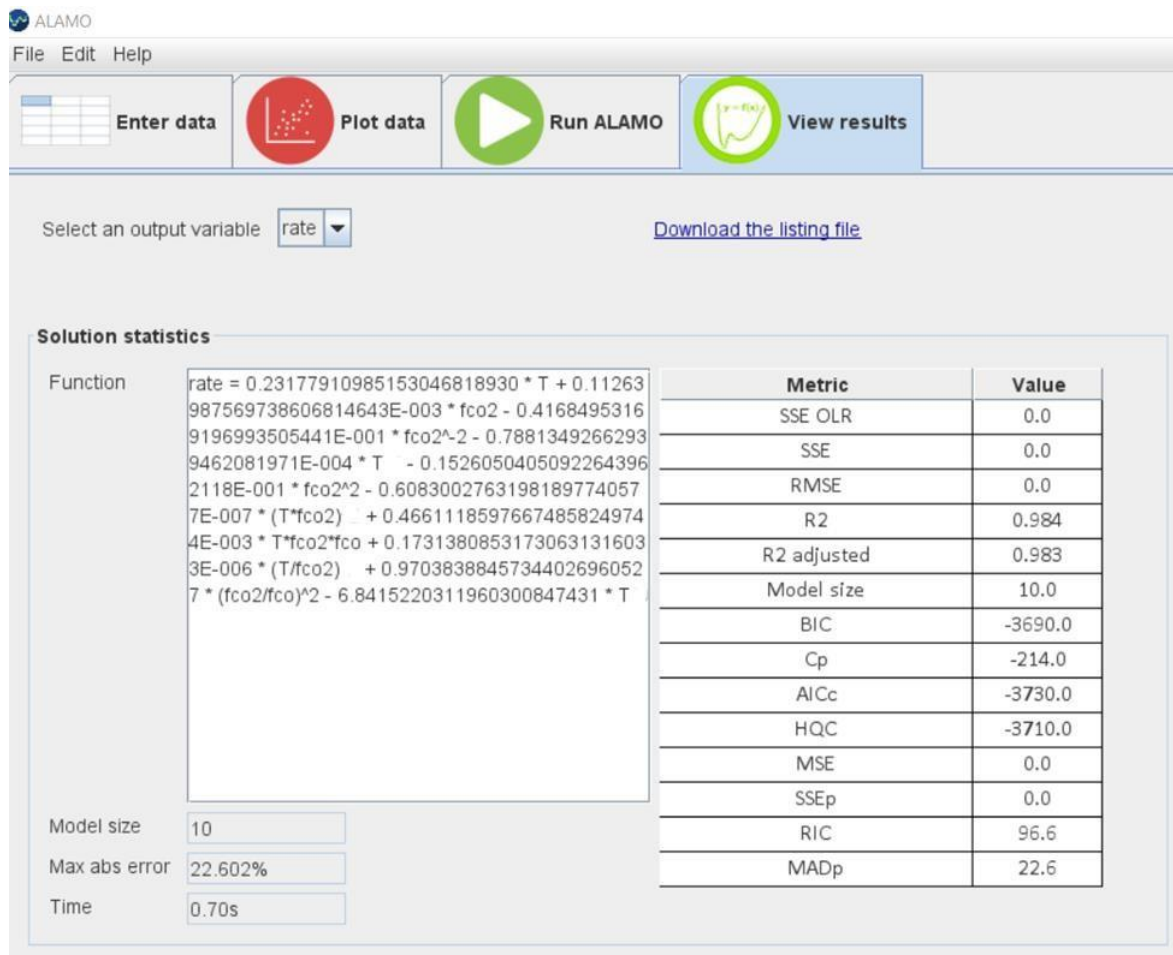


Figure 10. Generated surrogate model in ALAMO with quality metrics.

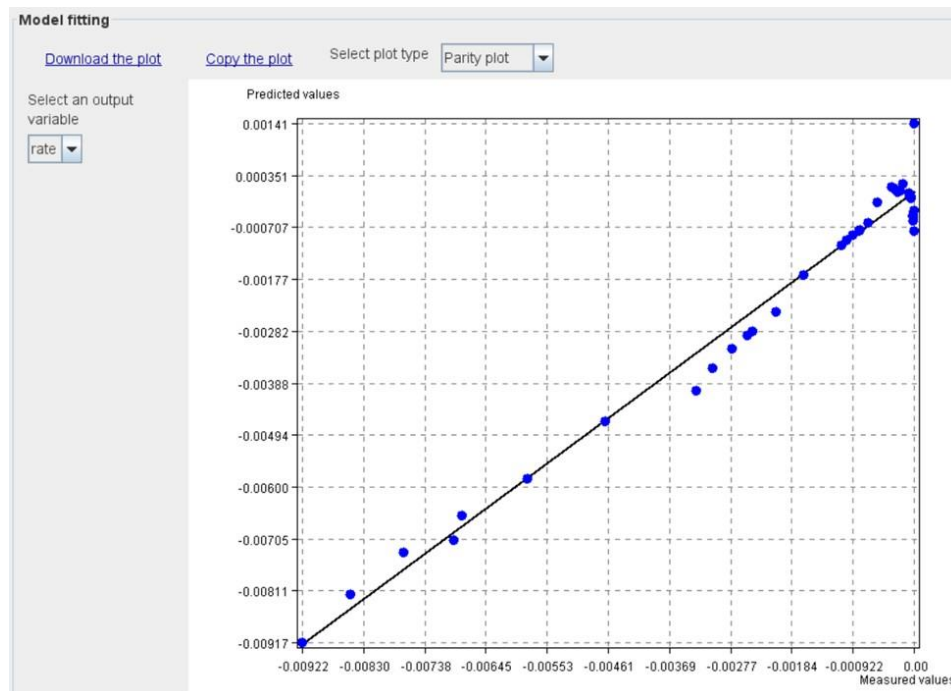


Figure 11. ALAMO generated parity plot depicting relationship between predicted and measured values.

The second approach we will attempt within ALAMO is a type of differential mapping using the governing reactor equation from part 1 of the project. The first ALAMO approach described above simply uses the outlet fugacity and the predicted rates from the already known rate determining step equations to determine the most accurate and simplest predicting model by ALAMO's output. However, the governing mass balance equation accounts for the change in fugacity through the length of the reactor that would require a method of formulating inputs to ALAMO

through some sort of differential mapping to arrive at a more accurate result. This equation varies in the number of rate equation terms by the summation of j for each model based on the underlying mechanisms proposed by the models' authors. For instance, the newer 6-parameter model and the Vanden Busshe et al. model do not incorporate the direct CO hydrogenation to methanol reaction, so the governing equation contains only 2 r_j terms upon fitting.

Differential mapping would utilize the older rate model described in the literature that accounts for all 3 of the governing stoichiometric methanol production equations to be regularized from solely the experimental data. In this way an older model would be tested on the outputs of ALAMO's surrogate model which would theoretically allow the data to bypass expert derivations for novel models. Formulating the inputs to ALAMO through this method is yet to be determined.