Support Information

A Novel AI-Driven Approach for Parameter Estimation in Gas-Phase Fixed-Bed Experiments

Rui D. G. Matias[a,b], Alexandre F.P. Ferreira[a,b], Idelfonso B.R. Nogueira[c],Ana M. Ribeiro[a,b,*]

[a]Laboratory of Separation and Reaction Engineering–Laboratory of Catalysis and Materials (LSRE-LCM), Department of Chemical Engineering, University of Porto, Porto, 4200-465, Portugal

[b]ALiCE–Associate Laboratory in Chemical Engineering, Faculty of Engineering, University of Porto, Porto, 4200-465, Portugal

[c]Chemical Engineering Department, Norwegian University of Science and Technology, Sem Sælandsvei 4, Kjemiblokk 5, Trondheim, 793101, Norway

* Corresponding Author: apeixoto@fe.up.pt

1. Experimental Dataset

The experimental dataset was obtained from [1] and includes 17 breakthrough experiments: 7 with pure CO_2, 7 with pure CH_4 and 7 binary experiments. The main experimental conditions and the adsorbed amount of each gas, obtained after mass balance, can be consulted in Table SI.1. Experiments 6,7, 13, 14, 15 and 17 were used as the training group to reduce the computational effort.

Table SI.1: List and Condition of Experiments

	Experiment	y_{CO2}	y_{CH4}	P_t (bar)	T (K)	q_{CO2} (mol/kg)	q_{CH4} (mol/kg)
Pure CO2	1	0.129	0	1.1	303	0.510	0
	2	0.259	0	1.1	303	0.850	0
	3	0.391	0	1.1	303	1.207	0
	4	0.485	0	1.1	303	1.340	0
	5	0.802	0	1.5	303	2.513	0
	6	0.802	0	2.5	303	3.298	0
	7	0.801	0	4.4	303	4.327	0
Pure CH4	8	0	0.129	1.1	303	0	0.130
	9	0	0.394	1.1	303	0	0.370
	10	0	0.487	1.1	303	0	0.440
	11	0	0.250	1.1	303	0	0.240
	12	0	0.799	1.25	303	0	0.656
	13	0	0.780	2.5	303	0	1.418
	14	0	0.840	4.4	303	0	2.140
Binary	**15**	0.117	0.119	1.1	303	0.451	0.072
	16	0.499	0.501	1.2	303	1.560	0.416
	17	0.389	0.433	4	303	2.844	0.767

2. Reinforcement Learning Algorithm

Agent

The Deep Deterministic Policy Gradient (DDPG) algorithm is a widely used off-policy actor-critic method designed for tasks involving continuous action spaces. It operates by learning a deterministic policy, using a Q-value function that evaluates the value of the policy. The algorithm incorporates essential components such as target networks for both the actor and critic, alongside an experience replay mechanism.

A key challenge in reinforcement learning arises from the sequential nature of data collection, which violates the assumption of independent and identically distributed samples that most optimization methods rely on. To address this, DDPG uses a replay buffer, a finite storage system that saves past interactions, including states, actions, rewards, and observations. When the buffer reaches its capacity, older entries are replaced with newer ones. During training, random batches of experiences are sampled from the buffer to update the actor and critic networks. Since DDPG is off-policy, the replay buffer can be large, allowing the algorithm to learn from a broad range of uncorrelated transitions, which helps reduce correlations between samples and improve training stability.

The DDPG agent implemented in this study builds on the framework outlined in [2], with some modifications aimed at simplifying the architecture while maintaining or improving the performance. While the neural networks for the actor, critic, and agent are continuously refined during training, careful selection of initial hyperparameters is critical for achieving optimal results. A sensitivity analysis was performed to fine-tune these parameters, and the ranges explored for each hyperparameter are detailed in Table SI.2.

Table SI.2: Sensitivity Analysis and Hyperparameter Selection for the Implemented DDPG Agent

	Agent			**Actor and Critic**			
	Min	**Max**	**Selected**		**Min**	**Max**	**Selected**
Discount Factor	0.90	0.99	0.99	**Learn Rate**	0.001	0.1	0.1
Batch Size	64	320	128	**Optimizer**	-		Adam
Experience Buffer Length	10	1000	1000	**Gradient Decay**	0.85	0.95	0.9
Smooth Factor	1×10^{-4}	1×10^{-2}	1×10^{-3}	**L2 Regularisation**	1×10^{-5}	1×10^{-3}	1×10^{-4}
				Number of hidden Layers	64	320	256

Exploration is one of the primary challenges when working with continuous action spaces. A key benefit of off-policy algorithms like DDPG is their ability to separate the exploration strategy from the learning process. To facilitate exploration, a noise-based strategy was employed by introducing noise sampled from a stochastic process. Specifically, Ornstein–Uhlenbeck noise was used, following a similar approach to that described in [2]. To optimize the noise parameters and ensure effective exploration, a sensitivity analysis was conducted, similar to the one performed for other hyperparameters (Table SI.3). This analysis aimed to identify the parameter settings that would provide the best performance for the given task.

Table SI.3: Sensitivity Analysis and Hyperparameter Selection for the Implemented Ornstein–Uhlenbeck function

	Ornstein–Uhlenbeck noise		
	Min	**Max**	**Selected**
Mean	-		0
Standard Deviation	0.05	0.3	0.15
Mean Attraction	0.01	0.5	0.15
Standard Deviation Decay	1×10^{-5}	1×10^{-3}	1×10^{-4}

Actions

The action applied in each trial must be constrained within specified lower and upper limits, with an initial guess provided. Table SI.4 presents the action limits and initial guess for the various approaches.

Table SI.4: Action Constraints Applied in the different Approaches

Parameter	Units	Approach 1 and 2			Approach 3		
		Act_{min}	Act_{max}	Act_0	Act_{min}	Act_{max}	Act_0
$q_{sat,CO2}$	[mmol g^{-1}]	5.10	7.65	6.38	2.00	9.00	5.50
$q_{sat,CH4}$	[mmol g^{-1}]	5.90	8.86	7.38	2.00	9.00	5.50
$b_{\infty,\mathrm{CO2}} \times 10^5$	[bar^{-1}]	1.24	1.86	1.55	0.10	10.00	5.05
$b_{\infty,\mathrm{CH4}} \times 10^5$	[bar^{-1}]	6.45	9.68	8.07	0.10	10.00	5.05

To maintain consistency in the rate of change for each parameter across trials, the action generated by the agent is first normalized (Act_{norm}) to a range of -1 to 1, as described by Equation SI.1. Once normalized, the action is then re-scaled (Act) before being passed to the environment (Equation SI.2). This re-scaling converts the action into a set of parameters that are appropriate for simulating the experiments. This process helps to mitigate the significant differences in the order of magnitude between the parameters being estimated, ensuring that they are treated on a more comparable scale throughout the trials.

$$Act_{norm} = 2 \times \frac{Act - Act_{min}}{Act_{max} - Act_{min}} - 1 \tag{SI.1}$$

$$Act = \frac{(Act_{norm} + 1) \times (Act_{max} - Act_{min})}{2} + Act_{min} \tag{SI.2}$$

3. Results

Different Approaches

The main results for approach 1, 2 and 3 can be consulted in Table SI. 5. This Table displays the value of the objective function (error between simulated and experimental data) using the reference parameters (f_{obj}^{Ref}) and the ones obtained on the different approaches (f_{obj}^{RL}) for the pure CO_2, pure CH_4 and binary experiments and for the full set of experiments. It can also be seen the percentual variation (calculated using Equation SI.3) of the objective function compared to the reference values.

$$\%\Delta = \frac{f_{obj}^{RL} - f_{obj}^{Ref}}{f_{obj}^{Ref}} \times 100 \qquad \textbf{(SI.3)}$$

The green values show an improvement on the adjustment between simulated and experimental data compared to simulated data using the reference parameters while the red values show a decrease in the capability of accurately simulate the process. Lastly, this Table also presents the required time to run every approach.

Table SI.5: Objective Function Values from Simulations for Pure CO_2, Pure CH_4, Binary Mixture, and All Experiments with Reference and Estimated Parameters via Approach (Ap.) 1, 2 and 3

f_{obj}	Reference	Approach 1	Approach 2	Approach 3
Pure CO_2 ($\times 10^4$)	7.06	6.43 (-9%)	4.12 (-42%)	3.95 (-44%)
Pure CH_4 ($\times 10^4$)	2.86	2.96 (+4%)	3.44 (+20%)	2.73 (-4%)
Binary ($\times 10^3$)	2.07	1.89 (-9%)	1.43 (-31%)	1.45 (-30%)
Global ($\times 10^3$)	3.06	2.82 (-8%)	2.19 (-28%)	2.11 (-31%)
Computational Time (h)		16.3	26.27	26.82

3.2. Approach 3

In order to better understand the behaviour of the agent during training, the action taken in each trial was registered. Figure SI.1 illustrates the normalized action taken by the agent during each trial throughout the search process, for the various parameters to be estimated.

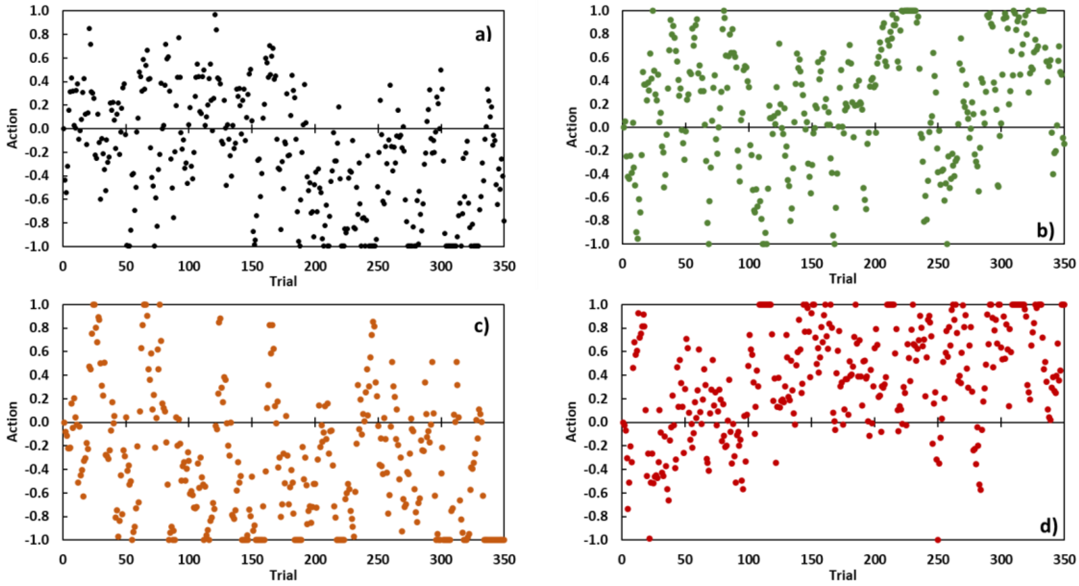

Figure SI.1: Given Normalized Actions Across Trials for All Parameters: a) $q_{sat,CO2}$ (•) ; b) $q_{sat,CH4}$ (•) ; c) $b_{\infty,\mathrm{CO2}}$(•) and d) $b_{\infty,\mathrm{CH4}}$(•)

To evaluate the influence that small changes on the taken actions have on the objective function, a sensitivity analysis was performed by multiplying the obtained set of parameters on approach 3 by different multiplication factors. The results are displayed in Table SI.6. Additionally, a percentage comparison highlights the increase or decrease in the objective function relative to the results obtained using the parameter set from approach 3 (f_{obj}^{Ap3}), calculated using Equation SI.4.

$$\%\Delta = \frac{f_{obj} - f_{obj}^{Ap3}}{f_{obj}^{Ap3}} \times 100 \quad \textbf{(SI.4)}$$

Table SI.6: Sensitivity Analysis of the system to small variations of all the parameters

Multiplication Factor	$f_{obj}\,(\times 10^3)$	$\%\Delta$
Approach 3	2.11	
+10%	9.29	+339.36%
+5%	3.71	+75.71%
+1%	2.06	-2.58%
+0.1%	2.08	-1.49%
+0.01%	2.11	-0.17%
-0.01%	2.12	+0.17%
-0.1%	2.16	+1.95%
-1%	2.91	+37.53%
-5%	5.80	+174.42%
-10%	1.13	+436.31%

The same annalyzis was done after multiplying each parameter by a factor of ± and the results are displayed in Table SI.7.

Table SI.7: Sensitivity Analysis of the system to 1% variations per parameter

θ	Multiplication Factor	$f_{obj}\,(\times 10^3)$	%Δ
	Approach 3	2.11	
$q_{sat,CO2}$	+1%	2.06	-2.58%
	-1%	2.20	+4.28%
$q_{sat,CH4}$	+1%	2.05	-3.13%
	-1%	2.34	+10.67%
$b_{\infty,\mathrm{CO2}}$	+1%	2.09	-1.26%
	-1%	2.16	+2.02%
$b_{\infty,\mathrm{CH4}}$	+1%	2.04	-3.33%
	-1%	2.31	+9.17%

To evaluate the capability of the estimated set of parameters to predict the adsorbed amount per experiment, the Extended Langmuir Isotherms were plotted using the obtained and the reference parameters and compared with the experimental adsorbed amount, obtained after mass balance in [1]. This was done for the pure (Figure SI.2) and binary (Figure SI.3) experiments and the MSE calculated for each type of experiment can be consulted in Table SI.8.

Table SI.8: Mean Squared Error for the adsorbed amount prediction with the Extended Langmuir Model

	q_i - MSE	
	Reference	**Approach 3**
Pure CO_2 ($\times 10^3$)	4.40	7.11 (+61%)
Pure CH_4 ($\times 10^3$)	2.77	2.52 (-9%)
Binary ($\times 10^3$)	4.86	8.33 (+71%)
Global ($\times 10^3$)	3.97	5.87 (+48%)

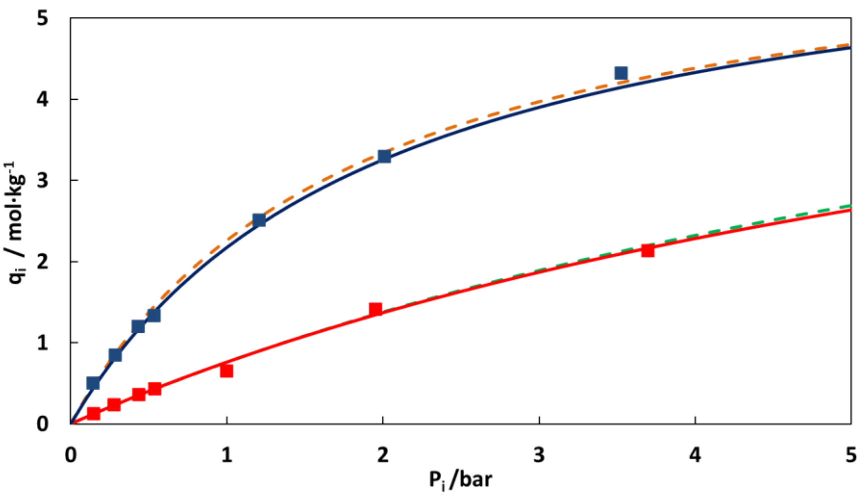

Figure SI.2: Adsorption isotherms for CO_2 and CH_4 on MIL-53(Al) at 303K for pure experiments. Experimental data points are shown as squares (■ for CO_2, ■ for CH_4). The curves represent the Extended Langmuir model fits: solid lines (— for CO_2, — for CH_4) using Reinforcement Learning-derived parameters, and dashed lines (--- for CO_2, --- for CH_4) using Reference parameters.

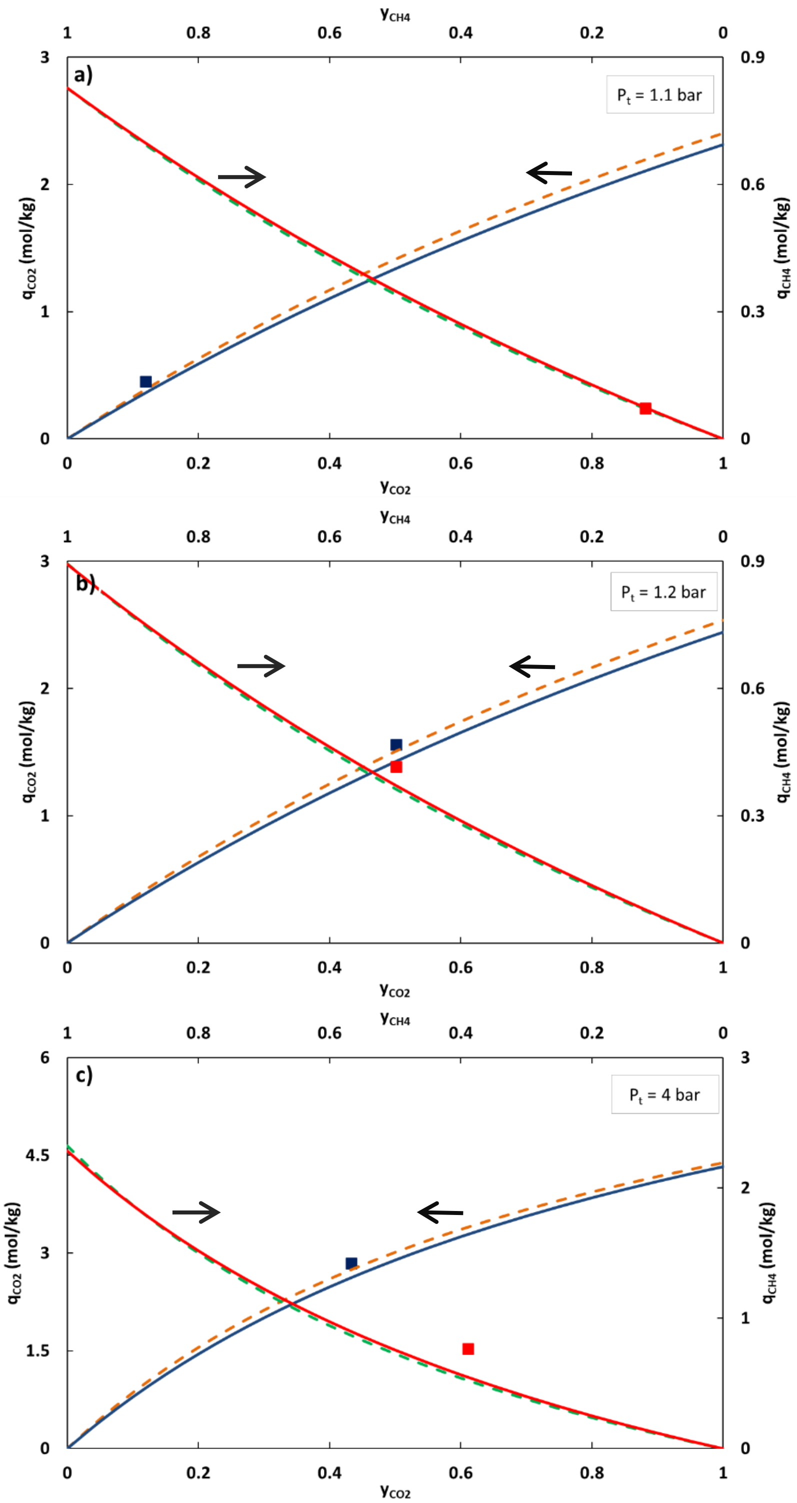

Figure SI.3: Adsorption isother ms for CO_2 and CH_4 on MIL-53(Al) at 303K for binary experiments: (a) Experiment 15, (b) Experiment 16, and (c) Experiment 17. Experimental data points are shown as squares (■ for CO_2, ■ for CH_4). The curves represent the Extended Langmuir model fits: solid lines (— for CO_2, — for CH_4) using Reinforcement Learning-derived parameters, and dashed lines (--- for CO_2, --- for CH_4) using Reference parameters.

To demonstrate the capability of the new parameters in accurately simulating the experiments, breakthrough curves were plotted to compare simulations using the new parameters with those using the reference parameters, as shown in Figure SI.4.

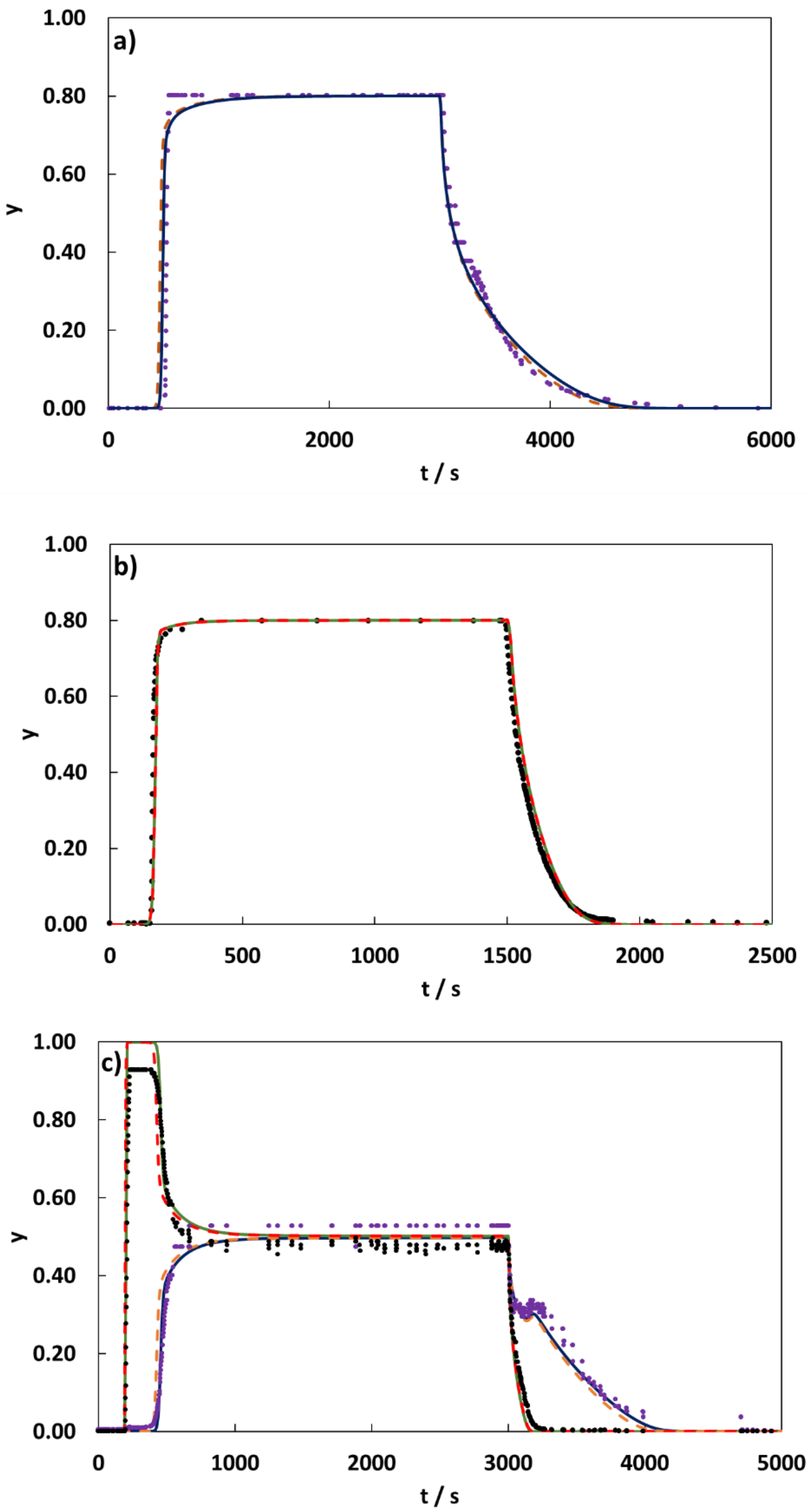

Figure SI.4: Comparison of experimental and simulated breakthrough curves: (a) Pure CO_2 (Experiment 5), (b) Pure CH_4 (Experiment 12), and (c) Binary mixture (Experiment 16). Simulations using reference parameters are shown as dashed lines (--- for CO_2, --- for CH_4). Simulations using RL-derived parameters (Approach 3) are shown as solid lines (– for CO_2, – for CH_4). Experimental data points are represented as filled circles (• for CO_2, • for CH_4).

References

1. Ferreira, A.F.P., et al., Methane purification by adsorptive processes on MIL-53(Al). Chemical Engineering Science, 2015. 124: p. 79-95.
2. Lillicrap, T., Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971, 2015.