**ANL 488 PROJECT PROPOSAL**

**Optimization of production portfolios in the**

**upstream oil and gas industry**

**to minimize emissions**

**using Python**

****

**Submitted by**

**KHOO WAN TING**

**SCHOOL OF BUSINESS**

**Singapore University of**

**Social Sciences**

**Presented to Singapore University of Social**

**Sciences in partial fulfillment of the**

**requirements for the**

**Degree of Bachelor of Science**

**in Business Analytics**

**2023**

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| Chapter 1 | Introduction | 1 |
| Chapter 2 | Literature reviews | 2 – 5 |
| Chapter 3 | Data understanding and preparation | 6 - 14 |
| Chapter 4 | Proposed modelling and evaluation | 15 |
| Chapter 5 | Proposed schedule | 16 |
| References |  | 17 - 18 |

1. Introduction

This study focuses on the optimization of production portfolios in the upstream oil and gas industry to minimize emissions using Python. The upstream oil and gas industry includes activities such as finding reservoirs, acquiring land rights, and onshore and offshore drilling of oil and gas wells for production. Around 15% of all energy-related emissions worldwide, or 5.1bn tons of greenhouse gas (GHG) emissions, are attributable to the operations of the oil and gas industry (International Energy Agency [IEA], 2023). Butterfield et al. (2005) researched that while the impacts of climate change could be positive in the short term, they would eventually turn negative in the long term once it gets severe. Climate change affects everyone and everywhere in various ways such as frequent changes in rainfall and temperature negatively impacting agriculture cultivation and water resources, health issues arising from thermal stress and disease burden, and ecological productivity and biodiversity due to sea-level rises leading to extinction of some species.

Therefore, to minimize the large emissions speeding up climate change, the business and data mining objective is to utilize optimization techniques to identify assets within the portfolio to divest. Optimization of production portfolio can help to minimize emissions by divesting assets releasing the highest emissions per production unit. The solution would thus include converting the objective into a mathematical equation, identifying the decision variables, and setting the constraints and parameters using optimization packages in Python.

2. Literature reviews

2.1 - Project Portfolio Optimization in a Changing Energy Landscape

Moubarak (2021) conducted a study aiming to identify assets to divest and the time to divest them, wanting to reduce carbon emissions while comparing the oil and gas (O&G) assets with the wind farms in a portfolio context. The dataset used was hypothetically created, consisting of a portfolio of 16 assets, “ten hydrocarbons, five wind farms, and one carbon capture and storage (CCS)”.

Using time-dependent stochastic aggregations in Python, the objective statement was to minimize the carbon emissions, by selecting the optimal combination of its 16 assets, with the primary constraint being annual expected Capital Expenditures (capex). Two scenarios are then considered for modelling. Firstly, by investing a 100% working interest in CCS, the O&G production targets would be maintained. Secondly, as it mentioned that wind energy production would release less emissions than oil production, changing the production portfolio by increasing the percentage of renewables, wind farms, with less working interest in CCS (Moubarak, 2021).

As this analysis pre-sets the percentage of working interest invested into the CCS assets, it would not be able to identify the optimal combination for all 16 assets, but only identify the best portfolio among those created. With the performance evaluation of three portfolios (C, D, and E) by assessing the expected payoffs and applying weights to the objectives, the research was concluded. Depending on the decision-maker’s preferences, Portfolio C would be optimal regardless of the weights on the emission and oil production targets, Portfolio D would be optimal if the weight on the gas production target increases, and Portfolio E would be optimal if the weight on the Net Present Value (NPV) is “high enough” (Moubarak, 2021).

This literature review is thus relevant to this study as it analyses the carbon emissions depending on the combination of its 16 hydrocarbon and CCS assets through the selection of the optimal portfolio by applying weights to the various objectives that the company could want to focus on.

2.2 - An improved portfolio optimization model for oil and gas investment selection

Dong et al. (2014) conducted a simulation using “19 overseas upstream assets owned by a large oil company in China” with the objective of maximizing expected utility at a pre-set level of risk under budget constraints.

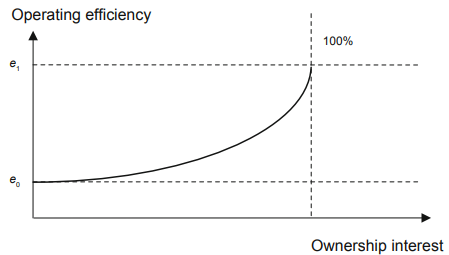


Figure 1: Ownership interest’s influence on assets’ operating efficiency (Dong et al, 2014)

The input-output ratio, return on investment (ROI) that a project will bring, was used to measure the operating efficiency. Since an asset’s output would be affected depending on the percentage of ownership interest, similar to the previous literature review, weights (*w*) were used to indicate the ownership interest of each asset. Based on management experience, the operating efficiency was expressed in a nonlinear form of ownership interest, *wa*, “where *a* indicates how much the ownership interest of the underlying asset affect the elasticity of output efficiency”. The convexity between the operating efficiency and ownership interest is also matched to the assumption that the control powers of the major and minor shareholders on the same project do not follow the proportion of equity allocation (Dong et al, 2014).

Even though this study uses a different optimization technique, simulation, it still focuses on optimizing assets within an upstream O&G industry capital budgeter’s portfolio. The constraints and assumptions set and how the objectives and constraints were converted into mathematical equations would still be relevant to this study. Figure 1 above also acts as a reiteration of how allocating weights would be important as it affects the selection of the optimum portfolio.

2.3 – A portfolio optimization model for a large number of hydrocarbon exploration projects

Bulai & Horobet (2018) used a fictional dataset consisting of “40 currently producing fields, 40 production enhancement projects, and 20 exploration projects” with the objective of increasing production in those fields. They made use of expected NPV calculations for each portfolio item for the different scenarios created before applying the Markowitz portfolio optimization theory. The objective is then to find the optimal vector of weights for each portfolio item, minimizing the portfolio variance, subject to the expected portfolio return and capex constraints.

This study not only reiterates the importance of weights on the objectives but also provides another angle to look at the assets. While they used “currently producing”, “production enhancement”, and “exploration” to categorize their projects, it is similar to the “field\_maturity” field in the dataset used for this study.

3. Data understanding and preparation

The dataset used is on the upstream oil and gas industry with fields such as but not limited to company names, hydrocarbon types, each field’s location, development type, onshore and offshore breakdown, and financial data such as capex, and operating expenses (opex).

A screenshot of a computer

Description automatically generated

Figure 2: Filtered and sorted company results from Microsoft Excel’s Stock function

The study focuses only on five international oil companies (IOC) and five national oil companies (NOC). Using Microsoft Excel’s built-in Stock function as seen in Figure 2 above, the industries of each company were identified and filtered to only show those under the O&G industry. The market cap was also shown using the function, then sorted from the largest to smallest, where the 5 IOCs were selected from there, namely, Reliance, INPEX Corporation, Chevron, Shell, and ConocoPhillips. The 5 NOCs were then selected based on 5 different countries, namely, Oil and Natural Gas Corporation (ONGC) from India, Sinopec Corp from China, PTT Exploration and Production (PTTEP) from Thailand, Petronas from Malaysia, and Rosneft from Russia. The first step of the data preparation was done in the three raw Microsoft Excel files, Production, Cost, and Revenue, by only keeping records of the companies of focus. While there were records in the raw dataset with similar company and subsidiary names that could be referring to the same entity, they were assumed to be separate entities unless both the company and subsidiary names were the same.

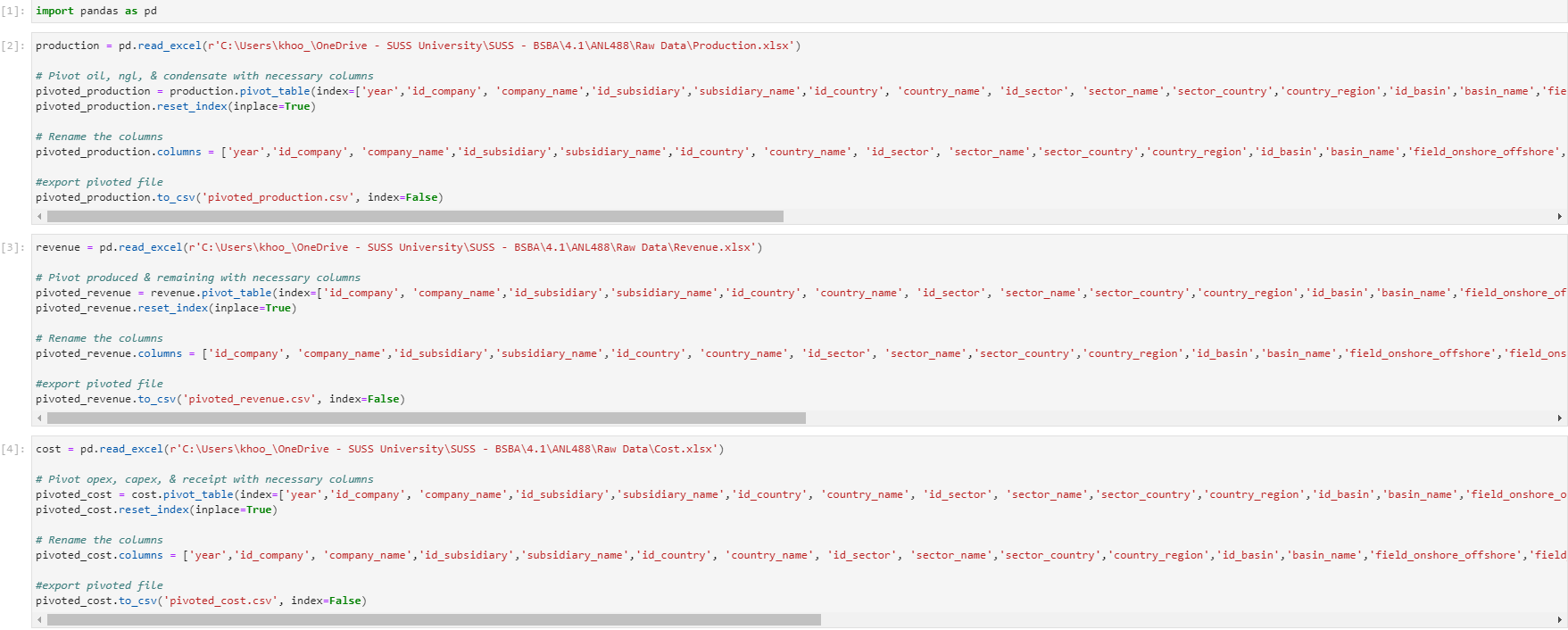


Figure 3: Data preparation – pivoting of data using Python\_1

A screenshot of a computer

Description automatically generated

Figure 4: Data preparation – pivoting of data using Python\_2

The next cleaning step was done in Python to select the fields wanted in each spreadsheet and to pivot the data. The pivoted datasets were then exported as CSV files thereafter, codes as seen in figures 3 and 4 above.

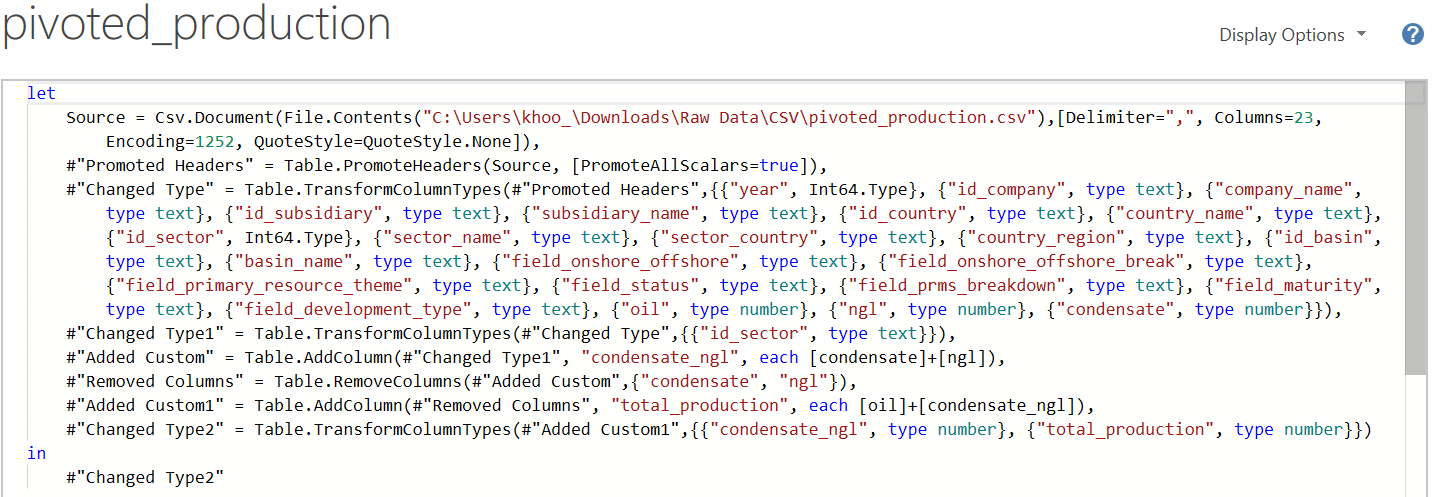


Figure 5: Data preparation – cleaning of data using Power Query

The three CSV files exported from Python were first imported into Power Query and edited separately by changing certain data types such as the year and the “id\_” fields from number to text. The “id\_sector” field in the revenue and production datasets was also set differently from the one in the cost dataset as they had one decimal place added when simply changing their data types to text. Their data types were thus changed twice, first to whole numbers, then to text, to remove the decimal place as the field would be used as one of the primary keys for merging. The “condensate” and “ngl” fields were combined into one field, “condensate\_ngl” by summing their respective metric values together. A custom field “total\_production” was also created by summing the oil, condensate, and Natural Gas Liquids (NGL) values. Figure 5 above shows the data cleaning steps done to the “production” dataset, as an example of the steps done to all three datasets.



Figure 6: Data preparation – merging of data using Power Query

Next, the production and cost datasets were merged first with their first 20 fields as the composite primary keys, resulting in the “Merge1” dataset. The “opex”, “capex”, and “receipt” fields were then expanded from the cost dataset to merge into the production dataset. Thereafter, “Merge1” was merged with the revenue dataset with 19 fields as the composite primary keys, excluding the “year” field since the revenue dataset does not have that field, resulting in the “merged\_data” dataset. The “remaining” and “produced” fields were then expanded from the revenue dataset to merge into the “Merge1” dataset. A custom column, “ncf”, was added using the formula (receipt + produced + remaining) – (capex + opex) to calculate its values.

The “onshore\_offshore\_breakdown” field was duplicated and then pivoted alongside the addition of two other custom columns for the estimated emissions’ fields to be added to the merged dataset, as seen in Figure 6 above. Research conducted by EPA (2023) stated that coal releases the largest amount of emissions per unit of useful energy, with oil releasing approximately 75% of coal’s, and natural gas releasing 55% of coal’s. The ratio was thus set at 1:0.73 for oil to condensate and gas, assuming that condensate’s emissions would be close to gases’ emissions. Research conducted by ICF (2023) mentioned that the emissions released from onshore production facilities could add up to around 3k tons of CO2e, equivalent to 200kg/ft of the measured well depth and around 9.7k tons of CO2e, equivalent to 600kg/ft of the measured well depth for offshore production facilities. The ratio was thus set at 1:0.33 for offshore production to onshore production. Concluding the data cleaning, the dataset now has a total of 32 fields and 2,616 records.

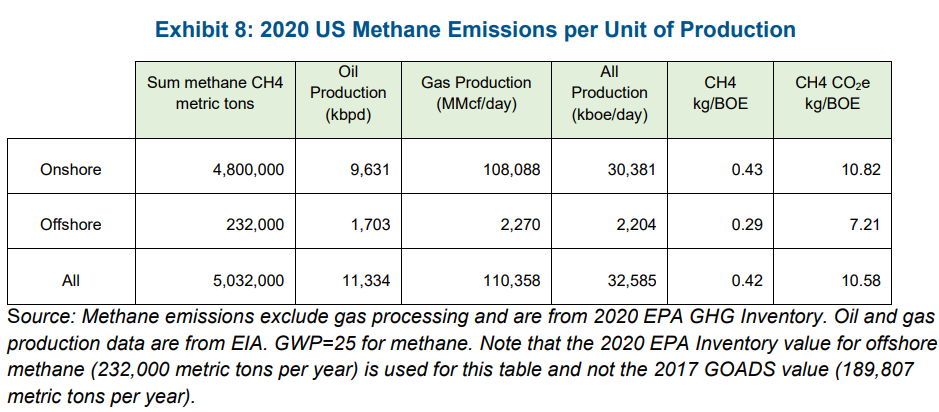


Figure 7: US methane emissions per unit of production in 2020 (ICF, 2023)

One of the angles to identify the types of assets to divest would be whether the field development is onshore or offshore. Even though the emissions from offshore facilities are around three times more than onshore facilities, due to their higher well productivity, the emissions per production unit should be examined instead of just the total emissions released. As per the example in Figure 7 above, ICF could divest more of their onshore facilities and invest more in their offshore facilities to minimize their emissions per production unit.

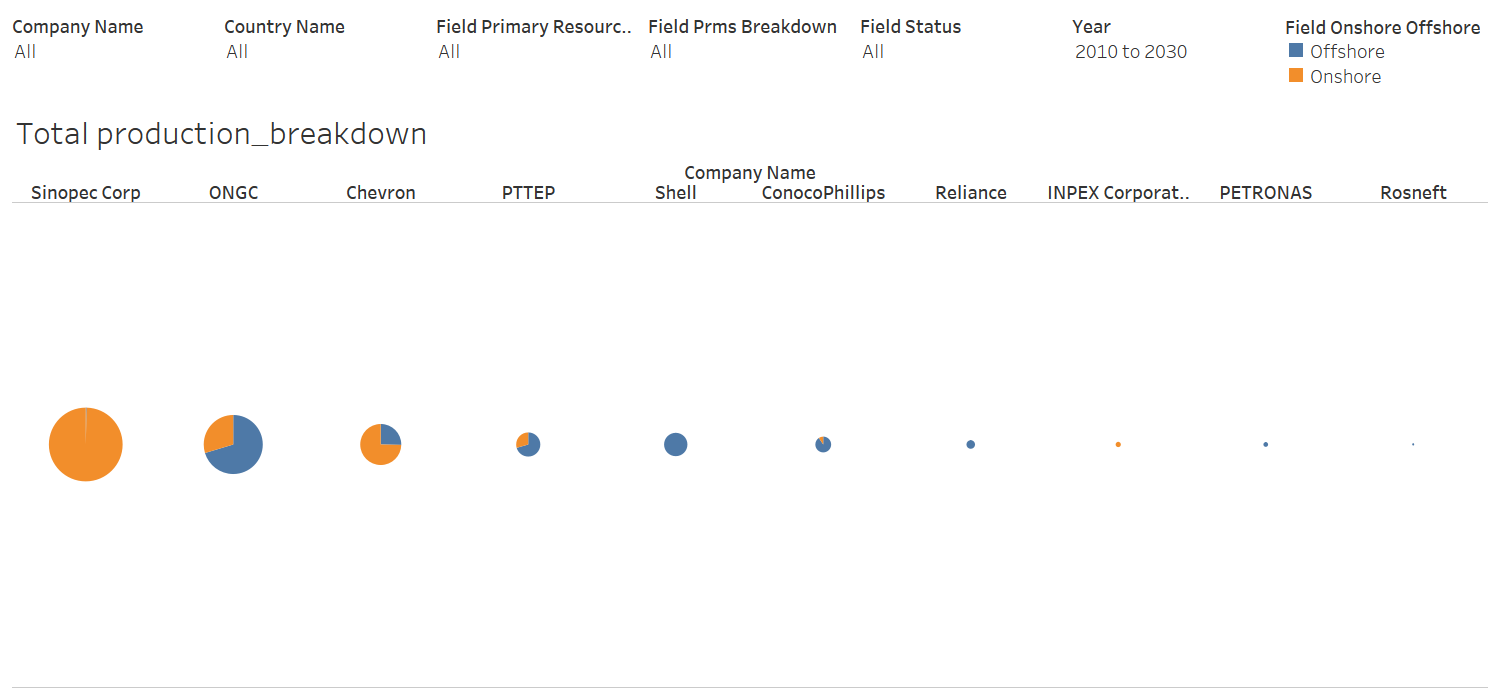


Figure 8: Data understanding – pie charts on total production breakdown for each company in Tableau

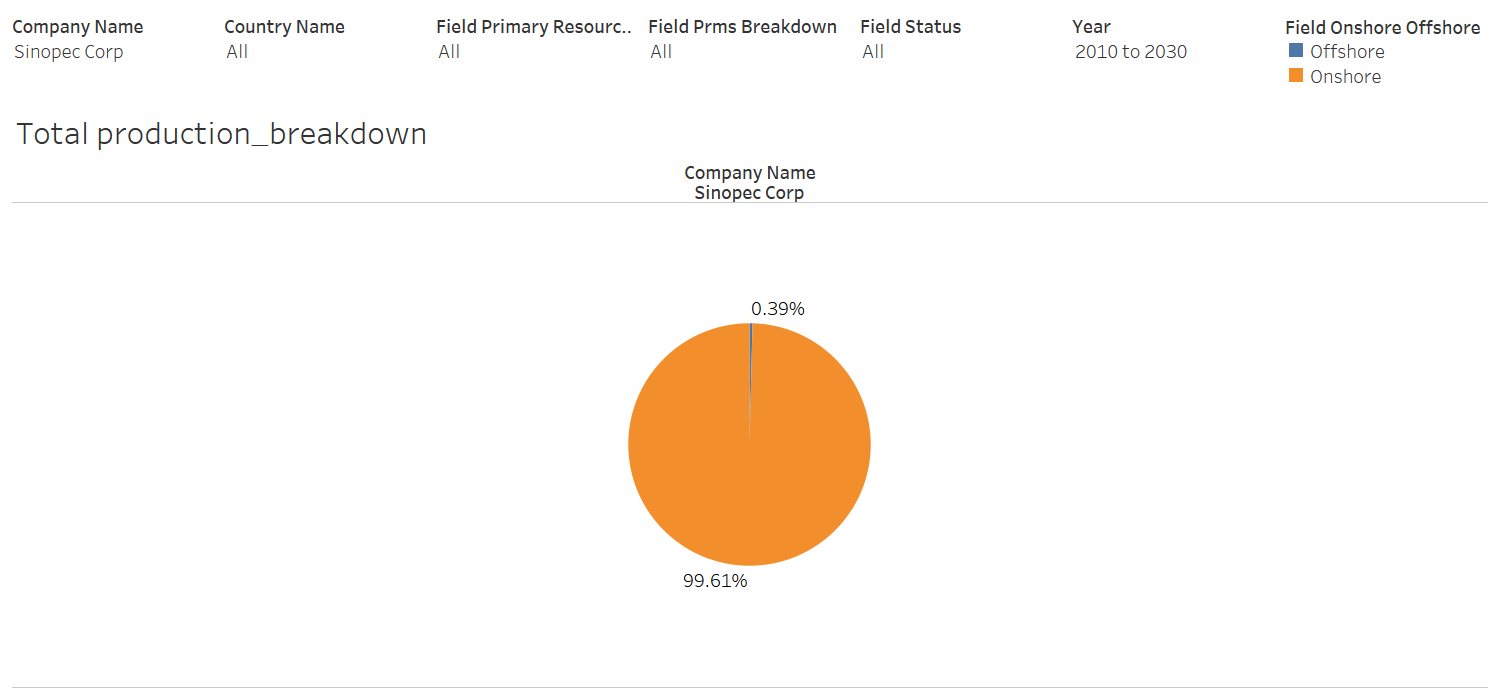


Figure 9: Data understanding – Sinopec Corp pie chart on total production breakdown in Tableau

Figure 8 above shows ten pie charts for each company’s breakdown of onshore and offshore facilities for their production. Taking Sinopec Corp as an example, as seen in Figure 9, it holds almost 100% of onshore facilities. Further analysis should be done to verify if the emission per production unit is low, following the relatively low emissions from onshore facilities.

Another method to identify assets to divest would be modifying the proportion of oil production and condensate or NGL production. EPA (2023) research found that the amount of carbon dioxide emission released by oil combustion is higher than gas combustion. While the emissions are calculated based on oil and gas combustion which occurs in the downstream oil and gas industry, this could still be a useful angle to look at when optimizing state-owned oil companies (NOC) because reducing the emissions for the industry in general is relevant as well. This would, however, not be relevant to international oil companies (IOC) as they, as individual companies, are not responsible for the emissions by other companies handling the downstream side of the industry.

A graph of different colored lines

Description automatically generated

Figure 10: Data understanding – line charts on emissions from Tableau\_1

The cleaned “merged\_data” dataset was then imported into Tableau for exploratory data analysis and data visualization. Tableau was used so that an interactive dashboard could be created. At first glance of Figure 10 above, it seems like the focus companies produce more gas than oil, but the y-axis range in the oil emission chart is 20 times larger than the gas emission chart’s.

A graph of different colored lines

Description automatically generated

Figure 11: Data understanding – line charts on emissions from Tableau\_2

A further analysis was thus conducted by using a filter only applicable to the oil emissions chart where the top oil producers, Chevron, ONGC, and Sinopec Corp, were excluded. It can be seen from Figure 11 above that including the three excluded companies, that most of the companies do produce more oil emissions than gas emissions.

4. Proposed modelling and evaluation

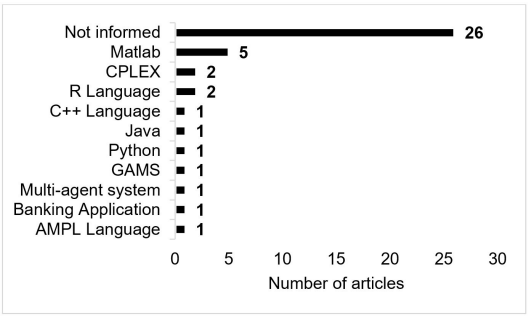


Figure 12: Software or programming language used for portfolio optimization (Milhomem & Dantas, 2020)

I intend to formulate parametric equations for the objective(s) and constraints, assign the decision variables, and then make use of Python to solve the optimization problem. There are many libraries and packages available in Python for optimization but there are not many projects that make use of Python for their business portfolio optimization, as seen in Figure 12 above (Milhomem & Dantas, 2020).

When setting the objective(s) and constraints, I would be assigning weights to them, as demonstrated in the literature reviews. With the various angles available to look at the assets to divest, it can also serve as a comparison for the company to have the final decision on which assets they would prefer to divest, due to other possible limitations such as the country’s political instability that would not be reflected in the optimization results.

For the evaluation of the results, I would drop the assets from the dataset and then plot visualization charts to see the improvement in carbon emissions that the recommended portfolio(s) would bring. Taking inspiration from the literature reviews, implementing the efficient frontier concept would be a good method to evaluate the model results as well.

Proposed schedule

|  |  |
| --- | --- |
| Week | Description |
| 0 (before 12 Aug) | * Choosing of project topic * Received dataset |
| 1-3 | * Project briefing seminar * 1st supervisor meeting session * Literature review and exploratory data analysis * Prepare for proposal submission |
| 4 | * Submission of proposal * Proceed with modelling |
| 5-8 | * 2nd supervisor meeting session * Edit proposal portions to be included in final report based on supervisor feedback * Prepare for oral presentation |
| 9-10 | * Oral presentation |
| 11-15 | * Edit final report draft based on feedback received during oral presentation * 3rd supervisor meeting session * Prepare for final report submission |
| 16 | * Final report submission |

(Word count: 2,249 words)

References

Bulai, V. & Horobet, A. (2018). A portfolio optimization model for a large number of hydrocarbon exploration projects. *Proceedings of the International Conference on Business Excellence, 12(1),* 171-181. doi: 10.2478/picbe-2018-0017

Butterfield, R., Downing, T. Handley, C., & Watkiss, P. (2005). *The Impacts and Costs of Climate Change*. Retrieved from: <https://climate.ec.europa.eu/system/files/2016-11/final_report2_en.pdf>

Dong, Z., Liu, S., Qing, X., & Zhen, W. (2014). An improved portfolio optimization model for oil and gas investment selection. *Petroleum Science, 11,* 181-188. doi: 10.1007/s12182-014-0331-8

EPA. (2023). *Inventory of U.S. Greenhouse Gas Emissions and Sinks.* Retrieved from: <https://www.epa.gov/system/files/documents/2023-04/US-GHG-Inventory-2023-Main-Text.pdf>

ICF. (2023). *GHG Emission Intensity of Crude Oil and Condensate Production.* Retrieved from: <https://www.noia.org/wp-content/uploads/2023/05/NOIA-Study-GHG-Emission-Intensity-of-Crude-Oil-and-Condensate-Production.pdf>

International Energy Agency. (2023). Emissions from Oil and Gas Operations in Net Zero Transitions. Retrieved from: <https://iea.blob.core.windows.net/assets/2f65984e-73ee-40ba-a4d5-bb2e2c94cecb/EmissionsfromOilandGasOperationinNetZeroTransitions.pdf>

Milhomem, D. A. & Dantas, M. J. P. (2020). Analysis of new approaches used in portfolio optimization: a systematic literature review. *Production*, 30, e20190144. doi: 10.1590/0103-6513.20190144

Moubarak, R. (2021). *Project Portfolio Optimization in a Changing Energy Landscape.* Retrieved from: <https://uis.brage.unit.no/uis-xmlui/bitstream/handle/11250/3038485/Project%20Portfolios%20Optimization%20in%20a%20Changing%20Energy%20Landscape_Racha%20Moubarak.pdf?sequence=2&isAllowed=y>