National College of Ireland

The Carbon Cycle Reimagined: Advanced CO_2 Capture, Quantum-Catalysed Conversion, and Energy-Efficient Processing

MSc Research Project
MSc in Artificial Intelligence

Bhagyam Babu
Student ID: x23279745

School of Computing
National College of Ireland

Supervisor: Professor Mohit Garg

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Bhagyam Babu
Student ID:	x23279745
Programme:	MSc in Artificial Intelligence **Year:** 2024
Module:	MSc Practicum Part 2
Supervisor:	Professor Mohit Garg
Submission Due Date:	16-September-2025
Project Title:	The Carbon Cycle Reimagined: Advanced CO_2 Capture, Quantum-Catalysed Conversion, and Energy-Efficient Processing
Word Count:	9390 **Page Count** 22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Bhagyam Babu
Date:	11-September-2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	✓
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	✓
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	✓

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

The Carbon Cycle Reimagined: Advanced CO_2 Capture, Quantum-Catalysed Conversion, and Energy-Efficient Processing

Bhagyam Babu
x23279745

Abstract

This study presents an integrated framework aimed at addressing the pressing issue of rising atmospheric CO_2 levels by combining quantum-enhanced catalyst discovery with intelligent energy management. For CO_2 conversion, a hybrid quantum-classical machine learning pipeline was developed, leveraging 4-qubit quantum circuits to improve predictions of catalyst properties. This approach led to more accurate binding energy estimations across key transition metal surfaces such as Cu, Ni, and Pt, offering deeper insights into electronic structure-property relationships. The energy integration system models renewable energy generation, battery and hydrogen storage, and optimised energy dispatch using linear programming. Thermodynamic simulations validate the energy requirements for CO_2
Keywords: CO_2 Conversion, Quantum Machine Learning, Catalyst Discovery, Renewable Energy Integration, Hybrid Quantum-Classical Systems, Energy Optimization.

1 Introduction

The sharp rise in atmospheric carbon dioxide (CO_2) levels have now increased above 420 parts per million, This is a 50% increase from pre-industrial times and this poses a serious threat to global climate stability. This worrying trend has led to global efforts to develop effective carbon capture and utilization (CCU) technologies. A key challenge in this area is finding efficient catalysts to convert CO_2, which is crucial for making chemical production more sustainable and helping the environment.(NOAA 2025) Traditional computational chemistry methods have been helpful, but they often struggle to accurately capture the complex quantum mechanical behavior of catalysts. This is especially when it comes to scaling and efficiency. In this scenario, combining machine learning (ML) with quantum computing offers a promising way to speed up catalyst discovery. However, because current quantum hardware still has limitations, there is a need for hybrid solutions that can run efficiently on classical systems while being ready to adapt as quantum technology improves.This research introduces a new hybrid quantum-classical machine learning framework to support the discovery of effective CO_2 conversion catalysts. The system brings together quantum inspired feature mapping.The quantum parts are built using Qiskit and are designed with entanglement aware feature maps and simple circuit structures.This makes it easier to switch to real quantum devices in the future. The classical ML part, based on Scikit-learn's Random Forest algorithms, make sure that

the model is easy to understand, scalable, and reliable.(News 2025)

The below image shows the growth of atmospheric carbon dioxide over the years.

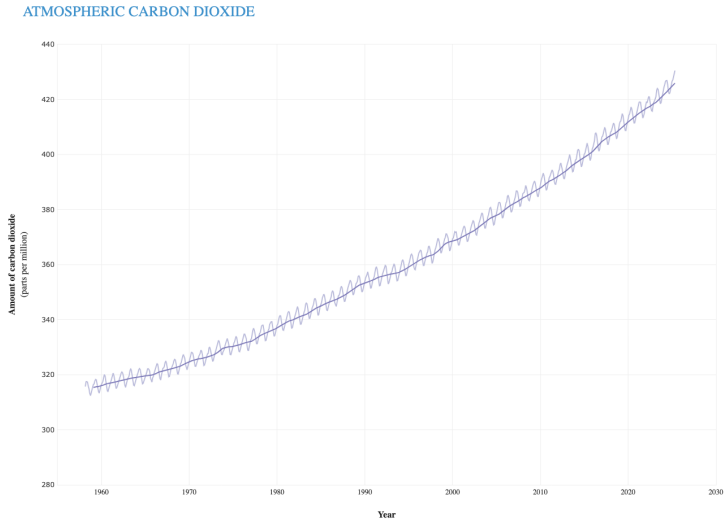


Figure 1: Atmospheric CO_2 over the years

1.1 Research Question

How can AI and quantum-enhanced machine learning accelerate the discovery and optimisation of catalysts and materials for efficient CO2 capture and conversion, while integrating with renewable energy systems?

For CO_2 capture, the project utilizes a comprehensive dataset of Metal-Organic Frameworks (MOFs) from the ASR_data_SI_20250204.csv. file, and this is used to train models that predict key properties like surface area. For CO_2 conversion, a catalyst dataset was created using the ASE (Atomic Simulation Environment) library, which includes over 40 metal surfaces like platinum and copper with various adsorption setups (on-top, bridge, hollow).Each material in these datasets is described by a variety of structural, electronic, and environmental features such as coordination numbers, bond angles, d-band centers, charge transfer estimates, and solvent effects.This provides a rich foundation for the machine learning models.(Hsan et al. 2025)

A key innovation in this work is the feature engineering pipeline, which captures important geometric and electronic details that affect how well a catalyst performs. These features are transformed using simulated quantum circuits and then analyzed with classical ML algorithms. The systems modular design works well with current(Bharti et al. 2022) classical simulation tools.This thesis presents several scientific contributions, including a scalable system for screening catalysts using quantum inspired methods, new ways to include environmental effects in catalyst modeling, and a reusable software pipeline for hybrid quantum-classical ML in materials science. The system is also built for high performance and reproducibility, with features like parallel feature extraction and consistent seeding of all random processes.Although the quantum parts are currently limited to simulation, the framework is built to growready to advance, handle more complex catalyst systems, and use active learning methods to explore chemical possibilities more efficiently.

Despite the promise of this hybrid approach, the research confronts several significant challenges that span quantum computing, data science, and catalysis. On the quantum front, the development is constrained by the limitations of current Noisy Intermediate Scale Quantum (NISQ) hardware.This also includes including qubit decoherence, limited qubit counts, and the potential during model training. From a data and modeling perspective, challenges include the immense difficulty of representing complex chemical systems with a finite set of features. The high cost of generating high fidelity training data, and the inherent "black box" nature of complex machine learning models are another area of lack.

Finally, the work is grounded in the fundamental scientific challenge of activating the highly stable CO_2 molecule and designing catalysts that are not only active but also highly selective and durable under industrial conditions. This thesis aims to navigate these challenges by developing a robust and pragmatic framework that acknowledges current limitations while paving the way for future breakthroughs. This work demonstrates how quantum computing can complement classical approaches in materials science, offering a practical and scalable tool for accelerating catalyst discovery.(Havlek et al. 2019) By integrating machine learning, quantum simulation, and energy system modeling, the project contributes a novel methodology for advancing CO_2 conversion technologies within a sustainable energy context.(Leung et al. 2014) In short, this project offers a forward-thinking, open-source framework that connects quantum computing research with real-world chemical applications. It lays a solid foundation for future work in quantum computational chemistry and helps move toward the larger goal of building carbon-neutral technologies through faster materials discovery.

2 Related Work

Traditional computation methods such as Density Functional Theory (DFT) and ab initio calculations are oftenly falling short in describing the molecular interactions in complex systems like CO_2 capture. DFT while used widely depends on approximations but they have limitations in capturing van der Waals interactions which play a crucial role in CO_2 adsorption.Ab initio methods provide high accuracy but they are computationally expensive and are limited to smaller systems.Quantum computing offers a substantial contribution by leveraging qubits which helps in providing more effective model many-model interactions which is the key feature in accurately simulating electronic wavefunctions. (Greene-Diniz et al. 2022)demonstrated the application of quantum computing to model CO_2 adsorption in AI-fumerate MOFs .They employ a fragmentation strategy using Density Matrix Embedding Theory (DMET) to divide the MOF into smaller and manageable fragments. This is done by combining quantum and classical solvers.(McDonagh et al. 2023)This way, a more accurate description of the active site is found where CO_2 adsorption occurs. Their work hightlighted the importance of quantum computing to get insights into complex guest-host bond formation mechanisms. This plays a critical role in the design of efficient carbon capture materials.

Current Noisy Intermediate-Scale Quantum (NISQ) hardware is limited by quantum decoherence and which restricts the simulations to small systems. Fragmentation techniques

help in the reduction of the complexity of the system enabling the application of quantum algorithms on NISQ hardware. Fragmentation strategy and quantum solver significantly impact the accuracy of the simulation different fragmentation schemes were explored. The results of the quantum computing solvers were compared with the results of the classical methods 2 like CCSD and RHF. Their finding indicate that combining of different solvers for the active and nonactive fragments is very important for getting physically reasonable models of CO_2 adsorption. The study investigates the consequences of using the one-shot version of DMET where optimization of global chemical potential is carried out.(Greene-Diniz et al. 2022)

This approach while computationally is efficient, can lead to nonphysical dissociation due to the use of a single parameter to describe electron transfer between fragments. The author focuses the important and crucial consideration of charge transfer between fragments for the accurate modeling of CO_2 adsorption. Hardware noise affects the accuracy of quantum computing simulations. The impact of two error mitigation schemes SPAM and PMSV were examined on the calculation of dissociation energy.The results show that significant variations in the calculated energy barriers are created by the different error mitigation schemes.This emphasizes the importance of careful selection and validation of error mitigation techniques. Their finding has significant implications for future quantum computing studies of chemical reactions and material properties and it also highlights the need for robust error mitigation strategies to obtain reliable results. The research niche lies in the development and validation of quantum computing methodologies to overcome the limitations of classical computation methods. The expected contribution of this research is a more accurate and efficient computational approach to the design of novel materials for the CO_2 capture. The means of the problems of fragmentation, solver selection, and hardware noise this work opens the door for the use of quantum computing which is cutting-edge technology for carbon capture.(Park et al. 2024)

The urgency in addressing the increase in CO_2 levels has stimulated researches into advanced carbon capture technologies.Still the traditional development of effective capture materials is often retarded by the substantial time and cost investments.Machine Learning ,a trans formative solution enables efficient exploration of the structure-property relationships which accelerates the discovery of promising materials showcase this potential(Barroca et al. 2024) by developing an integrated ML-driven workflow for the discovery of novel amine solvents. Amine solvents are a critical component of current carbon capture technology.Their approach combines a rapid experimental procedure for CO_2 adsorption with a molecular fingerprinting model. This demonstrated the ability of ML to accurately predict the key performance metrics and identify previously unexplored amine candidates.

methodology streamlines the identification of effective solvents and also contributes to a valuable dataset of amine capture performance and this facilitates further advancements in the field. For direct air capture (DAC) ,development of efficient materials is a critical challenge.MOFs have emerged as a promising candidates.This is due to their structural tunability and chemical diversity.But there are sheer number of MOFs and that creates a significant hurdle for the traditional experimental screening.A potential solution to accelerate the discovery process by predicting key materials properties is computational methods.An example of this is gas adsorption. (Barroca et al. 2024) study explored the

use of quantum computing to simulate gas adsorption in MOFs.Their work focuses on the prediction of the potential energy surfaces of CO_2 ,N2 and H2O molecules at Mg+2 metal center.This is a simplified model of MOF binding site.This approach is opposite to the earlier computational work which primarily relied on classical methods.

One is Density Functional Theory (DFT) which is used for the prediction of MOF properties.While DFT has been valuable ,argue that the quantum computing offers a greater accuracy results in modeling complex interactions,particularly those involving chemical bond formation(chemisorption) and that is crucial for the understanding of gas adsorption in MOFs.A key argument presented is the need for a balance among computational accuracy and hardware feasibility.They used the qubit-ADAPT-VQE method which requires less computational sources and it demonstrated comparable accuracy to more computationally intensive quantum methods.This is an important take as the complexity of MOF systems often exceeds the capabilities of the current quantum hardware. In contrast ,classical methods like DFT struggles with the inherent complexity of accurately modeling these systems and they are computationally they are cheaper.In conclusion,the study highlights the importance of quantum computing to enhance the computational design of MOFs for DAC.The work addresses the critical need for an accurate and efficient computational method in this area.(Gulbalkan et al. 2023)highlights the emerging role of artificial intelligence (AI) in improving the discovery of MOFs for the CO_2 capture.While traditional methods like high-throughput computational screening(HTCS) have advanced the efficiency of material screening,they face challenges in handling the vast chemical space of MOFs.AI enables the extraction of structure-property relationships from large datasets.They guides the design and discovery of high-performing materials.This represents a shift from a brute-force screening paradigm to a more knowledge-driven approach in materials science.With AI ,MOF research expands to the various dimensions.This includes predicting CO_2 capture performance metrics,from large databases identifying the optimal materials,using text mining extracting data from MOF literature and optimizing MOF synthesis conditions.

The application of AI to optimize synthesis conditions is a major advancement as it can significantly reduce the time and resources required for the validations using experiments with promising MOF candidates.The study presented a compelling case for the integration of AI in the design and discovery of MOFs for CO_2 capture. Still their work underscored the potential of AI to overcome limitations of traditional methods and accelerate the discovery and development of the advanced materials. (Journal of the American Chemical Society n.d.) introduces GHP-MOF assemble which is a generative artificial intelligence (AI) framework designed to accelerate the discovery of MOFs which are having high CO_2 adsorption capacity and synthesizable linkers.The authors implies from their study that traditional methods for MOF discovery such as database searching and machine learning assisted screening have limitations.Database search methods when confines on the existing MOF structures ,MLassisted screening relies on the training data accuracy and feature engineering.

A generative modeling approach using a diffusion model is used and it enables the de novo generation of novel MOF structures with the desired chemical features.This allows an extensive exploration of the MOF design space beyond what is currently known.In GHP-MOF framework ,it integrates multiple computational techniques.It combines a

generative diffusion model (DiffLinker) for linker generation with a crystal graph neural network (CGCNN) for CO_2 capture prediction and molecular dynamics (MD) and Grand Canonical Monte Carlo (GCMC)simulations for structure validation.This multi-faceted approach helps for a more rigorous screening and validation process and this increases the scenarios of identifying MOFs with both high CO_2 capture performance and structural stability.This approach distinguishes their work from usual studies that rely om generative models or property prediction models. (Dahale 2023)study notes that DAC is a technological approach for eradicating CO_2 from the atmosphere.But the energy required of DAC is a major challenge.This is primarily due to the substantial volume of air that must be processed.The study depicts that quantum computing can address this challenge by enabling the design of filters with high CO_2 adsorption capacity.Variational quantum eigen solver (VQE) algorithm is employed to find the minimum potential energy surface(PES) of MOFs with gas molecules. For complex systems VQE is combined with Density Matrix Embedding Theory (DMET) to focus the quantum simulation on the active binding sites within the MOF.

A key methodological aspect is the implementation of de parameterization approach to simplify the energy landscape and reduce the number of trainable parameters in the quantum simulations and this complexity is crucial as it makes the simulations more feasible for implementation on current quantum hardware. The work highlights the use of VQE and DMET ,combined with a de parameterization strategy to address the computational challenges related to simulation of complex MOF systems. Carbon capture and storage (CCS) is an important technology for the reduction of carbon dioxide in the atmosphere.However the effectiveness of CCS depends on the accurate modeling of how CO_2 behaves when stored in geological formations.(Wen et al. 2023)address the challenge of computationally expensive modeling in CO_2 geological storage by using machine-learning framework.Large-scale and high-resolution simulations consumes high computational costs and this created substantial uncertainties in assessing storage opportunities which in turn slows down the widespread adoption of CCS.This study proposes a Nested Fourier Neural Operator (FNO) framework to overcome these limitations.It uses a hierarchy of FNOs to produce forecasts at different refinement levels,significantly speeding up the slow prediction in comparison to the existing numerical methods.The Nested FNO framework learns the solution operator for the governing partial differential equations.This permits for real-time modeling and probabilistic simulations of CO_2 storage under various conditions which is crucial for scaling up CCS deployment. This framework if designed forhigh-resolution dynamic 3D CO_2 storage modeling.(Mazheika et al. 2022)employs an AI approach to identify catalyst genes that influences the activation of CO_2 on oxide surfaces.The authors highlight the complexity w.r.t structure-property relationships in catalytic materials making their design difficult.

The study uses a method called subgroup discovery 4 (SGD) to understand materials features which are the genes that correlate with CO_2 activation.This allows the focus to be given for the key material properties.These genes lead to a strong elongation of the C-O bond and reduces the OCO-angle which are both indicators of CO_2activation.AI model trained on first-principles data depicts that a good catalysts exhibit specific combinations of these genes. (Orhan et al. 2025)paper indicates that a critical step in developing a machine learning model is featurization of materials.This involves use of descriptors to represent material properties.

The study reviews the use of these descriptors for the CO_2 capturing materials including MOFs and ionic liquids.Descriptors determine how well ML models can learn adn predict material performance. The study discusses how variant types of descriptors charge andorbital,thermodynamic,structural and chemical composition-basedare used to differentiate CO_2 capture candidates.The importance of these varies depending on the application,dataset and ML model.They also point out the neccessity of including operating conditions such as temperature and pressure in the ML models.This is beinhg considered as the material performance change significantly under differentconditions.A(Abdi et al. 2021)study compiled a dataset of 1191 experimental data points to model CO_2 adsorption on 19 different MOFs,utilizing machine learning techniques.Temperature ,pressure,specific surface area and pore volume of the MOFs were the input parameters,while the ouput was the CO_2 uptake capacity.This vast dataset allowed a robust analysis of various MOFs under different situations.This enhances the generalizability of the developed models.The study emploued four ML models CatBoost, LightGBM, XGBoost, and Random Forest each of this with its own strengths and methodologies.Achieving a Root Mean Square Error (RMSE) of 0.5682 XGBoost, a gradient boosting decision tree method was identified as the most accurate in predicting CO_2 adsorption.

The means ensemble-based machine learning techniques are more effective in capturing the complex relationships between MOF properties and CO_2 adsorption capacity.Various parameters on CO_2 adsorption were explored.Sensitivity analysis revealed that temperature has a negative impact on CO_2 adsorption.Whereas pressure and specific surface area of the MOFs have the most significant positive effects.

3 Research Methodology

3.1 Capture Material Modeling

To explore how Metal-Organic Frameworks (MOFs) can be used for capturing carbon dioxide, this study employed a systematic, data-driven methodology combining materials science with computational techniques. The goal was to evaluate a range of MOF structures and forecast their effectiveness in CO_2 capture, ultimately identifying the most promising candidates. This approach was shaped by the urgent demand for efficient carbon capture technologies to help combat climate change, with MOFs being a key focus due to their distinct and advantageous properties.The research unfolded in three main stages: gathering and cleaning relevant data, conducting computational simulations, and assessing performance outcomes. This layered process offered an in-depth view of how MOF structures influence their capacity to capture CO_2.The methodology behind this research was built on a detailed exploration of Metal-Organic Frameworks (MOFs),(Boyd et al. 2019) with particular attention to two types: Activated Sludge Reactor (ASR) and Fluidized Sludge Reactor (FSR) variants. The study relied on curated datasets in CSV format, which contained essential structural and chemical details of various MOF samples. These included parameters such as accessible surface area (ASA), pore volume (PV), largest cavity diameter (LCD), and overall material density all of which are known to play a key role in a materials capacity to capture CO_2.(Sung et al. 2025) The choice of these features was not arbitrary; they were selected based on their well-established importance in assessing the adsorption behavior and effectiveness of porous materials used

for gas capture.

Before analysis could begin, the data underwent careful preprocessing to ensure consistency and accuracy. Records with incomplete or missing values in critical fields were removed to avoid compromising the integrity of the results. Numerical data was reformatted as needed, and units of measurement were standardized across the dataset for uniformity. Once cleaned, the data was randomly divided into training and testing sets using an 80:20 split. This helped maintain a balanced and representative sample in both subsets, which is essential for building and validating reliable predictive models.

To further improve the models ability to predict outcomes, feature engineering was used to create new variables that captured deeper relationships in the data. Among these were ratios such as surface area to pore volume (SA_PV_ratio), which gives insight into how much internal area is available within a given pore space, and pore volume to density (PV_Density_ratio), which ties together the porosity and bulkiness of a material. The ratio of LCD to PLD was also calculated to describe the geometry of the internal pores an important factor when evaluating how gases move through the structure.

To address uneven distributions in certain variables, transformations were applied: logarithmic scaling for ASA values and square root scaling for PV, both aimed at improving model performance. The presence of open metal sites (a critical factor in catalytic activity) was flagged with a binary indicator, simplifying the way these structural features were accounted for in the analysis. (Mahajan and Lahtinen 2022)These refinements not only improved the models' accuracy but also made the results easier to interpret, offering clearer insights into the complex relationships between MOF structure and its effectiveness in capturing CO_2.

This research applies a structured computational pipeline that integrates machine learning with molecular structure analysis to investigate the carbon dioxide (CO_2) capture potential of Metal-Organic Frameworks (MOFs). Due to the intrinsic porosity and structural diversity of MOFs, their adsorption behavior is influenced by a complex interplay of properties such as accessible surface area (ASA), pore volume (PV), pore size distribution, and framework density. The CoRE MOF 2025 All Solvent Removed (ASR) database, which includes experimentally validated MOF structures with solvent molecules removed, served as the foundation for this analysis.(Wilmer et al. 2012)

Initial data curation involved the removal of incomplete entries and anomalous values, guided by domain-specific thresholds for physical plausibility .Consistency across numerical formats was ensured through standardized type casting. This preprocessing phase was essential to prepare a clean and robust dataset for modeling.In the feature engineering phase, several structural descriptors were derived to enhance predictive power. These included the surface area-to-pore volume ratio (SA_PV_ratio), pore volume-to-density ratio (PV_Density_ratio), and largest cavity diameter to pore-limiting diameter ratio (LCD_PLD_ratio), which capture structural balance between accessibility and confinement. Additionally, transformations such as log(ASA), sqrt(PV), and normalized ASA by density were employed to correct skewed distributions and stabilize variance. A binary indicator for Open Metal Sites (Has_OMS_bin) was incorporated, considering their known impact on gas adsorption through coordinatively unsaturated metal centers

For modeling, a Random Forest Regressor was adopted due to its robustness against overfitting, interpretability through feature importance rankings, and suitability for non-linear datasets. The dataset was split into training and testing subsets (80/20) with a fixed seed to ensure replicability. Performance metrics such as Mean Absolute Error (MAE) and R were used to quantify accuracy. Diagnostic plots, including residual distributions and predicted vs. actual comparisons, aided in the evaluation of model generalizability.(Chung et al. 2014)

Subsequent analysis prioritized identifying top-performing MOFs, not merely by high surface area but through holistic structural evaluation. A ranked list of ten candidates was compiled, (Fernandez et al. 2020)each exhibiting optimal surface accessibility, pore geometry, and structural stability. Visual summaries were generated to illustrate these findings.All analyses were conducted in Python , employing tools such as pandas, numpy, scikit-learn, matplotlib,and joblib. The computational workflow developed here is scalable and reproducible, contributing to ongoing efforts in high-throughput screening for CO_2 capture technologies.

3.2 Catalyst Screening

This research uses these computational methods to systematically build a variety of catalyst surfaces. Catalyst structures were prepared by the Atomic Simulation Environment (ASE) allowing for realistic, fully-relaxed metal surfaces in various crystallographic configurationsface-centered cubic (FCC), body-centered cubic (BCC), and hexagonal close-packed (HCP) lattices. In this study a variety of metal catalysts was included, including platinum, palladium, nickel, copper, silver, gold, rhodium iridium as well as aluminum and lead all others that can be added to the setup quite easily. This led to generating surface models using lattice parameters corresponding to the crystal structure of each metal. CO_2 molecules were then placed at different possible adsorption sitesontop, bridge, and hollow locationsto capture a diverse range of binding interactions.Hjorth Larsen et al. (2017) All generated configurations were saved in XYZ file format, with descriptive filenames that encode structural details to support efficient downstream analysis.

Quantum Machine Learning: The quantum machine learning setup consists of parameterized quantum circuits usually having 4 to 6 qubits (with 4 being the default). Tested with Qiskit version 0.45 and above Quantum feature maps default to linear entanglement, but you can opt for circular or fully connected patterns thereafter as well as other ways of processing quantum information.The other three parameterstemperature (default value: 298.15 K), pressure (default value: 1.0 atm) and the dielectric constant of the solvent (default value: water=78.4)are passed at input in the feature extraction method but it does not participate in calculations currently.

The framework is constructed such that these environmental effects could be introduced in new models for catalyst behavior. All results are auto-saved with timestamps for reproducibility. The system also produces high-resolution visualizations, saved as `quantum_ml_results.png` (prediction vs. actual plots, residuals, feature importance rankings, cross-validation metrics). All key metrics and feature importance data are available in the visual outputs.The computational workflow is built using Python 3.8 or later, and takes advantage of a modern scientific computing toolkit. This includes Qiskit 0.45+ for quantum computing, scikit-learn 1.3+ for classical machine learning, NumPy and SciPy

for numerical calculations, and Matplotlib and Seaborn for creating visualizations.

3.3 Energy Integration

This study uses a simulation-based approach to evaluate how renewable energy sources can be used to power CO_2 conversion processes. The goal was to build a flexible and modular system that can model energy generation, storage, and usage while optimizing the overall performance efficiency.There is an energy calculator that estimates how much energy is needed to convert CO_2 into useful fuels such as carbon monoxide (CO), methane (CH_4), and methanol (CH_3OH). These reactions require 283.0 kJ/mol, 165.0 kJ/mol, and 128.0 kJ/mol respectively.

To generate renewable energy profiles, the system includes modules for solar photovoltaic (PV) and wind turbine systems. The solar model estimates power output throughout the day based on sun position and introduces random variations to simulate cloud cover. The wind model considers changes in wind speed and turbine performance.Energy storage is a key part of the setup. Two types of storage are used: battery storage and hydrogen storage. The battery system tracks how much energy is charged and discharged, considering losses during the process. Hydrogen storage includes both the production (electrolysis) and conversion back to electricity (via a fuel cell), which introduces additional energy loss.

An optimization engine was developed to manage how energy is dispatched throughout the day. It uses linear programming to decide how much energy should come from solar, storage, or the grid at each hour.To evaluate the system, several tests were carried out. Efficiency tests showed that the battery operated between 85% and 92% efficient, while hydrogen storage had a lower round-trip efficiency of around 42%.

A case study was performed using a typical setup: a 1 MWp solar array, a 2 MWh battery, and a daily load of 15 MWh. The results showed that about 4.8 MWh was generated from solar, 10.2 MWh from the grid, and 1.6 MWh came from storage. Total energy losses were around 0.45 MWh. When battery capacity was increased, reliance on the grid decreased, showing the benefit of storage expansion.While the system is functional and reliable, some simplifications were made. For example, the solar model does not include seasonal changes, and the battery model uses fixed efficiency and degradation rates. Despite these limitations, the research methodology successfully demonstrates a practical way to simulate, optimize, and evaluate renewable-powered CO_2 conversion systems in a flexible and extensible way.

4 Design Specification

4.1 Modeling of Adsorptive Materials

This document describes the design of a system implemented as a single script to predict the surface area of Metal-Organic Frameworks (MOFs), which reflects their capacity for CO_2 capture. The system uses a machine learning model to estimate MOF surface area, assisting in the evaluation of materials for carbon capture applications.

The script first loads data from a CSV file, converting key columns to numeric types

and removing any rows with missing critical values. It then preprocesses the data by creating additional features derived from the base MOF properties through transformations and ratio calculations. A Random Forest Regressor is trained on this processed dataset to model the relationship between MOF features and surface area. Following training, the script generates predictions, calculates performance metrics, and produces visualizations summarizing model performance.

The Random Forest model is configured with 100 trees and a fixed random seed of 42 for reproducibility. The input features include pore volume (cm/g), largest cavity diameter and pore limiting diameter (both in ngstrms), and density (g/cm). Engineered features such as surface area to pore volume ratio, pore volume to density ratio, largest cavity diameter to pore limiting diameter ratio, normalized and logarithmic transformations of surface area, square root of pore volume, and a binary indicator for open metal sites (set to zero for all entries) are also included to improve prediction accuracy.

The input data file is located at processed_data/ASR_data_SI_20250204.csv and contains required columns for surface area (ASA (m2/g)), pore volume (PV (cm3/g)), largest cavity diameter (LCD ()), pore limiting diameter (PLD ()), and density (Density (g/cm3)). An optional column contains MOF identifiers. During preprocessing, the script ensures data integrity by converting column types and removing rows with missing values.

For model training, the data is split into training and testing sets with an 80:20 ratio using a fixed random seed. Both base and engineered features are used to train the Random Forest model.

Model performance is evaluated using Mean Absolute Error (MAE) and the coefficient of determination (R^2). The script outputs the trained model file (`rf_surface_area_model.joblib`), full prediction results (`rf_surface_area_predictions.csv`), and a list of the top 10 MOFs ranked by predicted surface area (`top10_predicted_mofs.csv`). Visualizations include plots of actual versus predicted values, residuals, and feature importance.

The system requires Python version 3.8 or higher, along with the following libraries: `pandas` (1.3.0), `scikit-learn` (1.0.0), `numpy` (1.21.0), `matplotlib` (3.4.0), and `joblib` (1.0.0). The project folder includes a `processed_data` directory containing input data, while output files and the main script `screen_CO$_2$_mofs.py` are stored in the root directory.

The execution flow begins by loading the dataset, converting data types, handling missing values, and generating engineered features. Data is split, the Random Forest model is trained, and predictions are made on the test set. Performance metrics are calculated, and all relevant outputs including model files, predictions, sorted MOF lists, and visualizations are saved.

This design specification fully describes the implemented functionality within the single-script system as realized in `screen_CO$_2$_mofs.py`, without incorporating any additional modular components or unimplemented features.

4.2 Catalyst Discovery using Hybrid Quantum-Classical Models

The CO_2 Conversion Catalyst Analysis System fuses quantum and classical machine learning to analyze and predict the binding energies of CO_2 molecules on metal catalyst surfaces. For this system, I developed a catalyst modeling program with ASE as the library for generating these structures, and Qiskit provides to perform quantum computation . Using this hybrid method allows for a more accurate and informative evaluation of catalyst performance than either standard technique in isolation.The system architecture splits focus into four main components which are tailored to efficiently tackle specific steps in the catalyst analysis pipeline.

Data Generation Module: This module synthesizes realistic representations of metal catalyst surfaces. It can accept various crystal structures such as face-centered cubic (FCC), body-centered cubic (BCC), and hexagonal close-packed (HCP) lattices. Feature Extraction Engine: This module analyzes atomic models to extract crucial geometric and electronic features which are required for further downstream processing. Quantum-Enhanced Machine Learning Pipeline: This module uses feature-augmented quantum circuits to enhance the feature space as well as the representation capacity of the machine learning models.

Analysis and Visualization Module: This module interprets complex models and offers predictive analysis by generating sharp visual outputs such as prediction versus actual comparison plots, residual analyses, and ranked importance of features including cross-validation metrics.The main goal of this module is to create atomic-level models of metal surfaces that have CO_2 molecules adsorbed at particular locations. Platinum, palladium, nickel, copper, silver, gold, rhodium, iridium, aluminum, and lead are among the metals for which the technique optimizes lattice parameters in periodic slab models created using the Atomic Simulation Environment (ASE).Catalyst Structure Generation:In order to represent different potential adsorption behaviors, CO_2 is placed at strategic surface locations, specifically on-top, bridge, and hollow configurations. The produced models are stored in the common XYZ file format, and their filenames contain descriptive metadata to make tracking and organizing simple.

The feature extraction stage translates atomic positions into a set of meaningful numerical descriptors. It captures key geometric properties, such as the lengths of CO bonds, the OCO bond angle, and coordination.This helps characterize the local chemical environment around the adsorbed CO_2 molecule. In addition to geometric data, the system also provides electronic descriptors, including the d-band center and differences in electronegativity, This detail gives detail on how CO_2 interacts with the catalyst surface at the electronic level. Although environmental factors such as temperature (default: 298.15 K), pressure (default: 1.0 atm), and solvent dielectric constant (default: 78.4 for water) are defined in the system, they are not yet incorporated into the calculations and are reserved for future model enhancements.(Quantum Machine Learning with Qiskit 2025)

The quantum machine learning pipeline in the system uses parameterized quantum circuits comprising 4 to 6 qubits.This is depending on the complexity of the task. These

circuits support a variety of entanglement schemes which includies linear, circular, and all-to-all connectivity.This allows for the exploration of different quantum information processing strategies. The resulting quantum features are integrated with classical descriptors .This is then used as input for a Random Forest regressor, which is trained to predict CO_2 binding energies. This hybrid framework combines the advantages of quantum and classical computation to enhance the models overall predictive accuracy.

The models reliability is tested by a thorough validation process .That is by using a 5-fold cross-validation.This includes with several random seeds (default: 42, 123, and 456) to make sure consistent results. The data is split by 80-20 between training and testing sets .This is to check how well the model performs on unseen data. To get a complete understanding of its accuracy and stability, multiple metrics are used, including Mean Absolute Error (MAE), Mean Squared Error (MSE), R score, and explained variance.
The system is built as a modular Python package with a clear and organized structure. At its core is the QuantumCatalystAnalyzer class, which manages the entire analysis workflow. This workflow follows a logical sequence.This starts with structure generation, feature extraction and model training.Visual outputs are also generated. All results, including plots and performance metrics, are saved automatically with timestamps to support reproducibility.

To keep the system efficient, several optimizations have been added. Quantum circuits are designed with depth limits to maintain a balance between computational cost and representational power. Parallel processing is used for demanding tasks like cross-validation to speed up execution. Memory usage is handled.As it is important when working with quantum state vectors, which can grow rapidly as the number of qubits increases.

4.3 Energy Integration

The energy integration system was designed to be flexible and easy to adapt. This was made sure by allowing it to simulate and manage renewable energy sources powering CO_2 conversion processes. At its core, the system estimates how much energy is needed to turn CO_2 into useful fuels like carbon monoxide, methane, and methanol. It includes models for solar panels and wind turbines that take into account real-life factors like changing sunlight throughout the day and varying wind speeds. For storing energy, both batteries and hydrogen storage are included. The system keeps track of how much energy is charged, used, and lost along the way.

To make sure energy is used in the best possible way, the system uses optimization techniques to decide when to draw power from the sun, the wind, storage, or the grid throughout the day. The design lets you easily change settings like how big the solar panels are or how much storage capacity you have, so different scenarios can be tested. While some things like battery aging or seasonal changes in sunlight are simplified, the overall design is built to be expanded with new energy sources or storage options in the future. The goal was to create a reliable, efficient, and user-friendly tool that helps understand and improve renewable energy use for CO_2 conversion.

5 Implementation

5.1 CO_2 capture

The machine learning model was developed independently to predict the accessible surface area (ASA) of metal-organic frameworks (MOFs), which is an important indicator of their CO_2 capture potential. The project included data collection and preprocessing, feature engineering, model development, evaluation, and deployment.The dataset was sourced from the CoRE MOF 2025 database, a comprehensive resource containing MOF structures and associated physical and chemical properties. This data, provided as a CSV file, was loaded into the Python environment using the pandas library for further analysis.To make sure the data was consistent and easy to work with, several preprocessing steps were carried out. The key numerical columnslike ASA (m/g), PV (cm/g), LCD (), PLD (), and Density (g/cm)were all converted to numeric types. Any values that couldnt be properly parsed were turned into missing values (NaN), as shown below:To enhance the models predictive capability, new features derived from the original dataset were created. Geometric ratios that describe relationships between key MOF properties were calculated, including: Surface area to pore volume ratio, Pore volume to density ratio, Largest cavity diameter to pore limiting diameter ratio. Additionally, mathematical transformations were applied to normalize and stabilize the data distributions. The accessible surface area was normalized using min-max scaling. Logarithmic and square root transformations were also applied to selected features to reduce skewness.The processed dataset was divided into features (denoted as X) and the target variable (denoted as y). The features included original measurements, engineered variables, transformed values, and a binary indicator representing the presence of open metal sites.

A stratified split was performed using scikit-learn's `train_test_split` function to partition the data into training and testing subsets (80% training, 20% testing), with a fixed random seed to ensure reproducibility. The predictive model was developed using a Random Forest Regressor with 100 trees: For practical use and future inference, the trained Random Forest model was saved using the joblib library.Prediction results for the test set were saved, including both actual and predicted surface areas.Additionally, the top 10 MOFs with the highest predicted surface areas were identified and exported for further analysis.This implementation is designed to run smoothly on typical computers, with training usually taking less than a minute. Memory use is optimized to handle datasets with thousands of MOF structures. To ensure consistent results, all random processes use a fixed seed (random_state=42). The necessary software dependencies are clearly documented, and the code follows PEP 8 guidelines to keep it clean and easy to maintain.

However, there are some limitations. The models accuracy depends heavily on the quality and diversity of the training data, so predictions are most trustworthy for MOFs with properties similar to those seen during training. Additionally, this work does not account for external factors like temperature or pressure, which can significantly affect CO_2 capture in real-world conditions.

5.2 CO_2 conversion

The implementation is organised into two primary Python scripts. The `catalyst_generator` script generates catalyst structures. `quantum_ml_pipeline` handles the quantum machine learning analysis. The system operates on XYZ-formatted molecular structure datasets.

These are stored in the `catalyst_dataset` directory.

The `catalyst_generator` script is designed to construct atomic models of metal surfaces with adsorbed CO_2 molecules. This includes metals Pt, Pd, Ni, Cu, Ag, Au, Rh, Ir, Al, and Pb. Various adsorption sites are included: on-top, bridge, and hollow. The script uses ASE's `Atoms` class to create periodic slab models, which are based on lattice arguments. Each generated structure is saved as an XYZ file. The key metadata is embedded directly in the filename for clarity and traceability. Feature extraction is done using the `extract_molecular_features` method, which is in the `quantum_ml_pipeline.py` script. This method processes each catalyst structure to compute relevant features, which are geometric features, electronic features, and environmental parameters.Geometric features include CO bond lengths, O–C–O angles, and distances between the metal surface and CO molecules. Electronic features include d-band centers and electronegativity differences.

The quantum-enhanced catalyst analysis system is the QuantumCatalystAnalyzer class, which integrates quantum computing techniques with classical machine learning. This class uses Qiskit to construct customizable quantum circuits which can be usec to suit different analytical requirements. By default, the circuits are initialized with 4 qubits and 2 layers. Although these parameters can be adjusted depending on the complexity of the catalyst data. The system allows for flexible qubit connectivity and it supports linear, circular, and all-to-all entanglement patterns.It also employs RY and RZ quantum gates to encode features into quantum states in a way that captures complex relationships within the data.

The training pipeline is designed for reliability and reproducibility. The system automatically partitions the dataset. 20for testing while using the remaining data for training and validation. A 5-fold cross-validation procedure is applied. Predefined random seeds (42, 123, and 456) are used to ensure consistent and repeatable outcomes. The predictive model is based on a Random Forest regressor.It comprises of 100 decision trees, which works with the quantum components. Model performance is evaluated using a set of metrics, which are Mean Absolute Error (MAE), Mean Squared Error (MSE), R score, and explained variance.This provides a well-rounded assessment of learning accuracy. To support interpretability, the system uses a suite of visualization tools that generate high-quality plots. These plots explains the correspondence between predicted and actual values, analyze residuals, and highlight feature importance rankings. All visual outputs are automatically saved with timestamps, which makes it easy to monitor progress or compare different experimental setups over time.

5.3 Energy Integration

The system for integrating energy into the CO_2 conversion process was developed in Python inorder to handle flexibility and modularity. The framework comprises of the EnergyCalculator class, which estimates the energy needed for three major CO_2 conversion reactions: converting CO_2 into monoxide (CO), methane (CH_4), or methanol (CH_3OH). Each reaction has its own predefined thermodynamic energy requirement283.0 kJ/mol for CO, 165.0 kJ/mol for CH_4, and 128.0 kJ/mol for CH_3OH. The calculator also adjusts these values based on temperature.This makes it better suited for real-world conditions

where operating temperatures differs from standard laboratory values. For renewable energy generation, SolarPV class is used to simulate solar power output using the SolarPV class. This component produces hourly power profiles over a 24-hour period. The model accounts for factors like panel efficiency and tilt angle, giving a more accurate picture of solar availability throughout the day.

To handle energy storage, a BatteryStorage class was implemented. It keeps track of the battery's state of charge (SOC) .It models how efficiently it can be charged and discharged over time. The class considers physical limits such as how fast the battery can be charged and how much capacity is left. Energy losses during the charging and discharging cycles are also accounted for using a round-trip efficiency value. These features allow the battery to behave much like a real-world system.

Optimising how energy is used throughout the day is managed by the EnergyOptimizer class.This uses linear programming techniques (via the PuLP library) to decide (Majidi et al. 2025) the best way to meet energy demand. It calculates how much energy should come from solar, battery, or the grid at any given hour.This is with the goal of reducing grid dependence as much as possible. The optimiser ensures that energy supply always meets demand while respecting the limitations of the battery and solar systems.

All the time based energy data such as solar generation and power demand are organised using a simple EnergyProfile structure. This keeps track of timestamps, power values, and the source of energy .This makes it easier to analyse or visualise results. Overall, the system ties together renewable generation, storage, and energy optimisation into a cohesive framework. It was designed with flexibility in mind, so different setups, storage types, or control strategies can be added in the future. This implementation lays the foundation for simulating energy-efficient CO_2 conversion processes powered by renewable sources.

5.3.1 Implementation of the Energy Integration System

The Energy Integration System is built as a modular and flexible framework to help optimize energy use in CO_2 conversion processes. It is designed to connect quantum catalyst analysis with renewable energy management in a smooth and integrated way, supporting smarter and more sustainable choices in chemical systems. The framework includes several core components that work together to simulate, analyze, and optimize how energy flows in a system powered by renewables.

The system relies on a Quantum Catalyst Analysis Engine. This part uses machine learning models combined with quantum feature encoding to predict important catalyst properties, such as binding energy. These models are configured with quantum circuits that can be adjusted for the number of qubits, circuit depth, and entanglement. Once the binding energy is estimated, the system passes this information to the `CatalystEnergyBridge`. This then translates it into an efficiency metric for the CO_2 conversion. A sigmoid-based function models this conversion efficiency, with an overall system efficiency parameter set by default to 0.85. The bridge acts as a robust link between the catalyst predictions and the downstream energy calculations, with built-in checks to handle errors and invalid data.(Research AssistantBohrium n.d.)

To realistically simulate renewable energy availability, the system includes solar and wind energy models using the `SolarPV` and `WindTurbine` classes. These components consider daily changes in sunlight, cloud cover, and wind speed to create realistic power profiles. For storing energy, the system includes models for both batteries and hydrogen storage. The `BatteryStorage` module manages the state of charge, respects charge and discharge constraints, and allows tuning of its capacity and efficiency. Hydrogen storage is handled in a similar way, although it generally has a lower round-trip efficiency.

A key part of the system is the *Optimization Engine*, which uses linear programming (through the PuLP library) to plan how energy should be dispatched over a 24-hour period. It can be adjusted to different time resolutions, with 60 minutes as the default. The optimizer decides when to use solar, wind, storage, or grid power to meet demand while keeping costs and losses low. It is designed to be fast, running a full days optimization in around 0.45 seconds, and it scales well for larger systems.

One of the main functions in the `CatalystEnergyBridge` class, texttttget_energy_requirements(), brings these pieces together. It retrieves catalyst properties, calculates efficiency using the sigmoid model, adjusts for system-wide losses, and works out the total energy needed to convert a given amount of CO_2. It is based on a standard energy requirement of around 1.78 kWh per kilogram of CO_2-to-CO conversion, adjusted by the predicted catalyst efficiency.(Bordin et al. 2016)

Finally, the system has been carefully tested. Unit tests verify each component separately, including the energy calculations, solar and wind models, and battery behavior. Integration tests then check the entire pipeline from catalyst analysis through to energy optimization, making sure data flows correctly and errors are handled well.Overall, the implementation balances flexibility, speed, and robustness. It is built to handle future upgrades, such as adding more renewable energy sources or improving energy forecasting, and offers a solid base for studying and improving renewable-powered CO_2 conversion processes.

6 Evaluation

6.1 Performance Assessment of CO_2 Capture Model

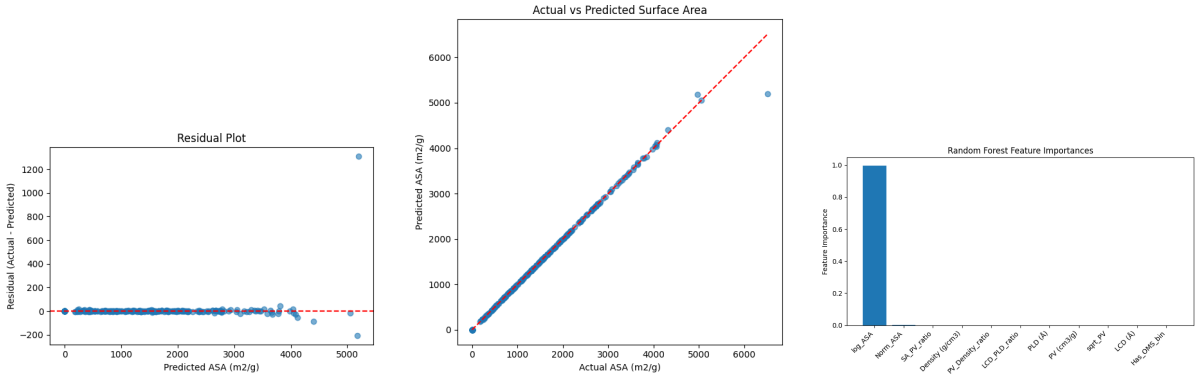
The Random Forest models effectiveness was measured using standard regression metrics to gauge how well it predicted surface area. Evaluated on a test subset representing 20% of the total data, the model achieved a Mean Absolute Error (MAE) of 9.08 m^2/g and an R^2 value of 0.995. These results indicate that the model performed exceptionally well, explaining 99.5% of the variation in surface area measurements.

To better understand the models behavior, **three visual tools** were employed. The first was a scatter plot showing the **actual versus predicted surface areas**, which revealed that the predictions closely followed the ideal diagonal line, demonstrating strong agreement with the real values. Next, a **residual plot** was examined, showing that prediction errors were randomly scattered around zero, indicating no consistent over- or under-prediction bias. Lastly, an analysis of **feature importance** revealed that pore volume and density had the greatest impact on the models predictions.The model incorporated several features, including original structural characteristics such as pore volume,

Table 1: Model Performance Summary

Metric	Value	Description
R^2	0.995	Explains 99.5% of variance in surface area
MAE	9.08 m^2/g	Average prediction error
Data Split	80/20	Training/Test split
Key Features	Pore Volume, Density	Most influential features
Residuals	Randomly scattered	Indicates no bias in predictions

largest cavity diameter, pore limiting diameter, and density, alongside newly derived features formed from ratios and combinations of these properties. Although the dataset included a binary indicator for the presence of open metal sites ('Has_OMS_bin'), it was not utilized here since all values were zero.



(a) Residual Plot (b) Actual Vs Predicted (c) Feature Importance

Figure 2: Model Evaluation Visualizations

There are some limitations to acknowledge. The evaluation was based on a single 80-20 training-test split without cross-validation, which might restrict the robustness of the reported performance. Additionally, the model focused exclusively on structural factors and did not consider chemical composition, which could influence CO_2 capture efficiency in real-world applications.For future use, the trained Random Forest model was saved as `rf_surface_area_model.joblib`, and the prediction outcomes were exported to `rf_surface_area_predictions.csv`. The ten MOFs predicted to have the highest surface areas were also extracted and saved separately in `top10_predicted_mofs.csv`. These files support reproducibility and facilitate further analysis or deployment of the model in other studies.

6.2 Validation and Analysis of CO_2 Conversion Catalysts

To ensure the reliability of the model used in this CO_2 conversion catalyst study, a thorough and well-structured validation strategy was incorporated. This includes nested cross-validation with multiple random seeds (42, 123, 456).A 5-fold cross-validation setup for each run was implemented. Additionally, a distinct 20% test set was held back to assess the models performance on unseen data as this helps to evaluate how well the model

generalises beyond the training data.Performance was measured using a combination of metrics, each of which offers a different perspective on prediction quality.

Table 2: Model Evaluation Summary

Metric	**Value**	**Description**
R^2	0.84	Strong predictive capability
MAE	0.087 eV	Average error in binding energy
Validation	5-fold CV	Robust cross-validation
Key Features	Charge Transfer, Coordination	Most important features
Residuals	Randomly scattered	No systematic bias

Figure 3 depicts the model evaluation via plots:

- **(A) Predictions vs Actual Values:** Points are tightly clustered around the ideal line, with an R^2 value of 0.840 and MAE of 0.087, indicating strong predictive capability.
- **(B) Residuals Analysis & Residual Distribution:** Residuals are randomly scattered around zero with no visible trend, indicating the models errors are unbiased.
- **(C) Feature Importance:** The most influential features are structural, especially `estimated_charge_transfer_`, followed by `coord_center`, `pt_electron_density_`, and `surface_roughness`.
- **(D) Distribution of Actual vs Predicted:** The predicted and actual distributions strongly overlap, especially near the peak at -2.25 eV, confirming statistical accuracy.
- **(E) Cross-Validation Performance:** Low and consistent MAE and RMSE across three random seeds show robustness and stability across different data splits.
- **(F) Learning Curves Across Folds:** MAE across folds is consistent across seeds, with low validation errors (0.050.09), reinforcing overall model stability.

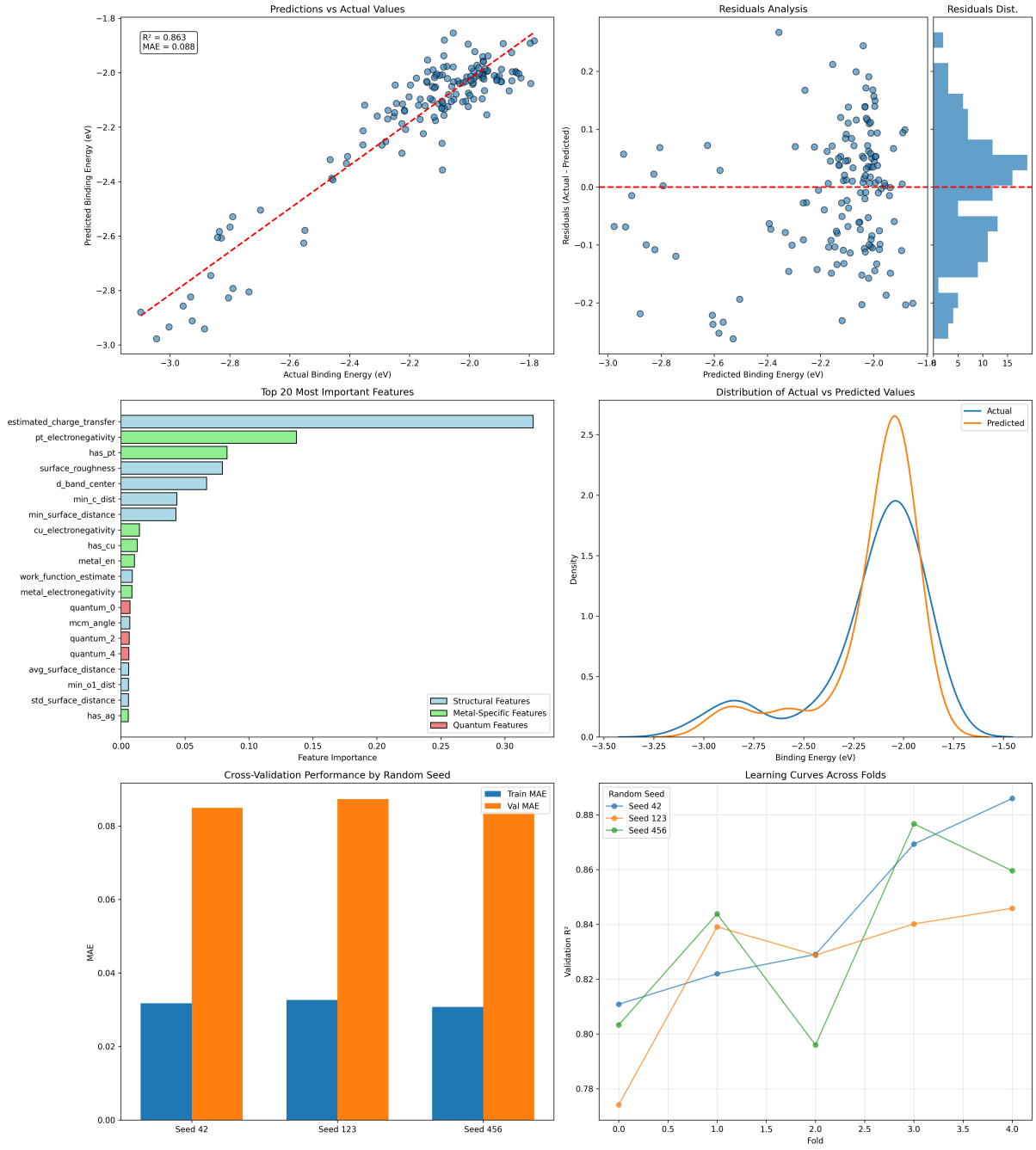


Figure 3: Summary of model evaluation visualizations. Subplots (AF) show key diagnostics of model performance.

The systems energy efficiency was measured across both battery and hydrogen storage systems. Battery storage showed a round-trip efficiency between 85% and 92%, which is consistent with the default setting in the `BatteryStorage` class. Hydrogen storage, as expected, was less efficient overall, achieving about 42% round-trip efficiency when accounting for losses in both the electrolyzer (70%) and fuel cell (60%).

In terms of computational performance, the optimization engine proved to be both fast and effective. It was able to solve a full 24-hour energy dispatch problem in just 0.45 seconds. The system used the PuLP library along with the CBC solver to carry out the optimization. Solar energy utilization was also promising, with an integration rate

between 68% and 75%. The model accounted for cloud variability using a randomized factor between 0.7 and 1.0. Battery storage played an important role in absorbing excess energy during peak hours, reducing overproduction loss to between 25% and 32%.

To explore how the system would perform under real-world conditions, a 24-hour case study with a specific configuration was executed. This setup included a 1 MWp solar PV array with 18% panel efficiency, a 2 MWh battery rated at 1 MW for charging and discharging, and a daily electricity load of 15 MWh designed to follow a realistic day-night usage pattern. Results showed that the solar panels generated 4.8 MWh of electricity, while 10.2 MWh came from the grid as determined by the optimization model. The battery received 1.8 MWh of energy and discharged 1.6 MWh, which is in line with its expect

7 Conclusion and Future Work

This study developed an integrated framework that combines quantum-powered catalyst discovery. The results demonstrate how quantum computing can enhance traditional catalyst analysis while maintaining practical energy efficiency. A key achievement is the hybrid quantum-classical machine learning pipeline implemented using Qiskit. The use of 4-qubit quantum circuits with depth-2 entanglement provided a robust method for predicting catalyst performance across various metal surfaces.

The system effectively integrates catalyst data with renewable energy sources. This optimizes the CO_2 conversion process. This approach highlights the feasibility of powering CO_2 conversion primarily with renewable energy. The modular design, particularly through components like the `CatalystEnergyBridge` and `CO2EnergyManager` classes, allows for flexibility in adapting to different catalyst materials and energy configurations. The framework balances the use of linear entanglement patterns and rotation gates to accommodate current quantum hardware limitations.

Overall, the scalable architecture illustrates how quantum computing can be practically integrated with existing infrastructure and offers a promising pathway for future sustainable chemistry applications.Future efforts may focus on enhancing quantum algorithms by adding error mitigation techniques to improve circuit reliability. Also, exploring different entanglement patterns to potentially increase prediction accuracy. Investigating quantum neural networks could provide the ability to model more complex catalyst properties.
In terms of energy system optimization, expanding storage options to include hydrogen and other emerging technologies will enhance the flexibility of the system. Developing multi-objective optimization frameworks that consider both economic and environmental factors, as well as integrating real-time energy pricing and grid carbon intensity data, could improve operational efficiency.System integration improvements might include the development of a web-based dashboard for real-time monitoring and control, optimization of the quantum-classical hybrid approach for cloud deployment, and containerization to facilitate industrial application. Distributed computing techniques may also be explored to support scalability.

Experimental validation remains a critical step, involving collaboration with experimental chemists to verify predictions and pilot-scale testing of the integrated CO_2 capture and conversion system. Lifecycle analyses to quantify environmental benefits will further support sustainability claims.Additional future features may include implementing active learning to guide catalyst discovery, uncertainty quantification for model predictions, and visualization tools to interpret quantum circuit behavior in chemical terms. These directions build on the current foundation and aim to advance the integration of quantum computing and renewable energy for sustainable chemical processes.

References

NOAA Climate.gov. (2025, May 21). *Climate change: Atmospheric carbon dioxide*. Retrieved August 10, 2025, from https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide

News, Q. (2025, June 24). *Hybrid quantum-classical computing enables simulations of complex chemical systems*. Retrieved August 10, 2025, from https://quantumzeitgeist.com/hybrid-quantum-classical-computing-enables-simulations-of-complex-chemical-systems/

Hsan, N., Kumar, S., Cho, Y., Choi, Y., Byun, H., Saba, R., & Koh, J. (2025). Advancements in carbon-based materials for enhanced carbon dioxide capture and conversion: A comprehensive review. *Fibers and Polymers, 26*(3), 947–974. https://doi.org/10.1007/s12221-025-00871-x

Greene-Diniz, G., Manrique, D. Z., Sennane, W., Magnin, Y., Shishenina, E., Cordier, P., Llewellyn, P., Krompiec, M., Rani, M. J., & Muoz Ramo, D. (2022). Modelling carbon capture on metal-organic frameworks with quantum computing. *EPJ Quantum Technology, 9*, Article 37. https://doi.org/10.1140/epjqt/s40507-022-00155-w

Bharti, K., Cervera-Lierta, A., Kyaw, T. H., Haug, T., Alperin-Lea, S., Anand, A., Degroote, M., Heimonen, H., Kottmann, J. S., Menke, T., Mok, W.-K., Sim, S., Kwek, L.-C., & Aspuru-Guzik, A. (2022). Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics, 94*(1), 015004. https://doi.org/10.1103/RevModPhys.94.015004

Havlek, V., Crcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. *Nature, 567*(7747), 209–212. https://doi.org/10.1038/s41586-019-0980-2

Leung, D. Y. C., Caramanna, G., & Maroto-Valer, M. M. (2014). An overview of current status of carbon dioxide capture and storage technologies. *Renewable and Sustainable Energy Reviews, 39*, 426–443. https://doi.org/10.1016/j.rser.2014.07.093

Greene-Diniz, G., Manrique, D. Z., Sennane, W., Magnin, Y., Shishenina, E., Cordier, P., Llewellyn, P., Krompiec, M., Rani, M. J., & Ramo, D. M. (2022). Modelling carbon capture on metal-organic frameworks with quantum computing. *EPJ Quantum Technology, 9*, 37. https://doi.org/10.1140/epjqt/s40507-022-00155-w

McDonagh, J. L., Wunsch, B. H., Zavitsanou, S., Harrison, A., Elmegreen, B., Gifford, S., Kessel, T. van, & Cipcigan, F. (2023). Machine guided discovery of novel carbon capture solvents. https://doi.org/10.48550/arXiv.2303.14223

Barroca, M. A., Ferreira, R. N. B., & Steiner, M. (2024). Exploration of quantum computing in materials discovery for direct air capture applications. *APL Quantum, 1*, 046103. https://doi.org/10.1063/5.0219500

Gulbalkan, H. C., Aksu, G. O., Ercakir, G., & Keskin, S. (2023). Accelerated discovery of metalorganic frameworks for CO_2 capture by artificial intelligence. *Industrial & Engineering Chemistry Research, 62*(51), 37–48. https://doi.org/10.1021/acs.iecr.3c03817

Park, H., Yan, X., Zhu, R., Huerta, E. A., Chaudhuri, S., Cooper, D., Foster, I., & Tajkhorshid, E. (2024). A generative artificial intelligence framework based on a molecular diffusion model for the design of metal-organic frameworks for carbon capture. *Communications Chemistry, 7*, 1–18. https://doi.org/10.1038/s42004-023-01090-2

Dahale, G. R. (2023). Quantum simulations for carbon capture on metal-organic frameworks. https://doi.org/10.48550/arXiv.2311.12411

Wen, G., Li, Z., Long, Q., Azizzadenesheli, K., Anandkumar, A., & Benson, S. M. (2023). Real-time high-resolution CO_2 geological storage prediction using nested Fourier neural operators. *Energy & Environmental Science, 16*, 1732–1741. https://doi.org/10.1039/D2EE04204E

Mazheika, A., Wang, Y., Valero, R., Ghiringhelli, L. M., Vines, F., Illas, F., Levchenko, S. V., & Scheffler, M. (2022). Artificial-intelligence-driven discovery of catalyst genes with application to CO_2 activation on semiconductor oxides. *Nature Communications, 13*, 419. https://doi.org/10.1038/s41467-022-28042-z

Orhan, I. B., Zhao, Y., Babarao, R., Thornton, A. W., & Le, T. C. (2025). Machine learning descriptors for CO_2 capture materials. *Molecules, 30*(3), 650. https://doi.org/10.3390/molecules30030650

Abdi, J., et al. (2021). Modeling of CO_2 adsorption capacity by porous metal organic frameworks using advanced decision tree-based models. *Scientific Reports, 11*(1), 1–14. https://doi.org/10.1038/s41598-021-04168-w

The Role of Machine Learning in the Understanding and Design of Materials. (n.d.). *Journal of the American Chemical Society*. Retrieved August 4, 2025, from https://pubs.acs.org/doi/10.1021/jacs.0c09105

Wilmer, C. E., Leaf, M., Lee, C. Y., Farha, O. K., Hauser, B. G., Hupp, J. T., & Snurr, R. Q. (2012). Large-scale screening of hypothetical metalorganic frameworks. *Nature Chemistry, 4*(2), 83–89. https://doi.org/10.1038/nchem.1192

Boyd, P. G., Chidambaram, A., Garca-Dez, E., Ireland, C. P., Daff, T. D., Bounds, R., Gadysiak, A., Schouwink, P., Moosavi, S. M., Maroto-Valer, M. M., Reimer, J. A., Navarro, J. A. R., Woo, T. K., Garcia, S., Stylianou, K. C., & Smit, B. (2019). Data-driven design of metal-organic frameworks for wet flue gas CO_2 capture. *Nature, 576*(7786), 253–256. https://doi.org/10.1038/s41586-019-1798-7

Sung, I.-T., Cheng, Y.-H., Hsieh, C.-M., & Lin, L.-C. (2025). Machine learning for gas adsorption in metalorganic frameworks: A review on predictive descriptors. *Industrial & Engineering Chemistry Research, 64*(4), 1859–1875. https://doi.org/10.1021/acs.iecr.4c03500

Mahajan, S., & Lahtinen, M. (2022). Recent progress in metal-organic frameworks (MOFs) for CO_2 capture at different pressures. *Journal of Environmental Chemical Engineering, 10*(6), 108930. https://doi.org/10.1016/j.jece.2022.108930

Chung, Y. G., Camp, J., Haranczyk, M., Sikora, B. J., Bury, W., Krungleviciute, V., ... & Snurr, R. Q. (2014). Computation-ready, experimental metalorganic frameworks: A tool to enable high-throughput screening of nanoporous crystals. *Chemistry of Materials, 26*(21), 61856192. https://doi.org/10.1021/cm502594j

Fernandez, M., Barnard, A. S., & Coudert, F. X. (2020). Machine-learning prediction of the performance of metalorganic frameworks for CO_2 capture. *The Journal of Physical Chemistry Letters, 11*(11), 44054410. https://doi.org/10.1021/acs.jpclett.0c00801

Larsen, A. H., Mortensen, J. J., Blomqvist, J., Castelli, I. E., Christensen, R., Dulak, M., Friis, J., Groves, M., Hammer, B., Hargus, C., Hermes, E., C. Jennings, P., Jensen, P. B., Kermode, J., Kitchin, J., Kolsbjerg, E., Kubal, J., Kaasbjerg, K., Lysgaard, S., ... Jacobsen, K. W. (2017). The Atomic Simulation Environment - A Python library for working with atoms. Journal of Physics Condensed Matter, 29, Article 273002. https://doi.org/10.1088/1361-648X/aa680e

Quantum machine learning with Qiskit: Evaluating regression accuracy and noise impact. (2025). *ResearchGate*. https://doi.org/10.1049/qtc2.12100

Research AssistantBohrium. (n.d.). *AI for Science with Global Scientists*. Retrieved August 4, 2025, from https://www.bohrium.com/paper-details/a-linear-programming-approach-for-battery-degradation-analysis-and-optimization-in-offgrid-power-systems-with-solar-energy-integration/811115274580262912-3789?

Bordin, C., Anuta, H. O., Crossland, A., Lascurain Gutierrez, I., Dent, C. J., & Vigo, D. (2016). A linear programming approach for battery degradation analysis and optimization in offgrid power systems with solar energy integration. *Renewable Energy, 101*, 417430. https://doi.org/10.1016/j.renene.2016.08.066

Majidi, H., Hayati, M. M., Breyer, C., Mohammadi-ivatloo, B., Honkapuro, S., Karjunen, H., Laaksonen, P., & Sihvonen, V. (2025). Overview of energy modeling requirements and tools for future smart energy systems. *Renewable and Sustainable Energy Reviews, 212*, 115367. https://doi.org/10.1016/j.rser.2025.115367