



PAULA NKULIKIYINKA

APPLICATION OF MACHINE LEARNING IN HYDROGEN
PRODUCTION VIA THE PROCESS OF SORBENT ENHANCED
STEAM METHANE REFORMING

SCHOOL OF WATER, ENERGY AND ENVIRONMENT
PhD in Energy and Power

PhD
Academic Year: 2019 - 2023

Supervisor: Dr Peter T. Clough
Associate Supervisors: Prof. Vasilije Manovic and Dr Stuart T.
Wagland
June 2023



SCHOOL OF WATER, ENERGY AND ENVIRONMENT
Energy and Power

PhD
Academic Year 2019 - 2023

PAULA NKULIKIYINKA

APPLICATION OF MACHINE LEARNING IN HYDROGEN
PRODUCTION VIA THE PROCESS OF SORBENT ENHANCED
STEAM METHANE REFORMING

Supervisor: Dr Peter T. Clough
Associate Supervisors: Prof. Vasilije Manovic and Dr Stuart T.
Wagland
June 2023

This thesis is submitted in partial fulfilment of the requirements for
the degree of PhD.

© Cranfield University 2023. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

Academic Integrity Declaration

I declare that:

- the thesis submitted has been written by me alone.
- the thesis submitted has not been previously submitted to this university or any other.
- that all content, including primary and/or secondary data, is true to the best of my knowledge.
- that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

This thesis is focused on the exploration of the use of machine learning and computational methods for modelling process conditions and for materials screening within the process of sorbent enhanced steam methane reforming (SE-SMR) for carbon-abated hydrogen production. Hydrogen is a clean, abundant and versatile energy carrier that can be used for a wide range of applications. However, the production of hydrogen is still largely dependent on fossil fuels, which presents a significant challenge for achieving a truly sustainable energy system.

The purpose of this study is to address this challenge by exploring novel approaches to hydrogen production, namely using machine learning, thermodynamic simulations, theoretical modelling, and the proposal of new methodologies and materials for low-carbon hydrogen production.

Three main areas of work were conducted within this thesis, which include 1) two surrogate models have been developed and used to predict and estimate variables that would otherwise be difficult direct measured.; 2) applying machine learning, namely quantitative structure–property relationship analysis (QSPR) has been employed in the exploration of combined sorbent catalyst material (CSCM) for SE-SMR; and 3) applying machine learning to screen suitable metal organic frameworks (MOFs) for the storage of the produced blue hydrogen.

Firstly, a surrogate model, was developed which was done by firstly simulating the model in Aspen Plus, applying a sensitivity analysis to gather a large dataset, then applying two multiple linear regression model, to observe the accuracy of predicting the gas concentration outputs. Two models were successfully developed with both models were accurate with high R^2 values, all above 98%.

Secondly, the novel approach of QSPR with inductive transfer learning and datamining, was applied to develop two large databases of sorbent and catalyst properties, respectively. Then the developed machine learning models from these databases were applied, to predict the optimal conditions and precursor

materials for the highest performing CSCM, in terms of last cycle capacity and methane conversion. Lastly, a similar approach was applied for the screening of MOFs for the storage of hydrogen by using multiple linear regression, simple geometric descriptors, and patterns in data to identify a better performing MOF than the currently reported experimental MOFs in literature.

Keywords:

Carbon capture and storage, combined sorbent catalyst materials (CSCM), sorption enhanced steam methane reforming (SE-SMR), quantitative structure property relationship (QSPR), calcium looping, metal organic frameworks (MOFs)

Acknowledgements

Firstly, I would like to express my immense gratitude to my supervisor Dr Peter Clough for giving me the opportunity to complete this PhD, along with his unwavering support, invaluable academic guidance, and most importantly patience throughout. It has not been an easy journey and I could not have done it without Dr Clough.

Additionally, I would like to thank Professor Vasilije Manovic and Dr Stuart T Wagland, for their support and motivation, valuable suggestions and wise words that have formed a great part in shaping me into the researcher I have become. I would like to extend my thanks to my reviewers, Professor Stephen Hallett and Dr Dawid Hanak, for their useful suggestions, direction and encouragement that kept me going throughout. I would like to acknowledge the financial support from the UK Engineering and Physical Sciences Research Council Doctoral Training Partnership (EPSRC DTP) grant No. EP/R513027/1 and the UKRI UK-Canada Globalink Research Award, grant No. NE/V019457/1. Without this support, none of this would be possible.

I would like to thank my co-authors Dr Yan (Harry) Yongliang and Dr Fatih Güleç for their continued support from the time I started right to the very end. Not only were they the best role models for me during my PhD, but I could always count on them for assistance and advice.

I would like to express my gratefulness to the Cranfield University Student Support team, particularly Emma Stevens. The support I received from the team is appreciated and I would not have completed this PhD without the additional support from Cranfield University. I would also like to extend this thanks to Nicola Lowe.

Lastly but most importantly, I would like to thank my family and friends for their continued love and support. To my parents, Viviane and Leopold, thank you for everything and for selflessly encouraging me to embark on this journey, I could not have done any of this without you. To my siblings, you kept me going when I needed you the most.

In memory of
Professor Henricus (Hans) J. Michels
Imperial College London
(1934-2019)

List of Publications

Published papers presented in this thesis:

- **Chapter 3:**

Yan, Y.L.; Borhani, T.N.; Subraveti, S.G.; Pai, K.N.; Prasad, V.; Rajendran, A.; **Nkulikiyinka, P.**; Asibor, J.O.; Zhang, Z.E.; Shao, D.; et al. Harnessing the Power of Machine Learning for Carbon Capture, Utilisation, and Storage (CCUS)—A State-of-the-art Review. *Energy Environ. Sci.* 2021, 14, 6122–6157.

- **Chapter 4:**

Nkulikiyinka P, Yan Y, Güleç F, Manovic V, Clough P.T. Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft-sensor Models. *Energy and AI* 2020, 2, 100037.

- **Chapter 5:**

Nkulikiyinka P, Wagland S.T, Manovic V, Clough P.T. Prediction of Combined Sorbent and Catalyst Materials for SE-SMR, Using QSPR and Multitask Learning. *Industrial & Engineering Chemistry Research* 2022 61 (26), 9218-9233.

Declaration

I declare that this thesis is an original report of my research, which has been written by myself and has not been submitted for any other degree or professional qualification.

The journal papers contained in this submission were published by Elsevier and American Chemical Society that allow reusing and reproducing the entire published material in an unchanged form in the PhD Thesis.

My contribution and those of the other authors to the collaborative work have been explicitly indicated in the text below and acknowledgements. All the co-authors and publishers allow me to use the joint publications in my thesis.

Paper 1: Yan, Y.L.; Borhani, T.N.; Subraveti, S.G.; Pai, K.N.; Prasad, V.; Rajendran, A.; **Nkulikiyinka, P.**; Asibor, J.O.; Zhang, Z.E.; Shao, D.; et al. Harnessing the Power of Machine Learning for Carbon Capture, Utilisation, and Storage (CCUS)—A State-of-the-art Review. *Energy Environ. Sci.* 2021, 14, 6122–6157. (Link to [chapter 3](#))

Contribution:

- Yongliang Yan: Conceptualization, Visualization, Writing – Original Draft, and writing – review and editing.
- Tohid N. Borhani, Sai Gokul Subraveti, Kasturi Nagesh Pai, Vinay Prasad, Arvind Rajendran, **Paula Nkulikiyinka**, Jude Odianosen Asibor, Zhien Zhang, Ding Shao, Lijuan Wang, Wenbiao Zhang, Yong Yan, William Ampomah, Junyu You, Meihong Wang: Witing – original draft, and writing – review and editing.
- Peter T. Clough, Vasilije Manovic and Edward J. Anthony: conceptualization, project administration, funding acquisition, supervision, and writing – review and editing.

Paper 2: **Nkulikiyinka P**, Yan Y, Güleç F, Manovic V, Clough P.T. Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft-sensor Models *Energy and AI* 2020, 2, 100037. (Link to [chapter 4](#))

Contribution:

- **Paula Nkulikiyinka**: Conceptualization, Methodology, Software, Validation, Investigation, Formal analysis, Data curation, Visualization, Writing - original draft.
- Peter T. Clough: Project administration, Supervision, Visualization, Writing - review & editing.
- Yongliang Yan: Writing - review & editing.
- Fatih Güleç: Writing - review & editing.
- Vasilije Manovic: Supervision, Writing - review & editing.

Paper 3: Nkulikiyinka P, Wagland S.T, Manovic V, Clough P.T. Prediction of Combined Sorbent and Catalyst Materials for SE-SMR, Using QSPR and Multitask Learning Industrial & Engineering Chemistry Research 2022 61 (26), 9218-9233. (Link to [chapter 5](#))

Contribution:

- **Paula Nkulikiyinka**: Conceptualization, Methodology, Software, Validation, Investigation, Formal analysis, Data curation, Visualization, Writing - original draft.
- Peter T. Clough: Project administration, Supervision, Visualization, Writing - review & editing.
- Stuart T. Wagland: Supervision, Writing - review & editing.
- Vasilije Manovic: Supervision, Writing - review & editing.

Although [chapter 6](#) was not a publication, the contributions were as follows:

Machine Learning for the Exploration of Metal-Organic Frameworks for Hydrogen Storage; **Paula Nkulikiyinka**^a, Stuart T. Wagland^a, Vasilije Manovic^a, Ashlee J. Howarth^b, Hudson A. Bicalho^b and Peter T. Clough^a

^a Energy and Power Theme, School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK.

^b Chemistry and Biochemistry, Concordia University, Montréal, QC H4B 1R6, Canada

Contribution:

- **Paula Nkulikiyinka** - Conceptualization, Methodology, Software, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition

- Stuart T. Wagland - Writing - Review & Editing, Supervision
- Vasilije Manovic - Writing - Review & Editing, Supervision
- Ashlee J. Howarth - Supervision, Writing - Review & Editing, Software
- Hudson Bicalho - Writing - Review & Editing
- Peter T. Clough - Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

ELSEVIER:

“Theses and dissertations which contain embedded PJAs (Published Journal Articles) as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect”
<https://www.elsevier.com/about/policies/sharing>

AMERICAN CHEMICAL SOCIETY:

“Authors may reuse all or part of the Submitted, Accepted or Published Work in a thesis or dissertation that the author writes and is required to submit to satisfy the criteria of degree-granting institutions.” Reprinted with permission from
<https://doi.org/10.1021/acs.iecr.2c00971>. Copyright 2022 American Chemical Society.

<https://pubs.acs.org/pb-assets/acspubs/Migrated/dissertation-1632927826810.pdf>

Table of Contents

Academic Integrity Declaration.....	3
Abstract	4
Acknowledgements	6
List of Publications	8
Declaration	9
List of Figures.....	15
List of Tables.....	20
List of Equations.....	23
List of Abbreviations	24
1 Introduction.....	25
1.1 Background.....	25
1.2 Project Aims and Objectives	31
1.3 Thesis Outline	32
1.4 References	35
2 General Literature Review.....	38
2.1 Sorbent Enhanced Steam Methane Reforming	38
2.2 Reactor Design and Operation	46
2.2.1 Fluidised Bed	46
2.2.2 Packed Bed.....	48
2.3 Operating Conditions and Materials.....	49
2.3.1 Temperature & Pressure	49
2.3.2 Steam to Carbon Ratio.....	54
2.3.3 Effect of Sorbent Ratio	56
2.3.4 Effect of Particle Size	60
2.3.5 Operating Conditions Conclusions	62
2.4 Sorbent and Catalyst Selection.....	62
2.4.1 Sorbent Material.....	62
2.4.2 Catalyst Material	70
2.4.3 Combined Sorbent Catalyst Material.....	71
2.5 Machine Learning	72
2.5.1 Machine Learning in SE-SMR	72
2.6 Literature Review Conclusion	79
2.7 References	80
3 Machine learning for Carbon Capture, Utilisation, And Storage (CCUS).....	97
3.1 Introduction	97
3.2 Machine Learning Algorithms	98
3.3 Machine Learning in CO ₂ Adsorption.....	101
3.3.1 Machine Learning in Calcium Looping	102
3.4 Conclusion	104

3.5 This Study	107
3.6 References	108
4 Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft Sensor Models.....	116
4.1 Introduction	117
4.2 Methodology	120
4.2.1 Process Configuration and Simulation	120
4.2.2 SE-SMR Process Model Validation	124
4.2.3 Machine Learning Model Selection	125
4.3 Machine Learning Theory and Calculations.....	126
4.3.1 Random Forest	126
4.3.2 Artificial Neural Network (ANN)	127
4.3.3 Multivariate Analysis.....	129
4.3.4 Machine Learning Model Setup.....	133
4.4 Results and Discussion.....	135
4.5 Conclusion	141
4.6 References	142
5 Prediction of Combined Sorbent Catalyst and Catalyst Materials for SE- SMR using QSPR and Multitask Learning.....	151
5.1 Introduction	152
5.2 Methodology	156
5.2.1 Database Development.....	157
5.2.2 Molecular Descriptors.....	160
5.2.3 Machine Learning Models	161
5.2.4 Single Task Learning (STL).....	162
5.2.5 Multitask Learning (MTL).....	163
5.2.6 Model Development and Performance Measurement	164
5.3 Results and Discussion.....	165
5.3.1 Comparison of the Models	165
5.3.2 Applicability Domain	174
5.3.3 Model Prediction Capability.....	177
5.3.4 Patterns in Data	179
5.4 Conclusion	187
5.5 References	189
6 Machine Learning for the Exploration of Metal-Organic Frameworks for Hydrogen Storage	200
6.1 Keywords	200
6.2 Abstract.....	200
6.3 Introduction	201
6.4 Methodology	207
6.5 Results and Discussion.....	210

6.6 Conclusion and Future Work	220
6.7 Acknowledgments.....	221
6.8 References	222
7 General Discussion	229
8 General Conclusion and Future Work	236
8.1 Final Conclusions.....	236
8.2 Future Work	242
8.2.1 Machine learning for the development of soft-sensor models	242
8.2.2 Quantitative structure–property relationship analysis for the development of a combined sorbent catalyst material	243
8.2.3 Machine learning for the Exploration of Metal-Organic Frameworks for Hydrogen Storage	243
Appendices.....	245
Appendix A - Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft Sensor Models.....	245
Appendix B - Prediction of Combined Sorbent and Catalyst Materials for SE-SMR, Using QSPR and Multitask Learning.....	249

List of Figures

Figure 1-1 Graphs as of November 2023, showing the monthly mean carbon dioxide measured at Mauna Loa Observatory, Hawaii a) over the last 5 years and b) since 1960. The red lines and symbols represent the monthly mean values, centred on the middle of each month. The black lines and symbols represent the same, after correction for the average seasonal cycle. *	26
Figure 1-2 Coal, oil, gas, cement: where do CO ₂ emissions come from	27
Figure 1-3 Global greenhouse gas emissions by sector.....	28
Figure 1-4 Structure and linkage of each project in this thesis	34
Figure 2-1 SMR process simplified from Clough 2017	39
Figure 2-2 Simplified diagram of calcination and carbonation	40
Figure 2-3 Diagram of the sequential SERP steps for the production of carbon monoxide by the reverse water-gas shift reaction.	42
Figure 2-4 Typical breakthrough curve for a completely developed SE-SMR ..	45
Figure 2-5 Schematic of a possible SE-SMR H ₂ production using dual circulating fluidised bed reactors with sorbent hydration.....	47
Figure 2-6(a) CH ₄ conversion (b) CO ₂ capture efficiency as a function of the reformer's temperature (P=1 atm, S/C molar ratio=3, CaO/C molar ratio=1)	49
Figure 2-7 Equilibrium H ₂ molar fraction (dry basis) in the product gas with and without the CaO-based acceptor as a function of temperature at fixed pressure and S/C molar ratio.....	50
Figure 2-8 Equilibrium CO ₂ pressure as a function of temperature	51
Figure 2-9 Equilibrium hydrogen and carbon monoxide concentrations as a function of temperature and pressure	52
Figure 2-10 CH ₄ conversion (a), H ₂ yield (b), H ₂ purity (c) and CO ₂ capture efficiency (d) as a function of the reformer's pressure (T=650°C, S/C=3, CaO/C ratio=1)	53
Figure 2-11 CH ₄ conversion (a), H ₂ yield (b), H ₂ purity (c) and CO ₂ capture efficiency (d) as a function of S/C ratio (T=650°C, P=1 atm, CaO/C ratio=1)	55
Figure 2-12 CH ₄ conversion and CO ₂ capture efficiency (a) and H ₂ yield and purity (b) as a function of CaO/C ratio (T=650°C, P=1atm, S/C ratio=3)	57
Figure 2-13 Methane conversion and hydrogen yield (a) 1:4 cat:sorb ratio (b) 1:1 ratio	59

Figure 2-14 Dynamic H ₂ and CH ₄ profiles on a dry basis at the reactor exit at different catalyst/sorbent ratios. Pt used for comparison.	60
Figure 2-15 Conversion curves vs. time for different particle size. Limestone: La Blanca, PCO ₂ 0.1bar, carbonation T:650°C, for 20 min; calcination T:850°C, for 15 min.(a) cycle 1; (b) cycle 20.....	61
Figure 2-16 Hydrothermal Stability Test on Hydrothermal Stability Testing Unit	64
Figure 2-17 Equilibrium CO ₂ pressure as a function of temperature for CaO, Li ₂ ZrO ₃ , Na ₂ ZrO ₃ , and Li ₄ SiO ₄ sorbents.....	66
Figure 2-18 - Representation of sintering phenomena in CaO (light grey) CaCO ₃ (dark grey)	67
Figure 2-19 Repeated calcination/carbonation cycles of limestone in a TGA...	68
Figure 3-1 Types and applications of machine learning	99
Figure 3-2 Application of the model for the evaluation of NO _x concentration if flue gas.	103
Figure 3-3 Structure of the artificial neural network used to map the thermodynamic performance of the calcium looping process retrofit.	104
Figure 4-1 Process flow diagram of the SE-SMR process simulated in Aspen Plus.....	122
Figure 4-2 Correlation heatmap of input features (Ref = Reformer, Reg = Regenerator, temp = Temperature).	131
Figure 4-3 Heatmap showing dependencies of 5 principal components on input features (Ref = Reformer, Reg = Regenerator, temp = Temperature).	132
Figure 4-4 Number of epochs vs MAE for reformer H ₂ O concentration: a) 30 epochs, b) 100 epochs.	134
Figure 4-5 Actual vs Predicted plots by the ANN algorithm for: a) reformer [H ₂ O], b) reformer [CO], c) reformer [H ₂], d) regenerator [H ₂ O], e) regenerator [CO ₂], f) reformer CH ₄ conversion reformer concentration modelled. Black lines in each figure represent y = x. Red lines indicate the line of best fit.	137
Figure 4-6 Actual vs Predicted plot by the random forest algorithm for: a) reformer [H ₂ O], b) reformer [CO], c) reformer [H ₂], d) regenerator [H ₂ O], e) regenerator [CO ₂], f) reformer CH ₄ conversion reformer concentration modelled. Black lines in each figure represent y = x. Red lines indicate the line of best fit.	138
Figure 5-1 Process flow diagram of the development of QSPR models	156
Figure 5-2 Correlation heatmap of sorbent input features.	159
Figure 5-3 Correlation heatmap of catalyst input features.	160

Figure 5-4 Schematic of a neural network configuration using different inductive transfer approaches (a) conventional single task learning (STL) and (b) multitask learning (MTL)	162
Figure 5-5 Process flow diagram showing the steps involved in unsupervised forward selection (UFS) to remove unneeded molecular descriptors	163
Figure 5-6 Comparison of machine learning models- least square support vector machine (LSSVM), deep neural network (DNN), associative neural network (ASNN) and molecular descriptors on the catalyst methane conversion database, (a) coefficient of determination (R^2), (b) root mean square error (RMSE) and (c) mean absolute error (MAE).....	167
Figure 5-7 Comparison of the effect of the inductive transfer approach and molecular descriptors, on the associative neural network (ASNN) algorithm, for the sorbent database (a) coefficient of determination (R^2), (b) root mean square error (RMSE), (c) mean absolute error (MAE).	168
Figure 5-8 Comparison of the effect of the inductive transfer approach and molecular descriptors, on the associative neural network (ASNN) algorithm, on the catalyst methane conversion database, (a) coefficient of determination (R^2), (b) root mean square error (RMSE) and (c) mean absolute error (MAE).	169
Figure 5-9 Williams plots for the model when applied to predict unseen CSCM molecules, with ASNN-Standard Deviation used as the distance to the model. Black lines represent the averaged RMSE over different distance to model intervals (a) Last cycle CO_2 capacity (g CO_2 /g Sorbent), and (b) methane Conversion (%).	176
Figure 5-10 Comparison of the percentage of data belonging to each average absolute relative error (AARE) range for the last cycle capacity and methane conversion prediction of CSCM molecules.	177
Figure 5-11 Actual vs Predicted scatter plots showing the effect of an influential input for (a) Last cycle capacity (g CO_2 /gSorbent) as a function of CaO concentration (%), and (b) methane conversion (%) as a function of nickel wt%.....	179
Figure 5-12 Actual vs Predicted scatter plot showing the effect of the calcium precursor on the Last cycle capacity (g CO_2 /gSorbent)	181
Figure 5-13 Actual vs Predicted scatter plot showing the effect of the inert support on the Last cycle capacity (g CO_2 /gSorbent)	181
Figure 5-14 Actual vs Predicted scatter plot showing the effect of the synthesis method on the Last cycle capacity (g CO_2 /gSorbent)	182
Figure 5-15 Actual vs Predicted scatter plots showing the effect of steam to carbon ratio on methane conversion (%)	184

Figure 5-16 Actual vs Predicted scatter plots showing the effect of catalyst calcination temperature ($^{\circ}\text{C}$) on methane conversion (%)	185
Figure 5-17 Actual vs Predicted scatter plots showing the effect of catalyst calcination time (a) maximum 100 hours, (b) 50 hours on methane conversion (%).....	186
Figure 6-1 Total hydrogen uptake for top candidates with the highest working capacity oof MOF-210, ZIF-8, and hypothetical MOF-5059389.....	205
Figure 6-2 Methodological approach for this study.....	207
Figure 6-3 Heatmap showing dependencies of 7 principal components on input features.....	209
Figure 6-4 Scatter plots showing correlation between the input features and H_2 capacity at TPS for a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%). Black lines represent $y = x$. Red lines indicate the line of best fit.	210
Figure 6-5 Actual vs Predicted plots using simple MLR using single crystal density (g/cm^3) to predict a) usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%)......	212
Figure 6-6 Actual vs Predicted plots using simple MLR using void fraction to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).	212
Figure 6-7 Actual vs Predicted plots using simple MLR using pore volume (cm^3/g) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%)......	213
Figure 6-8 Actual vs Predicted plots using simple MLR using largest cavity diameter ($\text{\AA}$) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).	213
Figure 6-9 Actual vs Predicted plots using simple MLR using pore limiting diameter ($\text{\AA}$) to predict a) Usable volumetric capacity ($\text{g H}_2/\text{L MOF}$) and b) usable gravimetric capacity (wt.%).	214
Figure 6-10 Actual vs Predicted plots using simple MLR using gravimetric surface area (m^2/g) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).	214
Figure 6-11 Actual vs Predicted plots using simple MLR using volumetric surface area (m^2/cm^3) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).	215
Figure 7-1 Actual vs Predicted plots by the ANN algorithm for: a) reformer [H_2O], b) reformer [CO], c) reformer [H_2], d) regenerator [H_2O], e) regenerator [CO_2],	

f) reformer CH ₄ conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.	230
Figure 7-2 Actual vs Predicted plot by the random forest algorithm for: a) reformer [H ₂ O], b) reformer [CO], c) reformer [H ₂], d) regenerator [H ₂ O], e) regenerator [CO ₂], f) reformer CH ₄ conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.	230
Figure 7-3 Comparison of the percentage of data belonging to each average absolute relative error (AARE) range for the last cycle capacity and methane conversion prediction of CSCM molecules.	232
Figure 8-1 Sizing and design schematic of the reactor.....	239
Figure 8-2 Process flow diagram of the proposed hydrogen production process	240
Figure 8-3 Constructed rig consisting of gas analysers on the left, mass flow controllers, a quartz reactor within a Carbolite furnace, joined via Swagelok fittings, and various other components.	240
Figure 8-4 Constructed rig, with software programming and process monitoring station on the right.	241
Figure_Apx A-1 Random Forest Model- close up annotation.....	246
Figure_Apx A-2 Epoch vs MAE for reformer [H ₂ O], using different activation functions: a) ReLU, b) hyperbolic tangent (tanh), c) Sigmoid, d) Combination of Sigmoid, ReLU, Sigmoid for each consecutive layer.	246
Figure_Apx B-1 Actual vs Predicted scatter plot showing the effect of the cycle number on the Last cycle capacity (g/g)	274
Figure_Apx B-2 Actual vs Predicted scatter plot showing the effect of the (a) carbonation time and (b) calcination time on the Last cycle capacity (g/g)	275
Figure_Apx B-3 Actual vs Predicted scatter plot showing the effect of the (a) carbonation temperature and (b) calcination temperature on the Last cycle capacity (g/g)	276
Figure_Apx B-4 Actual vs Predicted scatter plot showing the effect of the (a) SMR reformer temperature (°C) and (b) BET Surface Area (m ² /g) on the methane conversion (%).	277

List of Tables

Table 2-1 Best performance of SERP from Carvill et al. 1996	43
Table 2-2 Best performance of SERP from Hufton, Mayorga & Sircar, 1999 ...	43
Table 2-3 Thermodynamic equilibrium data	58
Table 2-4 Optimal operating conditions for sorption enhanced steam methane reforming.	62
Table 2-5 High performing sorbents in SE-SMR; *Capacity loss after 1 day of steam (10 atm H ₂ O/ 0.3 atm CO ₂) (Binary Desorption Unit); ** Air Products proprietary pelletised K ₂ CO ₃ -promoted hydrotalcite based material).....	64
Table 2-6 Comparison of natural and synthetic sorbents	65
Table 2-7 Properties and their respective experimental conditions within a CSCM	72
Table 2-8 The values of the model evaluating factors.	75
Table 3-1 Common machine learning algorithms	100
Table 4-1 Aspen Plus SE-SMR baseline process model parameters.....	121
Table 4-2 Parameter ranges used in the sensitivity analysis to build up the machine learning database. *In order to obtain a ratio, the flowrates for the steam, methane or sorbent, were altered between the given ranges, and a ratio was then calculated.	124
Table 4-3 Comparison of literature data against Aspen Plus data for validation purposes.....	124
Table 4-4 Random Forest feature importance for each output parameter. (Ref = Reformer, Reg = Regenerator, T = Temperature, P = Pressure).....	130
Table 4-5 Neural network vs random forest soft-sensor concentration predictions	136
Table 5-1 Properties and their respective experimental conditions per database.	158
Table 5-2 Descriptors used for comparison.....	161
Table 5-3 A comparison of R ² and R ² _{adj} for the ASNN STL and MTL models using GSfrag descriptors	173
Table 5-4 Properties of the final ML model used for prediction of unseen combined sorbent catalyst materials (CSCM).....	174
Table 5-5 Key for the synthesis method	183

Table 5-6 Data gathered from the colour plots, indicating parameters that lead to a well/bad performing CSCM.....	187
Table 6-1 Performance of ML models for the prediction of usable H ₂ storage capacities in MOFs under temperature & pressure swing (TPS) conditions, where d = Single crystal density (g/cm ³); pv = Pore volume (cm ³ /g); vf = void fraction; gsa = Gravimetric surface area (m ² /g); vsa = Volumetric surface area (m ² /cm ³); lcd = Largest cavity diameter (Å); pld = Pore limiting diameter (Å). Performance from [18] is ranked based on R ² , where green is the parameters identified as influential that match that of the original database study.....	216
Table 6-2 The ranges obtained from the colour plots, indicating the properties that a high performing MOF would have.....	217
Table 6-3 Top performing MOFs for Usable gravimetric capacity in weight percent (wt.%), where d = Single crystal density (g/cm ³); pv = Pore volume (cm ³ /g); vf = void fraction; gsa = Gravimetric surface area (m ² /g); vsa = Volumetric surface area (m ² /cm ³); lcd = Largest cavity diameter (Å); pld = Pore limiting diameter (Å), UG = Usable gravimetric capacity (wt%), and UV = Usable volumetric capacity (gH ₂ /LMOF).....	219
Table 6-4 Properties of previously reported high performing MOFs compared to the top performing hypothetical MOF identified in this study, where d = Single crystal density (g/cm ³); pv = Pore volume (cm ³ /g); vf = void fraction; gsa = Gravimetric surface area (m ² /g); vsa = Volumetric surface area (m ² /cm ³); lcd = Largest cavity diameter (Å); pld = Pore limiting diameter (Å), UG = Usable gravimetric capacity (wt%), and UV = Usable volumetric capacity (g H ₂ /L MOF). Blue is the UV and UG for the MOF identified in this study, and the green and red indicates better or worse performance from different studies, respectively.....	220
Table 7-1 Data gathered from the colour plots, indicating parameters that lead to a well/bad performing CSCM.....	233
Table_Apx A-1 Stream compositions under baseline conditions.....	245
Table_Apx A-2 Random Forest Feature Importance.....	247
Table_Apx A-3 Details of Activation functions.....	248
Table_Apx B-1 OECD guidelines for the development of a QSPR.....	249
Table_Apx B-2 PCA Variance	250
Table_Apx B-3 Feature importance.....	251
Table_Apx B-4 Sorbent Database- Property- Last cycle CO ₂ capacity (g/g) ..	252

Table_Apx B-5 Catalyst Database- Property- Methane Conversion (%)	261
Table_Apx B-6 Architecture of the ASNN model	269
Table_Apx B-7 Architecture of the DNN model	269
Table_Apx B-8 Architecture of the LSSVM model.....	269
Table_Apx B-9 CSCM database	270
Table_Apx B-10 Measured vs Predicted CSCM data using GSfrag/ASNN model, with ARE%.....	272

List of Equations

Equation (2-1) – Steam methane reforming reaction.....	38
Equation (2-2) – Water gas shift reaction	38
Equation (2-3) – Overall steam methane reforming reaction	39
Equation (2-4) – Carbonation and Calcination reaction.....	40
Equation (2-5) – Overall sorbent enhanced steam methane reforming	41
Equation (4-1) – Soft sensor prediction formula	119
Equation (4-2) – Average percentage error (%)	125
Equation (4-3) – CH ₄ conversion (%).....	125
Equation (4-4) – Random forest predicition formula.....	126
Equation (4-5) – Root mean square error (RMSE)	128
Equation (4-6) – Mean absolute error (MAE).....	128
Equation (5-1) – Coefficient of correlation formula	165
Equation (5-2) – Absolute relative error ARE)	165
Equation (5-3) – Average absolute relative error (AARE	165
Equation (5-4) – Adjusted coefficient of determination (R^2_{adj}).....	173
Equation (6-1) – Langmuir adsorption model	204

List of Abbreviations

AD	Applicability Domain
AI	Artificial Intelligence
ASNN	Associative Neural Network
CaL	Calcium Looping
CSCM	Combined Sorbent Catalyst Material
DFT	Density Functional Theory
GCMC	Grand Canonical Monte Carlo
GHG	Greenhouse Gases
GSA	Gravimetric Surface Area
LCD	Lowest Cavity Diameter
MAE	Mean Absolute Error
ML	Machine Learning
MLR	Multiple Linear Regression
MOF	Metal Organic Framework
MSE	Mean Square Error
MTL	Multi-Task Learning
PC	Principal Component
PCA	Principal Component Analysis
PLD	Pore Limiting Diameter
PV	Pore Volume
QSPR	Quantitative Structure–Property Relationship
RF	Random Forest
RMSE	Root Mean Square Error
S/C	Steam to Carbon Ratio
SE-SMR	Sorption-Enhanced Steam Methane Reforming
SMILES	Molecular-Input Line-Entry System
SMR	Steam Methane Reforming
STL	Single-Task Learning
TPS	Temperature and Pressure Swing
UG	Usable Gravimetric capacity
UV	Usable Volumetric capacity
VF	Void Fraction
VSA	Volumetric Surface Area
WGS	Water Gas Shift

1 Introduction

1.1 Background

The First and Second Industrial Revolutions brought many positive changes to human life including technological innovations in processes and development of modern medicine to name a few. However, as we now know, these came at a great cost. As the understanding of environmental chemistry has progressed, the burning of fossil fuels as an energy source has been led to a disproportionate amount of atmospheric greenhouse gases (GHG) particularly carbon dioxide (CO₂), and thus a disruption in the global carbon cycle, resulting in an increase in the average global temperature on Earth. The effects of global warming and climate change range from ecological, physical and health impacts, including extreme weather events rising sea-levels, crop growth changes, and disrupted water systems [1].

The goal set by The Paris Agreement was to restrict the global mean temperature increase to 2°C by the end of the 21st Century relative to the pre-industrial period, however from recent studies, projections for the increase could very well be as severe as 4°C [2]. This higher temperature increase, stems from global climate model projections, assessed by the Intergovernmental Panel on Climate Change (IPCC), which identify that a high-end warming of 4°C or more by the 2090s is likely, with the North African, Southern Europe and Central Asian regions possibly experiencing a 6-8°C increase [3]. Whilst these projections are extreme, they could unfortunately become the harsh reality if fossil fuel combustion doesn't ease at the current rate.

Despite vast changes being implemented to try to reduce anthropogenic greenhouse gas emissions, record high atmospheric CO₂ emissions were measured at The Mauna Loa Observatory in Hawaii, with a reading of 423.78 parts per million (ppm) in May 2023, as shown in Figure 1-1, the highest in over 3 million years [4]. The red line trend is due to measurements being highest in May of every year, the month when CO₂ peaks in the Northern Hemisphere, as a direct result of human activity. This amount shows that efforts to reduce

emissions is currently not effective enough in order to reach emission reduction goals, as the growth-rate has been steadily increasing, reaching 1.81 ± 0.26 ppm/y (mean $\pm$ 2 std dev) at the beginning of 2023. This means that currently, the concentration of carbon dioxide is growing by about 1.81 ppm per year [5].

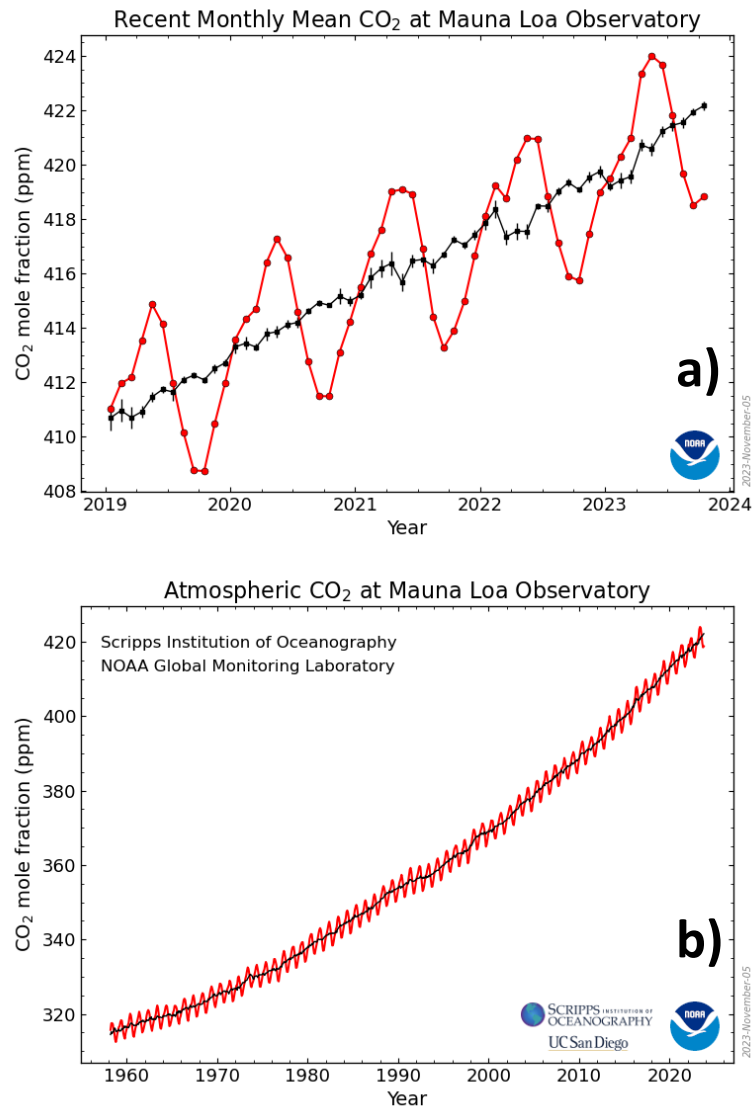


Figure 1-1 Graphs as of November 2023, showing the monthly mean carbon dioxide measured at Mauna Loa Observatory, Hawaii a) over the last 5 years and b) since 1960. The red lines and symbols represent the monthly mean values, centred on the middle of each month. The black lines and symbols represent the same, after correction for the average seasonal cycle. *

*It should be noted that due to the eruption of the Mauna Loa Volcano, measurements from Mauna Loa Observatory were suspended as of Nov. 29, 2022. Observations from December 2022 to July 4, 2023, were gathered from a site at the Maunakea Observatories, approximately 21 miles north of the Mauna Loa Observatory. Mauna Loa observations resumed in July 2023.

An important part of analysis of climate change is the allocation of CO₂ emissions, in order to assign resources or action to mitigate these emissions. Figure 1-2 [6] and Figure 1-3 [7], give data from the organisation 'Our World in Data', which is an online publication that presents research and data on global change, based at The University of Oxford. The results in these charts show that emissions by both fuel type and sector. From both graphs the clear message is that the main contributors to CO₂ emissions is energy production via fossil fuel combustion. Therefore, a simple conclusion would be that by tackling the use natural resources in energy production, the issue of CO₂ emissions would greatly be reduced. A vital step in doing this, is shifting from fossil fuels to low carbon technology, via the process of carbon capture and storage (CCS).

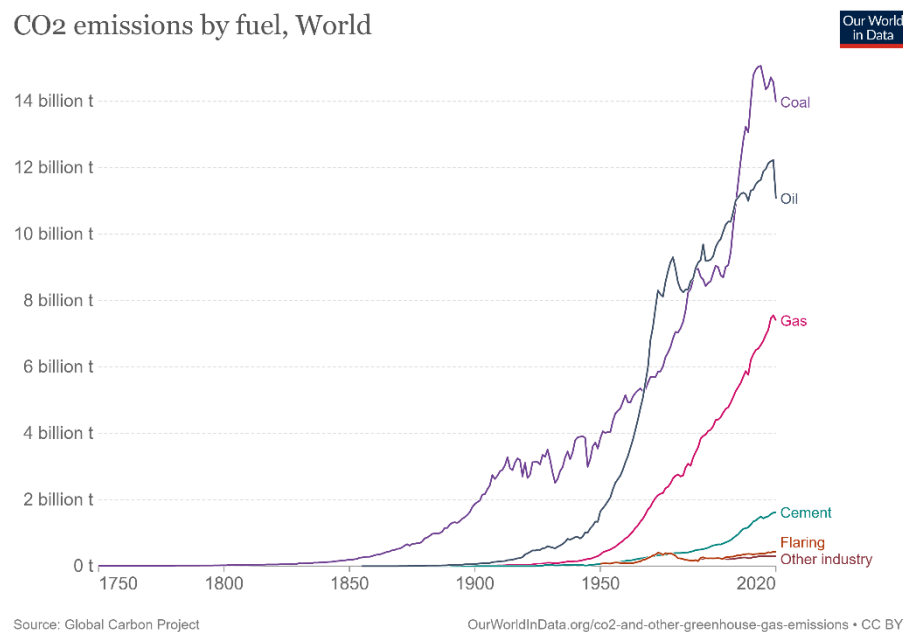
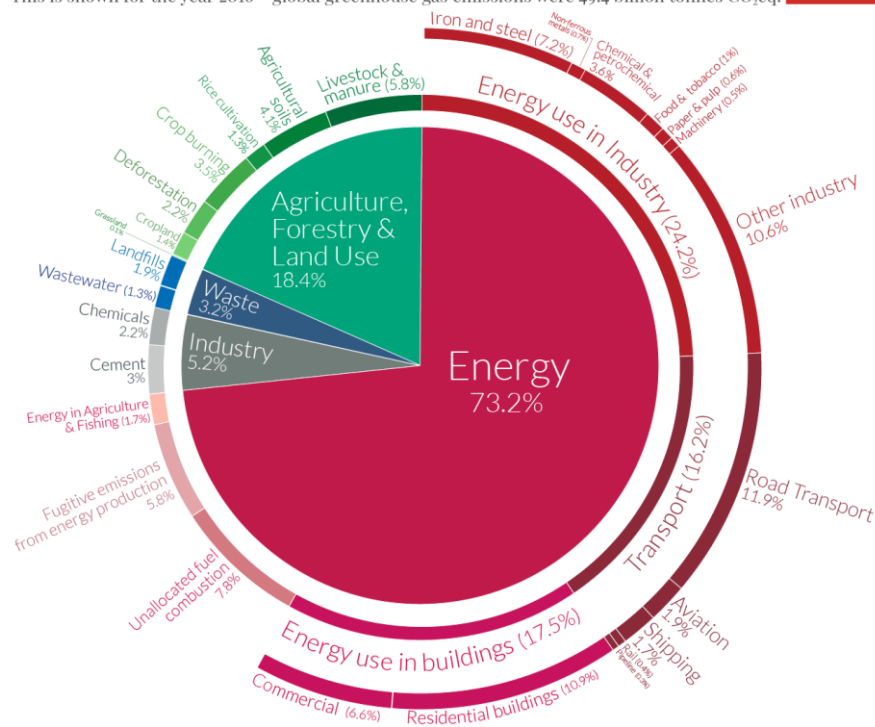


Figure 1-2 Coal, oil, gas, cement: where do CO₂ emissions come from

Global greenhouse gas emissions by sector

This is shown for the year 2016 – global greenhouse gas emissions were 49.4 billion tonnes CO₂eq.

Our World
in Data



OurWorldinData.org – Research and data to make progress against the world's largest problems.
Source: Climate Watch, the World Resources Institute (2020). Licensed under CC-BY by the author Hannah Ritchie (2020).

Figure 1-3 Global greenhouse gas emissions by sector

Carbon capture and storage technologies are the most promising route to reduce CO₂ emissions that have the potential to meet the climate change goals of the future. The three leading CO₂ capture methods are:

- Post-combustion capture
- Pre-Combustion capture
- And Oxy-fuel combustion capture.

The **post-combustion capture** process is usually amine based and the general idea is that CO₂ is separated from a CO₂ rich gas by absorption, using an aqueous amine solution, after the combustion reaction has occurred [8]. Once the CO₂ is removed by the amine solvent, the solvent is then regenerated to be re-used, then cooled and recycled continuously. The CO₂ removed from the solvent in the regeneration process is dried, compressed, and stored [9]. The main advantage

of this technology is ability for it to be applied to existing large, fixed-point sources of CO₂ for example to conventional power generation systems, such as coal fired power stations [10].

In the **pre-combustion storage** process, the fuel is converted into a syngas mixture mainly consisting of carbon monoxide and H₂, by gasification or reforming, prior to combustion. The syngas mixture is then reacted with steam to form carbon dioxide and H₂, through the 'water gas shift' (WGS) reaction. The products are then processed via desulphurisation and the CO₂ is removed, leaving an H₂ fuel stream which can be used for various applications, e.g., fuel cell, gas turbine etc. [11]. There are certain thermodynamic advantages to this technology, which lead to a lower-than-average efficiency penalty on a power station.

In **oxy-fuel combustion** processes, the fuel is combusted with nearly pure oxygen, instead of air, and CO₂. The flue gas produced consists mainly of CO₂ and H₂O, in which, the H₂O is readily removable by condensation, leaving a pure CO₂ product, to be compressed and stored [8].

Each of these technologies have their advantages and downsides, for example, although the post-combustion solvent scrubbing is the closest to market and most mature, the disadvantages include solvent poisoning, solvent degradation, and high solvent costs [12]. Oxy-fuel combustion is also fairly established and reasonably close to the market, with a Technology Readiness Level (TRL) of 7–8 for carbon capture from power generation and industry [13], however this process requires an air separation unit (ASU) in order to provide the large quantities of O₂ demanded, resulting in it being a high cost and energy-intensive process [14].

Due to these challenges, the motivation is to explore other viable options, in the aim to aid the acceleration of CCS technology to commercial scale. This study aims to do that via the application of artificial intelligence or more specifically machine learning, in the process of a pre-combustion technology, sorption enhanced steam methane reforming (SE-SMR).

The process of SE-SMR is the combination of calcium looping (CaL) and steam methane reforming (SMR), which involves an in-situ CO₂ capture process where methane is reacted with steam in the presence of a CO₂ sorbent (that is calcium based) and a reforming catalyst to generate decarbonised, high purity H₂ [15]. Advantages include the use of limestone (CaCO₃), which is a cheap, environmentally friendly, and widely available material. It can also be utilised at a wide range of temperature window and can keep CO₂ reliably locked away in a very stable form [16]. On the other hand, there are shortfalls in CaO as an adsorbent, including rapid loss of activity during the capture cycles, as a result of sintering, attrition and elutriation and additionally high operational conditions, such as higher temperatures and fewer carbonation and calcination cycles, prior to full capacity [17].

The production side of clean fuel is not the only challenge at hand to tackle the ever-increasing greenhouse gas emissions. Once carbon abated hydrogen is produced, a further obstacle is the ability to utilize hydrogen to its highest potential when used as an energy source, for example in vehicle fuel tank. A high-capacity material is required to store the gas, which can be difficult due to the molecular properties. Other than associated costs and practicality, the main issue is the amount of hydrogen that can safely be stored in a vehicle and delivered as required under operational conditions [18].

To overcome both the production side and end use challenges faced hydrogen from SE-SMR, the introduction of machine learning into this process will allow for accelerated streamlined screening of optimal materials and process conditions.

With recent government policies being implemented all over the world to achieve a net zero target by 2050, action is required now in the early 2020s to slow down the global warming increase to 1.5°C, thus reaching a stable global temperature [19]. In order to do this, the application of machine learning is increasingly being used in all industries as a means to mitigate associated costs and time-consuming methods such as practical experiments, in the physical chemistry field.

The process of SE-SMR is well-researched, however there are still areas unexplored and the complexity of H₂ production technology still poses some challenging tasks. Because of this, the use of simple machine learning models is of practical and theoretical significance, as it allows for the field to expand the wider knowledge of under-researched areas.

In the research of chemical processes such as steam methane reforming, the data and most of the results discussed in literature is normally found via expensive and/or difficult experiments and measurements on real projects. Additionally, high costs, time-consuming and other variable limitations make some data/results sometimes insufficient methods of data collection. A possible alternative solution to this is via machine learning, mathematical modelling and computational analysis, as a method of data collection, analysis and interpretation, thus contributing to a novel aspect of the SE-SMR process.

1.2 Project Aims and Objectives

The overall scientific aim of this thesis is to enhance the body of knowledge of the sorbent enhanced steam methane reforming process through the application of machine learning algorithms and computational methods that will be used to investigate and optimise the production and utilisation of blue hydrogen.

In order to achieve this, there are three linked aims involved in this research, described below:

Aim 1: Demonstrate the application of machine learning algorithms for the prediction of the SE-SMR process:

Objective 1a: Generate a representative simulation model of the SE-SMR process.

Objective 1b: Assess the performance of machine learning algorithms to predict SE-SMR outputs, with an accuracy relative to training data above 95%

Aim 2: Investigate the concept of combined sorbent catalyst materials (CSCM) for SE-SMR

Objective 2a: Conduct a detailed review of developments of sorbent and catalyst materials used for calcium looping and methane reforming, respectively.

Objective 2b: Develop a CSCM database and compare and examine the differences in performance of a bifunctional material and a two-particle system.

Objective 2c: Assess the effect of implementing quantitative structure property relationship analysis for the prediction of CSCM behaviour.

Objective 2d: Interpret the properties for the optimal CSCM performance from analytical methods.

Aim 3: Assess the use of machine learning for the discovery of Metal Organic Frameworks (MOFs) for H₂ storage.

Objective 3a: Assess the properties and suitability of MOFs for H₂ storage.

Objective 3b: Develop ML models for the prediction of H₂ storage capacity.

Objective 3c: Analyse and interpret data from ML models for the optimal performance of a H₂ storage MOF as a measure of their usable capacities.

1.3 Thesis Outline

In order to achieve the objectives proposed for this PhD project, **two first author** journal publications, **one second author** publication, and one further chapter have been completed to fulfil the intellectual contribution to the scientific body of knowledge.

The structure of this doctoral thesis with a brief description of the content of each chapter is detailed below with the linkage displayed in Figure 1-4:

- **Chapter 1- Introduction-** This chapter presents the background, aims and objectives of this project, as well as an overview of the novelty and linkages between publications and project outputs.
- **Chapter 2- General Literature Review** - This chapter is the general literature review on CCS, SE-SMR and machine learning in CCS- additionally on QSPR use.
- **Chapter 3- Machine Learning for Carbon Capture, Utilisation and Storage (CCUS)** - This chapter contains excerpts of my contribution to the published work (DOI: <https://doi.org/10.1039/D1EE02395K>) on ‘Harnessing the power of machine learning for carbon capture, utilisation, and storage (CCUS) – a state-of-the-art review’
- **Chapter 4- Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft Sensor Models** - This chapter is published work (DOI: <https://doi.org/10.1016/j.egyai.2020.100037>) on the investigation of using soft-sensor models developed from the use of artificial neural networks and random forest models to predict gas concentrations in the reformer and regenerator reactors of the SE-SMR process using temperature, pressure, steam to carbon ratio and sorbent to carbon ratio as input process features.
- **Chapter 5- Prediction of Combined Sorbent Catalyst and Catalyst Materials for SE-SMR using QSPR and Multitask Learning** – This chapter is published work (DOI: <https://doi.org/10.1021/acs.iecr.2c00971>). This is on the exploration of using QSPR to predict the behaviour of combined sorbent catalyst materials, based on data mining.
- **Chapter 6- Machine Learning for the Discovery of Metal-Organic Frameworks for Hydrogen Storage Applications** – This chapter is the short communication on the research that was carried out during the Doctoral exchange programme at Concordia University, Montreal, and is focused on using ML to find a MOF for H₂ storage.

- **Chapter 8- General Discussion** – This chapter presents a general discussion of all the work done through the course of this research linking all the chapters together.
- **Chapter 9- General Conclusion and Future Work** – This chapter provides a summary of the conclusions, discusses the contributions to knowledge and recommendations for future research.

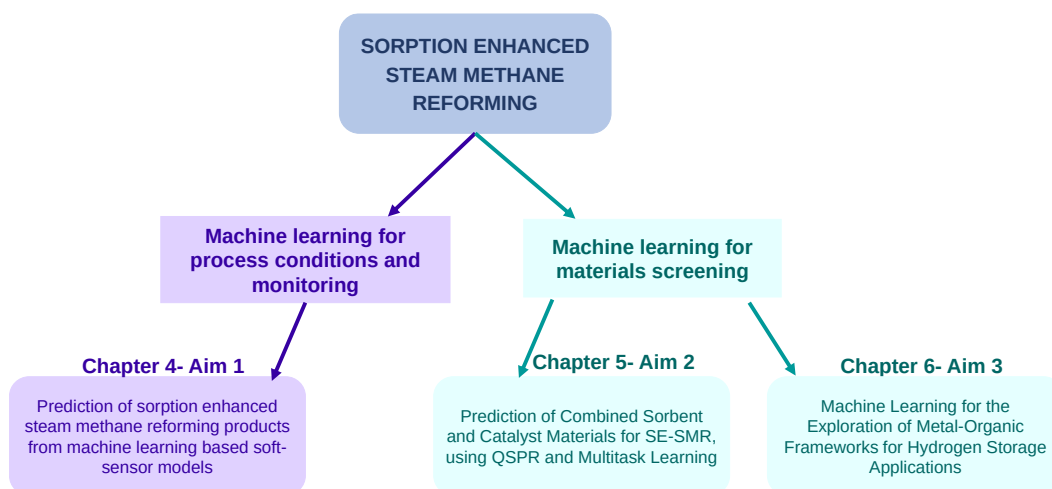


Figure 1-4 Structure and linkage of each project in this thesis

Overall, this thesis explores the application of different levels of computational modelling to the SE-SMR process, to accelerate the development process of optimal sorbent and catalyst materials. This is an area of significant importance as it helps tackle the race to obtain the net zero target. Using computational modelling tools is a time and cost-efficient solution to lead to the discovery of optimal materials to aid in the scale up of carbon-abated hydrogen production.

In terms of the three aims, the novelty applied within this thesis addresses the

- Limited work to date on the process of methane conversion in the SE-SMR process via a surrogate model
- Development of a novel predictive method for optimal materials and synthesis steps for CSCM

- And the identification of new materials for hydrogen storage, in an emerging scientific field. Highlighting a clear contribution to knowledge.

1.4 References

- [1] Ritchie H., Roser M. CO₂ and other Greenhouse Gas Emissions. OurWorldInData.org. 2015; Available at: DOI:10.1034/j.1399-0004.2002.610606.x
- [2] Wang X., Jiang D., Lang X. Future extreme climate changes linked to global warming intensity. Science Bulletin. Science China Press; 2017; 62(24): 1673–1680. Available at: DOI:10.1016/j.scib.2017.11.004
- [3] Sanderson MG., Hemming DL., Betts RA. Regional temperature and precipitation changes under warming. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2011; 369(1934): 85–98. Available at: DOI:10.1098/rsta.2010.0283
- [4] Earth System Research Laboratory., NOAA. Trends in Atmospheric Carbon Dioxide. Global Monitoring Laboratory. 2022.
- [5] Rasmussen CE. Atmospheric Carbon Dioxide Growth Rate. 2022. Available at: <http://mlg.eng.cam.ac.uk/carl/words/carbon.html> (Accessed: 26 July 2022)
- [6] Zhou P., Wang M. Carbon dioxide emissions allocation: A review. Ecological Economics. 2016. Available at: DOI:10.1016/j.ecolecon.2016.03.001
- [7] Ritchie H., Roser M., Rosado P. CO₂ and Greenhouse Gas Emissions. OurWorldInData.org. 2020.
- [8] MacDowell N., Florin N., Buchard A., Hallett J., Galindo A., Jackson G., et al. An overview of CO₂ capture technologies. 2010; : 1645–1669. Available at: DOI:10.1039/c004106h

- [9] Gibbins J., Chalmers H. Carbon capture and storage. *Energy Policy*. December 2008; 36(12): 4317–4322. Available at: DOI:10.1016/j.enpol.2008.09.058
- [10] Hanak DP., Michalski S., Manovic V. From post-combustion carbon capture to sorption-enhanced hydrogen production: A state-of-the-art review of carbonate looping process feasibility. *Energy Conversion and Management*. Elsevier; 2018; 177(October): 428–452. Available at: DOI:10.1016/j.enconman.2018.09.058
- [11] Theo WL., Lim JS., Hashim H., Mustaffa AA., Ho WS. Review of pre-combustion capture and ionic liquid in carbon capture and storage. *Applied Energy*. Elsevier Ltd; 2016; 183: 1633–1663. Available at: DOI:10.1016/j.apenergy.2016.09.103
- [12] Hanak DP., Biliyok C., Anthony EJ., Manovic V. Modelling and comparison of calcium looping and chemical solvent scrubbing retrofits for CO₂ capture from coal-fired power plant. *International Journal of Greenhouse Gas Control*. Elsevier Ltd; 1 November 2015; 42: 226–236. Available at: DOI:10.1016/j.ijggc.2015.08.003
- [13] Yan Y., Clough PT., Anthony EJ. Investigation of the apparent kinetics of air and oxy-fuel biomass combustion in a spout fluidised-bed reactor. *Chemical Engineering Research and Design*. Institution of Chemical Engineers; 1 January 2020; 153: 276–283. Available at: DOI:10.1016/j.cherd.2019.10.043
- [14] Yan Y., Thanganadar D., Clough PT., Mukherjee S., Patchigolla K., Manovic V., et al. Process simulations of blue hydrogen production by upgraded sorption enhanced steam methane reforming (SE-SMR) processes. *Energy Conversion and Management*. Elsevier Ltd; 15 October 2020; 222. Available at: DOI:10.1016/j.enconman.2020.113144
- [15] Clough PT., Boot-Handford ME., Zheng L., Zhang Z., Fennell PS. Hydrogen production by sorption enhanced steam reforming (SESR) of biomass

in a fluidised-bed reactor using combined multifunctional particles. *Materials*. MDPI AG; 21 May 2018; 11(5). Available at: DOI:10.3390/ma11050859

[16] Sun H., Wu C., Shen B., Zhang X., Zhang Y., Huang J. Progress in the development and application of CaO-based adsorbents for CO₂ capture—a review. *Materials Today Sustainability*. Elsevier Ltd; 2018. pp. 1–27. Available at: DOI:10.1016/j.mtsust.2018.08.001

[17] Manovic V., Anthony EJ. Lime-based sorbents for high-temperature CO₂ capture-a review of sorbent modification methods. *International Journal of Environmental Research and Public Health*. 2010; 7(8): 3129–3140. Available at: DOI:10.3390/ijerph7083129

[18] Thomas KM. Adsorption and desorption of hydrogen on metal-organic framework materials for storage applications: Comparison with other nanoporous materials. *Dalton Transactions*. 2009; (9): 1487–1505. Available at: DOI:10.1039/b815583f

[19] IPCC. The evidence is clear: the time for action is now. We can halve emissions by 2030. 2022.

2 General Literature Review

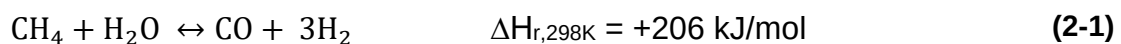
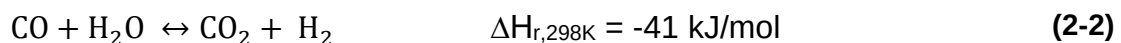
2.1 Sorbent Enhanced Steam Methane Reforming

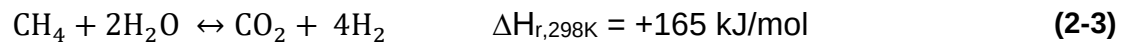
Aside from the clear imperative need to reduce energy production from fossil fuel combustion, there are many reasons why hydrogen as an energy source is at the forefront of energy and power production. For instance, it can be produced in several ways and from numerous resources- both renewable and non-renewable. In addition, it's a clean energy source producing no greenhouse gases or other particulates, and it is also a raw material for many other eco-friendly and/or high efficiency power systems for example fuel cells, for both vehicular transportation and distributed electricity generation [1].

As mentioned above, hydrogen can be produced from various resources be it renewable (hydro, wind, solar, biomass, geothermal) or non-renewable (natural gas, gasoline, nuclear or coal). The most viable and thoroughly researched method is production via steam methane reforming (SMR). In this process, methane is reacted with steam producing carbon monoxide and H₂, then the carbon monoxide is processed by the WGS reaction, with the overall SMR reactions producing H₂ and carbon dioxide [2].

H₂ production via SMR has been a topic of interest since the late 1980s, with a highly esteemed study on SMR and WGS intrinsic kinetics conducted by Xu & Froment in 1989. Throughout the years, as technology has advanced, the kinetic parameters derived from this research, using a nickel, magnesium, and aluminium based catalyst, remains the most cited and reputable work in this area of study.

The reactions that take place in order to produce H₂ from methane are reforming (reaction **(2-1)**) and the WGS Reaction (reaction **(2-2)**), giving the overall SMR reaction (reaction **(2-3)**):





The reforming reaction is endothermic and occurs out at high-temperatures; the methane enters the reactor at approximately 650°C and then heated up to ~850 - 950°C. The gas leaving the reformer is typically 50% pure H₂, therefore the gas is further reacted over iron oxide at lower temperatures under excess steam to undergo the water-gas-shift reaction. The exiting gas stream from WGS, which contains amounts of methane, CO and trace CO₂ is passed through a pressure swing adsorption unit which gives a product stream of very high purity hydrogen (99.9 % pure) and a separate waste gas stream, which can be recycled back into the reformer [2].

A simplified SMR process flow diagram is shown in Figure 2-1 below [2]:

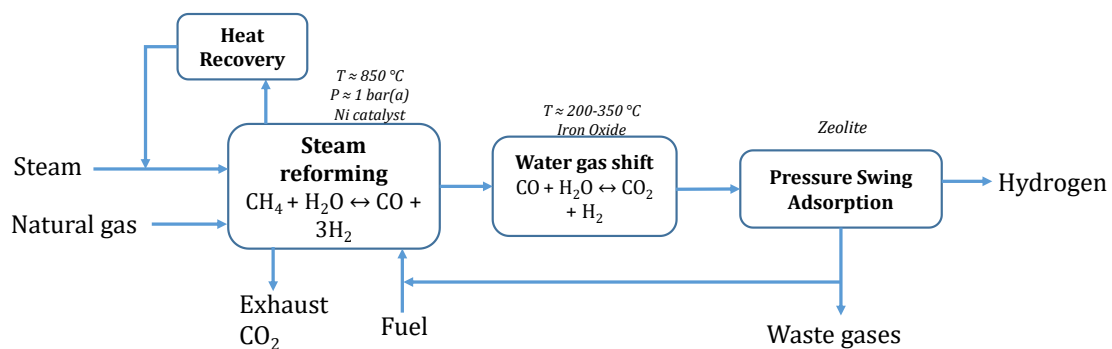


Figure 2-1 SMR process simplified from Clough 2017

Studies on this process has varied from the initial reaction kinetics and rate parameters to optimal operating conditions such as temperature, catalyst material as well as reactor type. Throughout this research, over the last few decades many of these factors have been founded and proven experimentally and theoretically, and there is a general consensus on the ideal conditions of SMR as it is now a well-established process. More recent research on SMR has focused on implementing other processes either before or after the reaction, with a novel example of the addition of a thermochemical energy storage reactor [3], or rather a more distinguished process such as calcium looping.

An additional CCS method which is relatively new compared to the previous three, is Calcium Looping (CaL) combustion, which can be applied to pre or post combustion capture systems. This process is a cyclic reversible reaction between calcium oxide and carbon dioxide, in which the sorbent (calcium oxide) acts as an ‘in-situ’ CO₂ absorbent, resulting in a separate & pure CO₂ product from a flue gas stream.

This process is fairly similar to the conventional pre-combustion amine-based absorption process mentioned above, with calcium oxide acting as the adsorbent. The CO₂ capture occurs first, in what’s known as the carbonator, then the “saturated” sorbent material, CaCO₃, is passed to a second reactor, known as the calciner, where the sorbent material is regenerated and recycled, whilst producing a pure stream of CO₂, as shown in Figure 2-2 [4].

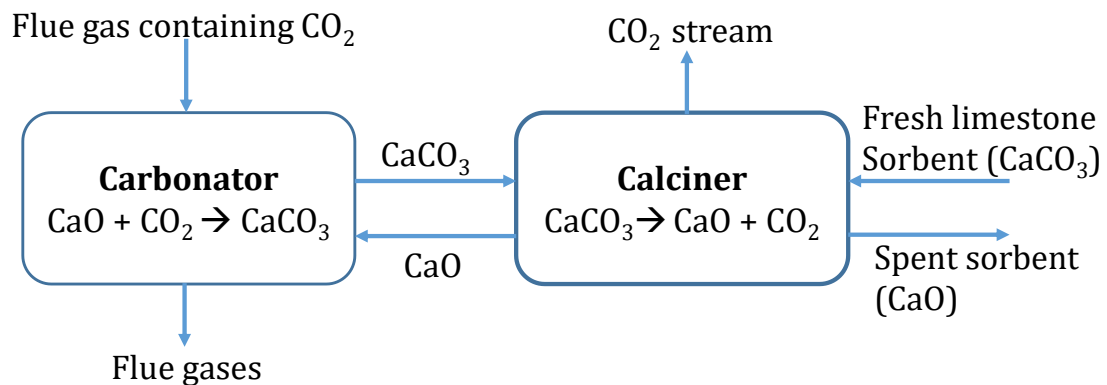


Figure 2-2 Simplified diagram of calcination and carbonation

The principal reactions in CaL are carbonation (forward reaction) and calcination (backwards reaction) described below:



The forward **carbonation** reaction is exothermic, typically carried out at 650°C. It has a fast-initial reaction rate, followed by an immediate slow reaction rate. During the backward **calcination** reaction, solid calcium carbonate is fed into a calciner, where it is heated to around 850-950 °C which allows it to thermally

decompose into gaseous carbon dioxide and solid calcium oxide CaO. Calcination is endothermic and occurs rapidly within a single step.

The operation conditions are what set it aside as a promising process in CCS technology. Thermodynamically, the reaction can occur at lower temperature (approx. 750°C), however, in order to achieve a pure CO₂ stream, higher temperatures are required, which then requires additional fuel to be combusted in the calciner. This additional energy need can be partially offset by the recuperation of heat in the form of the hot CaO and CO₂ streams, and heat produced from the exothermic carbonation reaction [5].

Sorbent Enhanced Steam Methane Reforming (SE-SMR) is the combination of CaL and SMR. The concept was first introduced by Air Products and Chemicals Inc. [6], back in the mid-late 90s and has since been subject to substantial research, branching off in many different directions.

Combining Equations (2-3)(SMR) and (2-4)(CaL) gives the overall SE-SMR reaction (2-5) Equation of:



In one of the early publications on SE-SMR, (referred to as Sorption Enhanced Reaction Process (SERP) in the older studies), the novel 'cyclic H₂ SERP' idea was introduced, having been verified experimentally (on a packed bed containing a mixture of reforming catalyst and K₂CO₃-promoted hydrotalcite) on laboratory-scale equipment, with the following steps, and illustrated in Figure 2-3 below [6]:

1. Sorption-reaction step: Steam and methane at a specified ratio are fed to the reactor and near pure H₂ product stream (>98% purity) exits as reactor effluent. The reaction step is repeated up to the point when the H₂ purity in the product decreases to a pre-set level. The feed is then diverted to a second identical reactor.
2. De-pressurisation step: The reactor is depressurized. The effluent gas can be recycled as feed to another reactor or used as fuel.

3. Purge step: The reactor is purged with a weak adsorbing gas such as methane to desorb the CO₂. The desorbed gas consists of CH₄, CO₂, H₂, and H₂O and is either separated for recycle of the CH₄ purge gas or used as fuel.
4. Product purge step: The weakly adsorbing purge gas (CH₄) is removed from the reactor with product H₂. The removed CH₄ is recycled as purge gas or used as fuel.
5. Product pressurisation: The reactor is pressurised to the reaction pressure with pure H₂. At this point, the regeneration of the reactor is complete, and it is ready to undergo a new cycle [6]

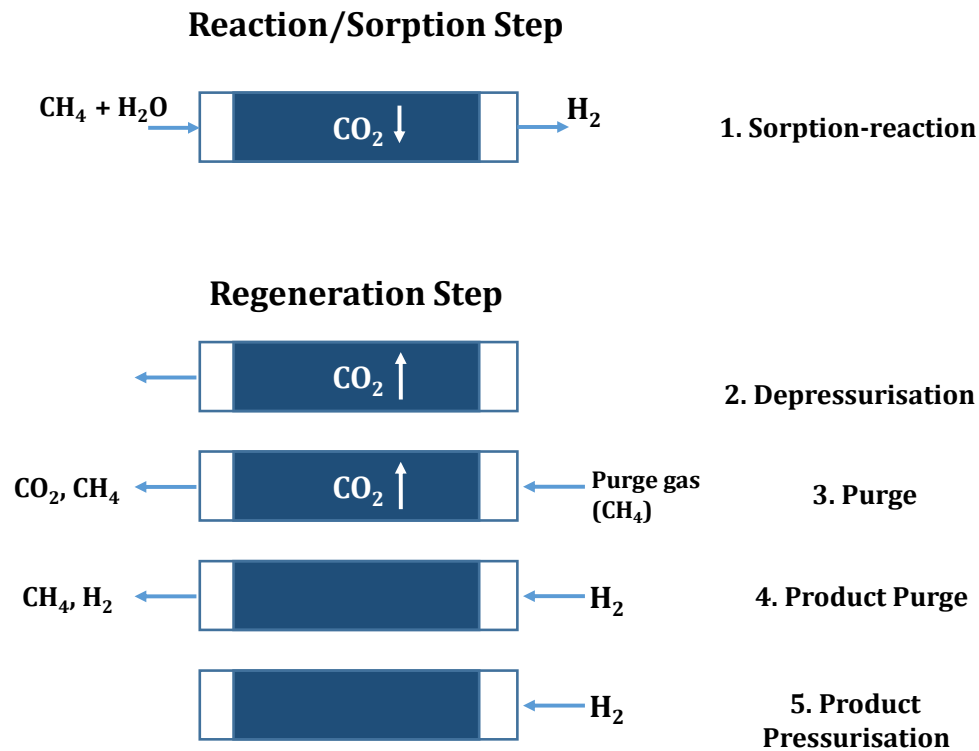


Figure 2-3 Diagram of the sequential SERP steps for the production of carbon monoxide by the reverse water-gas shift reaction.

Within this study the best performance conditions were proposed in Table 2-1, [7]:

Table 2-1 Best performance of SERP from Carvill et al. 1996

Parameter	Value
S:C	~8:1
Sorbent: Catalyst *	1:3
Operating temperature	450°C
Operating pressure	55 psig (~3.8 bar)
Methane conversion	68%
H₂ purity	95%

*It should be noted that catalyst and sorbent materials were not made explicit in this study and were referred to only as codes as this was still in the development stage. Subsequent studies revealed that the best performance catalyst was a commercial Ni-based catalyst and the sorbent, which was referred to as E-4, was found to be a proprietary pelletised K₂CO₃-promoted hydrotalcite based material [8]. Section 2.4.1 gives further details on the adsorbents that were evaluated at the early conception stage.

A few years later the same research group refined the operating conditions, shown in Table 2-2, [8], which give parameters relatively closer to figures seen in the present studies:

Table 2-2 Best performance of SERP from Hufton, Mayorga & Sircar, 1999

Parameter	Value
S:C	6:1
Sorbent: Catalyst *	1:1
Operating temperature	450°C
Operating pressure	55 psig (~3.8 bar)
Methane conversion	82%
H₂ purity	96%

Since these early studies, there has been many studies looking at varying certain parameters, in order to achieve optimal operating conditions, i.e., lower operating temperature and pressure, lower operating costs, all whilst maintaining a high H₂ product purity and CH₄ conversion rate.

A study, conducted fairly close to the introduction of SERP, is by Ding and Alpay [9], which focused on modelling and verifying the kinetics of the process in order to allow for the design, analysis or scale-up of pressure or concentration swing-based adsorption reactors [10]. This study references the main research at the time of publication which in this case was Hufton, Mayorga, Carvill and Sircar, and additionally two fundamental papers: Han and Harrison (1994) and Brun-Tsekhovoi, Zadorin, Katsobashvili and Kourdyumov (1986) [11]. The latter two were also closely referenced as being in Anand et al. 1995 [6] and have since been considered to have essentially paved the way for more recent studies on SE-SMR.

The Ding & Alpay study confirmed numerous kinetic and equilibrium parameters through the combination of various existing studies (Xu & Froment [12], Ergun [13], Edwards & Richardson [13], Kunii & Levenspiel [14] and others) along with the mathematical modelling tool gPROMS. The models from Ding & Alpay [9] and additional research from Xiu et al. [15] are realistic in the SERP description; however, with the nonlinear interaction involved in these models, the numerical results are not easy to be interpreted. The study did however confirm that the rate expressions proposed by Xu & Froment are suitable for both the transient and steady-state periods of the process. However as steady state conditions are rarely the case, and this model was only utilised as in the early days, as it does not involve complexities as compared to the unsteady state modelling.

Another experimental study, reported that the overall reaction profile can be described as a single step process, divided into 4 different subsections, as seen in Figure 2-4, [16]:

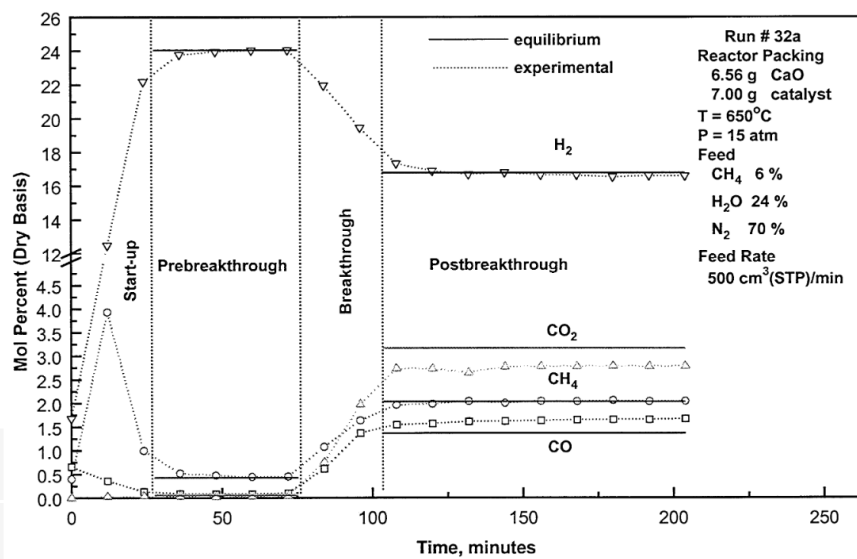


Figure 2-4 Typical breakthrough curve for a completely developed SE-SMR

1. **Start-up** – This is an unsteady time period that allows for the catalyst activation. In this stage there is a sharp rise in H₂ and a sharp fall in the CH₄ as the reforming reactions begin to take place.
2. **Pre-breakthrough** – The CaO is fully available, the SMR and WGS reactions are running at maximum efficiency thus, the maximum yield of H₂ is reached. The removal of H₂ and CO₂ in the course of the SMR and WGS reaction shifts the equilibrium to the forward reaction and improves the efficiency of H₂ production and CO₂ capture.
3. **Breakthrough** – Adsorption reaction efficiency starts decreasing as the sorbent reaches saturation, (i.e., CaO has been converted to CaCO₃), and the H₂ concentration begins to decrease. It is a gradual change (to the standard SMR reaction) as some internal CaO still reacts with CO₂ but at a slower rate due to product layer diffusion.
4. **Post-breakthrough** – All sorbent action ends at this point; thus, only the catalytic SMR reaction takes place, and the equilibrium level achieved is by conventional SMR [2,17]

The main cause of the varying 4 subsections is due to the structure of the sorbent material which is further explained in section 2.4.1.

2.2 Reactor Design and Operation

As previously mentioned, the first proposal of a singular reactor performing both the carbonation and calcination of SE-SMR, was by Hufton and Mayorga et al. 1996 [8]. In this publication, the reaction was first modelled in a fixed packed column consisting of both the catalyst and sorbent.

However, over the years, it has been competing with other design such as a circulating fluidised bed reactor (CFB), which is a mature technology at the large-scale. It was reported by [5] that it was the most suitable for the carbonation and calcination vessels due to a good gas-solid contacting and temperature uniformity across the reactor bed.

This section focuses on two designs: fluidised bed reactors and packed-bed reactors with alternating reaction-regeneration functions. These are of interest as various studies and literature indicate they are suitable to provide overall steady state H_2 production.

2.2.1 Fluidised Bed

The fluidised bed model is one of the favoured in literature concerning sorption enhanced steam methane reforming for various reasons. In addition to the fact that they are widely used in industry, they are preferred due to the favourable mass- and heat transfer characteristics [18].

In fluidised beds, the solid particles are suspended in an upward flowing stream of gas. At suitably high fluid velocities, the force between solid particle sands the fluid counterbalances the weight of the solid particles. In this state, the solid particles are referred to as fluidised because they exhibit fluid-like behaviour. Increasing the gas flow rate beyond the point of minimum fluidisation, the fluidised bed reactors can be operated in various fluidisation regimes. In view of chemical reaction processes, fluidised bed reactors are particularly suitable for highly

exothermic and temperature-sensitive reactions, such as steam methane reforming. The use of solids also means that fluidised bed systems represent optimal technology since they allow large amounts of solids to be transferred easily [18].

In particular, it has been proven that gas-solid contact in a laboratory-scale bubbling fluidised bed reactor is adequate to achieve close approximation to equilibrium conditions. Product hydrogen concentrations as high as 98% (dry basis) were attained during the steady-state pre-breakthrough period [19,20]

Whilst with a circulating fluidised bed (similar to bubbling fluidised bed but with a much higher gas velocity relative to the minimum fluidisation velocity), they also have the ability to divert a portion of the sorbent to a hydration reactor for reactivation if necessary. Figure 2-5, [20] gives an example of what a commercial process might look like using fluidised beds.

Another advantage that fluidised beds have over a packed bed reactor, is that they have lower operating and capital costs due to the packed bed material; a key for the large-scale deployment of sorption enhanced reforming is the sorbent development [20].

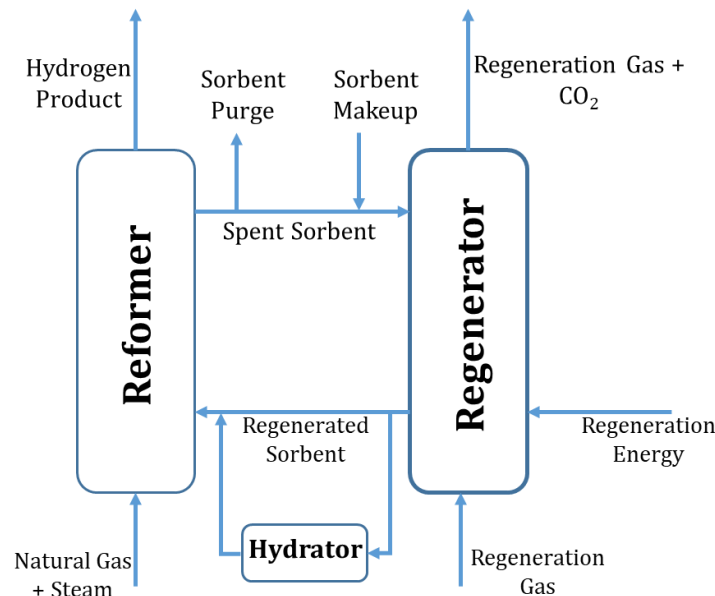


Figure 2-5 Schematic of a possible SE-SMR H₂ production using dual circulating fluidised bed reactors with sorbent hydration.

2.2.2 Packed Bed

More recently, packed beds were proposed as an alternative to fluidised beds as they allow for easier operation at higher pressures, as well as other operational issues such as removal of the gas/solid separation step, as well as fewer issues with attrition and elutriation of the solid material [21].

Because of these factors, the reactor bed design is important to consider the sorption capacity, as well as the sorbents structural characteristics specifically the surface area and porosity. Additionally, with the many cycles, that take place during a reaction, the implementation of the process on an industrial scale has to be robust enough to withstand the consecutive cycles. This is where the use of a packed bed would be beneficial because it could reduce the effect of attrition, as the sorbent would not be very mobile [21].

There are many proven studies of the sorption enhanced reforming process in a fluidised bed with particular natural sorbent materials such as calcined dolomite [22]. However as more recent publications are exploring the study of new synthetic Calcium Oxide based sorbents, there aren't many developed or validated studies within a fluidised bed. Typically, these has been conducted in fixed bed reactors. This is because they allow for the operating limits for both the sorbent particles and the combined sorbent- catalyst material (CSCM) to be determined [23,24,25].

The use of CSCM is a primary reason why a packed bed reactor would be favoured to model the SE-SMR reaction for this study. Bi-functional catalysts are a fast-emerging area of research and in order to stay at the forefront of further development of sorption enhanced reforming, a packed bed reactor is preferred, however due to the practicality of a lab-scale reactor, this isn't ideal [24].

The issues presented through using two packed beds operated in parallel, alternating feed and temperatures at constant pressure, is that due to the constant change in conditions the production isn't particularly stable. Temperature, pressure, and gas composition must be carefully controlled to

avoid catalyst deactivation due to coking [17]. Therefore, only small-scale batch production would be advisable using this system.

For the large scale, a network of packed beds was proposed by [26], reporting that the energetic self-sufficiency could be reached for SE-SMR, by burning part of the produced H_2 in an external combustor to cover the heat demand.

2.3 Operating Conditions and Materials

2.3.1 Temperature & Pressure

2.3.1.1 Carbonation

In conventional steam methane reforming, the increase of temperature greatly increases methane conversion, an expected result since the endothermic reforming reaction is favoured by high temperatures. However, the presence of the CO_2 sorbent results in high methane conversion levels even at temperatures below $650^\circ C$ as shown in Figure 2-6(a) [27].

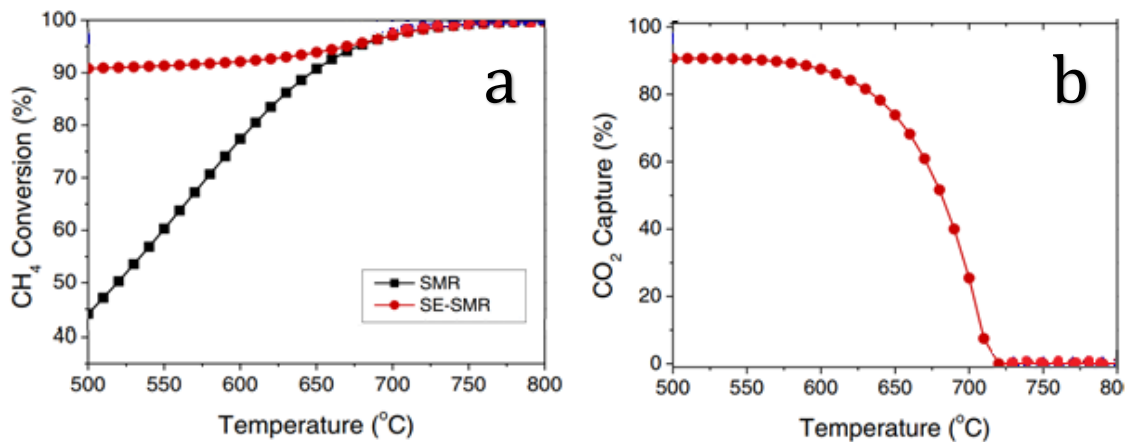


Figure 2-6(a) CH_4 conversion (b) CO_2 capture efficiency as a function of the reformer's temperature ($P=1$ atm, S/C molar ratio=3, CaO/C molar ratio=1)

The advantage that SE-SMR has over SMR at low temperatures is related to the in-situ capture of the produced CO_2 . At temperatures less than $\sim 650^\circ C$, the CO_2 capture efficiency is over 90% (Figure 2-6(b)). It starts to decrease rapidly at $\sim 670^\circ C$ and drops to zero at around $720^\circ C$.

When the conditions are increased to pressure of 5 atm, S/C ratio of 4 and CaO/C ratio of 2, the graph retains the similar curve shape, but at higher temperatures as shown in Figure 2-7[16]. When compared to the conventional SMR process, the temperature at which the maximum H_2 concentration is approximately 0.77 at 800°C, whereas the maximum of 0.98 is reached by SE-SMR at 580°C.

From Figure 2-7, it's clear that H_2 production is far greater in SE-SMR than traditional SMR, but only up to a certain temperature of approximately 820°C.

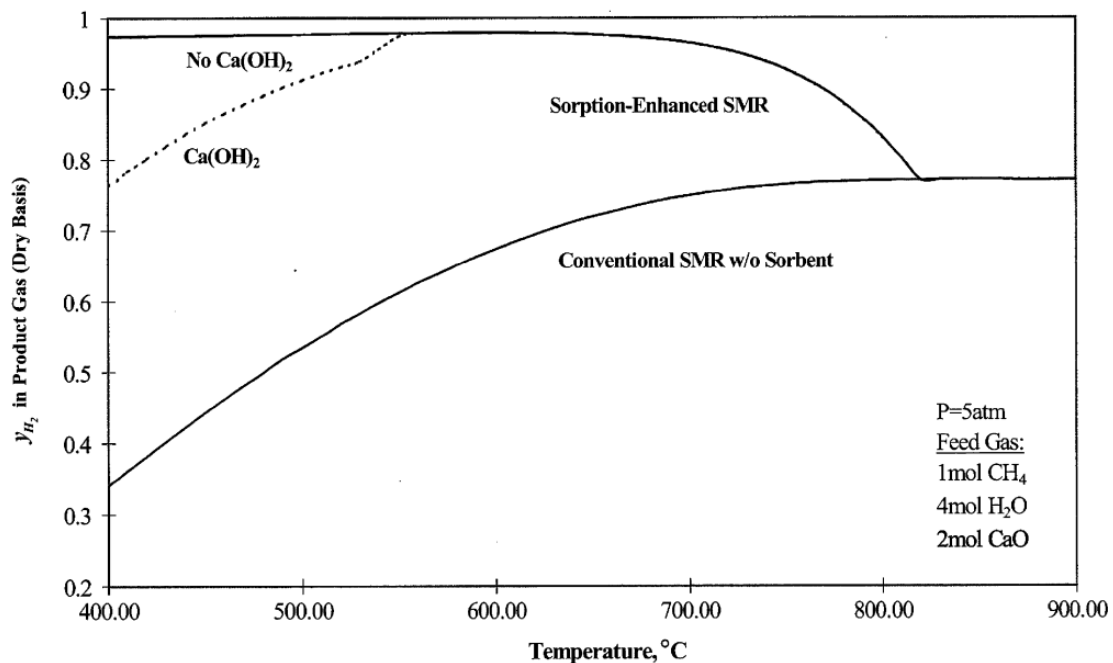


Figure 2-7 Equilibrium H_2 molar fraction (dry basis) in the product gas with and without the CaO-based acceptor as a function of temperature at fixed pressure and S/C molar ratio

At high temperatures the partial pressure of CO_2 in the system is lower than its equilibrium partial pressure which favours $CaCO_3$ decomposition (Figure 2-8, [28]). As a result, at temperatures above 820°C the sorption enhanced process slips into conventional steam reforming.

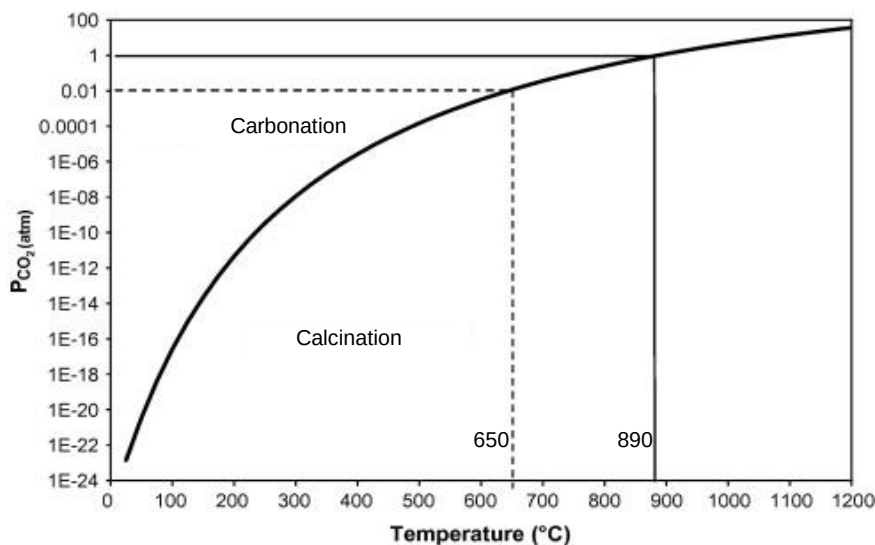


Figure 2-8 Equilibrium CO_2 pressure as a function of temperature

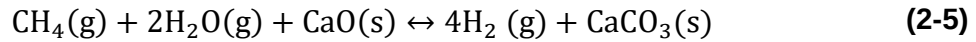
At the other end of the spectrum, SMR is not more efficient than SE-SMR at lower temperatures until around 400°C , where the benefits of CO_2 absorption become negligible. This is due to the methane conversion rate and CO_2 formation decreasing, as well as the $\text{Ca}(\text{OH})_2$ formation, leading to a low CO_2 molar concentration in the system [17].

So, whilst SE-SMR is thermodynamically more efficient, the addition of CaO introduces other issues. In order to meet the conditions of sorbent regeneration, the main factors to be met are appropriate pressure and temperature, as shown in Figure 2-8. Temperatures and pressures above and to the left of the equilibrium line favour carbonation, whereas carbonate decomposition is favoured at conditions below and to the right of the line [20].

In terms of the effect of pressure it has been stated that, the thermodynamic characteristics of the SMR and WGS reactions (Equations (2-1) and (2-2) respectively) suggest that larger H_2 and significantly lower CO concentrations can be achieved at lower pressure, although lower reaction temperature is also required [20].

The low-pressure argument was supported by [17], who stated that from a thermodynamic equilibrium point of view, a lower pressure promotes the

production of H₂. This is shown by the overall reaction expressed by which demands an increase in the total number of gas phase moles.



A study on the effects of pressure was presented in [29] and is shown in Figure 2-9. Equilibrium hydrogen content, which is almost independent of temperature and is favoured by low pressure, approaches 98% and 96% at 1 and 3 bars, respectively. In contrast, the equilibrium CO content increases with increasing temperature and decreasing pressure.

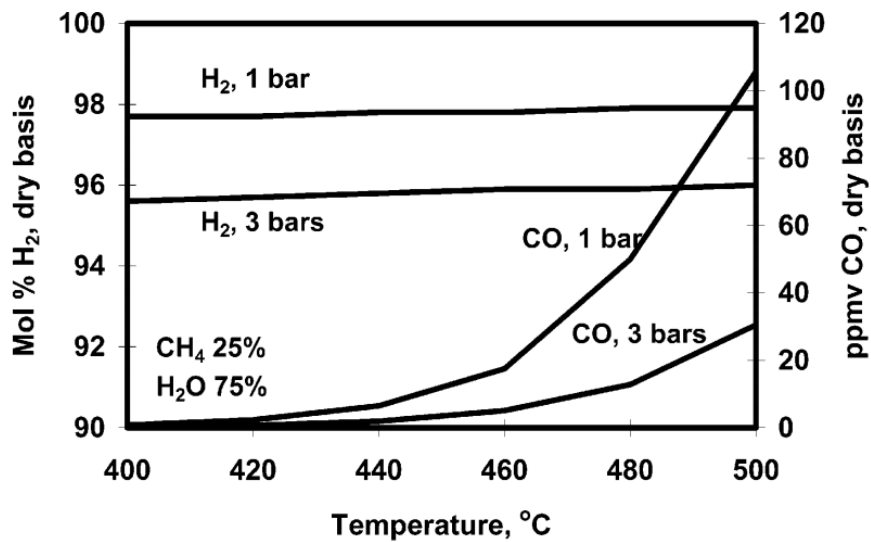


Figure 2-9 Equilibrium hydrogen and carbon monoxide concentrations as a function of temperature and pressure

Figure 2-9 suggests that keeping the temperature as low as 480-500°C is the only suitable operating temperatures to correspond with a low pressure. However, another study by [27] used the conventional SE-SMR temperature of 650°C, and, although it doesn't explicitly reference the CO make, the study confirms that lower pressures, close to atmospheric are optimal for SE-SMR- see Figure 2-10(a-d).

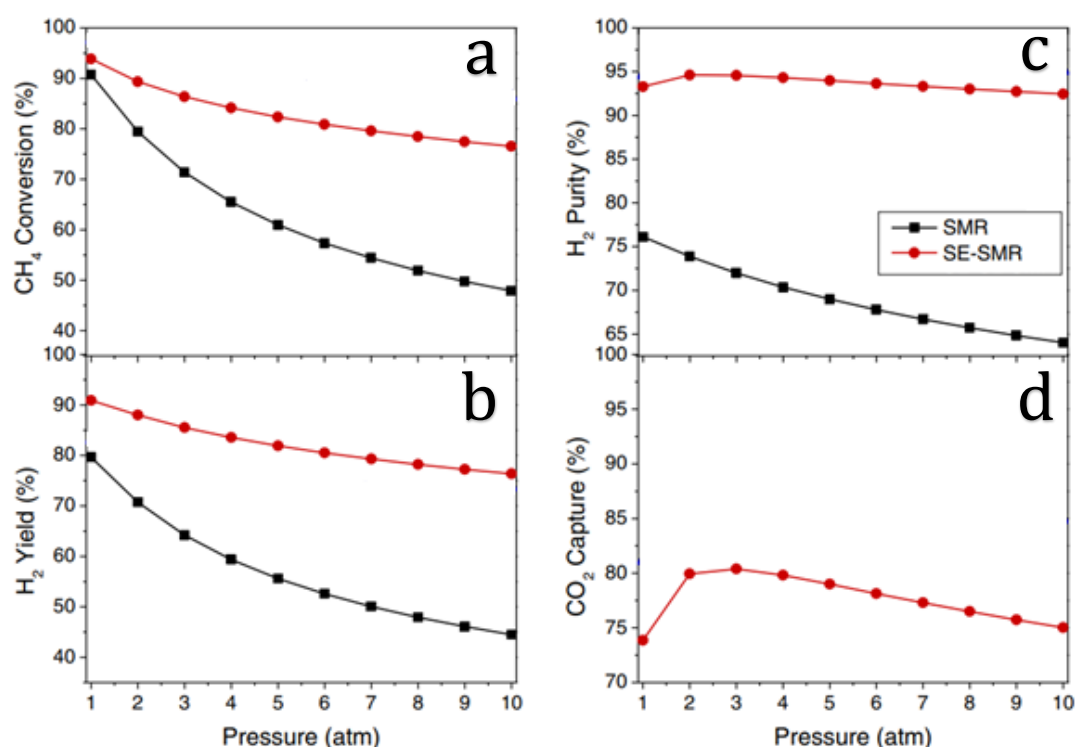


Figure 2-10 CH_4 conversion (a), H_2 yield (b), H_2 purity (c) and CO_2 capture efficiency (d) as a function of the reformer's pressure ($T=650^{\circ}\text{C}$, $S/C=3$, CaO/C ratio=1)

Figure 2-10(a) and (b) are in agreement that a higher pressure is not in the interest of the CH_4 conversion or H_2 yield, where CH_4 conversion is defined as the quantity of carbon converted from the fuel (methane) to the product (carbon dioxide) as a percent and H_2 yield is the quantity of hydrogen produced from methane, again as a percent. However, with H_2 purity and CO_2 conversion, the values increase up until around 4 atm. This happens because in this pressure range the partial pressure of CO_2 is higher than the equilibrium pressure and the carbonation reaction is shifted to the products side, resulting in greater CO_2 sorption and production of H_2 with higher purity. Further increase of pressure leads to the expected slight decrease in H_2 purity [27].

2.3.1.2 Calcination

Although CaCO_3 decomposition in a CO_2 -free atmosphere is thermodynamically favoured at all temperatures shown in the figure, a temperature of at least 800°C is needed to ensure adequate decomposition kinetics [28]

Complete regeneration of CaCO_3 occurs at temperatures near 850°C , at atmospheric pressure. Further increase of temperature has no effect on the thermodynamics of the system. Therefore, when operating near atmospheric pressure, the temperature of the regeneration reactor needs to be at least 850°C , in order to achieve complete regeneration of the sorbent.

Pressure increase adversely affects the sorbent regeneration and therefore a higher temperature for complete desorption of CO_2 is required. By increasing pressure up to 2 atm the minimum temperature for complete calcination of CaCO_3 is around 900°C . However, high regeneration temperatures lead to sintering and agglomeration of the material, reducing its sorption capacity. Therefore, these results indicate that it is necessary to operate the regeneration cycle at pressures as low as possible [27].

2.3.2 Steam to Carbon Ratio

An increase in the S/C ratio positively effects the process of SE-SMR in terms of CH_4 conversion, H_2 yield, H_2 purity as well as CO_2 capture. This can be seen in Figure 2-11(a-d) [27] and can be explained by the shift of the equilibrium towards the products in both the SMR and WGS reactions.

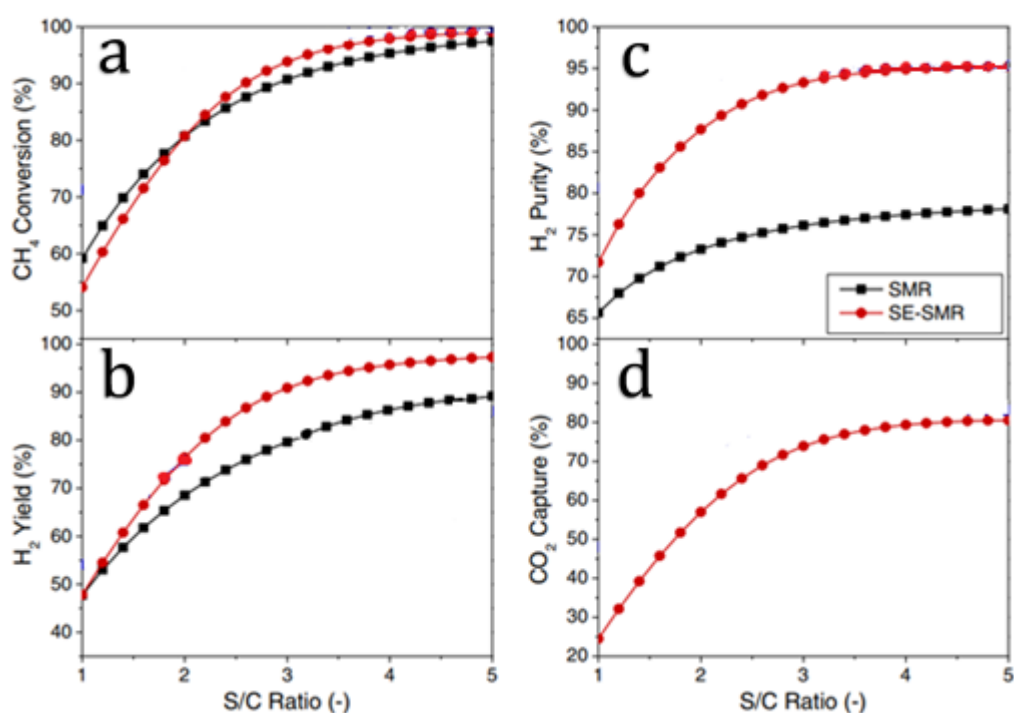


Figure 2-11 CH₄ conversion (a), H₂ yield (b), H₂ purity (c) and CO₂ capture efficiency (d) as a function of S/C ratio ($T=650^{\circ}\text{C}$, $P=1\text{ atm}$, $\text{CaO/C ratio}=1$)

A shift to the right in the in these reactions results in a greater methane conversion and H₂ yield. A shift to the right for the WGS reaction also increases CO₂ production and CO₂ partial pressure and therefore favouring the carbonation reaction of CaO, and thus the CO₂ capture by the sorbent, leading to a higher H₂ purity in the product stream.

From that logic, one would think, the higher the S/C ratio, the better, however a high demand of steam means increased energy demands, since heating of large quantities of steam is required. Additionally, as shown in Figure 2-11(c) and Figure 2-11(d), S/C ratios greater than four do not have any significant increase in parameter [30]. As you increase the amount of steam, you convert less methane per unit of duty, therefore resulting in a diminishing return, i.e. lower profits, from more money or energy invested. It's important to note, that higher conversion is achieved, but not per unit heat duty, i.e. the heat required to be added or be removed from the process fluids to create the required change in temperature.

An experimental study on the effects of S/C on the CH₄ yield and H₂ conversion was conducted with results in agreement with this principle [31]:

- **Conditions:** S/C: 2, S/C: 3 and S/C: 4 with 650°C reaction temperature and 300 minutes reaction time
- **Catalyst:** Coated Nickel Aluminide (Ni₃Al)
- **Reactor:** Microreactor
- **Analysis:** Gas chromatography

Results:

- A S/C of 4 produced the highest methane conversion of 65.56% and a S/C of 3 produced the highest hydrogen yield of 41.34%.
- The S/C of 2, showed faster and smoother stability trend conversion as early as 180 minutes from the start of the reaction.
- However, S/C of 3 showed the most optimum methane conversion and hydrogen yield and achieved stability trend conversion within the defined reaction time range of 300 minutes.
- It is inferred that the S/C of 3 is the best steam to carbon ratio for the developed catalyst in these conditions.

2.3.3 Effect of Sorbent Ratio

2.3.3.1 Sorbent to Carbon Ratio

If assuming the sorbent of choice is the popular CaO, the effect of the molar concentration relative to methane, is optimal at a roughly 1 to 1 ratio.

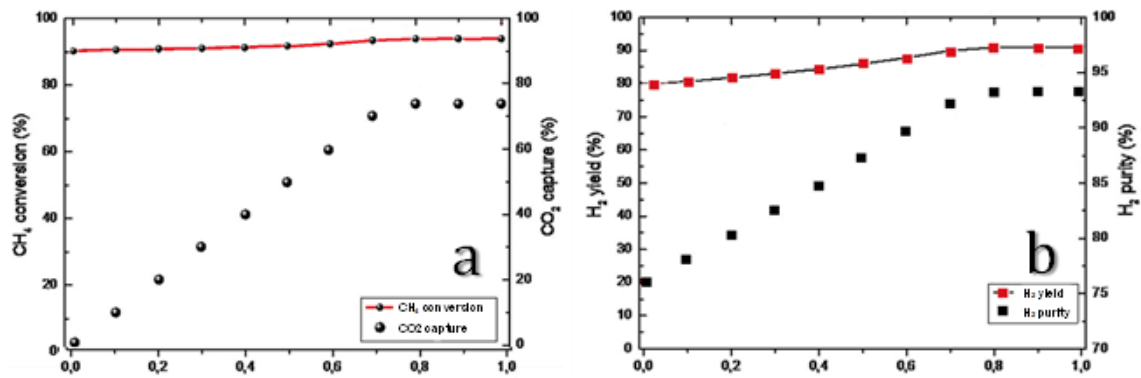


Figure 2-12 CH₄ conversion and CO₂ capture efficiency (a) and H₂ yield and purity (b) as a function of CaO/C ratio ($T=650^{\circ}\text{C}$, $P=1\text{atm}$, $S/C\text{ ratio}=3$)

This is confirmed in the same studies as referred to from the previous operating conditions [27][30] which report that CH₄ conversion and H₂ yield increase with increasing CaO/C ratio until it reaches ~ 0.8 at which the levels begin to level out, Figure 2-12, [27].

The maximum contribution of CaO in the process occurs for CaO/C ratios between 0.8 and 1. Further increase in CaO/C ratio above 1 has essentially no effect on the amount of captured CO₂, but instead has a negative impact on the energy efficiency of the process, since heating of extra CaO requires extra energy.

A study conducted on a circulating fluidised bed reported thermodynamic equilibrium data for SE-SMR at a reformer temperature and pressure of 575°C and 1atm respectively.

Table 2-3 Thermodynamic equilibrium data

CaO/CH₄ ratio	CH₄ conversion	H₂ yield	Gas composition to the reformer		
			CO flow	CO₂ flow	CaO/CaCO₃
0	98.2	0.90	16.1	33.0	-
0.5	98.8	0.95	8.6	15.8	0/0.5
1	99.5	0.98	2.9	4.8	0.16/0.84
10	99.5	0.98	2.9	4.8	9.16/0.84

As can be seen in Table 2-3, the methane conversion is high for all cases but the H₂ yield becomes high only at a CaO/CH₄ ratio of 1 [22]. The CaO/CH₄ ratio of 10 was included to provide a range of ratios, and it's clear that the results are as a CaO/CH₄ of 1. This supports the data from the studies by Antzara et al. [30] that a continuous increase of the ratio beyond, doesn't necessarily result in an improvement on conversion or yield.

2.3.3.2 Catalyst to Sorbent Ratio

For a reactor of the same size, decreasing the catalyst to sorbent ratio allows for operation at a higher space velocity since higher amount of CO₂ is possible to retain. At the same time, there is a minimum catalyst quantity requirement in order to maintain adequate reaction performance. It is also desirable to use the lowest catalyst amount to minimize the associated catalyst costs [24].

In the Arstad et al. [22] study mentioned above, the primary objective was observing the effect of the catalyst/sorbent ratio on SE-SMR results, in particular a calcined dolomite sorbent and a Ni based catalyst. As seen in Figure 2-13, it was reported that a lower percentage of catalyst in comparison to sorbent (catalyst/sorbent ratio of 20/80) performed worse compared to a 50/50 ratio.

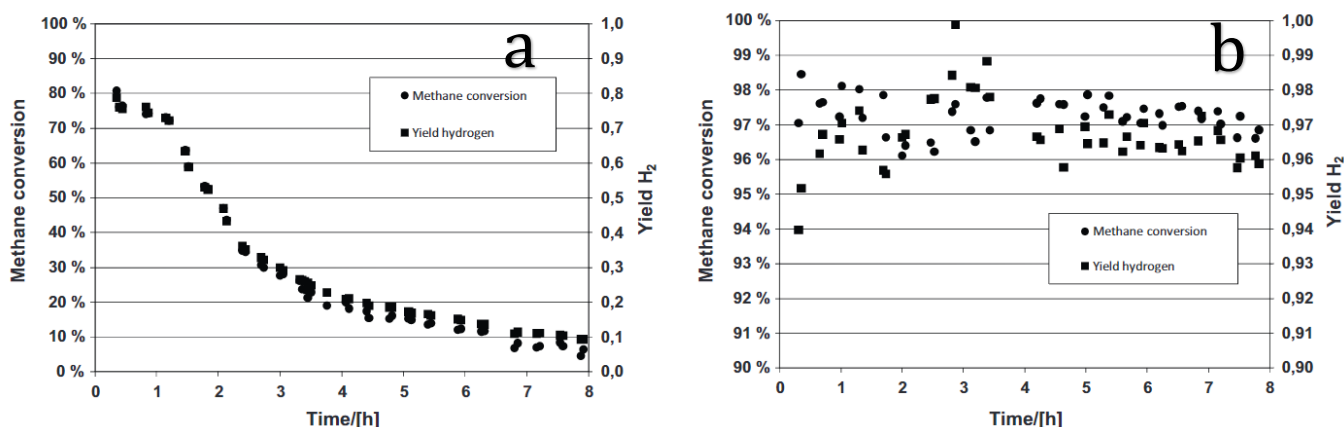


Figure 2-13 Methane conversion and hydrogen yield (a) 1:4 cat:sorb ratio (b) 1:1 ratio

At a 1:4 ratio, methane conversion and H₂ purity begin at around 80% and decline to a very insignificant conversion and yield after 8 hours, whereas at a 1:1 ratio starts at around 98% and is able to be maintained after the 8 hours [22]. Conditions for this comparison were: S/C ratio: 4. Reformer temp: 575°C, Regenerator temp: 890–900°C.

As there is very little research conducted on this ratio, in particular a chemical looping (CLC) study was analysed as it is a very similar process to CaL, difference being the use of a metal oxide oxygen carrier in CLC. Fernandez et al. [32] evaluated the effect of Cat/Sorb ratio on H₂ production and CH₄ conversion also using a calcium-based sorbent and a nickel-based catalyst on a Ca/Cu chemical loop process for hydrogen production, with results shown in Figure2-14, [32].

As the catalyst/sorbent ratio decreases from 0.3 to 0.1, the H₂ purity drops from 92% to 85% and likewise the methane conversion from 85% to 60%. The graph confirms that very low ratios result in lower methane conversion and therefore longer pre breakthrough periods and thus poor hydrogen production. An approximate ratio of 0.6-0.7 appears to be the minimum amount of reforming catalyst in order to achieve the maximum possible hydrogen production. Conditions for this comparison were 650°C, 35 bar, S/C molar ratio of 5 [32].

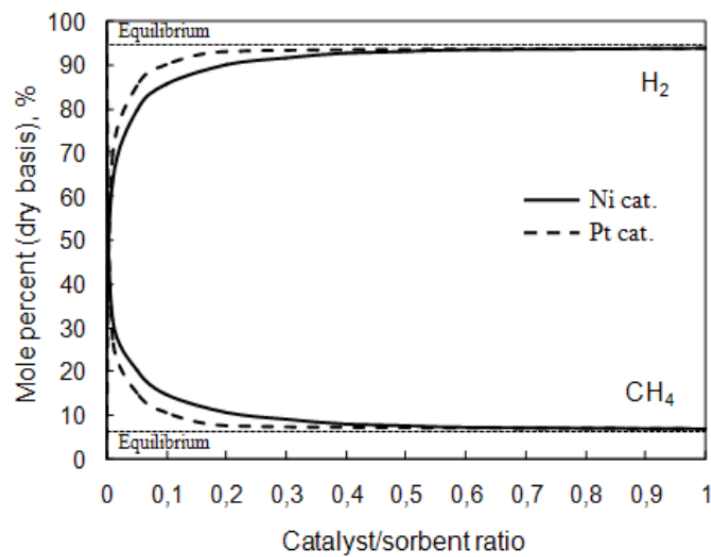


Figure 2-14 Dynamic H_2 and CH_4 profiles on a dry basis at the reactor exit at different catalyst/sorbent ratios. Pt used for comparison.

2.3.4 Effect of Particle Size

The effects of sorbent particle size on the carbonation process were explored in [33]. The particle size fractions explored were 0.25–0.4, 0.4–0.6, 0.6–0.8 and 0.8–1 mm) of a limestone named “La Blanca” and they were tested in a TGA apparatus. As shown in Figure 2-15(a), [33] with the increase in the grain size, the resistance of CO_2 diffusion inside CaO particles also increases and thus the degree of conversion is reduced. After 20 cycles, Figure 2-15(b), the sorbent capture capacity is lower, as expected, and the clear differences in the curves between the different particle sizes have disappeared. This indicates that there are no diffusion effects in the interior of the particles and the reaction rate must be controlled by the reaction mechanism taking place uniformly on a free surface of CaO [33].

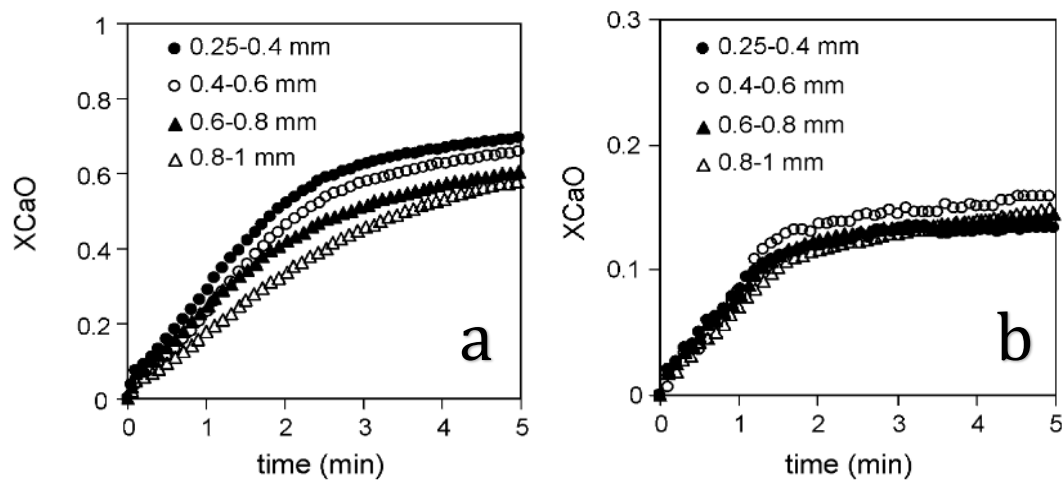


Figure 2-15 Conversion curves vs. time for different particle size. Limestone: La Blanca, PCO_2 0.1bar, carbonation T :650°C, for 20 min; calcination T :850°C, for 15 min.(a) cycle 1; (b) cycle 20

Various other literature has also confirmed that there is no significant impact in maximum carbonation conversions, with increased particle size for example [34] where the size was evaluated from 0.335-0.5 to 1-1.18 mm, and [35] which evaluated 0.1-0.25 mm to 0.8-1 mm.

However, particle size does have an influence on the rate of carbonation- the smaller the particle size, the more pronounced the shift between the fast and slow reaction stages is. [36] confirmed that with increasing cycle number this effect disappears as mentioned above in Grasa et al. 2008 [33]. Larger pores are formed, enhancing CO_2 diffusion through the particle porous structure.

2.3.5 Operating Conditions Conclusions

Based on section 2.3 and the study by Antzara et al 2014, 2015, [27][30], the optimal operating conditions of SE-SMR are concluded below.

Table 2-4 Optimal operating conditions for sorption enhanced steam methane reforming.

Reforming/ Carbonation	
Temperature	550 - 650°C
Pressure	1 - 4 atm
Steam/Carbon Ratio	~3
CaO/Carbon Ratio	0.8 - 1.0
Catalyst/Sorbent Ratio	0.7 - 1
Regeneration/ Calcination	
Temperature	850 - 900°C
Pressure	Atmospheric

The next section will discuss the sorbent, support, and catalyst type, which in turn will give an indication of particle size and a more conclusive and specific position on catalyst/sorbent ratio.

2.4 Sorbent and Catalyst Selection

2.4.1 Sorbent Material

As stated in an early Hufton et al. study [8], an important requirement for a sorbent in the use of SE-SMR is its availability and ability to selectively remove CO₂ at fairly high temperatures (300 - 500°C) in the presence of steam.

Other important factors include:

- The ability for the adsorbent to be regenerated
- the adsorbent having a high CO₂ capacity
- economically viable
- having fast reaction kinetics for the adsorption and desorption steps
- the material is stable in presence of steam
- and maintains adequate mechanical strength after carrying out certain number of cycles

These are the fairly basic properties that are required. Since the early studies, more knowledge has been gained on what makes an efficient and appropriate sorbent material. For example, durability comparable to the catalysts, to minimize the requirements of purge of the system to renew the spent sorbent; [25] and adequate pore size distribution with large pore volume and active surface of fresh sorbent; [37].

More specifically, [38] concurred that, a minimum amount of 10–15% wt. was required to maintain a stable and improved CO₂ carrying in the long-term performance.

Temperature, duration of carbonation and calcination, sorbent category, and sorbent particle size are all key players in determining the suitability of a sorbent for SE-SMR. However, the principal indicator of sorbent performance is the number of reaction cycles, and this is the key parameter in experimental models describing the change in carbonation levels [39].

One of the earliest players in the sorbent game was hydrotalcite (HTC), as explored in the Hufton et al. studies as the ‘best performing’ sorbent in their studies. Table 2-5, and Figure 2-16, both from [7], give a comparison of the sorbents tested in their studies, and it’s clear why the HTC sorbent (E4) was a clear front runner at the time.

Table 2-5 High performing sorbents in SE-SMR; *Capacity loss after 1 day of steam (10 atm H₂O/ 0.3 atm CO₂) (Binary Desorption Unit); ** Air Products proprietary pelletised K₂CO₃-promoted hydrotalcite based material)

Material	Reversible CO ₂ capacity at 400C (mmol/g) (dry CO ₂ -TGA)	Cyclic Stability	Hydrothermal Stability at 400°C *
MgO	0.20	Loses 20% after 20 cycles	-
CaO	0.05	-	-
C3 (promoted Al ₂ O ₃)	0.30 – 0.35	-	5% loss
C31 (promoted Al ₂ O ₃)	0.25 – 0.30	Loses 25% after 20 cycles	10% loss
B2 (promoted Al ₂ O ₃)	0.30 – 0.50	Loses 40% after 60 cycles	30% loss
E4**	0.50 – 0.60	Stabilizes at 0.45mmol/g after 20 cycles	No loss

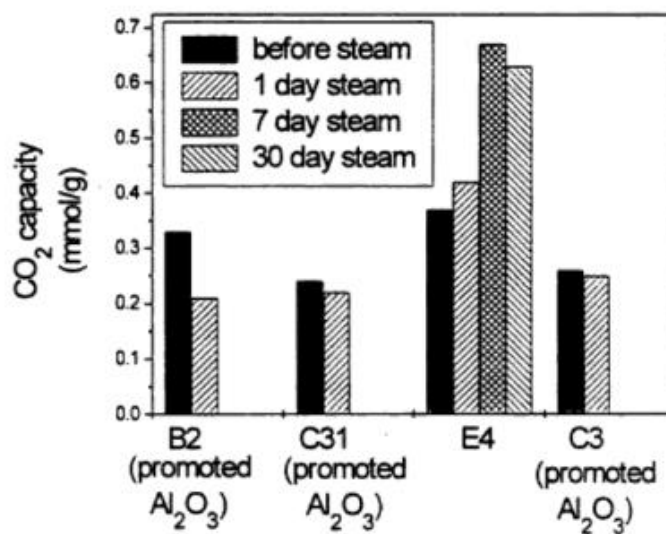


Figure 2-16 Hydrothermal Stability Test on Hydrothermal Stability Testing Unit

Hydrotalcite, aka magnesium–aluminium hydroxycarbonate ($\text{Mg}_6\text{Al}_2(\text{OH})_{16}\text{CO}_3 \cdot 4\text{H}_2\text{O}$), is a naturally occurring mineral with a layered crystal structure, and they are also referred to as anionic clays or layered double hydroxides [40].

The main reason why HTC based sorbents are suitable is the high stability after multiple cycles, along with other factors such as lower energy of regeneration, and suitable kinetics of carbon dioxide adsorption [41]. As it was introduced as a proprietary sorbent from the SERP patent holders, Air Products it can be understandable that some could argue there is bias towards HTC as a sorbent, however it has been subject to many studies throughout the years, and it still stands to be a considered sorbent.

Since the early studies, the popularity of HTC as a sorbent has been somewhat warped by use of calcium-based sorbents. A very clear side by side comparison of these sorbents is shown in Table 2-6, [17].

Table 2-6 Comparison of natural and synthetic sorbents

Type	Material	Stoichiometric adsorption ability	Regenerating temperature (°C)	Stoichiometric adsorption ability after 45 cycles
Natural	Calcium carbonate (Limestone- CaCO_3)	0.79	900	0.316
	Dolomite ($\text{CaMg}(\text{CO}_3)_2$)	0.46	900	0.16
	Huntite ($\text{Mg}_3\text{Ca}(\text{CO}_3)_4$)	0.25	900	0.20
	K_2CO_3 -promoted hydrotalcite	0.029 (0.65) [9]	400 (through PSA)	Stable after nearly 6000 cycles [8]
Synthetic	Lithium orthosilicate (Li_4SiO_4)	0.37	750	Stable until 100 cycles
	Lithium zirconate (Li_2ZrO_3)	0.29	690 (in N_2)	Stable until 100 cycles
	Sodium zirconate (Na_2ZrO_3)	0.24	790 (in air)	Stable until 100 cycles

From this comparison, whilst the synthetic sorbents have a higher stability than the natural based sorbents, their manufacturing costs are not economical enough at present, and in order to compete with natural sorbents, they are required to sustain at least 10 000 cycles, and have a generally higher rate of reaction [44].

Additionally, whilst the HTC sorbent has a much lower regeneration temperature as well as higher stability after cycles, the carrying capacity is the most important factor that sets aside CaO based sorbents as the optimal choice for SE-SMR especially if to be implemented on an industrial scale.

Other materials such as lithium or sodium-based sorbents, are not thermodynamically favourable as shown in Figure 2-17, [45]. From the plot, it was interpreted by authors that it translates into, at 10 bar, 850K and S/C of 5, using CaO the result is a product containing about 98.3% H₂ (dry basis), 90% with Na₂ZrO₃, 89% with Li₂ZrO₃, and only about 82% with Li₂SiO₄ [45].

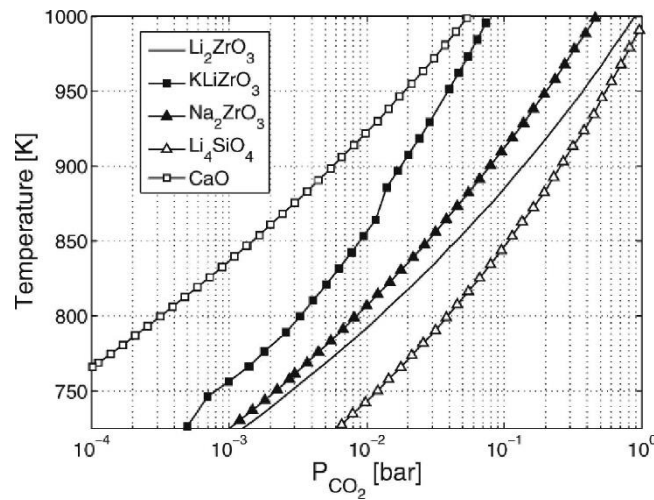


Figure 2-17 Equilibrium CO₂ pressure as a function of temperature for CaO, Li₂ZrO₃, Na₂ZrO₃, and Li₄SiO₄ sorbents.

CaO is also a promising material due to its low cost, wide availability in naturally occurring minerals, e.g., limestone or dolomite, and most importantly, its initially high absorption/ carrying capacity [4]. In more recent years, Di Giuliano and

Gallucci presented a refined study on this, with a review in 2018 focusing specifically on nickel-based catalysts and calcium looping [46].

In this study, reasons given for choice of CaO-based sorbents included:

- satisfactory CO₂ sorption capacity and rate
- full carbonation at 600–700°C and atmospheric pressure, conditions compatible with a simultaneous SMR conversion and at which CaO is thermodynamically favoured in comparison to other high-temperature sorbents.
- full calcination at atmospheric pressure and roughly 800-950°C. [17]

2.4.1.1 Sorbent Degradation & Mitigation

Despite many advantages of calcium-based sorbents, there are disadvantages which include attrition, sulphation and sintering, the latter of which is seen as the biggest issue to overcome [39], depicted in Figure 2-18; [47].

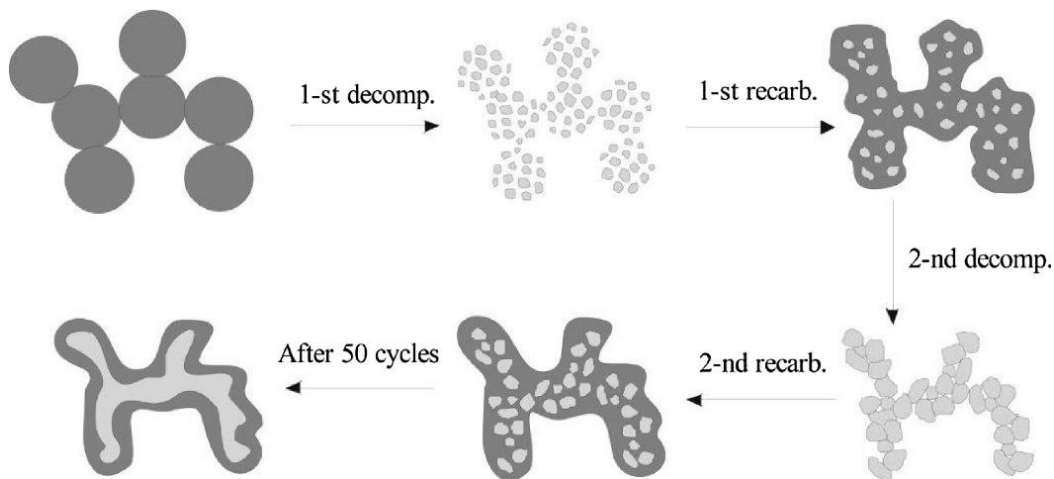


Figure 2-18 - Representation of sintering phenomena in CaO (light grey) CaCO₃ (dark grey)

Sorbent decay can be described through the rapidly fluctuating reaction kinetics within carbonation; the transition from rapid to slow phases is associated with the build-up of a CaCO₃ layer, thus resulting in obstruction of further conversion i.e.,

the rate determining step in the reaction is no longer the surface kinetics of the reaction, but instead becomes the diffusion of CO_2 through the solid CaCO_3 product [5]. Depicted graphically in Figure 2-19 [4].

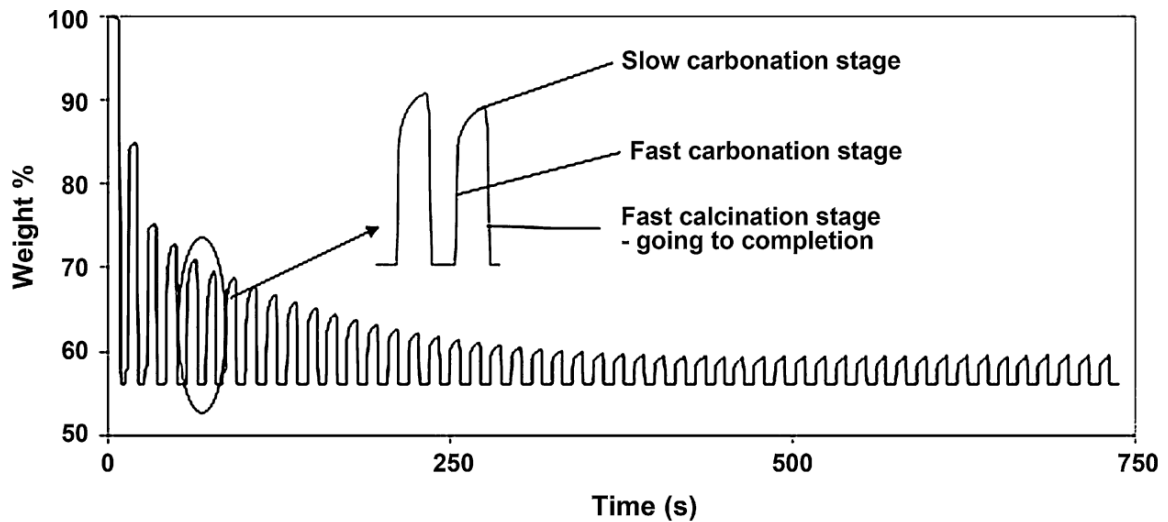


Figure 2-19 Repeated calcination/carbonation cycles of limestone in a TGA

Commonly referred to as the greatest contributing factor to sorbent degradation, sintering refers to changes in the pore shape and size distribution, and grain growth. As the surface of the CaO sorbent is coated with more and more solid CaCO_3 , the sorbent loses reactive surface area as well as total pore volume. This process is promoted with increase of temperature, long exposure at high temperatures, as well as exposure to steam, CO_2 , and impurities [5].

Less significant parameters that contribute to sorbent decay are attrition and sulphation (from the competing reaction of CaO with impurities found in fuel forming CaSO_4).

The main strategies being explored to counter the negatives of CaO based sorbents include [5][46]:

- Reactivation by steam hydration
- Thermal pre-activation
- Inert supports/stabilizers

- Pelletisation
- Nanomaterials
- Ion doping

Sorbent reactivation has been explored through many studies, via two main methods: water and steam [48]. Of the two, technically water is more efficient in the hydration of spent sorbent; however, the conditions are less than ideal due to the energy requirements for drying the wet hydrated material [49] therefore, steam hydration has been proposed as the preferred method for reactivation. In this process the calcined sorbent is cooled to ~250-450°C and subjected to high partial pressures of steam. This converts CaO to Ca(OH)₂ which is more reactive towards CO₂ and has improved pore surface area and distribution, leading to a faster carbonation reaction than that of un-hydrated CaO sorbents [50].

Thermal pre-activation was first explored by [47] and later honed by [39] in a widely cited study. Four different limestones were preheated at different temperatures (800-1300°C) and for different durations (6-48 hours) under nitrogen, then analysed via TGA. Initially the carrying capacity was lower than normal, however after about the first or second carbonation, there was an increase in carrying capacity, which steadily increases for a few cycles, until it starts to decrease as expected [4]. The trend of an increase in carrying capacity per cycle was particularly prominent in powdered samples preheated at 1000°C, which still showed an increase at the highest reported cycle of 30. The overall clear trend was that in comparison to the original, pre-treated sorbents showed better conversions over a longer series of CO₂ cycles, [39].

The use of an inert support material or a 'stabiliser' is a method that seeks to distribute inert particles in between CaO grains, thus preventing CaO sintering due to inert high Tammann temperature (temperature at which atoms or ions within a solid have bulk mobility, about one half of material's melting point in Kelvin [51]).

2.4.2 Catalyst Material

Catalyst materials used for SE-SMR should resist coke formation by steam reforming, be inactive for side-reactions, maintain the activity at high temperature, and have high mechanical strength as well as good heat transfer properties [53]. They must also be able to operate at low steam/methane ratios in order to improve the energy efficiency of the process [54].

The mechanisms of several metal-based (nickel, rhodium, platinum, ruthenium, and iridium) catalysts on the reactions of methane and steam, were explored by Wei and Iglesia [55]. They reported that platinum catalysts were the most reactive, however among these, nickel oxide has widely been proved to be preferable for use in SE-SMR due to its good conversion rate of methane and low cost ([12][17][56])

However, nickel does have its disadvantages [57]:

- Nickel nanoclusters (as opposed to a skeletal structure) are prone to sinter under catalytic reforming conditions of high temperatures and in the presence of steam.
- Severe carbon accumulation is also high particularly on nickel catalyst, known as coking.
- The production of hydrogen rich gas with a low concentration of CO is a challenge using nickel catalysts, as are not so active in the water gas shift reaction.

Other factors to consider in catalyst choice include the sensitivity to sulphur poisoning. It is assumed that the natural gas feed will have undergone desulfurization, however, most naturally occurring limestones and dolomites may still contain small quantities of sulphur. Therefore, in the reforming atmosphere of the reaction, with enough sulphur present, it can be transferred from the sorbent to the gas phase to quickly poison the nickel reforming catalyst [20].

The use of metal dispersion and alloying (particularly with Pt, Pd or Ru), are both effective methods to enhance certain properties of catalysts, including resistance

to sintering as well as resistance to nickel oxidation, which promotes its activation [58]. The addition of a metal such as iron, copper, or tin, has the possibility to improve the catalytic activity [57].

2.4.3 Combined Sorbent Catalyst Material

A combined sorbent catalyst material (CSCM), or a bifunctional material, is a one-particle system that contains a metal for the catalysis and a sorbent for the in-situ CO₂ sorption. Due to the close contact of the catalyst and sorbent in the solid, the CSCM particle system enables higher heat transfer and reduced mean path length for mass transfer of the reaction products, which enables a greater conversion efficiency compared to a two-particle system [57]. Additionally, the overall reactor volume can be reduced by using a sorbent that acts as a support [58].

The application of CSCM also isn't limited to the process of SE-SMR with methane as a carbon source, as there are many studies conducted on CSCM applied in sorption enhanced steam reforming using other raw materials such as glycerol [59][60], phenol [61], coal [62], biomass [57], ethanol [63] and methanol [64].

The different parameters of a CSCM that contribute to an optimal material- in the sense of capturing the highest CO₂ capacity after multiple cycles, whilst maintaining high methane conversion, are shown in Table 2-7.

Table 2-7 Properties and their respective experimental conditions within a CSCM

Material	Property (units)	Experimental conditions (units)
Sorbent	Last cycle capacity (gCO ₂ /gSorbent)	CaO concentration (%), cycle number, calcination, and carbonation temperatures (°C) and times / (mins), synthesis method, CaO precursor, initial cycle capacity (gCO ₂ /gSorbent)
Catalyst	Methane conversion (%)	Nickel concentration (wt%), calcination temperature (C), calcination duration (h), SMR reaction temperature (°C) fresh (unaged catalyst) BET surface area (m ² /g), steam/carbon ratio

The use of an inert support material or a ‘stabiliser’ is a method that seeks to distribute inert particles in between CaO grains, thus preventing CaO sintering due to inert high Tammann temperature (temperature at which atoms or ions within a solid have bulk mobility, about one half of material’s melting point in Kelvin [65]).

The purpose of a support for a catalyst and sorbent particle, is to increase the surface area [66] and thus, to offer more stability, ability to withstand destruction of pore volume, as well as possibly increasing the catalyst activity and/or the sorbent carrying capacity, whilst having little effect on the products from the reactions. The commonly used supports of choice are Al₂O₃ or calcium aluminates (Ca_xAl_yO_z) [46], or silica, SiO₂.

2.5 Machine Learning

2.5.1 Machine Learning in SE-SMR

2.5.1.1 ML in Hydrogen Production and Analysis

In the research of chemical processes such as steam methane reforming, the data and most of the results discussed in literature is normally found via

expensive and/or difficult experiments and measurements on real projects. Additionally, high costs, time-consuming and other variable limitations make some data/results sometimes insufficient methods of data collection. A possible alternative solution to this is via mathematical modelling as a method of data collection and analysis. In these studies, programs such as Aspen or MATLAB or gPROMS, allow for certain processes to be modelled using kinetic parameters etc. and subsequently influences of changes to the system to be predicted. There are various shortcomings of using this method such as that its time-consuming, sometimes quite complicated, and includes assumptions, data from other experiments/literature, extrapolation, and adjustments/ changes to allow it to be applicable to the current study. Therefore, the application of AI is a necessary tool to overcome these issues.

There are numerous types of AI that can be applied to better understand and control H₂ production, for example artificial neural network (ANN) [67][68], genetic algorithms (GA) [69], random forest (RF) [70], deep neural network (DNN) [71][72], to name a few.

Applying machine learning can be alternative methods of data handling, overcoming the complexity of numerical and analytical methods and high costs of empirical experiments [73]. These techniques allow use of non-iterative methods, where the only inputs necessary are inlet parameters then using the developed AI model allows the process parameter to be generated as answers to the input data set.

These models have the ability to reproduce the process from training examples so specific knowledge of the process is not essential [74]. They are deemed to be tools which sometimes can overcome the shortcomings of the programmed computing approach and the experimental procedures, or in some cases used alongside the other approaches.

Whilst the production of H₂ via SE-SMR is a fairly well researched process, there are still areas unknown and the complexity of H₂ production technology still poses

some challenging tasks. Because of this, the development of a simple, non-iterative model of a reactor is of practical significance.

As the automation of processes is expanding across most industries, it is optimal to apply the use of computational intelligence (CI) to optimise or control the production of H₂.

A review of the CI approaches applied to H₂ production was published in 2018, and it gave an overview of 21 recent leading-edge articles for H₂ production [75]. The techniques reviewed were ANN, fuzzy methods such as ANFIS, GA and one case using support vector machine (SVM) method. The studies range between 2007 and 2017, with the more recent studies, focusing on ANN and ANFIS. The objectives targeted in these studies include modelling the thermochemical conversion process (ANN), optimising H₂ production via fermentation (GA) and estimating the H₂ production from jet fuel and optimising the production for maximum yield and energy efficiency (ANFIS).

Through this review it was shown that MLP (Multi layered perceptron- a type of ANN) and ANFIS presented the highest correlation coefficient and the lowest modelling error encountered in the prediction of the H₂ production process (as of 2017). MLP is a feed-forward ANN that uses backpropagation, and supervised learning for the training of the network. The method contains input, hidden and output layers and the model aims to map input data onto an output space [76]. Table 2-8 depicts some of the values of the model evaluating factors from this study [77]. Overall, the study also concluded that using hybrid CI methods (particularly the GA-ANN method) led to an improved and optimized opportunity for the production of H₂. It should be noted that this review, covered various H₂ production methods and was not exclusive to SE-SMR.

Table 2-8 The values of the model evaluating factors.

Paper	Method	Efficiency factor	Value
1	ANN-MLP	R^2	0.99 for H ₂ production yield
	RSM	R^2	0.90 for H ₂ production yield
2	ANN-MLP	R^2	0.9726, 0.8597, 0.9638 & 0.9394 for H ₂ yield, CO yield, CH ₄ conversion
	ANN-RBF	R^2	0.9218, 0.7759, 0.8307 & 0.7425 for H ₂ yield, CO yield, CH ₄ conversion
6	BP ANN	r	0.955 for H ₂ production yield
7	ANFIS	R^2	0.998, 0.998, 0.999 & 0.999, respectively for H ₂ yield & conversion efficiency of jet fuel & butanol.
8	ANFIS	R^2	0.9, 0.97 & 0.92 for normalized exergy destruction, rational exergy, & process exergy efficiencies
		MAPE	0.019, 0.0029, 0.0228 for normalized exergy destruction, rational exergy, & process exergy efficiencies
		MSE	1.15, 0.245 & 1.39 for normalized exergy destruction, rational exergy, & process exergy efficiencies
9	Modified fuzzy algorithm	frequency	8% of successful long-term conduction cooling
12	Fuzzy AHP	Overall weights of each prod. method based on risk, cost & opportunities	0.529, 0.46, 0.449, 0.331, 0.395 & 0.366 for SMR, CGP, BG, WE, NPE, & H ₂ yield, respectively
13	Fuzzy AHP	Final score of each prod. method based on GHG emission, raw material, energy efficiency, scalability, wastes & emission	0.1945, 0.1648, 0.1474, 0.1385, 0.1151, 0.0908, 0.0863 & 0.0627 for WS-CL, H-EL, SMR, W-EL, POX, PV-EL, CG & BG, respectively
14	GA- ANN	R^2 for Training	0.966, 0.81, 0.882 & 0.92 for H ₂ concentration in the biogas, prod. rate, yield, & effluent TOC
		R^2 for validation	0.719, 0.806, 0.843 & 0.854 for H ₂ concentration in the biogas, prod. rate, yield, & effluent TOC
15	ANN-MLP	RMSE, & SEP	16.6% & 38.4%, respectively for RMSE & SEP for H ₂ prod.
	RSM	RMSE, & SEP	7.7% & 17.8% respectively for RMSE & SEP for H ₂ prod.
	GA- ANN	Max H ₂ prod.	360.5 ml/g of glucose
	GA- RSM	Max H ₂ prod.	289.8 ml/g of glucose

2.5.1.2 ML in Sorbent Materials; Quantitative Structure Property Relationship (QSPR)

There have been different approaches taken to develop and predict the behaviour of optimal properties for suitable sorbents, catalysts and CSCMs, such as the determination of kinetic and diffusion parameters and computational modelling [78-83], as well as material synthesis and experimental studies in a fluidised bed reactor [84] or fixed bed reactors [85-87].

The development of an 'optimal' CSCM has proven to be a complicated and monotonous one, however this can be overcome with the application of machine learning, specifically the process of quantitative structure–property relationship analysis (QSPR). QSPR is a modelling method to predict different physical and thermodynamic properties using the knowledge about the chemical structure of the molecules [88]. These chemical structure and properties are known as descriptors. Molecular descriptors are defined as mathematical representations of a molecule, from characteristics related to the structure of the chemical [89]. They are calculated quantum chemistry methods and allow for the modelling of many different properties in fields such as physical chemistry, pharmaceuticals, and analytical chemistry [90].

QSPR modelling has long been used particularly in the pharmaceutical industry for drug development, and it has recently seen expansion into various physical chemistry fields, including adsorption of metal organic frameworks [91-94]. Additionally, the use of QSPR in combination with other computational methods such as DFT or process simulation has also garnered effective prediction within research in the energy field [95,96], however, to the best of our knowledge, QSPR has not been applied to the development of a CSCM material for use in SE-SMR.

2.5.1.3 ML in Hydrogen Storage; Metal Organic Frameworks (MOFs)

Storing hydrogen can be difficult due to its molecular properties. Whilst hydrogen has almost three times the energy content of gasoline with 120 MJ/kg versus 44

MJ/kg respectively, this is also inversely reflected on a volume basis. The equivalent amount of hydrogen requires more storage space than gasoline with liquid hydrogen having a density of 8 MJ/L whereas gasoline has 32 MJ/L [97]. Other than associated costs and practicality, the main issue is the amount of hydrogen that can safely be stored in a vehicle and delivered as required under operational conditions [98].

Some solutions include high-pressure gas or liquid hydrogen, physical adsorption at cryogenic temperature (~ 77 K) in porous materials, primarily on the internal and external surfaces of a nano-porous solid, is an attractive option due to the speed and reversibility, whilst being maintained at relatively low storage pressures ($\sim$ up to 100 bar) [99]. Hydrogen physisorption also has a low enthalpy of adsorption, compared to chemisorption, which therefore reduces thermal management issues that can be problematic [100]. For these reasons, a lot of research has gone into developing an optimal material, particularly metal organic frameworks (MOFs).

MOFs are made from inorganic nodes and organic linkers and are a class of often crystalline materials that can be tuned to have ultrahigh porosity and high surface areas, made from infinitely extendable metal ligand networks [101,102]. These properties make MOFs an effective candidate for storage and transportation applications in clean energy, for gases such as hydrogen and methane, and as high-capacity adsorbents to meet various separation needs. As advantageous as MOF are, there are also some downsides to MOFs. These challenges vary greatly depending on the MOF class, but can include laborious or expensive synthesis, lack of electrical conductivity, and chemical, thermal, and mechanical stability to name a few [103]. For example, in terms of synthesis, difficulties can include understanding and controlling both structural and compositional complexity.

Therefore, owing to the difficulty in exhaustively assessing large libraries of these complex materials experimentally, the use of computational chemistry has been widely employed to accelerate the evaluation of such materials, particularly the

use of GCMC [104]. GCMC is a statistical-mechanical method, in which a thorough molecular-level modelling of adsorption is solved accurately [105]. It is an approach that is widely used for the simulation of adsorption equilibrium, thus making it beneficial to model H₂ adsorption on a MOF. However, a main downside is computational time which is why ML is sometimes favoured.

There are a number of studies that have explored MOFs for H₂ uptake, using either ML, GCMC or both with interesting results [106–111]. A largely cited paper in the field of MOFs for H₂ uptake is by Ahmed et al. [112], which looks at using computational screening via GCMC to examine 493,458 MOFs based on the crystallographic properties and the Chahine rule. The conditions used were a pressure swing between 100 bar and 5 bar at 77 K and the benchmark used was the best performing MOF at the time which was identified to be IRMOF-20, which had been found to have the best combination of gravimetric & volumetric capacities under these pressure-temperature conditions [113]- a tricky balance to achieve. From this study, three MOFs surpassing IRMOF-20 were found, and experimentally validated with positive results, yielding the current record-holders for balanced hydrogen capacity- SNU-70, UMCM-9, and PCN-610/NU-100, with a ceiling of ~40 g H₂/L. It should be noted that although there are MOFs with higher volumetric capacities, this study focused on high volumetric along with gravimetric capacity, thus these are the current highest performing MOFs in both aspects.

Bucior et al. [114] used ML to train on MOF databases, and in addition, they calculated new descriptors for gas adsorption derived from the energetics of MOF–guest interactions. The findings from this showed that their ML method was three times faster than conventional molecular simulations, which allows for rapid exploration of large numbers of MOFs. In order to validate their method, it was applied to screen experimental MOFs and the selected highly performing MOF was chosen to be experimentally validated, with positive results. Another study that applied the use of machine learning in combination with other methods is by Giappa et al., [97] where they employed multiscale theoretical calculations such as binding energy, to identify and explore the effect of the functional group

composition on the hydrogen uptake profile. Consequently, GCMC was then applied to three MOF candidates, with some functional group modifications, and the total gravimetric and volumetric uptake profile was obtained both at cryogenic (77 K) and ambient temperature (300 K). The results obtained, showed a fair prediction capability of the H₂ binding energies even with a limited amount of data, and that the use of ML can largely bypass the need for computational methods, by extrapolating for valuable physical-chemical insights.

Anderson et al. [115], successfully used both machine learning and molecular simulation for the exploration of hydrogen storage of porous crystals, including MOFs, amongst others. They gave insights into how inputs influenced the adsorption output, namely volumetric surface area (VSA) and gravimetric surface area (GSA), again two parameters that lead to difficulty in finding a good balance between volumetric and gravimetric uptake. In this study, over 18,000 adsorption simulations using the topologically based crystal constructor code, ToBaCCo, was used, alongside the application of a neural network model to predict hydrogen uptake and various conditions. Results showed that the combined application of simulation data at various temperature and pressures, as inputs to a neural network, allowed for the prediction of hydrogen adsorption at different thermodynamic conditions, with final results determining that 62 g/L is an attainable deliverable hydrogen capacity for 100 bar/77 K ↔ 5 bar/160 K swing conditions.

2.6 Literature Review Conclusion

From the findings presented in the literature review, there have been areas that have given an insight into the future development of this process along with machine learning. There are gaps in the research which allows for the research presented in this study going forward to address and pave the way for future work. As mentioned, prior, the process of SE-SMR is a well-established one, and little or very novel work can result in more unknowns, therefore the main focus will be on the AI aspect, with some exceptions.

An emerging development in the field of SE-SMR is the concept of combined catalyst and sorbent material (CSCM) or also known as hybrid/ bifunctional catalyst sorbent and catalyst. There are already substantial studies focused on this developed, and although the separate materials in their own right have been subject to years of analysis and investigation, with a general consensus, it can't be said the same for a CSCM combination, therefore that will be an incorporated part of this study.

From this literature review however other aspects such as the operating conditions, reactor type, and reactor material properties were explored, and it was concluded that there is an agreement on these parameter conditions amongst esteemed researchers.

Certainly, within the modelling aspect of the SE-SMR process, there is larger scope, such as research on the machine learning control, namely neural networks. Whilst other processes have had these models applied for various outcomes (optimisations, predictions, verification), the sorption enhanced steam methane reforming process as a whole, has not had NN applied. Moreover, QSPR is a very novel approach to this process, therefore this will be applied in our study into the exploration of hydrogen production via SE-SMR, from process control to the materials used (CSCM or doped sorbents) to the final product i.e., the storage of the blue hydrogen produced.

2.7 References

- [1] Edwards PP., Kuznetsov VL., David WIF., Brandon NP. Hydrogen and fuel cells: Towards a sustainable energy future. Energy Policy. December 2008; 36(12): 4356–4362. Available at: DOI:10.1016/j.enpol.2008.09.036
- [2] Clough P. An investigation into H₂ production by sorbent enhanced steam reforming of biomass. Chemical Engineering. 2017; PhD (June).
- [3] Gu R., Ding J., Wang Y., Yuan Q., Wang W., Lu J. Heat transfer and storage performance of steam methane reforming in tubular reactor with focused

solar simulator. *Applied Energy*. Elsevier; 2019; 233–234(January 2018): 789–801. Available at: DOI:10.1016/j.apenergy.2018.10.072

[4] Blamey J., Anthony E.J., Wang J., Fennell P.S. The calcium looping cycle for large-scale CO₂ capture. *Progress in Energy and Combustion Science*. Elsevier Ltd; 2010; 36(2): 260–279. Available at: DOI:10.1016/j.pecs.2009.10.001

[5] MacDowell N., Florin N., Buchard A., Hallett J., Galindo A., Jackson G., et al. An overview of CO₂ capture technologies. 2010; : 1645–1669. Available at: DOI:10.1039/c004106h

[6] Anand M., Hufton J., Mayorga S., Nataraj S., Sircar S., Gaffney T. Sorption enhanced reaction (SERP) for the production of hydrogen. APCI report for DOE. 1995. pp. 537–551.

[7] Carvill B.T., Hufton J.R., Anand M., Sircar S. Sorption-enhanced reaction process. *AIChE Journal*. October 1996; 42(10): 2765–2772. Available at: DOI:10.1002/aic.690421008

[8] Hufton J.R., Mayorga S., Sircar S. Sorption-enhanced reaction process for hydrogen production. *AIChE Journal*. February 1999; 45(2): 248–256. Available at: DOI:10.1002/aic.690450205

[9] Ding Y., Alpay E. High temperature recovery of CO₂ from flue gases using hydrotalcite adsorbent. *Process Safety and Environmental Protection*. 2001; 79(1): 45–51. Available at: DOI:10.1205/095758201531130

[10] Ding Y., Alpay E. Adsorption-enhanced steam-methane reforming. *Chemical Engineering Science*. 2000; 55(18): 3929–3940. Available at: DOI:10.1016/S0009-2509(99)00597-7

[11] Brun-Tsekhovoi A.R., Zadorin A.N., Katsobashvili Y.R., Kourdyumov S.S. The process of catalytic steam reforming of hydrocarbons in the presence of carbon dioxide acceptor. *Hydrocarbon Energy Progress VII, Proceedings of the 7th World Hydrogen Energy Conference, Moscow, Russia, 25–29 September*

1988. Oxford: Pergamon Press; 1988. p. 885.

[12] Xu J., Froment GF. Methane steam reforming: Intrinsic Kinetics. *AIChE J.* 1989; 35(1): 88–96. Available at: DOI:10.1002/aic.690350109

[13] Ergun S. Kinetics of the Reaction of Carbon with Carbon Dioxide. *The Journal of Physical Chemistry.* 1 April 1956; 60(4): 480–485. Available at: DOI:10.1021/j150538a022

[14] Kunii D., Levenspiel O. Bubbling Bed Model for Kinetic Processes in Fluidized Beds. *Gas-Solid Mass and Heat Transfer and Catalytic Reactions. Industrial & Engineering Chemistry Process Design and Development.* 1 October 1968; 7(4): 481–492. Available at: DOI:10.1021/i260028a001

[15] Xiu G hua., Li P., Rodrigues AE. Sorption-enhanced reaction process with reactive regeneration. *Chemical Engineering Science.* 2002; 57(18): 3893–3908. Available at: DOI:10.1016/S0009-2509(02)00245-2

[16] Balasubramanian B., Ortiz AL., Kaytakoglu S., Harrison DP. Hydrogen from methane in a single-step process. *Chemical Engineering Science.* 1999; 54(15–16): 3543–3552. Available at: DOI:10.1016/S0009-2509(98)00425-4

[17] Barelli L., Bidini G., Gallorini F., Servili S. Hydrogen production through sorption-enhanced steam methane reforming and membrane technology: A review. *Energy.* April 2008; 33(4): 554–570. Available at: DOI:10.1016/j.energy.2007.10.018

[18] Manovic V., Anthony EJ. Lime-based sorbents for high-temperature CO₂ capture-a review of sorbent modification methods. *International Journal of Environmental Research and Public Health.* 2010; 7(8): 3129–3140. Available at: DOI:10.3390/ijerph7083129

[19] Johnsen K., Ryu HJ., Grace JR., Lim CJ. Sorption-enhanced steam reforming of methane in a fluidized bed reactor with dolomite as CO₂-acceptor. *Chemical Engineering Science.* 2006; 61(4): 1195–1202. Available at:

DOI:10.1016/j.ces.2005.08.022

[20] Harrison DP. Sorption-enhanced hydrogen production: A review. *Industrial and Engineering Chemistry Research*. 2008; 47(17): 6486–6501. Available at: DOI:10.1021/ie800298z

[21] Diglio G., Hanak DP., Bareschino P., Pepe F., Montagnaro F., Manovic V. Modelling of sorption-enhanced steam methane reforming in a fixed bed reactor network integrated with fuel cell. *Applied Energy*. Elsevier; 2018; 210(October 2017): 1–15. Available at: DOI:10.1016/j.apenergy.2017.10.101

[22] Arstad B., Probst J., Blom R. Continuous hydrogen production by sorption enhanced steam methane reforming (SE-SMR) in a circulating fluidized bed reactor: Sorbent to catalyst ratio dependencies. *Chemical Engineering Journal*. Elsevier B.V.; May 2012; 189–190: 413–421. Available at: DOI:10.1016/j.cej.2012.02.057

[23] Aloisi I., di Giuliano A., di Carlo A., Foscolo PU., Courson C., Gallucci K. Sorption enhanced catalytic Steam Methane Reforming: Experimental data and simulations describing the behaviour of bi-functional particles. *Chemical Engineering Journal*. Elsevier B.V.; 2017; 314: 570–582. Available at: DOI:10.1016/j.cej.2016.12.014

[24] García-Lario AL., Aznar M., Martínez I., Grasa GS., Murillo R. Experimental study of the application of a NiO/NiAl₂O₄ catalyst and a CaO-based synthetic sorbent on the Sorption Enhanced Reforming process. *International Journal of Hydrogen Energy*. 2015; 40(1): 219–232. Available at: DOI:10.1016/j.ijhydene.2014.10.033

[25] García-Lario AL., Grasa GS., Murillo R. Performance of a combined CaO-based sorbent and catalyst on H₂ production, via sorption enhanced methane steam reforming. *Chemical Engineering Journal*. 2015; 264: 697–705. Available at: DOI:10.1016/j.cej.2014.11.116

- [26] Diglio G., Hanak DP., Bareschino P., Mancusi E., Pepe F., Montagnaro F., et al. Techno-economic analysis of sorption-enhanced steam methane reforming in a fixed bed reactor network integrated with fuel cell. *Journal of Power Sources*. Elsevier B.V; 2017; 364: 41–51. Available at: DOI:10.1016/j.jpowsour.2017.08.005
- [27] Antzara A., Heracleous E., Bukur DB., Lemonidou AA. Thermodynamic Analysis of Hydrogen Production via Chemical Looping Steam Methane Reforming Coupled with in Situ CO₂ Capture. *Energy Procedia*. Elsevier B.V.; 2014; 63: 6576–6589. Available at: DOI:10.1016/j.egypro.2014.11.694
- [28] Harrison., Ortiz L. Hydrogen production using sorption-enhanced reaction. *Industrial and Engineering Chemistry Research*. 2001; 40(23): 5102–5109.
- [29] Yi BK., Harrison DP. Low-pressure sorption-enhanced hydrogen production. *Industrial and Engineering Chemistry Research*. 2005; 44(6): 1665–1669. Available at: DOI:10.1021/ie048883g
- [30] Antzara A., Heracleous E., Bukur DB., Lemonidou AA. Thermodynamic analysis of hydrogen production via chemical looping steam methane reforming coupled with in situ CO₂ capture. *International Journal of Greenhouse Gas Control*. Elsevier Ltd; January 2015; 32: 115–128. Available at: DOI:10.1016/j.ijggc.2014.11.010
- [31] Madon RH., Fawzi M., Sarwani KI., Osman SA., Razali MA., Mohammad AW. Effect of Steam to Carbon Ratio (S:C) on Steam Methane Reforming's yield over Coated Nickel Aluminide (Ni₃Al) Catalyst in Micro Reactor. *Jurnal Kejuruteraan*. 30 November 2020; 32(4): 657–662. Available at: DOI:10.17576/jkukm-2020-32(4)-14
- [32] Fernandez JR., Abanades JC., Grasa G. Modeling of sorption enhanced steam methane reforming-Part II: Simulation within a novel Ca/Cu chemical loop process for hydrogen production. *Chemical Engineering Science*. 2012; 84: 12–20. Available at: DOI:10.1016/j.ces.2012.07.050

- [33] Grasa GS., Abanades JC., Alonso M., González B. Reactivity of highly cycled particles of CaO in a carbonation/calcination loop. *Chemical Engineering Journal*. 2008; 137(3): 561–567. Available at: DOI:10.1016/j.cej.2007.05.017
- [34] Fennell PS., Pacciani R., Dennis JS., Davidson JF., Hayhurst AN. The effects of repeated cycles of calcination and carbonation on a variety of different limestones, as measured in a hot fluidized bed of sand. *Energy and Fuels*. 2007; 21(4): 2072–2081. Available at: DOI:10.1021/ef060506o
- [35] Grasa GS., Abanades JC. CO₂ capture capacity of CaO in long series of carbonation/calcination cycles. *Industrial and Engineering Chemistry Research*. 2006; 45(26): 8846–8851. Available at: DOI:10.1021/ie0606946
- [36] Manovic V., Anthony EJ. Parametric study on the CO₂ capture capacity of CaO-based sorbents in looping cycles. *Energy and Fuels*. 2008; 22(3): 1851–1857. Available at: DOI:10.1021/ef800011z
- [37] Dou B., Wang C., Song Y., Chen H., Jiang B., Yang M., et al. Solid sorbents for in-situ CO₂ removal during sorption-enhanced steam reforming process: A review. *Renewable and Sustainable Energy Reviews*. Elsevier; 2016; 53: 536–546. Available at: DOI:10.1016/j.rser.2015.08.068
- [38] López JM., Grasa G., Murillo R. Evaluation of the effect of inert support on the carbonation reaction of synthetic CaO-based CO₂ sorbents. *Chemical Engineering Journal*. Elsevier; 2018; 350(May): 559–572. Available at: DOI:10.1016/j.cej.2018.05.014
- [39] Manovic V., Anthony EJ. Thermal activation of CaO-based sorbent and self-reactivation during CO₂ capture looping cycles. *Environmental Science & Technology*. 2008; 42(11): 4170–4174. Available at: DOI:10.1021/es800152s
- [40] Kovanda F., Kováčsová E., Koloušek D. Removal of anions from solution by calcined hydrotalcite and regeneration of used sorbent in repeated calcination-rehydration-anion exchange processes. *Collection of Czechoslovak Chemical*

Communications. 1999; 64(9): 1517–1528. Available at:
DOI:10.1135/cccc19991517

[41] Hanif A., Dasgupta S., Divekar S., Arya A., Garg MO., Nanoti A. A study on high temperature CO₂ capture by improved hydrotalcite sorbents. Chemical Engineering Journal. Elsevier B.V.; 2014; 236: 91–99. Available at: DOI:10.1016/j.cej.2013.09.076

[42] Ding Y., Alpay E. Equilibria and kinetics of CO₂ adsorption on hydrotalcite adsorbent. Chemical Engineering Science. 2000; 55(17): 3461–3474. Available at: DOI:10.1016/S0009-2509(99)00596-5

[43] Hufton JR., Mayorga S., Gaffhey T., Sircar S., Nataraj S. Sorption Enhanced Reaction Process (SERP) for the production of hydrogen. Hydrogen Program Review; United States Department of Energy. 1997. pp. 178–194.

[44] Abanades JC., Rubin ES., Anthony EJ. Sorbent cost and performance in CO₂ capture systems. Industrial and Engineering Chemistry Research. 2004; 43(13): 3462–3466. Available at: DOI:10.1021/ie049962v

[45] Ochoa-Fernández E., Haugen G., Zhao T., Rønning M., Aartun I., Børresen B., et al. Process design simulation of H₂ production by sorption enhanced steam methane reforming: Evaluation of potential CO₂ acceptors. Green Chemistry. 2007; 9(6): 654–662. Available at: DOI:10.1039/b614270b

[46] di Giuliano A., Gallucci K. Sorption enhanced steam methane reforming based on nickel and calcium looping: a review. Chemical Engineering and Processing - Process Intensification. Elsevier; 2018; 130(June): 240–252. Available at: DOI:10.1016/j.cep.2018.06.021

[47] Lysikov AI., Salanov AN., Okunev AG. Change of CO₂ carrying capacity of CaO in isothermal recarbonation-decomposition cycles. Industrial and Engineering Chemistry Research. 2007; 46(13): 4633–4638. Available at: DOI:10.1021/ie0702328

- [48] Manovic V., Anthony EJ. Steam Reactivation of Spent CaO-Based Sorbent for Multiple CO₂ Capture Cycles. *Environmental Science & Technology*. February 2007; 41(4): 1420–1425. Available at: DOI:10.1021/es0621344
- [49] Erans M., Manovic V., Anthony EJ. Calcium looping sorbents for CO₂ capture. *Applied Energy*. 2016; 180: 722–742. Available at: DOI:10.1016/j.apenergy.2016.07.074
- [50] Manovic V., Lu D., Anthony EJ. Steam hydration of sorbents from a dual fluidized bed CO₂ looping cycle reactor. *Fuel*. 2008; 87(15–16): 3344–3352. Available at: DOI:10.1016/j.fuel.2008.04.035
- [51] Froment GF. Modeling of catalyst deactivation. *Applied Catalysis A: General*. 2001; 212(1–2): 117–128. Available at: DOI:10.1016/S0926-860X(00)00850-4
- [52] García-Lario AL., Aznar M., Grasa GS., García T., Murillo R. Study of nickel catalysts for hydrogen production in sorption enhanced reforming process. *Journal of Power Sources*. Elsevier B.V; 2013; 242: 371–379. Available at: DOI:10.1016/j.jpowsour.2013.05.069
- [53] Wei J., Iglesia E. Mechanism and Site Requirements for Activation and Chemical Conversion of Methane on Supported Pt Clusters and Turnover Rate Comparisons among Noble Metals. *Journal of Physical Chemistry B*. 2004; 108(13): 4094–4103. Available at: DOI:10.1021/jp036985z
- [54] Broda M., Manovic V., Imtiaz Q., Kierzkowska AM., Anthony EJ., Müller CR. High-purity hydrogen via the sorption-enhanced steam methane reforming reaction over a synthetic CaO-based sorbent and a Ni catalyst. *Environmental Science & Technology*. 2013; 47(11): 6007–6014. Available at: DOI:10.1021/es305113p
- [55] Li S., Gong J. Strategies for improving the performance and stability of Ni-based catalysts for reforming reactions. *Chemical Society Reviews*. Royal Society

of Chemistry; 2014; 43(21): 7245–7256. Available at: DOI:10.1039/c4cs00223g

[56] Li D., Nakagawa Y., Tomishige K. Methane reforming to synthesis gas over Ni catalysts modified with noble metals. *Applied Catalysis A: General*. Elsevier B.V.; 2011; 408(1–2): 1–24. Available at: DOI:10.1016/j.apcata.2011.09.018

[57] Clough PT., Boot-Handford ME., Zheng L., Zhang Z., Fennell PS. Hydrogen production by sorption enhanced steam reforming (SESR) of biomass in a fluidised-bed reactor using combined multifunctional particles. *Materials*. MDPI AG; 21 May 2018; 11(5). Available at: DOI:10.3390/ma11050859

[58] Nkulikiyinka P., Wagland ST., Manovic V., Clough PT. Prediction of Combined Sorbent and Catalyst Materials for SE-SMR, Using QSPR and Multitask Learning. *Industrial & Engineering Chemistry Research*. 6 July 2022; 61(26): 9218–9233. Available at: DOI:10.1021/acs.iecr.2c00971

[59] Dang C., Liu L., Yang G., Cai W., Long J., Yu H. Mg-promoted Ni-CaO microsphere as bi-functional catalyst for hydrogen production from sorption-enhanced steam reforming of glycerol. *Chemical Engineering Journal*. Elsevier; 2020; 383(July 2019): 123204. Available at: DOI:10.1016/j.cej.2019.123204

[60] Wang C., Dou B., Jiang B., Song Y., Du B., Zhang C., et al. Sorption-enhanced steam reforming of glycerol on Ni-based multifunctional catalysts. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2015; 40(22): 7037–7044. Available at: DOI:10.1016/j.ijhydene.2015.04.023

[61] Wang X., He Y., Xu T., Xiao B., Liu S., Hu Z., et al. CO₂ sorption-enhanced steam reforming of phenol using Ni–M/CaO–Ca₁₂Al₁₄O₃₃ (M = Cu, Co, and Ce) as catalytic sorbents. *Chemical Engineering Journal*. Elsevier; 2020; 393(November 2019): 124769. Available at: DOI:10.1016/j.cej.2020.124769

[62] Satrio JA., Shanks BH., Wheelock TD. A combined catalyst and sorbent for enhancing hydrogen production from coal or biomass. *Energy and Fuels*.

2007; 21(1): 322–326. Available at: DOI:10.1021/ef060132z

[63] Rahmanzadeh L., Taghizadeh M. Sorption-enhanced ethanol steam reforming on Ce-Ni/MCM-41 with simultaneous CO₂ adsorption over Na- and Zr-promoted CaO based sorbent. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2019; 44(39): 21238–21250. Available at: DOI:10.1016/j.ijhydene.2019.04.289

[64] Hosseini Abbandanak M., Taghizadeh M., Fallah N. High-purity hydrogen production by sorption-enhanced methanol steam reforming over a combination of Cu–Zn–CeO₂–ZrO₂/MCM-41 catalyst and (Li–Na–K) NO₃·MgO adsorbent. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2020; (3): 7099–7112. Available at: DOI:10.1016/j.ijhydene.2020.11.250

[65] Ertl G., Knözinger H., Weitkamp J. (eds.) *Preparation of Solid Catalysts*. Wiley; 1999. Available at: DOI:10.1002/9783527619528

[66] Li C., Chen YW. Temperature-programmed-reduction studies of nickel oxide/alumina catalysts: effects of the preparation method. *Thermochimica Acta*. 1995; 256(2): 457–465. Available at: DOI:10.1016/0040-6031(94)02177-P

[67] Lee J., Hong S., Cho H., Lyu B., Kim M., Kim J., et al. Machine learning-based energy optimization for on-site SMR hydrogen production. *Energy Conversion and Management*. Elsevier Ltd; 15 September 2021; 244. Available at: DOI:10.1016/j.enconman.2021.114438

[68] Yu X., Shen Y., Guan Z., Zhang D., Tang Z., Li W. Multi-objective optimization of ANN-based PSA model for hydrogen purification from steam-methane reforming gas. *International Journal of Hydrogen Energy*. March 2021; 46(21): 11740–11755. Available at: DOI:10.1016/j.ijhydene.2021.01.107

[69] Pajak M., Buchaniec S., Kimijima S., Szmyd JS., Brus G. A multiobjective optimization of a catalyst distribution in a methane/steam reforming reactor using a genetic algorithm. *International Journal of Hydrogen Energy*. June 2021; 46(38):

20183–20197. Available at: DOI:10.1016/j.ijhydene.2020.02.228

[70] Nkulikiyinka P., Yan Y., Güleç F., Manovic V., Clough PT. Prediction of sorption enhanced steam methane reforming products from machine learning based soft-sensor models. *Energy and AI*. November 2020; 2: 100037. Available at: DOI:10.1016/j.egyai.2020.100037

[71] Ke C., He W., Liu S., Ru X., Liu S., Lin Z. Multiscale Catalyst Design for Steam Methane Reforming Assisted by Deep Learning. *Journal of Physical Chemistry C*. American Chemical Society; 27 May 2021; 125(20): 10860–10867. Available at: DOI:10.1021/acs.jpcc.1c01632

[72] Hong S., Lee J., Cho H., Kim M., Moon I., Kim J. Multi-objective optimization of CO₂ emission and thermal efficiency for on-site steam methane reforming hydrogen production process using machine learning. *Journal of Cleaner Production*. Elsevier Ltd; 20 July 2022; 359. Available at: DOI:10.1016/j.jclepro.2022.132133

[73] Kesgin U. Genetic algorithm and artificial neural network for engine optimisation of efficiency and emission. 2004; 83: 885–895. Available at: DOI:10.1016/j.fuel.2003.10.025

[74] Nowak W. Neurocomputing approach for the prediction of NO_x emissions from CFBC in air-fired and oxygen-enriched atmospheres. 2017; 97(2): 75–84.

[75] Ardabili SF., Najafi B., Shamshirband S., Bidgoli BM., Deo RC., Chau KW. Computational intelligence approach for modeling hydrogen production: A review. *Engineering Applications of Computational Fluid Mechanics*. 2018; 12(1): 438–458. Available at: DOI:10.1080/19942060.2018.1452296

[76] Rosenblatt F. Principles of neurodynamics. Perceptrons and the theory of brain mechanisms. New York: CORNELL AERONAUTICAL LAB INC; 1961.

[77] Ardabili SF., Najafi B., Shamshirband S., Bidgoli BM., Deo RC., Chau KW. Computational intelligence approach for modeling hydrogen production: A review.

Engineering Applications of Computational Fluid Mechanics. 2018; 12(1): 438–458. Available at: DOI:10.1080/19942060.2018.1452296

[78] Brouwer RK. A Hybrid Neural Network for Input That Is Both Categorical and Quantitative. 2004; 19: 979–1001.

[79] di Giuliano A., Gallucci K., Foscolo PU. Determination of Kinetic and Diffusion Parameters Needed to Predict the Behavior of CaO-Based CO₂ Sorbent and Sorbent-Catalyst Materials. Industrial and Engineering Chemistry Research. 2020; 59(15): 6840–6854. Available at: DOI:10.1021/acs.iecr.9b05383

[80] Aloisi I., Jand N., Stendardo S., Foscolo PU. Hydrogen by sorption enhanced methane reforming: A grain model to study the behavior of bi-functional sorbent-catalyst particles. Chemical Engineering Science. Elsevier; 2016; 149: 22–34. Available at: DOI:10.1016/j.ces.2016.03.042

[81] di Giuliano A., Gallucci K., Giancaterino F., Courson C., Foscolo PU. Multicycle sorption enhanced steam methane reforming with different sorbent regeneration conditions: Experimental and modelling study. Chemical Engineering Journal. Elsevier; 2019; 377(September 2018): 119874. Available at: DOI:10.1016/j.cej.2018.09.035

[82] Rout KR., Jakobsen HA. A numerical study of pellets having both catalytic- and capture properties for SE-SMR process: Kinetic- and product layer diffusion-controlled regimes. Fuel Processing Technology. Elsevier B.V.; 2013; 106: 231–246. Available at: DOI:10.1016/j.fuproc.2012.07.029

[83] di Carlo A., Aloisi I., Jand N., Stendardo S., Foscolo PU. Sorption enhanced steam methane reforming on catalyst-sorbent bifunctional particles: A CFD fluidized bed reactor model. Chemical Engineering Science. Elsevier Ltd; 2017; 173: 428–442. Available at: DOI:10.1016/j.ces.2017.08.014

[84] Martínez I., Grasa G., Meyer J., di Felice L., Kazi S., Sanz C., et al. Performance and operating limits of a sorbent-catalyst system for sorption-

enhanced reforming (SER) in a fluidized bed reactor. *Chemical Engineering Science*. 2019; 205: 94–105. Available at: DOI:10.1016/j.ces.2019.04.029

[85] Albrecht KO., Satrio JA., Shanks BH., Wheelock TD. Application of a combined catalyst and sorbent for steam reforming of methane. *Industrial and Engineering Chemistry Research*. 2010; 49(9): 4091–4098. Available at: DOI:10.1021/ie901914c

[86] García-Lario AL., Grasa GS., Murillo R. Performance of a combined CaO-based sorbent and catalyst on H₂ production, via sorption enhanced methane steam reforming. *Chemical Engineering Journal*. 2015; 264: 697–705. Available at: DOI:10.1016/j.cej.2014.11.116

[87] Chanburanasiri N., Ribeiro AM., Rodrigues AE., Arpornwichanop A., Laosiripojana N., Praserttham P., et al. Hydrogen production via sorption enhanced steam methane reforming process using Ni/CaO multifunctional catalyst. *Industrial and Engineering Chemistry Research*. 2011; 50(24): 13662–13671. Available at: DOI:10.1021/ie201226j

[88] Borhani TN., García-Muñoz S., Vanesa Luciani C., Galindo A., Adjiman CS. Hybrid QSPR models for the prediction of the free energy of solvation of organic solute/solvent pairs. *Physical Chemistry Chemical Physics*. 2019; 21(25): 13706–13720. Available at: DOI:10.1039/C8CP07562J

[89] Borhani TNG., Saniedanesh M., Bagheri M., Lim JS. QSPR prediction of the hydroxyl radical rate constant of water contaminants. *Water Research*. Elsevier Ltd; 2016; 98: 344–353. Available at: DOI:10.1016/j.watres.2016.04.038

[90] Consonni V., Todeschini R. Molecular descriptors. *Challenges and Advances in Computational Chemistry and Physics*. 2010. 29–102 p. Available at: DOI:10.1007/978-1-4020-9783-6_3

[91] Borhani TN., García-Muñoz S., Vanesa Luciani C., Galindo A., Adjiman CS. Hybrid QSPR models for the prediction of the free energy of solvation of

organic solute/solvent pairs. *Physical Chemistry Chemical Physics*. 2019; 21(25): 13706–13720. Available at: DOI:10.1039/c8cp07562j

[92] Borhani TNG., Bagheri M., Manan ZA. Molecular modeling of the ideal gas enthalpy of formation of hydrocarbons. *Fluid Phase Equilibria*. Elsevier B.V.; 2013; 360: 423–434. Available at: DOI:10.1016/j.fluid.2013.09.066

[93] Burner J., Schwiedrzik L., Krykunov M., Luo J., Boyd PG., Woo TK. High-Performing Deep Learning Regression Models for Predicting Low-Pressure CO₂ Adsorption Properties of Metal-Organic Frameworks. *Journal of Physical Chemistry C*. 2020; Available at: DOI:10.1021/acs.jpcc.0c06334

[94] Yang P., Lu G., Yang Q., Liu L., Lai X., Yu D. Analyzing Acetylene Adsorption of Metal-Organic Frameworks Based on Machine Learning. *Green Energy & Environment*. Institute of Process Engineering, Chinese Academy of Sciences; January 2021; : 108947. Available at: DOI:10.1016/j.gee.2021.01.006

[95] Li J., Xu K., Yao X., Chen S. Prediction and optimization of syngas production from steam gasification: Numerical study of operating conditions and biomass composition. *Energy Conversion and Management*. Elsevier Ltd; 15 May 2021; 236. Available at: DOI:10.1016/j.enconman.2021.114077

[96] Wang B., Zhou L., Xu K., Wang Q. Prediction of minimum ignition energy from molecular structure using quantitative structure-property relationship (QSPR) models. *Industrial and Engineering Chemistry Research*. American Chemical Society; 11 January 2017; 56(1): 47–51. Available at: DOI:10.1021/acs.iecr.6b04347

[97] Giappa RM., Tylanakis E., di Gennaro M., Gkagkas K., Froudakis GE. A combination of multi-scale calculations with machine learning for investigating hydrogen storage in metal organic frameworks. *International Journal of Hydrogen Energy*. Elsevier Ltd; 5 August 2021; 46(54): 27612–27621. Available at: DOI:10.1016/j.ijhydene.2021.06.021

- [98] Thomas KM. Adsorption and desorption of hydrogen on metal-organic framework materials for storage applications: Comparison with other nanoporous materials. *Dalton Transactions*. 2009; (9): 1487–1505. Available at: DOI:10.1039/b815583f
- [99] Villajos JA. Experimental Volumetric Hydrogen Uptake Determination at 77 K of Commercially Available Metal-Organic Framework Materials. *C. MDPI AG*; 5 January 2022; 8(1): 5. Available at: DOI:10.3390/c8010005
- [100] Broom DP., Webb CJ., Fanourgakis GS., Froudakis GE., Trikalitis PN., Hirscher M. Concepts for improving hydrogen storage in nanoporous materials. *International Journal of Hydrogen Energy*. Elsevier Ltd; 22 March 2019; 44(15): 7768–7779. Available at: DOI:10.1016/j.ijhydene.2019.01.224
- [101] Saraci F., Quezada-Novoa V., Donnarumma PR., Howarth AJ. Rare-earth metal-organic frameworks: From structure to applications. *Chemical Society Reviews*. Royal Society of Chemistry; 2020. pp. 7949–7977. Available at: DOI:10.1039/d0cs00292e
- [102] Howarth AJ., Liu Y., Li P., Li Z., Wang TC., Hupp JT., et al. Chemical, thermal and mechanical stabilities of metal-organic frameworks. *Nature Reviews Materials*. Nature Publishing Group; 2016. Available at: DOI:10.1038/natrevmats.2015.18
- [103] Hendon CH., Rieth AJ., Korzyński MD., Dincă M. Grand Challenges and Future Opportunities for Metal-Organic Frameworks. *ACS Central Science*. American Chemical Society; 28 June 2017; 3(6): 554–563. Available at: DOI:10.1021/acscentsci.7b00197
- [104] Colón YJ., Gómez-Gualdrón DA., Snurr RQ. Topologically Guided, Automated Construction of Metal-Organic Frameworks and Their Evaluation for Energy-Related Applications. *Crystal Growth and Design*. American Chemical Society; 1 November 2017; 17(11): 5801–5810. Available at: DOI:10.1021/acs.cgd.7b00848

- [105] Yun J-H., He Y., Otero M., Düren T., Seaton NA. Adsorption equilibrium of polar/non-polar mixtures on MCM-41: experiments and Monte Carlo simulation. 2002. pp. 685–692. Available at: DOI:10.1016/S0167-2991(02)80197-5
- [106] Gómez-Gualdrón DA., Colón YJ., Zhang X., Wang TC., Chen YS., Hupp JT., et al. Evaluating topologically diverse metal-organic frameworks for cryo-adsorbed hydrogen storage. *Energy and Environmental Science*. Royal Society of Chemistry; 1 October 2016; 9(10): 3279–3289. Available at: DOI:10.1039/c6ee02104b
- [107] Bobbitt NS., Chen J., Snurr RQ. High-Throughput Screening of Metal-Organic Frameworks for Hydrogen Storage at Cryogenic Temperature. *Journal of Physical Chemistry C*. American Chemical Society; 8 December 2016; 120(48): 27328–27341. Available at: DOI:10.1021/acs.jpcc.6b08729
- [108] Liu B., Yang Q., Xue C., Zhong C., Chen B., Smit B. Enhanced adsorption selectivity of hydrogen/methane mixtures in metal-organic frameworks with interpenetration: A molecular simulation study. *Journal of Physical Chemistry C*. 3 July 2008; 112(26): 9854–9860. Available at: DOI:10.1021/jp802343n
- [109] Glasby LT., Moghadam PZ. Hydrogen storage in MOFs: Machine learning for finding a needle in a haystack. *Patterns*. Cell Press; 9 July 2021; 2(7). Available at: DOI:10.1016/j.patter.2021.100305
- [110] García-Holley P., Schweitzer B., Islamoglu T., Liu Y., Lin L., Rodriguez S., et al. Benchmark Study of Hydrogen Storage in Metal-Organic Frameworks under Temperature and Pressure Swing Conditions. *ACS Energy Letters*. American Chemical Society; 2018. pp. 748–754. Available at: DOI:10.1021/acsenergylett.8b00154
- [111] Altintas C., Altundal OF., Keskin S., Yildirim R. Machine Learning Meets with Metal Organic Frameworks for Gas Storage and Separation. *Journal of Chemical Information and Modeling*. American Chemical Society; 2021. pp. 2131–2146. Available at: DOI:10.1021/acs.jcim.1c00191

- [112] Ahmed A., Seth S., Purewal J., Wong-Foy AG., Veenstra M., Matzger AJ., et al. Exceptional hydrogen storage achieved by screening nearly half a million metal-organic frameworks. *Nature Communications*. Nature Publishing Group; 1 December 2019; 10(1). Available at: DOI:10.1038/s41467-019-09365-w
- [113] Ahmed A., Liu Y., Purewal J., Tran LD., Wong-Foy AG., Veenstra M., et al. Balancing gravimetric and volumetric hydrogen density in MOFs. *Energy and Environmental Science*. Royal Society of Chemistry; 1 November 2017; 10(11): 2459–2471. Available at: DOI:10.1039/c7ee02477k
- [114] Bucior BJ., Bobbitt NS., Islamoglu T., Goswami S., Gopalan A., Yildirim T., et al. Energy-based descriptors to rapidly predict hydrogen storage in metal-organic frameworks. *Molecular Systems Design and Engineering*. Royal Society of Chemistry; 1 February 2019; 4(1): 162–174. Available at: DOI:10.1039/c8me00050f
- [115] Anderson G., Schweitzer B., Anderson R., Gómez-Gualdrón DA. Attainable Volumetric Targets for Adsorption-Based Hydrogen Storage in Porous Crystals: Molecular Simulation and Machine Learning. *Journal of Physical Chemistry C*. American Chemical Society; 1 October 2019; 123(1): 120–130. Available at: DOI:10.1021/acs.jpcc.8b09420

3 Machine learning for Carbon Capture, Utilisation, And Storage (CCUS)

This chapter contains excerpts the publication (DOI: <https://doi.org/10.1039/D1EE02395K>) on ‘**Harnessing the power of machine learning for carbon capture, utilisation, and storage (CCUS) – a state-of-the-art review**’ Yongliang Yan^{1,2}, Tohid N. Borhani³, Sai Gokul Subraveti⁴, Kasturi Nagesh Pai⁴, Vinay Prasad⁴, Arvind Rajendran⁴, **Paula Nkulikiyinka**¹, Jude Odianosen Asibor¹, Zhien Zhang⁵, Ding Shao⁶, Lijuan Wang⁶, Wenbiao Zhang⁶, Yong Yan⁷, William Ampomah⁸, Junyu You^{8,9}, Meihong Wang¹⁰, Edward J. Anthony¹, Vasilije Manovic¹, Peter T. Clough^{1*}.

¹ Energy and Power Theme, Cranfield University, Bedfordshire, MK43 0AL

² Materials, Concept and Reaction Engineering (MatCoRE) Group, School of Engineering, Newcastle University, Merz Court, Newcastle Upon Tyne, NE1 7RU, UK

³ School of Engineering, Division of Chemical Engineering, University of Wolverhampton, Wolverhampton, WV1 1LY, United Kingdom

⁴ Department of Chemical and Materials Engineering, Donadeo Innovation Centre for Engineering, University of Alberta, 9211-116 Street NW, Edmonton, Alberta, Canada T6G 1H9

⁵ Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV 26506, USA

⁶ School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, P.R. China

⁷ School of Engineering and Digital Arts, University of Kent, Canterbury, Kent CT2 7NT, U.K.

⁸ Petroleum Recovery Research Centre, New Mexico Tech, Socorro NM, 87801, USA

⁹ School of Petroleum and Natural Gas Engineering, Chongqing University of Science and Technology, Chongqing, 401331, China

¹⁰ Department of Chemical and Biological Engineering, University of Sheffield, Sheffield S1 3JD, United Kingdom

* Corresponding Author: Dr Peter T. Clough, Email: p.t.clough@cranfield.ac.uk

Text from the publication by co-authors is denoted in quotation marks “ ”, otherwise the work is my own.

3.1 Introduction

“Carbon Capture and Storage (CCS) will play a critical role in future decarbonisation efforts to meet the Paris Agreement targets and mitigate the worst effects of climate change. Whilst there are many well developed CCS technologies there is the potential for improvement that can encourage CCS deployment. A time and cost-efficient way of advancing CCS is through the

application of machine learning (ML). ML is a collective term for high-level statistical tools and algorithms that can be used to classify, predict, optimise and cluster data.

The use of machine learning (ML) has increased for a multitude of applications due to the growth in computing power in recent years, as well as for CCS applications as well. ML offers the potential to identify links between data/results that aren't readily identifiable, and it also provides alternative lower computing cost pathways. Within the field of CCS, ML has begun to be utilised to evaluate new CO₂ sorbents and oxygen carrier materials [1], simulate, control and operate capture processes [2–7], simplify process economics, predict CO₂ solubilities in solvents and CO₂ capture capacities in adsorbents [8–10], improve the accuracy of multiphase flowmeters used for CO₂ pipelines [11], and predict leaks from CO₂ wells [12]; each with the aim of advancing the field of CCS in a cost and time effective manner. Nevertheless, ML in CCS is still relatively new and there is much yet to be studied.”

3.2 Machine Learning Algorithms

“ML is a subset of artificial intelligence (AI) that involves the study of computer algorithms that allow programs to automatically improve through experience [13,14]. Advantages include ease of trends and pattern identification, minimal human intervention (automation), ability to improve continuously, as well as high efficiency in the handling of multi-dimensional and multi-variety data [13,15]. Its application is however sometimes limited by factors such as ethics, lack of physical constraint, data availability and quality, misapplication as well as interpretability [16]. Figure 3-1 presents the types of ML and respective areas of application, [17].”

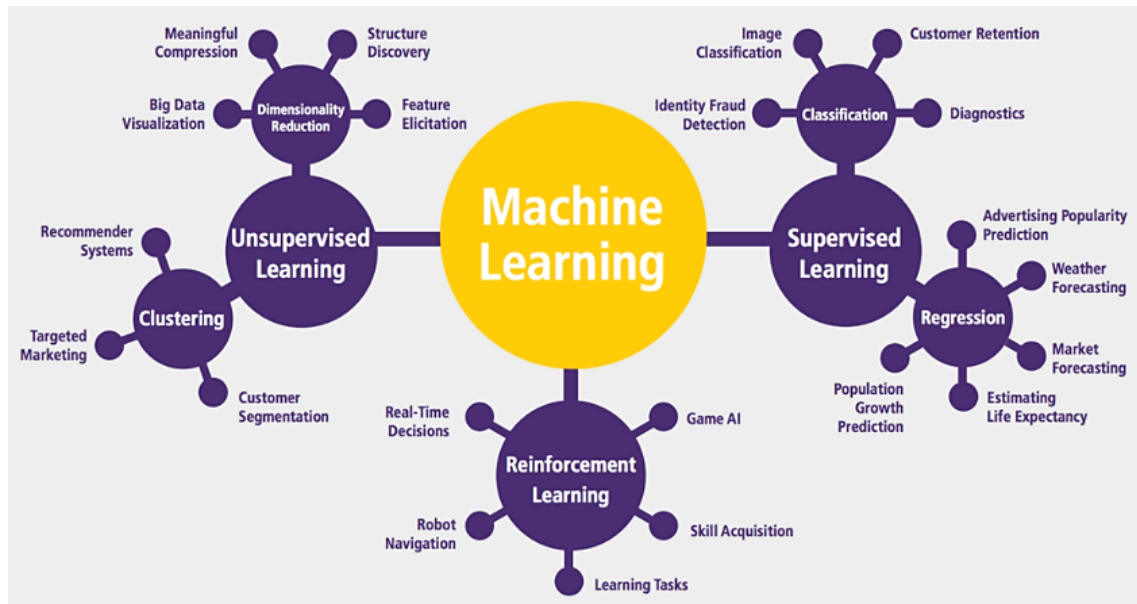


Figure 3-1 Types and applications of machine learning

“There are three main types of ML: supervised, unsupervised and reinforcement learning. The supervised ML, which is the most commonly used of the three is usually applied when the input-output data is known. It involves training the ML models to learn the relationship between the given inputs and associated output values [18]. If the available dataset consists of only input values (no labels), unsupervised ML can be used in an attempt to identify trends, structure, patterns or clustering in the input data [19]. Reinforcement learning is an ML technique that enables an agent to learn in an interactive environment by trial and error using feedback from its actions and experiences [20]. The execution of any of the types of ML can be done through the application of the appropriate algorithm. A brief description of common ML algorithms is presented in Table 3-1, [17,21].

Table 3-1 Common machine learning algorithms

Algorithm name	Task type	Description
Linear regression	Regression	By fitting a linear model with coefficients, this algorithm correlates each data feature to the output, thus assisting in predicting future values.
Logistic regression	Classification	A classification algorithm that predicts the likelihood of a dependent variable (usually binary) belonging to a category.
Decision tree	Regression and classification	This interpretable algorithm performs by splitting values of data features into branches at decision nodes until a final decision output is established.
Naïve Bayes	Regression and classification	This algorithm is based on the Bayesian theorem which updates the prior knowledge of an event with the independent probability of each feature that can affect the event.
Support Vector Machines (SVMs)	Regression, classification, and outlier detection	This algorithm operates by transforming the required data and determining the optimal boundary (hyperplane) between the various outputs.
Random forest	Regression and classification	The algorithm is an ensemble of decision trees characterised by improved accuracy. It operates by generating a multitude of decision trees and uses either the modal vote or average prediction for classification or regression tasks respectively.
Artificial Neural Network (ANN)	Regression, classification, and clustering	This algorithm which is modelled after the biological neurons of the brain consists of several layers with interconnected artificial neurons performing various data transformations to obtain the required output.
K-means clustering	Clustering	This centroid-based algorithm clusters unlabelled data points by their similarity of characteristics determined by the model without human interference.
Hierarchical clustering	Clustering	This algorithm splits clusters along a hierarchical tree to form a classification system.
Gaussian mixture model	Clustering	This unsupervised algorithm clusters data by estimating the density distribution of the dataset.
AdaBoost	Regression and classification	This is an ensemble algorithm that combines multiple weak algorithms to obtain an improved output.
Principal Component Analysis (PCA)	Dimension reduction	This algorithm is often used to reduce the dimensionality of large data sets without distorting its characteristics. (Though not strictly a ML algorithm on its own.)

Other ML algorithms include K-Nearest neighbour, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), recommender systems, genetic algorithm, gradient boosting trees, particle swarm algorithms. The choice of algorithm is dependent on the characteristics of the available dataset and required output or task [18]. It is important to note that in many applications, multiple algorithms are usually combined (referred to as ensemble algorithms) to improve model performance accuracy and robustness.”

3.3 Machine Learning in CO₂ Adsorption

“Adsorbents are micro-porous structures with a characteristically large surface area and the ability to capture large amounts of gases on their surface [22]. They generally have a selective affinity for specific gases in a mixture of gases, making them ideal for gas separation applications such as CO₂ capture [22,23]. One of the primary considerations when designing an adsorbent-based CO₂ capture process is the choice of the adsorbent media [24]. This field has gone through a renaissance in recent years with the advent of organometallic chemistry [8,25,26]. There are several new classes of adsorbents such as Metal-Organic Frameworks (MOFs) [25–27], Covalent Organic Frameworks (COFs) [28], Zeolitic Imidazolate Frameworks (ZIFs) [29], Porous Organic Cages (POCs) [30] along with the classical Zeolites [22,31] and activated carbons. Many of these porous structures are chemically and physically tuneable and can be reverse engineered to provide the process designer with tailor-made options [32–34]. This means an effectively infinite number of possible structures can be theorised [8,35]. Exploring the entire adsorbent material design space is computationally restrictive, and traditional adsorbent characterisation techniques are time-consuming, adding to the complexity [35,36]. Large databases with over one million of such real and *in-silico* hypothetical porous structures are available to process designers that are already partially characterised for the application of CO₂ capture [8,37–43].”

3.3.1 Machine Learning in Calcium Looping

Calcium looping is a CO₂ capture process, that uses calcium oxide-based sorbents to separate and remove CO₂ from flue gases. The process is based on the reversible reaction of lime with CO₂ and is considered as an emerging CO₂ capture technology. This process has been well researched with findings focusing on optimal CaO based sorbents to achieve the best capture efficiency, however the application of ML to this field is relatively new, with very few studies on this.

Chen *et al.* [44] proposed the use of BPNN to predict the performance of Ca-based sorbents in the calcination/carbonation cycles, based on TGA experimental data. This study investigated the factors that affected the sorbent performance, namely sample particle diameter, calcination temperature, calcination duration, calcination atmosphere and carbonation duration. The feed-forward multilayer ANN, which had the architecture of 5-34-1, had the five aforementioned factors as inputs, and the carbonation conversion degree as the output parameter, calculated with the assumption that the decomposition of calcium carbonation was the only reason for sample weight change. Here, 75% of the data was used for training while the remaining 25% was accounted as the test data. The model proposed showed strong correlation with TGA results and proved the validity for the approximation of Ca-based sorbent in the carbonation process even when conducted at extreme reaction condition.

A recent application of ML to the calcium looping process was developed by Nkulikiyinka *et al.* [45]. Here, the authors developed an ANN and random forest (RF) model to act as soft sensor models, for the prediction of gas concentrations for the reaction of steam methane reforming coupled with calcium looping, also known as sorption enhanced steam methane reforming (SE-SMR). In this study, the data was obtained using the Aspen Plus software, where input parameters, regenerator and reformer temperatures, pressure, steam-to-carbon ratio and sorbent-to-carbon ratio, were varied to obtain a wide range of data for the process. The Aspen Plus data was validated against literature data, and was then

split into training, validation and test data. Various gas concentrations in the reformer and regenerator, as well as methane conversion were used as the output parameters. The models developed displayed high accuracy prediction for the reactor gas concentrations and confirmed that ANN and RF algorithms can successfully model a nonlinear process such as SE-SMR, and therefore can act as a suitable data-driven soft sensor for the process.

Krzywanski *et al.* [46] explored a method of predicting the NO_x emissions produced from the regenerator of a calcium looping system, coupled with oxyfuel combustion of coal to provide heat of decomposition, using a regression analysis-based modelling technique. The authors conducted the experiment in a dual-fluidized bed (DFB), with the effects of fuel type, oxygen feed, and NO addition to primary or secondary feed gas, being evaluated. The authors provided limited detail on the regression model, however Figure 3-2, [46], shows the flowchart of the model application, and the only inputs necessary are the fixed carbon, the molar ratio of nitrogen to carbon content in fuel (N/C), and the O₂ concentration in the flue gas from the regenerator, leading to the NO_x emission as the output parameter. The results obtained from the model were in good agreement with experimental results, with a correlation coefficient equal to 0.925. They confirmed that the developed model could readily be applied by scientists as well as engineers for future optimization purposes.



Figure 3-2 Application of the model for the evaluation of NO_x concentration if flue gas.

An alternate purpose that ML has been applied to in the calcium looping field, is on the study of the economic feasibility of the post-combustion calcium looping process on a 580 MW coal fired power plant, by Hanak and Manovic [47]. In this study an ANN was developed using data from Aspen Plus simulations, and this model was then combined with results from an economic model developed from a Monte Carlo (MC) simulation. The ANN model was used to connect the process inputs of the process model with the process inputs of the economic model. A

two-layer feedforward ANN with ten sigmoid hidden neurons and linear output neurons was developed, with 70% of the data obtained from the Aspen Plus model, used for training, 15% used for validation and 15% used for testing. Figure 3-3, [47], shows that the ANN used in this study is capable of depicting the thermodynamic performance of the calcium looping retrofit accurately, despite the nonlinear characteristic. The study concluded that the stochastic approach, and incorporation of the ANN model in the economic feasibility assessment enables a more accurate and reliable comparison of different calcium looping retrofit configurations.”

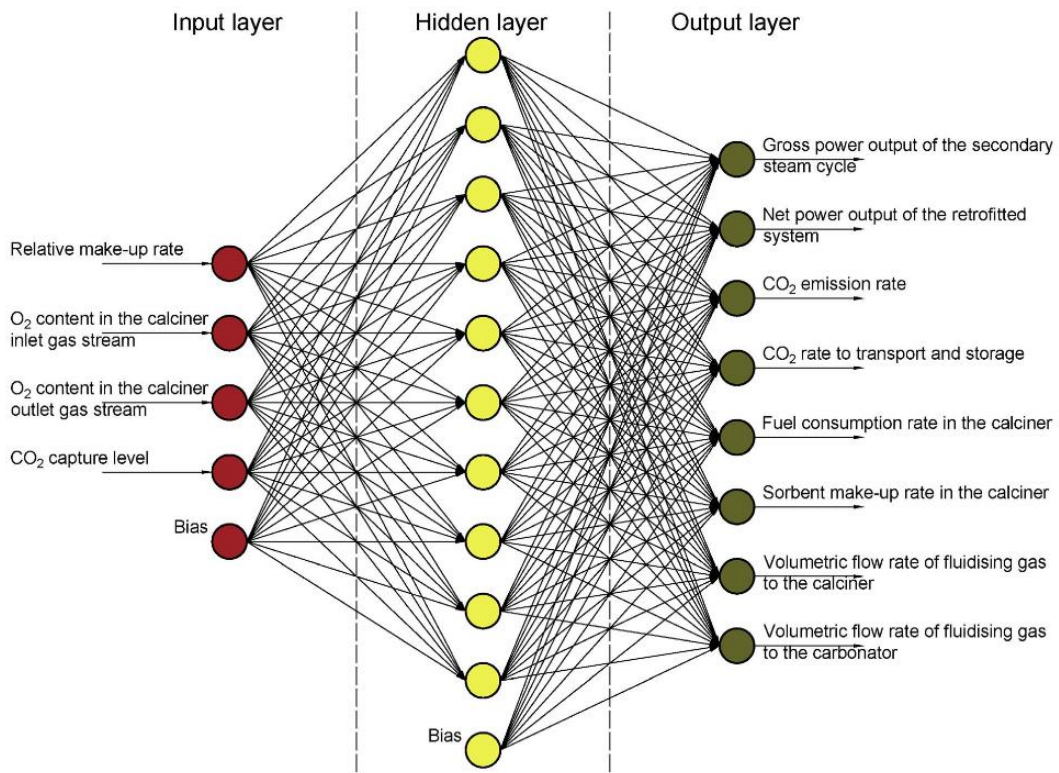


Figure 3-3 Structure of the artificial neural network used to map the thermodynamic performance of the calcium looping process retrofit.

3.4 Conclusion

“For ML in CO₂ adsorption, the research is focused on applying ML in adsorbent synthesis and characterisation, process modelling and optimisation, and process inversion. It suggests that ML is a powerful tool for screening solvents and adsorbents, as well as process modelling and optimization, which can reduce the

development time, capital and operating costs for CO₂ capture. There are some studies available that utilise ML models for calcium looping and/or chemical looping combustion for CO₂ capture and is an area that requires more work. Some researchers have started to apply ML to predict the performance of oxygen carriers and Ca-based sorbents, as well as process control and techno-economic assessments. The findings so far in ML use for CO₂ absorption and adsorption, is that it can be adapted to the calcium looping and chemical looping combustion for CO₂ capture. For instance, using QSPR to find the optimal properties of oxygen carriers and Ca-based sorbents for CO₂ separation. ML is also expected to play a vital role in the development of CO₂ utilisation technologies, such as screening the catalysts for CO₂ catalytic or electro-catalytic conversion, combined with the DFT calculations, and predicting suitable microalgae types and optimal cultivation conditions for carbon fixation.

The distinct advantages of applying ML in CCS are that it provides the potential to identify links between data/results that aren't readily identifiable, and it also provides alternative lower computing cost pathways. Researchers in CCS can apply ML to accelerate the design and development of materials for CO₂ separation and conversion, measure the multiphase CO₂ flow, evaluate the trapping mechanisms for CO₂ storage, and develop the surrogate model for process optimisation and uncertainty analysis. It is important to mention that ML is a data-driven method, which always requires a large quantity of data to develop generalised and robust models. The quality of training dataset, the selections of input-output features and the type of ML algorithms play a vital role to develop a comprehensive model. As mentioned before, some researchers have illustrated suitable methods for feature selection, avoiding overfitting, and issues with small datasets, when applying ML in CCS. With the development of ML in CCS, it is expected that ML will be an efficient tool to accelerate the development of cost-effective CCS systems to tackle the climate change.

Regarding further topics discussed in the full review, the authors made the following recommendations for future work and research:

1. “Education of future generations in ML techniques and CCS at undergraduate and graduate levels is incredibly important and something that is not always part of mainstream curriculums in engineering. We recommend ML and CCS take a greater emphasis in Higher Education practice.
2. Greater emphasis should also be placed on transferable learning-focused methods, so that models do not need to be retrained for each material and/or process. Generalised models, which can infer functional information should be explored in CCS.
3. The development of combined models for materials and process optimisation (performed simultaneously) would prove useful for deployment of CCS technologies at commercial scale.
4. Utilising uncertainty assessment methods to accurately estimate the potential value proposition of CCS systems.
5. Performing more detailed investigations on the effect of process scale (in capture/utilisation). Do models/designs/optimisation conducted at lab/pilot scales hold at industrial scales, or will models need to be retrained and optimisation redone during scaling up? Can ML models be truly multi-scale (accounting for chemical properties of materials to overall reactor performance) in their CCS applications? This will need increased collaboration with industrial partners.
6. Developing hybrid ML methods that find ways to incorporate intuition/domain knowledge to compensate for a lack of data.
7. Process control is challenging in many CCS (and other chemical) processes, more work needs to be conducted to understand if ML can be applied to improve process control.
8. Finally, we recommend that when ML research is conducted in CCS, then the training data and ML models should be made publicly accessible in the open domain to enable greater take up and deployment.”

3.5 This Study

Considering these recommendations, this thesis has attempted to address some, for example, points 2, 3, 6, and 7.

More specifically, Chapter 4 addresses point 7, in producing a machine learning methodology with the ability to predict gas concentrations in a reactor when hardware sensors are not available, therefore allowing for better process control where necessary.

Chapter 5 partially addresses point 3, and point 6, in terms of developing a machine learning model that can predict the behaviour of a combined material using single and multitask learning, via datamining to create two separate databases.

Chapter 6 also partially addresses point 3, by utilising machine learning and patterns in data, for the identification of optimal hydrogen storage MOF materials.

3.6 References

- [1] Yan Y., Mattisson T., Moldenhauer P., Anthony EJ., Clough PT. Applying machine learning algorithms in estimating the performance of heterogeneous, multi-component materials as oxygen carriers for chemical-looping processes. *Chemical Engineering Journal*. Elsevier B.V.; 1 May 2020; 387: 124072. Available at: DOI:10.1016/j.cej.2020.124072
- [2] Bai Z., Li F., Zhang J., Oko E., Wang M., Xiong Z., et al. Modelling of a Post-combustion CO₂ Capture Process Using Bootstrap Aggregated Extreme Learning Machines. *Computer Aided Chemical Engineering*. Elsevier; 1 January 2016; 38: 2007–2012. Available at: DOI:10.1016/B978-0-444-63428-3.50339-8
- [3] Li F., Zhang J., Oko E., Wang M. Modelling of a post-combustion CO₂ capture process using extreme learning machine. *International Journal of Coal Science & Technology* 2017 4:1. Springer; 23 February 2017; 4(1): 33–40. Available at: DOI:10.1007/S40789-017-0158-1
- [4] Li F., Zhang J., Shang C., Huang D., Oko E., Wang M. Modelling of a post-combustion CO₂ capture process using deep belief network. *Applied Thermal Engineering*. Pergamon; 5 February 2018; 130: 997–1003. Available at: DOI:10.1016/J.APPLTHERMALENG.2017.11.078
- [5] Li Z., Ding Z., Wang M. Optimal Bidding and Operation of a Power Plant with Solvent-Based Carbon Capture under a CO₂ Allowance Market: A Solution with a Reinforcement Learning-Based Sarsa Temporal-Difference Algorithm. *Engineering*. Elsevier; 1 April 2017; 3(2): 257–265. Available at: DOI:10.1016/J.ENG.2017.02.014
- [6] Wu X., Wang M., Shen J., Li Y., Lawal A., Lee KY. Reinforced coordinated control of coal-fired power plant retrofitted with solvent-based CO₂ capture using model predictive controls. *Applied Energy*. Elsevier; 15 March 2019; 238: 495–515. Available at: DOI:10.1016/J.APENERGY.2019.01.082

- [7] Wu X., Shen J., Wang M., Lee KY. Intelligent predictive control of large-scale solvent-based CO₂ capture plant using artificial neural network and particle swarm optimization. *Energy*. Pergamon; 1 April 2020; 196: 117070. Available at: DOI:10.1016/J.ENERGY.2020.117070
- [8] Farmahini AH., Krishnamurthy S., Friedrich D., Brandani S., Sarkisov L. Performance-based screening of porous materials for carbon capture. *arXiv*. arXiv; 25 September 2020.
- [9] Burns TD., Pai KN., Subraveti SG., Collins SP., Krykunov M., Rajendran A., et al. Prediction of MOF Performance in Vacuum Swing Adsorption Systems for Post combustion CO₂ Capture Based on Integrated Molecular Simulations, Process Optimizations, and Machine Learning Models. *Environmental Science and Technology*. American Chemical Society; 7 April 2020; 54(7): 4536–4544. Available at: DOI:10.1021/acs.est.9b07407
- [10] Pai KN., Prasad V., Rajendran A. Generalized, Adsorbent-Agnostic, Artificial Neural Network Framework for Rapid Simulation, Optimization, and Adsorbent Screening of Adsorption Processes. *Industrial and Engineering Chemistry Research*. American Chemical Society; 23 September 2020; 59(38): 16730–16740. Available at: DOI:10.1021/acs.iecr.0c02339
- [11] Shao D., Yan Y., Zhang W., Sun S., Sun C., Xu L. Dynamic measurement of gas volume fraction in a CO₂ pipeline through capacitive sensing and data driven modelling. *International Journal of Greenhouse Gas Control*. Elsevier Ltd; 1 March 2020; 94: 102950. Available at: DOI:10.1016/j.ijggc.2019.102950
- [12] Sinha S., de Lima RP., Lin Y., Sun AY., Symon N., Pawar R., et al. Leak detection in carbon sequestration projects using machine learning methods: Cranfield site, Mississippi, USA. *Proceedings - SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers (SPE); 2020. Available at: DOI:10.2118/201552-ms

- [13] Jordan MI., Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. American Association for the Advancement of Science; 2015. pp. 255–260. Available at: DOI:10.1126/science.aaa8415
- [14] Garbade DrMJ. Clearing the Confusion: AI vs Machine Learning vs Deep Learning Differences | by Dr. Michael J. Garbade | Towards Data Science. 2018. pp. 1–6. Available at: <https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb> (Accessed: 7 April 2021)
- [15] Sunil Kumar. Advantages and Disadvantages of Artificial Intelligence | by sunil kumar | Towards Data Science. Towards Data Science. 2019. Available at: <https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-182a5ef6588c> (Accessed: 7 April 2021)
- [16] Stewart M. The Limitations of Machine Learning | by Matthew Stewart, PhD Researcher | Towards Data Science. 2019. Available at: <https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6> (Accessed: 7 April 2021)
- [17] Pareek P. Machine learning Algorithms and where they are used? Samurai's Sakura. 2019. Available at: <https://medium.com/@priyapareek0205/machine-learning-algorithms-and-where-they-are-used-c74de1441e1> (Accessed: 7 April 2021)
- [18] Amanpreet Singh; Narina Thakur; Aakanksha Sharma. A review of supervised machine learning algorithms - IEEE Conference Publication. Ieee. 2016; : 1310–1315.
- [19] Butler KT., Davies DW., Cartwright H., Isayev O., Walsh A. Machine learning for molecular and materials science. Nature. Nature Publishing Group; 2018. pp. 547–555. Available at: DOI:10.1038/s41586-018-0337-2

- [20] Shweta Bhatt. Reinforcement Learning 101. Learn the essentials of Reinforcement. Available at: <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292> (Accessed: 7 April 2021)
- [21] Dey A. Machine Learning Algorithms: A Review. International Journal of Computer Science and Information Technologies. 2016; 7(3): 1174–1179.
- [22] Principles of Adsorption and Adsorption Processes - Douglas M. Ruthven - Google Books. Available at: <https://books.google.co.uk/books?hl=en&lr=&id=u7wq21njR3UC&oi=fnd&pg=PR17&dq=1.%09D.+M.+Ruthven,+Principles+of+adsorption+and+adsorption+processes,+John+Wiley,+New+York,+1984.&ots=wcTuRuzjxX&sig=qoElCOM1JqWpmOakwxFzmJzy2Qo#v=onepage&q=1.%2509D. M. Ruthven%252> (Accessed: 7 April 2021)
- [23] Ruthven D., Publishers FP-NYV., 1994 undefined. Swing Adsorption. sutlib2.sut.ac.th.
- [24] Bui M., Adjiman CS., Bardow A., Anthony EJ., Boston A., Brown S., et al. Carbon capture and storage (CCS): The way forward. Energy and Environmental Science. Royal Society of Chemistry; 2018. pp. 1062–1176. Available at: DOI:10.1039/c7ee02342a
- [25] Long JR., Yaghi OM. The pervasive chemistry of metal-organic frameworks. Chemical Society Reviews. Royal Society of Chemistry; 2009. pp. 1213–1214. Available at: DOI:10.1039/b903811f
- [26] Li JR., Sculley J., Zhou HC. Metal-organic frameworks for separations. Chemical Reviews. American Chemical Society; 2012. pp. 869–932. Available at: DOI:10.1021/cr200190s
- [27] Yaghi OM., Li H. Hydrothermal Synthesis of a Metal-Organic Framework Containing Large Rectangular Channels. Journal of the American Chemical

Society. American Chemical Society; 1995; 117(41): 10401–10402. Available at: DOI:10.1021/ja00146a033

[28] Geng K., He T., Liu R., Dalapati S., Tan K.T., Li Z., et al. Covalent Organic Frameworks: Design, Synthesis, and Functions. Chemical Reviews. American Chemical Society; 2020. pp. 8814–8933. Available at: DOI:10.1021/acs.chemrev.9b00550

[29] Chen B., Yang Z., Zhu Y., Xia Y. Zeolitic imidazolate framework materials: Recent progress in synthesis and applications. Journal of Materials Chemistry A. Royal Society of Chemistry; 2014. pp. 16811–16831. Available at: DOI:10.1039/c4ta02984d

[30] Hasell T., Cooper A.I. Porous organic cages: Soluble, modular and molecular pores. Nature Reviews Materials. Nature Publishing Group; 2016. pp. 1–14. Available at: DOI:10.1038/natrevmats.2016.53

[31] Lin L.C., Berger A.H., Martin R.L., Kim J., Swisher J.A., Jariwala K., et al. In silico screening of carbon-capture materials. Nature Materials. Nature Publishing Group; 27 May 2012; 11(7): 633–641. Available at: DOI:10.1038/nmat3336

[32] McDonald T.M., Mason J.A., Kong X., Bloch E.D., Gygi D., Dani A., et al. Cooperative insertion of CO₂ in diamine-appended metal-organic frameworks. Nature. Nature Publishing Group; 19 March 2015; 519(7543): 303–308. Available at: DOI:10.1038/nature14327

[33] Siegelman R.L., McDonald T.M., Gonzalez M.I., Martell J.D., Milner P.J., Mason J.A., et al. Controlling Cooperative CO₂ Adsorption in Diamine-Appended Mg₂(dobpdc) Metal-Organic Frameworks. Journal of the American Chemical Society. American Chemical Society; 2 August 2017; 139(30): 10526–10538. Available at: DOI:10.1021/jacs.7b05858

[34] Rosen A.S., Mian M.R., Islamoglu T., Chen H., Farha O.K., Notestein J.M., et al. Tuning the Redox Activity of Metal-Organic Frameworks for Enhanced,

Selective O₂ Binding: Design Rules and Ambient Temperature O₂ Chemisorption in a Cobalt-Triazolate Framework. *Journal of the American Chemical Society*. American Chemical Society; 4 March 2020; 142(9): 4317–4328. Available at: DOI:10.1021/jacs.9b12401

[35] Moosavi SM., Jablonka KM., Smit B. The Role of Machine Learning in the Understanding and Design of Materials. *Journal of the American Chemical Society*. American Chemical Society; 2020. pp. 20273–20287. Available at: DOI:10.1021/jacs.0c09105

[36] Jablonka KM., Ongari D., Moosavi SM., Smit B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chemical Reviews*. American Chemical Society; 2020. pp. 8066–8129. Available at: DOI:10.1021/acs.chemrev.0c00004

[37] Wilmer CE., Leaf M., Lee CY., Farha OK., Hauser BG., Hupp JT., et al. Large-scale screening of hypothetical metal-organic frameworks. *NATURE CHEMISTRY* |. 2012; 4. Available at: DOI:10.1038/NCHEM.1192

[38] Moghadam PZ., Li A., Wiggins SB., Tao A., Maloney AGP., Wood PA., et al. Development of a Cambridge Structural Database Subset: A Collection of Metal-Organic Frameworks for Past, Present, and Future. *Chemistry of Materials*. American Chemical Society; 2017. pp. 2618–2625. Available at: DOI:10.1021/acs.chemmater.7b00441

[39] Boyd PG., Woo TK. A generalized method for constructing hypothetical nanoporous materials of any net topology from graph theory. *CrystEngComm*. Royal Society of Chemistry; 23 May 2016; 18(21): 3777–3792. Available at: DOI:10.1039/c6ce00407e

[40] Colón YJ., Gómez-Gualdrón DA., Snurr RQ. Topologically Guided, Automated Construction of Metal-Organic Frameworks and Their Evaluation for Energy-Related Applications. *Crystal Growth and Design*. American Chemical

Society; 1 November 2017; 17(11): 5801–5810. Available at: DOI:10.1021/acs.cgd.7b00848

[41] Cole JC., Groom CR., Read MG., Giangreco I., McCabe P., Reilly AM., et al. Generation of crystal structures using known crystal structures as analogues. *Acta Crystallographica Section B: Structural Science, Crystal Engineering and Materials*. International Union of Crystallography; 1 August 2016; 72(4): 530–541. Available at: DOI:10.1107/S2052520616006533

[42] Chung YG., Haldoupis E., Bucior BJ., Haranczyk M., Lee S., Zhang H., et al. Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal-Organic Framework Database: CoRE MOF 2019. *Journal of Chemical and Engineering Data*. American Chemical Society; 12 December 2019; 64(12): 5985–5998. Available at: DOI:10.1021/acs.jced.9b00835

[43] Pophale R., Cheeseman PA., Deem MW. A database of new zeolite-like materials. *Physical Chemistry Chemical Physics*. Royal Society of Chemistry; 21 July 2011; 13(27): 12407–12412. Available at: DOI:10.1039/c0cp02255a

[44] Chen H., Yan J., Wei R., Gao J., Lian J., Huang X. BP neural network prediction of calcium-based sorbent calcination/ carbonation cycle. *Asia-Pacific Power and Energy Engineering Conference, APPEEC*. 2011. Available at: DOI:10.1109/APPEEC.2011.5747710

[45] Nkulikiyinka P., Yan Y., Güleç F., Manovic V., Clough PT. Prediction of sorption enhanced steam methane reforming products from machine learning based soft-sensor models. *Energy and AI*. Elsevier BV; 1 November 2020; 2: 100037. Available at: DOI:10.1016/j.egyai.2020.100037

[46] Krzywanski J., Czakiert T., Shimizu T., Majchrzak-Kuceba I., Shimazaki Y., Zylka A., et al. NO_x Emissions from Regenerator of Calcium Looping Process. *Energy and Fuels*. American Chemical Society; 17 May 2018; 32(5): 6355–6362. Available at: DOI:10.1021/acs.energyfuels.8b00944

[47] Hanak DP., Manovic V. Economic feasibility of calcium looping under uncertainty. *Applied Energy*. Elsevier Ltd; 15 December 2017; 208: 691–702. Available at: DOI:10.1016/j.apenergy.2017.09.078

4 Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft Sensor Models

Paula Nkulikiyinka ^a, Yongliang Yan ^a, Fatih Güleç ^{a, b}, Vasilije Manovic ^a, Peter T. Clough ^a

^aEnergy and Power Theme, School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK.

^b Advanced Materials Research Group, Faculty of Engineering, University of Nottingham, Nottingham, NG7 2RD, UK.

DOI: <https://doi.org/10.1016/j.egyai.2020.100037>

Abstract

Carbon dioxide-abated hydrogen can be synthesised via various processes, one of which is sorption enhanced steam methane reforming (SE-SMR), which produces separated streams of high purity H₂ and CO₂. Properties of hydrogen and the sorbent material hinder the ability to rapidly upscale SE-SMR, therefore the use of artificial intelligence models, that are a realistic representation of plant performance, is useful in order to assist scale up. Advantages of a data driven soft-sensor model over thermodynamic simulation software is the ability to obtain real time information dependent on actual process conditions. A soft sensor model using machine learning has been developed and used to predict and estimate variables that are otherwise difficult to be directly measured. In this study, artificial neural networks and the random forest models were developed as soft sensor prediction models. Both models were shown to provide good performance predictions for gas concentrations in the reformer and regenerator reactors of the SE-SMR process using temperature, pressure, steam to carbon ratio and sorbent to carbon ratio as input process features. Additionally, the CH₄ conversion within the reformer, that was predicted with both algorithms was able to reflect the efficiency of the reaction very accurately. Both models were accurate with high R² values, all above 98%, however the random forest model was more

precise in the predictions, and thus has a good ability to act as a soft sensor, with consistently higher R^2 values and lower mean absolute error values (0.002-0.014) compared to the neural network model (0.005-0.024).

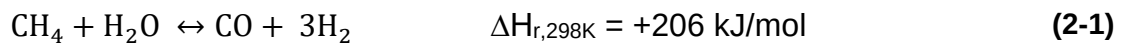
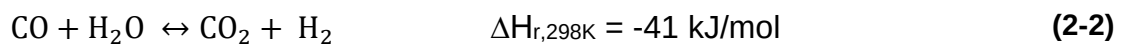
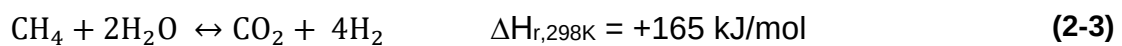
Keywords: Machine learning; artificial neural network; sorption enhanced steam methane reforming; soft sensor; calcium looping

4.1 Introduction

Aside from the imperative need to reduce greenhouse gas emissions from fossil fuel use, there are many other reasons why blue hydrogen, produced from fossil fuels with CO_2 capture, is likely to be at the forefront of future energy systems. For instance, it can be produced in several ways and from numerous resources, both renewable and non-renewable, and upon combustion, no greenhouse gases or other emissions are produced. Hydrogen can also be used to decarbonise several sectors that are otherwise difficult to reduce greenhouse gas emissions from [1].

Although hydrogen can be produced from various resources, both renewables and non-renewables, the most common method at present for production of H_2 is via steam methane reforming (SMR). In this process, methane is reacted with steam producing carbon monoxide and H_2 , then the carbon monoxide is reacted with steam in water gas shift (WGS) reactors to produce CO_2 and more H_2 .

The reactions that take place in order to produce H_2 from methane are reforming (reaction (2-1)) and the WGS Reaction (reaction (2-2)), giving the overall SMR reaction (reaction (2-3)):



Sorbent Enhanced Steam Methane Reforming (SE-SMR) is a process which integrates calcium looping (CaL) and SMR. The principal reactions in CaL are carbonation (forward reaction) and calcination (backward reaction) described in Equation 2-4. The combination of Equations (2-3) (SMR) and (2-4) (CaL) gives the overall SE-SMR reaction shown in Equation (2-5):



One of the challenges faced in SE-SMR, and in various other thermochemical processes, is the inability to measure and thus control important variables in a reactor and process control systems due to a variety of physical conditional limitations. These variables range from parameters such as product composition, concentration of reactants, spatial-temporal temperature and pressure profiles, and the limitations stem from reasons including financial, physical, inaccuracies, or competing influences [2]. A solution to this challenge is to incorporate artificial intelligence (AI) into the energy sector through the use of soft sensors, which are inferential estimators using predictions based on data to provide real time digital approximations as to the conditions inside the reactor [3,4]. A simpler definition of a soft sensor is that it is a predictive model based on large quantities of data available on an industrial process, which can either be first principle models (white-box models) or data driven models (black-box models). White box models depend on actual mechanical data of the process, whereas the latter uses historical collected process data, which makes black-box models far more practical and readily applicable to process plants [5].

The principle on which soft sensors work is based on the estimation of quality through a mathematical model that uses all available measured process variables [6]. In further detail, the idea of inference estimation (soft sensing) is to approximate the values of primary variables (e.g., gas composition) by the easily measured secondary variables (e.g., temperature), which are correlated to the primary variables.

For example, if Y is the dependent variable (soft sensor output, e.g. concentration) and $X_1, X_2 \dots X_n$ are the independent inputs (e.g. temperature, flow), the mathematical model for the output prediction can be described with Equation (4-1), where $\tilde{X} = \Sigma(X_1, X_2 \dots X_n)$ and is the vector of the inputs, B is the vector of the model coefficients, and ε is the output error [6].

$$Y = F(\tilde{X}, B) + \varepsilon \quad (4-1)$$

In recent years, soft sensors have been widely studied and used in industrial process control to improve the quality of product and assure safety in production, and additionally act as backup when hardware sensors are unavailable or unsuitable [7,8]. In addition, the desired measurements which are the key indicators of process performance, are normally ascertained by off-line sample analyses or on-line product quality analysers, which are often costly and require frequent and high-cost maintenance, and sometimes add delay to the process as the measurements cannot be used as feedback signals for quality control systems [9].

The use of soft sensors is not new and has been extensively implemented in process industries for the past three decades. Recently, machine learning, as a whole, has been used in carbon capture technologies, mainly focused on the process modelling and control of amine-based post-combustion capture systems [10,11]. The most up to date literature that is somewhat comparable to this study includes use of a surrogate model to optimise the energy within steam methane reforming [12], domain reduction for efficient surrogate model generation in the production of hydrogen [13] and data-driven process design on SMR [14].

However, to the best of our knowledge, machine learning algorithms have not been applied to sorption enhanced steam methane reforming to develop a soft sensor model of the process. In this study, artificial neural network (ANN) and random forest (RF) algorithms are employed to construct soft sensor models. A comparison of the models is then conducted against the SE-SMR reactions taking place to evaluate their ability to predict operating parameters including reactor

gas concentrations, reformer CH₄ conversion and overall process H₂ purity at specified conditions.

4.2 Methodology

The development of the data-driven soft sensor followed the typical systematic approach as found in literature, with the following being the main elements of the process [5][15]:

1. Data collection, with the secondary variables selected according to knowledge of the process.
2. Data pre-processing
 - a. Outlier detection using univariate analysis.
 - b. Normalisation where necessary
 - c. Correlation/redundancy elimination using multivariate principal component analysis (PCA)
3. Regression model development
4. Validation of the soft sensor on independent process data

For the first point, data collection, a database of 13,756 data points were obtained for the machine learning soft sensor, which were produced from an Aspen Plus model that was developed to simulate the SE-SMR process and then applying a sensitivity analysis to the process across a range of commonly employed operating ranges.

4.2.1 Process Configuration and Simulation

The SE-SMR process was simulated using Aspen Plus V10 (Figure 4-1) using the Peng-Robinson property method and the equilibrium calculations were performed, initially using the baseline conditions (Table 4-1). Peng-Robinson was chosen as it's a key aspect of understanding real fluid behaviour, phase equilibrium and thermodynamic properties. It is defined as a mathematical model that calculates and predicts the behaviour of pure components and mixtures in the gas, liquid and supercritical fluid state, and is one of the most commonly used Equation of State models used within chemical engineering for this reason.

The sorbent carrying capacity is defined as the CaO molar conversion to CaCO₃ at the end of the fast carbonation stage of CaO in each carbonation–calcination cycle. The cyclone separation efficiency is defined as how well a cyclone can remove particulate matter in flue gas, which is usually dependant on particle size.

Table 4-1 Aspen Plus SE-SMR baseline process model parameters.

Baseline Conditions	
Steam to carbon molar ratio	4
Sorbent to carbon molar ratio	1
Pressure	1 bar
Reformer temperature	650°C
Regenerator temperature	850°C
Pressure Swing Adsorber (PSA) efficiency	95%
Sorbent carrying capacity	20%
Cyclone separation efficiency	100%

The reformer and regenerator reactors were simulated with RGibbs blocks. CO₂-abated hydrogen (stream 18) was obtained from separating hydrogen from the other flue gas components using a pressure swing adsorber with an efficiency of 95%. Water was separated from the system, using flash separators, producing two product streams of water (streams 16 and 20) and the product stream of CO₂ (stream 19). The temperature and pressure for FL1 and FL2 were both at 25°C and 1 bar. The method of operation within these units is to heat the feed up at a certain pressure and temperature, wherein some of the mixture will be vaporised. This results in two phases: a vapor phase, saturated with the more volatile components, and a liquid phase, with the less volatile components, resulting into two exit streams in different phases.

An oxyfired calciner was added to the process in order to provide heat for the regeneration of the carbonated sorbent. In this case the heat requirement is met by combustion of methane (stream 2) and PSA tail gas (stream 17- consisting of CH₄, H₂O, CO, CO₂ and H₂ all in trace amounts i.e., 0.02-0.19kmol/hr), using

oxygen (stream 22) from an air separation unit, producing carbon dioxide and steam. To ensure these were the only gases formed, excess oxygen was supplied to ensure complete combustion of methane. Fresh sorbent i.e., limestone (CaCO_3), was entered into the process via stream 24. Two heat exchangers were used in order to elevate as well as maintain the heat in the streams to and from the reactors, thus achieving an efficient heat integration of the reactors' thermal content. With these baseline conditions, the H_2 purity was found to be 91.65%. Stream compositions under baseline conditions are shown in the appendices (Table_Apx A-1).

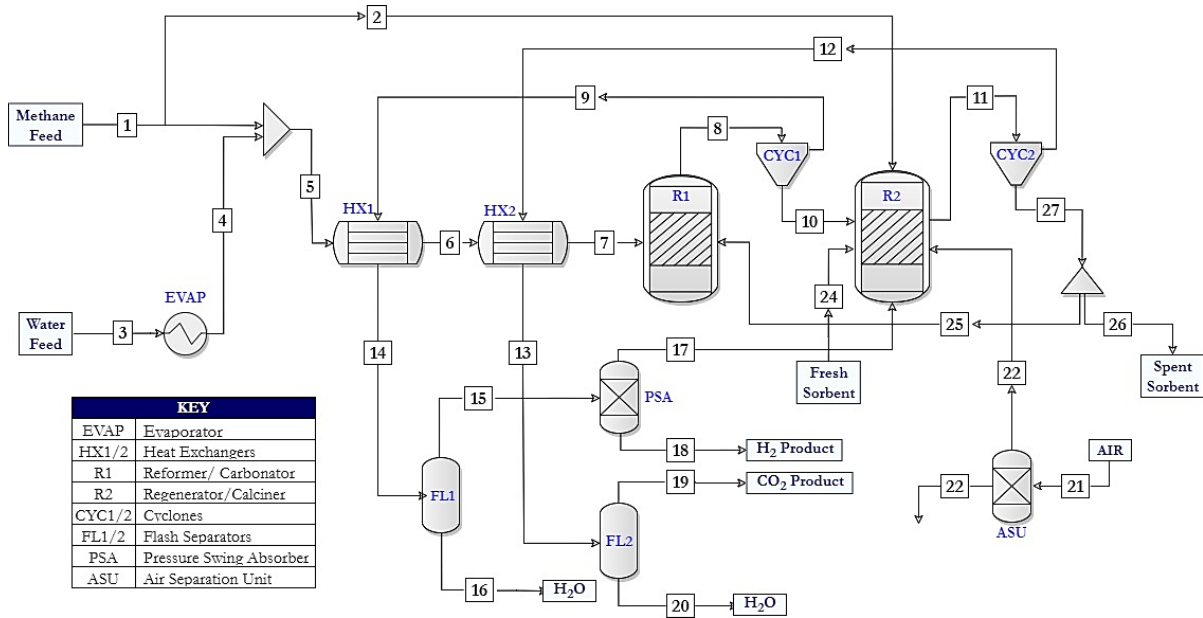


Figure 4-1 Process flow diagram of the SE-SMR process simulated in Aspen Plus.

A sensitivity analysis was performed to generate a large database (13,756) that was implemented in the machine learning models. The parameters that were chosen to vary were: steam to carbon ratio, sorbent to carbon ratio, reformer pressure, as well as reformer and regenerator temperatures. They were varied over the ranges shown in Table 4-2. These parameters were chosen as they form the independent control/performance variables for the SE-SMR process, and the ranges were chosen as they were typically reported in the available literature [16–19]. The data sampling was arranged with one change per datapoint and the rationale for this was to act as if it were an actual reactor and therefore

experiencing slight changes in the conditions in every and any direction. For example, for one datapoint the reformer temperature, regenerator temperature, S/C (steam to carbon ratio) and reformer pressure were all kept constant with changing CaO/C (sorbent to carbon ratios). Then for the next datapoint, the next reformer temperature was chosen, and again the other three parameters were kept constant, for changing CaO/C, and so forth so that each parameter was altered and was combined with all changes in the other parameters as well. Additionally, as seen in Table 4-2, whilst the pressures and temperatures had set intervals, the ratios varied between the given ranges as they were calculated from changing flowrates. Again, this was chosen to replicate practical operation as the flowrates would be the parameters altered as opposed to picking specific ratios. What was ensured however was that the ratios were kept within the specified ranges.

Although oxy-fuel combustion usually occurs in 70-80% CO₂ and 30-20% O₂ since burning methane in O₂ releases a lot of heat, instead of recycling CO₂ from stream 12, only CO₂ from the PSA off-gas was recycled. This is because it has been shown in literature, that increasing the oxygen concentration in the calciner reduces the amount of recycled gas (CO₂, H₂O and impurities); thus, removing the energy necessary to heat the recycled gas stream [20]. To ensure excess air was supplied to the process at all times, a calculator block was used, which took into account the methane fed to and the CO₂ from the PSA off-gas recycled to the calciner and made appropriate adjustments to the flow from the ASU. In addition, a design spec was used to ensure the quantity of fresh sorbent supplied was equal to the spent sorbent being purged from the calciner. The specification also ensured the relative sorbent make up (Fresh limestone/Sorbent circulation rate) was 0.04 mol% which is in the range of typical rates presented in the literature [21].

Table 4-2 Parameter ranges used in the sensitivity analysis to build up the machine learning database. *In order to obtain a ratio, the flowrates for the steam, methane or sorbent, were altered between the given ranges, and a ratio was then calculated.

Parameter	Range	Intervals
Steam to carbon ratio	1 - 7	Various*
Sorbent to carbon ratio	0.1 - 5	Various*
Reformer pressure	1 - 10 bar	1, 5, 10 bar
Reformer temperature	500 - 800°C	Every 50°C
Regenerator temperature	700 - 1000°C	Every 50°C

4.2.2 SE-SMR Process Model Validation

In order to validate the SE-SMR process model and thus to ensure that the results from the machine learning models were an accurate representation of the system, the literature data were used, as a comparison with the model predictions is presented in Table 4-3.

Table 4-3 Comparison of literature data against Aspen Plus data for validation purposes.

	Li et al. [22]	Johnsen et al. [23]	Arstad et al. [24]
Process parameters			
S/C ratio	5	3	4
CaO/C ratio	0.125	0.18	1.1
Reformer pressure (bar)	1	1	1
Reformer temperature (°C)	630	600	575
Regenerator temperature (°C)	850	850	895
Literature data			
H ₂ concentration (%)	95.6	98.4	95.8
CO ₂ concentration (%)	2.7	0.4	0.7
Model predictions			
H ₂ concentration (%)	97.3	96.9	98.9
CO ₂ concentration (%)	1.5	0.5	0.3

Table 4-3 shows that the model predictions are in good agreement with the experimental data found in literature. The average percentage error (Equation (4-2)), where n is the number of times the summation iteration happens, A_t is the actual value and F_t is the forecast value, across the H₂ concentration and CO₂ concentration is 2.2% and 42.2%, respectively. The deviation for H₂ concentration is deemed acceptable to validate the model as 2.2% margin of error for experimental results, particularly for this sample size. On the other hand, whilst the CO₂ concentration error is much higher, it should be considered that CO₂ concentrations are relatively low, and a small deviation in process performance can make a significant difference.

$$\text{Average percentage error (\%)} = \frac{1}{n} \sum_{t=1}^n \frac{A_t - F_t}{A_t} \quad (4-2)$$

4.2.3 Machine Learning Model Selection

The input variables, or features, chosen were S/C (steam to carbon ratio), CaO/C (sorbent to carbon ratio), reformer pressure and the reformer and regenerator temperatures. The desired soft sensor outputs to be obtained from these features were the gaseous concentrations within the reformer (H₂O, CO, H₂) and regenerator (H₂O, CO₂). Another valued output to be found during process operation was the reformer CH₄ conversion efficiency (Equation (4-3)) defined as the moles of CH₄ reacted, divided by the moles of CH₄ going entering the reformer [25].

$$\text{CH}_4 \text{ conversion (\%)} = \frac{(n_{\text{CH}_4, \text{in}} - n_{\text{CH}_4, \text{out}})}{n_{\text{CH}_4, \text{in}}} \times 100 \quad (4-3)$$

With the five input features and the desired output performance indicators selected, a machine learning algorithm was developed. As the information gathered was all numerical, very little data pre-processing was required at this stage. One action taken on the data was randomising the order of the data as it had been collected in order of increasing variable e.g., S/C, CaO/C, temperature, pressure. This was an action taken to prevent the models from overfitting the

data, by recognising a sequential pattern. Additionally, the data was scaled once it was implemented into Python, more specifically normalised to have a value between zero and one, which prevented the models attributing a higher 'importance' to larger values, and thus giving it higher influence on the overall results.

The next step in the process to develop the soft sensor, was to select appropriate machine learning models that were suitable for the type of data collected. The type of machine learning chosen was 'supervised learning' since the system was presented with labelled input-output data in order to learn from it. Of the two types of supervised learning categories- classification and regression, regression was the type to be used as the data was continuous and numerical. The two models chosen were RF and ANN.

4.3 Machine Learning Theory and Calculations

4.3.1 Random Forest

The random forest model is a combination of decision tree predictors, where each tree depends on the values of a random vector sampled independently, as well as with an equal distribution across all trees [26]. They are able to overcome the problem of overfitting as they introduce the randomness into each individual tree by growing each tree on a bootstrap sample of the training data and only using random subsets of the available predictors when creating the split. The formula for the predictions is given in Equation (4-4), where x' is unseen samples, f_b is each tree and B is the number of trees after the split.

$$f(x) = \frac{1}{B} \sum_{b=1}^B f_b(x') \quad (4-4)$$

Two principal features of the RF algorithm are out-of-bag (OOB) error estimates and feature importance rankings. Out-of-bag samples refers to a third of samples not being used for fitting a particular regression tree in the forest. The OOB score provides a similar measure of the model's generalization error and is calculated by constructing the predictor for each observation with the trees that apply for

that specific observation that was out-of-bag. The overall OOB score is then computed as the average error of all OOB predictions [27]. The feature importance is calculated by recording the improvement in the split-criterion at each split and in each tree. These improvements are then attributed to the feature that were used to split these particular nodes and summed over all trees and for all feature separately [28].

The primary reasons for the random forest algorithm being beneficial for this type of dataset includes that they are quite resistant to overfitting, with the addition of more trees and they do not require complicated or time-consuming processes of variable selection [29]. Additionally, RF can handle missing values and maintain the accuracy of a large proportion of data. An annotation of how this random forest works through the ‘trees’ is shown in supplementary materials (Figure_Apx A-1)

4.3.2 Artificial Neural Network (ANN)

ANN is a non-parametric, information processing, statistical paradigm, which does not require any pre-assumption of the input–output relationship. ANNs consist of three principal layers: an input layer, several hidden layers and an output layer, which are made up of various numbers of connected ‘neurons’ or ‘nodes’ in each layer. The quantity of hidden layers defines the depth of the architecture. The data is transferred from the input node, along to the hidden layer nodes, eventually reaching the output layer to achieve the results or predictions [30]. It has also been shown that ANNs have the ability to approximate any non-linear system with high interpolation capacity. ANNs have a high ability to effectively approximate non-linear systems which is down to the use of one or more hidden layers and non-linear transfer functions in the hidden layer’s neurons. A unique attribute of ANNs is that they are adaptive, but with the requirement that the input data is presented to the network in order to recognize the process i.e., it must be trained first, therefore, ANN is a ‘supervised learning’ algorithm. The training process of the ANNs aims to achieve the highest coefficient of determination (R^2) and the lowest root mean square error (RMSE)

and mean absolute error (MAE) (Equations (4-5) and (4-6)), where n is the quantity of data required for the network training, y_i is the predicted output and t_i is the target output. This is done by fine-tuning the weights and biases according to the training algorithms [10][31]. The reason both RMSE and MAE were used as accuracy metrics include, that they can be used together to diagnose the variation in the errors in a set of forecasts.

The MAE measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables and is a linear score which means that all the individual differences are weighted equally in the average.

The RMSE is a quadratic scoring rule which measures the average magnitude of the error. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable.

$$RMSE = \sqrt{\frac{1}{n} \left[\sum_{i=1}^n (t_i - y_i)^2 \right]} \quad (4-5)$$

$$MAE = \frac{1}{n} \left[\sum_{i=1}^n |t_i - y_i| \right] \quad (4-6)$$

The use of ANN as a tool for nonlinear soft sensing modelling has been widely used in the recent years, particularly regarding the prediction accuracy and saving of computational costs in many industrial fields [32]. Some examples of the applications include estimation of outputs based on industrial data [11], reduction of costs involved in computational methods [33], and prediction of behaviour for control and optimization systems [34,35]. Benefits of this model include that it has the ability to model highly non-linear functions and can be trained to accurately generalise when presented with new, unseen data [36].

4.3.3 Multivariate Analysis

As previously mentioned in the section 4.3.1, a feature commonly used within the random forest model is the feature importance (Table 4-4), which is found by the OOB data. When the training set for the current tree is drawn by sampling with replacement, as explained in section 4.3.1, only a third of the cases are left out of the sample. This OOB data is used to get a running unbiased estimate of the classification error as trees are added to the forest, and also used to get estimates of feature importance, which has proven to be unbiased in many tests [26], (full list of data in Appendix A, Table_Apx A-2). In addition, another popular method to quantify the importance of the features on the impact it has on a model, is to permute the values of each feature and measure the effect of permutation on the accuracy of the model. This essentially creates a cut-off point in the feature importance values [37]. An example of this is applied to the CO₂ concentration in the regenerator. Table 4-4 indicates that regenerator temperature and the sorbent to carbon ratio are the most influential in the CO₂ concentration, and it is clear from permuting the values of the other three features, that the former two are the only important features in predicting CO₂ concentration from the calciner, which is expected. From this it can also be inferred that the value of 0.1 for variable importance acts as a good cut-off point as the R² value remained at >0.99 with the permuted features. This was confirmed by testing the other concentrations and permuting values until a significant change in accuracy was found, and across each parameter it was clear that features with an importance greater than 0.1 was considered 'important'.

The equivalence of this application in ANN is backward stepwise elimination to the ANN model (assessing change to RMSE with each input node removal) [38]. However, it has been reported in the literature [39,40] that the method of backward stepwise elimination is not always a preferable method of quantifying variable importance, as the method is usually only successful in measuring the most influential variable as opposed to the whole input dataset.

Table 4-4 Random Forest feature importance for each output parameter. (Ref = Reformer, Reg = Regenerator, T = Temperature, P = Pressure)

Level of importance	Reformer [H ₂ O]	Reformer [CO]	Reformer [H ₂]	Regenerator [H ₂ O]	Regenerator [CO ₂]	[CH ₄] reformer conversion
Most	S/C	Ref T	S/C	CaO/C	Reg T	S/C
	Reg T	Ref P	Reg T	S/C	CaO/C	Ref T
	Ref T	S/C	Ref T	Reg T	Ref T	Ref P
	Ref P	Reg T	Ref P	Ref T	Ref P	Reg T
Least	CaO/C	CaO/C	CaO/C	Ref P	S/C	CaO/C

ANN is often seen as a black box, from which it is very difficult to extract useful information for another purpose like feature importance/explanations [41]. Although it is possible, for this reason, to analyse the input features for the ANN model and compare against the random forest variable importance, principal component analysis (PCA) was performed as an effective procedure for the determination of input parameters [42]. PCA is a mathematical model that is used to reduce the dimensionality of a dataset, whilst retaining most of the variation in the dataset, which allows for better interpretation of the data. Some features in a dataset, measure related properties, and are effectively redundant, and therefore should be removed for effective analysis, which is what PCA aims to do. The reduction of features is achieved by identifying the directions (principal components) that the variation is highest [43,44].

The PCA was calculated in Python using 'scikit-learn's PCA package', firstly to check the interaction and redundancies of the five input variables.

Figure 4-2 Figure 4-2 shows the correlation plot of the inputs. From this, it is inferred that the reformer pressure and regenerator temperature, are inversely correlated, meaning as one increases the other decreases. It can also be seen that regenerator temperature has little linear correlation with reformer temperature, but is positively correlated to steam to carbon ratio, as the former value is close to zero, and the latter, close to one. Additionally, it is evident that none of the input variables are redundant as none of the variables correlate perfectly

throughout, i.e., none of the values are significantly close to one, so they are independent of each other, therefore all inputs are retained for the model development.

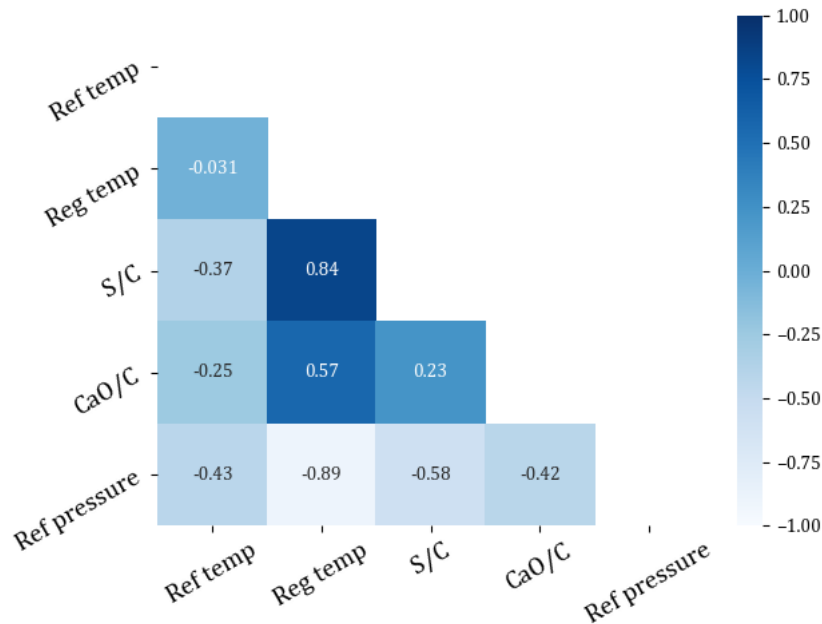


Figure 4-2 Correlation heatmap of input features (Ref = Reformer, Reg = Regenerator, temp = Temperature).

Secondly, Figure 4-3 was generated to examine the interrelation among a set of variables in order to identify the underlying structure of those variables. As mentioned before, PCA allows a reduction in components by looking at the independence (Figure 4-2) as well as how they contribute to the variance, i.e., how the variables combine to form the principal components. This can be found by measuring the variance ratio of the principal components.

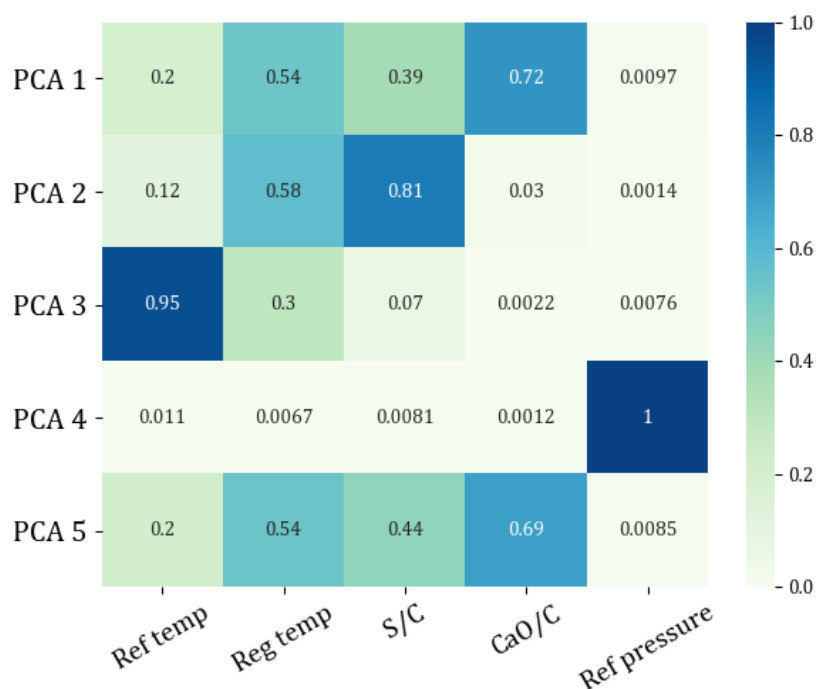


Figure 4-3 Heatmap showing dependencies of 5 principal components on input features (Ref = Reformer, Reg = Regenerator, temp = Temperature).

The results for this dataset was found to be that the first PC explains 32.1% of the variance, PC2 explains, 21.2%, PC3 and PC4 explain 20% each and PC5 explains 0.07%. This indicates a fairly even distribution of variable influence across the first four principal components, and therefore not essentially necessary to remove any parameters. On relatively small data sets, and in examples from literature [40], the first two components normally contribute to around 70-95% of the total variance, hence why PCA is normally depicted with two principal components. PCA becomes a valid method for feature selection, only when the most important features, are the ones that happen to have to most variation in them, which is usually not true, and therefore features with low variation should not automatically be disregarded [45].

In addition to explaining the variance, Figure 4-3 also explains feature importance. From this plot it can be calculated that the top three features that contribute to the five PCs of reformer methane conversion, are the reformer and

regenerator temperature, and the steam to carbon ratio, which corresponds to the random forest variable importance results (Table 4-4).

4.3.4 Machine Learning Model Setup

The models were coded in Python 3.8 using the PyCharm integrated development environment with the Keras and TensorFlow libraries as the backend. Keras is a library of open sources of the neural network developed in Python, which is focused on minimization, modularity, and scalability. Additionally, TensorFlow is an open-source software library which provides an interface for expressing and executing various machine learning algorithms [46]. Firstly, with the ANN model, various network configurations were tested trialled and compared with their R^2 and MAE values. The elbow-curve method was employed to find the configuration with the least number of nodes per layer that still resulted in high prediction accuracy metrics and minimised computational time.

Alongside the number of nodes per layer, the number of epochs and batch size are two important variables that influence the run time of the model as well as the accuracy of prediction. Simple descriptions of these parameters are:

- Batch size: The number of samples worked through by the model before updating the internal node parameters.
- Epoch: The number of times that the machine learning model works through the full dataset provided whilst training.

The optimal combination was found, by methods recommended in the literature, through multiple iterations and by plotting network accuracy vs epochs [47].

The metrics used to give the accuracy of the models were the coefficient of determination (R^2), the MAE and the root mean square error (RMSE). An indication of a well fitted and highly accurate model was that with low RMSE and MAE values as well as a high R^2 . Figure 4-4 shows a plot of accuracy using the MAE, on the training and validation datasets over training epochs, for CH_4 conversion. Whilst it is clear the variation is lesser with more epochs, the actual

MAE at the end of both iterations is very close and the difference is almost negligible after ~10 epochs. Due to this, 10 epochs were tested against the dataset to further reduce computation time, however, it was found that this resulted in error and therefore there is a minimum limit of epochs for the size of dataset used, which in this case was 30.

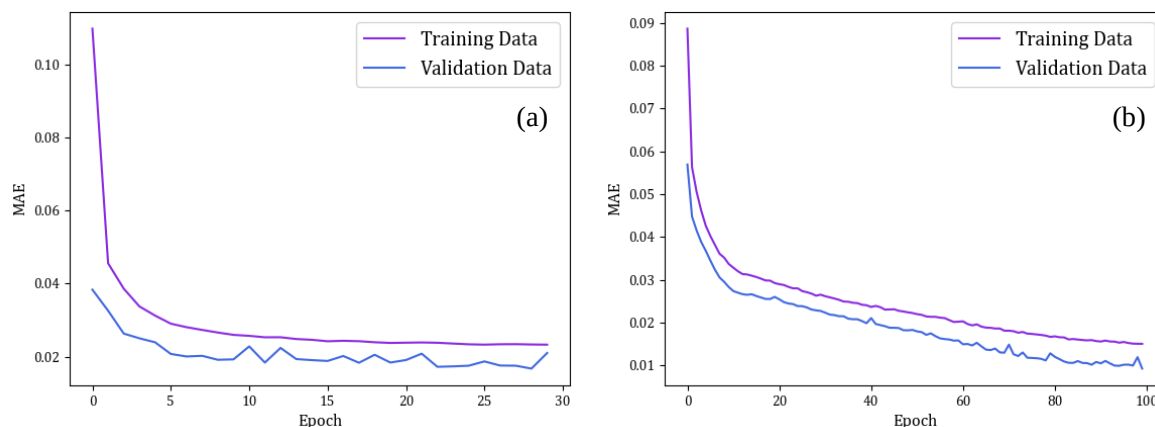


Figure 4-4 Number of epochs vs MAE for reformer H₂O concentration: a) 30 epochs, b) 100 epochs.

In the literature, it is suggested that a larger number of epochs gives higher accuracy for prediction as the model is able to go through the dataset more times. However, with this dataset and model scenario, it was found that a sufficiently high R^2 and MSE were obtained with low batch sizes and epochs. Reasons for this outcome is thought to be down to the use of a relatively simple and clean dataset obtained from thermodynamic calculations in Aspen Plus, as opposed to using experimental data that will contain experimental error leading to deviations in the dataset. A model with 100 epochs (Figure 4-4b) and batch size of 10 gave an R^2 value of 0.95 whereas a model with 30 epochs (Figure 4-4a) and batch size of 10, and a lower model run time, gave R^2 value of 0.92. The small difference in error from these two examples hides the computational time required, thus, to minimise computational time (from 907 seconds to 245 seconds), and error a low epoch size (30) was applied.

For the model developed, the batch size chosen was 2 and 30 epochs, and the way in which this was found is explained in the results section.

- To give an insight into what that meant for the data provided:
- With a dataset with 13,756 samples, the dataset was divided into 6,878 batches, each with 2 samples.
- The model weights were updated after each batch of 2 samples. One epoch will involve 6878 batches i.e., 6878 updates to the model.
- With 30 epochs, the model will go through the whole dataset 30 times.
- Resulting in a total of 206,340 batches during the entire training process.

An additional factor in the ANN setup to be considered was the activation function used, which defines the output signal of that node, given an input or set of inputs [48]. From literature [11][42], it was found that whilst the Rectified Linear Unit (ReLU), was the most commonly used due to its speed, functions such as hyperbolic tangent (tanh) and sigmoid functions were more appropriate functions to be used with the dataset provided. This was investigated in order to reach the optimal function and it was found that a combination of both the ReLU and sigmoid functions in between the layers of nodes gave the highest accuracy metrics. A comparative view of the functions is shown in supplementary materials (Figure_Apx A-2) and details on the activation functions are shown in Table_Apx A-3.

The calculations involved in the random forest model are less dependent on the inputs from the user as the model has a more automatic analysis process [49]. What does have to be specified however is the subsample size (the number of preselected directions for splitting the data) the tree depth (which can be specified in different ways) and the number of trees. However, there is a general consensus that implemented default values for these parameters produce good empirical performance in prediction, which is partially why random forests are so popular as a machine learning tool [50].

4.4 Results and Discussion

The results are graphically presented in the form of predicted against actual plots, which shows the accuracy of the model. These plots indicate how well the predicted data fits with actual data, and whether the patterns in the actual data

be reproduced by the model [51]. A model that performs well would produce a scatter of data that are close to the $y = x$ line. Figure 4-5 and Figure 4-6 depict the calculated Aspen data as ‘Actual’ data plotted against the ‘Predicted’ data from the two machine learning models. From both figures and the supporting metrics, it is evident that the random forest performed better than the neural network in the ability to predict the concentration of the gas in each reactor, due to achieving higher R^2 and lower RMSE and MAE values (presented in Table 4-5).

Table 4-5 Neural network vs random forest soft-sensor concentration predictions

	Neural Network			Random Forest		
	R^2	MAE	RMSE	R^2	MAE	RMSE
Reformer [H₂O]	0.98	0.012	0.018	0.99	0.007	0.011
Reformer [CO]	0.99	0.005	0.008	0.99	0.003	0.007
Reformer [H₂]	0.98	0.016	0.020	0.98	0.014	0.023
Regenerator [H₂O]	0.99	0.011	0.015	0.99	0.010	0.015
Regenerator [CO₂]	1.00	0.007	0.008	1.00	0.002	0.003
CH₄ Conversion	0.98	0.024	0.043	0.99	0.012	0.017

Another comparison of the two models is the prediction of the conversion of CH₄ within the reformer, Figure 4-5(f) (ANN) and Figure 4-6(f) (RF). In this comparison, both the neural network and random forest give a high R^2 value (0.98 and 0.99, respectively), indicating that they produce an accurate prediction of the efficiency of the reaction taking place in the reformer, at any given moment. In practical terms, this translates to allowing an operator to observe if the set conditions (pressure, temperature, flowrates) are giving a sufficient conversion rate and product purity in real time and make appropriate changes to achieve the desired product.

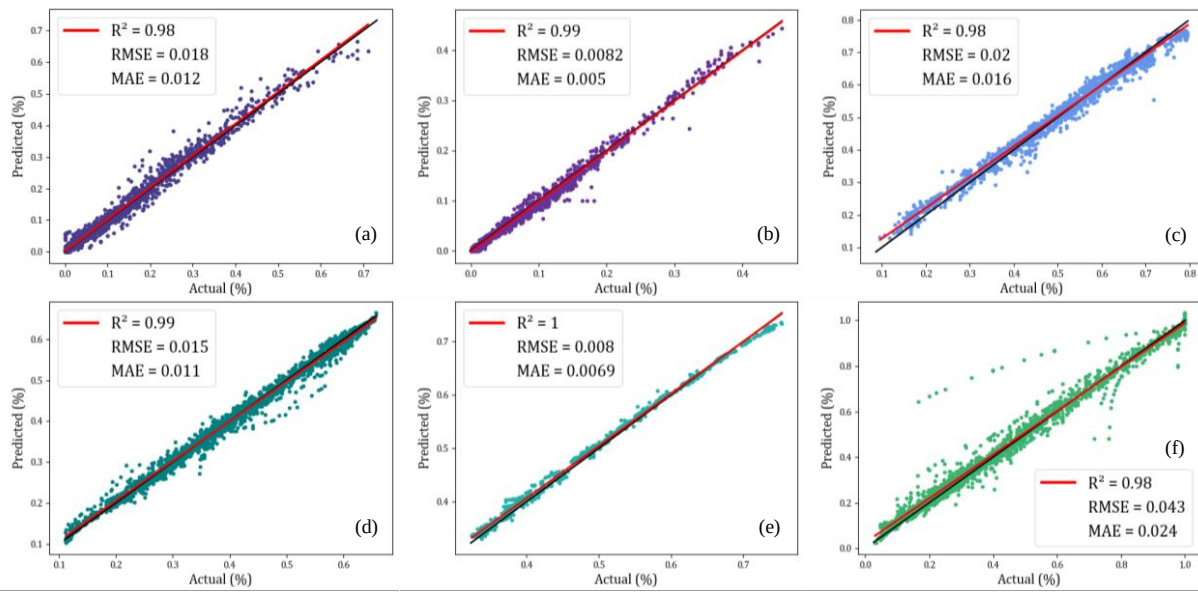


Figure 4-5 Actual vs Predicted plots by the ANN algorithm for: a) reformer $[H_2O]$, b) reformer $[CO]$, c) reformer $[H_2]$, d) regenerator $[H_2O]$, e) regenerator $[CO_2]$, f) reformer CH_4 conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.

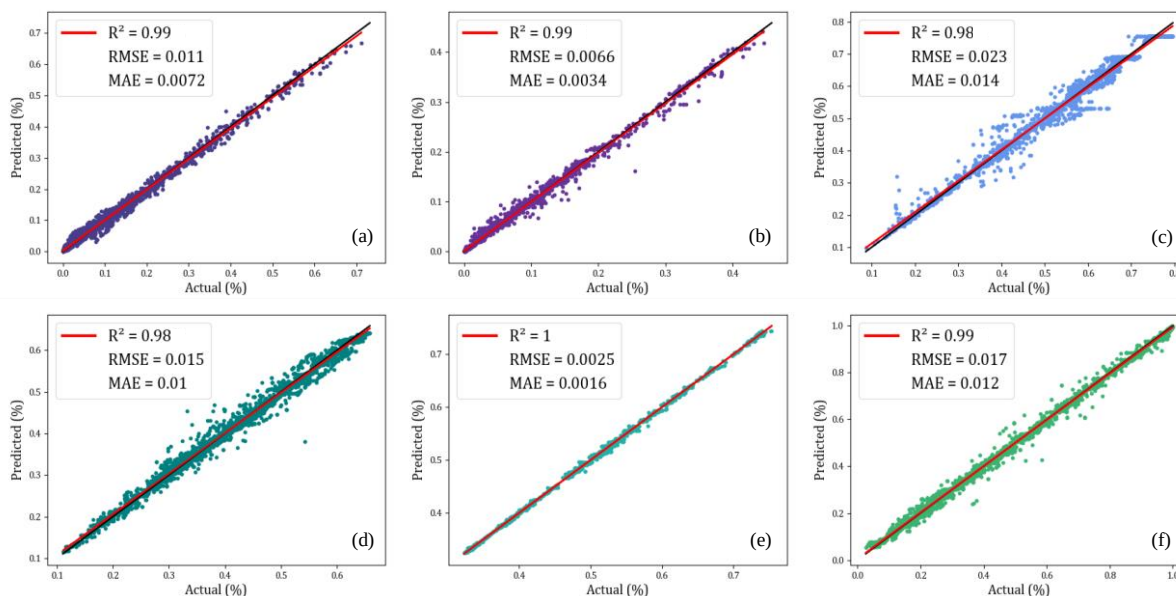


Figure 4-6 Actual vs Predicted plot by the random forest algorithm for: a) reformer $[H_2O]$, b) reformer $[CO]$, c) reformer $[H_2]$, d) regenerator $[H_2O]$, e) regenerator $[CO_2]$, f) reformer CH_4 conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.

In the plots, the red line represents the line of best fit for the data, with the black line representing $y = x$. For the neural network, the one main difference is that there is a higher spread of data, and this can be explained by MAE value, which gives the magnitude of model prediction error, and is higher for the neural network. For all plots it is clear that that line of best fit almost overlaps the $y = x$ line for most plots, specifically regenerator CO_2 concentration (Figure 4-6(e) and Figure 4-5(e)). This is due to the Aspen model calculation, which is able to calculate the most stable species in a gas stream at each temperature, which in this case at the given temperature range, is CO_2 . This is translated into the Python models, as the set of five features used from Aspen, allows the model to obtain a more accurate prediction.

Whilst the majority of the graphs are fairly uniform in the points plotted around the line of best fit and $y = x$ line, there are some apparently anomalous datapoints seen in both Figure 4-5(f) and Figure 4-6(c). For Figure 4-5(f), neural network CH_4 conversion, there is a clear curve of data points above, and a few points below

the line of best fit and $y = x$ line. To eliminate this being a mistake, this plot was run in Python multiple times, with different epochs and number of nodes in the layers, with the same pattern being detected. An explanation for this is down to the raw data and potentially, that an insufficient set of features was selected to identify all reactions occurring, for example more components than necessary were used leading to over or underfitting. When analysing the raw data, it is evident that whilst 0.99% was the maximum for the conversion, at the lower end, 0.02% was the minimum however this was for a specific set of inputs, i.e. the results go from 0.99 to 0.02% in a set pattern, with little to no results at 0-0.4% for higher input feature conditions, and the same at 0.6-0.99% for lower input feature data, creating a curve in the overall raw dataset.

For the Figure 4-6(c), random forest H_2 concentration in the reformer, an explanation for the plateau of datapoints at around 0.5% and again between 0.7-0.8%, is that since H_2 is the final and main product gas from all the reactions taking place in the reformer, it is essentially least affected by slight changes in input features. This is because it is the agglomeration of 3 different reactions - Equations (2-1), (2-2) and (2-4) thus a lot of the different combinations of inputs (S/C, CaO/C, temperature, pressure), result in the same H_2 concentrations, which can be seen when analysing the raw data, and can also be seen practically. Additionally, reasons the plateau of data points occurs at the aforementioned concentrations percentages, can be attributed to the concentration increasing from Equation 2-1 (SMR), then again with Equation 2-2 (WGS). The data in this study is theoretical and manually adjusted in the Aspen model however, in literature concerning experimental studies on SE-SMR [52–54], there are a number of different conditions that are accepted as ‘optimal’, thus, the reason for different conditions resulting in a high H_2 concentration. This pattern in data can vaguely be seen in the neural network plot as well (Figure 4-5(c)), particularly around 0.7%, but not as clearly. Reasons for this fall under the fact that the ANN is a connected network of neurons, that are grouped in layers and process data in each layer sequentially before passing forward onto next layers, therefore, each message is passed on until the very last end results, and the output is a

more holistic representation of the input features. On the other hand, random forests are made up of decision trees, where within each tree, the input is processed and an output is predicted, independent of all the other trees [55]. Additionally, an assumption that RF takes on is that it relies on sampling being representative. For example, if one class consists of two components and in the dataset one component is represented by 100 samples, and another component is represented by one sample - most individual decision trees will see only the first component and RF will misclassify the second one. RF on the whole does a good job at classification but not as much for regression in some cases, as it does not give precise continuous nature prediction. In the case of regression, it does not predict beyond the range in the training data and can sometimes lead to nonsensical predictions if applied to extrapolation domains, as seen in Figure 4-6(c) [56]. Advantages of a data driven soft-sensor model, as opposed to a mathematical modelling program such as Aspen Plus, is that one is able to obtain real time variable measurements dependent on actual process conditions instead of manually inputting each component conditions in order to obtain a parameter output. Despite how thermodynamically accurate an Aspen Plus model is over a soft sensor, in practical terms the use of a representative model is more beneficial.

The use of an ANN or RF model acts as the code within a control system which in turn predicts the concentration of the gases, thus bypassing the use of hardware components, and with the metrics shown in Figure 4-5 and Figure 4-6, the predictions are highly accurate. Mathematical programming software such as Aspen Plus, have the feature of performing reactions at equilibrium, whereas actual experimental processes experience various limitations such as diffusional limitations and axial dispersion. For this reason, the use of a soft sensor is highly advantageous over a simulation program, due to the use of actual historical process data.

In addition, in practical terms, a soft sensor can be used as a backup sensor, when the hardware sensor is faulty or is removed for maintenance or

replacement, or it can be used in the process of quality control, in place of laboratory analysis, which is often not exhaustive enough [6,57].

4.5 Conclusion

In this work, a soft sensor model was developed through the application of artificial neural network (ANN) and random forest (RF) algorithms for the hydrogen production process by sorption enhanced steam methane reforming (SE-SMR). The choice and development of the models were discussed as well as parameters that were necessary to produce accurate predictions of the dataset provided. From the results, both models were shown to provide good performance predictions for gas concentrations in the reformer and regenerator reactors using five specified process features, in particular those of high concentrations. Additionally, the CH₄ conversion within the reformer was calculated and predicted with both models able to reflect the efficiency of the reaction very accurately. The random forest model indicated higher levels of accuracy with consistently higher R² values and lower mean absolute error (0.002-0.014) compared to the neural network model (0.005-0.024). Overall, this study concludes that ANN and RF algorithms can successfully model a nonlinear process such as SE-SMR. The developed models can be applied as a soft sensor for analysis and evaluation of a SE-SMR based hydrogen production plant and to predict key process performance indicators. The chemical properties of hydrogen and the sorbent material hinder the ability to rapidly upscale this process, therefore the use of machine learning models, that are a realistic representation of experimental data, is very useful in order to overcome such limitations. With hydrogen fast becoming a leader in clean energy systems, it is essential that all efforts are made to simplify and ensure the safety and feasibility in the process of up-scaling hydrogen production. The use of a data-driven soft sensor does exactly that and allows for further exploration and incorporation of artificial intelligence into this energy field. Future directions for this study would be to implement the soft sensor experimentally in order to compare the success rate of the soft sensor against hardware sensors, and thus further validating this machine learning application for the process of SE-SMR.

4.6 References

- [1] Edwards PP., Kuznetsov VL., David WIF., Brandon NP. Hydrogen and fuel cells: Towards a sustainable energy future. *Energy Policy*. December 2008; 36(12): 4356–4362. Available at: DOI:10.1016/j.enpol.2008.09.036
- [2] Yan W., Tang D., Lin Y. A Data-Driven Soft Sensor Modeling Method Based on Deep Learning and its Application. *IEEE Transactions on Industrial Electronics*. IEEE; May 2017; 64(5): 4237–4245. Available at: DOI:10.1109/TIE.2016.2622668
- [3] Fortuna L., Graziani S., Rizzo A., Xibilia MG. *Soft Sensors for Monitoring and Control of Industrial Processes*. London: Springer London; 2007. Available at: DOI:10.1007/978-1-84628-480-9
- [4] Wang K., Shang C., Liu L., Jiang Y., Huang D., Yang F. Dynamic Soft Sensor Development Based on Convolutional Neural Networks. *Industrial & Engineering Chemistry Research*. 3 July 2019; 58(26): 11521–11531. Available at: DOI:10.1021/acs.iecr.9b02513
- [5] Shang C., Yang F., Huang D., Lyu W. Data-driven soft sensor development based on deep learning technique. *Journal of Process Control*. Elsevier Ltd; March 2014; 24(3): 223–233. Available at: DOI:10.1016/j.jprocont.2014.01.012
- [6] Torgashov A., Zmeu K. Nonparametric soft sensor evaluation for industrial distillation plant. *Computer Aided Chemical Engineering*. Elsevier; 2015. pp. 1487–1492. Available at: DOI:10.1016/B978-0-444-63577-8.50093-0
- [7] Yuan X., Huang B., Wang Y., Yang C., Gui W. Deep Learning-Based Feature Representation and Its Application for Soft Sensor Modeling With Variable-Wise Weighted SAE. *IEEE Transactions on Industrial Informatics*. IEEE; July 2018; 14(7): 3235–3243. Available at: DOI:10.1109/TII.2018.2809730
- [8] Yao L., Ge Z. Deep Learning of Semisupervised Process Data With

Hierarchical Extreme Learning Machine and Soft Sensor Application. IEEE Transactions on Industrial Electronics. IEEE; February 2018; 65(2): 1490–1498. Available at: DOI:10.1109/TIE.2017.2733448

[9] Yan W., Shao H., Wang X. Soft sensing modeling based on support vector machine and Bayesian model selection. Computers and Chemical Engineering. 2004; 28(8): 1489–1498. Available at: DOI:10.1016/j.compchemeng.2003.11.004

[10] Sipöcz N., Tobiesen FA., Assadi M. The use of Artificial Neural Network models for CO₂ capture plants. Applied Energy. July 2011; 88(7): 2368–2376. Available at: DOI:10.1016/j.apenergy.2011.01.013

[11] Zamaniyan A., Joda F., Behroozsarand A., Ebrahimi H. Application of artificial neural networks (ANN) for modeling of industrial hydrogen plant. International Journal of Hydrogen Energy. Elsevier Ltd; May 2013; 38(15): 6289–6297. Available at: DOI:10.1016/j.ijhydene.2013.02.136

[12] Kim M., Han A., Lee J., Cho S., Moon I., Na J. Comparison of Derivative-Free Optimization: Energy Optimization of Steam Methane Reforming Process. International Journal of Energy Research. 3 June 2023; 2023: 1–20. Available at: DOI:10.1155/2023/8868540

[13] Lueg L., Schack D., Örs E., Schmidt R., Bickert P., von Kurnatowski M., et al. Data-driven Process Design Exemplified on the Steam Methane Reforming Process. 2021. pp. 1013–1019. Available at: DOI:10.1016/B978-0-323-88506-5.50156-X

[14] Schack D., Lueg L., Schmidt R., von Kurnatowski M., Ludl PO., Bortz M. Data-Driven Process Simulation Using Connected Surrogate Unit Models Exemplified on a Steam Methane Reforming Process. Chemie Ingenieur Technik. 29 December 2021; 93(12): 2052–2062. Available at: DOI:10.1002/cite.202100087

- [15] Lin B., Recke B., Knudsen JKH., Jørgensen SB. A systematic approach for soft sensor development. *Computers and Chemical Engineering*. 2007; 31(5–6): 419–425. Available at: DOI:10.1016/j.compchemeng.2006.05.030
- [16] Barelli L., Bidini G., Gallorini F., Servili S. Hydrogen production through sorption-enhanced steam methane reforming and membrane technology: A review. *Energy*. 2008; 33(4): 554–570. Available at: DOI:http://dx.doi.org/10.1016/j.energy.2007.10.018
- [17] Abbas SZ., Dupont V., Mahmud T. Modelling of H₂ production in a packed bed reactor via sorption enhanced steam methane reforming process. *International Journal of Hydrogen Energy*. Elsevier Ltd; July 2017; 42(30): 18910–18921. Available at: DOI:10.1016/j.ijhydene.2017.05.222
- [18] Antzara A., Heracleous E., Bukur DB., Lemonidou AA. Thermodynamic Analysis of Hydrogen Production via Chemical Looping Steam Methane Reforming Coupled with in Situ CO₂ Capture. *Energy Procedia*. Elsevier B.V.; 2014; 63: 6576–6589. Available at: DOI:10.1016/j.egypro.2014.11.694
- [19] García-Lario AL., Aznar M., Grasa GS., García T., Murillo R. Study of nickel catalysts for hydrogen production in sorption enhanced reforming process. *Journal of Power Sources*. Elsevier B.V; 2013; 242: 371–379. Available at: DOI:10.1016/j.jpowsour.2013.05.069
- [20] Erans M., Jeremias M., Manovic V., Anthony EJ. Operation of a 25 KWth calcium looping pilot-plant with high oxygen concentrations in the calciner. *Journal of Visualized Experiments*. 2017; 2017(128): 1–10. Available at: DOI:10.3791/56112
- [21] Hanak DP., Manovic V. Calcium looping combustion for high-efficiency low-emission power generation. *Journal of Cleaner Production*. Elsevier Ltd; 2017; 161: 245–255. Available at: DOI:10.1016/j.jclepro.2017.05.080
- [22] Li Z., Cai N., Yang J. Continuous Production of Hydrogen from Sorption-

Enhanced Steam Methane Reforming in Two Parallel Fixed-Bed Reactors Operated in a Cyclic Manner. *Industrial & Engineering Chemistry Research*. December 2006; 45(26): 8788–8793. Available at: DOI:10.1021/ie061010x

[23] Johnsen K., Ryu HJ., Grace JR., Lim CJ. Sorption-enhanced steam reforming of methane in a fluidized bed reactor with dolomite as CO₂-acceptor. *Chemical Engineering Science*. 2006; 61(4): 1195–1202. Available at: DOI:10.1016/j.ces.2005.08.022

[24] Arstad B., Probst J., Blom R. Continuous hydrogen production by sorption enhanced steam methane reforming (SE-SMR) in a circulating fluidized bed reactor: Sorbent to catalyst ratio dependencies. *Chemical Engineering Journal*. Elsevier B.V.; May 2012; 189–190: 413–421. Available at: DOI:10.1016/j.cej.2012.02.057

[25] Antzara A., Heracleous E., Bukur DB., Lemonidou AA. Thermodynamic analysis of hydrogen production via chemical looping steam methane reforming coupled with in situ CO₂ capture. *International Journal of Greenhouse Gas Control*. 2015; 32: 115–128. Available at: DOI:https://doi.org/10.1016/j.ijggc.2014.11.010

[26] Breiman L. Random Forests. *Machine Learning*. 2001; 45(1): 5–32. Available at: DOI:https://doi.org/10.1023/A:1010933404324

[27] James G., Witten D., Hastie T., Tibshirani R. Tree-Based Methods. *An Introduction to Statistical Learning*. 2013. pp. 303–335. Available at: DOI:10.1007/978-1-4614-7138-7_8

[28] Hutengs C., Vohland M. Downscaling land surface temperatures at regional scales with random forest regression. *Remote Sensing of Environment*. Elsevier Inc.; June 2016; 178: 127–141. Available at: DOI:10.1016/j.rse.2016.03.006

[29] Polishchuk PG., Muratov EN., Artemenko AG., Kolumbin OG., Muratov

NN., Kuz'min VE. Application of Random Forest Approach to QSAR Prediction of Aquatic Toxicity. *Journal of Chemical Information and Modeling*. 23 November 2009; 49(11): 2481–2488. Available at: DOI:10.1021/ci900203n

[30] Yan Y., Mattisson T., Moldenhauer P., Anthony EJ., Clough PT. Applying machine learning algorithms in estimating the performance of heterogeneous, multi-component materials as oxygen carriers for chemical-looping processes. *Chemical Engineering Journal*. Elsevier; May 2020; 387(September 2019): 124072. Available at: DOI:10.1016/j.cej.2020.124072

[31] Ayodele B V., Cheng CK. Modelling and optimization of syngas production from methane dry reforming over ceria-supported cobalt catalyst using artificial neural networks and Box–Behnken design. *Journal of Industrial and Engineering Chemistry*. The Korean Society of Industrial and Engineering Chemistry; December 2015; 32: 246–258. Available at: DOI:10.1016/j.jiec.2015.08.021

[32] Vo ND., Oh DH., Hong SH., Oh M., Lee CH. Combined approach using mathematical modelling and artificial neural network for chemical industries: Steam methane reformer. *Applied Energy*. Elsevier; 2019; 255(July): 113809. Available at: DOI:10.1016/j.apenergy.2019.113809

[33] Lang Y., Malacina A., Biegler LT., Munteanu S., Madsen JL., Zitney SE. Reduced Order Model Based on Principal Component Analysis for Process Simulation and Optimization. *Energy & Fuels*. 19 March 2009; 23(3): 1695–1706. Available at: DOI:10.1021/ef800984v

[34] Ye F., Ma S., Tong L., Xiao J., Bénard P., Chahine R. Artificial neural network-based optimization for hydrogen purification performance of pressure swing adsorption. *International Journal of Hydrogen Energy*. February 2019; 44(11): 5334–5344. Available at: DOI:10.1016/j.ijhydene.2018.08.104

[35] Xiao J., Li C., Fang L., Böwer P., Wark M., Bénard P., et al. Machine learning–based optimization for hydrogen purification performance of layered bed pressure swing adsorption. *International Journal of Energy Research*. May 2020;

44(6): 4475–4492. Available at: DOI:10.1002/er.5225

[36] Gardner MW., Dorling SR. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment*. August 1998; 32(14–15): 2627–2636. Available at: DOI:10.1016/S1352-2310(97)00447-0

[37] Matsuo R., Yamazaki T., Suzuki M., Toyama H., Araki K. A random forest algorithm-based approach to capture latent decision variables and their cutoff values. *Journal of Biomedical Informatics*. Elsevier Inc.; October 2020; 110(August): 103548. Available at: DOI:10.1016/j.jbi.2020.103548

[38] Şener AN., Günay ME., Leba A., Yıldırım R. Statistical review of dry reforming of methane literature using decision tree and artificial neural network analysis. *Catalysis Today*. January 2018; 299(May 2017): 289–302. Available at: DOI:10.1016/j.cattod.2017.05.012

[39] Gevrey M., Dimopoulos I., Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. *Ecological Modelling*. February 2003; 160(3): 249–264. Available at: DOI:10.1016/S0304-3800(02)00257-0

[40] Olden JD., Joy MK., Death RG. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. *Ecological Modelling*. November 2004; 178(3–4): 389–397. Available at: DOI:10.1016/j.ecolmodel.2004.03.013

[41] Zhang Z., Beck MW., Winkler DA., Huang B., Sibanda W., Goyal H. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. *Annals of Translational Medicine*. June 2018; 6(11): 216–216. Available at: DOI:10.21037/atm.2018.05.32

[42] Elmolla ES., Chaudhuri M., Eltoukhy MM. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by

the Fenton process. *Journal of Hazardous Materials*. Elsevier B.V.; July 2010; 179(1–3): 127–134. Available at: DOI:10.1016/j.jhazmat.2010.02.068

[43] Jolliffe IT., Cadima J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 13 April 2016; 374(2065): 20150202. Available at: DOI:10.1098/rsta.2015.0202

[44] Ringnér M. What is principal component analysis? *Nature Biotechnology*. March 2008; 26(3): 303–304. Available at: DOI:10.1038/nbt0308-303

[45] Song F., Guo Z., Mei D. Feature Selection Using Principal Component Analysis. 2010 International Conference on System Science, Engineering Design and Manufacturing Informatization. IEEE; 2010. pp. 27–30. Available at: DOI:10.1109/ICSEM.2010.14

[46] Luo XJ., Oyedele LO., Ajayi AO., Akinade OO., Delgado JMD., Owolabi HA., et al. Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings. *Energy and AI*. Elsevier Ltd; November 2020; 2: 100015. Available at: DOI:10.1016/j.egyai.2020.100015

[47] Brownlee J. What is the Difference Between a Batch and an Epoch in a Neural Network? *Machine Learning Mastery*. 2018. pp. 3–4. Available at: https://deeplearning.lipingyang.org/wp-content/uploads/2018/07/What-is-the-Difference-Between-a-Batch-and-an-Epoch-in-a-Neural-Network_.pdf
(Accessed: 6 April 2020)

[48] Feng J., Lu S. Performance Analysis of Various Activation Functions in Artificial Neural Networks. *Journal of Physics: Conference Series*. June 2019; 1237(4): 022030. Available at: DOI:10.1088/1742-6596/1237/2/022030

[49] Gao X., Wen J., Zhang C. An Improved Random Forest Algorithm for Predicting Employee Turnover. *Mathematical Problems in Engineering*. 17 April

2019; 2019: 1–12. Available at: DOI:10.1155/2019/4140707

[50] Scornet E. Tuning parameters in random forests. Coeurjolly J-F, Leclercq-Samson A (eds.) ESAIM: Proceedings and Surveys. 14 December 2017; 60: 144–162. Available at: DOI:10.1051/proc/201760144

[51] Gotelli NJ. Research frontiers in null model analysis. Global Ecology and Biogeography. 2001; 10(4): 337–343. Available at: DOI:10.1046/j.1466-822X.2001.00249.x

[52] Ding Y., Alpay E. Adsorption-enhanced steam-methane reforming. Chemical Engineering Science. 2000; 55(18): 3929–3940. Available at: DOI:10.1016/S0009-2509(99)00597-7

[53] Di Giuliano A., Gallucci K. Sorption enhanced steam methane reforming based on nickel and calcium looping: a review. Chemical Engineering and Processing - Process Intensification. Elsevier; 2018; 130(June): 240–252. Available at: DOI:10.1016/j.cep.2018.06.021

[54] Anderson DM., Kottke PA., Fedorov AG. Thermodynamic analysis of hydrogen production via sorption-enhanced steam methane reforming in a new class of variable volume batch-membrane reactor. International Journal of Hydrogen Energy. Elsevier Ltd; 2014; 39(31): 17985–17997. Available at: DOI:10.1016/j.ijhydene.2014.03.127

[55] Liaw A., Wiener M. Classification and Regression by Random Forest. R News. 2002; 2(3): 18–22.

[56] Hengl T., Nussbaum M., Wright MN., Heuvelink GBM., Gräler B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ. 29 August 2018; 6(8): e5518. Available at: DOI:10.7717/peerj.5518

[57] Slišković D., Grbić R., Hocenski Ž. Methods for plant data-based process modeling in soft-sensor development. Automatika. 2011; 52(4): 306–318.

Available at: DOI:10.1080/00051144.2011.11828430

5 Prediction of Combined Sorbent Catalyst and Catalyst Materials for SE-SMR using QSPR and Multitask Learning

Paula Nkulikiyinka, Stuart T. Wagland, Vasilije Manovic, Peter T. Clough

Energy and Power Theme, School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK.

DOI: <https://doi.org/10.1021/acs.iecr.2c00971>

Abstract

The process of Sorption Enhanced Steam Methane Reforming (SE-SMR) is a popular emerging technology for the production of low carbon hydrogen. Upscaling this process faces issues in the development of a suitable catalytic material as well as a CO₂ adsorbent with high capture capacity. In this study, to aid the development of a Combined Sorbent Catalyst Material (CSCM) for SE-SMR, a novel approach involving quantitative structure–property relationship analysis (QSPR) has been proposed. Through data mining, two databases have been developed for the prediction of the last cycle capacity (gCO₂/gSorbent) and methane conversion (%), wherein 'last' cycle refers to the cycle once the sorbent is spent i.e., saturated. Multi-task learning (MTL) was applied for the prediction of CSCM properties. Further information yielded from this study is patterns in data; Colour scatter plots were able to show certain patterns in the input data as well as suggestions on how to develop an optimal material. With all the results from the Actual vs Predicted plots collated raw materials and synthesis conditions were proposed that could lead to the development of a CSCM that has good performance with respect to both the last cycle capacity and the methane conversion.

5.1 Introduction

Hydrogen is seen as an attractive energy source for many reasons, including the high potential to reduce carbon emissions by decarbonising multiple sectors, being more efficient due to the high gravimetric energy density, compared to hydrocarbon fuels such as diesel or gasoline [1,2]. This therefore gives hydrogen the ability to store a lot more energy for every unit weight of fuel in comparison. Hydrogen production as a fuel source is vital in the race towards the 2050 Net-Zero Target and one of the most promising methods for the future production of low carbon H₂ is via sorbent enhanced steam methane reforming (SE-SMR) [3], (Equation (2-5)), which combines the two processes of steam methane reforming (SMR) (Equation (2-3)) and calcium looping (CaL) (Equation (2-4)).



In the SE-SMR process, the reforming reaction takes place in the presence of a calcium oxide (CaO)-based sorbent, which allows it to capture in-situ the produced CO₂, which then drives the reaction equilibrium of reaction 1 to the right, leading to higher hydrogen yields. So, whilst the introduction of a sorbent in the conventional steam reforming process is beneficial for those two reasons, there are also some challenges that come with the use of a sorbent.

Firstly, the selection of sorbent itself poses an issue as the material of choice must possess several properties in order for efficient operation in terms of economics as well and operating conditions. One of the important requirements of a sorbent in the use of SE-SMR is its' carrying capacity over multiple cycles [4,5]. Other important factors include:

- The ability for the adsorbent to be regenerated.
- The adsorbent having a high CO₂ capacity in the kinetic limited regime.
- Economically viable with a potential export market for the spent material

- Having fast reaction kinetics for the absorption and desorption steps
- And maintains adequate mechanical strength after multiple cycles.

From early studies on sorbents used in steam methane reforming, further knowledge has been gained on what makes an efficient and appropriate sorbent material. For example, additional factors include durability comparable to the catalysts, to minimize the purge requirements for the spent sorbent [6]; and an adequate pore size distribution with large pore volume in the 50-100 μm range and active surface of fresh sorbent [7]. Reaction temperature, carbonation and calcination duration times and atmospheric conditions, sorbent precursor, and sorbent particle size are key influencers in determining the suitability of a sorbent for SE-SMR. A principal indicator of sorbent performance is the number of reaction cycles and the capacity capability at the last cycle, which essentially suggests the lifetime of a sorbent before replenishment is required [8].

Quicklime (CaO) is a particularly favourable material due to its low cost and wide availability in naturally occurring minerals, e.g., limestone or dolomite. Despite calcium-based sorbents being a favoured material, a disadvantage of naturally occurring calcium-based materials is the rapid decrease of their CO_2 uptake capacity with cycle number due to sintering or pore plugging. Some proposed mechanisms to improve the CO_2 capture characteristics of CaO , include methods such as hydration or thermal pre-treatment [9]. The choice of the CaO precursor has also been reported to have an effect on the sorption properties of the final synthesised material, with calcium acetate proven to a high performer [10].

Another challenge faced in the SE-SMR process is the catalyst selection. Catalyst materials used for SE-SMR should resist coke formation and sulphation poisoning, be inactive for side-reactions, maintain the activity at high temperature, and have high mechanical strength as well as good heat transfer properties [11]. Preferably, they should also be able to operate at low steam/carbon ratios in order to improve the energy efficiency of the process [12]. The mechanisms of several catalysts (nickel, rhodium, platinum, ruthenium, and iridium) for SMR reactions were investigated [13], where it was reported that

platinum catalysts were the most reactive however, among these, nickel oxide has widely been proved to be preferable for use in SE-SMR due to its high catalytic activity i.e., high conversion rate of methane and lower cost compared to rare earth elements [14-16]. However, nickel does have its disadvantages [17]:

- Nickel nanoclusters (as opposed to a skeletal structure) are prone to sintering from the high temperatures.
- A tendency towards coking leading to particle fracturing and active site loss.
- The production of hydrogen rich gas with a low concentration of CO is a challenge using nickel catalysts, as they aren't as active in the water gas shift reaction as other catalysts.

Other factors to consider in catalyst choice include the sensitivity to sulphur poisoning. It is assumed that the natural gas feed will have undergone desulfurization, however, most naturally occurring limestones and dolomites may still contain small quantities of sulphur. Therefore, in the reforming reaction, with enough sulphur present, it can be transferred from the sorbent to the gas phase to quickly poison the nickel reforming catalyst [18]. The use of metal dispersion and alloying (particularly with platinum, palladium, or ruthenium), are both effective methods to enhance certain properties of catalysts, including resistance to sintering as well as resistance to nickel oxidation, which promotes its activation [19]. The addition of a metal such as iron, copper or tin, also has the possibility to improve the catalytic activity [17].

The concept of a material that possess both catalytic and adsorbing properties is not a new one and has been of interest for the last 20 years. A combined sorbent catalyst material (CSCM), or a bifunctional material, is a one-particle system that contains a metal for the catalysis and a sorbent for the in-situ CO₂ sorption. Due to the intimate contact of catalyst and sorbent in the solid, the CSCM particle system enables greater heat transfer and reduced mean path length for mass transfer of the reaction products, which enables a greater conversion efficiency

compared to a two-particle system [20]. Additionally, the overall reactor volume can be reduced by using a sorbent that acts as a support.

There have been different approaches taken to develop and predict the behaviour of optimal properties for a suitable CSCM, such as the determination of kinetic and diffusion parameters and computational modelling [21-25], as well as material synthesis and experimental studies in a fluidised bed reactor [26] or fixed bed reactors [27-29]. The application of CSCM also isn't limited to the process of SE-SMR with methane as a carbon source, as there are many studies conducted on CSCM applied in sorption enhanced steam reforming using other raw materials such as glycerol [30,31], phenol [32], coal [33], biomass [20], ethanol [34], methanol [35].

The development of an 'optimal' CSCM has proven to be a complicated and monotonous one, however this can be overcome with the application of machine learning, specifically the process of quantitative structure–property relationship analysis (QSPR). This analysis is based on using the molecular structure of a material to predict its' physical behaviour. QSPR modelling has long been used particularly in the pharmaceutical industry for drug development, and it has recently seen expansion into various physical chemistry fields, including adsorption of metal organic frameworks [36-39]. Additionally, the use of QSPR in combination with other computational methods such as DFT or process simulation has also garnered effective prediction within research in the energy field [40,41], however, to the best of our knowledge, QSPR has not been applied to the development of a CSCM material for use in SE-SMR.

In this study, we propose a new QSPR machine learning approach for the prediction of a CSCM, through the development of a calcium-based sorbent database and nickel-based catalyst database, consisting of SE-SMR and SMR experimental data from literature. The properties of interest used as a measure of the performance for the CSCM, is the methane conversion and the CO₂ adsorption capacity for the last cycle calcination/carbonation. The individual databases were firstly developed into QSPR models for the prediction of separate

sorbent and catalyst materials, then eventually integrated for the prediction of unseen CSCM materials.

5.2 Methodology

The methodology typically used in QSPR studies consist of six steps that fall into 4 main categories outlined in Figure 5-1.

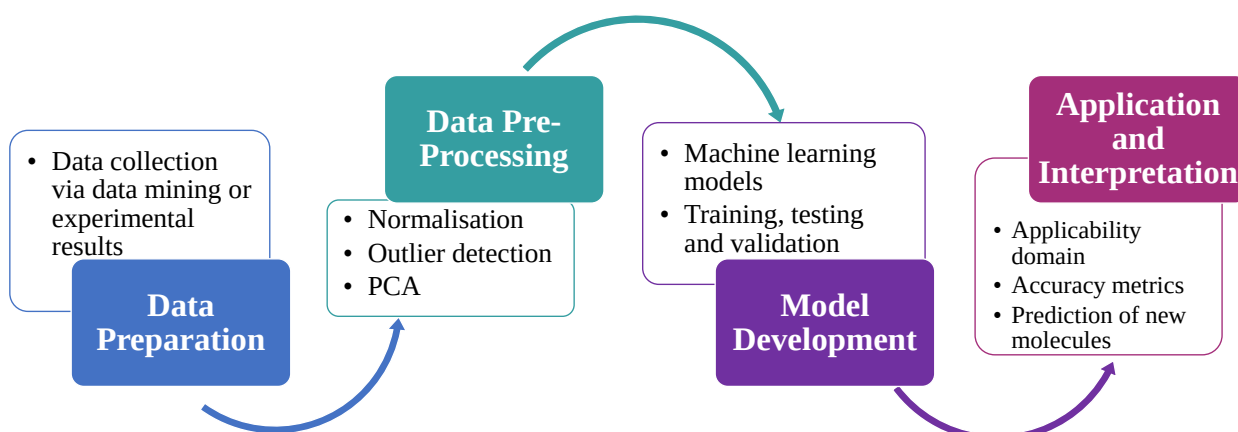


Figure 5-1 Process flow diagram of the development of QSPR models

Preparation of data:

1. The collection or measurement of property data points (here, the last cycle capacity and the methane conversion),
2. Collection of descriptor data, (descriptor data was selected via the OCHEM software. Different descriptor sets were trialled and compared. Examples include constitutional or topological descriptors)

Data pre-processing:

3. Analysis of the data to ensure its suitability; application of processes such as feature selection, normalisation, standardisation, and outlier detection. (Specifically in this study, the pre-processing included standardization, neutralization, removing salts (where applicable since salts are not typically considered in QSPR models as most software doesn't calculate molecular descriptors on non-full connected structure) and cleaning the structure, which were processes were applied in the OCHEM software, prior to model predictions.)

Model development

4. Training and validation of the model; development of various machine learning models to obtain high prediction metrics.

Model application and interpretation

5. Recognition of the applicability domain
6. Statistical evaluation and interpretation of the model

In addition to using examples from literature [42,43], care was made to ensure the development followed the principles outlined by the Organisation for Economic Co-operation and Development (OECD), which are guidelines that were formulated in 2002 by QSPR experts to regulate the use and development of these models [44]. Further details are given in Appendix B, Table_Apx B-1.

The stages involved in QSPR development are described in the following sections:

5.2.1 Database Development

The data collection method took the form of a data-mining process, often used in model building, where data is gathered through sifting through large amounts of data from literature in order to obtain new patterns, correlations, or structures in data [45]. Literature ranging from around 20 years ago to the present was studied to build up two separate databases, and eventually an overall combined database, consisting of the properties of interest (last cycle sorbent capacity and catalyst methane conversion) and well as the experimental conditions used to obtain these results. Care was taken to ensure the data collected was based on experimental measured data as opposed to computational modelling work. Additionally, some literature was omitted where the properties were present in the work, however with a large percentage of the details of the conditions missing, as this would cause a skew in the ability to predict the properties. For instance, the sorbent database went from a size of 248, to 239, and the catalyst database was reduced from 249 datapoints to 183, due to the aforementioned reasoning. The conditions chosen to aide in the predictions of the properties are shown in

Table 5-1. These were chosen for two reasons; they have been proven to have an influence on the properties' outcome. Additionally, principal component analysis (PCA) was conducted to confirm that none of the input categories were not redundant (Figure 5-2 and Figure 5-3).

Table 5-1 Properties and their respective experimental conditions per database.

Database	Property (units)	Experimental conditions (units)
Sorbent	Last cycle capacity (gCO ₂ /gSorbent)	CaO concentration (%) of the whole sorbent mass, cycle number, calcination and carbonation temperatures (°C), synthesis method, CaO precursor, initial cycle capacity (gCO ₂ /gSorbent)
Catalyst	Methane conversion (%)	Nickel concentration (wt%) with respect to the whole catalyst monolith weight, calcination temperature (C), calcination duration (h), SMR reaction temperature (°C) fresh BET surface area (m ² /g), steam/carbon ratio

PCA is a mathematical model that is used to reduce the dimensionality of a dataset, whilst retaining most of the variation in the dataset and as an effective procedure for the determination of input parameters [46]. The reduction of features is achieved by identifying the directions (principal components) that the variation is highest [47,48]. From the PCA heatmaps, the variance across the 9 and 6 principal components (PCs) for the sorbent and catalyst databases respectively, was shown to be fairly evenly distributed across the first few PCs (e.g., for the catalyst database, the first PC explains 32% of the variance, PC2, PC3, PC4, PC5 and PC6 explain 20.3%, 15.7%, 13.3% 11.6% and 7.1% respectively). Therefore, it was not essentially necessary to remove any parameters. That would be necessary if the first two components contributed to around 70-95% of the total variance. (See Appendix B, Table_Apx B-2 for sorbent variance data).

Additional data garnered from the PCA heatmaps was feature importance. From Figure 5-2 it can be calculated that the top three features that contribute to the PCs of last cycle capacity, are the precursor type, CaO concentration and calcination temperature, and from Figure 5-3, the top three to contribute to methane conversion are calcination temperature, S/C ratio and BET surface area. (See Appendix B, Table_Apx B-3 for full feature importance data). This is evident as they are the PCs that have the largest total.

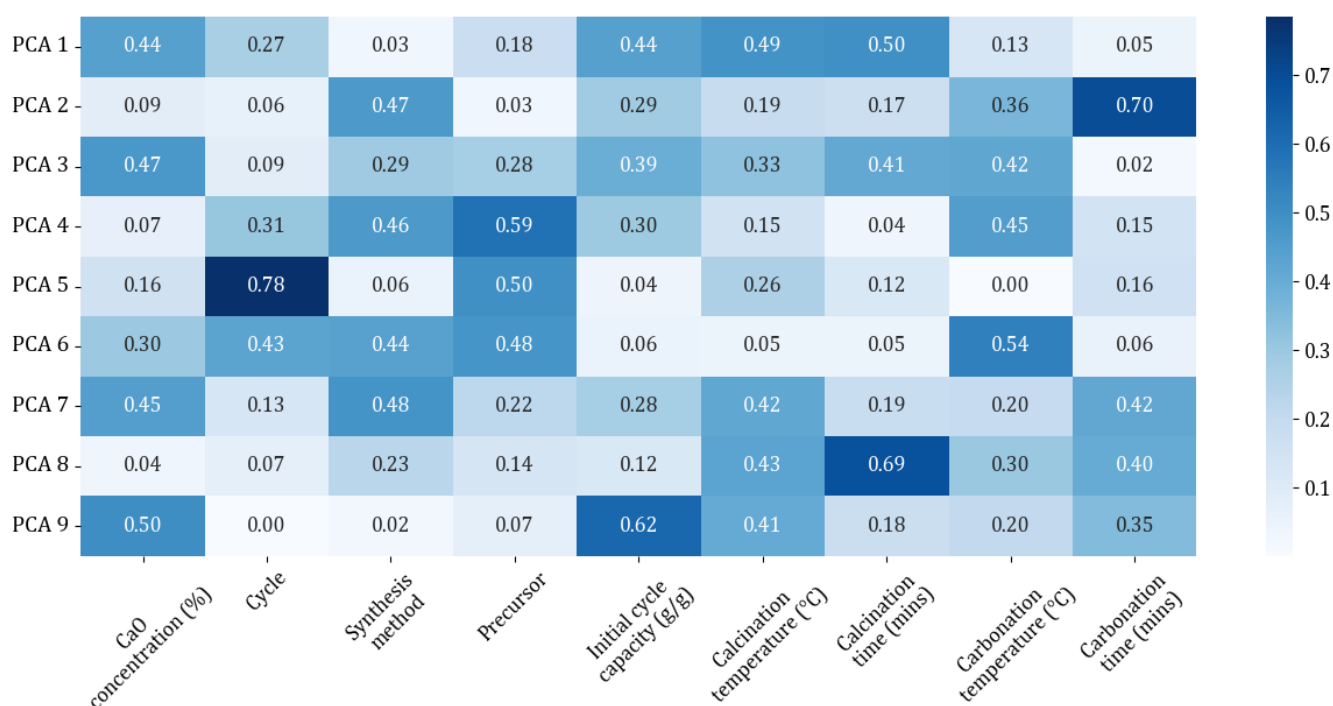


Figure 5-2 Correlation heatmap of sorbent input features.

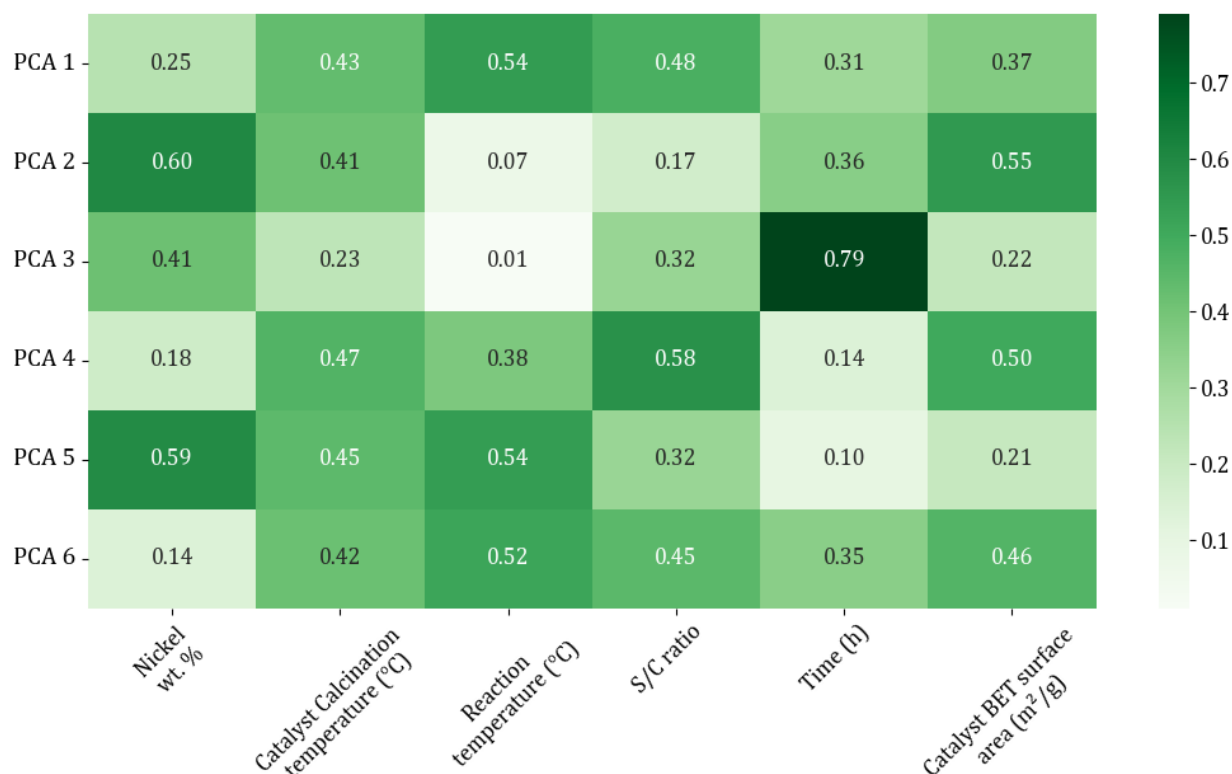


Figure 5-3 Correlation heatmap of catalyst input features.

See Appendix B, Table_Apx B-4 and B-5 for the individual sorbent and catalyst databases respectively, and Table_Apx B-9 for the CSCM literature data.

5.2.2 Molecular Descriptors

Once the experimental conditions were confirmed, the molecular descriptors were evaluated. Molecular descriptors are defined as mathematical representations of a molecule, from characteristics related to the structure of the chemical [43]. They are calculated quantum chemistry methods and allow for the modelling of many different properties in fields such as physical chemistry, pharmaceuticals, and analytical chemistry [49].

The use of different descriptor sets has been shown to result in different performance in the modelling of the properties of interest [50]. Therefore, a selection of descriptors was trialled by measuring the accuracy metrics with varying descriptors using the OCHEM platform [51], which is an online chemical database with a QSPR/AR modelling environment. A list of the descriptors trialled

is given in Table 5-2, [50-55]. Alongside the molecular descriptors and experimental conditions from literature, an additional fundamental input was the structures of the molecules in SMILES format (simplified molecular-input line-entry system), i.e., CaO would be represented as O=[Ca].

Table 5-2 Descriptors used for comparison.

Descriptor	Description
ALogPS, OEstate	Prediction of logP by ALogPS2.1 program, calculated from each atoms' intrinsic electronic properties and the influence of other atoms in the molecule
Fragmentor	Molecular fragments which contain from 2-4 atoms generated by the ISIDA module in OCHEM
GSFrag	Descriptors based on the occurrence of certain special fragments
alvaDesc v.2.0.2 (3D)	Calculates and analyses molecular descriptors and fingerprints, such as constitutional, topological, and geometrical descriptors
QNPR	(Quantitative Name Property Relationship) Uses substrings of SMILES or IUPAC name as a representation of molecules

5.2.3 Machine Learning Models

Several types of models were used in this study along with the application of the inductive transfer approach. This is the use of a combination of several conventional single task learning (STL) machine learning models, used in parallel, also known as multitask learning (MTL), to act as additional nodes of the original model as opposed to separate models, as shown in Figure 5-4, [56,57].

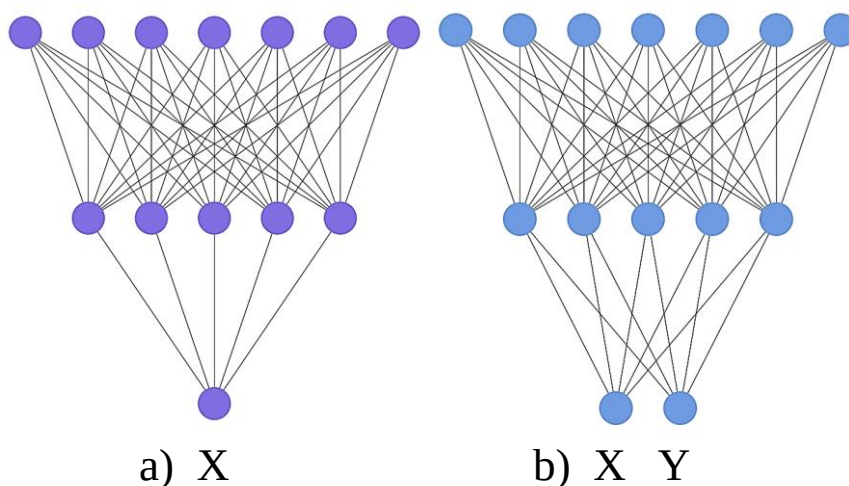


Figure 5-4 Schematic of a neural network configuration using different inductive transfer approaches (a) conventional single task learning (STL) and (b) multitask learning (MTL)

5.2.4 Single Task Learning (STL)

The conventional approach of STL in QSPR, is to focus on obtaining the target of one property, using only descriptors and conditions relating to that single property. This was the first approach taken to predict the properties of last cycle capacity and methane conversion using their respective databases, alongside comparing performance with descriptors. The machine learning models trialled were associative neural network (ASNN), deep neural network (DNN) and least-squares support-vector machine (LSSVM). These models were chosen as they have the ability to be applied as MTL models, as well as with the use of molecular descriptors on the OCHEM platform.

A built-in function of OCHEM called unsupervised forward selection (UFS) was implemented. This is a method used for eliminating redundant descriptors and is specific for use in the development of QSPR models. Redundancy can be common in QSPR datasets due to high or exact linear dependencies between subsets of the variables and high multiple correlations between subsets of the variables. These factors hinder the development of models that have the ability to effectively predict new data. UFS produces a reduced dataset that contains no redundancy and has a minimal amount of multicollinearity, via the method

depicted in Figure 5-5 [58]. See Supporting Information for other descriptor filter settings in addition to UFS.

ASNN, DNN and LSSVM were implemented in OCHEM, and then compared against each other to firstly observe the effect on ability to predict the properties using molecular descriptors with experimental conditions, then the best performing algorithms were used to further predict unseen properties using more data from literature. Details on the architectures of the machine learning models can be found in Appendix B, Table_Apx B-6 – B-8).

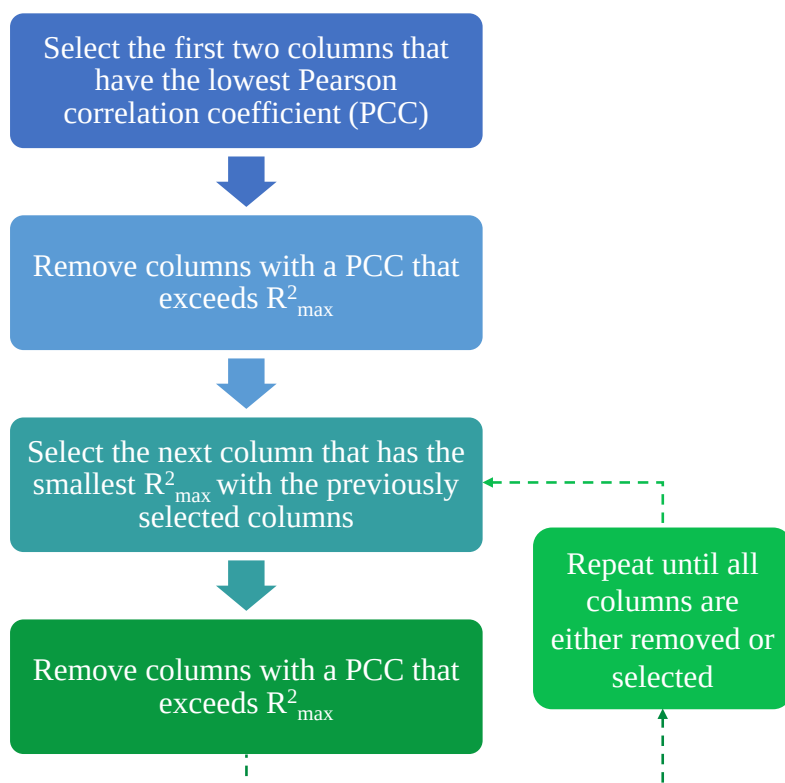


Figure 5-5 Process flow diagram showing the steps involved in unsupervised forward selection (UFS) to remove unneeded molecular descriptors

5.2.5 Multitask Learning (MTL)

With MTL, multiple properties are trained and learned in parallel [59]. This approach improves the performance of STL due to the training data from the extra tasks acting as a suggestive bias, adding in effect constraints for the other task at hand, which aides in the accuracy and the speed of learning [60].

Two types of comparisons were conducted, the first being a combination of finding the machine learning model and molecular descriptor set, resulted in the highest prediction accuracy. For efficiency, this was trialled using STL data only, and the models were conducted using OCHEM's automatic multiple models setting. The results from this were used in the second comparison which was STL compared against MTL, all using default model settings, in order to observe which inductive transfer approach was the best performing for this data.

Following this, the approach that yielded the best accuracy metrics, was then taken and the model settings were optimised to obtain the best predicting model, which was then used to predict unseen CSCM molecules.

5.2.6 Model Development and Performance Measurement

Each model was validated using a grid search method based on 5-fold cross-validation procedure, which is useful to mitigate overfitting, parameter optimisation as well as evaluating the predictive validity of linear regression models [61]. On the OCHEM platform, the 5-cross validation process, develops a new model on each validation step without the use of known information about the molecule, as they are only calculated after completion of the model developed. This approach is reported to be the correct validation approach, as no prior information about the test molecules is used to skew the models' development [50]. Additionally, the models themselves had internal validation which can be seen in the Supporting Information.

To assess the predictive performance, the accuracy metrics used were root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2), shown in Equations (4-5) to (5-1). The aims of the machine learning models was to obtain the highest R^2 and lowest RMSE and MAE, where n is the quantity of data required for the network training, $\hat{y}$ is the predicted value of y , y_i is the target output, and $\bar{y}$ is the mean value of y .

$$RMSE = \sqrt{\frac{1}{n} \left[\sum_{i=1}^n (y_i - \hat{y})^2 \right]} \quad (4-5)$$

$$MAE = \frac{1}{n} \left[\sum_{i=1}^n |y_i - \hat{y}| \right] \quad (4-6)$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2} \quad (5-1)$$

Additionally, a recommended tool when dealing with large datasets, in order to evaluate the models' prediction capabilities of new molecules, the absolute relative error (ARE) for each molecule was calculated using Equation (5-2) as well as the models' average (AARE), Equation (5-3) where n is the number of molecules predicted per model, and i is the predicted molecule [62].

$$ARE_i = \left| \frac{y_i - \hat{y}}{y_i} \right| \quad (5-2)$$

$$AARE = \frac{\sum_{i=1}^n ARE_i}{n} \quad (5-3)$$

5.3 Results and Discussion

5.3.1 Comparison of the Models

A comparison of molecular descriptors and machine learning models for the sorbent database, Figure 5-6, shows that the ASNN model gave the highest R^2 and lowest RMSE, across all models. In terms of the 'most optimal' set of descriptors, although not clear in the figure, AlogPS/OEstate, Fragmentor and GSFrag resulted in equally strong values whilst QNPR and alvaDesc consistently underperformed for this database. On the other hand, interestingly, as also shown in Figure 5-6, a comparison of machine learning models and descriptors for the catalyst database, whilst the best performing model was in agreement with the sorbent database (ASNN), the descriptors that performed best were the opposite.

Here, alvaDesc and QNPR were the top two and Fragmentor was the least accurate predictor.

ASNN stood out with both the catalyst and sorbent results, and suggestions for this include the learning method behind this algorithm. It is a combination of an ensemble of the feedforward artificial neural network model (ANN) and the K-nearest neighbour technique (kNN). ANN is a supervised learning algorithm, which does not require any pre-assumption of the input-output relationship [63]. This would have a positive effect on the prediction capability as it would allow the model to rely more on the molecular descriptors calculated to provide a prediction. The model uses the correlation between ensemble responses as a measure of distance within the analysed cases for the kNN [57], which corrects the bias of the neural network ensemble [64], and the combination of two algorithms, proved to be optimal in this case.

As all five descriptors resulted in high predictions for either of the databases, it was too ambiguous to decide which was optimal at this stage, therefore it was decided to take all sets to the next step of applying the inductive transfer approach of multitask learning, to see if what effect, if any, this approach had on the results, which yielded interesting results.

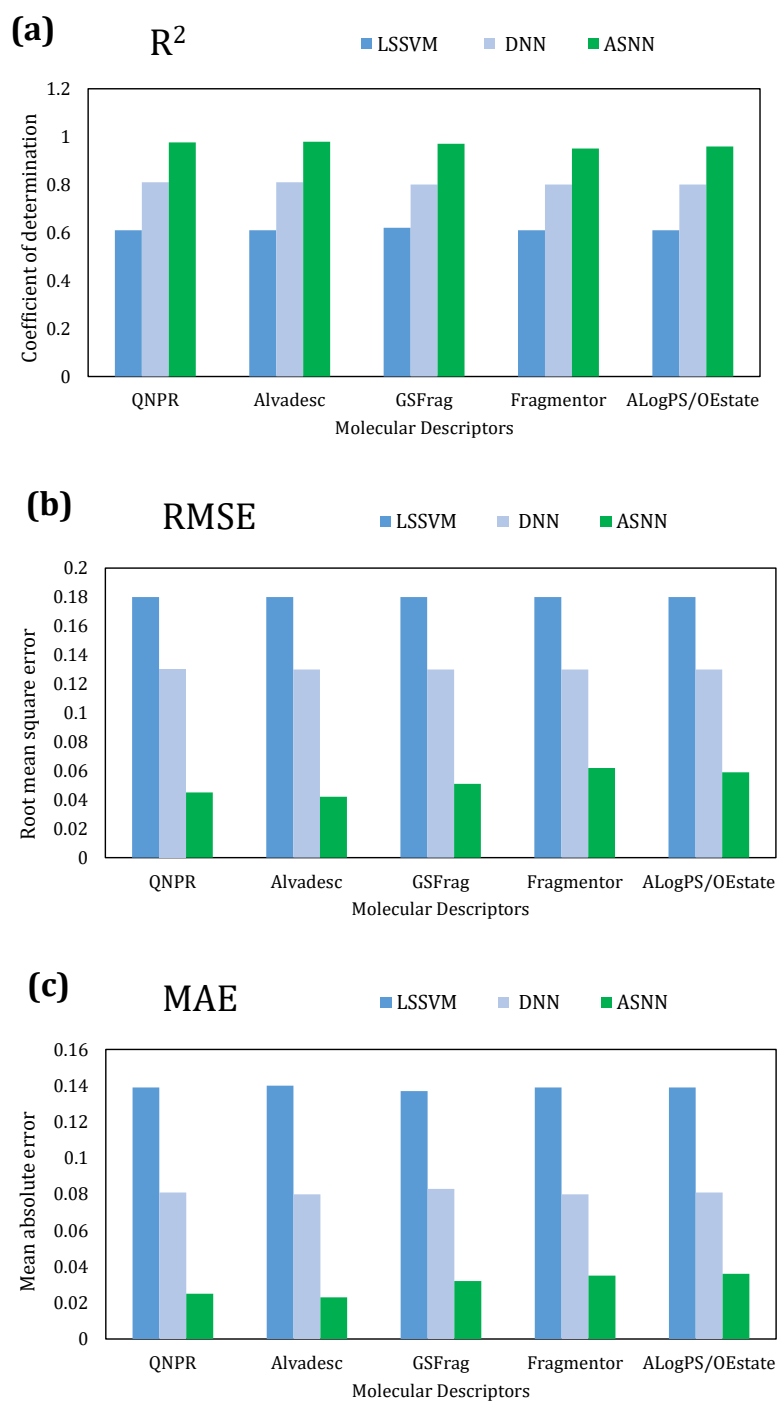


Figure 5-6 Comparison of machine learning models- least square support vector machine (LSSVM), deep neural network (DNN), associative neural network (ASNN) and molecular descriptors on the catalyst methane conversion database, (a) coefficient of determination (R^2), (b) root mean square error (RMSE) and (c) mean absolute error (MAE).

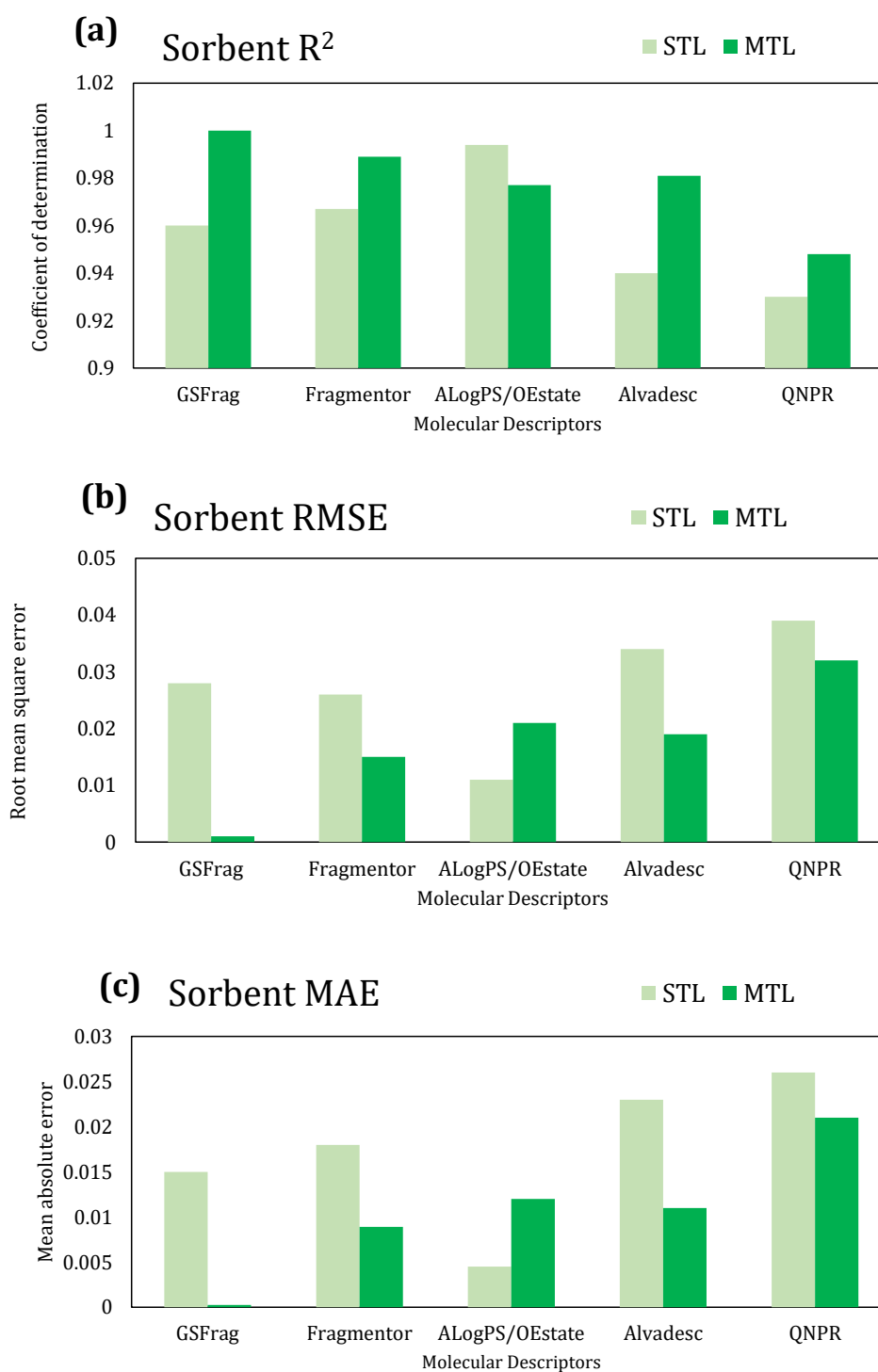


Figure 5-7 Comparison of the effect of the inductive transfer approach and molecular descriptors, on the associative neural network (ASNN) algorithm, for the sorbent database (a) coefficient of determination (R^2), (b) root mean square error (RMSE), (c) mean absolute error (MAE).

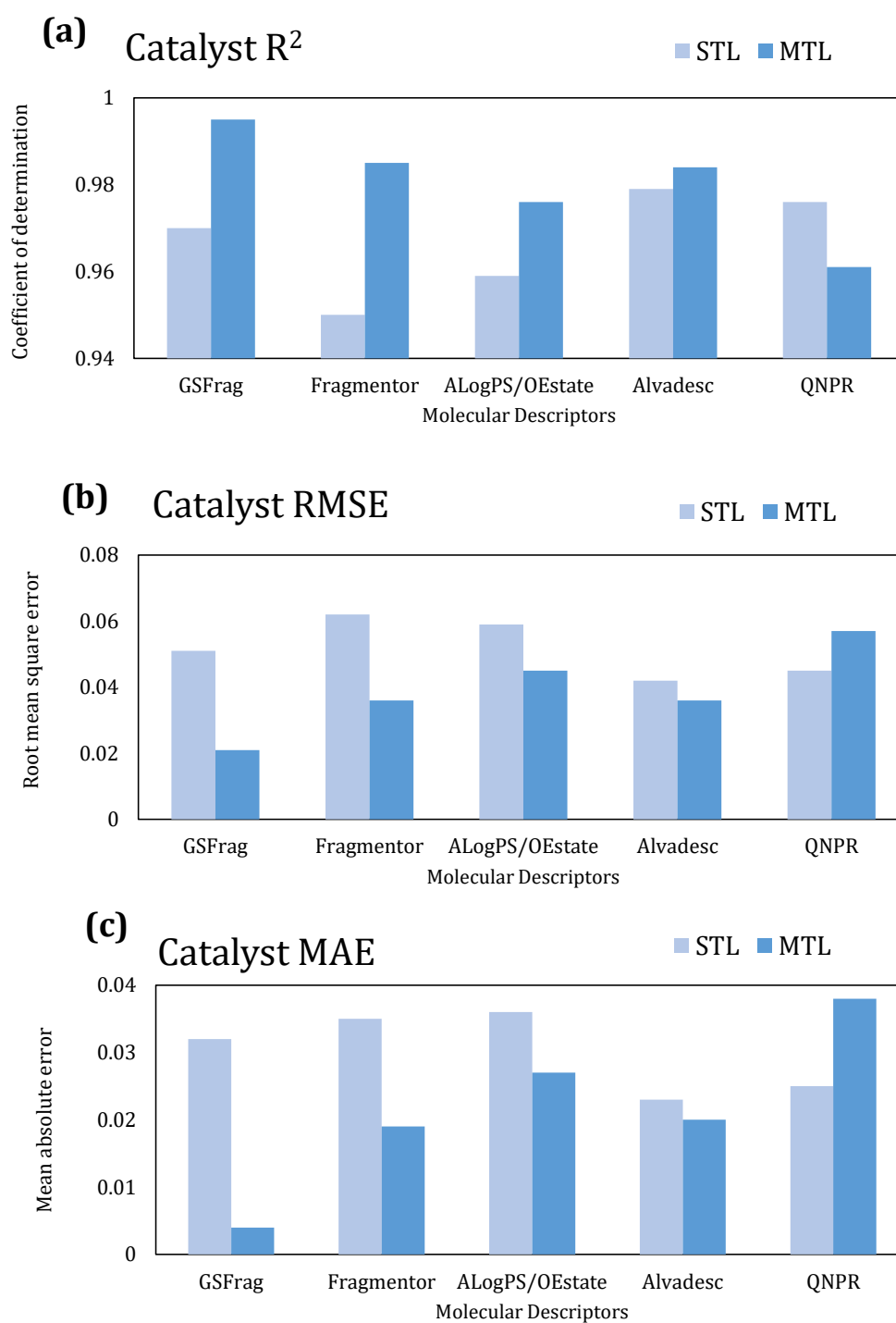


Figure 5-8 Comparison of the effect of the inductive transfer approach and molecular descriptors, on the associative neural network (ASNN) algorithm, on the catalyst methane conversion database, (a) coefficient of determination (R^2), (b) root mean square error (RMSE) and (c) mean absolute error (MAE).

Figure 5-7 gives the accuracy metrics for sorbent database and Figure 5-8, the results for the catalyst database, using the ASNN model. The results show that MTL gave the highest R^2 , and the lowest MAE and RMSE. Reasons why the MTL improved accuracy predictions over STL, is mainly down to having additional datapoints, with 422 for the MTL compared to 239 and 183 for the sorbent and catalyst STL models, respectively. Extra data isn't always a good thing and instead the quality of data has more bearing, due to the effect of overfitting. If the model is tuned with too much data, it starts to memorise data instead of 'learning', which leads to high errors for unseen data [65]. However, it has also been shown that more data can lead to lower estimation variance and therefore better predictive performance. More data increases the probability that it contains useful information, which is advantageous [66]. With these opposing arguments in mind, for this particular dataset, it is suggested that the additional datapoints improved the generalization of the prediction, by using the domain-specific information contained in the training signals of the added related property, which allowed for these training signals to act as an inductive bias [56].

As mentioned before the molecular descriptors gave interesting results between the two databases, with AlogPS/OEstate and Fragmentor descriptors performing well with the sorbent database and alvaDesc and QNPR giving good results for the catalyst database. As seen in Table 5-2, AlogPS/OEstate calculates descriptors based on each atom's intrinsic electronic properties, as well as looking at the counts of atom or bond types and Fragmentor set of descriptors focus on molecular fragments which contain between 2 - 4 atoms. The QNPR set of descriptors uses substrings of SMILES as a representation of molecules [50], and alvaDesc calculates descriptors such as constitutional indices, charge descriptors and atom pairs for example, with the majority of descriptors being 2D based. When looking at the two databases, a main difference is firstly the number of experimental conditions as inputs, however this is unlikely to affect the way the descriptors behaved with the databases, and another difference is the SMILES format. As a percentage, more of the sorbent database had a SMILES format that incorporated bonds and functional groups (31% compared to catalyst 22%),

whereas the catalyst database had more data that was a 'mixture'. This could explain why it yielded the good results for the catalyst dataset. Typical molecules used in QSPR are standard compounds and therefore, the SMILES input often involve structures such as aromatic rings, functional groups, and the bond type for example. As most of the catalyst molecules included in this study aren't whole compounds but rather a physical mixture, the SMILES input format in this case was simply the element in the molecule along with its charge, as a string. Therefore, considering that QNPR relies on the substrings to calculate descriptors as opposed to focusing on bonds or functional group, it suggested that it could readily calculate more accurate and relevant descriptors to the databases used in this study. Additionally, alvaDesc looks at aspects such as charge descriptors, and 2D matrix-based descriptors for example which would favour the catalyst database based on the SMILES input. Conversely, looking at the sorbent database, reasoning behind why AlogPS/OEstate and Fragmentor would be favourable here, is that it specifically looks at the bond types, as well as the molecular fragments and neighbouring atoms, which would allow for better prediction if the SMILES input data consisted of this information.

In terms of STL vs MTL, as seen in Figure 5-7 and Figure 5-8, the descriptors that favour the STL approach for the sorbent database, i.e., the set that yields higher R^2 , and lower RMSE/MAE in order is AlogPS/OEstate>Fragmentor>GSFrag>alvaDesc>QNPR. For the STL catalyst database the performance is alvaDesc>QNPR>GSFrag>AlogPS/OEstate>Fragmentor, which is in line with the aforementioned reasoning. Interestingly, the set of molecular descriptors that sits in the middle for both databases in terms of predictive ability, is GSFrag. For this reason, it is the best performing set of descriptors for both databases when used with the MTL approach, as it provides equally strong performance for each database, thus resulting in an overall high performing prediction capability when a combined sorbent and catalyst database is the input data.

The GSFrag descriptors are the occurrence numbers of certain special fragments containing 2–10 non-hydrogen atoms, and it's been proven that the occurrence

of specific fragments produces a unique code of a chemical structure for wide sets of compounds [67]. As most QSPR descriptors sets are predominately developed on the field of drug discovery, it is unlikely that other descriptors such as Fragmentor which is based on the pharma industry, to have fragments common to sorbents and nickel catalysts. QSPR equations constructed from these descriptors usually provide good statistical characteristics and high predictive ability. Molecular fragments of this type provide good correlations between properties and chemical structure for many classes of compounds [68].

Consequently, as the results indicated that the GSfrag was favourable, along with the MTL ASNN approach, these settings were taken forward to predict new unseen CSCM molecules.

When predicting new unseen molecules using this ML model set up, although the accuracy metrics were almost perfect, the prediction capability was less than optimal, giving a range of accuracy between 11-300% between the actual vs predicted values for both the methane conversion and last cycle capacity. This can be down to a number of reasons, namely overfitting and having the experimental conditions skew the effects of the molecular descriptors [65]. In this case what may have occurred can be due to either the limits of the data, either from being limited in size or having too much 'noise' (which can be caused by various factors, including incorrect collection, from either human error or instruments, or both), or it may be down to the constraints of algorithm itself, which may too complicated or too simplified to this type of data [69]. It can also be caused when additional independent variables do not add any significant information to the dependent variable.

A way to overcome this issue is the use of an adjusted coefficient of determination, (R^2_{adj} , Equation (5-4)), instead of R^2 as an accuracy metric, where R^2 is the coefficient of determination, n is the number of observations i.e. datapoints, and k is the number of independent variables, i.e. experimental conditions. In doing so, adjusted coefficient takes into consideration the number of independent variables in the model in order to avoid overestimating, the impact

of adding independent variables to the model. i.e. reduces the chances of overfitting the model.

$$R_{adj}^2 = 1 - \left(\frac{(n-1)(1-R^2)}{n-k-1} \right) \quad (5-4)$$

Adjusted R^2 is always less than or equal to R^2 . R_{adj}^2 increases when a new datapoint improves the model more than would be expected by chance and it decreases when a predictor improves the model by less than expected [69].

As demonstrated in Table 5-3, the difference in R^2 and R_{adj}^2 is almost negligible, therefore the likelihood of overfitting due to the quantity of datapoints can be ignored, and be attributed to other factors, such as the aforementioned molecular descriptors or incorrect collection, from either human error or instruments.

Table 5-3 A comparison of R^2 and R_{adj}^2 for the ASNN STL and MTL models using GSfrag descriptors

GSfrag ASNN	MTL R^2	STL R^2	n	k	MTL R_{adj}^2	STL R_{adj}^2
Sorbent Data	0.99995	0.9700	183	6	0.999948	0.9690
Catalyst Data	0.99500	0.9600	239	9	0.994803	0.9584

So, whilst it's difficult to confirm, the molecular descriptors may have a limitation on the prediction abilities as they are seldom used in sorbent and catalyst materials, therefore, the ability to 'learn' is only as good as the past research they had been used in.

Because of these various reasons, in order to reduce the cumulative discrepancies on the end results the machine learning model development stage was repeated with some omitted steps. Firstly, the only model taken forward was the ASNN as that outperformed the other two clearly, additionally GSfrag was the molecular descriptor set taken forward for the reason of resulting in equally good results for both databases. What was changed however, was the dataset inputs, the validation and the model architecture. After trial and error and using

data from the PCA information, a new model that resulted in more accurate predictions, was developed, with the input data and ASNN structure given in Table 5-4.

Table 5-4 Properties of the final ML model used for prediction of unseen combined sorbent catalyst materials (CSCM)

Database	Property (units)	Experimental conditions (units)
Sorbent	Last cycle capacity (gCO ₂ /gSorbent)	CaO concentration (%), cycle number, calcination and carbonation temperatures and times (°C/ mins), synthesis method, CaO precursor, initial cycle capacity (gCO ₂ /gSorbent)
Catalyst	Methane conversion (%)	Nickel concentration (wt%), calcination temperature (C), calcination duration (h), SMR reaction temperature (°C) fresh BET surface area (m ² /g), steam/carbon ratio
Validation		5 cross-validation
ASNN architecture		
Training method		SuperSAB
Number of neurons in hidden layer		32
Learning iterations		1000
Ensemble		256
Additional Parameters		PARTITION=3, SELECTION=2

5.3.2 Applicability Domain

The Applicability Domain (AD) is a theoretical region in ‘physicochemical space’ on which the training set of the model has been developed, and for which for which a QSPR model should makes predictions with a given reliability [70,71]. Due to this, its recommended that new data is predicted within the AD by interpolation as opposed to extrapolation. In this study the AD was well defined with validation data falling within the AD, however the prediction results for the unseen CSCM molecules did not fall within the AD, therefore, they were calculated by extrapolation. To visualise the spread of data, a distance to model plot (Williams plot) was generated, which assesses how ‘far’ a molecule is from

the model. Compounds that are 'further from the model', are by definition expected to have lower prediction accuracy than compounds that are 'closer' [72].

Figure 5-9 gives the Williams plot for the graphical visualization of outliers for the predicted results in the developed model. The red dots indicate the prediction capability, where its X-position is the 'distance to model' given by ASNN-Standard Deviation (STDEV) and its Y-position is the expected RMSE. From this, it can be suggested that OECD guideline 3 is adhered to as the predicted data for the model development largely falls within the AD.

Conversely, when applying this model to new unseen data, the results are not as fruitful. The green datapoints in Figure 5-9, which are the new predicted compounds, all indicate an RMSE value of, or around, 0.086 and 0.4 for the sorbent and catalyst properties, respectively. This is expected as when looking at the original metrics, the RMSE values for the developed models were 0.09 ± 0.03 for the sorbent training data and 0.23 ± 0.08 for catalyst training data. Therefore, the expected RMSE for the predicted CSCM molecules is in line with the developed model training data, with the catalyst RMSE being significantly higher. Reasons for this include the fact that the catalyst database has less data in quantity, as well as using fewer parameters for prediction. In terms of distance to model, they are on the high end, meaning the predictions aren't as reliable as they could be. Reasons for why this occurred could be due to the nature of the input data vs the CSCM predictions.

However, distance to model plots estimate the reliability of predictions, and whilst accuracy is an objective measure that has a fixed calculation procedure, reliability is subjective and can be estimated in numerous ways [73]. Because of this, data that falls out of the AD isn't necessarily invalid, however is less reliable. Due to this, further analysis, whether practical or theoretical i.e. more computational modelling, would be recommended to validate or confirm the findings of developed models via this QSPR approach, prior to commercial synthesis.

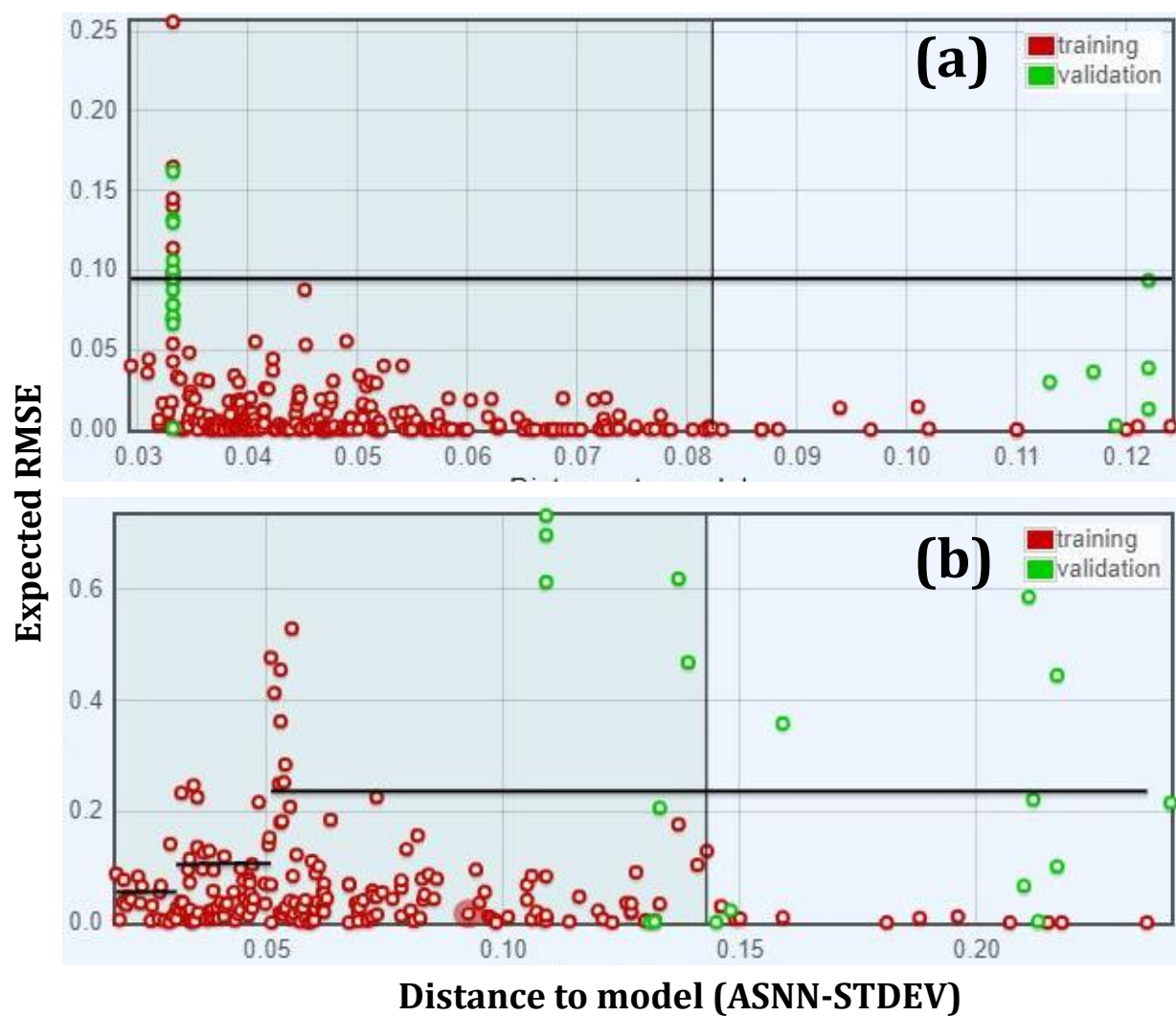


Figure 5-9 Williams plots for the model when applied to predict unseen CSCM molecules, with ASNN-Standard Deviation used as the distance to the model. Black lines represent the averaged RMSE over different distance to model intervals (a) Last cycle CO₂ capacity (g CO₂/g Sorbent), and (b) methane Conversion (%).

5.3.3 Model Prediction Capability

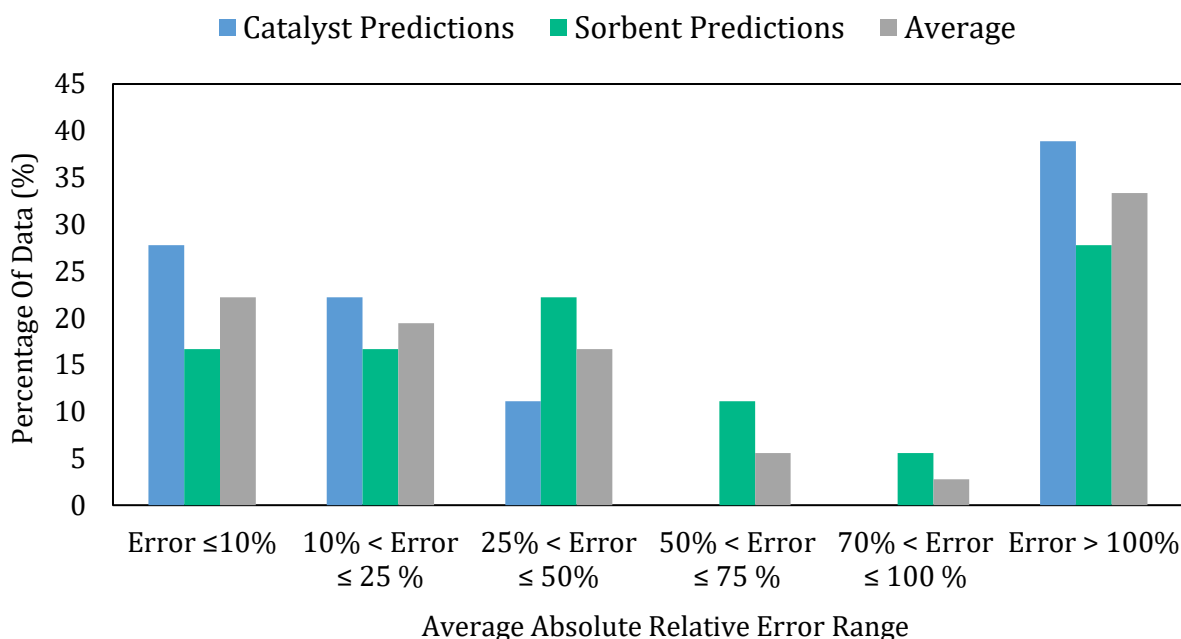


Figure 5-10 Comparison of the percentage of data belonging to each average absolute relative error (AARE) range for the last cycle capacity and methane conversion prediction of CSCM molecules.

The GSfrag/ASNN model was used to predict the last cycle capacity of CSCM molecules, again where the data was collected through literature data mining, difference being that the new molecules possessed both catalytic and sorbent properties, compared to the molecules used to create the predictive model, which had one or the other. As the study of CSCMs isn't as exhaustive as nickel-based catalysts and calcium-based sorbents, there was less data available to apply the model to, with only 41 datapoints. Furthermore, with some of the data collected, some literature didn't provide both last cycle capacity and methane conversion despite reporting on a CSCM. The database along with the predicted properties can be found in Appendix B, Table_Apx B-9 and B-10. Lastly, through trial and error a pattern was observed in the type of input data fed into the machine learning model. A combined sorbent and catalyst database made up of just the 239 and 183 datapoints (422 datapoints) yielded a model that could not predict the output behaviour at all, i.e., no clear regression pattern was seen in the Actual

vs predicted graphs. Consequently, splitting the CSCM database of 41 datapoints into two and adding some to the overall MTL database (resulting in 446 datapoints overall) and using some as the validation unseen data (17 datapoints) was successful and the model was able to 'learn' and predict some of the CSCM behaviours.

Figure 5-10 shows the percentage of molecules' predications that fell in specific AARE ranges (%), for the last cycle capacity prediction, the methane conversion prediction, and an average of the two for the unseen CSCM molecules. From this figure, it's clear that although good amount of unseen CSCM were predicted well (around 30-35% of data falling in an error range of 10% or less), a significant amount also fell in the error range above 100%.

The molecules with the highest ARE (%), were those at the extremities, in this case the molecules with very low sorbent capacities, the molecules missing highly influential conditions (e.g., CaO concentration %), and molecules with conditions that were not commonly occurring in the databases that developed the model. In these cases, the information provided was not sufficient for the model to extrapolate well, based on existing data.

The average prediction capability of the overall model isn't tremendously high, with only 58% of the data resulting in an AARE of 50% or less, however given this new approach, and the fact that high ARE (%) values have an explanation behind them, the models' application on new data is a fairly positive one.

Looking at the prediction capability of the individual properties, the catalyst properties are predicted ever so slightly better at 61% of data falling in and AARE range of 50% or less, compared to 58% for the sorbent properties. Reasons for this can be attributed to the fact the catalyst database uses less experimental conditions and relies on the molecule more to predict the behaviour. This can be a good and bad thing; it allows for less literature data to be gathered in order to predict an output from ML and QSPR. On the other hand, it prevents the ability to observe correlations in the input data, which can help further the development of a CSCM, as described in the next section.

5.3.4 Patterns in Data

Further to obtaining a measure of how well the model could predict new molecules, additional information was also garnered from the results. Patterns in the output prediction data were seen as a function of input data, specifically the two main influencers CaO concentration (%) and Nickel %wt, Figure 5-11.

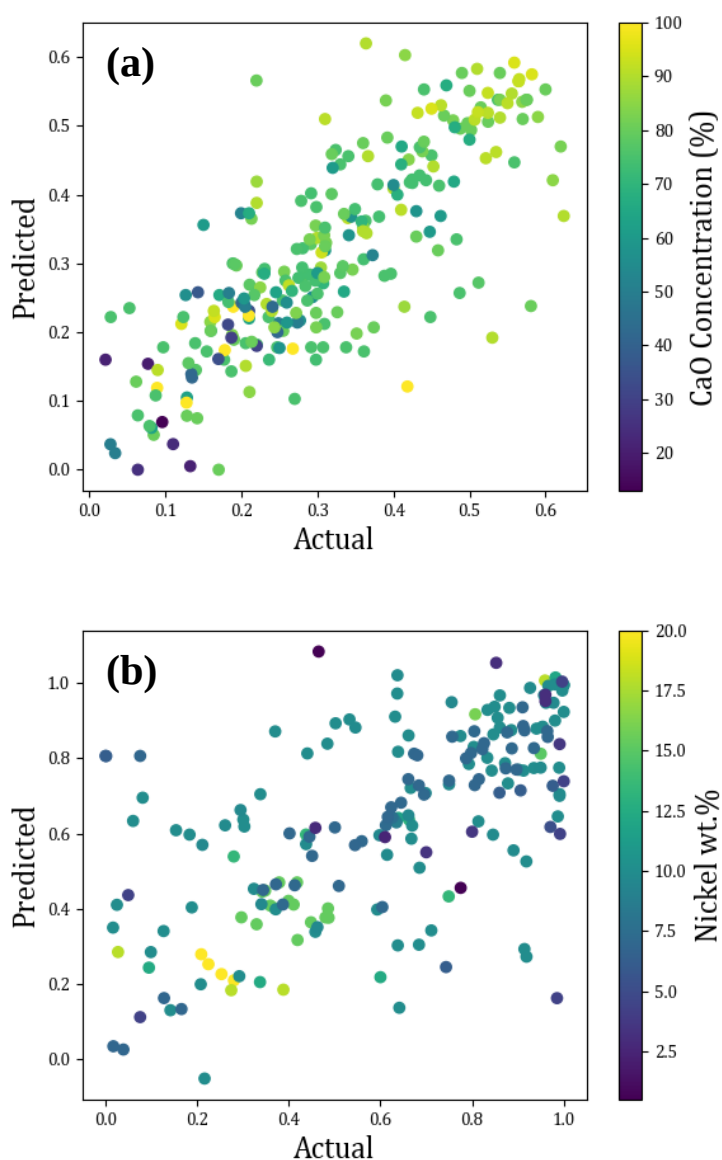


Figure 5-11 Actual vs Predicted scatter plots showing the effect of an influential input for (a) Last cycle capacity (gCO₂/gSorbent) as a function of CaO concentration (%), and (b) methane conversion (%) as a function of nickel wt%.

Figure 5-11(a and b) show certain patterns in the input data; a low concentration of calcium oxide in a sorbent material yields lower CO₂ capacities in the last cycle, as expected, but there is also an optimal concentration as too much CaO % also does not yield the maximum CO₂ capacity in the last cycle. Similarly, for the wt% of Nickel in a catalyst is shown to yield low conversion if too high. In both cases, this is likely to be due to dispersion and availability of active sites.

Additional parameters have an effect on where a datapoint sits on an actual vs prediction plot, therefore colour plots such as these, prove helpful to narrow down the best direction to go in CSCM development. Further information garnered from colour plots generated from this study include the calcium precursor and the inert support in the sorbent. As you can see from Figure 5-12, there is an indication that the naturally occurring precursors of calcium carbonate and calcium oxide, yield a lower last cycle capacity. Conversely the best performing precursors are shown to be calcium gluconate and calcium nitrate. Other precursors that gave good performance but equally underperformed were calcium acetate and calcium chloride. Reasons behind this, is that it is well known that calcium oxide from natural precursors undergo a rapid loss of reactivity after several carbonation/calcination cycles. Additionally, it has been reported that natural sorbents, such as limestones, form product layers with higher cohesion than those formed by the synthetic sorbent. Thus, once the small pores have largely filled and a thin layer of product has been deposited the natural sorbents CO₂ uptake rapidly decreases [74].

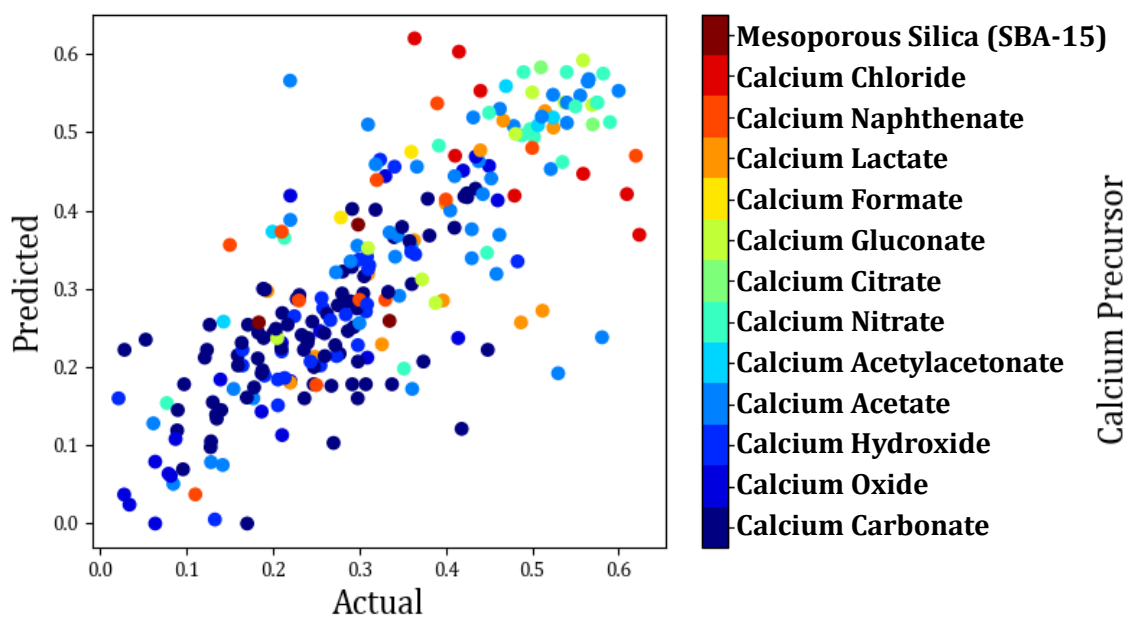


Figure 5-12 Actual vs Predicted scatter plot showing the effect of the calcium precursor on the Last cycle capacity ($\text{gCO}_2/\text{gSorbent}$)

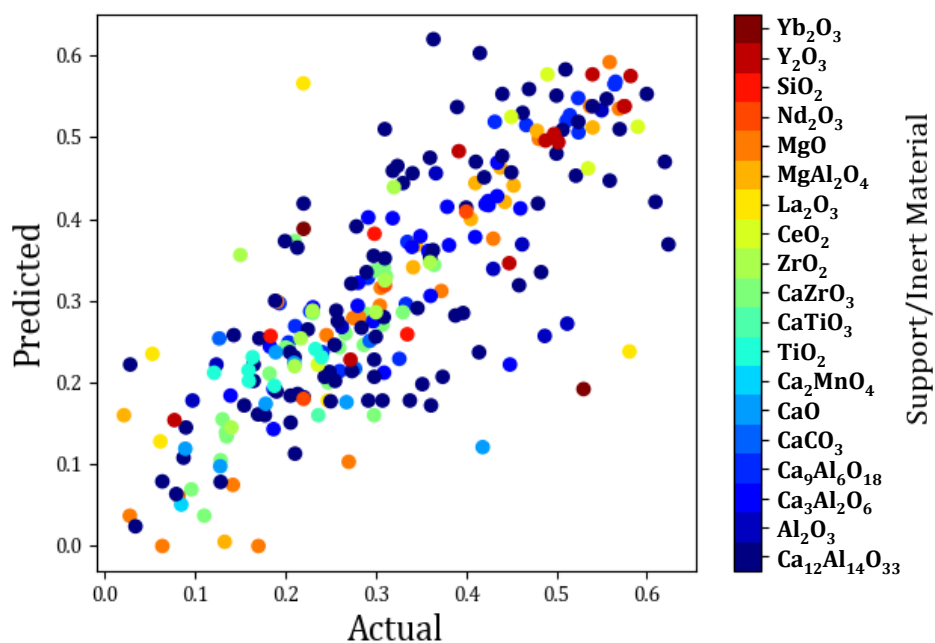


Figure 5-13 Actual vs Predicted scatter plot showing the effect of the inert support on the Last cycle capacity ($\text{gCO}_2/\text{gSorbent}$)

From Figure 5-13, it can be seen that Mayenite ($\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$), provides both good and under par performance. From this colour plot, very little information can be gathered, in terms of the type of support is optimal for the sorbent development. This can be down to the fact that there are other more influential factors, and that the inert support, as inferred from the name, has little effect on the overall performance. A colour plot that did provide a definitive output is the synthesis method, Figure 5-14 and Table 5-5. From this colour plot it can clearly be seen that the sol-gel method outperforms the other methods consistently. A method that consistently underperformed was the surfactant template method/ultrasound assisted technique. Reasons for this include that the former is a well-researched method and has had time to be perfected in time whereas the latter is a fairly new and novel method.

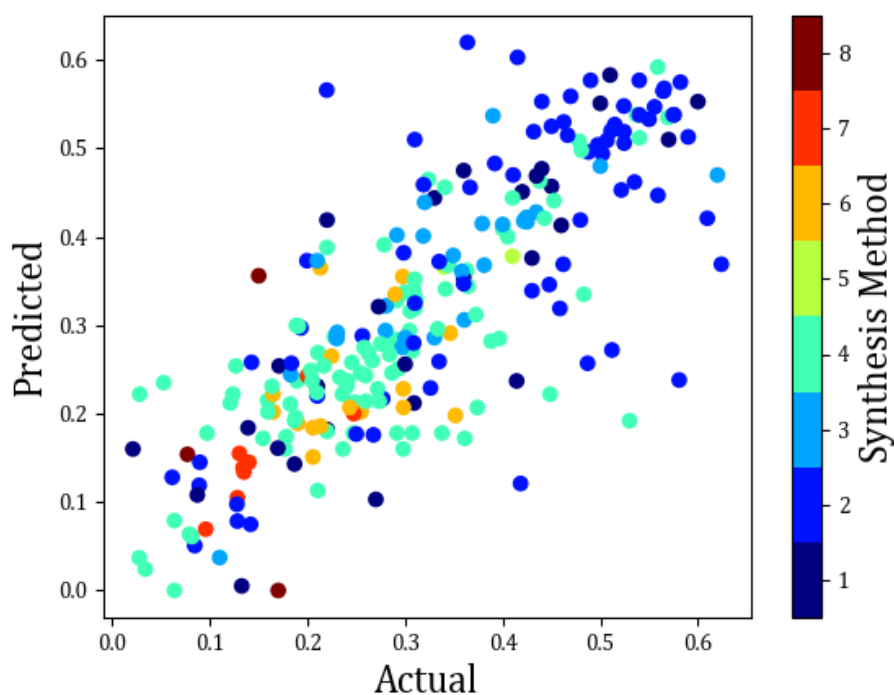


Figure 5-14 Actual vs Predicted scatter plot showing the effect of the synthesis method on the Last cycle capacity ($\text{gCO}_2/\text{gSorbent}$)

Table 5-5 Key for the synthesis method

	Synthesis method
Method-1	Mixing
Method-2	Sol-gel/sol gel combustion/ sol mixing/ gel template
Method-3	Flame spray pyrolysis (FSP)/Flame synthesis/Combustion synthesis
Method-4	Wet mixing/Mixing + palletisation
Method-5	Four-step heating method
Method-6	Coprecipitation
Method-7	Surfactant template/ultrasound-assisted technique
Method-8	Wet impregnation

Additional information gathered from other colour plots for the sorbent properties (see Appendix B, Figure_Apx B-1 to B-3.) was data such as the optimal cycle number, calcination/carbonation time and calcination/carbonation temperature.

Data gathered from the catalyst colour plots equally gave some conclusive and some ambiguous results. Figure 5-15 represents the optimal steam to carbon ratio during the reforming step. Clearly its shown that the higher the S/C the higher the methane conversion (%). Likewise, from Figure 5-16, it can be seen that an average calcination temperature for the catalyst, at around 600°C, results in a high methane conversion (%). On the other hand, a colour plot that seemed inconclusive is Figure 5-17, the catalyst calcination time. Even with changing the y-axis to a smaller time range, the data didn't provide a definitive time, as shown in both colour plots. What can be inferred from this plot is that a relatively low time e.g., 10 hours, is sufficient.

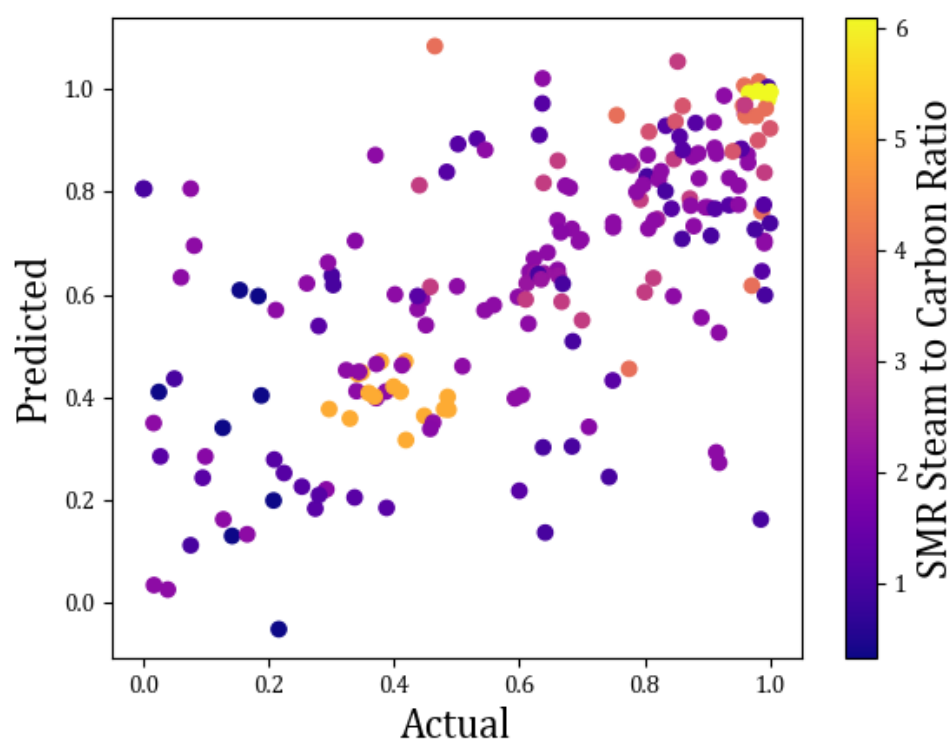


Figure 5-15 Actual vs Predicted scatter plots showing the effect of steam to carbon ratio on methane conversion (%)

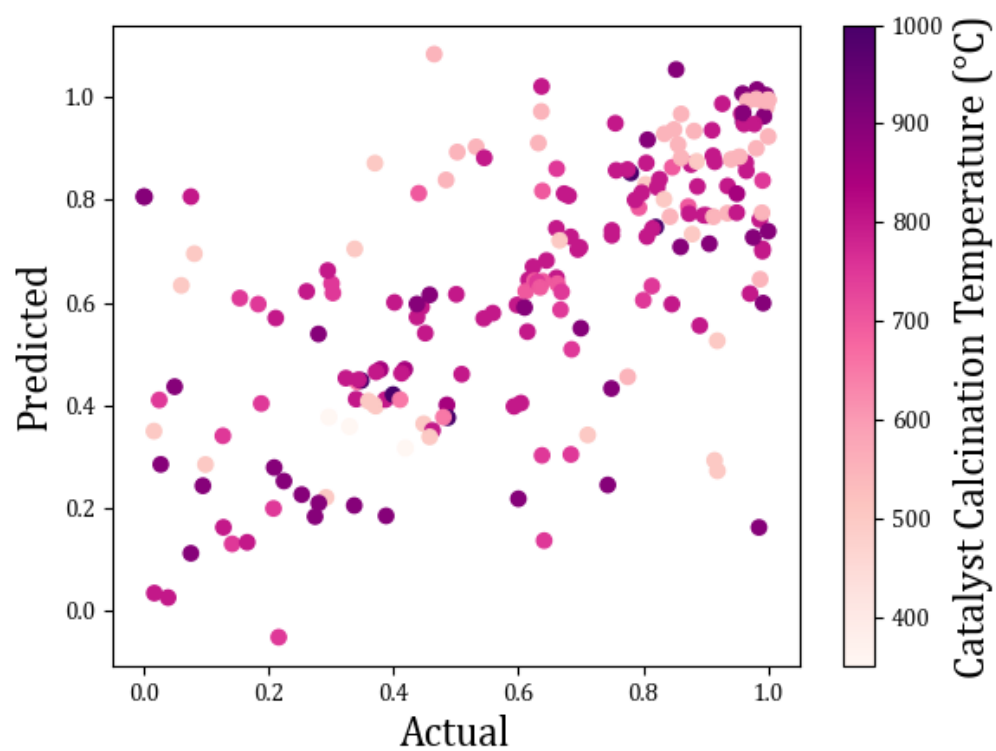


Figure 5-16 Actual vs Predicted scatter plots showing the effect of catalyst calcination temperature (°C) on methane conversion (%)

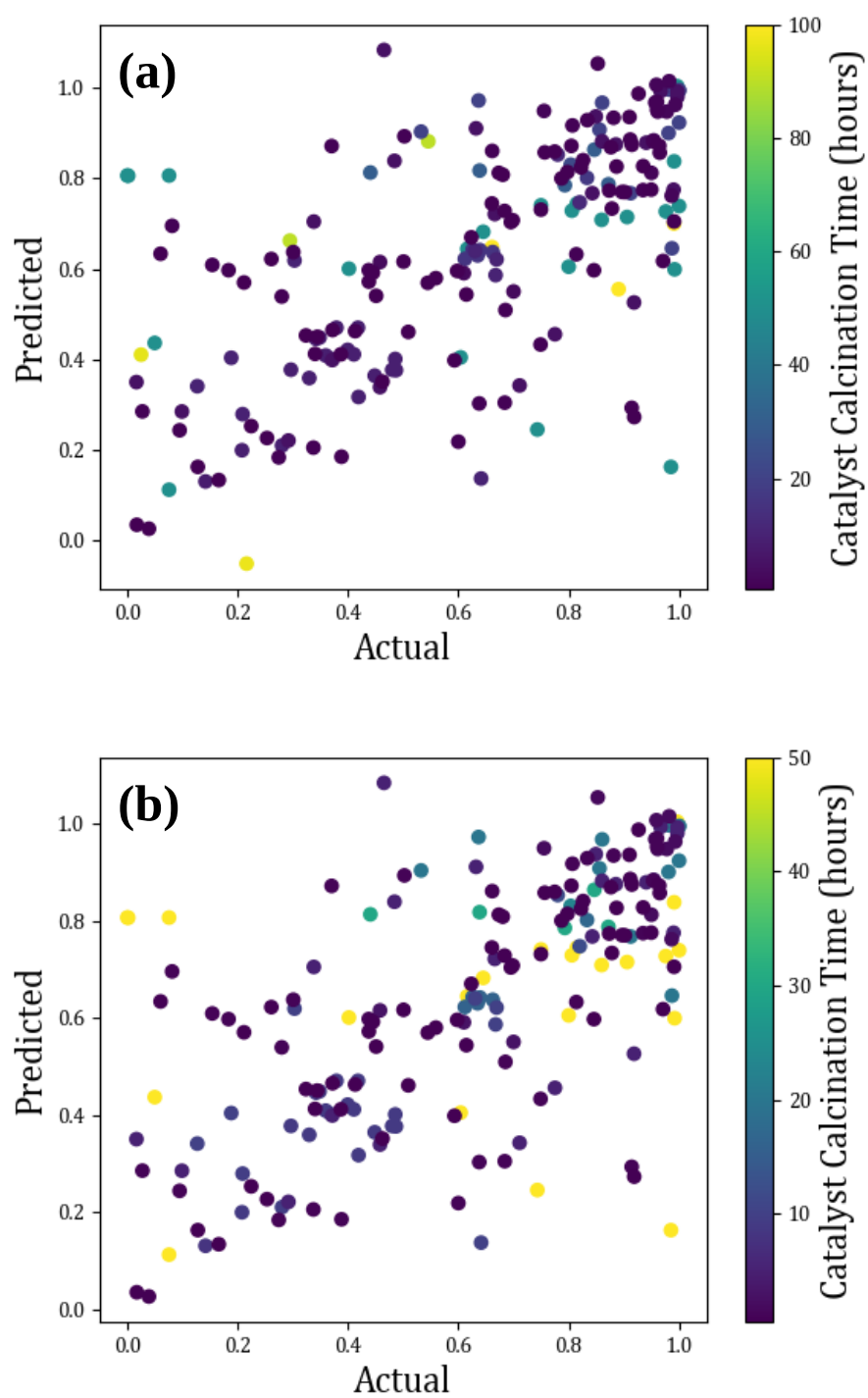


Figure 5-17 Actual vs Predicted scatter plots showing the effect of catalyst calcination time (a) maximum 100 hours, (b) 50 hours on methane conversion (%)

Additional information gathered from other colour plots include sorbent properties such as cycle number, carbonation/calcination time and temperature, and catalyst properties such as BET surface area and the SMR reformer temperature (see Appendix B, Figure_Apx B-1 to B-4.) When all the results from the Actual vs Predicted plots are collated, the results of the route that that should be explored to develop a CSCM that has good performance with respect to both the last cycle capacity and the methane conversion, is shown in Table 5-6.

Table 5-6 Data gathered from the colour plots, indicating parameters that lead to a well/bad performing CSCM.

Parameter	Good performance	Bad performance
Calcium precursor	Calcium nitrate, Calcium gluconate	Calcium carbonate/oxide
Inert support	Y ₂ O ₃ , CeO ₂	Ca ₁₂ Al ₁₄ O ₃₃ , CaZrO ₃
Synthesis method	Sol gel	Surfactant template
Optimal synthesis details:		
CaO concentration (%)	70-90%	
Cycles	10-15	
SMR reformer temperature	650°C	
SMR S/C ratio	6	
Carbonation time/ temperature	15-20mins / 650-700°C	
Calcination time/ temperature	5-10mins / 900°C	
Nickel wt%	2.5-10%	
Catalyst calcination temperature	600-700°C	
Catalyst calcination time	10 hours	

5.4 Conclusion

In this work, a database for the prediction of Combined Sorbent Catalyst Material (CSCM) properties, was developed through the application of data mining, quantitative structure–property relationship analysis (QSPR) and multi-task learning (MTL).

The development of the databases was discussed as well as the parameters chosen to produce accurate predictions for the properties of interest- last cycle capacity (gSorbent/ gCO₂) and methane conversion (%). A thorough comparison of machine learning models, molecular descriptors and their respective settings was also detailed. From the results, the application of MTL was shown to improve the prediction capability of new molecules, compared to single-output models. Consequently, the MTL model was used to predict properties of unseen CSCM molecules somewhat effectively, 58% of the data resulting in an AARE of 50% or less.

Overall, this study has presented a methodology for the prediction of CSCM that hasn't been conducted before, with promising results. The study concludes that a combination of machine learning models and application of molecular descriptors has the capability to predict properties for a CSCM for the process of SE-SMR. The results from this aim to streamline and accelerate the experimental discovery of an optimal CSCM, by reducing the repetitive trial and error processes involved in material development. This study also possibly paves a route into expanding the QSPR approach for applications beyond drug discovery or biochemistry, which is the main use of molecular descriptors and QSPR.

Future directions for this study would be to use information from *Table 5-6* to synthesise a new material using the model as is, then make changes to the models as appropriate, from the findings of this study, namely the reduction of experimental conditions. This would allow for, firstly, the synthesis of a material that would be assumed to possess the desired properties for a CSCM, but with the removal of the conditions after this data is obtained, the prediction of how it would perform is assumed to be more accurate with fewer conditions specified. The material could then be synthesised, and performance compared against the models' predictions, and thus further validating this QSPR and MTL application for the process of SE-SMR. This could also pave a route for the application of multitask learning to be more utilised to predict behaviour of materials that rely on more than one type of property feature.

5.5 References

- [1] Balat M. Potential importance of hydrogen as a future solution to environmental and transportation problems. *International Journal of Hydrogen Energy*. 2008; 33(15): 4013–4029. Available at: DOI:10.1016/j.ijhydene.2008.05.047
- [2] Kapdan IK., Kargi F. Bio-hydrogen production from waste materials. *Enzyme and Microbial Technology*. 2006; 38(5): 569–582. Available at: DOI:10.1016/j.enzmictec.2005.09.015
- [3] Babcock D (GTI). Phase 1 SBRI Hydrogen Supply Competition Bulk Hydrogen Production by Sorbent Enhanced Steam Reforming (HyPER) Project the Gas Technology Institute. 2019; 20. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/866373/Phase_1_-_Cranfield_-_HyPER.pdf
- [4] Hufton JR., Mayorga S., Sircar S. Sorption-enhanced reaction process for hydrogen production. *AIChE Journal*. February 1999; 45(2): 248–256. Available at: DOI:10.1002/aic.690450205
- [5] García-Lario AL., Aznar M., Martinez I., Grasa GS., Murillo R. Experimental study of the application of a NiO/NiAl₂O₄ catalyst and a CaO-based synthetic sorbent on the Sorption Enhanced Reforming process. *International Journal of Hydrogen Energy*. 2015; 40(1): 219–232. Available at: DOI:10.1016/j.ijhydene.2014.10.033
- [6] García-Lario AL., Grasa GS., Murillo R. Performance of a combined CaO-based sorbent and catalyst on H₂ production, via sorption enhanced methane steam reforming. *Chemical Engineering Journal*. 2015; 264: 697–705. Available at: DOI:10.1016/j.cej.2014.11.116
- [7] Dou B., Wang C., Song Y., Chen H., Jiang B., Yang M., et al. Solid sorbents for in-situ CO₂ removal during sorption-enhanced steam reforming process: A review. *Renewable and Sustainable Energy Reviews*. Elsevier; 2016;

53: 536–546. Available at: DOI:10.1016/j.rser.2015.08.068

[8] Manovic V., Anthony EJ. Thermal activation of CaO-based sorbent and self-reactivation during CO₂ capture looping cycles. *Environmental Science & Technology*. 2008; 42(11): 4170–4174. Available at: DOI:10.1021/es800152s

[9] Broda M., Kierzkowska AM., Müller CR. Influence of the calcination and carbonation conditions on the CO₂ uptake of synthetic Ca-based CO₂ sorbents. *Environmental Science & Technology*. 2012; 46(19): 10849–10856. Available at: DOI:10.1021/es302757e

[10] Lu H., Reddy EP., Smirniotis PG. Calcium oxide-based sorbents for capture of carbon dioxide at high temperatures. *Industrial and Engineering Chemistry Research*. 2006; 45(11): 3944–3949. Available at: DOI:10.1021/ie051325x

[11] Froment GF. Modelling of catalyst deactivation. *Applied Catalysis A: General*. 2001; 212(1–2): 117–128. Available at: DOI:10.1016/S0926-860X(00)00850-4

[12] García-Lario AL., Aznar M., Grasa GS., García T., Murillo R. Study of nickel catalysts for hydrogen production in sorption enhanced reforming process. *Journal of Power Sources*. Elsevier B.V; 2013; 242: 371–379. Available at: DOI:10.1016/j.jpowsour.2013.05.069

[13] Wei J., Iglesia E. Mechanism and Site Requirements for Activation and Chemical Conversion of Methane on Supported Pt Clusters and Turnover Rate Comparisons among Noble Metals. *Journal of Physical Chemistry B*. 2004; 108(13): 4094–4103. Available at: DOI:10.1021/jp036985z

[14] Xu J., Froment GF. Methane steam reforming: Intrinsic Kinetics. *AIChE J*. 1989; 35(1): 88–96. Available at: DOI:10.1002/aic.690350109

[15] Barelli L., Bidini G., Gallorini F., Servili S. Hydrogen production through sorption-enhanced steam methane reforming and membrane technology: A

review. *Energy*. April 2008; 33(4): 554–570. Available at: DOI:10.1016/j.energy.2007.10.018

[16] Broda M., Manovic V., Imtiaz Q., Kierzkowska AM., Anthony EJ., Müller CR. High-purity hydrogen via the sorption-enhanced steam methane reforming reaction over a synthetic CaO-based sorbent and a Ni catalyst. *Environmental Science & Technology*. 2013; 47(11): 6007–6014. Available at: DOI:10.1021/es305113p

[17] Li S., Gong J. Strategies for improving the performance and stability of Ni-based catalysts for reforming reactions. *Chemical Society Reviews*. Royal Society of Chemistry; 2014; 43(21): 7245–7256. Available at: DOI:10.1039/c4cs00223g

[18] Harrison DP. Sorption-enhanced hydrogen production: A review. *Industrial and Engineering Chemistry Research*. 2008; 47(17): 6486–6501. Available at: DOI:10.1021/ie800298z

[19] Li D., Nakagawa Y., Tomishige K. Methane reforming to synthesis gas over Ni catalysts modified with noble metals. *Applied Catalysis A: General*. Elsevier B.V.; 2011; 408(1–2): 1–24. Available at: DOI:10.1016/j.apcata.2011.09.018

[20] Clough PT., Boot-Handford ME., Zheng L., Zhang Z., Fennell PS. Hydrogen production by sorption enhanced steam reforming (SESR) of biomass in a fluidised-bed reactor using combined multifunctional particles. *Materials*. MDPI AG; 21 May 2018; 11(5). Available at: DOI:10.3390/ma11050859

[21] di Giuliano A., Gallucci K., Foscolo PU. Determination of Kinetic and Diffusion Parameters Needed to Predict the Behavior of CaO-Based CO₂ Sorbent and Sorbent-Catalyst Materials. *Industrial and Engineering Chemistry Research*. 2020; 59(15): 6840–6854. Available at: DOI:10.1021/acs.iecr.9b05383

[22] Aloisi I., Jand N., Stendardo S., Foscolo PU. Hydrogen by sorption

enhanced methane reforming: A grain model to study the behavior of bi-functional sorbent-catalyst particles. *Chemical Engineering Science*. Elsevier; 2016; 149: 22–34. Available at: DOI:10.1016/j.ces.2016.03.042

[23] di Giuliano A., Gallucci K., Giancaterino F., Courson C., Foscolo PU. Multicycle sorption enhanced steam methane reforming with different sorbent regeneration conditions: Experimental and modelling study. *Chemical Engineering Journal*. Elsevier; 2019; 377(September 2018): 119874. Available at: DOI:10.1016/j.cej.2018.09.035

[24] Rout KR., Jakobsen HA. A numerical study of pellets having both catalytic and capture properties for SE-SMR process: Kinetic- and product layer diffusion-controlled regimes. *Fuel Processing Technology*. Elsevier B.V.; 2013; 106: 231–246. Available at: DOI:10.1016/j.fuproc.2012.07.029

[25] di Carlo A., Aloisi I., Jand N., Stendardo S., Foscolo PU. Sorption enhanced steam methane reforming on catalyst-sorbent bifunctional particles: A CFD fluidized bed reactor model. *Chemical Engineering Science*. Elsevier Ltd; 2017; 173: 428–442. Available at: DOI:10.1016/j.ces.2017.08.014

[26] Martínez I., Grasa G., Meyer J., di Felice L., Kazi S., Sanz C., et al. Performance and operating limits of a sorbent-catalyst system for sorption-enhanced reforming (SER) in a fluidized bed reactor. *Chemical Engineering Science*. 2019; 205: 94–105. Available at: DOI:10.1016/j.ces.2019.04.029

[27] Albrecht KO., Satrio JA., Shanks BH., Wheelock TD. Application of a combined catalyst and sorbent for steam reforming of methane. *Industrial and Engineering Chemistry Research*. 2010; 49(9): 4091–4098. Available at: DOI:10.1021/ie901914c

[28] García-Lario AL., Grasa GS., Murillo R. Performance of a combined CaO-based sorbent and catalyst on H₂ production, via sorption enhanced methane steam reforming. *Chemical Engineering Journal*. 2015; 264: 697–705. Available at: DOI:10.1016/j.cej.2014.11.116

- [29] Chanburanasiri N., Ribeiro AM., Rodrigues AE., Arpornwichanop A., Laosiripojana N., Praserthdam P., et al. Hydrogen production via sorption enhanced steam methane reforming process using Ni/CaO multifunctional catalyst. *Industrial and Engineering Chemistry Research*. 2011; 50(24): 13662–13671. Available at: DOI:10.1021/ie201226j
- [30] Dang C., Liu L., Yang G., Cai W., Long J., Yu H. Mg-promoted Ni-CaO microsphere as bi-functional catalyst for hydrogen production from sorption-enhanced steam reforming of glycerol. *Chemical Engineering Journal*. Elsevier; 2020; 383(July 2019): 123204. Available at: DOI:10.1016/j.cej.2019.123204
- [31] Wang C., Dou B., Jiang B., Song Y., Du B., Zhang C., et al. Sorption-enhanced steam reforming of glycerol on Ni-based multifunctional catalysts. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2015; 40(22): 7037–7044. Available at: DOI:10.1016/j.ijhydene.2015.04.023
- [32] Wang X., He Y., Xu T., Xiao B., Liu S., Hu Z., et al. CO₂ sorption-enhanced steam reforming of phenol using Ni–M/CaO–Ca₁₂Al₁₄O₃₃ (M = Cu, Co, and Ce) as catalytic sorbents. *Chemical Engineering Journal*. Elsevier; 2020; 393(November 2019): 124769. Available at: DOI:10.1016/j.cej.2020.124769
- [33] Satrio JA., Shanks BH., Wheelock TD. A combined catalyst and sorbent for enhancing hydrogen production from coal or biomass. *Energy and Fuels*. 2007; 21(1): 322–326. Available at: DOI:10.1021/ef060132z
- [34] Rahmanzadeh L., Taghizadeh M. Sorption-enhanced ethanol steam reforming on Ce-Ni/MCM-41 with simultaneous CO₂ adsorption over Na- and Zr-promoted CaO based sorbent. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2019; 44(39): 21238–21250. Available at: DOI:10.1016/j.ijhydene.2019.04.289
- [35] Hosseini Abbandanak M., Taghizadeh M., Fallah N. High-purity hydrogen production by sorption-enhanced methanol steam reforming over a combination of Cu–Zn–CeO₂–ZrO₂/MCM-41 catalyst and (Li–Na–K) NO₃·MgO adsorbent.

International Journal of Hydrogen Energy. Elsevier Ltd; 2020; (3): 7099–7112.
Available at: DOI:10.1016/j.ijhydene.2020.11.250

[36] Borhani TN., García-Muñoz S., Vanesa Luciani C., Galindo A., Adjiman CS. Hybrid QSPR models for the prediction of the free energy of solvation of organic solute/solvent pairs. *Physical Chemistry Chemical Physics*. 2019; 21(25): 13706–13720. Available at: DOI:10.1039/c8cp07562j

[37] Borhani TNG., Bagheri M., Manan ZA. Molecular modeling of the ideal gas enthalpy of formation of hydrocarbons. *Fluid Phase Equilibria*. Elsevier B.V.; 2013; 360: 423–434. Available at: DOI:10.1016/j.fluid.2013.09.066

[38] Burner J., Schwiedrzik L., Krykunov M., Luo J., Boyd PG., Woo TK. High-Performing Deep Learning Regression Models for Predicting Low-Pressure CO₂ Adsorption Properties of Metal-Organic Frameworks. *Journal of Physical Chemistry C*. 2020; Available at: DOI:10.1021/acs.jpcc.0c06334

[39] Yang P., Lu G., Yang Q., Liu L., Lai X., Yu D. Analyzing Acetylene Adsorption of Metal-Organic Frameworks Based on Machine Learning. *Green Energy & Environment*. Institute of Process Engineering, Chinese Academy of Sciences; January 2021; : 108947. Available at: DOI:10.1016/j.gee.2021.01.006

[40] Li J., Xu K., Yao X., Chen S. Prediction and optimization of syngas production from steam gasification: Numerical study of operating conditions and biomass composition. *Energy Conversion and Management*. Elsevier Ltd; 15 May 2021; 236. Available at: DOI:10.1016/j.enconman.2021.114077

[41] Wang B., Zhou L., Xu K., Wang Q. Prediction of minimum ignition energy from molecular structure using quantitative structure-property relationship (QSPR) models. *Industrial and Engineering Chemistry Research*. American Chemical Society; 11 January 2017; 56(1): 47–51. Available at: DOI:10.1021/acs.iecr.6b04347

[42] Nantasenamat C., Isarankura-Na-Ayudhya C., Naenna T.,

Prachayasittikul V. A practical overview of quantitative structure-activity relationship. EXCLI Journal. 2009; 8: 74–88.

[43] Borhani TNG., Saniedanesh M., Bagheri M., Lim JS. QSPR prediction of the hydroxyl radical rate constant of water contaminants. Water Research. Elsevier Ltd; 2016; 98: 344–353. Available at: DOI:10.1016/j.watres.2016.04.038

[44] Dearden JC., Cronin MTD., Kaiser KLE. How not to develop a quantitative structure-activity or structure-property relationship (QSAR/QSPR). SAR and QSAR in Environmental Research. 2009; 20(3–4): 241–266. Available at: DOI:10.1080/10629360902949567

[45] Hand DJ., Adams NM. Data Mining. Wiley StatsRef: Statistics Reference Online. Chichester, UK: John Wiley & Sons, Ltd; 2015. pp. 1–7. Available at: DOI:10.1002/9781118445112.stat06466.pub2

[46] Elmolla ES., Chaudhuri M., Eltoukhy MM. The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process. Journal of Hazardous Materials. Elsevier B.V.; July 2010; 179(1–3): 127–134. Available at: DOI:10.1016/j.jhazmat.2010.02.068

[47] Jolliffe IT., Cadima J. Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 13 April 2016; 374(2065): 20150202. Available at: DOI:10.1098/rsta.2015.0202

[48] Ringnér M. What is principal component analysis? Nature Biotechnology. March 2008; 26(3): 303–304. Available at: DOI:10.1038/nbt0308-303

[49] Consonni V., Todeschini R. Molecular descriptors. Challenges and Advances in Computational Chemistry and Physics. 2010. 29–102 p. Available at: DOI:10.1007/978-1-4020-9783-6_3

[50] Sosnin S., Karlov D., Tetko I v., Fedorov M v. Comparative Study of Multitask Toxicity Modeling on a Broad Chemical Space. Journal of Chemical

Information and Modeling. 2019; 59(3): 1062–1072. Available at: DOI:10.1021/acs.jcim.8b00685

[51] Sushko I., Novotarskyi S., Körner R., Pandey AK., Rupp M., Teetz W., et al. Online chemical modeling environment (OCHEM): web platform for data storage, model development and publishing of chemical information. *Journal of Computer-Aided Molecular Design*. 10 June 2011; 25(6): 533–554. Available at: DOI:10.1007/s10822-011-9440-2

[52] Tetko I v., Tanchuk VY. Application of associative neural networks for prediction of lipophilicity in ALOGPS 2.1 program. *Journal of Chemical Information and Computer Sciences*. 2002; 42(5): 1136–1145. Available at: DOI:10.1021/ci025515j

[53] Hall LH., Kier LB. Electrotopological State Indices for Atom Types: A Novel Combination of Electronic, Topological, and Valence State Information. *Journal of Chemical Information and Computer Sciences*. 1995; 35(6): 1039–1045. Available at: DOI:10.1021/ci00028a014

[54] Varnek A., Fourches D., Horvath D., Klimchuk O., Gaudin C., Vayer P., et al. ISIDA - Platform for Virtual Screening Based on Fragment and Pharmacophoric Descriptors. *Current Computer Aided-Drug Design*. 1 September 2008; 4(3): 191–198. Available at: DOI:10.2174/157340908785747465

[55] Mauri A. alvaDesc: A Tool to Calculate and Analyze Molecular Descriptors and Fingerprints. *Ecotoxicological QSARs*. 2020. pp. 801–820. Available at: DOI:10.1007/978-1-0716-0150-1_32

[56] Sadawi N., Olier I., Vanschoren J., van Rijn JN., Besnard J., Bickerton R., et al. Multi-task learning with a natural metric for quantitative structure activity relationship learning. *Journal of Cheminformatics*. Springer International Publishing; 2019; 11(1): 1–13. Available at: DOI:10.1186/s13321-019-0392-1

- [57] Varnek A., Gaudin C., Marcou G., Baskin I., Pandey AK., Tetko I v. Inductive transfer of knowledge: Application of multi-task learning and Feature Net approaches to model tissue-air partition coefficients. *Journal of Chemical Information and Modeling*. 2009; 49(1): 133–144. Available at: DOI:10.1021/ci8002914
- [58] Whitley DC., Ford MG., Livingstone DJ. Unsupervised forward selection: A method for eliminating redundant variables. *Journal of Chemical Information and Computer Sciences*. 2000; 40(5): 1160–1168. Available at: DOI:10.1021/ci000384c
- [59] Dahl GE., Jaitly N., Salakhutdinov R. Multi-task Neural Networks for QSAR Predictions. 2014; : 1–21. Available at: <http://arxiv.org/abs/1406.1231>
- [60] Sosnin S., Vashurina M., Withnall M., Karpov P., Fedorov M., Tetko I v. A Survey of Multi-task Learning Methods in Chemoinformatics. *Molecular Informatics*. 2019; 38(4): 1800108. Available at: DOI:10.1002/minf.201800108
- [61] Browne MW. Cross-Validation Methods. *Journal of Mathematical Psychology*. March 2000; 44(1): 108–132. Available at: DOI:10.1006/jmps.1999.1279
- [62] Bagheri M., Yerramsetty K., Gasem KAM., Neely BJ. Molecular modeling of the standard state heat of formation. *Energy Conversion and Management*. Elsevier Ltd; 2013; 65: 587–596. Available at: DOI:10.1016/j.enconman.2012.07.014
- [63] Yan Y., Mattisson T., Moldenhauer P., Anthony EJ., Clough PT. Applying machine learning algorithms in estimating the performance of heterogeneous, multi-component materials as oxygen carriers for chemical-looping processes. *Chemical Engineering Journal*. Elsevier; May 2020; 387(September 2019): 124072. Available at: DOI:10.1016/j.cej.2020.124072
- [64] Tetko I v. Associative neural network. *Neural Processing Letters*. 2002;

16(2): 187–199. Available at: DOI:10.1023/A:1019903710291

[65] Dietterich T. Overfitting and Undercomputing in Machine Learning. *ACM Computing Surveys (CSUR)*. 1995; 27(3): 326–327. Available at: DOI:10.1145/212094.212114

[66] Martens D., Provost F. Pseudo-social network targeting from consumer transaction data. *NYU Working Paper*. 2011; : 1–31.

[67] Skvortsova MI., Baskin II., Skvortsov LA., Palyulin VA., Zefirov NS., Stankevich I v., et al. Chemical graphs and their basis invariants.

[68] Tetko I v., Gasteiger J., Todeschini R., Mauri A., Livingstone D., Ertl P., et al. Virtual Computational Chemistry Laboratory – Design and Description. *Journal of Computer-Aided Molecular Design*. June 2005; 19(6): 453–463. Available at: DOI:10.1007/s10822-005-8694-y

[69] Ying X. An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*. Institute of Physics Publishing; 2019. Available at: DOI:10.1088/1742-6596/1168/2/022022

[70] Jaworska J., Nikolova-Jeliazkova N., Aldenberg T. QSAR applicability domain estimation by projection of the training set in descriptor space: A review. *ATLA Alternatives to Laboratory Animals*. 2005; 33(5): 445–459. Available at: DOI:10.1177/026119290503300508

[71] Gramatica P. On the Development and Validation of QSAR Models. In: Reisfeld B, Mayeno AN (eds.) *Alternatives to laboratory animals: ATLA*. Totowa, NJ: Humana Press; 2013. pp. 499–526. Available at: DOI:10.1007/978-1-62703-059-5_21

[72] Tetko I v., Sushko I., Pandey AK., Zhu H., Tropsha A., Papa E., et al. Critical assessment of QSAR models of environmental toxicity against *tetrahymena pyriformis*: Focusing on applicability domain and overfitting by variable selection. *Journal of Chemical Information and Modeling*. 2008; 48(9):

1733–1746. Available at: DOI:10.1021/ci800151m

[73] Sushko I. Applicability domain of QSAR models. The Technical University of Munich; 2011.

[74] Dennis JS., Pacciani R. The rate and extent of uptake of CO₂ by a synthetic, CaO-containing sorbent. Chemical Engineering Science. 1 May 2009; 64(9): 2147–2157. Available at: DOI:10.1016/j.ces.2009.01.051

6 Machine Learning for the Exploration of Metal-Organic Frameworks for Hydrogen Storage

6.1 Keywords

Machine learning, hydrogen, metal organic frameworks, Grand Canonical Monte Carlo

6.2 Abstract

Using machine learning in the development of metal organic frameworks (MOFs) for various applications, is a field that is gaining momentum, particularly for gas storage, namely hydrogen (H_2). One of the main reasons behind this is the promising nature of H_2 as an automotive fuel, and in order to adopt this new technology, a suitable material is required. MOFs are posed as a viable solution as they have properties that counter the main challenge with H_2 which is the difficulty in storage and transportation due to the low volumetric density. The design and synthesis of MOFs is usually highly experimentally based, with some promising work in using high-throughput computational screening, which can be equally expensive and time consuming. Alternatively, using machine learning to screen existing and hypothetical MOFs for applications of interest can avoid these issues. In our study we have taken elements from the different approaches currently available to screen and identify properties of a high performing H_2 storage MOF, further stream-lining the computational application in material discovery. Our approach eliminates the longer computational methods such as GCMC or DFT, whilst showing patterns in data i.e., the influential factors of several input parameters. In using multiple linear regression and patterns in data, a MOF that had high usable gravitational and volumetric uptake (mof_10846), was identified albeit a hypothetical MOF. This further allows for the exploration of optimal MOFs to be accelerated experimentally via extracting vital data from the graphs produced in this study.

6.3 Introduction

Computational methods in the development and discovery of new materials, is an increasingly popular tool, as it allows for not only the interpretation of experimental results but to also act as a guide into which areas to investigate further. Hydrogen production and use is also rapidly developing its market potential and scalability as it is a promising replacement for fossil fuels because it reduces harmful greenhouse gas emissions into the environment. Additionally, it is an attractive automotive fuel source for many reasons, including the high potential to reduce carbon emissions by decarbonising multiple sectors, being more efficient due to the high gravimetric energy density, compared to hydrocarbon fuels such as diesel or gasoline [1, 2]. This therefore gives H₂ the ability to store a lot more energy for every unit weight of fuel in comparison.

However, to store hydrogen under relatively mild conditions upon a solid material, a high-capacity material is required to store the gas, which can be difficult due to the molecular properties of hydrogen. Whilst hydrogen has almost three times the energy content of gasoline with 120 MJ/kg versus 44 MJ/kg respectively, this is also inversely reflected on a volume basis. The equivalent amount of hydrogen requires more storage space than gasoline with liquid hydrogen having a density of 8 MJ/L whereas gasoline has 32 MJ/L [3]. Other than associated costs and practicality, the main issue is the amount of hydrogen that can safely be stored in a vehicle and delivered as required under operational conditions [4].

Among different proposed solutions, such as high-pressure gas or liquid hydrogen, physical adsorption at cryogenic temperature (~77 K) in porous materials, primarily on the internal and external surfaces of a nano-porous solid, is an attractive option due to the speed and reversibility, whilst being maintained at relatively low storage pressures (~up to 100 bar) [5]. Hydrogen physisorption also has a low enthalpy of adsorption, compared to chemisorption, which therefore reduces thermal management issues that can be problematic [6]. For these reasons, a lot of research has gone into developing an optimal material, particularly metal organic frameworks (MOFs).

MOFs are made from inorganic nodes and organic linkers and are a class of often crystalline materials that can be tuned to have ultrahigh porosity and high surface areas, made from infinitely extendable metal ligand networks [7][8]. These properties make MOFs an effective candidate for storage and transportation applications in clean energy, for gases such as hydrogen and methane, and as high-capacity adsorbents to meet various separation needs. As advantageous as MOF are, there are also some downsides to MOFs. These challenges vary greatly depending on the MOF class, but can include laborious or expensive synthesis, lack of electrical conductivity, and chemical, thermal, and mechanical stability to name a few [9]. For example, in terms of synthesis, difficulties can include understanding and controlling both structural and compositional complexity.

Consequently, owing to the difficulty in exhaustively assessing large libraries of these complex materials experimentally, the use of computational chemistry has been widely employed to accelerate the evaluation of such materials, particularly the use of GCMC [10]. GCMC is a statistical-mechanical method, in which a thorough molecular-level modelling of adsorption is solved accurately [11]. It is an approach that is widely used for the simulation of adsorption equilibrium, thus making it beneficial to model H₂ adsorption on a MOF. However, a main downside is computational time which is why ML is sometimes favoured.

There are a number of studies that have explored MOFs for H₂ uptake, using either ML, GCMC or both with interesting results [12–17]. A largely cited paper in the field of MOFs for H₂ uptake is by Ahmed et al. [18], which looks at using computational screening via GCMC to examine 493,458 MOFs based on the crystallographic properties and the Chahine rule. The conditions used were a pressure swing between 100 bar and 5 bar at 77 K and the benchmark used was the best performing MOF at the time which was identified to be IRMOF-20, which had been found to have the best combination of gravimetric & volumetric capacities under these pressure-temperature conditions [19]- a tricky balance to achieve. From this study, three MOFs surpassing IRMOF-20 were found, and experimentally validated with positive results, yielding the current record-holders

for balanced hydrogen capacity- SNU-70, UMCM-9, and PCN-610/NU-100, with a ceiling of ~ 40 g H₂/L. It should be noted that although there are MOFs with higher volumetric capacities, this study focused on high volumetric along with gravimetric capacity, thus these are the current highest performing MOFs in both aspects.

Bucior et al. [20] used ML to train on MOF databases, and in addition, they calculated new descriptors for gas adsorption derived from the energetics of MOF–guest interactions. The findings from this showed that their ML method was three times faster than conventional molecular simulations, which allows for rapid exploration of large numbers of MOFs. In order to validate their method, it was applied to screen experimental MOFs and the selected highly performing MOF was chosen to be experimentally validated, with positive results. Another study that applied the use of machine learning in combination with other methods is by Giappa et al., [3] where they employed multiscale theoretical calculations such as binding energy, to identify and explore the effect of the functional group composition on the hydrogen uptake profile. Consequently, GCMC was then applied to three MOF candidates, with some functional group modifications, and the total gravimetric and volumetric uptake profile was obtained both at cryogenic (77 K) and ambient temperature (300 K). The results obtained, showed a fair prediction capability of the H₂ binding energies even with a limited amount of data, and that the use of ML can largely bypass the need for computational methods, by extrapolating for valuable physical-chemical insights.

Anderson et al. [21], successfully utilized both machine learning and molecular simulation for the exploration of hydrogen storage of porous crystals, including MOFs, amongst others. They gave insights into how inputs influenced the adsorption output, namely volumetric surface area (VSA) and gravimetric surface area (GSA), again two parameters that lead to difficulty in finding a good balance between volumetric and gravimetric uptake. In this study, over 18,000 adsorption simulations using the topologically based crystal constructor code, ToBaCCo, was used, alongside the application of a neural network model to predict hydrogen uptake and various conditions. Results showed that the combined

application of simulation data at various temperature and pressures, as inputs to a neural network, allowed for the prediction of hydrogen adsorption at different thermodynamic conditions, with final results determining that 62 g/L is an attainable deliverable hydrogen capacity for 100 bar/77 K ↔ 5 bar/160 K swing conditions.

Thornton et al. [22], looked at the use of computational tools to explore physical hydrogen storage. They applied a combination of thermodynamic models coupled with a neural network machine learning as well as data mining for validation, to address the limits of hydrogen storage using porous materials and also to identify new additional promising candidate materials to enhance the viability of a global hydrogen fuel economy. The database screened consisted of 850,000 materials, of which around 20 % were MOFs - either Computationally Ready Experimental (CoRE) MOFs [23] or Hypothetical MOFs [24]. In this study, they used the Langmuir adsorption model, to predict hydrogen uptake (N) at room temperature, using Equation(6-1, where P is the pressure, K is the Langmuir equilibrium constant, and M is the saturation capacity:

$$N = \frac{KP}{1+KP}M \quad (6-1)$$

From this they found that at cryogenic temperatures, the Langmuir model failed to adequately predict adsorption capacity because of the poor correlation between hydrogen density and pore volume at high pressures. Therefore, instead GCMC simulations were adopted to predict hydrogen uptake. Consequently, neural network ML models for gas adsorption were trained on the data generated by GCMC to predict hydrogen adsorption volumetric working capacity based on the following structural descriptors:

- Adsorption energy
- Void fraction
- Volumetric surface area (VSA)
- Gravimetric surface area (GSA)
- Pore diameter
- Density

Results from this study showed reasonably good agreement is found between GCMC and data-mined experimental data for some materials, particularly with hypothetical MOF-5059389 which had the highest working capacity of 40 H₂ g/L, along with MOF-210 and ZIF-8 with a similar uptake, with a comparison shown in Figure 6-1 [22]. Similarities across the top performing candidates included a void fraction close to 0.5 and a major pore diameter of around 10 Å, along with high surface areas above 3000 m²/cm³ and 5000 m²/g.

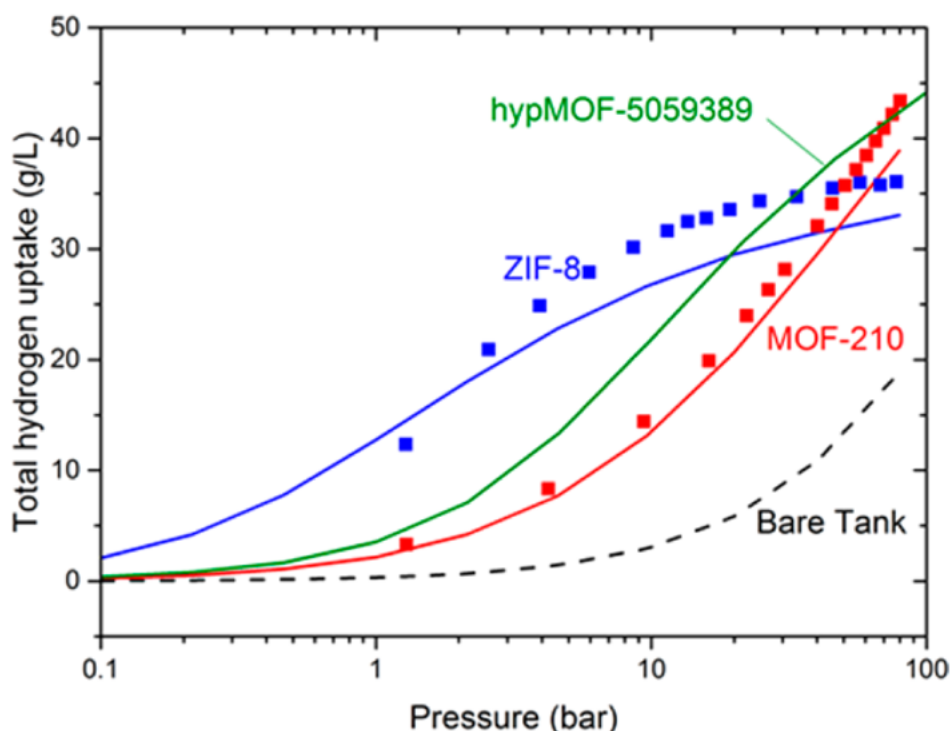


Figure 6-1 Total hydrogen uptake for top candidates with the highest working capacity of MOF-210, ZIF-8, and hypothetical MOF-5059389

An early study that adopted the accuracy of the ab initio methods and the speed of the classical approaches is by Borboudakis et al. [25]. Their proposed approach was to introduce machine learning for the prediction of gas adsorption in MOFs. In this study, data from literature of 100 MOFs that had been experimentally synthesized was collected, with conditions of 1 bar and 77 K. Across all selected MOFs, 62 organic linkers, 18 different types of metal nodes, and 12 functional groups were identified, and these were inputted into the JustAddData (JAD) Machine learning tool. Results from this study showed that

chemical properties of MOFs are predictable (stochastically, not deterministically) and that machine learning algorithms can learn predictive models that can generalize unseen MOFs (i.e., MOFs that have not been used to train the ML model), and guide experiments in material discovery by predicting properties of new materials, which leads us onto the aim of this study. The most recent study within this field in using machine learning for H₂ MOFs was the application of SHapley Additive exPlanations (SHAP) and contours to improve the accuracy prediction of a support vector regression (SVR) to observe the effects of temperature, pressure, isosteric heat of adsorption, BET surface area, on hydrogen uptake [26].

Whilst there is no known 'acceptable' storage capacity for storage use, the highest excess H₂ storage capacity reported so far for MOFs is 99.5 mg/g at 56 bar and 77K in NU-100, which has a total capacity of 164 mg/g at 77 K and 70 bar. The highest total H₂ storage capacity reported is 176 mg/g (excess 86 mg/g) in MOF-210 at 77 K and 80 bar [18].

In this study, we employ the application of machine learning, to screen and identify high performing MOFs for H₂ uptake, using available databases on the computational prediction of hydrogen storage capacities in MOFs. Starting from 98,694 MOFs, we apply multivariate analysis as well as looking at patterns in data to narrow down the database to the 'best' performing MOFs in terms of volumetric and gravimetric capacities.

6.4 Methodology

The approach for this study is shown in Figure 6-2.



Figure 6-2 Methodological approach for this study

In more detail, the process is:

1. Use the HyMARC databases consisting of hydrogen storage capacities of 98,695 MOFs in both usable **gravimetric** capacity and usable **volumetric** capacity [26].
 - These are calculated via GCMC.
2. Find the correlation between the desired output (H_2 uptake) of the GCMC simulated MOFs and their structural properties – using multiple linear regression (MLR) in Python to find correlation with simple geometric descriptors e.g., surface area, density etc. as inputs.
 - Plot **Actual** (GCMC from literature) vs. **Predicted** (ML), then use colour plots to find patterns in data.
3. Using the CCDC database, to identify high-performing hypothetical MOFs.
 - Use the actual vs predicted plot, to predict the deliverable capacity of hydrogen for the selected materials.
 - Using the ranges found from the predictive plots, apply this to available H_2 MOF databases to filter through and narrow down the best performing MOFs, based on highest uptake only in this study. Other KPIs that could be used include stability and reproducibility.

For step one, the input parameters were seven common crystallographic features that describe the MOF structure:

1. Single crystal density, **d** (g/cm³); range = 0.03 – 5.18
2. Gravimetric surface area, **gsa** (m²/g); range = 0 – 9750
3. Volumetric surface area, **vsa** (m²/cm³); range = 0 – 3995
4. Void fraction, **vf** (unitless, range of 0-1); range = 0 – 0.99
5. Pore volume, **pv** (cm³/g); range = 0 – 35.7
6. Largest cavity diameter, **lcd** (Å); range = 0.4 – 71.6
7. Pore limiting diameter, **pld** (Å); range = 0 – 71.5

Additional assumptions made were:

- The predictions of volumetric capacity assume that the MOF is a single crystal monolith.
- Hence, losses due to imperfect packing (i.e., void space) or structural defects (i.e., missing linkers and nodes) that might be present in a storage system where the MOF is stored as a powder or in an experimentally synthesized bulk material are not accounted for.
- The present volumetric predictions should therefore be considered as an upper bound on the practical volumetric capacity of the MOF storage medium [28], though it should be noted that in some cases structural defects could lead to experimental values higher than predicted [29].

Prior to commencing further, principal component analysis (PCA) was applied to ensure non redundancy via variance analysis. From the PCA heatmap, Figure 6-3, the variance across the 7 principal components, was shown to be fairly evenly distributed across the first few PCs (e.g., the first PC explains 71.8% of the variance, PC2, PC3, PC4, PC5, PC6 and PC7 explain 17.9%, 3.5%, 2.7%, 2.0%, 1.1% and 1.0% respectively. This indicates that 2 PCs highly influence the output (they explain 89.7% of the variance) and the latter 5 are almost redundant. With that said, all the 7 features were used as inputs to further analyse the influence on the output, in order to maximise the accuracy prediction and to see what the effect was with each input feature.

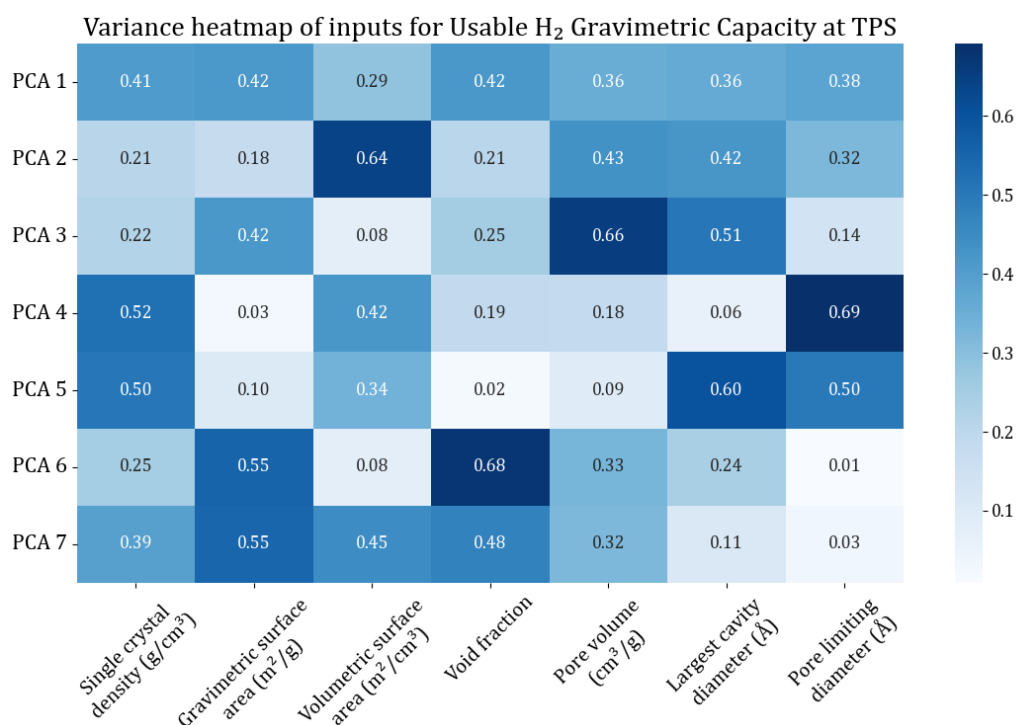


Figure 6-3 Heatmap showing dependencies of 7 principal components on input features.

The models developed were based on machine learning algorithms that were trained on computed H₂ uptake data (i.e., isotherms from Grand Canonical Monte Carlo calculations) on a diverse set of MOFs. The capacities were predicted using operating conditions [27]:

- Temperature + Pressure Swing (**TPS**) between P = 100 bar / T = 77 K (full condition) and P = 5 bar / T = 160 K (empty condition)

And two types of capacities are reported:

- Usable **gravimetric** capacity in weight percent (wt.%), defined as the mass of stored H₂ divided by the summed mass of H₂ and the MOF
- Usable **volumetric** capacity in grams of H₂ stored per litre of MOF (g H₂/L MOF)

6.5 Results and Discussion

The first models that were developed are shown in Figure 6-4, where a) is the usable volumetric capacity (gH_2/LMOF) and b) is usable gravimetric capacity (wt.%).

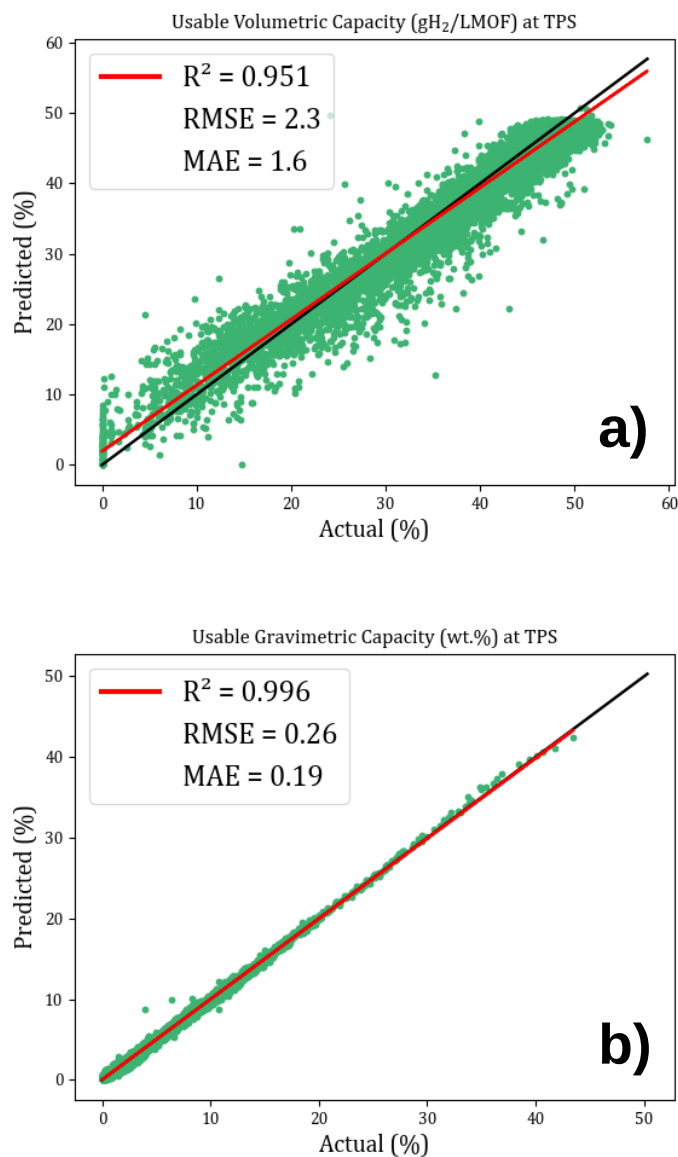


Figure 6-4 Scatter plots showing correlation between the input features and H_2 capacity at TPS for a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%). Black lines represent $y = x$. Red lines indicate the line of best fit.

From these plots, it is clear that the ability to predict the gravimetric capacity is much higher than that of volumetric capacity. A reason behind this includes that calculating the volumetric capacity of a material requires knowledge of its density, or bulk volume. More specifically, for crystalline materials, such as MOFs, this can be calculated from the amount of material bounded by the faces of the crystal (single crystal density); however, a downfall of this is that this ignores interparticle voids and structural defects present in real bulk samples, and therefore, results in an overestimation of the practical volumetric H₂ storage capacity, hence the variation [6], to overcome this, bulk density could be considered. Additionally, it is clear that the relationships between output capacities both volumetric and gravimetric, and the input features behave differently under different pressure swing and temperature pressure swing conditions [27], although the exact patterns are fairly ambiguous due to the complex nature of different MOFs; thus, it's more on a type of MOF basis as opposed to a holistic trend across all H₂ uptake MOF behaviour [30].

Following the predictions of the usable capacities from the seven inputs, colour plots were configured in order to observe patterns in data, which are shown from Figure 6-5 to Figure 6-11. It should be noted that the PLD of a MOF is defined as the aperture size of the material and is often used as a key structural property to evaluate the size exclusion capability for a gas mixture [31].

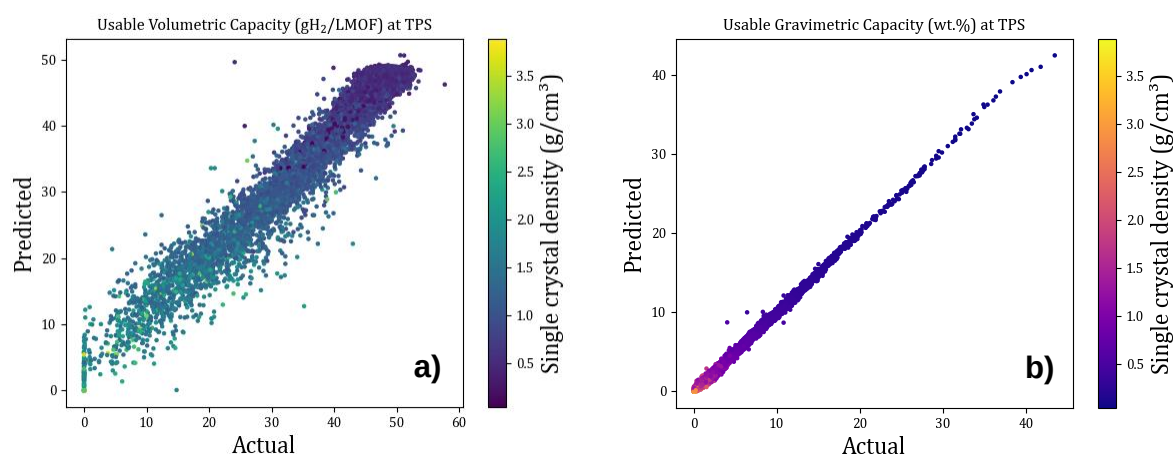


Figure 6-5 Actual vs Predicted plots using simple MLR using single crystal density (g/cm^3) to predict a) usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).

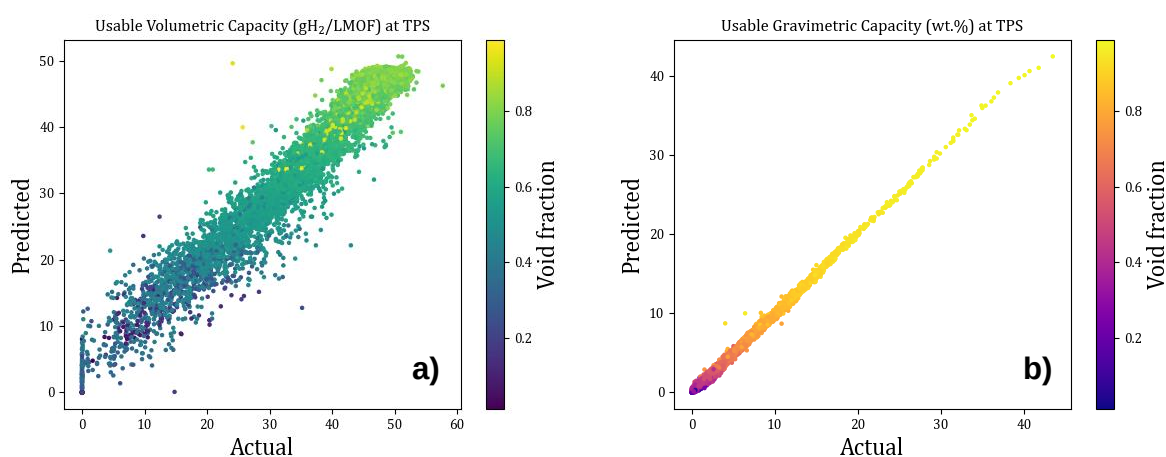


Figure 6-6 Actual vs Predicted plots using simple MLR using void fraction to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).

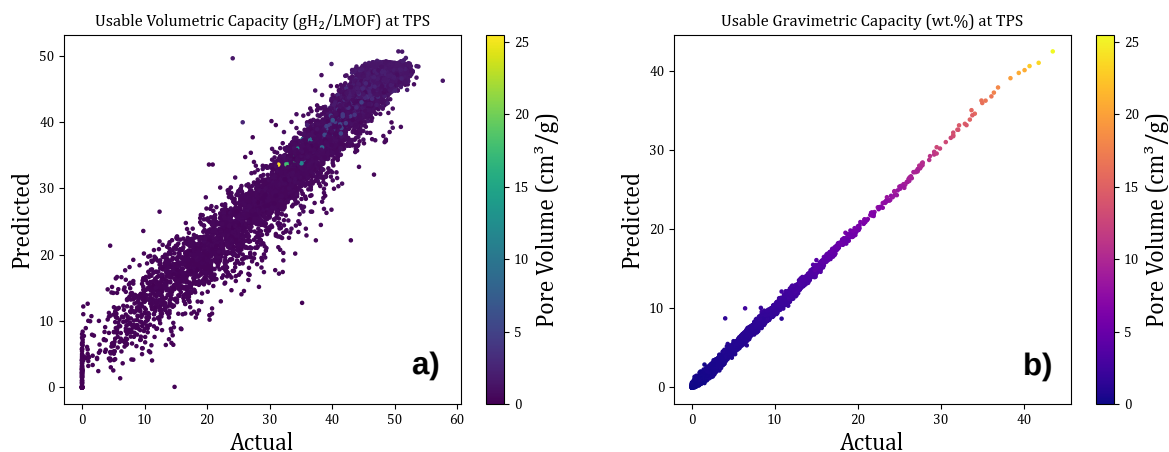


Figure 6-7 Actual vs Predicted plots using simple MLR using pore volume (cm^3/g) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).

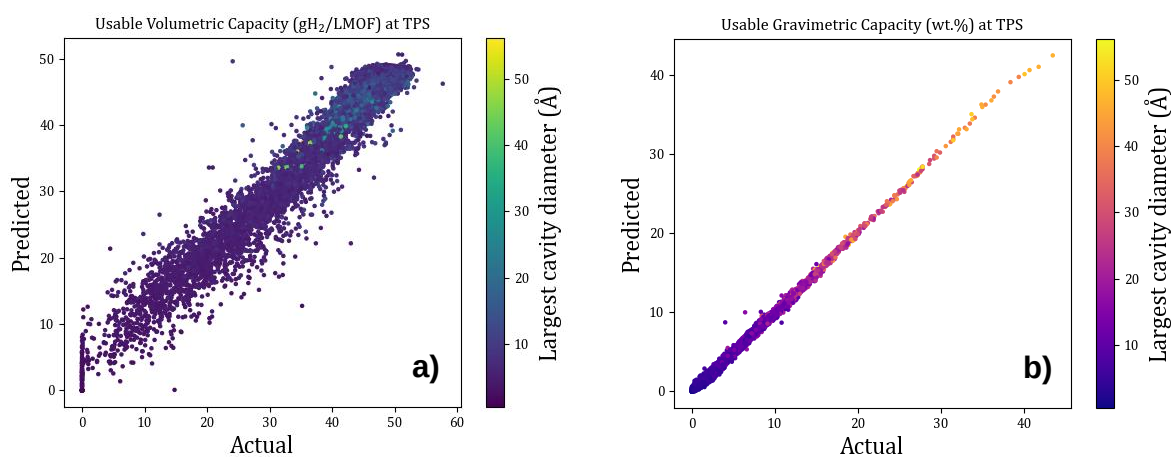


Figure 6-8 Actual vs Predicted plots using simple MLR using largest cavity diameter ($\text{\AA}$) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).

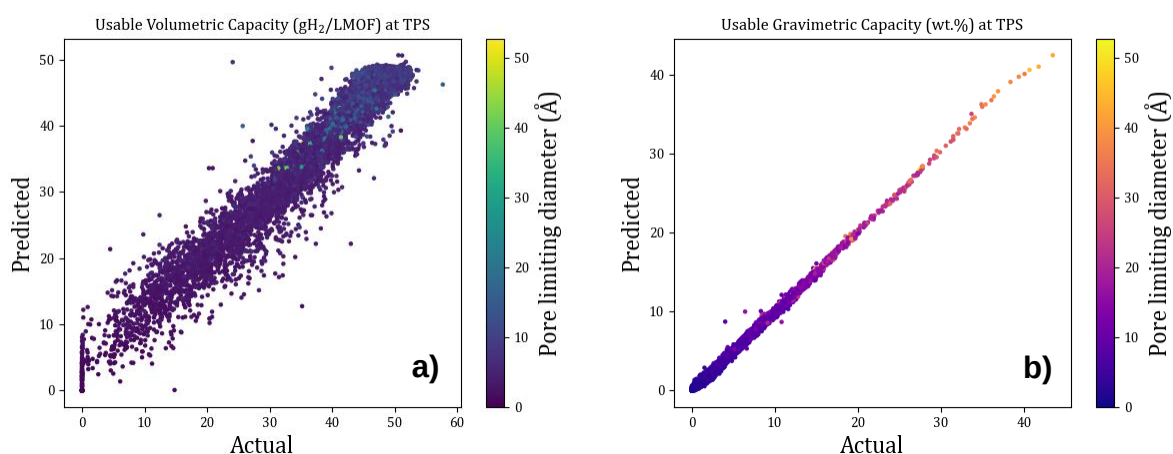


Figure 6-9 Actual vs Predicted plots using simple MLR using pore limiting diameter (Å) to predict a) Usable volumetric capacity (g H₂/L MOF) and b) usable gravimetric capacity (wt.%).

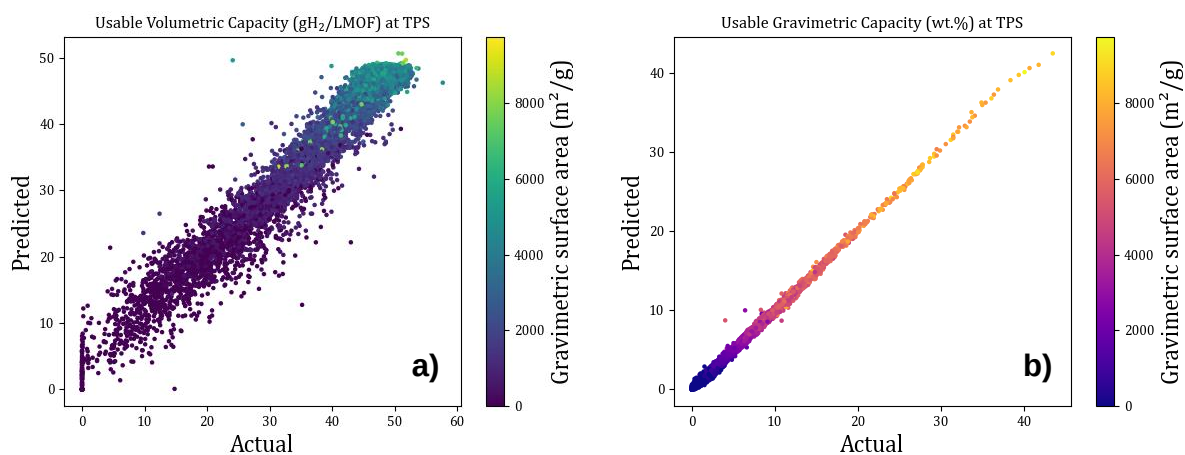


Figure 6-10 Actual vs Predicted plots using simple MLR using gravimetric surface area (m²/g) to predict a) Usable volumetric capacity (gH₂/LMOF) and b) usable gravimetric capacity (wt.%).

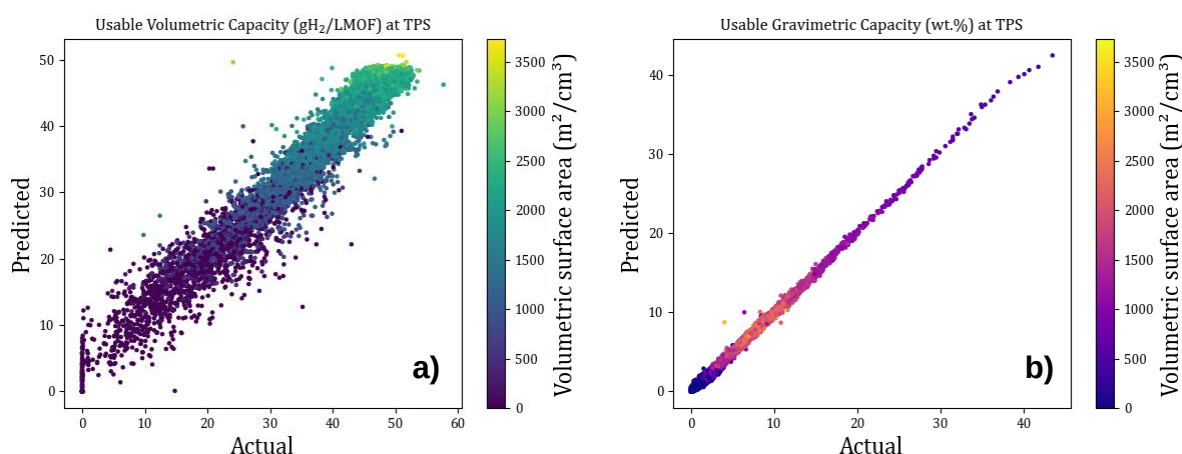


Figure 6-11 Actual vs Predicted plots using simple MLR using volumetric surface area (m^2/cm^3) to predict a) Usable volumetric capacity (gH_2/LMOF) and b) usable gravimetric capacity (wt.%).

From these plots, the clearest correlations in usable volumetric capacity, is shown from the following plots: single crystal density (Figure 6-5(a)), gravimetric surface area (Figure 6-10(a)), volumetric surface area (Figure 6-11(a)) and void fraction (Figure 6-6(a)). And for usable gravimetric capacity, the clearest correlations were seen in pore volume (Figure 6-7(b)), gravimetric surface area (Figure 6-10(b)), pore limiting diameter (Figure 6-9(b)) and largest cavity diameter (Figure 6-8(b)).

Whilst the aforementioned parameters gave the clearest correlations in the colour plots, this shouldn't be mistaken with the most influential for the given usable capacities, the clearer correlations are merely used as an indication as to what range to narrow down the database to for a high performing MOF. If anything, the plots where the colour gradient is spread across the scattered dots or there is a clear contrast between high and low, are more of an indicator as to which are most influential. So, in this case, for the usable volumetric capacity, the plots that have a wider spread of the coloured scattered dots are single crystal density (Figure 6-5(a)), volumetric surface area (Figure 6-11(a)) and void fraction (Figure 6-6(a)), and for usable gravimetric capacity, they are void fraction (Figure 6-6(b))

and gravimetric surface area (Figure 6-10(b)). Then looking at the plots that have a clear divide between high and low for both capacities this is pore volume (Figure 6-7), and additionally single crystal density for the gravimetric capacity (Figure 6-5(b)). Having said that, the research from which this data is from [27], gives the most influential parameters in order, having calculated it via ML models with one parameter as the input for the output capacity. Then results of this approach is shown in Table 6-1 and are in line with the colour plot indications, perfectly for the top 4 for usable gravimetric capacity, whilst only identifying two out of the top 4 for volumetric capacity. Again, this is down to the varying effect of temperature and pressure along with the property relationship with the input features on the output capacities.

Table 6-1 Performance of ML models for the prediction of usable H_2 storage capacities in MOFs under temperature & pressure swing (TPS) conditions, where d = Single crystal density (g/cm^3); pv = Pore volume (cm^3/g); vf = void fraction; gsa = Gravimetric surface area (m^2/g); vsa = Volumetric surface area (m^2/cm^3); lcd = Largest cavity diameter ($\text{\AA}$); pld = Pore limiting diameter ($\text{\AA}$). Performance from [18] is ranked based on R^2 , where green is the parameters identified as influential that match that of the original database study.

Usable volumetric capacity (g H_2 /LMOF)		Usable gravimetric capacity (wt%)	
From study [18]	This study	From study [18]	This study
vf	pv	pv	pv
pv	d	d	d
gsa	vsa	vf	vf
vsa	vf	gsa	gsa

By looking at patterns in data the properties that can lead to a high performing MOF can be narrowed down, and on using this method, the figures in Table 6-2 are what indicated a higher uptake for both volumetric and gravimetric capacities.

Table 6-2 The ranges obtained from the colour plots, indicating the properties that a high performing MOF would have.

Property	Range from Volumetric Capacity plot (gH ₂ /LMOF)	Range from Gravimetric Capacity plot (wt.%)
Single crystal density (g/cm ³)	0.03 – 0.75	
Gravimetric SA (m ² /g)	7000 – 9000	
Volumetric SA (m ² /cm ³)	2250 – 3250	2250 – 3250
Void fraction (-)	0.8 – 0.9	0.8 – 0.9
Pore volume (cm ³ /g)	20 – 25	
LCD (Å)	40 – 71.6	
PLD (Å)	40 – 71.5	

Note:

**VSA as a parameter was removed as the range didn't apply to both plots*

*** The range encompassing both was used: 0.8-1*

Interestingly and also expectedly, using the conditions in the table the database was narrowed down from 98,694 datapoints to 0, and the main reason for this is that although the gravimetric capacity plots were largely in agreement, using volumetric surface area as an input to predict this, largely contradicts the other parameters, and this is seen clearly in Figure 6-11(a). For this plot, there is a 'sweet spot' for optimal capacity, ranging from 2250-3250 m²/cm³ for volumetric surface area. However, the gravimetric plot indicates that the exact same range, results in a capacity of approx. 10 wt%, compared to the optimal capacity of around 50 wt% based on the data currently available. This is in line with literature as mentioned before, finding the balance between the usable gravimetric and volumetric capacities has long been one of the main targets within the MOF gas adsorption field [18,19,32,33]. Due to this, the parameter of volumetric surface area (vsa) was removed as an input, as it was less influential on the performance compared to gravimetric surface area (gsa) in either capacity output. In removing this parameter, the database was further amended from 0 datapoints, up to 34 MOFs of interest. Another interesting point as noted in Table 6-1, is that this study manage to predict the gravimetric capacity in line with the performance from study

[18], however not the volumetric capacity, which is expected due to the effect of volumetric surface area having contradictory results from the other input factors. All the 34 narrowed down MOFs were from the same database which was gathered from [10], which was a study on the use of GCMC to evaluate over 10,000 structures for the storage of gaseous fuels such as hydrogen and methane at various operating conditions, using the top-down algorithm via the code Topologically Based Crystal Constructor (ToBaCCo). They were all available as crystallographic information files (**CIF**) files, which give geometry information on molecular building blocks, as well as cell geometry, space group, cell coordinates etc. From this, similarities and differences were noted, and some are given below:

Similarities

- All MOFs have **Cu** present.
- All MOFs consist of **trinuclear Cu clusters** (Cu₃) connected by linkers containing chelating N groups.
- The majority of the MOF's have 21 or 33 Cu per unit cell.

Differences/ Patterns

- Only 1 MOF has **Zn** present. This results in the lowest **volumetric capacity**.
- The higher performing MOFs in terms of **volumetric capacity** consist of a metal count of **24** whereas lower performing have a count of **12**. The opposite is true to higher performing MOFs in terms of **gravimetric capacity**.
- In the unit cell parameters, larger **alpha** (α) angle sizes favour **gravimetric capacity** (where α is the geometric mean between sides ca and ab of a unit cell),
- Longer unit cell lengths favour **volumetric capacity**, expectedly.

Based on the original colour plots and the additional patterns in data found from CIF files, a MOF leaning towards a higher gravimetric capacity is the route to explore further. This is because the original prediction plots show that more of the geometrical descriptors are in agreement with a high performing MOF relative to

gravimetric capacity. Applying all parameters except vsa, results in (expectedly) a MOF with high gravimetric capacity, and a medium level volumetric capacity, seemingly giving a good balance. When ranking the 34 MOFs in terms of gravimetric capacity, the top 5 performing MOFs are shown in Table 6-3, all of which are hypothetical MOFs yet to be experimentally validated, but only computationally calculated using GCMC.

Table 6-3 Top performing MOFs for Usable gravimetric capacity in weight percent (wt.%), where d = Single crystal density (g/cm^3); pv = Pore volume (cm^3/g); vf = void fraction; gsa = Gravimetric surface area (m^2/g); vsa = Volumetric surface area (m^2/cm^3); lcd = Largest cavity diameter ($\text{\AA}$); pld = Pore limiting diameter ($\text{\AA}$), UG = Usable gravimetric capacity (wt%), and UV = Usable volumetric capacity (gH_2/LMOF)

Name	d	gsa	vsa	vf	pv	lcd	pld	UG	UV
mof_10846	0.04	7753	313	0.99	24.36	48.2	46.1	42.4	31.4
mof_10842	0.04	8467	346	0.99	24.02	44.6	42.7	42.3	31.6
mof_10900	0.04	7973	317	0.99	24.77	48.3	45.8	42.3	30.7
mof_10857	0.04	7613	309	0.99	24.24	47.9	45.5	42.1	31.2
mof_10807	0.04	8747	360	0.99	23.86	44.4	42.4	42.1	31.6

From further analysis, the topology of all five MOFs was found to be **srs**, which is a chiral net with 3-connected vertices. Many MOFs with **srs** topology have metal ion nodes, and on looking at the structure of other reported MOFs in literature with this topology via the Northwestern MOF database [34] and the Cambridge Structural Database, a database of small molecule crystal structures [35], the types of linkers required to synthesize these, are those that are not readily available and are extremely long, further making the task of experimentally validating it, very time consuming. When looking at the current highest performing MOFs for H_2 uptake, Table 6-4 shows how the top identified MOF from this study compares.

Table 6-4 Properties of previously reported high performing MOFs compared to the top performing hypothetical MOF identified in this study, where d = Single crystal density (g/cm^3); p_v = Pore volume (cm^3/g); vf = void fraction; gsa = Gravimetric surface area (m^2/g); vsa = Volumetric surface area (m^2/cm^3); l_{cd} = Largest cavity diameter ($\text{\AA}$); p_{ld} = Pore limiting diameter ($\text{\AA}$), UG = Usable gravimetric capacity (wt%), and UV = Usable volumetric capacity ($\text{g H}_2/\text{L MOF}$). Blue is the UV and UG for the MOF identified in this study, and the green and red indicates better or worse performance from different studies, respectively.

Property	mof_10846	MOF-5	IRMOF-20	SNU-70	UMCM-9	NU-100
Year	This study	2003	2017		2019	
Reference	[10]	[36]	[19]		[18]	
gsa	7753	3512	4073	4944	5039	6050
vsa	313	2172	2000	1905	1805	1603
vf	0.99	0.81	0.84	0.86	0.86	0.88
p_v	24.36	1.36	1.65	2.14	2.31	3.17
UG	42.4	51.9	51	47.9	47.4	47.6
UV	31.4	7.8	9.1	10.6	11.3	13.9

Whilst its clear the MOF identified in this study is just a hypothetical MOF, what sets its aside from the other ‘high performing’ MOFs is that it manages to achieve the balance between the two opposing capacities, thus making it more optimal, if it were to be validated experimentally, over the other MOFs that have much higher usable volumetric capacities, but very low usable volumetric capacities.

6.6 Conclusion and Future Work

In this study a literature review on the different approaches that have been applied to the development and discovery of hydrogen uptake MOFs was conducted. Consequently, a simple machine learning algorithm (multiple linear regression) was applied using Python and available databases consisting of seven input parameters to predict the usable gravimetric and usable volumetric capacities. The aim was to find a good balance between the two outputs, which is defined as the ideal properties of a gas adsorption MOF in terms of volumetric and gravimetric capacity, in various literature. The application of colour plots was

applied as way to streamline and bypass the complex computational approaches, with promising results. The top performing MOFs were identified and compared to current high performing MOFs. In doing this, our approach eliminates the longer computational methods such as GCMC or DFT, whilst showing patterns in data i.e., the influential factors of several input parameters. This further allows for the exploration of optimal MOFs to be accelerated via extracting vital data from the graphs produced in this study.

Future directions for this study would be to consider practical aspects such as ease of synthesis of the MOF, expected MOF robustness, and other heuristics based on experimental literature, to go on and select one MOF for experimental preparation. Additionally, an area of research that is worth exploring further in MOF discovery, is the application of quantitative structure property relationship analysis (QSPR). The use of QSPR is an emerging and helpful mathematical tool that allows the link between physical or chemical properties to predict the behaviour or desired characteristics of a molecule. The use of this application within the study of MOFs, particularly for carbon capture, is not particularly new, for example there are studies focused on using QSPR for screening of MOFs for CO₂/H₂ selectivity [37], CO₂ capture [38,39] and methane storage [40–42] to name a few, with most of them published in very recent years. However, despite the clear interest in this field, the novel aspect is that there are very few studies on the application of QSPR to screen and explore the use of suitable MOFs for hydrogen storage specifically. In fact, with just one key study [43], this is a novel area worth exploring further as there is high potential to obtain newer findings and predictions from updated descriptors and materials.

6.7 Acknowledgments

The authors acknowledge the financial support from the UK Engineering and Physical Sciences Research Council Doctoral Training Partnership (EPSRC DTP) grant No. EP/R513027/1, and the UKRI UK-Canada Globalink Research Award, grant No. NE/V019457/1

CRedit author statement

- **Paula Nkulikiyinka** - Conceptualization, Methodology, Software, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Funding acquisition
- **Stuart T. Wagland** - Writing - Review & Editing, Supervision
- **Vasilije Manovic** - Writing - Review & Editing, Supervision
- **Ashlee J. Howarth** - Supervision, Writing - Review & Editing, Software
- **Hudson Bicalho** - Writing - Review & Editing
- **Peter T. Clough** - Conceptualization, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

6.8 References

- [1] Balat M. Potential importance of hydrogen as a future solution to environmental and transportation problems. *International Journal of Hydrogen Energy*. 2008; 33(15): 4013–4029. Available at: DOI:10.1016/j.ijhydene.2008.05.047
- [2] Kapdan IK., Kargi F. Bio-hydrogen production from waste materials. *Enzyme and Microbial Technology*. 2006; 38(5): 569–582. Available at: DOI:10.1016/j.enzmictec.2005.09.015
- [3] Giappa RM., Tylianakis E., di Gennaro M., Gkagkas K., Froudakis GE. A combination of multi-scale calculations with machine learning for investigating hydrogen storage in metal organic frameworks. *International Journal of Hydrogen Energy*. Elsevier Ltd; 5 August 2021; 46(54): 27612–27621. Available at: DOI:10.1016/j.ijhydene.2021.06.021
- [4] Thomas KM. Adsorption and desorption of hydrogen on metal-organic framework materials for storage applications: Comparison with other nanoporous materials. *Dalton Transactions*. 2009; (9): 1487–1505. Available at: DOI:10.1039/b815583f

- [5] Villajos JA. Experimental Volumetric Hydrogen Uptake Determination at 77 K of Commercially Available Metal-Organic Framework Materials. C. MDPI AG; 5 January 2022; 8(1): 5. Available at: DOI:10.3390/c8010005
- [6] Broom DP., Webb CJ., Fanourgakis GS., Froudakis GE., Trikalitis PN., Hirscher M. Concepts for improving hydrogen storage in nanoporous materials. *International Journal of Hydrogen Energy*. Elsevier Ltd; 22 March 2019; 44(15): 7768–7779. Available at: DOI:10.1016/j.ijhydene.2019.01.224
- [7] Saraci F., Quezada-Novoa V., Donnarumma PR., Howarth AJ. Rare-earth metal-organic frameworks: From structure to applications. *Chemical Society Reviews*. Royal Society of Chemistry; 2020. pp. 7949–7977. Available at: DOI:10.1039/d0cs00292e
- [8] Howarth AJ., Liu Y., Li P., Li Z., Wang TC., Hupp JT., et al. Chemical, thermal and mechanical stabilities of metal-organic frameworks. *Nature Reviews Materials*. Nature Publishing Group; 2016. Available at: DOI:10.1038/natrevmats.2015.18
- [9] Hendon CH., Rieth AJ., Korzyński MD., Dincă M. Grand Challenges and Future Opportunities for Metal-Organic Frameworks. *ACS Central Science*. American Chemical Society; 28 June 2017; 3(6): 554–563. Available at: DOI:10.1021/acscentsci.7b00197
- [10] Colón YJ., Gómez-Gualdrón DA., Snurr RQ. Topologically Guided, Automated Construction of Metal-Organic Frameworks and Their Evaluation for Energy-Related Applications. *Crystal Growth and Design*. American Chemical Society; 1 November 2017; 17(11): 5801–5810. Available at: DOI:10.1021/acs.cgd.7b00848
- [11] Yun J-H., He Y., Otero M., Düren T., Seaton NA. Adsorption equilibrium of polar/non-polar mixtures on MCM-41: experiments and Monte Carlo simulation. 2002. pp. 685–692. Available at: DOI:10.1016/S0167-2991(02)80197-5

- [12] Gómez-Gualdrón DA., Colón YJ., Zhang X., Wang TC., Chen YS., Hupp JT., et al. Evaluating topologically diverse metal-organic frameworks for cryo-adsorbed hydrogen storage. *Energy and Environmental Science*. Royal Society of Chemistry; 1 October 2016; 9(10): 3279–3289. Available at: DOI:10.1039/c6ee02104b
- [13] Bobbitt NS., Chen J., Snurr RQ. High-Throughput Screening of Metal-Organic Frameworks for Hydrogen Storage at Cryogenic Temperature. *Journal of Physical Chemistry C*. American Chemical Society; 8 December 2016; 120(48): 27328–27341. Available at: DOI:10.1021/acs.jpcc.6b08729
- [14] Liu B., Yang Q., Xue C., Zhong C., Chen B., Smit B. Enhanced adsorption selectivity of hydrogen/methane mixtures in metal-organic frameworks with interpenetration: A molecular simulation study. *Journal of Physical Chemistry C*. 3 July 2008; 112(26): 9854–9860. Available at: DOI:10.1021/jp802343n
- [15] Glasby LT., Moghadam PZ. Hydrogen storage in MOFs: Machine learning for finding a needle in a haystack. *Patterns*. Cell Press; 9 July 2021; 2(7). Available at: DOI:10.1016/j.patter.2021.100305
- [16] García-Holley P., Schweitzer B., Islamoglu T., Liu Y., Lin L., Rodriguez S., et al. Benchmark Study of Hydrogen Storage in Metal-Organic Frameworks under Temperature and Pressure Swing Conditions. *ACS Energy Letters*. American Chemical Society; 2018. pp. 748–754. Available at: DOI:10.1021/acsenergylett.8b00154
- [17] Altintas C., Altundal OF., Keskin S., Yildirim R. Machine Learning Meets with Metal Organic Frameworks for Gas Storage and Separation. *Journal of Chemical Information and Modeling*. American Chemical Society; 2021. pp. 2131–2146. Available at: DOI:10.1021/acs.jcim.1c00191
- [18] Ahmed A., Seth S., Purewal J., Wong-Foy AG., Veenstra M., Matzger AJ., et al. Exceptional hydrogen storage achieved by screening nearly half a million

metal-organic frameworks. *Nature Communications*. Nature Publishing Group; 1 December 2019; 10(1). Available at: DOI:10.1038/s41467-019-09365-w

[19] Ahmed A., Liu Y., Purewal J., Tran LD., Wong-Foy AG., Veenstra M., et al. Balancing gravimetric and volumetric hydrogen density in MOFs. *Energy and Environmental Science*. Royal Society of Chemistry; 1 November 2017; 10(11): 2459–2471. Available at: DOI:10.1039/c7ee02477k

[20] Bucior BJ., Bobbitt NS., Islamoglu T., Goswami S., Gopalan A., Yildirim T., et al. Energy-based descriptors to rapidly predict hydrogen storage in metal-organic frameworks. *Molecular Systems Design and Engineering*. Royal Society of Chemistry; 1 February 2019; 4(1): 162–174. Available at: DOI:10.1039/c8me00050f

[21] Anderson G., Schweitzer B., Anderson R., Gómez-Gualdrón DA. Attainable Volumetric Targets for Adsorption-Based Hydrogen Storage in Porous Crystals: Molecular Simulation and Machine Learning. *Journal of Physical Chemistry C*. American Chemical Society; 1 October 2019; 123(1): 120–130. Available at: DOI:10.1021/acs.jpcc.8b09420

[22] Thornton AW., Simon CM., Kim J., Kwon O., Deeg KS., Konostas K., et al. Materials Genome in Action: Identifying the Performance Limits of Physical Hydrogen Storage. *Chemistry of Materials*. American Chemical Society; 11 April 2017; 29(7): 2844–2854. Available at: DOI:10.1021/acs.chemmater.6b04933

[23] Chung YG., Camp J., Haranczyk M., Sikora BJ., Bury W., Krungleviciute V., et al. Computation-ready, experimental metal-organic frameworks: A tool to enable high-throughput screening of nanoporous crystals. *Chemistry of Materials*. American Chemical Society; 11 November 2014; 26(21): 6185–6192. Available at: DOI:10.1021/cm502594j

[24] Wilmer CE., Leaf M., Lee CY., Farha OK., Hauser BG., Hupp JT., et al. Large-scale screening of hypothetical metal-organic frameworks. *Nature Chemistry*. February 2012; 4(2): 83–89. Available at: DOI:10.1038/nchem.1192

- [25] Borboudakis G., Stergiannakos T., Frysali M., Klontzas E., Tsamardinos I., Froudakis GE. Chemically intuited, large-scale screening of MOFs by machine learning techniques. *npj Computational Materials*. Nature Publishing Group; 1 December 2017; 3(1). Available at: DOI:10.1038/s41524-017-0045-8
- [26] Meduri S., Nandanavanam J. Prediction of hydrogen uptake of metal organic frameworks using explainable machine learning. *Energy and AI*. April 2023; 12: 100230. Available at: DOI:10.1016/j.egyai.2023.100230
- [27] Ahmed A., Siegel DJ. Predicting hydrogen storage in MOFs via machine learning. *Patterns*. Cell Press; 9 July 2021; 2(7). Available at: DOI:10.1016/j.patter.2021.100291
- [28] Müller K., Vankova N., Schöttner L., Heine T., Heinke L. Dissolving uptake-hindering surface defects in metal-organic frameworks. *Chemical Science*. Royal Society of Chemistry; 2019; 10(1): 153–160. Available at: DOI:10.1039/c8sc03735c
- [29] Basdogan Y., Keskin S. Simulation and modelling of MOFs for hydrogen storage. *CrystEngComm*. Royal Society of Chemistry; 14 January 2015; 17(2): 261–275. Available at: DOI:10.1039/c4ce01711k
- [30] Allendorf MD., Hulvey Z., Gennett T., Ahmed A., Autrey T., Camp J., et al. An assessment of strategies for the development of solid-state adsorbents for vehicular hydrogen storage. *Energy and Environmental Science*. Royal Society of Chemistry; 1 October 2018; 11(10): 2784–2812. Available at: DOI:10.1039/c8ee01085d
- [31] Hung TH., Lyu Q., Lin LC., Kang DY. Transport-Relevant Pore Limiting Diameter for Molecular Separations in Metal-Organic Framework Membranes. *Journal of Physical Chemistry C*. American Chemical Society; 23 September 2021; 125(37): 20416–20425. Available at: DOI:10.1021/acs.jpcc.1c05959

- [32] Chen Z., Li P., Anderson R., Wang X., Zhang X., Robison L., et al. Balancing volumetric and gravimetric uptake in highly porous materials for clean energy. Available at: <https://www.science.org>
- [33] Gómez-Gualdrón DA., Wang TC., García-Holley P., Sawelewa RM., Argueta E., Snurr RQ., et al. Understanding volumetric and gravimetric hydrogen adsorption trade-off in metal-organic frameworks. *ACS Applied Materials and Interfaces*. American Chemical Society; 4 October 2017; 9(39): 33419–33428. Available at: DOI:10.1021/acsami.7b01190
- [34] Bucior BJ., Rosen AS., Haranczyk M., Yao Z., Ziebel ME., Farha OK., et al. Identification Schemes for Metal-Organic Frameworks To Enable Rapid Search and Cheminformatics Analysis. *Crystal Growth and Design*. American Chemical Society; 6 November 2019; 19(11): 6682–6697. Available at: DOI:10.1021/acs.cgd.9b01050
- [35] Groom CR., Bruno IJ., Lightfoot MP., Ward SC. The Cambridge Structural Database. *Acta Crystallographica Section B Structural Science, Crystal Engineering and Materials*. 1 April 2016; 72(2): 171–179. Available at: DOI:10.1107/S2052520616003954
- [36] Rosi NL., Eckert J., Eddaoudi M., Vodak DT., Kim J., O’Keeffe M., et al. Hydrogen Storage in Microporous Metal-Organic Frameworks. *Science*. 16 May 2003; 300(5622): 1127–1129. Available at: DOI:10.1126/science.1083440
- [37] Dureckova H., Krykunov M., Aghaji MZ., Woo TK. Robust machine learning models for predicting high CO₂ working capacity and CO₂/H₂ selectivity of gas adsorption in metal organic frameworks for pre-combustion carbon capture. *Journal of Physical Chemistry C*. 2019; 123(7): 4133–4139. Available at: DOI:10.1021/acs.jpcc.8b10644
- [38] Burns TD., Pai KN., Subraveti SG., Collins SP., Krykunov M., Rajendran A., et al. Prediction of MOF Performance in Vacuum Swing Adsorption Systems for Post-combustion CO₂ Capture Based on Integrated Molecular Simulations,

Process Optimizations, and Machine Learning Models. *Environmental Science and Technology*. 2020; 54(7): 4536–4544. Available at: DOI:10.1021/acs.est.9b07407

[39] Fernandez M., Boyd PG., Daff TD., Aghaji MZ., Woo TK. Rapid and accurate machine learning recognition of high performing metal organic frameworks for CO₂ capture. *Journal of Physical Chemistry Letters*. 2014; 5(17): 3056–3060. Available at: DOI:10.1021/jz501331m

[40] Mahmoud E. Quantitative structure–property relationships from experiments for CH₄ storage and delivery by metal–organic frameworks. *Crystals*. 2020; 10(8): 1–15. Available at: DOI:10.3390/cryst10080700

[41] Wang R., Zhong Y., Bi L., Yang M., Xu D. Accelerating Discovery of Metal–Organic Frameworks for Methane Adsorption with Hierarchical Screening and Deep Learning. *ACS Applied Materials & Interfaces*. 2020; : acsami.0c16516. Available at: DOI:10.1021/acsami.0c16516

[42] Fernandez M., Woo TK., Wilmer CE., Snurr RQ. Large-scale quantitative structure-property relationship (QSPR) analysis of methane storage in metal-organic frameworks. *Journal of Physical Chemistry C*. 2013; 117(15): 7681–7689. Available at: DOI:10.1021/jp4006422

[43] Kim D., Kim J., Jung DH., Lee TB., Choi SB., Yoon JH., et al. Quantitative structure-uptake relationship of metal-organic frameworks as hydrogen storage material. *Catalysis Today*. 2007; 120(3-4 SPEC. ISS.): 317–323. Available at: DOI:10.1016/j.cattod.2006.09.029

7 General Discussion

The overall aim of this study was to successfully apply the use of machine learning algorithms and computational methods for the prediction of properties, materials and conditions of the sorption enhanced steam methane reforming process (SE-SMR). This has had varying success rates within the three applications outlined in this thesis.

Firstly, the application of machine learning to produce a surrogate model for the accessibility of process conditions within a reactor had accurate levels of prediction, with the exception of some concentrations and conditions. As previously discussed in Chapter 4, from the accuracy metrics for both algorithms, it is evident that the random forest performed better than the neural network in the ability to predict the concentration of the gas in each reactor, due to achieving higher R^2 and lower RMSE and MAE values.

The main takeaway points from this publication were which concentrations gave the best prediction and which conditions didn't perform as well, whilst looking at what the physical implications were. Specifically, regenerator CO_2 concentration, (Figure 7-1(e) and 7-2(e)). This was deducted to be due to the Aspen model calculation, which is able to calculate the most stable species in a gas stream at each temperature, which in this case at the given temperature range, is CO_2 . This is translated into the Python models, as the set of five features used from Aspen, allows the model to obtain a more accurate prediction.

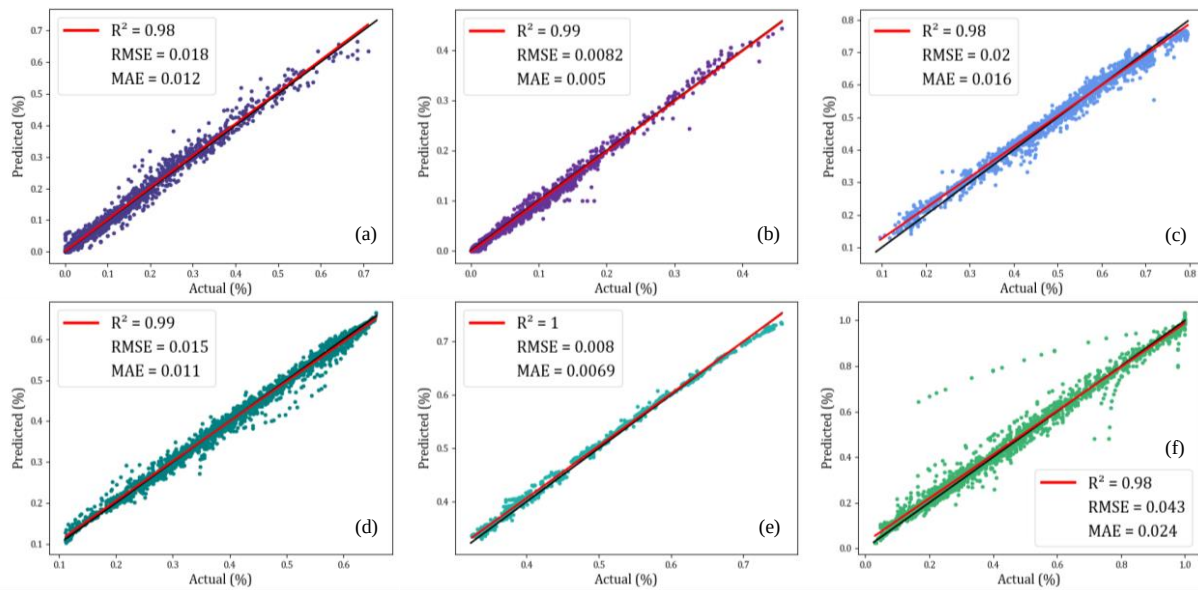


Figure 7-1 Actual vs Predicted plots by the ANN algorithm for: a) reformer $[H_2O]$, b) reformer $[CO]$, c) reformer $[H_2]$, d) regenerator $[H_2O]$, e) regenerator $[CO_2]$, f) reformer CH_4 conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.

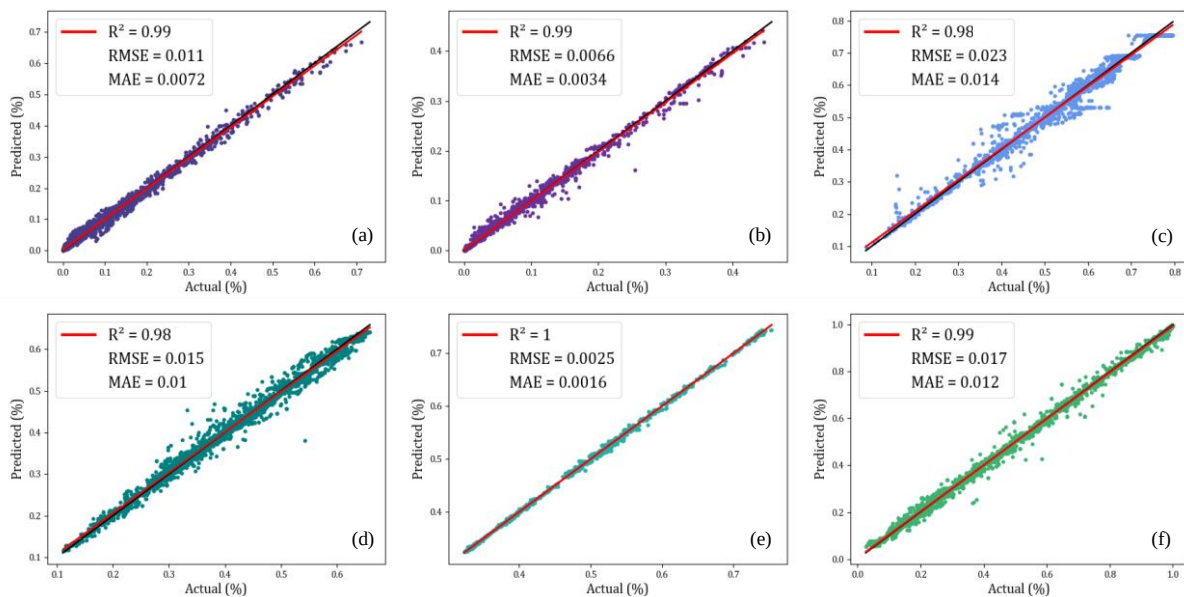


Figure 7-2 Actual vs Predicted plot by the random forest algorithm for: a) reformer $[H_2O]$, b) reformer $[CO]$, c) reformer $[H_2]$, d) regenerator $[H_2O]$, e) regenerator $[CO_2]$, f) reformer CH_4 conversion reformer concentration modelled. Black lines in each figure represent $y = x$. Red lines indicate the line of best fit.

Adversely, there are some apparently anomalistic datapoints seen in the neural network CH_4 conversion (Figure 7-1(f)) and the random forest H_2 concentration in the reformer (Figure 7-2(c)), with reasons behind these effects being suggested to be due to an insufficient set of features was selected to identify all reactions occurring, and H_2 being final and main product gas from all the reactions taking place in the reformer, therefore it is essentially least affected by slight changes in input features. This is because it is the agglomeration of 3 different reactions, however overall, both models were very accurate with high R^2 values, all above 98%. However, the random forest model was more precise in the predictions, with consistently higher R^2 values and lower mean absolute error (0.002-0.014) compared to the neural network model (0.005-0.024). Whilst the difference is small, due to the low values, an increment of even 0.001 makes a difference in this case.

The second publication was the use of machine learning, specifically quantitative structure–property relationship analysis (QSPR) alongside the use of inductive transfer learning, and datamining, for the development of combined sorbent and catalyst materials (CSCM) (Chapter 5). The accuracy success rate with this study was slightly less than the surrogate model development, in terms of the ability to predict the output based on a multiple linear regression (MLR) model. From the results, the application of multitask learning, (MTL), was shown to improve the prediction capability of new molecules, compared to single-output models. Consequently, the MTL model was used to predict properties of unseen CSCM molecules somewhat effectively, 58% of the data resulting in an average absolute relative error (AARE) of 50% or less (Figure 7-3).

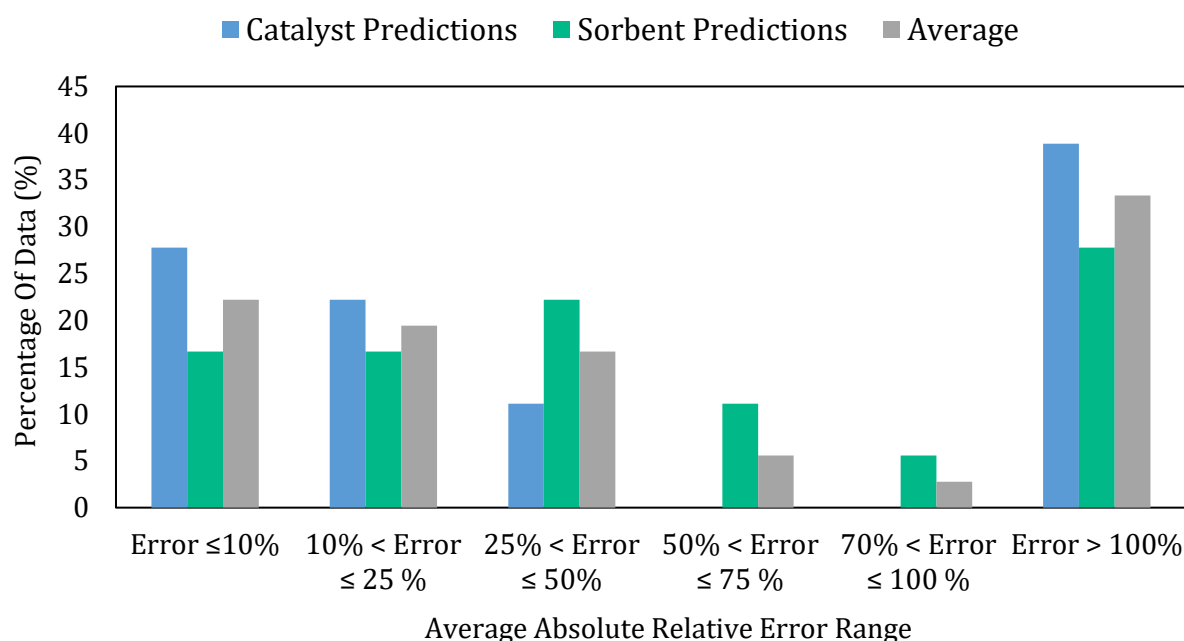


Figure 7-3 Comparison of the percentage of data belonging to each average absolute relative error (AARE) range for the last cycle capacity and methane conversion prediction of CSCM molecules.

Overall, this study presented a methodology for the prediction of CSCM that hasn't been conducted before, with promising results. The study concludes that a combination of machine learning models and application of molecular descriptors has the capability to predict properties for a CSCM for the process of SE-SMR. The results from this study had the aim of streamlining and accelerating the experimental discovery of an optimal CSCM, by reducing the repetitive trial and error processes involved in material development. This study also possibly paves a route into expanding the QSPR approach for applications beyond drug discovery or biochemistry, which is the main use of molecular descriptors and QSPR. The main outcome of this study was the use of colour plots to ascertain the best experimental conditions and material precursors that would result in a high capture rate and an optimal catalytic rate, for SE-SMR, with results shown in Table 7-1.

Table 7-1 Data gathered from the colour plots, indicating parameters that lead to a well/bad performing CSCM.

Parameter	Good performance	Bad performance
Calcium precursor	Calcium nitrate, Calcium gluconate	Calcium carbonate/ oxide
Inert support	Y ₂ O ₃ , CeO ₂	Ca ₁₂ Al ₁₄ O ₃₃ , CaZrO ₃
Synthesis method	Sol gel	Surfactant template
Optimal synthesis details:		
CaO concentration (%)	70-90%	
Cycles	10-15	
SMR reformer temperature	650°C	
SMR S/C ratio	6	
Carbonation time/ temperature	15-20mins / 650-700°C	
Calcination time/ temperature	5-10mins / 900°C	
Nickel wt%	2.5-10%	
Catalyst calcination temperature	600-700°C	
Catalyst calcination time	10 hours	

Lastly, the final study within this thesis was using machine learning to screen real and hypothetical MOFs and to use patterns in data, to identify a high-capacity MOF for hydrogen storage (Chapter 6).

From the plots produced, it was clear that the ability to predict the gravimetric capacity is much higher than that of volumetric capacity. A reason behind this includes that calculating the volumetric capacity of a material requires knowledge of its bulk volume or density, so for crystalline materials, such as MOFs, this can be calculated from the amount of material bounded by the faces of the crystal (single crystal density) however a downfall of this is that this ignores interparticle voids and structural defects present in real bulk samples, and therefore results in an overestimation of the practical volumetric H₂ storage capacity, hence the variation.

An important takeaway from this study was which parameters showed the clearest correlations in the colour plots and which were most influential in the ability to predict the usable capacity, which aren't interchangeable. From these plots, the clearest correlations in usable volumetric capacity, shown from, single crystal density, gravimetric surface area, volumetric surface area and void fraction. And for usable gravimetric capacity, the clearest correlations were seen in pore volume, gravimetric surface area, pore limiting diameter and largest cavity diameter.

Whilst the aforementioned parameters gave the clearest correlations in the colour plots, this shouldn't be mistaken with the most influential for the given usable capacities, the clearer correlations are merely used as an indication as to what range to narrow down the database to for a high performing MOF. If anything, the plots where the colour gradient is spread across the scattered dots or there is a clear contrast between high and low, are more of an indicator as to which are most influential. So, in this case, for the usable volumetric capacity, the plots that have a wider spread of the coloured scattered dots are single crystal density, volumetric surface area and void fraction, and for usable gravimetric capacity, they are void fraction and gravimetric surface area. Then looking at the plots that have a clear divide between high and low for both capacities this is pore volume, and additionally single crystal density for the gravimetric capacity.

The results from this colour plots, specifically by looking at patterns in data, the properties that can lead to a high performing MOF can be narrowed down, and on using this method.

The clear common interlinking within these chapters has been shown to be the application of multiple linear regression in order to test the predictive accuracy of the models to predict the process conditions and materials. Additionally, there is the common theme of employing graphical presentations of observing patterns in data via the use of colour plots, either narrow down an optimal range, or observe the conditions/parameters with the best predictability, or to identify the most influential factors, to name a few.

Despite the benefits of the colour plots, there are clearly still some inabilities to fully be able to predict unseen data, namely within the study regarding the use of QSPR for CSCM. From the results, the application MTL was shown to improve the prediction capability of new molecules, compared to single-output models, with the capability of predicting 58% unseen molecules with an AARE of 50% or less.

The training and test sets were applied to build and validate the models, while the external evaluation 'unseen' sets are used to assess the true predictive ability of the model. Since the obtained models rely on the molecule structures alone, i.e., the inputted data, the estimated values by the presented approach were only as good as the databases developed and the molecular structure descriptors which are not specifically built for the types of molecules evaluated, i.e., the structures sets are more suitable to drug discovery or biochemistry, which is the main use of molecular descriptors and QSPR. The application of QSPR in physical chemistry needs to be further explored in order to build up more suitable molecular descriptors, in this field of research. However, overall, this study has presented a methodology for the prediction of CSCM that hasn't been conducted before, with promising results.

8 General Conclusion and Future Work

8.1 Final Conclusions

The scope of this work has been the application of machine learning and computational methods, within the process of sorption enhanced steam methane reforming (SE-SMR) for the production of blue hydrogen. This approach aimed to streamline the overall process by allowing for accelerated and more holistic screening of materials for capture, catalysis and storage, as well as looking at the ability to monitor the process in circumstances where hardware sensors are not available or applicable.

Chapter 4 was to develop a surrogate/ soft sensor model which was done by firstly simulating the model in Aspen Plus, addressing Objective 1a, applying a sensitivity analysis to gather a large dataset, then applying two multiple linear regression model, to observe the accuracy of predicting the gas concentration outputs. Objective 1b was achieved, with two models being successfully developed with both models were accurate with high R^2 values, all above 98%.

Secondly, in chapter 5 a novel approach of QSPR with inductive transfer learning and datamining, was applied to develop two large databases of sorbent and catalyst properties, respectively, which addressed Objective 2a and 2b. Then the developed machine learning models from these databases were applied, to predict the optimal conditions and precursor materials for the highest performing CSCM, in terms of last cycle capacity and methane conversion. Lastly, in chapter 6, a similar approach was applied for the screening of MOFs for the storage of hydrogen by using multiple linear regression, simple geometric descriptors, and patterns in data to identify a better performing MOF than the currently reported experimental MOFs in literature, addressing Aim 3.

In conclusion, this PhD project addressed some of the key challenges of upscaling the process of sorption enhanced steam methane reforming. More specifically, tackling the issues that arise in all aspects of the SE-SMR value chain:

- The actual process monitoring itself,
- The sorbent and catalyst materials used within the process,
- and lastly the materials to store the produced hydrogen at the end of the process.

This study has explored the application of different levels of computational modelling to SE-SMR, to accelerate the scale up of this process to commercial scale. This is an area of significant importance as it helps tackle the race to obtain the net zero target. Using computational modelling tools is a time and cost-efficient solution to lead to the discovery of optimal materials to aid in the scale up of carbon-abated hydrogen production.

The intellectual contributions to this field include:

1. A soft sensor/surrogate model has been developed, with the ability to offer accurate gas concentration readings and methane conversion within the SE-SMR process.
2. A QSPR model has also been developed, with the uses of inductive transfer learning, that is able to predict the behaviour of unseen CSCM. The model in this study was the first to use feature nets and multitask learning in the SE-SMR field, and one of very few to use QSPR. It presented a methodology for the prediction of CSCM that hasn't been conducted before, with promising results.
3. And lastly MOF screening for H₂ storage using MLR and patterns in data has streamlined a somewhat complex task.

Process optimisation and optimal materials are some the key factors to accelerate the development of sorption enhanced steam methane reforming. The use of machine learning can accelerate the finding and designing of cost-effective materials for large-scale SE-SMR for carbon abated hydrogen, which in turn can lead us closer to a net zero future.

8.2 Limitations

Limitations that were faced within this thesis include the following:

- Time constraints
- Inability to conduct experiments.
- Lack of data

These limitations can be seen in all the in studies in this thesis, but each chapter was affected differently.

For the first study that produced a soft sensor/ surrogate model, the main limitation was the COVID-19 pandemic. This affected the study, as the preliminary hypothesis was to develop the model then prove it out experimentally, as the data used in the study was theoretical and gathered from Aspen Plus. The original goal was to conduct experiments on a rig that had been designed and constructed (see Figures 8-1 to 8-4). The software's to be used to observe the model's results vs the experimental results was to be Agilent VEE and LabVIEW.

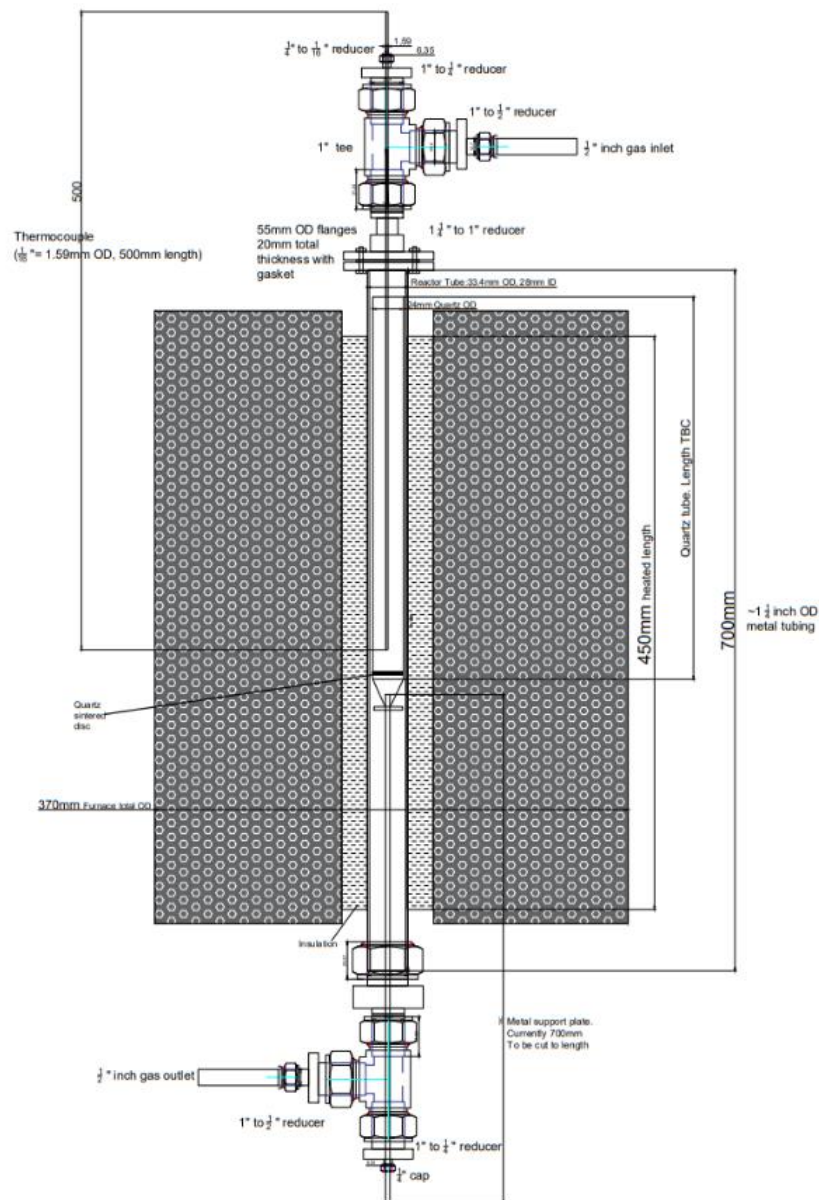


Figure 8-1 Sizing and design schematic of the reactor

However due to certain policies during the pandemic, the study was amended to be theoretical only, and to use existing literature as validation.

The positive side of this study was that, although the rig was not used in this study, an operational hydrogen reactor was built, and various skills were honed and applied.

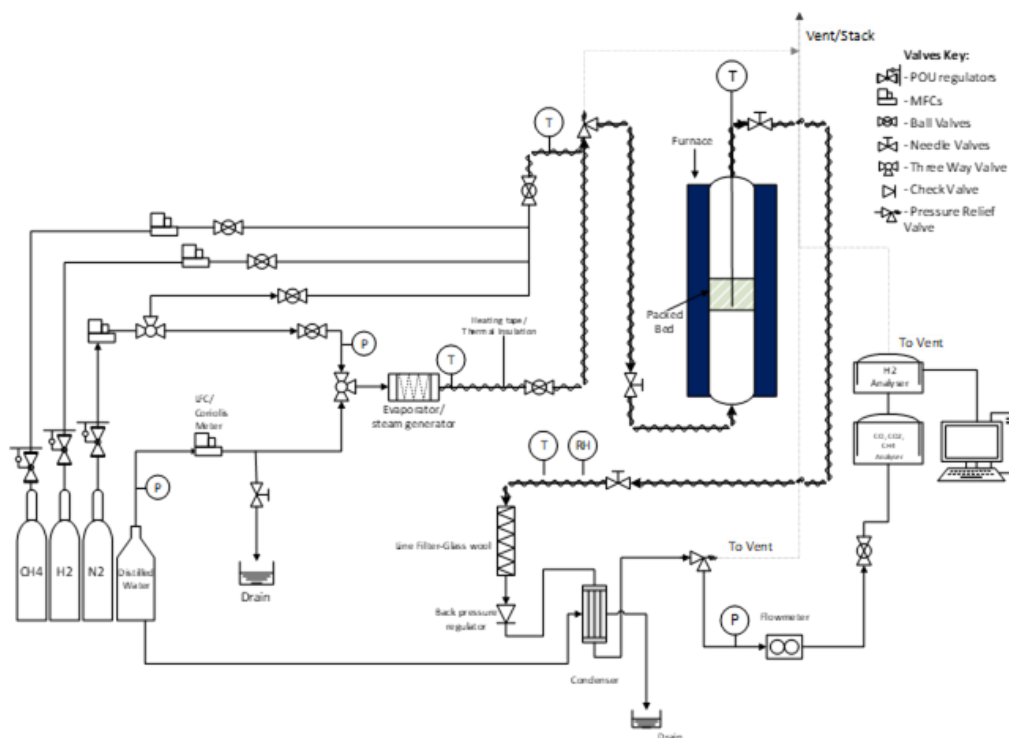


Figure 8-2 Process flow diagram of the proposed hydrogen production process



Figure 8-3 Constructed rig consisting of gas analysers on the left, mass flow controllers, a quartz reactor within a Carbolite furnace, joined via Swagelok fittings, and various other components.



Figure 8-4 Constructed rig, with software programming and process monitoring station on the right.

The second study was based on a very novel take on the use of machine learning, QSPR and SE-SMR, to explore and develop an optimal combine sorbent catalyst material (CSCM). The main restriction in this study was lack of data, as this was a new methodology, therefore quantity of data was limited.

The approach taken was to gather a large amount of data, then to narrow down certain sources that was more relevant to the study, i.e., the type of catalyst was methane based and the absorbent was calcium based. In doing this the data size was further reduced but this ensured that the end results were more accurate and in line with the scope of study, and ensured there was less noise in the results.

Additionally, an aim was to synthesise the ‘optimal’ material found from the study, however this was not completed, due to time limitations, however this is something that is still a goal despite completing the study.

The final study was the exploration of metal organic frameworks (MOFs) for hydrogen storage, through the application of machine learning. The main restriction in this study was time constraints as this was a study in collaboration with Concordia University, and the time was limited to 12 weeks. Again, the goal was to attempt to synthesise the hypothetical MOF identified in the study, however this is a very long process that can take months to years, especially when the linker used is long and complex and not readily commercially available.

The use of machine learning in SE-SMR is not new, however the different approaches employed in this study was able to look at the production process, materials used and end product (H₂) storage, in a different way and offer a methodology not used before.

Overall, despite the various obstacles encountered, the end results of each study in this thesis, was still able to meet the set objectives and offer novelty in an under- researched area. A lot of the original aims were quite ambitious and time willing, the theoretical results found, would have been experimentally applied for further validation. The ability to adapt based on time and practicality, is what has allowed this study to be successfully completed.

8.3 Future Work

All recommendations presented following each study were all practical based, that is, to experimentally validate the theoretical results presented in each piece of research.

8.3.1 Machine learning for the development of soft-sensor models

Future suggestions for this study would be to implement the soft sensor experimentally in order to compare the success rate of the soft sensor against hardware sensors, and thus further validating this machine learning application for the process of SE-SMR. The use of the term surrogate model and soft sensor model is sometimes interchangeable with the main difference of a soft sensor being 'online' whereas a surrogate model is 'offline'. Therefore, the next step of

this study would be essentially to put the developed surrogate model online, and effectively use the soft sensor practically to observe and control the process conditions.

8.3.2 Quantitative structure–property relationship analysis for the development of a combined sorbent catalyst material

Future directions for this study would be to use information from Table 7-1 to synthesise a new material using the model as is, then make changes to the models as appropriate, from the findings of this study, namely the reduction of experimental conditions. This would allow for, firstly, the synthesis of a material that would be assumed to possess the desired properties for a CSCM, but with the removal of the conditions after this data is obtained, the prediction of how it would perform is assumed to be more accurate with fewer conditions specified. The material could then be synthesised, and performance compared against the models' predictions, and thus further validating this QSPR and MTL application for the process of SE-SMR. This could also pave a route for the application of multitask learning to be more utilised to predict behaviour of materials that rely on more than one type of property feature.

8.3.3 Machine learning for the Exploration of Metal-Organic Frameworks for Hydrogen Storage

Future development of this study would be to consider practical aspects such as ease of synthesis of the MOF, expected MOF robustness, and other heuristics based on experimental literature, to go on and select one MOF for experimental preparation. the types of linkers required are not readily available and are extremely long, further making the task of making it very time consuming to the point of years as opposed to weeks or months. However, the proposed MOFs presented in this study, if successfully synthesised, would absolutely be better than the current physically available MOFs, the limitations are the linkers and other physical parameters, hence why all the top performing MOFs are hypothetical MOFs.

Additionally, an area of research that is worth exploring further in MOF discovery, is the application of quantitative structure property relationship analysis (QSPR). The use of QSPR is an emerging and helpful mathematical tool that allows the link between physical or chemical properties to predict the behaviour or desired characteristics of a molecule. The use of this application within the study of MOFs, particularly for carbon capture, is not particularly new, for example there are studies focused on using QSPR for screening of MOFs for CO₂/H₂ selectivity, CO₂ capture and methane storage to name few, with most of them published in very recent years.

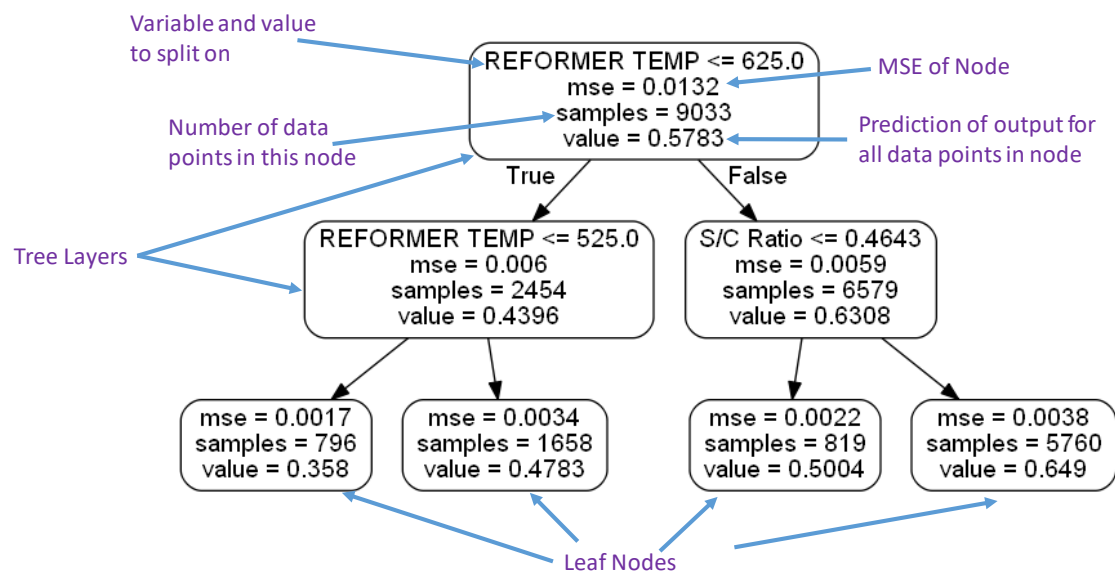
However, despite the clear interest in this field and the novelty that this topic presents, there are very few studies on the application of QSPR to screen and to explore the use of suitable MOFs for hydrogen storage specifically. With just one key study between the two topics, this is a niche area worth exploring further as there is a strong potential to obtain newer findings and predictions from updated descriptors and materials.

Appendices

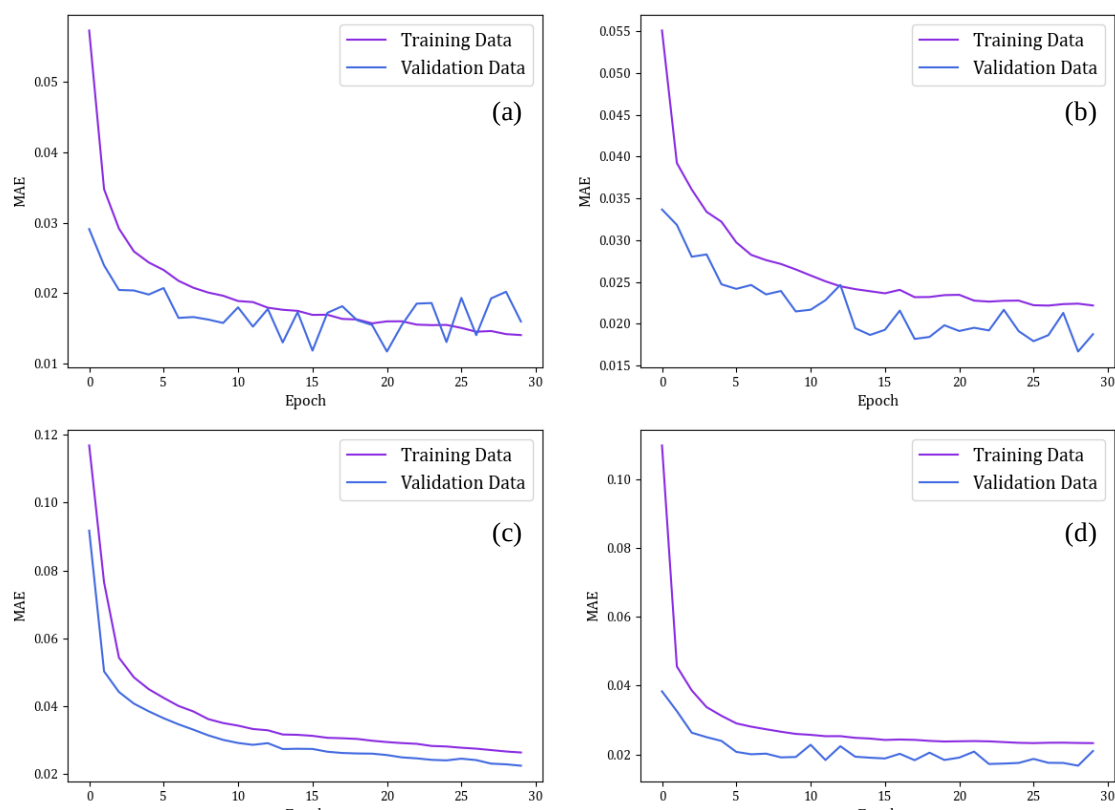
Appendix A - Prediction of Sorption Enhanced Steam Methane Reforming Products from Machine Learning Based Soft Sensor Models

Table_Apx A-1 Stream compositions under baseline conditions

Stream Name	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Temperature (°C)	25	25	25	150	123	382	599	650	650	650	850	850	619	407	25	25	25	25	25	25	25	25	25	850	850	850	850
Mole flows (kmol/hr)																											
CH ₄	1	0.5			1	1	1	0.02	0.02					0.02	0.02		0.02										
H ₂ O			4	4	4	4	4	2.12	2.12		1.34	1.34	1.34	2.12	0.11	2.01	0.11		0.15	1.19							
CO								0.08	0.08					0.08	0.08		0.08										
CO ₂								0.09	0.09		3.74	3.74	3.74	0.09	0.09		0.09		3.74								
H ₂								3.84	3.84					3.84	3.84		0.19	3.65									
CaO								52.70		52.70	55.76														53.53	2.23	55.76
CaCO ₃								0.81		0.81														2.24			
O ₂											1.83	1.83	1.83						1.83		3.00	3.00					
N ₂																					11.29		11.29				
Total	1	0.5	4	4	5	5	5	59.66	6.15	53.51	62.67	6.91	6.91	6.15	4.14	2.01	0.49	3.65	5.72	1.19	14.29	3	11.29	2.24	53.53	2.23	55.76



Figure_Apx A-1 Random Forest Model- close up annotation

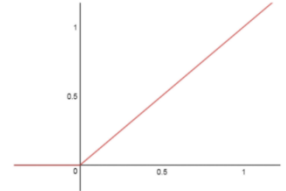
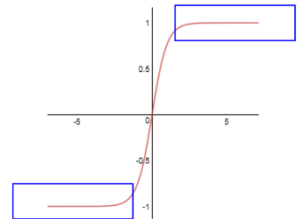
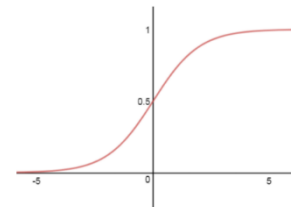


Figure_Apx A-2 Epoch vs MAE for reformer [H₂O], using different activation functions: a) ReLU, b) hyperbolic tangent (tanh), c) Sigmoid, d) Combination of Sigmoid, ReLU, Sigmoid for each consecutive layer.

Table_Apx A-2 Random Forest Feature Importance

Level of importance	Reformer [H ₂ O]		Reformer [CO]		Reformer [H ₂]		Regenerator [H ₂ O]		Regenerator [CO ₂]		[CH ₄] reformer conversion	
Most	S/C	0.7454	Ref T	0.5244	S/C	0.4339	CaO/C	0.5194	Reg T	0.6727	S/C	0.5124
	Reg T	0.1194	Ref P	0.1872	Reg T	0.3154	S/C	0.1779	CaO/C	0.308	Ref T	0.2469
	Ref T	0.086	S/C	0.1322	Ref T	0.1303	Reg T	0.1367	Ref T	0.0088	Ref P	0.1392
	Ref P	0.0489	Reg T	0.0828	Ref P	0.1187	Ref T	0.1052	Ref P	0.0085	Reg T	0.0918
Least	CaO/C	0.0003	CaO/C	0.0734	CaO/C	0.0017	Ref P	0.0609	S/C	0.002	CaO/C	0.0096

Table_Apx A-3 Details of Activation functions

Activation function	Description	Mathematical definition and diagram
Rectified Linear Unit (ReLU)	The activation is linear for input values greater than zero: $R(x) = \max(0, x)$. Other activation functions require very computationally expensive operations such as exponents, etc., but ReLU, on the other hand, can be implemented simply by thresholding a value vector at zero.	$R(x) = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$ 
Hyperbolic tangent (tanh)	The function maps a real-valued number to the range $[-1, 1]$. Its outputs are zero-centred.	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ 
Sigmoid	The sigmoid function maps the incoming inputs to a range between 0 and 1.	$f(x) = \frac{1}{1 + e^{-x}}$ 

Appendix B - Prediction of Combined Sorbent and Catalyst Materials for SE-SMR, Using QSPR and Multitask Learning

Table_Apx B-1 OECD guidelines for the development of a QSPR

Guideline	Examples/ Application
A defined endpoint	An example of this is having consistency in the data used, for the calculation of the endpoint property a specific instance of this is if data is collected from multiple literature sources. As this was the case for this study, care was taken to ensure all data was inputted in heterogenous units, and with consideration regarding the type of experiment conducted.
An unambiguous algorithm	An indicator of a good QSPR model is the transferability of the model, i.e., the ability to replicate the descriptor values, the accuracy metrics and ability for a new user to predict new models using the developed model easily. In this case, it has been confirmed that all the models developed are readily reproducible with the given input data.
A defined domain of applicability	The applicability domain (AD) is defined as the physico-chemical, structural, or biological space, knowledge or information on which the training set of the model has been developed, and for which it is applicable to make predictions for new compounds' [1]. Due to this its recommended that new data is predicted within the AD by interpolation as opposed to extrapolation. In this study the AD was well defined and further detail is given in Section 5.3.3.
Appropriate measures of goodness of fit, robustness and predictivity	Two examples of errors for this include use of excessive numbers of descriptors and failure to validate the model properly. Because of this, an evaluation of the number of descriptors was trialled. Furthermore, in accordance with the 'principle of parsimony', unnecessary additional descriptors were not utilised, as it often did not improve the fit. In terms of validation, its stated that use of external validation, i.e., using data not included in the training set is the only acceptable way to ensure good predictivity for a QSPR model.
A mechanistic interpretation, if possible	This can be validated by checking if the descriptors either have physiochemical interpretation that is consistent with a known mechanism or if data from literature can support the mechanistic basis of the model. In this case the latter was applied, in using experimental data from literature to validate the model against.

Table_Apx B-2 PCA Variance

Sorbent Variance	
PCA1	21.5%
PCA2	15.7%
PCA3	14.3%
PCA4	12.2%
PCA5	8.5%
PCA6	7.5%
PCA7	6.2%
PCA8	4.7%
PCA9	4.0%
PCA10	3.2%
PCA11	2.3%
Catalyst Variance	
PCA1	32.0%
PCA2	20.3%
PCA3	15.7%
PCA4	13.3%
PCA5	11.6%
PCA6	7.1%

Table_Apx B-3 Feature importance

Sorbent – Last cycle capacity (gCO₂/gSorbent)	
1	Precursor
2	CaO concentration (%)
3	Calcination temperature (°C)
4	Carbonation temperature (°C)
5	Synthesis Method
6	Calcination time(mins)
7	Initial cycle capacity (g/g)
8	Carbonation time (mins)
9	Cycle number
Catalyst – Methane conversion (%)	
1	Calcination temperature (°C)
2	SMR S/C ratio
3	BET surface area (m ² /g)
4	Nickel wt. %
5	SMR Reaction temperature (°C)
6	Time (h)

Table_Apx B-4 Sorbent Database- Property- Last cycle CO₂ capacity (g/g)

Molecule	CaO Concentration (%)	Calcium Precursor	Cycle Number	Method	Calcination Temperature (°C)	Calcination Time (mins)	Carbonation Temperature (°C)	Carbonation Time (mins)	Initial Capacity (gCO ₂ /gSorbent)	Cycle Last cycle CO ₂ capacity (gCO ₂ /gSorbent)	Ref
Ca ₁₂ Al ₁₄ O ₃₃	91	5	10	2	750	20	750	20	0.5562	0.5242	(2)
Ca ₁₂ Al ₁₄ O ₃₃	82.16	4	30	2	750	20	750	20	0.5283	0.54	(3)
Ca ₁₂ Al ₁₄ O ₃₃	80	8	7	1	800	10	650	30	0.523916	0.5	(4)
Ca ₁₂ Al ₁₄ O ₃₃	90	4	30	2	750	20	750	20	0.5272	0.5556	(3)
Ca ₁₂ Al ₁₄ O ₃₃	82	11	32	3	950	10	850	10	0.62	0.62	(5)
Ca ₁₂ Al ₁₄ O ₃₃	82	11	100	3	700	10	700	10	0.39	0.39	(5)
Ca ₁₂ Al ₁₄ O ₃₃	84	2	25	1	700	30	850	10	0.493344	0.42	(6)
Ca ₁₂ Al ₁₄ O ₃₃	80	1	30	4	750	10	750	10	0.55704	0.374	(7)
Ca ₁₂ Al ₁₄ O ₃₃	75	2	13	1	850	10	690	30	0.4	0.45	(8)
Ca ₁₂ Al ₁₄ O ₃₃	55	11	100	3	950	10	850	10	0.4	0.4	(5)
Ca ₁₂ Al ₁₄ O ₃₃	67	11	82	3	950	10	850	10	0.5	0.5	(5)
Ca ₁₂ Al ₁₄ O ₃₃	67	5	10	2	750	20	750	20	0.503	0.4696	(2)
Ca ₁₂ Al ₁₄ O ₃₃	34	1	50	1	800	5	650	40	0.258144	0.169598	(9)
Ca ₁₂ Al ₁₄ O ₃₃	80	10	7	1	800	10	650	30	0.495815	0.44	(4)
Ca ₁₂ Al ₁₄ O ₃₃	80	4	7	1	800	10	650	30	0.621946	0.6	(4)
Ca ₁₂ Al ₁₄ O ₃₃	35	2	13	1	850	10	690	30	0.215659	0.21	(8)
Ca ₁₂ Al ₁₄ O ₃₃	80	9	7	1	800	10	650	30	0.410043	0.36	[3]
Ca ₁₂ Al ₁₄ O ₃₃	80	7	7	1	800	10	650	30	0.581804	0.57	[3]
Ca ₁₂ Al ₁₄ O ₃₃	37	5	10	2	750	20	750	20	0.1467	0.1427	(2)
Ca ₁₂ Al ₁₄ O ₃₃	58.35	11	100	2	950	10	850	10	0.36	0.25	[4]
Ca ₁₂ Al ₁₄ O ₃₃	82	6	10	6	750	20	750	20	0.4421	0.3515	(10)
Ca ₁₂ Al ₁₄ O ₃₃	75	2	56	1	980	5	650	30	0.52009	0.220173	[7]

$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	80	4	100	2	950	2.5	650	2.5	0.409	0.1282	(11)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	75	2	20	4	750	8.33333	750	8.33333	0.3009	0.2543	(12)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	52	5	10	2	750	20	750	20	0.213	0.1994	(2)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	75	2	200	1	1000	15	600	25	0.1518	0.0871	(13)
Al_2O_3	80	4	11	2	850	10	650	20	0.429264	0.43	(11)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	75	4	45	1	850	10	690	30	0.3256	0.2728	(14)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82.16	3	30	2	900	10	650	30	0.3952	0.2562	(2)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	75	4	45	1	850	5	690	30	0.35	0.3	(15)
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	75	2	13	1	950	10	690	30	0.42	0.33	[7]
Ca_2MnO_4	80	4	100	2	950	2.5	650	2.5	0.271	0.0845	(11)
TiO_2	71.85	1	40	4	750	10	600	10	0.18743	0.2363	(16)
CaZrO_3	70	3	30	4	950	0	650	15	0.3555	0.3075	(17)
CaZrO_3	23.83	11	1200	3	700	30	700	30	0.11	0.11	[4]
CeO_2	85	6	18	2	700	20	600	45	0.59	0.59	(18)
La_2O_3	80	4	100	2	950	2.5	650	2.5	0.3474	0.0614	(11)
MgO	58	4	50	1	758	30	758	30	0.45	0.43	(19)
MgO	75	1	51	1	850	6	650	10	0.3345	0.27	(20)
MgO	80	4	100	2	950	2.5	650	2.5	0.3923	0.1417	(11)
MgO	80	1	1250	8	750	30	750	20	0.36916	0.17	(21)
SiO_2	65.1	13	50	2	700	5	600	60	0.4861	0.3348	(22)
TiO_2	90	1	10	4	750	10	600	10	0.2	0.24	(16)
TiO_2	83.44	1	40	4	750	10	600	10	0.310737	0.187695	(16)
Y_2O_3	19.9	6	190	8	740	10	740	10	0.06	0.077	(23)
Y_2O_3	80	6	10	2	950	5	650	30	0.5426	0.4879	(24)
Yb_2O_3	90	4	50	4	900	5	650	25	0.4	0.22	(25)
Yb_2O_3	90	4	15	4	800	5	650	25	0.51	0.53	(25)
ZrO_2	90	3	10	2	900	10	650	20	0.57	0.36	[2]
ZrO_2	64	3	10	2	900	10	650	20	0.33	0.21	[2]

ZrO ₂	80	3	10	2	900	10	650	20	0.49	0.31	[2]
ZrO ₂	76.9	1	15	7	750	15	600	30	0.19	0.14	(26)
CaZrO ₃	76.9	11	100	3	700	30	700	30	0.34	0.33	[4]
ZrO ₂	60.3	11	50	8	750	10	750	10	0.31	0.15	(27)
CaZrO ₃	66.7	11	100	3	950	10	850	10	0.21	0.21	[4]
ZrO ₂	60.3	11	50	3	750	10	750	10	0.31	0.32	(27)
ZrO ₂	60.3	11	23	3	950	0	850	10	0.25	0.23	(27)
CaZrO ₃	66.7	11	100	3	700	30	700	30	0.23	0.23	[4]
ZrO ₂	60.3	11	100	3	700	30	700	30	0.3	0.3	(27)
Ca ₁₂ Al ₁₄ O ₃₃	85	4	45	1	850	5	690	30	0.45	0.36	(15)
Al ₂ O ₃	91	6	30	2	750	20	750	20	0.52	0.55	[2]
CeO ₂	91	6	18	2	700	20	600	45	0.55	0.535	(18)
CeO ₂	80	6	18	2	700	20	600	45	0.57	0.49	(18)
CeO ²	95	6	18	2	700	20	600	45	0.44	0.45	(18)
Nd ₂ O ₃	90	10	24	4	900	5	650	30	0.56	0.4	(25)
Nd ₂ O ₃	60	10	24	4	900	5	650	30	0.46	0.31	(25)
Nd ₂ O ₃	30	10	24	4	1000	5	650	30	0.23	0.22	(25)
CaCO ₃	56	1	50	2	700	5	600	60	0.6683	0.277	(22)
SiO ₂	50	13	50	2	700	5	600	60	0.2825	0.1835	(22)
SiO ₂	75	13	50	2	700	5	600	60	0.5367	0.2985	(22)
CaO	100	1	50	2	700	5	600	60	0.5135	0.2675	(22)
Ca ₉ Al ₆ O ₁₈	77.9	1	25	4	750	30	650	30	0.2626	0.3333	(26)
Ca ₉ Al ₆ O ₁₈	59.2	1	25	4	750	30	650	30	0.1547	0.2033	(26)
Ca ₉ Al ₆ O ₁₈	77.9	1	25	4	750	30	650	30	0.2371	0.2912	(26)
Ca ₉ Al ₆ O ₁₈	59.2	1	25	4	750	30	650	30	0.2062	0.2599	(26)
Ca ₉ Al ₆ O ₁₈	43.2	1	25	4	750	30	650	30	0.1661	0.2408	(26)
Ca ₉ Al ₆ O ₁₈	29.3	1	25	4	750	30	650	30	0.1199	0.1867	(26)
Al ₂ O ₃	91.7	1	25	4	750	30	650	30	0.2198	0.2624	(26)

MgO	93.3	1	25	4	750	30	650	30	0.4148	0.3054	(26)
MgO	87.4	1	25	4	750	30	650	30	0.405	0.3045	(26)
MgO	77.7	1	25	4	750	30	650	30	0.4015	0.2814	(26)
MgO	73.6	1	25	4	750	30	650	30	0.4235	0.275	(26)
MgO	63.5	1	25	4	750	30	650	30	0.41	0.2454	(26)
CaCO ₃	56	1	25	4	750	30	650	30	0.265	0.2933	(26)
Y ₂ O ₃	83.2	1	25	4	750	30	650	30	0.1701	0.2719	(26)
ZrO ₂	76.9	1	25	4	750	30	650	30	0.2335	0.2168	(26)
CaZrO ₃	73.8	1	25	4	750	30	650	30	0.3439	0.2893	(28)
CaZrO ₃	55.6	1	25	4	750	30	650	30	0.2872	0.2595	(28)
CaZrO ₃	42.2	1	25	4	750	30	650	30	0.2267	0.2033	(28)
CaZrO ₃	31.9	1	25	4	750	30	650	30	0.1869	0.1823	(28)
MgO	77.7	1	25	4	930	10	650	30	0.4307	0.1907	(26)
CaZrO ₃	73.8	1	25	4	930	10	650	30	0.3248	0.298	(28)
Ca ₉ Al ₆ O ₁₈	77.9	1	25	4	930	10	650	30	0.3067	0.2273	(26)
CaCO ₃	56	1	25	4	930	10	650	30	0.2683	0.1266	(26)
Ca ₉ Al ₆ O ₁₈	77.9	1	25	4	930	10	650	30	0.2232	0.2107	(26)
Y ₂ O ₃	80	6	10	2	850	5	650	30	0.5965	0.5399	(24)
Y ₂ O ₃	80	6	10	2	950	5	650	30	0.5901	0.4974	(24)
Y ₂ O ₃	80	6	10	2	850	5	650	30	0.5838	0.5748	(24)
Y ₂ O ₃	80	6	10	2	1050	5	650	30	0.5843	0.3919	(24)
Y ₂ O ₃	80	6	10	2	950	5	650	30	0.5721	0.5021	(24)
Y ₂ O ₃	95	6	10	2	850	5	650	30	0.6058	0.5821	(24)
Y ₂ O ₃	80	6	10	2	850	5	650	30	0.5847	0.5756	(24)
CaO	100	1	10	2	950	5	650	30	0.5292	0.4183	(24)
Y ₂ O ₃	60	6	10	2	850	5	650	30	0.4365	0.4477	(24)
CaO	100	1	115	4	758	30	758	30	0.687016	0.178031	(29)
MgAl ₂ O ₄	90	4	131	4	758	30	758	30	0.613301	0.45249	(29)

MgAl ₂ O ₄	82	4	126	4	758	30	758	30	0.48762	0.437944	(29)
MgAl ₂ O ₄	74	4	126	4	758	30	758	30	0.48762	0.442778	(29)
MgAl ₂ O ₄	68	4	106	4	758	30	758	30	0.389737	0.405143	(29)
MgAl ₂ O ₄	58	4	132	4	758	30	758	30	0.322062	0.34132	(29)
CaO	100	1	60	4	758	30	758	30	0.687016	0.209582	[28]
MgAl ₂ O ₄	90	4	60	4	758	30	758	30	0.613301	0.540095	[28]
MgAl ₂ O ₄	82	4	60	4	758	30	758	30	0.48762	0.478465	[28]
MgAl ₂ O ₄	68	4	60	4	758	30	758	30	0.389737	0.410186	[28]
MgAl ₂ O ₄	58	4	60	4	758	30	758	30	0.322062	0.344325	[28]
Ca ₁₂ Al ₁₄ O ₃₃	87	2	50	1	850	5	700	30	0.448624	0.414285	[7]
Ca ₁₂ Al ₁₄ O ₃₃	80	4	11	2	850	10	850	30	0.4696	0.4586	(11)
Ca ₁₂ Al ₁₄ O ₃₃	66	1	50	1	800	5	650	40	0.25795	0.17097	(9)
Ca ₁₂ Al ₁₄ O ₃₃	90	7	28	1	800	10	650	30	0.59	0.51	[3]
Ca ₃ Al ₂ O ₆	91	1	50	5	850	10	650	30	0.6	0.41	(30)
Ca ₃ Al ₂ O ₆	91	1	100	5	850	10	650	30	0.6	0.34	(30)
La ₂ O ₃	80	4	11	2	850	10	850	30	0.4987	0.2197	(11)
La ₂ O ₃	80	4	11	2	850	10	850	30	0.6009	0.5808	(11)
Ca ₁₂ Al ₁₄ O ₃₃	80	4	11	2	850	10	850	30	0.4691	0.3188	(11)
Ca ₁₂ Al ₁₄ O ₃₃	87	2	50	1	980	5	700	30	0.52	0.22	[7]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	30	0.4919	0.4248	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	10	0.4739	0.3606	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	5	0.4357	0.2805	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	2	0.3258	0.1828	[29]
Ca ₃ Al ₂ O ₆	70	1	50	3	850	10	700	30	0.4508	0.358	[29]
Ca ₃ Al ₂ O ₆	74	1	50	3	850	10	700	30	0.4919	0.381	[29]
Ca ₃ Al ₂ O ₆	78	1	50	3	850	10	700	30	0.5671	0.3186	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	900	10	700	30	0.4858	0.3493	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	950	10	700	30	0.4865	0.2969	[29]

Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	30	0.4927	0.4243	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	950	10	700	30	0.4729	0.2799	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	950	10	700	30	0.4593	0.2304	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	680	30	0.5143	0.4343	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	700	30	0.4929	0.4215	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	725	30	0.4702	0.379	[29]
Ca ₃ Al ₂ O ₆	74	1	20	3	850	10	650	30	0.4582	0.2916	[29]
Ca ₁₂ Al ₁₄ O ₃₃	90	12	30	2	900	10	650	30	0.6827	0.624	[32]
Ca ₁₂ Al ₁₄ O ₃₃	85	12	30	2	900	10	650	30	0.6417	0.6096	[32]
Ca ₁₂ Al ₁₄ O ₃₃	75	12	30	2	900	10	650	30	0.5823	0.559	[32]
CaO	100	1	30	2	900	10	650	30	0.555	0.1277	[32]
Ca ₁₂ Al ₁₄ O ₃₃	65	12	30	2	900	10	650	30	0.4891	0.4795	[32]
Ca ₁₂ Al ₁₄ O ₃₃	90	12	30	2	950	10	650	30	0.6793	0.3638	[32]
Ca ₁₂ Al ₁₄ O ₃₃	85	12	30	2	950	10	650	30	0.6407	0.4152	[32]
Ca ₁₂ Al ₁₄ O ₃₃	75	12	30	2	950	10	650	30	0.5802	0.4401	[32]
CaO	100	1	30	2	950	10	650	30	0.552	0.089	[32]
Ca ₁₂ Al ₁₄ O ₃₃	65	12	30	2	950	10	650	30	0.4898	0.4107	[32]
Ca ₉ Al ₆ O ₁₈	80	10	35	2	900	10	650	30	0.5486	0.4665	[33]
Ca ₉ Al ₆ O ₁₈	80	10	35	2	800	10	650	30	0.5482	0.5148	[33]
Ca ₉ Al ₆ O ₁₈	80	10	35	2	1000	5	650	10	0.534	0.1933	[33]
Ca ₉ Al ₆ O ₁₈	80	10	35	2	800	10	650	30	0.5272	0.5244	[33]
Ca ₉ Al ₆ O ₁₈	80	10	35	2	900	10	650	10	0.5185	0.326	[33]
MgO	75	8	24	4	900	10	650	10	0.5758	0.5373	[34]
MgO	67	8	24	4	900	10	650	10	0.499	0.4806	[34]
CaO	100	1	24	4	900	10	650	10	0.4863	0.1891	[34]
MgO	52	8	24	4	900	10	650	10	0.391	0.3725	[34]
Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	800	10	700	30	0.587734	0.565443	[35]
Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	800	10	650	30	0.553211	0.56465	[35]

Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	800	10	600	30	0.460751	0.431713	[35]
Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	800	10	650	30	0.55382	0.564516	[35]
Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	850	10	650	30	0.552235	0.524109	[35]
Ca ₉ Al ₆ O ₁₈	92.5	4	31	2	900	10	650	30	0.58472	0.511036	[35]
Ca ₉ Al ₆ O ₁₈	80	4	31	2	930	5	650	30	0.519672	0.334915	[35]
CaZrO ₃	90	3	30	4	800	0	650	15	0.3101	0.3645	[17]
CaZrO ₃	82	3	30	4	800	0	650	15	0.2503	0.3113	[17]
CaZrO ₃	70	3	30	4	800	0	650	15	0.1743	0.2664	[17]
CaZrO ₃	90	3	30	4	950	0	650	15	0.4622	0.2992	[17]
CaZrO ₃	82	3	30	4	950	0	650	15	0.4344	0.3087	[17]
TiO ₂	90	1	10	4	750	10	600	10	0.206206	0.23313	[16]
TiO ₂	85	1	10	4	750	10	600	10	0.183203	0.159056	[16]
TiO ₂	92	1	10	4	750	10	600	10	0.172168	0.163632	[16]
TiO ₂	80	1	10	4	750	10	600	10	0.168648	0.159408	[16]
TiO ₂	95	1	10	4	750	10	600	10	0.113714	0.120908	[16]
Ca ₁₂ Al ₁₄ O ₃₃	82.16	3	30	2	900	10	650	30	0.4918	0.3088	[2]
MgO	85	8	24	4	900	10	650	10	0.6282	0.5693	[34]
MgO	95	8	24	4	900	10	650	10	0.6577	0.5589	[34]
Ca ₁₂ Al ₁₄ O ₃₃	85	2	21	4	750	8.33333	750	8.33333	0.0812	0.2105	[12]
MgO	25	2	21	4	750	8.33333	750	8.33333	0.0665	0.0636	[12]
Ca ₁₂ Al ₁₄ O ₃₃	75	2	20	4	750	8.33333	750	8.33333	0.043	0.0636	[12]
Ca ₁₂ Al ₁₄ O ₃₃	75	2	20	4	750	8.33333	750	8.33333	0.043	0.079	[12]
MgO	50	2	21	4	750	8.33333	750	8.33333	0.0323	0.0276	[12]
Ca ₁₂ Al ₁₄ O ₃₃	50	2	21	4	750	8.33333	750	8.33333	0.0125	0.0338	[12]
MgO	63	2	21	4	750	8.33333	750	8.33333	0.0724	0.0816	[12]
Ca ₁₂ Al ₁₄ O ₃₃	80	3	30	6	800	5	650	10	0.499695	0.298191	[36]
Ca ₁₂ Al ₁₄ O ₃₃	80	3	30	6	900	5	650	10	0.472178	0.254909	[36]
Ca ₁₂ Al ₁₄ O ₃₃	80	3	30	6	900	5	650	10	0.428466	0.189125	[36]

Ca ₁₂ Al ₁₄ O ₃₃	80	3	30	6	900	5	650	10	0.410982	0.213346	[36]
Ca ₁₂ Al ₁₄ O ₃₃	95	3	30	6	900	5	650	10	0.549851	0.164672	[36]
Ca ₁₂ Al ₁₄ O ₃₃	90	3	30	6	900	5	650	10	0.53663	0.205625	[36]
Ca ₁₂ Al ₁₄ O ₃₃	85	3	30	6	900	5	650	10	0.498256	0.22465	[36]
Ca ₁₂ Al ₁₄ O ₃₃	80	3	30	6	900	5	650	10	0.428765	0.189985	[36]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	30	6	900	5	650	10	0.397003	0.164672	[36]
Ca ₁₂ Al ₁₄ O ₃₃	85	3	15	6	900	5	650	10	0.412839	0.243346	[36]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	15	6	900	5	650	10	0.36416	0.20554	[36]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	18	4	900	10	650	30	0.5764	0.483	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	18	4	900	10	650	30	0.4963	0.3239	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	18	4	900	10	650	30	0.4295	0.2579	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	1	18	4	900	10	650	30	0.4188	0.1885	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	10	18	4	900	10	650	30	0.4772	0.3966	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	4	18	4	900	10	650	30	0.4705	0.3611	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	18	4	900	10	650	30	0.4682	0.2841	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	8	18	4	900	10	650	30	0.4283	0.3881	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	9	18	4	900	10	650	30	0.4184	0.2785	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	4	18	4	900	10	650	30	0.4028	0.1771	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	4	18	4	900	10	650	30	0.3736	0.1543	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	3	62	4	900	10	650	30	0.5764	0.3406	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	10	68	4	900	10	650	30	0.4779	0.3632	[37]
Ca ₁₂ Al ₁₄ O ₃₃	75	8	66	4	900	10	650	30	0.4298	0.3099	[37]
Ca ₁₂ Al ₁₄ O ₃₃	50	3	50	4	900	10	650	30	0.3315	0.2735	[37]
Ca ₁₂ Al ₁₄ O ₃₃	50	10	70	4	900	10	650	30	0.2719	0.2486	[37]
Ca ₁₂ Al ₁₄ O ₃₃	50	8	66	4	900	10	650	30	0.2068	0.2048	[37]
Al ₂ O ₃	90	4	11	2	850	10	650	20	0.461639	0.366498	[11]
Al ₂ O ₃	60	4	11	2	850	10	650	20	0.384814	0.462	[11]
Ca ₁₂ Al ₁₄ O ₃₃	91	4	10	2	750	20	750	20	0.6158	0.4625	[2]

$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	91	4	10	2	750	20	750	20	0.5516	0.5215	[2]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	91	1	10	2	750	20	750	20	0.0423	0.0897	[2]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	91	4	30	2	750	20	750	20	0.6188	0.3099	[2]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	91	5	30	2	750	20	750	20	0.5537	0.5067	[2]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82	4	10	6	750	20	750	20	0.5385	0.2976	[10]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82	4	10	6	750	20	750	20	0.4198	0.2897	[10]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82	1	10	6	750	20	750	20	0.5377	0.2983	[10]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82	6	10	6	750	20	750	20	0.5215	0.2131	[10]
$\text{Ca}_{12}\text{Al}_{14}\text{O}_{33}$	82	4	10	6	750	20	750	20	0.4853	0.3461	[10]
$\text{Ca}_3\text{Al}_2\text{O}_6$	75	2	30	1	850	5	700	30	0.4479	0.4346	[7]
$\text{Ca}_3\text{Al}_2\text{O}_6$	75	2	30	1	850	5	650	30	0.3617	0.46	[7]
$\text{Ca}_3\text{Al}_2\text{O}_6$	75	2	15	1	850	5	650	30	0.1746	0.3091	[7]
$\text{Ca}_3\text{Al}_2\text{O}_6$	75	2	15	1	850	5	650	30	0.0482	0.1867	[7]
$\text{Ca}_3\text{Al}_2\text{O}_6$	75	2	15	1	850	5	650	30	0.0341	0.1391	[7]
CaZrO_3	62.7	1	15	7	750	30	600	30	0.193194	0.12817	[28]
CaZrO_3	42	1	15	7	750	30	600	30	0.190553	0.134856	[28]
CaZrO_3	78.2	1	15	7	750	30	600	30	0.166702	0.130234	[28]
CaZrO_3	13	1	15	7	750	30	600	30	0.136413	0.095735	[28]
CaZrO_3	42	1	15	7	750	30	700	30	0.245857	0.247321	[28]
CaZrO_3	42	1	15	7	750	30	650	30	0.215933	0.200011	[28]
CaZrO_3	42	1	15	7	750	30	600	30	0.190568	0.134032	[28]

Table_Apx B-5 Catalyst Database- Property- Methane Conversion (%)

Molecule	Nickel wt. %	Calcination temperature catalyst	Reaction temperature (°C)	S/C ratio	Time (h)	BET Surface Area (m ² /g)	Methane Conversion (%)	Ref
SiO ₂ Al ₂ O ₃	4.34	900	650	1	50	105.8	0.075663	[38]
SiO ₂ Al ₂ O ₃	4.34	900	700	1	50	105.8	0.049667	[38]
SiO ₂ Al ₂ O ₃	4.34	900	750	1	50	105.8	0.985249	[38]
SiO ₂ Al ₂ O ₃	4.34	900	800	1	50	105.8	0.991883	[38]
SiO ₂ Al ₂ O ₃	4.34	900	850	1	50	105.8	1	[38]
SiO ₂ Al ₂ O ₃	4.34	900	900	1	50	105.8	0.996566	[38]
SiO ₂ Al ₂ O ₃	6.2	900	650	1	50	73.4	0.00011	[38]
SiO ₂ Al ₂ O ₃	6.2	900	700	1	50	73.4	0.001592	[38]
SiO ₂ Al ₂ O ₃	6.2	900	750	1	50	73.4	0.743141	[38]
SiO ₂ Al ₂ O ₃	6.2	900	800	1	50	73.4	0.859668	[38]
SiO ₂ Al ₂ O ₃	6.2	900	850	1	50	73.4	0.905794	[38]
SiO ₂ Al ₂ O ₃	6.2	900	900	1	50	73.4	0.97596	[38]
MgAl ₂ O ₄	15.3	350	600	5	9	70.2	0.419067375	[39]
MgAl ₂ O ₄	15.3	500	600	5	9	64.8	0.44841325	[39]
MgAl ₂ O ₄	15.3	650	600	5	9	50.2	0.4800525	[39]
MgAl ₂ O ₄	15.3	850	600	5	9	43.2	0.485568	[39]
MgAl ₂ O ₄	15.3	1000	600	5	9	42.3	0.486923375	[39]
MgAl ₂ O ₄	15.3	350	600	5	9	70.2	0.32975025	[39]
MgAl ₂ O ₄	15.3	500	600	5	9	64.8	0.3691085	[39]
MgAl ₂ O ₄	15.3	650	600	5	9	50.2	0.410592625	[39]

MgAl ₂ O ₄	15.3	850	600	5	9	43.2	0.418064	[39]
MgAl ₂ O ₄	15.3	1000	600	5	9	42.3	0.399417	[39]
MgAl ₂ O ₄	15.3	350	600	5	9	70.2	0.296565125	[39]
MgAl ₂ O ₄	15.3	500	600	5	9	64.8	0.3596845	[39]
MgAl ₂ O ₄	15.3	650	600	5	9	50.2	0.342117125	[39]
MgAl ₂ O ₄	15.3	850	600	5	9	43.2	0.379053875	[39]
MgAl ₂ O ₄	15.3	1000	600	5	9	42.3	0.348423375	[39]
Al ₂ O ₃	34	1000	700	2.4	12	37	0.78	[40]
NiAl ₂ O ₅	45	1000	700	2.4	12	25	0.82	[40]
Al ₂ O ₃	10	700	600	3	30	93	0.440502	[41]
Al ₂ O ₃	10	700	650	3	30	93	0.638497	[41]
Al ₂ O ₃	10	700	700	3	30	93	0.792961	[41]
Al ₂ O ₃	10	700	750	3	30	93	0.846399	[41]
Al ₂ O ₃	10	700	800	3	30	93	0.871973	[41]
Y ₂ Zr ₂ O ₇	9.7	800	550	2	0.5	1.5	0.211503	[42]
Y ₂ Zr ₂ O ₇	9.7	800	600	2	0.5	1.5	0.340291	[42]
Y ₂ Zr ₂ O ₇	9.7	800	650	2	0.5	1.5	0.462472	[42]
Y ₂ Zr ₂ O ₇	9.7	800	700	2	0.5	1.5	0.597867	[42]
Y ₂ Zr ₂ O ₇	9.7	800	750	2	0.5	1.5	0.698571	[42]
Y ₂ Zr ₂ O ₇	9.7	800	800	2	0.5	1.5	0.749708	[42]
Y ₂ Zr ₂ O ₇	9.8	800	550	2	0.5	20.5	0.323846	[42]
Y ₂ Zr ₂ O ₇	9.8	800	600	2	0.5	20.5	0.437764	[42]
Y ₂ Zr ₂ O ₇	9.8	800	650	2	0.5	20.5	0.614463	[42]
Y ₂ Zr ₂ O ₇	9.8	800	700	2	0.5	20.5	0.845681	[42]
Y ₂ Zr ₂ O ₇	9.8	800	750	2	0.5	20.5	0.926558	[42]

Y ₂ Zr ₂ O ₇	9.8	800	800	2	0.5	20.5	0.990913	[42]
Y ₂ Zr ₂ O ₇	9.8	800	550	2	0.5	62.6	0.261065	[42]
Y ₂ Zr ₂ O ₇	9.8	800	650	2	0.5	62.6	0.592986	[42]
Y ₂ Zr ₂ O ₇	9.8	800	700	2	0.5	62.6	0.77464	[42]
Y ₂ Zr ₂ O ₇	9.8	800	750	2	0.5	62.6	0.901776	[42]
Y ₂ Zr ₂ O ₇	9.8	800	800	2	0.5	62.6	0.949612	[42]
Y ₂ Zr ₂ O ₇	9.8	800	800	2	100	20.5	0.990792667	[42]
Y ₂ Zr ₂ O ₇	9.8	800	800	2	100	62.6	0.89036	[42]
Y ₂ Zr ₂ O ₇	9.7	800	800	2	100	1.5	0.661074	[42]
Y ₂ Zr ₂ O ₇	9.8	800	800	2	200	20.5	0.63712775	[42]
Y ₂ Zr ₂ O ₇	9.8	800	800	2	90	62.6	0.545453667	[42]
Y ₂ Zr ₂ O ₇	9.7	800	800	2	90	1.5	0.294691333	[42]
Al ₂ O ₃ .C ₄ H ₈ O ₂	7.2	700	550	2.2	16	257	0.611779333	[43]
Al ₂ O ₃ .C ₄ H ₈ O ₂	7.6	700	550	2.2	16	276	0.625771333	[43]
Al ₂ O ₃ .C ₄ H ₈ O ₂	13	700	550	2.2	16	265	0.662265	[43]
Al ₂ O ₃ .C ₄ H ₈ O ₂	11.9	700	550	2.2	16	262	0.639041667	[43]
Al ₂ O ₃ .C ₄ H ₈ O ₂	8.8	700	550	2.2	16	274	0.634857667	[43]
Y ₂ Ti ₂ O ₇	7	800	600	2	0.5	21	0.509268	[44]
Y ₂ Ti ₂ O ₇	7	800	650	2	0.5	21	0.683814	[44]
Y ₂ Ti ₂ O ₇	7	800	700	2	0.5	21	0.826027	[44]
Y ₂ Ti ₂ O ₇	7	800	750	2	0.5	21	0.914368	[44]
Y ₂ Ti ₂ O ₇	7	800	800	2	0.5	21	0.96499	[44]
Y ₂ Sn ₂ O ₇	6.9	800	600	2	0.5	38	0.01718	[44]
Y ₂ Sn ₂ O ₇	6.9	800	650	2	0.5	38	0.039073	[44]
Y ₂ Sn ₂ O ₇	6.9	800	750	2	0.5	38	0.127747	[44]

Y ₂ Sn ₂ O ₇	6.9	800	800	2	0.5	38	0.165798	[44]
Y ₂ Zr ₂ O ₇	7	800	600	2	0.5	2	0.34404	[44]
Y ₂ Zr ₂ O ₇	7	800	650	2	0.5	2	0.500623	[44]
Y ₂ Zr ₂ O ₇	7	800	700	2	0.5	2	0.660801	[44]
Y ₂ Zr ₂ O ₇	7	800	750	2	0.5	2	0.786855	[44]
Y ₂ Zr ₂ O ₇	7	800	800	2	0.5	2	0.896742	[44]
Y ₂ Ce ₂ O ₇	7.1	800	600	2	0.5	4	0.387142	[44]
Y ₂ Ce ₂ O ₇	7.1	800	700	2	0.5	4	0.673375	[44]
Y ₂ Ce ₂ O ₇	7.1	800	750	2	0.5	4	0.797629	[44]
Y ₂ Ce ₂ O ₇	7.1	800	800	2	0.5	4	0.8734	[44]
Y ₂ Ti ₂ O ₇	6.9	800	750	2	50	21	0.814813	[44]
Y ₂ Ce ₂ O ₇	7.1	800	750	2	50	4	0.604116	[44]
Y ₂ Zr ₂ O ₇	7	800	750	2	50	2	0.401646	[44]
Y ₂ Sn ₂ O ₇	6.9	800	750	2	50	38	0.07572	[44]
La ₂ Ti ₂ O ₇	7.1	800	600	2	0.5	19.2	0.412934	[45]
Pr ₂ Ti ₂ O ₇	6.9	800	600	2	0.5	18.5	0.371955	[45]
Sm ₂ Ti ₂ O ₇	7.1	800	600	2	0.5	16.9	0.445095	[45]
Y ₂ Ti ₂ O ₇	7	800	600	2	0.5	15.8	0.450934	[45]
La ₂ Ti ₂ O ₇	7.1	800	650	2	0.5	19.2	0.559158	[45]
Pr ₂ Ti ₂ O ₇	6.9	800	650	2	0.5	18.5	0.544514	[45]
Y ₂ Ti ₂ O ₇	7	800	650	2	0.5	15.8	0.623494	[45]
La ₂ Ti ₂ O ₇	7.1	800	700	2	0.5	19.2	0.695141	[45]
Pr ₂ Ti ₂ O ₇	6.9	800	700	2	0.5	18.5	0.680497	[45]
Sm ₂ Ti ₂ O ₇	7.1	800	700	2	0.5	16.9	0.756562	[45]
Y ₂ Ti ₂ O ₇	7	800	700	2	0.5	15.8	0.822388	[45]

La ₂ Ti ₂ O ₇	7.1	800	750	2	0.5	19.2	0.804787	[45]
Sm ₂ Ti ₂ O ₇	7.1	800	750	2	0.5	16.9	0.876452	[45]
Y ₂ Ti ₂ O ₇	7	800	750	2	0.5	15.8	0.910089	[45]
La ₂ Ti ₂ O ₇	7.1	800	800	2	0.5	19.2	0.886637	[45]
Pr ₂ Ti ₂ O ₇	6.9	800	800	2	0.5	18.5	0.911496	[45]
Sm ₂ Ti ₂ O ₇	7.1	800	800	2	0.5	16.9	0.934893	[45]
Y ₂ Ti ₂ O ₇	7	800	800	2	0.5	15.8	0.964139	[45]
Pr ₂ Ti ₂ O ₇	6.9	800	800	2	50	18.5	0.805779	[45]
La ₂ Ti ₂ O ₇	7.1	800	800	2	50	19.2	0.749774	[45]
Sm ₂ Ti ₂ O ₇	7.1	800	800	2	50	16.9	0.644985	[45]
Y ₂ Ti ₂ O ₇	7	800	800	2	50	15.8	0.616078	[45]
Al ₂ O ₃	10	550	600	1.2	1	101	0.502171	[46]
Al ₂ O ₃	10	550	650	1.2	1	101	0.833715	[46]
Al ₂ O ₃	10	550	700	1.2	1	101	0.934697	[46]
Al ₂ O ₃	10	550	650	1.2	1	205	0.881292	[46]
Al ₂ O ₃	10	550	700	1.2	1	205	0.952997	[46]
Al ₂ O ₃	10	550	600	1.2	20	101	0.532417	[46]
Al ₂ O ₃	10	550	600	1.2	5	101	0.484269	[46]
Al ₂ O ₃	10	550	650	1.2	20	101	0.855784	[46]
Al ₂ O ₃	10	550	650	1.2	5	101	0.842772	[46]
Al ₂ O ₃	10	550	700	1.2	20	101	0.912283	[46]
Al ₂ O ₃	10	550	700	1.2	5	101	0.954138	[46]
Al ₂ O ₃	10	550	600	1.2	5	205	0.632041	[46]
Al ₂ O ₃	10	550	600	1.2	20	205	0.636707	[46]
Al ₂ O ₃	10	550	650	1.2	5	205	0.860525	[46]

Al ₂ O ₃	10	550	700	1.2	5	205	0.990004	[46]
Al ₂ O ₃	10	550	700	1.2	20	205	0.98705	[46]
Al ₂ O ₃	10	550	600	3.5	5	101	0.849106	[46]
Al ₂ O ₃	10	550	600	3.5	20	101	0.860623	[46]
Al ₂ O ₃	10	550	650	3.5	5	101	0.9405	[46]
Al ₂ O ₃	10	550	650	3.5	20	101	0.980956	[46]
Al ₂ O ₃	10	550	700	3.5	20	101	1	[46]
Al ₂ O ₃	10	550	600	6.1	5	101	0.966395	[46]
Al ₂ O ₃	10	550	600	6.1	20	101	0.980963	[46]
Al ₂ O ₃	10	550	650	6.1	5	101	0.996862	[46]
Al ₂ O ₃	10	550	650	6.1	20	101	1	[46]
Al ₂ O ₃	10	550	700	6.1	5	101	0.996862	[46]
Al ₂ O ₃	10	550	700	6.1	20	101	1	[46]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	1	7.4117	0.300809	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	10	7.4117	0.302993	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	1	8.8919	0.637843	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	10	8.8919	0.630293	[47]
Al ₂ O ₃	10	750	700	1	1	4.067	0.685295	[47]
Al ₂ O ₃	10	750	700	1	10	4.067	0.641244	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	1	10.7274	0.684079	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	1	10	10.7274	0.669228	[47]
Al ₂ O ₃	10	750	700	0.33	1	4.067	0.154028	[47]
Al ₂ O ₃	10	750	700	0.33	10	4.067	0.126861	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	0.33	1	10.7274	0.183494	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	0.33	10	10.7274	0.188356	[47]

CaZrO ₃ .Al ₂ O ₃	10	750	700	3	1	10.7274	0.661359	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	3	10	10.7274	0.667502	[47]
Al ₂ O ₃	10	750	700	3	1	4.067	0.813815	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	0.33	9	8.8919	0.207896	[47]
CaZrO ₃ .Al ₂ O ₃	10	750	700	0.33	97	8.8919	0.216223	[47]
Al ₂ O ₃	10	750	700	0.33	8	4.067	0.141751	[47]
Al ₂ O ₃	10	750	700	0.33	96	4.067	0.024964	[47]
MgAl ₂ O ₄	17.9	900	750	1.24	1	64	0.027388	[48]
MgAl ₂ O ₄	17.9	900	800	1.24	1	64	0.274466	[48]
MgAl ₂ O ₄	17.9	900	850	1.24	1	64	0.388216	[48]
MgAl ₂ O ₄	12.3	900	750	1.24	1	31	0.094834	[48]
MgAl ₂ O ₄	12.3	900	800	1.24	1	31	0.337217	[48]
MgAl ₂ O ₄	12.3	900	850	1.24	1	31	0.599987	[48]
MgAl ₂ O ₄	13.4	900	750	1.24	1	71	0.279938	[48]
MgAl ₂ O ₄	13.4	900	800	1.24	1	71	0.437609	[48]
MgAl ₂ O ₄	13.4	900	850	1.24	1	71	0.749001	[48]
CeO ₂ -ZrO ₂	10	500	500	1.94	5	63.6	0.099	[49]
CeO ₂ -ZrO ₂	10	500	600	1.94	5	63.6	0.292	[49]
CeO ₂ -ZrO ₂	10	500	700	1.94	5	63.6	0.458	[49]
CeO ₂ -ZrO ₂	10	500	800	1.94	5	63.6	0.666	[49]
Al ₂ O ₃	10	500	700	1.94	5	123	0.338	[49]
Al ₂ O ₃	10	500	800	1.94	5	123	0.711	[49]
LaAlO ₃	10	500	800	2	0.5	15.9	0.918695	[50]
LaAlO ₃	10	500	800	2	1	15.9	0.914169	[50]
LaAlO ₃	10	500	800	2	5	15.9	0.918412	[50]

SrTiO ₃	10	500	800	2	0.5	13.4	0.878246	[50]
SrTiO ₃	10	500	800	2	5	13.4	0.88545	[50]
LaFeO ₃	10	500	800	2	1	15.9	0.370349	[50]
LaFeO ₃	10	500	800	2	5	15.9	0.371596	[50]
BaTiO ₃	10	500	800	2	0.5	15	0.06027	[50]
BaTiO ₃	10	500	800	2	1	15	0.081212	[50]
BaTiO ₃	10	500	800	2	5	15	0.016542	[50]
LaAlO ₃	10	500	800	1	17	15.9	0.833139	[50]
LaAlO ₃	10	500	800	1	23	15.9	0.803568	[50]
RuAl ₂ O ₃	20	900	650	1.2	0.5	1.05	0.253151	[51]
Al ₂ O ₃	20	900	650	1.2	0.5	0.32	0.224613	[51]
RuAl ₂ O ₃	20	900	650	1.2	9.5	1.05	0.280361	[51]
Al ₂ O ₃	20	900	650	1.2	9.5	0.32	0.20901	[51]

Filters of descriptors

- Eliminate descriptors with less than 2 unique values
- Delete descriptors that have absolute values larger than 999999
- Delete descriptors that have variance smaller than 0.01
- Group descriptors, that have pair-wise correlations Pearson's correlation coefficient R larger than 0.95

Table_Apx B-6 Architecture of the ASNN model

Setting	Value
Training method	SuperSAB
Number of neurons in hidden layer	3
Learning iterations	1000
Ensemble	64
Additional Parameters	PARTITION=3, SELECTION=2

Table_Apx B-7 Architecture of the DNN model

Setting	Value
Epochs	1024
Batch size	10000
Training/internal validation set ratio	0.8
Model	Dense7 (seven layers) Net
Optimisation method	Adam
Activation function	relu

Table_Apx B-8 Architecture of the LSSVM model

Setting	Value
Kernel	RBF Kernel
Internal CV (cross validation) folds	5

Table_Apx B-9 CSCM database

N	Molecule	Nickel wt. %	Calcination Temperature Catalyst (°C)	Reaction Temperature (°C)	S/C Ratio	Time (h)	BET Surface Area (m ² /g)	Methane Conversion (%)	CaO Concentration (%)	Cycle Number	Calcium Precursor	Method	Calcination Temperature (°C)	Calcination Time (mins)	Carbonation Temperature (°C)	Carbonation Time (mins)	Initial Cycle Capacity (gCO ₂ /gSorbent)	Last Cycle CO ₂ Capacity (gCO ₂ /gSorbent)	Reference
1	Ca ₅ Al ₆ O ₁₄	15	850	650	2	2		0.95	61.7	50	10	2	800	10	650	30	0.484	0.449	[52]
2	CaZrO ₃	15	850	650	3	1	12.8	0.83	60.3	50	6	2	800	10	650	20	0.399	0.355	[53]
3	Ca ₁₂ Al ₁₄ O ₃₃	10	900	600	3	1	7.1	0.93	30	200	2	4	900	30	600	30	0.206	0.192	[54]
4	Ca ₁₂ Al ₁₄ O ₃₃	5.2	800	650	4	0.67	10.5	0.97	18	200	3	4	925	15	650	20	0.116	0.028	[55]
5	Ca ₅ Al ₆ O ₁₄	15	850	650	3	0.83	17.5	0.94	18	40	10	2	800	45	650	30	0.130	0.115	[56]
6	Ca ₁₂ Al ₁₄ O ₃₃	3	900	650	3	3	12.7	0.10	54	50	4	4	900	60	650	60	0.195	0.063	[57]
7	Ca ₁₂ Al ₁₄ O ₃₃	3	900	650	3	3	9.2	0.12	30	50	4	4	900	60	650	60	0.136	0.127	[57]
8	Ca ₁₂ Al ₁₄ O ₃₃	3.6	750	650	3	50	11.5	0.99	75	30	6	3	800	10	650	30	0.450	0.338	[58]
9	Ca ₁₂ Al ₁₄ O ₃₃	3.6	750	650	3	50	12.6	0.80	75	30	1	4	800	10	650	30	0.401	0.292	[58]
10	Ca ₁₂ Al ₁₄ O ₃₃	13.2	900	650	3	0.5	12.44	0.96	10	200	2	8	850	10	650	10	0.104	0.099	[59]
11	Ca ₁₂ Al ₁₄ O ₃₃	20	900	650	3.4	1	9.5	0.36	70	44	4	1	850	5	690	30	0.303	0.264	[60]
12	Ca ₁₂ Al ₁₄ O ₃₃	16	900	650	3.4	1	8.8	0.81	70	44	4	1	850	5	690	30	0.292	0.307	[60]
13	Ca ₁₂ Al ₁₄ O ₃₃	11	900	650	3.4	1	10	0.59	70	45	4	1	850	5	690	30	0.317	0.271	[60]
14	Ca ₁₂ Al ₁₄ O ₃₃	8	900	650	3.4	1	11	0.22	70	44	4	1	850	5	690	30	0.378	0.281	[60]
15	CaZrO ₃	3.4	900	600	3	3	17.7	0.70	95.4	10	2	4	850	30	600	30	0.308	0.287	[61]
16	CeO ₂	2.5	900	600	3	3	8.2	0.61	96.3	10	2	4	850	30	600	30	0.226	0.236	[61]
17	La ₂ O ₃	2.8	900	600	3	3	13.1	0.46	92.6	10	2	4	850	30	600	30	0.295	0.247	[61]
18	Al ₂ O ₃	3	900	650	3	1	10.1	0.96	29	50		4	800	30	650	30	0.099	0.087	[62]

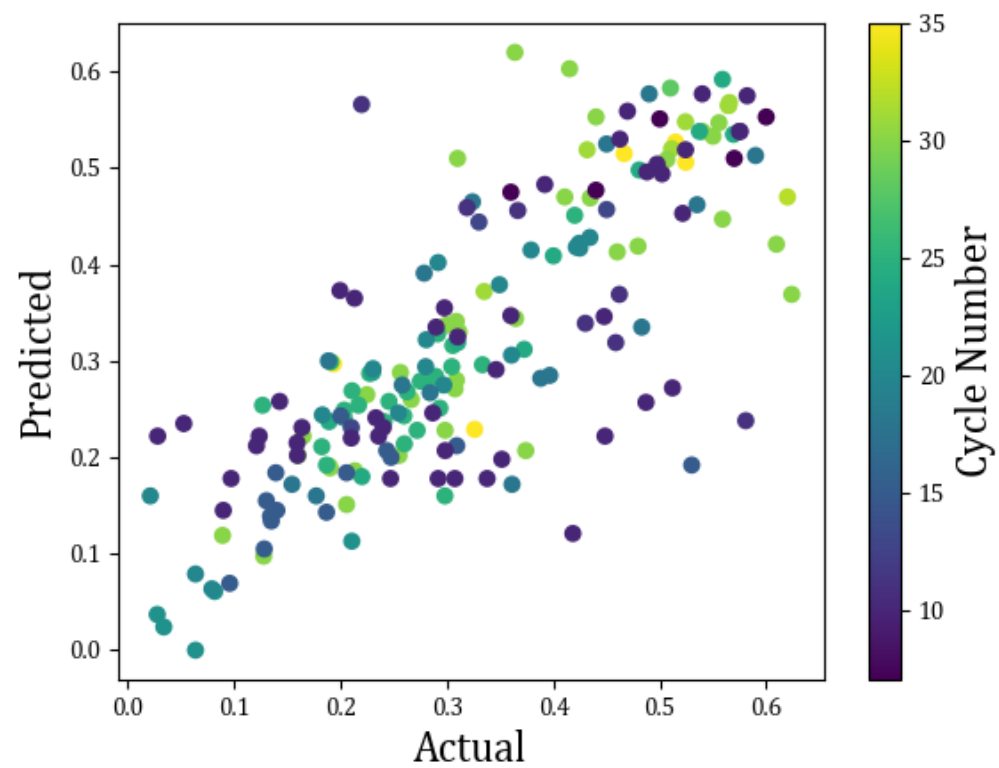
19	Al ₂ O ₃	3	900	650	3	1	9.2	0.96	29	50		4	800	30	650	30	0.135	0.123	[62]
20	Al ₂ O ₃	3	900	650	3	1	14.9	0.96	52	50		4	800	30	650	30	0.257	0.097	[62]
21	Al ₂ O ₃	3	900	650	3	1	12.7	0.96	52	50		4	800	30	650	30	0.194	0.061	[62]
22	Al ₂ O ₃	15	900	650	4	2	9.9		56	10	10	2	800	10	650	30	0.218	0.233	[63]
23	Al ₂ O ₃	15	900	650	4	2	11.5		68	10	10	2	800	10	650	30	0.396	0.397	[63]
24	Al ₂ O ₃	15	900	650	4	2	12.1		72.3	10	10	2	800	10	650	30	0.457	0.456	[63]
25	Al ₂ O ₃	15	900	650	4	2	12.3		75	10	10	2	800	10	650	30	0.492	0.487	[63]
26	Al ₂ O ₃	15	900	650	4	2	10.9		77.2	10	10	2	800	10	650	30	0.510	0.512	[63]
27	Al ₂ O ₃	15	900	650	4	2	9.6		78.5	10	10	2	800	10	650	30	0.525	0.526	[63]
28	Al ₂ O ₃	15	900	650	2	2	12.1		72.3	50	10	2	800	10	650	30	0.395	0.376	[63]
29	CaZrO ₃	18	800	650	4	2	15.9	0.97	34.3	10	4	4	800	20	650	30			[64]
30	CaZrO ₃	20.5	800	650	4	2	9.6	0.99	33.2	10	4	4	800	20	650	30			[64]
31	CaZrO ₃	13	800	650	4	2	8.3	0.89	36.4	2	4	4	800	20	650	30			[64]
32	Ca ₉ Al ₆ O ₁₈	25	900	650	4	2	10	0.99	58.4	10	1	1	800	20	650	30			[65]
33	Ca ₉ Al ₆ O ₁₈	18	900	650	4	2	12.9	0.96	63.9	10	1	1	800	20	650	30			[65]
34	La ₂ O ₃ Al ₂ O ₃		550	600	4	5	29.18	0.78		30	1	8	725	10	600	10	0.136	0.053	[66]
35	La ₂ O ₃ Al ₂ O ₃		550	600	4	5	33.37			30	1	8	725	10	600	10	0.091	0.037	[66]
36	MgAl ₂ O ₄ Ca ₁₂ Al ₁₄ O ₃₃	5.3	800	650	3	3	14.62		21	20	3	1	925	35	650	35	0.164	0.021	[67]
37	MgAl ₂ O ₄ Ca ₁₂ Al ₁₄ O ₃₃	5.3	800	650	3	3	14.62		21	90	3	1	925	35	650	35	0.211	0.133	[67]
38	Ca ₁₂ Al ₁₄ O ₃₃	3	900	640	3	1.5	13	0.85	30	4	2	4	700	40	640	16			[68]
39	PdCH ₂₄ Al ₂ Mg ₆ O ₂₃	9.68	800	650	4	2	29.14	0.755	65	1		6	700	20	700	20			[69]
40	RhCH ₂₄ Al ₂ Mg ₆ O ₂₃	9.68	800	650	4	2		0.266	65	1		6	700	20	700	20			[69]
41	RuCH ₂₄ Al ₂ Mg ₆ O ₂₃	9.68	800	650	4	2		0.957	65	10		6	700	20	700	20			[69]

Data in black, is input data added to the overall combined sorbent (239) and catalyst (183) database used to predict unseen molecules (24) (446 datapoints overall for MTL). Data in green is the unseen molecules (17).

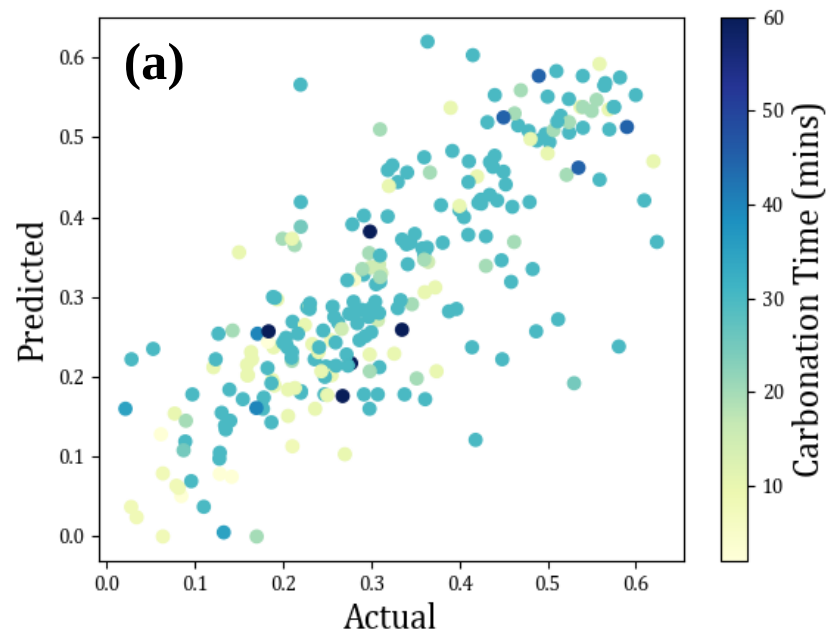
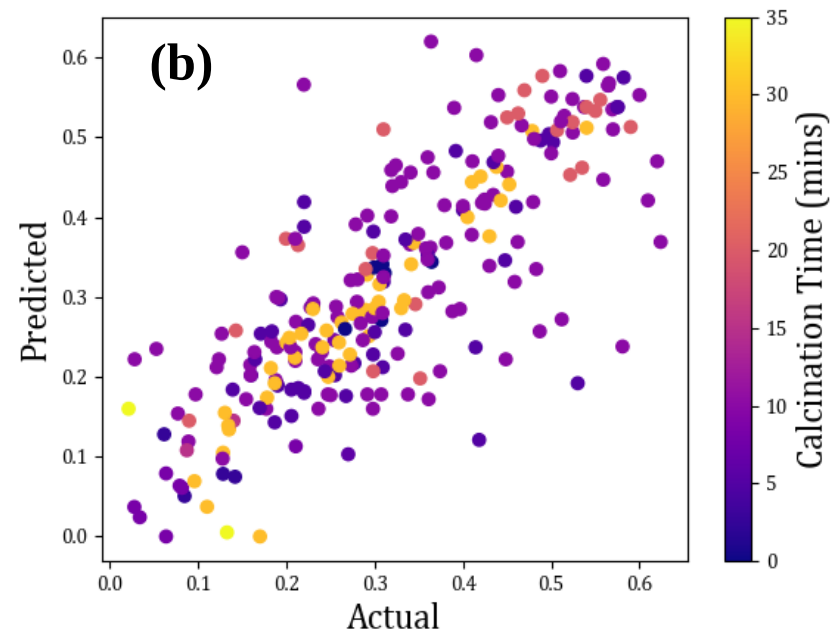
Table_Apx B-10 Measured vs Predicted CSCM data using GSfrag/ASNN model, with ARE%

N	Methane Conversion (%) {measured}	Methane Conversion (%) {predicted}	ARE (%)	Last cycle CO ₂ capacity (gCO ₂ /gSorbent) {measured}	Last cycle CO ₂ capacity (gCO ₂ /gSorbent) {predicted}	ARE (%)
5	0.9352	0.918	1.84			
1				0.1146	0.184	60.56
21	0.96	0.959	0.10			
2				0.060649	0.184	203.39
23				0.3967	0.454	14.44
29	0.972245	0.997	2.55			
38				0.037431	-0.0504	234.65
24				0.456	0.474	3.95
43	0.2305	0.965	318.66			
22				0.2332	0.385	65.09
27				0.5257	0.525	0.13
46	0.9956	0.755	24.17			
35	0.937	0.797	14.94			
11				0.09284	0.278	199.44
10	0.964	0.73	24.27			
12				0.099	0.184	85.86
3	0.930637	0.495	46.81			
13				0.1921	0.184	4.22
11	0.358614	0.802	123.64			
14				0.263876	0.184	30.27
14	0.224135	0.819	265.40			
15				0.280825	0.184	34.48

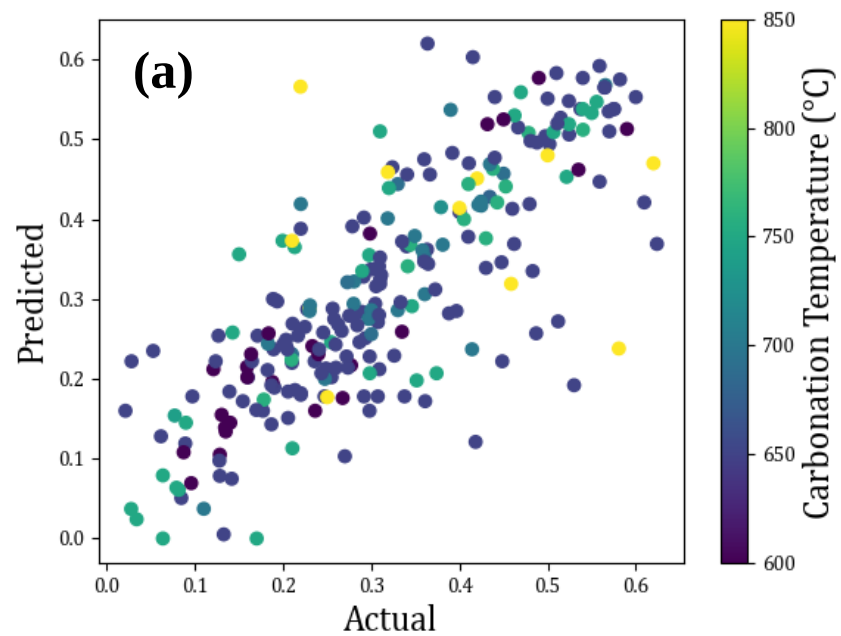
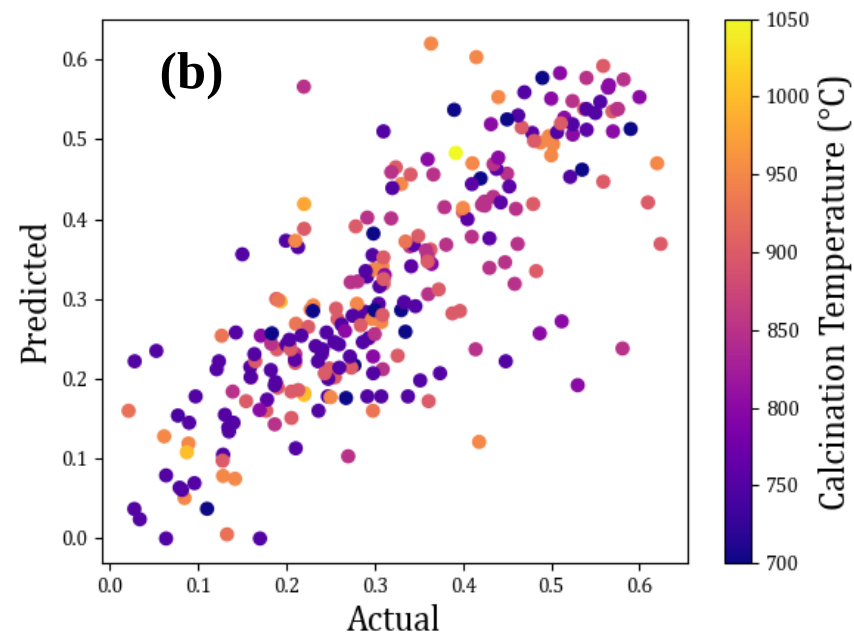
31	0.88621	0.996	12.39			
7	0.119	0.63	429.41			
17				0.126654	0.184	45.28
2	0.825943	0.849	2.79			
18				0.3554	0.287	19.25
18	0.96	0.96	0.00			
19				0.086673	0.184	112.29
44	0.2663	0.965	262.37			
28				0.3761	0.43	14.33
6	0.101	0.535	429.70			
22				0.063326	0.184	190.56
45	0.3513	0.965	174.69			
13	0.588014	0.815	38.60			
24				0.271215	0.184	32.16



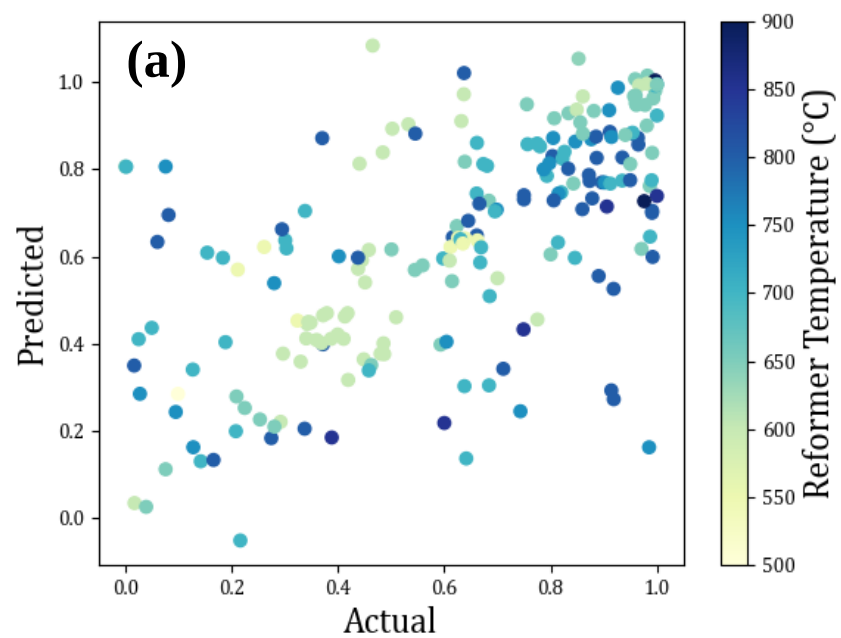
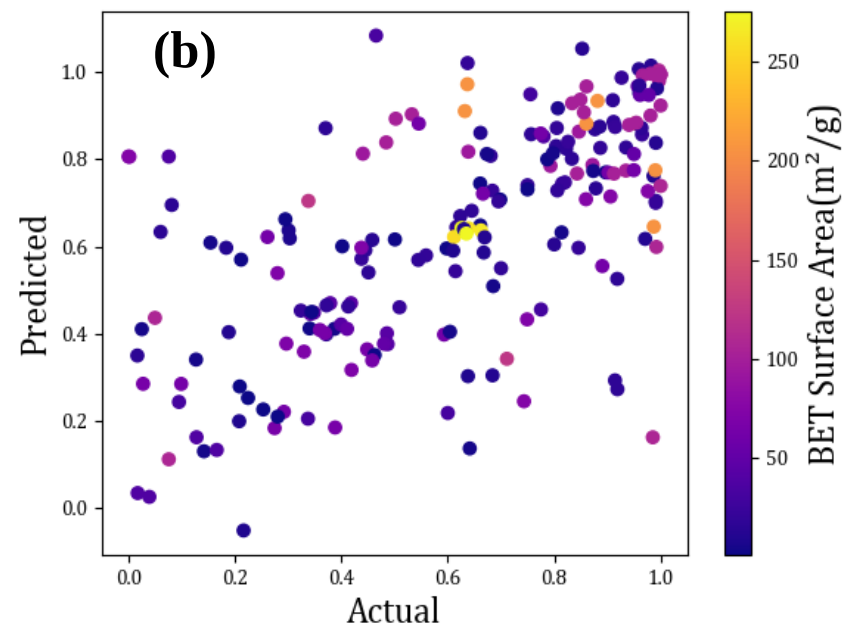
Figure_Apx B-1 Actual vs Predicted scatter plot showing the effect of the cycle number on the Last cycle capacity (g/g)



Figure_Apx B-2 Actual vs Predicted scatter plot showing the effect of the (a) carbonation time and (b) calcination time on the Last cycle capacity (g/g)



Figure_Apx B-3 Actual vs Predicted scatter plot showing the effect of the (a) carbonation temperature and (b) calcination temperature on the Last cycle capacity (g/g)



Figure_Apx B-4 Actual vs Predicted scatter plot showing the effect of the (a) SMR reformer temperature ($^{\circ}\text{C}$) and (b) BET Surface Area (m^2/g) on the methane conversion (%)

B.1.1 Database References

- [1] Jaworska J., Nikolova-Jeliazkova N., Aldenberg T. QSAR applicability domain estimation by projection of the training set in descriptor space: A review. *ATLA Alternatives to Laboratory Animals*. 2005; 33(5): 445–459. Available at: DOI:10.1177/026119290503300508
- [2] Broda M., Kierzkowska AM., Müller CR. Application of the sol-gel technique to develop synthetic calcium-based sorbents with excellent carbon dioxide capture characteristics. *ChemSusChem*. 2012; 5(2): 411–418. Available at: DOI:10.1002/cssc.201100468
- [3] Broda M., Müller CR. Synthesis of highly efficient, Ca-based, Al₂O₃-stabilized, carbon gel-templated CO₂ sorbents. *Advanced Materials*. 2012; 24(22): 3059–3064. Available at: DOI:10.1002/adma.201104787
- [4] Zhou Z., Qi Y., Xie M., Cheng Z., Yuan W. Synthesis of CaO-based sorbents through incorporation of alumina/aluminate and their CO₂ capture performance. *Chemical Engineering Science*. May 2012; 74: 172–180. Available at: DOI:10.1016/j.ces.2012.02.042
- [5] Koirala R., Reddy GK., Smirniotis PG. Single nozzle flame-made highly durable metal doped Ca-based sorbents for CO₂ capture at high temperature. *Energy and Fuels*. 2012; 26(5): 3103–3109. Available at: DOI:10.1021/ef3004015
- [6] Wang K., Guo X., Zhao P., Zheng C. Cyclic CO₂ capture of CaO-based sorbent in the presence of metakaolin and aluminum (hydr)oxides. *Applied Clay Science*. September 2010; 50(1): 41–46. Available at: DOI:10.1016/j.clay.2010.06.021
- [7] Manovic V., Anthony EJ. CaO-based pellets supported by calcium aluminate cements for high-temperature CO₂ capture. *Environmental Science & Technology*. 2009; 43(18): 7117–7122. Available at: DOI:10.1021/es901258w

- [8] Li Z., Cai N., Huang Y., Han H. Synthesis, Experimental Studies, and Analysis of a New Calcium-Based Carbon Dioxide Absorbent. *Energy & Fuels*. 1 July 2005; 19(4): 1447–1452. Available at: DOI:10.1021/ef0496799
- [9] Wu SF., Li QH., Kim JN., Yi KB. Properties of a Nano CaO/Al₂O₃ CO₂ Sorbent. *Industrial & Engineering Chemistry Research*. 1 January 2008; 47(1): 180–184. Available at: DOI:10.1021/ie0704748
- [10] Broda M., Kierzkowska A., Müller C. Development of synthetic, Ca-based sorbents for CO₂ capture using the co-precipitation technique. *Materials Science*. 2011.
- [11] Luo C., Zheng Y., Ding N., Zheng C. Enhanced cyclic stability of CO₂ adsorption capacity of CaO-based sorbents using La₂O₃ or Ca₁₂Al₁₄O₃₃ as additives. *Korean Journal of Chemical Engineering*. 2011; 28(4): 1042–1046. Available at: DOI:10.1007/s11814-010-0469-z
- [12] Pacciani R., Müller CR., Davidson JF., Dennis JS., Hayhurst AN. Synthetic Ca-based solid sorbents suitable for capturing CO₂ in a fluidized bed. *The Canadian Journal of Chemical Engineering*. June 2008; 86(3): 356–366. Available at: DOI:10.1002/cjce.20060
- [13] Stendardo S., Andersen LK., Hecce C. Self-activation and effect of regeneration conditions in CO₂–carbonate looping with CaO–Ca₁₂Al₁₄O₃₃ sorbent. *Chemical Engineering Journal*. March 2013; 220: 383–394. Available at: DOI:10.1016/j.cej.2013.01.045
- [14] Martavaltzi CS., Lemonidou AA. Parametric study of the CaO–Ca₁₂Al₁₄O₃₃ synthesis with respect to high CO₂ sorption capacity and stability on multicycle operation. *Industrial and Engineering Chemistry Research*. 2008; 47(23): 9537–9543. Available at: DOI:10.1021/ie800882d
- [15] Martavaltzi CS., Lemonidou AA. Development of new CaO based sorbent materials for CO₂ removal at high temperature. Microporous and Mesoporous

Materials. April 2008; 110(1): 119–127. Available at:
DOI:10.1016/j.micromeso.2007.10.006

[16] Wu SF., Zhu YQ. Behavior of CaTiO₃ /Nano-CaO as a CO₂ Reactive Adsorbent. Industrial & Engineering Chemistry Research. 17 March 2010; 49(6): 2701–2706. Available at: DOI:10.1021/ie900900r

[17] Zhao M., Bilton M., Brown AP., Cunliffe AM., Dvinnikov E., Dupont V., et al. Durability of CaO–CaZrO₃ Sorbents for High-Temperature CO₂ Capture Prepared by a Wet Chemical Method. Energy & Fuels. 20 February 2014; 28(2): 1275–1283. Available at: DOI:10.1021/ef4020845

[18] Wang S., Fan S., Fan L., Zhao Y., Ma X. Effect of Cerium Oxide Doping on the Performance of CaO-Based Sorbents during Calcium Looping Cycles. Environmental Science & Technology. 21 April 2015; 49(8): 5021–5027. Available at: DOI:10.1021/es5052843

[19] Li L., King DL., Nie Z., Howard C. Magnesia-Stabilized Calcium Oxide Absorbents with Improved Durability for High Temperature CO₂ Capture. Industrial & Engineering Chemistry Research. 2 December 2009; 48(23): 10604–10613. Available at: DOI:10.1021/ie901166b

[20] Lan P., Wu S. Synthesis of a Porous Nano-CaO/MgO-Based CO₂ Adsorbent. Chemical Engineering & Technology. April 2014; 37(4): 580–586. Available at: DOI:10.1002/ceat.201300709

[21] Albrecht KO., Wagenbach KS., Satrio JA., Shanks BH., Wheelock TD. Development of a CaO-Based CO₂ Sorbent with Improved Cyclic Stability. Industrial & Engineering Chemistry Research. 15 October 2008; 47(20): 7841–7848. Available at: DOI:10.1021/ie8007743

[22] Zhao M., Yang X., Church TL., Harris AT. Novel CaO–SiO₂ Sorbent and Bifunctional Ni/Co–CaO/SiO₂ Complex for Selective H₂ Synthesis from

Cellulose. Environmental Science & Technology. 6 March 2012; 46(5): 2976–2983. Available at: DOI:10.1021/es300135d

[23] Derevschikov VS., Lysikov AI., Okunev AG. High Temperature CaO/Y₂O₃ Carbon Dioxide Absorbent with Enhanced Stability for Sorption-Enhanced Reforming Applications. Industrial & Engineering Chemistry Research. 16 November 2011; 50(22): 12741–12749. Available at: DOI:10.1021/ie2015334

[24] Zhang X., Li Z., Peng Y., Su W., Sun X., Li J. Investigation on a novel CaO–Y₂O₃ sorbent for efficient CO₂ mitigation. Chemical Engineering Journal. May 2014; 243: 297–304. Available at: DOI:10.1016/j.cej.2014.01.017

[25] Hu Y., Liu W., Chen H., Zhou Z., Wang W., Sun J., et al. Screening of inert solid supports for CaO-based sorbents for high temperature CO₂ capture. Fuel. Elsevier Ltd; 2016; 181: 199–206. Available at: DOI:10.1016/j.fuel.2016.04.138

[26] Radfarnia HR., Iliuta MC. Metal oxide-stabilized calcium oxide CO₂ sorbent for multicycle operation. Chemical Engineering Journal. October 2013; 232: 280–289. Available at: DOI:10.1016/j.cej.2013.07.049

[27] Lu H., Khan A., Pratsinis SE., Smirniotis PG. Flame-Made Durable Doped-CaO Nanosorbents for CO₂ Capture. Energy & Fuels. 19 February 2009; 23(2): 1093–1100. Available at: DOI:10.1021/ef8007882

[28] Radfarnia HR., Iliuta MC. Development of Zirconium-Stabilized Calcium Oxide Absorbent for Cyclic High-Temperature CO₂ Capture. Industrial & Engineering Chemistry Research. 8 August 2012; 51(31): 10390–10398. Available at: DOI:10.1021/ie301287k

[29] Li L., King DL., Nie Z., Li XS., Howard C. MgAl₂O₄ Spinel-Stabilized Calcium Oxide Absorbents with Improved Durability for High-Temperature CO₂ Capture. Energy & Fuels. 17 June 2010; 24(6): 3698–3703. Available at: DOI:10.1021/ef100245q

- [30] Zhang M., Peng Y., Sun Y., Li P., Yu J. Preparation of CaO-Al₂O₃ sorbent and CO₂ capture performance at high temperature. *Fuel*. Elsevier Ltd; 2013; 111: 636–642. Available at: DOI:10.1016/j.fuel.2013.03.078
- [31] Li Y., Su M., Xie X., Wu S., Liu C. CO₂ capture performance of synthetic sorbent prepared from carbide slag and aluminum nitrate hydrate by combustion synthesis. *Applied Energy*. May 2015; 145: 60–68. Available at: DOI:10.1016/j.apenergy.2015.01.061
- [32] Liu FQ., Li WH., Liu BC., Li RX. Synthesis, characterization, and high temperature CO₂ capture of new CaO based hollow sphere sorbents. *Journal of Materials Chemistry A*. 2013; 1(27): 8037–8044. Available at: DOI:10.1039/c3ta11369h
- [33] Xu P., Xie M., Cheng Z., Zhou Z. CO₂ capture performance of CaO-based sorbents prepared by a sol-gel method. *Industrial and Engineering Chemistry Research*. 2013; 52(34): 12161–12169. Available at: DOI:10.1021/ie401600e
- [34] Liu W., Feng B., Wu Y., Wang G., Barry J., Diniz da Costa JC. Synthesis of Sintering-Resistant Sorbents for CO₂ Capture. *Environmental Science & Technology*. 15 April 2010; 44(8): 3093–3097. Available at: DOI:10.1021/es903436v
- [35] Radfarnia HR., Sayari A. A highly efficient CaO-based CO₂ sorbent prepared by a citrate-assisted sol–gel technique. *Chemical Engineering Journal*. February 2015; 262: 913–920. Available at: DOI:10.1016/j.cej.2014.09.074
- [36] Florin NH., Blamey J., Fennell PS. Synthetic CaO-based sorbent for CO₂ capture from large-point sources. *Energy and Fuels*. 2010; 24(8): 4598–4604. Available at: DOI:10.1021/ef100447c
- [37] Qin C., Liu W., An H., Yin J., Feng B. Fabrication of CaO-Based Sorbents for CO₂ Capture by a Mixing Method. *Environmental Science & Technology*. 7 February 2012; 46(3): 1932–1939. Available at: DOI:10.1021/es203525y

- [38] Ali S., Al-Marri MJ., Abdelmoneim AG., Kumar A., Khader MM. Catalytic evaluation of nickel nanoparticles in methane steam reforming. *International Journal of Hydrogen Energy*. December 2016; 41(48): 22876–22885. Available at: DOI:10.1016/j.ijhydene.2016.08.200
- [39] Katheria S., Gupta A., Deo G., Kunzru D. Effect of calcination temperature on stability and activity of Ni/MgAl₂O₄ catalyst for steam reforming of methane at high pressure condition. *International Journal of Hydrogen Energy*. Elsevier Ltd; 24 August 2016; 41(32): 14123–14132. Available at: DOI:10.1016/j.ijhydene.2016.05.109
- [40] Rogers JL., Mangarella MC., D'Amico AD., Gallagher JR., Dutzer MR., Stavitski E., et al. Differences in the Nature of Active Sites for Methane Dry Reforming and Methane Steam Reforming over Nickel Aluminate Catalysts. *ACS Catalysis*. American Chemical Society; 2 September 2016; 6(9): 5873–5886. Available at: DOI:10.1021/acscatal.6b01133
- [41] Khani Y., Shariatnia Z., Bahadoran F. High catalytic activity and stability of ZnLaAlO₄ supported Ni, Pt and Ru nanocatalysts applied in the dry, steam and combined dry-steam reforming of methane. *Chemical Engineering Journal*. Elsevier B.V.; 1 September 2016; 299: 353–366. Available at: DOI:10.1016/j.cej.2016.04.108
- [42] Fang X., Zhang X., Guo Y., Chen M., Liu W., Xu X., et al. Highly active and stable Ni/Y₂Zr₂O₇ catalysts for methane steam reforming: On the nature and effective preparation method of the pyrochlore support. *International Journal of Hydrogen Energy*. Elsevier Ltd; 13 July 2016; 41(26): 11141–11153. Available at: DOI:10.1016/j.ijhydene.2016.04.038
- [43] Yoo J., Park S., Song JH., Yoo S., Song IK. Hydrogen production by steam reforming of natural gas over butyric acid-assisted nickel/alumina catalyst. *International Journal of Hydrogen Energy*. Elsevier Ltd; 23 November 2017; 42(47): 28377–28385. Available at: DOI:10.1016/j.ijhydene.2017.09.148

- [44] Zhang X., Peng L., Fang X., Cheng Q., Liu W., Peng H., et al. Ni/Y₂B₂O₇ (B=Ti, Sn, Zr and Ce) catalysts for methane steam reforming: On the effects of B site replacement. *International Journal of Hydrogen Energy*. Elsevier Ltd; 26 April 2018; 43(17): 8298–8312. Available at: DOI:10.1016/j.ijhydene.2018.03.086
- [45] Fang X., Xu L., Zhang X., Zhang K., Dai H., Liu W., et al. Effect of rare earth element (Ln = La, Pr, Sm, and Y) on physicochemical properties of the Ni/Ln₂Ti₂O₇ catalysts for the steam reforming of methane. *Molecular Catalysis*. May 2019; 468: 130–138. Available at: DOI:10.1016/j.mcat.2019.02.022
- [46] Dan M., Mihet M., Lazar MD. Hydrogen and/or syngas production by combined steam and dry reforming of methane on nickel catalysts. *International Journal of Hydrogen Energy*. Elsevier Ltd; 2 October 2020; 45(49): 26254–26264. Available at: DOI:10.1016/j.ijhydene.2019.12.158
- [47] Lertwittayanon K., Youravong W., Lau WJ. Enhanced catalytic performance of Ni/ α -Al₂O₃ catalyst modified with CaZrO₃ nanoparticles in steam-methane reforming. *International Journal of Hydrogen Energy*. November 2017; 42(47): 28254–28265. Available at: DOI:10.1016/j.ijhydene.2017.09.030
- [48] Azancot L., Bobadilla LF., Santos JL., Córdoba JM., Centeno MA., Odriozola JA. Influence of the preparation method in the metal-support interaction and reducibility of Ni-Mg-Al based catalysts for methane steam reforming. *International Journal of Hydrogen Energy*. Elsevier Ltd; 26 July 2019; 44(36): 19827–19840. Available at: DOI:10.1016/j.ijhydene.2019.05.167
- [49] Kusakabe K., Sotowa KI., Eda T., Iwamoto Y. Methane steam reforming over Ce-ZrO₂-supported noble metal catalysts at low temperature. *Fuel Processing Technology*. 15 December 2004; 86(3): 319–326. Available at: DOI:10.1016/j.fuproc.2004.05.003
- [50] Urasaki K., Sekine Y., Kawabe S., Kikuchi E., Matsukata M. Catalytic activities and coking resistance of Ni/perovskites in steam reforming of methane.

Applied Catalysis A: General. May 2005; 286(1): 23–29. Available at: DOI:10.1016/j.apcata.2005.02.020

[51] Jeong JH., Lee JW., Seo DJ., Seo Y., Yoon WL., Lee DK., et al. Ru-doped Ni catalysts effective for the steam reforming of methane without the pre-reduction treatment with H₂. Applied Catalysis A: General. April 2006; 302(2): 151–156. Available at: DOI:10.1016/j.apcata.2005.12.007

[52] Xu P., Zhou Z., Zhao C., Cheng Z. Catalytic performance of Ni/CaO-Ca₅Al₆O₁₄ bifunctional catalyst extrudate in sorption-enhanced steam methane reforming. Catalysis Today. Elsevier; 3 November 2016; 259: 347–353. Available at: DOI:10.1016/j.cattod.2015.05.026

[53] Zhao C., Zhou Z., Cheng Z., Fang X. Sol-gel-derived, CaZrO₃-stabilized Ni/CaO-CaZrO₃ bifunctional catalyst for sorption-enhanced steam methane reforming. Applied Catalysis B: Environmental. Elsevier B.V.; 5 November 2016; 196: 16–26. Available at: DOI:10.1016/j.apcatb.2016.05.021

[54] Vanga G., Gattia DM., Stendardo S., Scaccia S. Novel synthesis of combined CaO-Ca₁₂Al₁₄O₃₃-Ni sorbent-catalyst material for sorption enhanced steam reforming processes. Ceramics International. Elsevier Ltd; 15 April 2019; 45(6): 7594–7605. Available at: DOI:10.1016/j.ceramint.2019.01.054

[55] di Giuliano A., Gallucci K., Kazi SS., Giancaterino F., di Carlo A., Courson C., et al. Development of Ni- and CaO-based mono- and bi-functional catalyst and sorbent materials for Sorption Enhanced Steam Methane Reforming: Performance over 200 cycles and attrition tests. Fuel Processing Technology. Elsevier B.V.; 1 December 2019; 195. Available at: DOI:10.1016/j.fuproc.2019.106160

[56] Chen X., Yang L., Zhou Z., Cheng Z. Core-shell structured CaO-Ca₉Al₆O₁₈@Ca₅Al₆O₁₄/Ni bifunctional material for sorption-enhanced steam methane reforming. Chemical Engineering Science. Elsevier Ltd; 2017; 163: 114–122. Available at: DOI:10.1016/j.ces.2017.01.036

- [57] di Giuliano A., Girr J., Massacesi R., Gallucci K., Courson C. Sorption enhanced steam methane reforming by Ni–CaO materials supported on mayenite. *International Journal of Hydrogen Energy*. Elsevier Ltd; 11 May 2017; 42(19): 13661–13680. Available at: DOI:10.1016/j.ijhydene.2016.11.198
- [58] Cesário MR., Barros BS., Courson C., Melo DMA., Kiennemann A. Catalytic performances of Ni-CaO-mayenite in CO₂ sorption enhanced steam methane reforming. *Fuel Processing Technology*. Elsevier; 2015; 131: 247–253. Available at: DOI:10.1016/j.fuproc.2014.11.028
- [59] di Giuliano A., Gallucci K., di Carlo A., Stendardo S., Courson C., Foscolo PU. Sorption enhanced steam methane reforming by Ni/CaO/mayenite combined systems: Overview of experimental results from European research project ASCENT. *Canadian Journal of Chemical Engineering*. Wiley-Liss Inc.; 1 September 2020; 98(9): 1907–1923. Available at: DOI:10.1002/cjce.23779
- [60] Martavaltzi CS., Lemonidou AA. Hydrogen production via sorption enhanced reforming of methane: Development of a novel hybrid material-reforming catalyst and CO₂ sorbent. *Chemical Engineering Science*. 15 July 2010; 65(14): 4134–4140. Available at: DOI:10.1016/j.ces.2010.04.038
- [61] Phromprasit J., Powell J., Wongsakulphasatch S., Kiatkittipong W., Bumroongsakulsawat P., Assabumrungrat S. H₂ production from sorption enhanced steam reforming of biogas using multifunctional catalysts of Ni over Zr-, Ce- and La-modified CaO sorbents. *Chemical Engineering Journal*. Elsevier B.V.; 2017; 313: 1415–1425. Available at: DOI:10.1016/j.cej.2016.11.051
- [62] Aloisi I., di Giuliano A., di Carlo A., Foscolo PU., Courson C., Gallucci K. Sorption enhanced catalytic Steam Methane Reforming: Experimental data and simulations describing the behaviour of bi-functional particles. *Chemical Engineering Journal*. Elsevier B.V.; 2017; 314: 570–582. Available at: DOI:10.1016/j.cej.2016.12.014

- [63] Xu P., Zhou ZM., Zhao C., Cheng Z. Ni/CaO-Al₂O₃ bifunctional catalysts for sorption-enhanced steam methane reforming. *AIChE Journal*. John Wiley and Sons Inc.; 2014; 60(10): 3547–3556. Available at: DOI:10.1002/aic.14543
- [64] Radfarnia HR., Iliuta MC. Hydrogen production by sorption-enhanced steam methane reforming process using CaO-Zr/Ni bifunctional sorbent-catalyst. *Chemical Engineering and Processing: Process Intensification*. Elsevier; 1 December 2014; 86: 96–103. Available at: DOI:10.1016/j.cep.2014.10.014
- [65] Radfarnia HR., Iliuta MC. Development of Al-stabilized CaO-nickel hybrid sorbent-catalyst for sorption-enhanced steam methane reforming. *Chemical Engineering Science*. Elsevier Ltd; 19 April 2014; 109: 212–219. Available at: DOI:10.1016/j.ces.2014.01.033
- [66] Feng HZ., Lan PQ., Wu SF. A study on the stability of a NiO-CaO/Al₂O₃ complex catalyst by La₂O₃ modification for hydrogen production. *International Journal of Hydrogen Energy*. October 2012; 37(19): 14161–14166. Available at: DOI:10.1016/j.ijhydene.2012.06.099
- [67] di Felice L., Kazi SS., Sørby MH., Martinez I., Grasa G., Maury D., et al. Combined sorbent and catalyst material for sorption enhanced reforming of methane under cyclic regeneration in presence of H₂O and CO₂. *Fuel Processing Technology*. Elsevier B.V.; 1 January 2019; 183: 35–47. Available at: DOI:10.1016/j.fuproc.2018.10.012
- [68] Micheli F., Sciarra M., Courson C., Gallucci K. Catalytic steam methane reforming enhanced by CO₂ capture on CaO based bi-functional compounds. *Journal of Energy Chemistry*. Elsevier B.V.; 1 November 2017; 26(5): 1014–1025. Available at: DOI:10.1016/j.jechem.2017.09.001
- [69] Dang C., Long J., Li H., Cai W., Yu H. Pd-promoted Ni-Ca-Al bi-functional catalyst for integrated sorption-enhanced steam reforming of glycerol and methane reforming of carbonate. *Chemical Engineering Science*. Elsevier Ltd; 2 February 2021; 230. Available at: DOI:10.1016/j.ces.2020.116226